

الجمهورية الجزائرية الديمقراطية الشعبية
REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
وزارة التعليم العالي والبحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche
Scientifique

Université MUSTAPHA STAMBOULI de
Mascara
Faculté des sciences Humaines et sociales



جامعة مصطفى اسطمبولي معسكر

كلية العلوم الإنسانية والاجتماعية

قسم : علم النفس وعلوم التربية

الدكتورة: سليمان فاطمة الزهراء

السند البيداغوجي الخاص بمقياس:
بناء الاختبارات

موجه لطلبة السنة: الثانية ليسانس

فرع: علم النفس

ميدان: علوم إنسانية واجتماعية

السنة الجامعية: 2023-2024

الفهرس:

05	مقدمة.....
06	1 - التعريف بالمقياس.....
10	المحاضرة الأولى التطور التاريخي للاختبارات منذ نشأتها إلى الوقت الحالي.....
10	1- نشأة القياس النفسي وتطوره قبل (بينيه، Binet):.....
12	2- أعمال (بينيه، Binet) في قياس القدرة العقلية.....
12	3- تطور حركة القياس النفسي بعد أعمال بينيه.....
14	4- أسس القياس النفسي.....
15	خلاصة.....
17	المحاضرة 02: أنواع الاختبارات.....
17	1- تعريف الاختبار.....
18	2- تصنيف الاختبارات النفسية.....
18	2-1- الاختبارات الإسقاطية.....
20	2-2- الاختبارات الموضوعية.....
22	خلاصة.....
25	المحاضرة 03: معايير تصديق الاختبار.....
25	1- صدق الاختبار:.....
25	2- أنواع الصدق.....
26	2-1- الصدق الخارجي أو السطحي (Face Validity):.....
27	2-2- الصدق الداخلي.....
32	3- العوامل التي تؤثر في صدق الاختبار.....
33	خلاصة.....
35	المحاضرة 04: ثبات الاختبار.....
35	1- تعريف الثبات (Reliability):.....

35	2- طرق قياس ثبات الاختبار النفسي
38	3 -العوامل المؤثرة على ثبات الاختبار
38	خلاصة
41	المحاضرة 05: التمييز
41	1-المقصود بالتمييز
42	2-مؤشر صعوبة البنود في حالة الإجابة ببدلين
42	3- مؤشر صعوبة البنود في حالة الإجابة بعدة بدائل:
42	4- حساب معامل الصعوبة والسهولة للاختبار ككل:
43	5- مؤشر التمييز(Discrimination Coefficient):
44	خلاصة
46	المحاضرة 06: المعايرة
46	1- تعريف المعيار والمعايرة
46	2- الشروط الواجب توفرها في المعيار
47	3- أنواع المعايير
49	خلاصة
51	المحاضرة رقم 07:المعاينة
51	1- تعريف المعاينة والعينة
51	2- شروط المعاينة
51	3- أسلوب العينة وأسلوب التعداد
53	4- خطوات اختيار العينة
54	5- أنواع العينات
61	خلاصة
63	المحاضرة 08: التقنين
63	1 - تعريف التقنين (Standarization)
63	2- خطوات تقنين اختبار

65	3- خصائص عينة التقنين ودليل الاختبار
65	خلاصة
66	خاتمة
67	قائمة المراجع

قائمة الجداول:

14	1- أدوات القياس النفسي الأكثر استخداما
22	2- تصنيف الاختبارات النفسية
27	3- النسب المتوقعة لاتفاق المحكمين حسب لوشي
44	4- دلالات معامل التمييز
52	5- جوانب القوة والضعف في استعمال كل من التعداد والعينة
55	6- جوانب القوة والضعف للعينات الاحتمالية وغير الاحتمالية
57	7- نموذج مصغر عن جدول الأرقام العشوائية
58	8- نموذج عن العينة التطبيقية

قائمة الأشكال:

06	1- تقديم مقياس بناء الاختبارات
08	2- أهداف المقياس في علاقتها بالتعلم
15	3- خريطة مفاهيمية لتطور حركة القياس النفسي
18	4- ماهية الاختبار النفسي وخصائصه
22	5- تصنيف الاختبارات
64	6- خطوات تقنين اختبار

مقدمة:

تطور العلوم في مختلف الميادين قائم على مدى قدرتها على قياس مختلف الظواهر الطبيعية أو الفيزيائية أو البيولوجية بدقة تعكس تناول الموضوعي لها; ومع أن الظاهرة النفسية ظاهرة هجينة غير صافية، إلا أن العلوم النفسية بعد انفصالها عن الفلسفة أخذت هي الأخرى منحى التجريب وخضعت للتجريب والقياس حتى تكتم نتائجها وتقرب من الدقة والموضوعية، كان هذا من منطلق مقولة (ثورندايك، Thorndike) المشهورة التي مفادها أن كل شيء يوجد بمقدار يمكن قياسه، لذا شهدت فترة التسعينات من القرن الماضي اهتماما واضحا بالاختبارات النفسية من خلال تطوير اختبارات جديدة يمثل بعضها مناهج حديثة في تناول الظاهرة النفسية، من خلال تشخيص المشاكل الانفعالية الخطيرة والسلوك المتعلق بالمشاكل اللغوية، وظواهر أخرى فرضتها الحاجة، كتصنيف الأطفال في المدارس بناء على قدرتهم العقلية، والتوجيه التربوي للمرحلة الثانوية، واختيار المهن. (Anastasi&Susana,2006).

يحمل القياس معنى التقدير الكمي للأشياء والمستويات، كما يدل كذلك على عملية مقارنة نتائج القياس بنتائج أخرى على عينة ممثلة للسلوك المراد قياسه، وبالتالي فهو لا يخرج عن كونه يفسر الظاهرة إحصائيا مع عملية المقارنة المحكية أو المرجعية التي ترافق التقدير العددي للخاصية موضوع القياس.

القياس لم يعد مقتصرًا على العلوم الدقيقة والعلوم التجريبية، وإنما اقتحم مجال العلوم النفسية وأضحى القياس النفسي من أهم فروع علم النفس وأساس البحوث العلمية في المجال النفسي، لذا فهو يرتبط بالعديد من الفروع كعلم النفس التربوي، علم النفس العيادي، علم النفس المدرسي، علم نفس عمل وتنظيم، علم نفس النمو، وعلم النفس الاجتماعي.

ومنه فإن القياس النفسي والاختبارات النفسية تعتبر نقطة تحول في تاريخ تطور علم النفس، وأعطت دفعة قوية لتطور وانتشار الاختبارات النفسية في إطار خضوعها للسيكومترية بشكل يسهل عملية البحث العلمي وبالتالي وصف الظاهرة النفسية والتنبؤ بها، ومن ثم التحكم وإحداث التغيير.

سيتم في هذا السداسي تناول مقياس بناء الاختبارات من خلال سلسلة من المحاضرات التي تشرح ماهية الاختبارات وأنواعها، طرق التأكد من خصائصها السيكومترية من صدق وثبات وكذا معامل السهولة والصعوبة، والقدرة التمييزية لبنوده، كذا أنواع المعايرة، طرق حسابها، وأساليب المعاينة الصحيحة، وصولا إلى التعرف على ماهية وشروط تقنين الاختبار.

سيتمكن الطالب من خلال سلسلة المحاضرات هذه من خوض غمار البحث العلمي لاسيما الجانب القياسي منه بكل ثقة، ومعرفة أدواته في إطار منهج علمي سليم، حتى يقترب من تحقيق الموضوعية المرجوة في العلوم الاجتماعية.

1-التعريف بالمقياس:

المقياس: بناء الاختبارات

الفئة المستهدفة: سنة ثانية علم النفس

الوحدة: منهجية

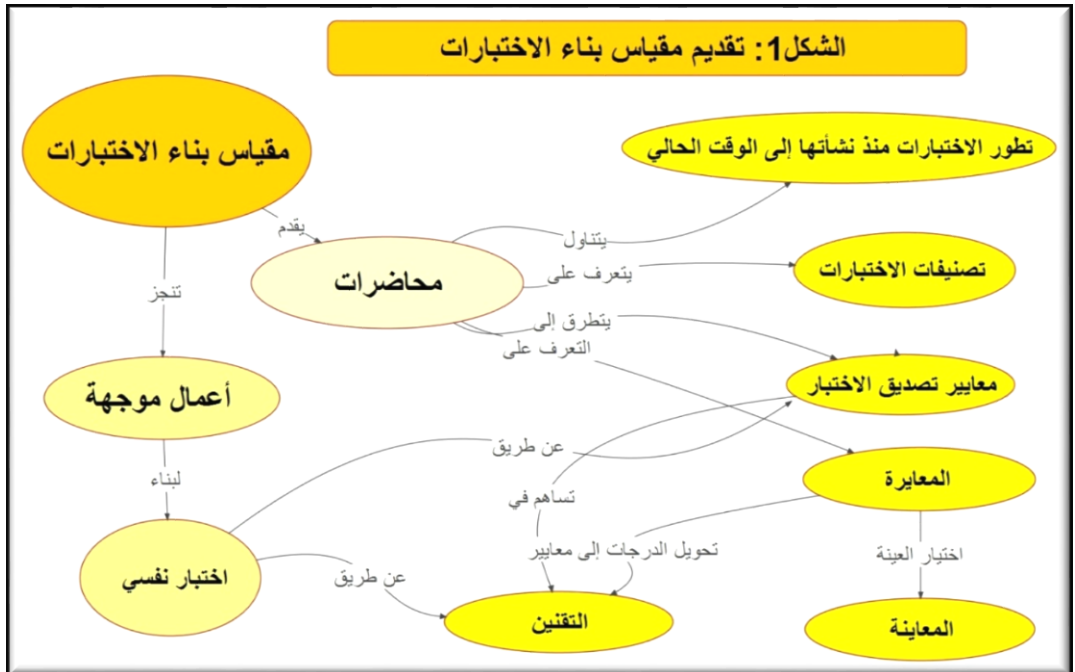
المعامل: 01

الرصيد: 02

مقياس بناء الاختبارات الموجه إلى طلبة السنة الثانية علم النفس، هو تكملة لمقياس القياس النفسي الذي تم تناوله سابقا في السداسي الأول، يؤسس هذا المقياس للتعرف على نشأة الاختبارات والمقاييس النفسية عموما، أنواعها وتصنيفاتها، ومن ثم طرق تحقيق معايير التصديق من حسابات للصدق والثبات، ثم المعايرة والمعاينة برصد لمختلف أنواع العينات وطرق اختيارها، ضمنا بعدها لتقنين أمثل للاختبار أو المقياس في بيئة ثقافية جديدة مع مراعاة حساب خصائصه السيكمومترية كلما اختلف مجتمع الدراسة، والشكل أدناه يلخص مكونات المقياس:

شكل 01

تقديم مقياس بناء الاختبارات



2-أهداف المقياس:

يهدف هذا المقياس إلى تحقيق ملمح خروج يتجلى في الكفاءات التي يكتسبها المتعلم والتي تلمس الجوانب المعرفية، المعرفة الفعلية والمعرفة الوجدانية تتمثل فيما يلي:

الكفاءة الختامية المستهدفة هي قدرة الطالب على بناء وتقنين الاختبارات النفسية وفق معايير الصدق والثبات مع التمكن من المعايرة والمعاينة حسب منهجية البحث العلمي السليمة والمعتمدة.

يمكن اكتساب هذه الكفاءة بالتدرج المستمر والتطور التدريجي بين مستوياتها بدءا باكتساب المعارف حول المادة، ثم المعارف الفعلية والتي تخص تطوير الممارسات ومهارات الأداء، ثم المعارف الوجدانية حيث يصبح هذا التعلم ملزما وراسخا في تفكير الطالب وسلوكه المستقبلي تمكنهم من تقويم وإصدار أحكام في مواقف تعلم جديدة بشكل احترافي.

وعليه، فإن أهداف تعلم هذا المقياس عموما تستهدف ملمح خروج تعبر عنه الكفاءات التالية:

2-1-على مستوى المعرفة: أن يتمكن الطالب من خلال دراسة هذا المقياس من:

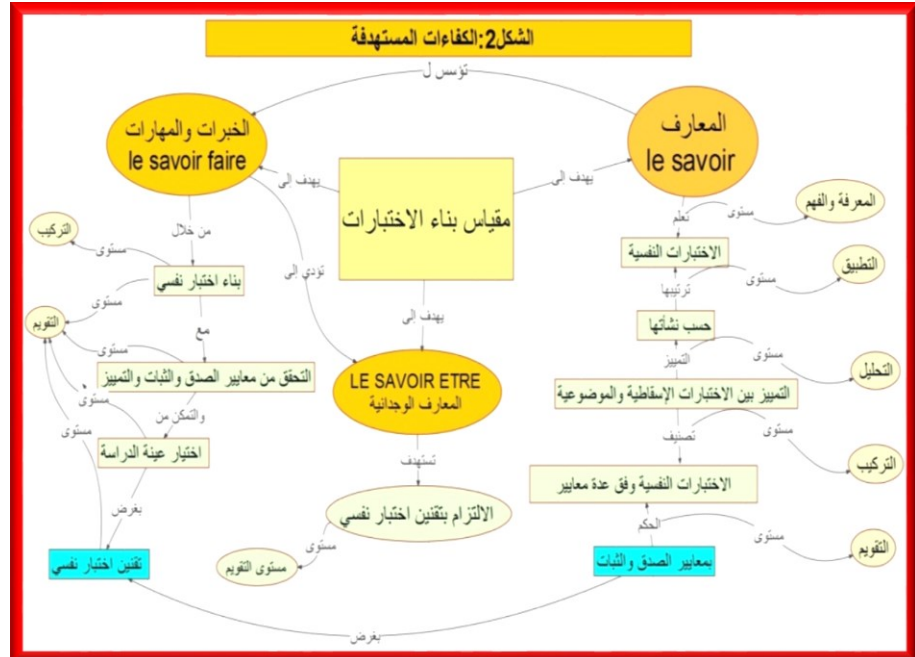
- معرفة الاختبارات النفسية.
- ترتيب الاختبارات النفسية حسب نشأتها.
- التمييز بين الاختبارات النفسية الإسقاطية والموضوعية.
- تصنيف الاختبارات النفسية حسب معايير الأداء، طريقة التصحيح وطبيعة السمة.
- الحكم على معايير الصدق والثبات.

2-2-على مستوى المعارف الفعلية: أن يتمكن الطالب من خلال دراسة مقياس بناء الاختبارات من:

- بناء اختبار نفسي وفق الخطوات المحددة.
- التحقق من الخصائص السيكومترية بتطبيق معاملات الصدق والثبات والتمييز.
- اختيار عينة الدراسة بأكثر من طريقة حسب فرضيات الدراسة.
- تقنين اختبار نفسي.

2-3-على مستوى المعارف الوجدانية: أن يتمكن الطالب في هذا المستوى من:

- الالتزام بتقنين الاختبار النفسي قبل تطبيقه في بيئة ثقافية جديدة.
- والشكل أدناه يعرض هذه الأهداف من خلال ربطها مع مواضيع المقياس:



3- المكتسبات القبلية:

اكتساب المعرفة في منهاج المقاربة بالكفاءات يتأسس لوليا باستثمار للمكتسبات القبلية بمعنى ربط للتعلم السابق بالتعلم اللاحق، فالمعرفة تراكمية، ولا يمكن فهم دروس هذا المقياس دون التمهيد لها من خلال بناء أرضية صلبة لتعلم سابق ركز على تحصيل المعارف التالية:

- 1- ماهية القياس النفسي
- 2- مفهوم الخاصية أو السمة
- 3- مستويات القياس النفسي
- 4- أدوات القياس النفسي
- 5- الاختبارات النفسية: البناء والاستخدام

4- محتويات المقياس الحالي (بناء الاختبارات):

يهتم المقياس الحالي أولا بعرض تاريخ القياس النفسي والاختبارات النفسية وتطورها التدريجي منذ نشأتها إلى الوقت الحالي، ثم تصنيفها، وأنواعها وفق عدة محكات، ثم البحث في خصائصها السيكمومترية من صدق وثبات، وكذا تحليل فقرات الاختبار لحساب معامل صعوبتها وسهولتها وقوتها التمييزية. يتطرق المقياس كذلك إلى المعايرة وأنواعها وطريقة حساب المعايير وتحويل درجات الاختبار الخام إلى درجات معيارية، ثم أساليب المعاينة، شروطها وأنواعها وصولا إلى تقنين الاختبارات.

الأهداف

في نهاية المحاضرة الأولى يجب أن يكون الطالب قادراً على:

- 1- يستفيد من مكتسباته القبلية في تحديد دقيق لماهية الاختبار النفسي دون الخلط بينه وبين مصطلحات أخرى.
- 2- يوضح سيرورة تطور الاختبارات النفسية بناء على تسلسل زمني
- 3- يوظف ما تعلمه في ترتيب تطور الاختبارات حسب منشئها
- 4- يميز بين اختبارات الفروق واختبارات الشخصية.
- 5- يصنف الاختبار النفسي حسب طبيعة السمة المقاسة
- 6- يدرك الحاجة إلى ظهور اختبارات جديدة

المحتويات

تمهيد

- 1- نشأة القياس النفسي وتطوره قبل أعمال بينيه
 - 2- أعمال بينيه وزملائه في قياس القدرة العقلية
 - 3- تطور حركة القياس النفسي بعد أعمال بينيه
 - 4- أسس القياس النفسي
- خلاصة

المحاضرة الأولى

التطور التاريخي للاختبارات منذ نشأتها إلى الوقت الحالي

تمهيد:

شكلت حركة القياس النفسي نقلة نوعية في مجال علم النفس وعلم النفس المعرفي، لكن مع هذا يتعذر تحديد نقطة زمنية بذاتها لبدايات هذا المجال، غير أنه هناك شبه اتفاق أنها وليدة القرن 19، وأنها جاءت بمنأى عن الجانب التقليدي الاستبطاني كونها ظاهرة جديدة اهتمت بدراسة الظاهرة النفسية من منظور علمي تجريبي، ويمكن تقسيم الخلفية التاريخية للاختبارات النفسية إلى 3 مراحل: نشأة القياس النفسي قبل بينيه، أعمال بينيه وأخيرا تطور حركة القياس النفسي بعده.

1- نشأة القياس النفسي وتطوره قبل (بينيه، Binet):

يكاد يجمع معظم الباحثون أن الدراسات الخاصة بكشف الفروق الفردية في زمن الاستجابة أو ما يسمى بدراسات المعادلة الشخصية جرت خارج إطار علم النفس على يد عالم فلكي منذ القرن الثامن عشر، مهدت بدورها لظهور حركة الاختبارات النفسية والقياس النفسي، وكان قد حدث في عام 1976 أن طرد أحد الفلكيين (Maskelyne) مساعده (Kinne-brook) لأنه أخطأ في رصد أحد النجوم بمقدار ثانية، هذا ما دفع العالم (Bessel) بعد هذه الحادثة إلى دراسة الفروق في تقديرات الفلكيين لحركة النجوم التي أثبتت أن الزمن بين المثير والاستجابة يختلف من فرد لآخر أي اختلاف زمن الرجوع وهو ما سمي لاحقا بمعادلة الشخصية (مقدم، 2010). أخذ بعدها القياس النفسي في تطور آخذا في الأول منحى الفروق الفردية، وفيما يلي أهم المساهمين في تطور هذه الحركة:

1-1- في فرنسا:

تم الاهتمام خلال النصف الأول من القرن التاسع عشر بمسألة الذكاء والتخلف العقلي وبرزت الحاجة إلى اختبارات تميز بين القدرة العقلية للأفراد، كانت بداياتها مع الطبيب الفرنسي (إتارد، Etard) في محاولته تدريب طفل متوحش عثر عليه في غابة. (مقدم، 2010).

وقدم الطبيب جان أسكيرو (Esquirol, 1838) إسهاما في هذا المجال من خلال أعماله في التمييز بين مفهومي المرض العقلي والتخلف العقلي، وكذا مستويات التخلف العقلي (البله والعتة) استنادا على ملامح الوجه ومقاييس الجمجمة، وكذا محك اللغة اللفظية سابقا بذلك عصره، حيث ظهرت اختبارات الذكاء بعده التي تعتمد على القدرة اللفظية في تحديد مستويات الذكاء. (ميخائيل، 2016).

في ذات السياق، كانت أعمال سيجان (Seguin, 1837) مشاهمة لأعمال أسكيرول حيث أنشأ معهدا لتدريب العقول بفرنسا عام 1837، وابتكر لوحة سيجان المشهورة التي تعتبر اختبار أدائي للقدرة العقلية للصح والاسيما الأميين وضعاف العقول، وهي لوحة خشبية محفورة تتيح إدخال أشكال هندسية متنوعة، ومنه اهتمامات كل من (أسكيرول وسيجان) دعت إلى الاهتمام بالفروقات الفردية في القدرات العقلية. (مقدم، 2010).

1-2- في ألمانيا:

النقطة الحاسمة التي شكلت انطلاقة للقياس النفسي هو إنشاء (فونت، Wundt) لأول مخبر لعلم النفس التجريبي في مدينة لايبزيغ في ألمانيا سنة 1879 حيث اتجه هو وزملاؤه إلى استخلاص القوانين العامة التي يخضع لها السلوك البشري دون الاهتمام فقط بالفروق الفردية. (ميخائيل، 2016).

وقد ركزوا بذلك على توفير ظروف فيزيقية مشابهة لأفراد العينة الواحدة وتطبيق بعض الاختبارات الخاصة بالظواهر الحسية، والنفس جسمية كظواهر الإحساس والعتبات الفارقة وردود الفعل، مما شكل ذلك نقلة نوعية في حركة القياس النفسي، حيث يرجع إليه الفضل في وضع أسس المنهج التجريبي في علم النفس. (مقدم، 2010).

1-3- في بريطانيا:

أجرى عالم البيولوجيا الإنجليزي (فرانسيس جالتون Galton، 1911-1982) اهتماما بالوراثة، انصبت أعماله لدراسة صفات التوائم والأقارب والأشخاص الذين لا تربطهم صلة قرابة، وكذا دراسة الموهوبين، وقد أنشأ سنة 1882 مخبرا لعلم الإنسان القياسي يقيس فيه القدرات الحس حركية، وقياس حدة السمع والبصر وزمن الرجوع. (مقدم، 2010).

يعتبر جالتون المكتشف الحقيقي للفروق الفردية، حيث صمم بنفسه مجموعة من الاختبارات مثل قضيب جالتون للتمييز البصري للأطوال، وصفارة جالتون لتحديد أعلى مقام سمعي، وهو أول من استخدم الإحصاء في تحليل الاختبارات، بمجموعة من الأساليب طورها بعده بارسون (Pearson) في و م أ. (ميخائيل، 2016).

1-4- في أمريكا:

أسهمت كذلك أعمال (جيمس كاتل، J.M.Cattel) هو الآخر في و م أ والذي تتلمذ على يد فونت في و م أ في تطوير حركة القياس، ولكنه اهتم بالفروق الفردية وزمن الرجوع، وأنشأ معملا لعلم النفس التجريبي في أمريكا وأنشأ مجموعة من الاختبارات النفسية مثل اختبارات قوة السمع وحدة الإبصار والتمييز بين الأوزان، وسرعة الحركة، والقوة العضلية وزمن الرجوع، حيث أعدها عام 1885، لكنه لم يتمكن من نشرها إلا في عام 1896. (ميخائيل، 2016).

وهذا بسبب معارضة أستاذه فونت (Wundt) له، وكان (كاتل، Catell) أول من استخدم عبارة الاختبار العقلي سنة 1890، حيث حدد الفروق العقلية من خلال المدركات الحسية (سرعة الحركة، حدة الإبصار، قوة السمع التمييز بين الأوزان، الحساسية للألم) وزمن الرجوع. (مقدم، 2010).

هذا ما اعترض عليه بينيه في فرنسا، الذي أسس أول انطلاقة لاختبارات الذكاء من خلال اعتباره أن قياس القدرات العقلية فقط من خلال الإدراك الحسي وزمن الرجوع لا يمكنه قياس القدرات العقلية الكبرى كالتذكر، الانتباه والتفكير. (ميخائيل، 2016).

يتفق أغلب المتخصصين أن فترة قبل الحرب العالمية الأولى كانت من أخصب الفترات التي شهدت تطورا ملحوظا للقياس النفسي، فهي شهدت أعمال (بيرسون، سبيرمان وثورندايك، Pearson Spearman & Thorndike)، حيث وضع بيرسون عام 1901 نظرية الارتباط التي ساهمت في تطور القياس النفسي وعلم

الإحصاء، كما قام بعدها "تشارلز سبيرمان" عام 1904 بوصف نظرية العاملين للقدرات العقلية التي طورت أسلوب التحليل العملي الذي أثر بشكل مباشر في بناء الاختبارات والحكم على صدق المقاييس النفسية. (بركات، 2008).

2- أعمال (بينه، Binet) في قياس القدرة العقلية:

يعتبر (بينه، Binet) وزملاؤه (سيمون وهنري، Simon & Henry) من المساهمين الأوائل في قياس الذكاء، وظهر على إثر أعمالهم مقياس بينه- سيمون في الذكاء سنة 1905، وأضاف عليه تعديلا سنة 1911 يضم بنودا لمستويات عمرية أعلى تصل إلى الراشدين، انتهت إلى نشر المقياس بصورته المعدلة سنة 1916 من طرف (تيرمان وتشايلدز، Terman & Childs) أطلق عليه اسم "مقياس ستانفورد بينه"، وأعيد تعديله من طرف (ميريل، Mirell) عام 1937، ثم عام 1960. (عباس، 1996).

كما كان (أرثر أوتيس، Arthur Otis) قد وضع أول اختبار جمعي للذكاء عام 1914 بالاعتماد على مراجعة (تيرمان، Terman) لمقياس بينه- سيمون التي تم إعدادها في جامعة ستانفورد، وفي نفس السنة ظهرت سلاسل (نوكس، Nox) للاختبارات المعلمية، وكذا اختبارات متاهة (بورتوس، Porteous). (بركات، 2008).

عرف القياس النفسي خلال الحرب العالمية الأولى دفعا جديدا حيث اتجهت أنظار العلماء إلى إنشاء اختبارات خاصة لاقتناء الجنود والضباط الصالحين للخدمة العسكرية كاختبار ألفا للجيش وبيتا للأميين وغير ناطقين بالإنجليزية لفرز المجندين وانتقاء القادة واستبعاد المتخلفين عقليا (مقدم، 2010)، استخدمت على إثرها في الثلاثينات من القرن الماضي اختبارات الذكاء في المدارس والمعاهد والجامعات وللحصول على وظيفة على نطاق واسع في العالم، لاسيما فرنسا، أمريكا والاتحاد السوفياتي. (ميخائيل، 2016).

وشهد عام 1910 حدثا هاما في تاريخ القياس تمثل في ظهور "نظرية القياس" التي وضع أسسها (سبيرمان، Spearman). (بركات، 2008).

طور عام 1939 (وكسلر، Wechsler) مقياس الذكاء للراشدين يولي اهتماما بالجوانب الأدائية على غرار اللفظية، كما وضع مقياسا خاصا بالأطفال من 04-06 سنوات، ومن 05-15 سنة.

وكان لمنهج التحليل العملي دور لا يستهان به في تطور اختبارات الذكاء حيث استخلص (سبيرمان، Spearman) عامل الذكاء العام استنادا إلى التحليل العملي، لكن (ثيرستون، Thurston) رفض هذا التوجه، حيث صمم بطارية القدرات العقلية (PMA) التي تستند على أكثر من عامل في تفسير الذكاء، واكتشف 07 قدرات عقلية أولية هي: الفهم اللفظي - الطلاقة اللفظية - القدرة العددية - الإدراك المكاني - التذكر - السرعة الإدراكية - والاستدلال المنطقي. (أحمد الخطيب و حامد الخطيب، 2010).

3- تطور حركة القياس النفسي بعد أعمال بينه:

3-1- حركة قياس الشخصية:

ظهرت ملامح جديدة للقياس النفسي غير القدرات العقلية بدأت أثناء الحرب العالمية الأولى وانتشرت بعدها، بهدف الكشف عن سمات الشخصية والتمييز بين السواء واللاسواء، من منطلق أن شخصية الفرد وسماته المزاجية قد تؤثر على الأداء العقلي للفرد، من رواد حركة قياس الشخصية (وودورث، Woodworth) الذي وضع أثناء الحرب

العالمية الأولى قائمة لقياس الشخصية طبقت على عديد من المجندين، وظهر سنة 1918 مقياس "وودورث للعصاب" (أحمد الخطيب و حامد الخطيب، 2010)، وقائمة روبرت للسيطرة والخضوع سنة 1928، ثم شكل اختبار (مينوسيتا، MMPI) متعدد الأوجه للشخصية لكل من (هاتاوي وماكينلي، Hathaway & Mckinley) سنة 1943 (ميخائيل، 2016).

كما ظهر أول استبيان مقنن للشخصية سنة 1917 سمي ب "استمارة البيانات الشخصية"، وضعها (روبرت وودورث، Robert Woodworth)، وتم استخدامها في عملية الانتقاء العسكري (بركات، 2008).

الاتجاه الثاني في قياس الشخصية هو الطرائق الإسقاطية التي بدأها الطبيب النفسي السويسري (رورشاخ، Rorschach) منذ 1911، وقد لاحظ أن المرضى يسقطون اتجاهاتهم وعقدتهم الشخصية على أشكال غامضة، وقد وضع اختباره الشهير لبقع الحر سنة 1921، ويستخدم خاصة لأغراض التشخيص الإكلينيكي.

ظهر اختبار إسقاطي آخر عام 1926، تمثل في اختبار (جودانف، Goodenough) لرسم الرجل، وهو اختبار يقيس القدرة العقلية كما يكشف عن شخصية الأطفال. (بركات، 2008).

ومن الاختبارات الإسقاطية كذلك اختبار تقبل الموضوع الذي نشره كل من (موراي ومورجون، Murray & Morgan) سنة 1935، وهو يختلف عن الرورشاخ كون بطاقته تمثل مواقف شبيهة بمواقف الحياة الطبيعية العادية. (ميخائيل، 2016).

3-2- مقاييس الاتجاهات والميول:

ظهر اتجاه آخر لقياس جوانب الشخصية المتعلقة بالاتجاهات والميول والقيم والآراء، وقام (ثيرستون، Thurston) بمحاولات مبكرة في هذا المجال، حيث وضع مع مساعديه نحو ثلاثين سلما للاتجاهات اعتمد من خلالها على آراء المحكمين، تناولت مواضيع خاصة بالمواقف من الحرب، الزواج، الشيوعية، الرأسمالية، الكنيسة، كما وضع (ليكرت، Likert) عديدا من مقاييس الاتجاهات بشكل آخر من خلال تقديم جمل وبدائل للإجابة تعبر عن شدة الاتجاهات. أما عن اختبارات الميول فقد كانت بداياتها سنة 1907 أين وضع (ستالني هول، Stanley Hall) مقياسا لقياس ميول الأطفال حول الاستجمام، تلتها محاولات أخرى أبرزها مقياس (سترونغ، Strong) للميول المهنية سنة 1927، وتم تعديله سنة 1974 حيث ظهر باسم مقياس "سترونغ كامبل" للميول (ميخائيل، 2016).

كما ظهر كذلك مقياسا للقيم سنة 1931 من تصميم كل من (ألبرتوفرنون، Allport & Vernon)، وأشار (أبت، Abt) إلى أن علم النفس بدأ يتجه لوضع مفاهيمه وفروضه في صورة قابلة للاختبار والفحص متماشيا مع أدق الأساليب المنهجية. (أحمد الخطيب و حامد الخطيب، 2010).

وأعد (بركات، 2008) قائمة يبين فيها أدوات القياس النفسي الأكثر استخداما في مجال الإرشاد النفسي نقلا عن كل من (واتكنز، كامبل وماكجروجر، 1988) مبينة في الجدول أدناه:

جدول 01

أدوات القياس النفسي الأكثر استخداماً إلى غاية 1988

الاختبار	رتبة الاستخدام
إختبار الشخصية المتعدد الأوجه (MMPI)	01
مقياس وكسلر المعدل لذكاء الراشدين (WAIS-R)	02
اختبار سترونج- كامبل للاهتمامات	03
استمارات تكملة الجملة	04
اختبار بندر جشطلت	05
اختبار وكسلر المعدل لذكاء الأطفال (WISC-R)	06
اختبار المنزل والشجرة والشخص (HTP)	07
اختبار رسم الرجل	08
اختبار تفهم الموضوع (TAT)	09
اختبار بقع الحبر لرورشاخ	10.5
استبيان عوامل للشخصية الستة عشر (16PF)	10.5
اختبار المدى الواسع للتحصيل	12
مقياس أدوارز للتفضيل الشخصي	13.5
مقياس كور للاهتمام المهني	13.5
اختبار كاليفورنيا النفسي (CPI)	15
مقياس وكسلر للذاكرة	16
مقياس وكسلر لذكاء ما قبل المدرسة والابتدائي (WPPSI)	17
اختبار الاستعدادات الفارقة	18

استمرت الاختبارات تسير تطورات العلمية، وقد ساهمت تطورات الإحصاء والقياس النفسي في تسهيل بنائها وتقنينها حتى تحافظ على موضوعيتها وموثوقيتها وفق البيئة الثقافية التي تطبق فيها.

4-أسس القياس النفسي:

يمكن القول أن القياس النفسي يقوم على أساسين (معمرية، 2013):

4-1-مبدأ الفروق الفردية:

تمس عادة الخصائص المعرفية، الوجدانية والنزوعية، وبالتالي الفروق في خصائص الشخصية المتعلقة بالقدرات، السمات النفسية، والاتجاهات، يشمل هذا نوع الفروق في الخاصية من جهة ودرجة وجودها من جهة أخرى، فهناك فروق في ذات الفرد متعلقة باختلاف خصائصه المعرفية والوجدانية والنزوعية من حيث قوتها وضعفها، وهناك فروق بين الأفراد من حيث اختلافهم في الخاصية الواحدة أو عدة خصائص.

كما ذكر معمريه (2013) أن الحديث عن الفروق الفردية يقودنا بالضرورة إلى الحديث عن الفروق بين الجماعات، وقياسها يضمن فهم سيكولوجية الجماعات وخصائصها، وكذلك الفروق بين المهن وما تتطلبه هذه الأخيرة من قدرات واستعدادات ومهارات، يسهل قياسها لا محالة بغرض الانتقاء والتوجيه الدراسي والمهني.

وبالتالي الفروق الفردية في الشخصية تخص جانبين: الأداء الأقصى (قياس القدرات والاستعدادات)، والأداء النمطي (مثل قياس السمات والاتجاهات).

4-2- مبدأ الكم:

في القياس النفسي تقدر الظاهرة النفسية من خلال تحويل كل ما هو نوعي إلى كمي وعددي، بغرض مقارنة درجات المفحوصين مع أدائهم السابق أو مع مجموعة معيارية أو محك مرجعي تفسر على أساسه نتائج القياس.

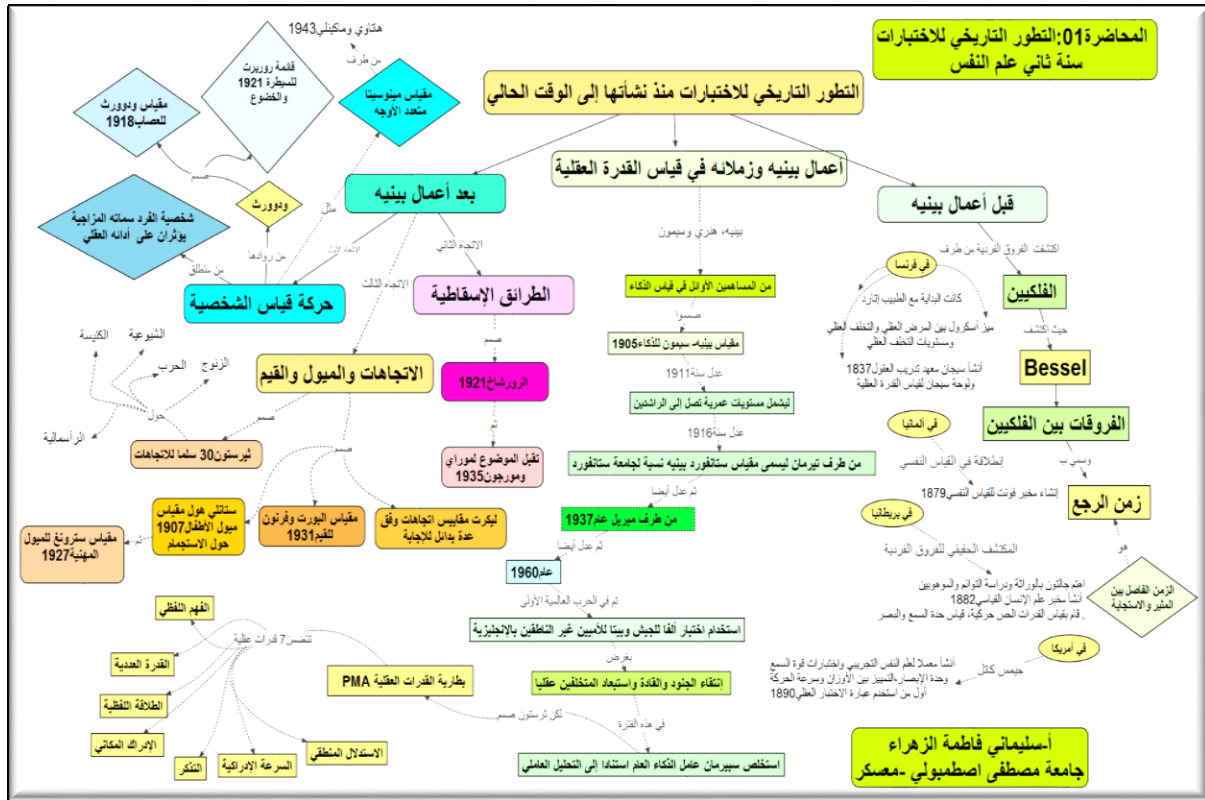
خلاصة:

تطورت حركة القياس النفسي بالبدء أساسا في البحث في الفروق الفردية بين الأفراد في زمن الرجوع، منتقلة بالتدريج إلى تعميم البحث حول القدرات العقلية وجميع جوانب الشخصية لتتطور إلى إعداد اختبارات تحيط بجميع جوانب الشخصية المعرفية منها، الوجدانية والنزوعية، بتطور الحاجة إلى قياسها لفهم أكبر لطبيعة الفرد والجماعة، وبالتالي استثمار أفضل في رأس المال البشري.

الشكل أدناه يلخص النشأة التاريخية للقياس النفسي:

شكل 03

خريطة مفاهيمية لتطور حركة القياس النفسي



الأهداف

في نهاية المحاضرة على الطالب أن يتمكن من أن:

- 1- يتذكر محكات تصنيف الاختبارات النفسية
- 2- يوضح خصائص الاختبارات الموضوعية والاختبارات الإسقاطية
- 3- يوظف مكتسباته في التمييز بين الاختبارات حسب نمطها
- 4- يصنف الاختبارات بناء على محكات مختلفة
- 5- يوظف محكات تصنيف الاختبارات في الحصول على تصنيفات مركبة جديدة لكل نوع

المحتويات

تمهيد

- 1- تعريف الاختبار
- 2- تصنيف الاختبارات النفسية
 - 1-2- الاختبارات الإسقاطية
 - 2-2- الاختبارات الموضوعية
- خلاصة

المحاضرة 02: أنواع الاختبارات

تمهيد:

نلجأ إلى الاختبارات والقياس في حياتنا بغرض تقدير خصائص الأشياء سواء المادية منها كتلك المتعلقة بالأطوال أو الأوزان أو المساحات؛ أو المعنوية كالميل والانطباع، والشخصية والانفعالات والاستعدادات والقدرة على حل المشكلات وأداء المهام، حيث تكون نتيجة الاختبار في أي حالة من هذه الحالات عبارة عن تقدير كمي أو كيفي حول الظاهرة المقصودة بالقياس.

لذا سنتطرق من خلال هذه المحاضرة للمقاييس والاختبارات النفسية وتصنيفها:

1-تعريف الاختبار:

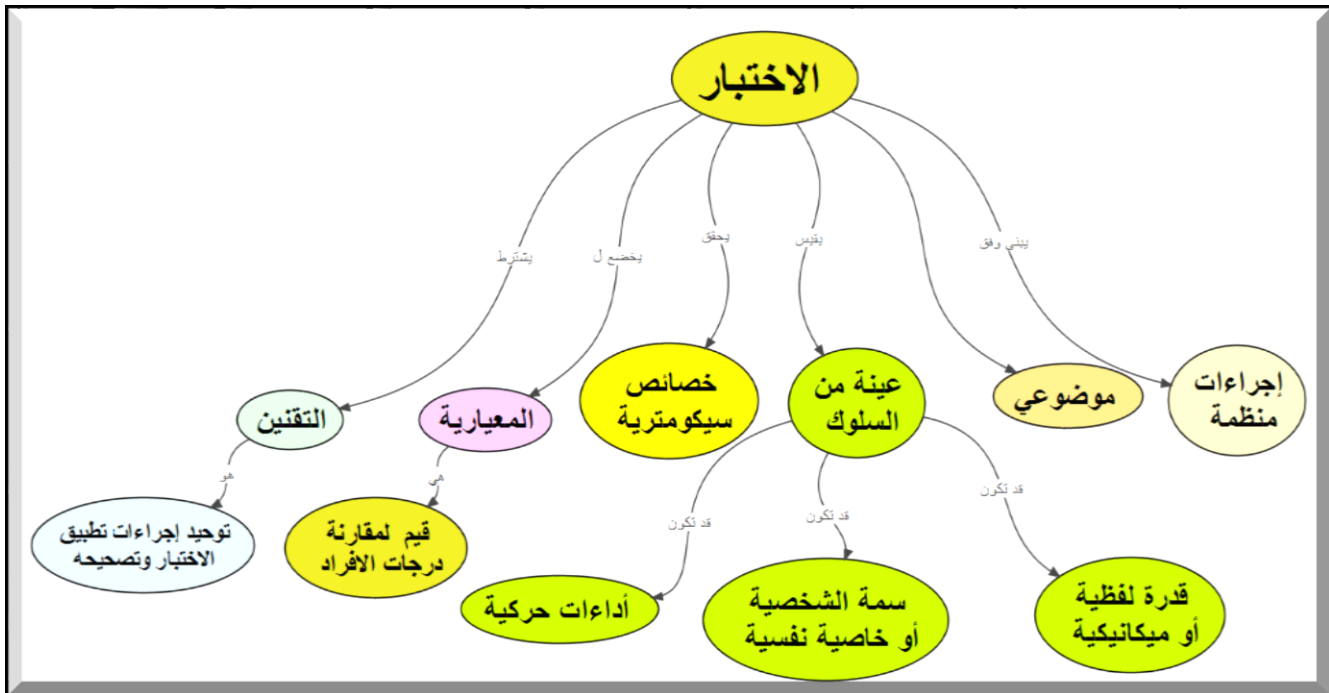
يرتبط بمصطلح الاختبار النفسي مصطلح المقياس ويشتركان في كونهما أداتان لجمع البيانات، يستلزمان خطوات محددة في البناء، كما يشترطان المصدقية والموثوقية، أي شروط الصدق والثبات، غير أن المقياس يعني بالوصف الكمي للسلوك، بينما يمكن للاختبار أن يقدم قياساً كمياً أو نوعياً للسمة المراد قياسها، وعليه فإن الاختبار يتضمن مفهوم المقياس.

يعرف (بين، Bean) الاختبار على أنه "مجموعة منظمة من المثيرات أعدت لتقيس بطريقة كمية أو كيفية بعض العمليات العقلية أو الخصائص النفسية بطريقة كمية أو كيفية". (عبد الرحمن، 2011، ص.13) وتعرف (أنستازي، Anastasi) الاختبار النفسي بأنه: "مقياس موضوعي مقنن لعينة من السلوك." (مجيد، 2014، ص.27).

والموضوعية هنا تعني توحيد تطبيق وتصحيح الاختبار في جميع مجالات وحالات التطبيق. ويعرفه (شيس، Chace 1985) نقلاً عن علام (2001، ص.28) بأنه: "أداة قياس مقننة أو أسلوب منظم يصمم للحصول على قياس موضوعي لعينة من السلوك بهدف موازنة أداء الفرد بمعيار أو بمستوى أداء معينة." يقيس الاختبار النفسي إذن السلوك الإنساني سواء كان سلوكاً ظاهراً قابلاً للقياس المباشر والملاحظة المباشرة، أو سلوكاً كامناً نستدل عليه من خلال مؤشرات السلوك الملاحظة الدالة عليه.

الاختبار يتضمن أيضاً معنى المقارنة، فكرونباخ (Cronbach, 1966) يعرفه على أنه: "طريقة منظمة للمقارنة بين سلوك فردين أو أكثر" (مجيد، 2014، ص.27).

ويمكن تلخيص ماهية الاختبار النفسي حسب ما يبينه الشكل التالي:



2- تصنيف الاختبارات النفسية:

تختلف الاختبارات النفسية باختلاف مرجعيتها النظرية، فمنها ما أسس على مقاربات تحليلية، وأخرى معرفية، وعلى هذا الأساس يمكن تصنيف الاختبارات النفسية إلى اختبارات إسقاطية واختبارات موضوعية.

2-1- الاختبارات الإسقاطية:

هي نوع من الاختبارات الشخصية تقوم على فكرة مؤداها أن الأشكال والمواقف الفارغة تدفع الشخص إلى تكملتها أو توضيحها، وفي حالة صعوبة الأمر تنشط العوامل اللاشعورية لاستكمال العملية، ومنه تتمكن هذه الاختبارات من إخراج جوانب الشخصية اللاشعورية المكبوتة بإسقاطها على موضوع الاختبار، حيث الغرض من هذه الاختبارات هو معرفة هيكل وعمل الشخص، بالإضافة إلى اكتشاف العواطف والصراعات الداخلية، فكلما كانت خصائص الاختبار أكثر غموضاً كلما انعكست شخصية المفحوص، حيث لا يدرك هذا الأخير العلاقة بين إجاباته والعالم الداخلي. (سليمان، 2010).

تعتمد الاختبارات الإسقاطية على الفروق بين الأفراد في إدراكهم للمثيرات الغامضة التي تعكس الجوانب الأساسية للتكوين النفسي، وهذا ما جعل أنزيو وشابير (Anzieu & Chabert, 1983) يحددان المنهجية الإسقاطية في إطار علم النفس الشكلي (المشطلت)، والتحليل النفسي، فإذا كان علم النفس الشكلي يستخدم غموض الأداة كوسيلة لتناول الظروف الخارجية للإدراك وعلاقة الإنسان بالمحيط الخارجي، فإن التحليل النفسي يركز على الغموض ويستعمله كوسيلة للتعرف على الظروف الداخلية والعوامل الكامنة في الشخصية، وبالتالي يمكن القول أن علم النفس الإسقاطي يحاول الجمع بين علاقة الإنسان بالآخرين وعلاقته بذاته. (سي موسى، 2010).

معالجة الظاهرة النفسية من خلال التحليل النفسي يقود بالضرورة إلى الجذور الفرويدية في هذا المجال والمبادئ الأساسية للتحليل النفسي بغية تشخيص دقيق للتوظيف النفسي، ذلك أن التنظيم النفسي غير ثابت، فهو في تطور مستمر متأرجح بين الصحة والمرض اللذان يخضعان بدورهما إلى تعديل في النظام الدفاعي، وهو كذلك يدفع إلى الاهتمام بما وراء النفس الذي ينقل المعطيات السريرية إلى وقائع أخرى تتعلق بالفكر المبدع والأحلام. (Chabert, 2018).

كما تجدر الإشارة أن فهم وتفسير التقنيات الإسقاطية يجب أن يصاحبه نوع من التحفظ الذي دعت إليه دراسات تناولت هذا الموضوع من خلال طبيعة المقابلة أثناء تطبيق التقنية الإسقاطية، ملاحظة السلوك أثناء التطبيق وفي وضعيات أخرى، وكذا في إدراك السلوك المتعلم في الاختبارات المدرسية، من طرف كل من (Pieron, 1952; Chauffard Bornnard, 1949; CF.Nahoun, 1958) على التوالي، حيث أشاروا إلى استعمال التأويلات والتفسيرات مع نوع من الحذر، وهذا لا يعني عدم صحتها، وإنما قد تتأثر بعوامل مختلفة مثلها أثناء التطبيق أو الملاحظة، لذا التحفظ في الحكم على أداء الفرد خلال تطبيق التقنية الإسقاطية يدعو إلى تكملة التشخيص بإقرانها بطرق أخرى أكثر موضوعية للحصول على نتائج دقيقة. (Cambon, 1967). من بين هذه الاختبارات الإسقاطية :

2-1-1- اختبار تفهم الموضوع (Tematic Apperception Test, TAT):

أعدده كل من (موراي ومساعدة مورجان، Morgan & Murray) سنة 1935، يتكون هذا الاختبار من 31 بطاقة تحتوي على أحداث درامية، منها بطاقة واحدة فارغة. يعتبر اختبار تفهم الموضوع أول اختبار مستوحى من تقنية القمص الحرة، حيث أخذت فكرة معرفة الشخص انطلاقاً من أسلوب إنتاجه الفني، ونشر الشكل النهائي للاختبار سنة 1943، ويعني سياق (TAT) مجموع الآليات العقلية الملتزمة بوضعية يطلب فيها من الشخص أن يتخيل قصة انطلاقاً من اللوحة (سي موسى وخليفة، 2010).

2-1-2- اختبار بقع الحبر-رورشاخ (Rorschach):

يتكون من 10 بطاقات تحمل كل منها شكل معين مماثل متشكل من بقعة حبر، يقوم المختبر بعرض الصور على المفحوص حسب ترتيب محدد وفي كل مرة يسلم له فيها البطاقة يطلب منه أن يذكر ما يراه، ويسجل الفاحص زمن الرجوع بين تسلم البطاقة والاستجابة، ثم الزمن المستغرق في كل بطاقة ثم الزمن الكلي للاختبار.

بعد انتهاء المفحوص من جميع البطاقات يقوم الفاحص بتوجيه الأسئلة عن كل بطاقة بدءاً من البطاقة العاشرة نزولاً إلى البطاقة الأولى مع تحديد الاستجابات العامة وطلب تفصيل عنها، كأن يسأل عن جنس الشخص الذي يراه مثلاً، وكذا مكان الاستجابة إن خص به كل اللوحة أو جزء منها.

يسجل الفاحص حرفياً كل أقوال المفحوص، وكذا التوقفات والمشاعر المصاحبة، السرعة واللفظات والتعابير، وتصنف الاستجابات إلى ثلاث أنواع: المكان جزؤه أو كله، المحددات وتشمل خصائص البقعة من شكل ولون وحركة وظلال، المحتوى ويشير إلى مضمون الاستجابة، ويتم التفسير حسب الدرجات الكلية التي يحصل عليها المفحوص وفق التصنيفات المذكورة، لكن التفسير يعتمد على مدى خبرة السيكولوجي وتدريبه. (مقدم، 2010).

2-1-3- اختبارات الرسم:

يشير الرسم إلى اتجاهات نفسية محددة، هناك العديد من الاختبارات النفسية التي تعتمد على الرسم كأسلوب إسقاطي منها اختبار رسم الشجرة للعالم السويسري (كارل كوك، Karl Koch, 1949) والعالمية الفرنسية (رينيه ستورا، Renée Stora)، ورسم العائلة على طريقة (لويس كورمان، Louis Corman)، ورسم الشخص للعالمية الأمريكية (كارين ماكوفر، Karen Machover). (عطية، 1982).

تحتاج الاختبارات الإسقاطية عموماً إلى تدريب مكثف للفاحصين، ومع ذلك فإنه يمكن أن تتأثر الاستجابات بمواقف الفاحص وسياق الاختبار وحتى المشاعر أو الحالة الذهنية التي يمر بها الشخص أثناء تطبيق الاختبار، كما أن مثل هذه الاختبارات تفتقد إلى تعليمات موحدة في التصحيح والتفسير وقد ينتج عنها التحيز في التفسير.

2-2-2- الاختبارات الموضوعية:

سميت بالموضوعية لأنها عكس الإسقاطية تهدف إلى قياس سمة من السمات من خلال السلوك الظاهر وتتحرى الموضوعية في تفسير استجابات المفحوص، وبالتالي لا تتغير النتائج بتغير الفاحص، حيث تقيس جوانب الشخصية المتعلقة بالقدرات، الانفعالات والسمات والجوانب النزوعية كالميول والاتجاهات والقيم، نذكر ما يلي:

2-2-1- اختبارات القدرات:

تهدف اختبارات القدرات إذن إلى قياس قدرة الفرد وموهبته للقيام بنشاط معين، لا تقتصر فقط على قياس مستوى الشخص فحسب وإنما يمكنها ذلك قياس المستوى الذي يمكن أن يصل إليه الفرد الممتحن، ولهذا فهي من الوسائل التي يقاس بها مدى الاستعداد للقيام بعمل ما في المستقبل، وقد تشمل قياس مجموعة من القدرات كالقدرة اللغوية والقدرة الميكانيكية، والقدرة الحسائية، القدرة الكتابية والقدرات الحركية النفسية. (العظيم، 2013).

عادة ما يجمع هذا النوع من الاختبارات في بطاريات بحيث تقيس كل بطارية عدداً من القدرات الخاصة، ومن بين الاختبارات التي تقيس القدرات:

2-2-2- اختبارات الذكاء:

تقيس القدرة العقلية التي تنعكس في سرعة الفهم، القدرة على التعلم، سرعة الإدراك، القدرة على التكيف، ومن بين اختبارات الذكاء اختبار ستانفورد بنيه واختبار وكسلر. (مقدم، 2010).

ومن أهم الاختبارات المتحررة من اللغة في قياسها للذكاء اختبار تكملة الصورة ولوحات الأشكال (Paterson, 1917)، واختبار المصفوفات المتتابعة لرافن 1938. (السيد، 1994).

2-2-3- اختبارات الاستعدادات:

تفيد كمؤشرات تنبئ بما يمكن للفرد أن يقوم به، وهي مطلب أساسي لظهور القدرات، لذا في الواقع العملي هي تقاس في صورة قدرات. (الفاقي، 2005).

2-2-4- اختبارات التحصيل:

هي الاختبارات التي تقيس جوانب الأداء لإظهار ما تعلم الفرد فعلاً أن يقوم به في مجال أو أكثر من مجالات التعلم.

تعدد أنواع الاختبارات التحصيلية ويمكن تقسيمها هي الأخرى إلى أنواع هي حسب ما ورد عن (العظيم، 2013):

أ-اختبارات شفوية:

ب-اختبارات تحريرية: وهي نوعان اختبارات مقالية، يطلب فيها من المفحوص الإجابة في شكل مقال أو فقرة، واختبارات موضوعية تتضمن الإجابة بصح أو خطأ، أو الاختيار من متعدد، أو المزاوجة أو الإكمال.

ج-اختبارات الأداء: يقصد بها قياس أداء الفرد في مجال يتطلب عملاً كتمرين رياضي، تجارب علمية، استخدام أجهزة، أي قياس مهارات تتعلق بالجانب النفس حركي.

2-3-اختبارات الشخصية:

يقصد بها الاختبارات الشخصية التي تقيس التوافق والطباع والمزاج، وكذلك الجوانب الانفعالية من السلوك كمقاييس التوافق الانفعالي التي تعرف بقوائم الشخصية مثل قائمة (مينسوتا المتعددة الأوجه للشخصية، MMPI)، قائمة (جيلفورد وزيمرمان، Guilford & Zimmerman)، ومقاييس السمات كالحضوع والسيطرة والانطواء والانبساط.

من الاستبيانات المعروفة استبيان وودورث للكشف عن النزاعات العصبية، واستبيان قياس العصبية ل (شاير وكاتل، Shier&Cattel). (مقدم، 2010).

2-4-اختبارات قياس الميول:

يقصد بالميل نزعة الفرد نحو تفضيل شيء معين على الأخرسواءً نشاط، مهنة هواية،دراسة، رياضة، واختبارات الميول تقيس اهتمامات الأفراد وميولهم نحو أنشطة أو مهن معينة، ومن بين مقاييس الميول مقياس "سترونج" الذي عدل سنة 1974 وأصبحت تعرف بقائمة "سترونج وكامبل"، واختبار التفضيل لكودر. (مقدم، 2010).

2-5-اختبارات قياس الاتجاهات والقيم:

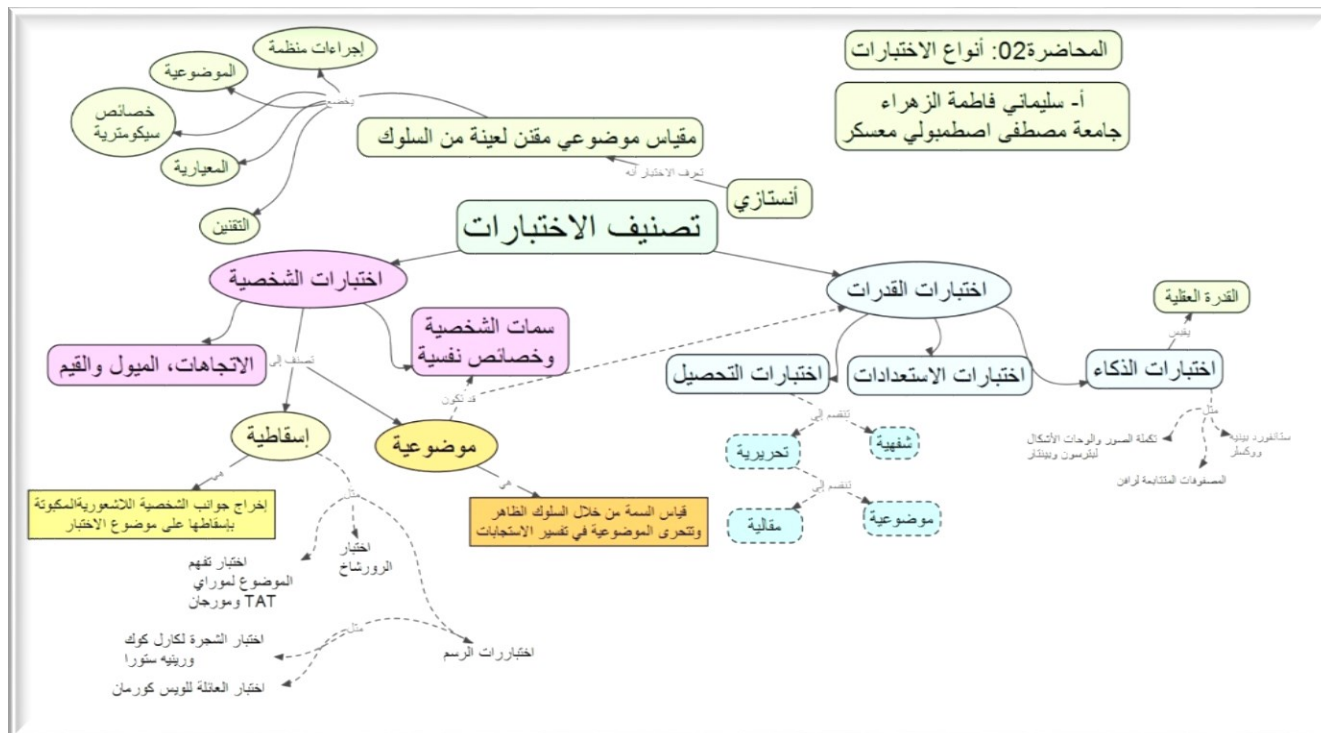
تقيس هذه الاختبارات طبيعة وأبعاد الاتجاهات والمعتقدات إزاء الأفراد أو القضايا والأنشطة المجتمعية، كقياس اتجاهات الطلبة نحو التعليم الإلكتروني، الاتجاهات حول الإصلاحات التربوية، القيم السائدة في المجتمع، الاتجاهات حول التعليم الإلكتروني، ومن بين الاختبارات المعروفة مقياس (ليكرت وثورستون، Likert & Thurston) للاتجاهات ومقياس (ألبورت، Allport) للقيم.

على العموم تشترط هذه الاختبارات الموضوعية توافرها على شروط الصدق والثبات، كما أن نتائجها لا تخلو من خطأ القياس الذي قد يتعلق بعدم فهم التعليم، أو طول زمن الاختبار، أو الحالة النفسية والمزاجية للمفحوص، لذا على الفاحص في تطبيقه لها أن يراعي جميع الجوانب المتعلقة بالاختبار نفسه، بالمفحوص وبالشروط الفيزيائية التي تتعلق بإمكان التطبيق من تهوية وإنارة وتدفئة وغيرها من الشروط المؤثرة.

ويمكن حوصلة تصنيف الاختبارات النفسية في الشكل أدناه:

شكل 05

تصنيف الاختبارات



خلاصة:

تم في هذه المحاضرة تقديم تصنيف للاختبارات النفسية بشقيها الموضوعي والإسقاطي، ولا يمكن الجزم قطعا بتصنيف واحد، كما يمكن لاختبار ما أن يتبع أكثر من تصنيف، فقد ورد عن الباحثين عدة تصنيفات وفق مرجعيات متعددة، نلخص منها ما جاء في (أحمد الخطيب و حامد الخطيب، 2010)، (مقدم، 2010)، (العظيم، 2013)، (عباس، 1996)، (مجيد، 2014)، (ميخائيل، 2016)، (الفقي، 2005) في الجدول التالي:

جدول 02

تصنيف الاختبارات النفسية

طبيعة الاختبار	مرجعية التصنيف
اختبارات إسقاطية: لا يحدد المثير والاستجابة مثل الرورشاخ	درجة تحديد المثير والاستجابة
اختبارات محددة البناء: المثير ومفتاح الإجابة محددان مثل اختبارات التحصيل والاستعداد.	
اختبارات اقصى أداء: اختبارات القدرات والاستعدادات	طبيعة الأداء
اختبارات الأداء النمطي أو الطبيعي: اختبار الذكاء	
اختبارات فردية	عدد المفحوصين
اختبارات جماعية	
اختبارات السرعة: سرعة الإجابة عامل حاسم في تحديد الأداء	سرعة الإجابة

-اختبارات القوة: قدرة المتعلم على الإجابة هي التي تحدد الأداء	طريقة تفسير النتائج
- اختبارات معيارية المرجع: مقارنة أداء المفحوص بأداء أقرانه. -اختبارات محكية المرجع: مقارنة أداء المفحوص بمحك أو مستويات أداء محددة	شكل الفقرة
-المطابقة- الصواب والخطأ- الاختيار من متعدد-التكملة- أسئلة مقالية	الجهة التي تعد الاختبار
-اختبارات مقننة: رسمية يعدها فريق متخصص - غير مقننة: يعدها المعلم مثلاً	كيفية الأداء
-اختبارات لفظية-إجابة تحريرية مكتوبة أو شفوية-اختبارات الأداء المبرهن علمياً	السمات التي تقيسها
-القدرات الاستعدادات- التحصيل - متغيرات الشخصية: التوافق-الطباع-الميول-الاتجاهات والقيم...	

الأهداف

في نهاية المحاضرة على الطالب أن يتمكن من:

- 1- التعرف على ماهية الصدق في الاختبارات النفسية
- 2- يوضح الاختلافات القائمة بين أنواع الصدق
- 3- يربط علاقة بين نوع الصدق ودلالته
- 4- يوظف أنواع الصدق حسب طبيعة السمة المقاسة وخصائص الاختبار
- 5- يتمكن من طريقة حساب الصدق حسب نوعه
- 6- يحرص على تطبيق أمثل للاختبار بمراعاة العوامل المؤثرة في صدقه

المحتويات

تمهيد

- 1- صدق الاختبار
 - 2- أنواع الصدق
 - 1-2 الصدق الخارجي (السطحي)
 - 2-2 الصدق الداخلي:
 - 1-2-2 صدق المحتوى
 - 3-2-2 الصدق المرتبط بمحك
 - 4-2-2 صدق المفهوم (صدق التكوين الفرضي)
 - 3- العوامل التي تؤثر في صدق الاختبار
- خلاصة

المحاضرة 03: معايير تصديق الاختبار

تمهيد:

يُصمم الاختبار لقياس خصائص وصفات معينة في الشخصية، قد تتعلق بالقدرات، المهارات، الاتجاهات، القيم، أو جوانب من الشخصية، وقد يتعلق الأمر بخاصية واحدة في الاختبار ذاته أو بعدة خصائص مثل الأبعاد المركزية للشخصية، وبالتالي فإن بنود الاختبار يجب أن تكون شاملة للإحاطة بجميع الجوانب المقاسة، ويمكن التأكد من ذلك من خلال معايير لتصديق الاختبار.

1- صدق الاختبار:

يمكن تعريف صدق الاختبار على أنه الاطمئنان إلى قياس السمة المقصودة دون غيرها، غير أن هذا الأخير لا يمكن مجال من الأحوال أن يكون مائة بالمائة، فالصدق عموماً أمر نسبي تماماً (Relative)، فمن غير المعقول منطقياً أن نجزم بصدق المقياس تماماً، لأن الظاهرة النفسية ظاهرة غير نقية وغير مستقرة أيضاً، ولا يمكن مجال من الأحوال عزلها عن بقية الظواهر والسمات، إلا أنه من باب القياس نسعى نظرياً إلى تكميمها (غانم، 2007، ص. 119).

كما أن مؤشرات الصدق تتعلق بطبيعة العينة التي طبق عليها الاختبار وخصائص المجتمع الأصلي الذي اشتقت منه معاملات الصدق، فما ثبت صدقه في مجتمعات غربية قد لا يثبت صدقه في مجتمعات عربية.

تجدر الإشارة أيضاً أن الصدق نوعي (Specific)، أي أنه يكون صادقاً لأنه يقيس جانباً وضع لقياسه، حيث لا يقيس جانباً آخر وبالتالي فهو دليل صلاحية أداة القياس لما وضعت لقياسه. (أحمد، 1960).

يجب الأخذ بعين الاعتبار كذلك أن الصدق يختلف من اختبار إلى آخر، فلا يمكن الجزم أن الاختبار صادق أو غير صادق، في حين يمكن اعتباره صادقاً بقدر معين أو درجة معينة، وبالتالي يمكن الأخذ هنا بتعريف (ثورندايك وهاجان، Thorndike & Hagan) للصدق باعتباره تقدير لمعرفة ما إذا كان الاختبار يقيس ما وضع لقياسه فقط مع الإحاطة بجميع جوانب الخاصية ولا شيء غيرها (غانم، 2007)، بمعنى أننا نسعى من خلال التأكد من صدق الاختبار إلى ضمان إمامه بجميع جوانب ومظاهر الخاصية فهو: " يقيس ما نريد أن نقيسه به، وكل ما نريد أن نقيسه به، ولا شيء غير ما نريد أن نقيسه به"، وأنه نجح إلى حد كبير في وصفها دون غيرها من السمات (غانم، 2007، ص. 120).

2- أنواع الصدق:

ظهرت عدة تصنيفات وأدلة تصديق الاختبار، مما دفع إلى تشكيل لجنة مشتركة سنة 1985 تكونت من 3 هيئات هي:

رابطة البحوث التربوية الأمريكية (AERA) (American Education resarsh Association)

رابطة علم النفس الأمريكية (APA) American Psychological Association .

التربوي للقياس التربوي (NCME) National Doucil of Measurement in education.

تم في الاجتماع الذي تم سنة 1985 فحص تعريفات الصدق الموجودة والخروج باتفاق على تعريف شامل للتعريفات المفحوصة، مفاده أن الصدق هو الدليل على الاستنتاجات المستخلصة حول الدرجة على الاختبار، وتوجد 3 أنواع لأدلته هي: أدلة مرتبطة بالمفهوم، أدلة مرتبطة بالمحك، أدلة مرتبطة بالمحتوى، حيث تنطوي جميع أنواع الصدق تحت المجموعات الثلاث هذه مهما تعددت مسمياتها. (بركات، 2008).

ومع ذلك يزخر الأدب النظري بأدلة كثيرة عن صدق الاختبار أو المقياس ستناولها ببعض الشرح كآلاتي:

2-1 الصدق الخارجي أو السطحي (Face Validity):

يسمى أيضا الصدق الظاهري، أي صدق المظهر العام للاختبار، ومدى مناسبة الاختبار للمفحوصين من حيث وضوح تعليماته ودقتها، وارتباط البنود بالخاصية، وانطباق موضوعه على الهدف منه، يعتمد هذا الأخير على التقدير الذاتي على القائمين عليه، وهو ينصب على ما يبدو الاختبار يقيسه وليس ما يقيسه فعلا. (غانم، 2007). يتضح هذا النوع من الصدق بالفحص المبدئي لمحتويات الاختبار، أي بفحص مؤشرات والنظر في مدى مطابقتها لوظيفة الاختبار، فإذا اقترب الاثنان كان ذلك دالا على أن الاختبار صادق صدقا سطحيا، وإذا حدث العكس تستبعد هذه الفقرات.

غير أن الفحص الظاهري للاختبار قد لا يدل دائما على صدقه، ففي الاختبارات الإسقاطية قد تبدو ظاهريا أنها لا تقيس السمة المراد قياسها (يخفى على المفحوصين الغرض من استخدام الاختبار)، وبرغم هذا قد يكون صدق الاختبار مقبولا، فقد يصدق الاختبار علميا وعمليا، في حين قد لا يصدق ظاهريا. (أحمد، 1960). هذا ما يدعو إلى عدم كفايته للتأكد من موثوقية الاختبار وضرورة الاستعانة بطرق تقدير أخرى أكثر موضوعية ودقة، ثم إن اللجنة المشتركة المشكلة بتاريخ 1985 من رابطة البحوث التربوية الأمريكية (American Educational Research Association) AERA، رابطة علم النفس الأمريكية (APA) (National Council of Measurement and Testing) (National Council of Measurement and Testing)، والمجلس الوطني للقياس التربوي (National Council of Measurement and Testing) (National Council of Measurement and Testing) رفضت اعتبار الصدق السطحي نوعا معترفا به من أنواع الصدق (بركات، 2008)، لكن هذا لا يلغي أهميته ولا ينفي ضرورة التقييد به في بناء الاختبار كمؤشر أولي وضمان إضافي للحصول على تقديرات أخرى للصدق تضمن صلاحيته.

2-2- الصدق الداخلي:

يدل على مدى تمثيل أداة القياس للنطاق السلوكي للسمة المراد الاستدلال عليها ويتضمن 03 أنواع من الصدق هي صدق المحتوى، صدق المرتبط بمحك، وصدق التكوين الفرضي.

2-2-1- صدق المحتوى (Content validity):

يسمى أيضا الصدق المنطقي أو صدق المضمون، ذلك أنه يعتمد على المنطق أكثر من اعتماده على الإحصاء. تحقيق صدق المحتوى يعتمد أساسا على تحديد درجة تمثيل بنود الاختبار للمحتوى الشامل أو السمة المقاسة من خلال الفحص المنظم لبنود الاختبار، وأشار عبد الخالق (2000) إلى ضرورة التفريق بين هذا النوع من الصدق والصدق

الظاهري باعتبار هذا الأخير ليس صدقا بالمعنى الفني، فهو لا يشير إلى ما يقيسه الاختبار في الحقيقة، بل إلى ما يبدو أنه يقيسه (ص.132).

يشيع استخدام صدق المحتوى في الاختبارات التحصيلية من خلال تحليل مضمون الاختبار في إطار مدى قياسه للأهداف السلوكية للتعلم في تطابقها مع محتواه من خلال جدول مواصفات.

أشار (عبد الخالق، 2000) أن هذا النوع من الصدق قد يكون مضللا في اختبارات الاستعدادات واختبارات الشخصية، حيث أوضح بركات (2008) أن هذا يرجع إلى أن مجال الاختبار غير محدد أو معروف مثل الاختبارات التحصيلية التي تعتمد على ما تم تدريسه.

لكن لا بد من الإشارة أنه قد يختلف الأمر طبعاً إذا ما توفرت معلومات مفسرة للسمة، من خلال اعتماد التزام نظري معين، أو دليل تشخيص مثلاً، فالاعتبارات الخاصة بصدق المحتوى وكفاءته يجب أن تدخل في المراحل الأولى لبناء الاختبار، حيث يتم الأمر حسب غانم (2007) من خلال صياغة وإعداد فقرات الاختبار بناء على تحليل المجال السلوكي المراد قياسه تحليلاً يتيح الكشف عن عناصره.

كما يمكن كذلك إشراك المختصين في بناء الاختبار منذ البداية أو الاستعانة بمؤشرات ثبت اقترانها كخاصية ببعض المتغيرات النفسية، أو تقدير هذا النوع من الصدق بعد بناء الاختبار مباشرة عن طريق استطلاع آراء المحكمين، على ألا يتعلق الأمر بفحص تطابق الفقرات ظاهرياً مع السمة فقط، وإنما بفحص دقيق لمجالها السلوكي.

على هذا الأساس يمكن لمجموعة من المتخصصين أن يحددوا مدى صلاحية كل بند من بنود الاختبار أو المقياس في قياس السمة أو القدرة المطلوب قياسها، وهذا بعد توضيح السمة أو القدرة توضيحاً إجرائياً، يقوم المتخصصون بعدها بقبول بنود أو رفضها أو اقتراح صياغات جديدة لها، ويمكن حساب هذا النوع من الصدق بتحويل الملاحظات الكيفية إلى كمية بحساب النسبة المئوية لاتفاق المحكمين على الفقرة، كما يحسب كذلك بطريقة "الوشي" (Lawshe) من خلال المعادلة التالية التي وردت في (محمد، 2016، ص.222):

درجة صدق العبارة = م-0.5/ن/0.5، حيث م هي عدد الحكام الذين اتفقوا على صدق العبارة، و"ن" عدد جميع الحكام، كما يمكن كتابة المعادلة بالصيغة التالية: $Cvr = ne - 0.5n / 0.5n$.

يتم مقارنة النتيجة بالحد الأدنى للدرجة المطلوبة على عدد الحكام كما يوضحها الجدول الذي ورد في (محمد، 2016، ص.223):

جدول 03

النسب المئوية لاتفاق المحكمين حسب "الوشي"

عدد الحكام	الحد الأدنى المطلوب لصدق العبارة
05	0.99
06	0.99
07	0.99

0.75	08
0.78	09
0.62	10
0.59	11
0.56	12
0.54	13
0.51	14
0.49	15
0,42	16
0.37	17
0.33	18
0.31	19
0.29	20

من النماذج المبكرة للاستخبارات متعددة الأبعاد التي استخدم في تأليفها صدق المحتوى كل من قائمة (بييل، Bal) للتوافق والتي تقيس توافق الفرد في مجالات منزلية واجتماعية وانفعالية، وقائمة (بيرنرويتز، Pernreuter) للشخصية (BPI) والتي تقيس سمات الانطواء والعصابية، وكذلك هو الحال بالنسبة لقائمة (موني، Moony broblem-cheeklist) للمشكلات حيث تعتبر المثال الأحدث لاستعمال صدق المحتوى في إعدادها سنة 1950، حيث تم اشتقاق بنود للقائمة من سجلات الحالات وجلسات المقابلة الإرشادية والتقارير المكتوبة عن المشكلات الشخصية لـ 4000 شخص لطلاب المدارس العليا. (Freeman, 1962).

2-2-2- الصدق المرتبط بالمحك:

نتعرف من خلال دليل صدق المحك على مدى تماثل الاختبار أو المقياس المقصود بالدراسة أو البناء مع محك مماثل، إذا ما تحقق ارتباط مرتفع بين الاختبار والمقياس المحكي الذي يعتبر بدوره المعيار الذي نقارن به الاختبار (بركات، 2008)، وهو بدوره يتضمن نوعين من الصدق:

أ- الصدق التنبؤي (Predictive Validity):

هو اختبار مدى قدرة الاختبار النفسي على التنبؤ بنتيجة معينة في المستقبل، من خلال مقارنة درجات المفحوصين على الاختبار بسلوكهم التالي على الاختبار، حيث يشكل دليلا للصدق التنبؤي إذا ما اتفقت النتائج المتحصل عليها مع النتائج المتحصل عليها. (غانم، 2007).

تتم المقارنة من خلال محك يدل على أدائهم في المستقبل، حيث الاتفاق بين درجات الاختبار ودرجات هذا المحك مؤشر على مدى قدرة الاختبار على التنبؤ بنتائج المحك الذي ينبغي أن يكون مقياسا أثبت صدقه سابقا. (محمد عبد الحفيظ و حسين باهي، 2002).

يدل هذا النوع من الصدق على مدى كفاءة درجات الاختبار في التنبؤ بالسلوك في المستقبل بفواصل زمني لا يقل عن 06 أشهر بين تطبيق الاختبار الأول وتطبيق الاختبار الثاني المتمثل في المحك الذي يقيس السلوك المتنبئ به. (مراد و أمين، 2005).

ومثال ذلك أن نطبق الاختبار الأول الخاص بالذكاء مثلا على مجموعة طلبة يتبين حصولهم على درجات مرتفعة على الاختبار، مما يمكن أن نتنبأ بارتفاع تحصيلهم الدراسي مستقبلا، وبعد مدة معينة، نطبق اختبارا تحصيليا، فإذا كانت نتائج التحصيل الدراسي مرتفعة، فهذا دليل على صدق اختبار الذكاء الذي تنبأنا من خلاله بارتفاعه في نتائج التحصيل الدراسي.

أو أن نطبق مقياسا للمفارقة بين القيم المتصورة والقيم الواقعية على مجموعة أفراد، فتظهر النتائج وجود مفارقة كبيرة بين ما يتصورونه من قيم وبين القيم التي يمارسونها كسلوك فعلي، ومنه يمكن أن يتنبأ الاختبار السابق بتغير في القيم لدى أفراد العينة مستقبلا، وبعد مدة زمنية، يتم تطبيق مقياسا لتغير القيم، فيتبين صدق تنبأ الاختبار الأول، أي تغير القيم لدى أفراد العينة، وهذا دليل على الصدق التنبؤي لمقياس المفارقة القيمية.

ب-الصدق التلازمي (Concurrent Validity):

هو عبارة عن درجة ارتباط الدرجات التي يتحصل عليها المفحوص في اختبار أو مقياس ما مع درجاته على اختبار موضوعي آخر ثبت صدقه وطبق في نفس وقت إجراء الاختبار. (محمد عبد الحفيظ و حسين باهي، 2002). إذن شرط الفاصل الزمني وارد هنا، فلا يجب أن يكون الفاصل بين تطبيق الاختبار والمحك كبيرا، وإلا اختلفت النتائج بسبب أثر التدريب، التعلم أو أثر التخمين.

إجرائيا يتم تطبيق الاختبار الأول، ثم تطبيق الاختبار المحك يقيس نفس السمة في حينه أو بعد مدة قصيرة، وتحسب درجة الارتباط بين الاختبارين، فإذا كان الارتباط موجودا دل ذلك على تحقق الصدق التلازمي، وإلا العكس.

مثال ذلك كأن نطبق اختبار وكسلر للذكاء، ثم اختبار بينيه للذكاء تلازما مع الاختبار الأول.

كان هذا عن الصدق المرتبط بالمحك، لكن ما الفرق بين الصدق التنبؤي والصدق التلازمي؟

- تعتمد كلا الطريقتان على محك خارجي ثبت صدقه وثباته، غير أن الفاصل الزمني

- في الصدق التلازمي يجب أن يكون قصيرا أو في نفس الوقت، أما في الصدق التنبؤي فالفاصل الزمني طويل لا يجب أن يقل عن 06 أشهر للحصول على درجات المحك. (مراد و أمين، 2005).

- من حيث طبيعة القياس والجانب المنطقي في القياس، يمكن التفريق بين كل من الصدق التلازمي والصدق التنبؤي من حيث أن هذا الأخير يهدف إلى قياس صلاحية الاختبار في التنبؤ بدرجات الأفراد في المستقبل، أما الصدق التلازمي فهو يهدف إلى قياس صلاحية درجات الاختبار لاسيما إذا ارتبطت بدرجات المحك الذي يجب أن يكون مماثلا، أي يقيس نفس السمة. (مراد و أمين، 2005).

مع هذا يثير الصدق التلازمي من الناحية المنهجية مشكلة معاملات الارتباط بين درجات الاختبارين، حيث ارتفاعها أو انخفاضها لا يقصد به السببية، وإنما مجرد علاقة اقتران، بينما الصدق التنبؤي أقرب للسببية، من حيث أن نتائج الأفراد على فقراته، تمكن من التنبؤ بنتائج مستقبلية. (غانم، 2007).

2-2-3- صدق التكوين الفرضي (Construct Validity) :

يطلق عليه أحيانا صدق المفهوم، وهو مدى قياس اختبار لتكوينات فرضية معينة، أي قياس الخصائص المجردة التي لا يمكن الاستدلال عليها بشكل ملموس والتي تسمى بالعوامل الكامنة، إلا من خلال مجموعة من المؤشرات المقاسة التي تفسرها وتعكس السلوك المصاحب لها كقياس الدافعية، الضغط والاكتئاب وغيرها من التكوينات من خلال مؤشرات تفسرها وتدل عليها.

لا يمكن التأكد من صدق التكوين الفرضي إلا بعد تجريب الاختبار على عينة من المفحوصين، وإجراء تحليل عاملي لتبيان عدد العوامل أو التكوينات الفرضية التي تفسر من خلال مؤشرات مقاسة. لا يمكن بالضرورة تفسير درجات الاختبار كقياس لتكوين فرضي واحد، بمعنى انتماء الفقرات لعامل واحد فقط، فأي اختبار يحتوي عادة على عدة عوامل. (مراد و أمين، 2005).

فمقياس الدافعية للإنجاز لـ "عبد اللطيف خليفة" مثلا يحتوي على خمسة تكوينات فرضية هي: الشعور بالمسؤولية، السعي نحو التفوق والنجاح، المثابرة، الشعور بأهمية الزمن والتخطيط للمستقبل، وهي بالتالي تشكل عوامل كامنة، يتم قياسها من خلال مؤشرات تفسرها.

يجمع صدق التكوين الفرضيتين التحليل المنطقي والتحليل الإمبريقي للاختبار، بمعنى يمكن الاعتماد في قياسه عن طريق: **أ- مؤشرات فمائية:** مثلما هو الحال في اختبار نسبة الذكاء التي من المفروض تنمو وتتطور مع الوقت، فإذا كانت درجات الاختبار تتزايد مع السن، يمكن اعتبار الاختبار صادقا. (غانم، 2007).

ب- التحليل العاملي (Factor Analysis):

تم تطوير هذا الأسلوب لتحديد السمات النفسية، يهتم أساسا بتحليل العلاقات الداخلية للبيانات الخاصة بالسلوك، حيث يسعى إلى الكشف عن عدد صغير من المتغيرات الكامنة غير المشاهدة، حيث تشعب على كل متغير كامن مجموعة من المتغيرات المقاسة أو الملاحظة.

فالعوامل هي متغيرات كامنة افتراضية مشتقة من تحليل بيانات مجموعة متغيرات تم قياسها قياسا مباشرا، ويقصد بالمتغيرات المقاسة فقرات استبيان أو اختبار أو مقياس بحيث تمثل كل فقرة متغيرا. (تيغزة، 2012).

وهو يتضمن كذلك استخدام شكل من أشكال الاتساق الداخلي (internal consistenay) بهدف تحديد ما إذا كانت كل بنود الاختبار ترتبط بدرجات المفحوص الكلية، وكذلك مدى ارتباطها بدرجة كبيرة مع بعضها البعض، أو عدم ارتباطها تماما أو بدرجة منخفضة مع مجموعة بنود أخرى. (عبد الخالق، 2000).
ينقسم التحليل العاملي إلى نوعين أساسيين:

ب-1- التحليل العاملي الاستكشافي (Exploratory factor Analysis):

يمكن التحليل العاملي الاستكشافي من اشتقاق متغيرات كامنة أو عوامل تعكس البنية العلائقية المشتركة بين عدد كبير من المتغيرات المقاسة، فإذا كان الباحث يتعامل مع فقرات، فهي متغيرات مقاسة مباشرة يتم اختزالها إلى عوامل كامنة تشكل القاسم المشترك بين هذه الفقرات من خلال هذا النوع من التحليل (تيغزة، 2012)، الأمر الذي يسمح بجمع شتات الدراسة والتركيز أكثر على المتغيرات الكامنة بدل العدد الكبير من المؤشرات أو المتغيرات المقاسة.

في التحليل العاملي الاستكشافي لا يفترض الباحث بنية عاملية بذاتها، وإنما يقوم باكتشافها بعد الانتهاء من التحليل العاملي، حيث سيظهر له عددا من العوامل الكامنة التي تفسرها عددا من المؤشرات أو المتغيرات المقاسة. يهدف التحليل العاملي الاستكشافي حسب ما ورد عن (تيغزة، 2012) إلى اختزال عدد كبير من المتغيرات المقاسة إلى عدد قليل من المتغيرات الكامنة، والتي تفسر العلاقة التي تجمع بين المتغيرات المقاسة، حيث تحلل معاملات الارتباط الناتجة عامليا لتحديد أي من البنود تتجمع معا بتشعبات مرتفعة كي تكون وتساهم في تفسير عامل، أو تقسم في مجموعات أخرى لتقيس سماتا بعينها، وفي هذه الحالة يتوجب الربط بين التحليل الكمي للبيانات والتحليل الكيفي بعد فحص المحتوى طبعاً في كل تجمع للبنود بغرض تسمية العوامل وتحديد السمات والخصائص التي تفسر من خلالها.

ب-2- التحليل العاملي التوكيدي (Confirmatory factor Analysis):

على عكس التحليل العاملي الاستكشافي الذي يفترض الباحث فيه سلفاً لتصوير واضح لعدد العوامل وطبيعتها، لا ينطلق الباحث في مثل هذا النوع من التحليل من متغيرات مقاسة (مؤشرات أو بنود أو فقرات)، بل من بنية عاملية، حيث يفترض مسبقاً عدداً من العوامل تتشعب عليها مجموعة من المؤشرات المقاسة، أو ينطلق من بنية عاملية تم اكتشافها من خلال التحليل العاملي الاستكشافي ليتم التأكد من صحتها ومن تشعب الفقرات على عواملها الكامنة. (تيغزة، 2012).

ورد عن تيغزة (2012) كذلك أن المنظور التوكيدي أو التثبتي يقصد به أن الباحث لا يمارس التحليل الإحصائي على البيانات التي يجمعها إلا بعد بناء نموذج نظري، ويستعمل استراتيجية التحليل الإحصائي التي تنتمي إلى المعادلات البنائية للتأكد من مطابقة النموذج المفترض للبيانات التي جمعت من عينة تمثل المجتمع. وبالتالي فالتحليل العاملي التوكيدي هو أسلوب إحصائي ينتمي إلى النمذجة بالمعادلات البنائية.

ت-الصدق التقاربي والصدق التمييزي (Canvergent and Discriminant Validity):

يقصد بالصدق التمييزي قدرة الاختبار على التمييز بين الأفراد الذين يتمتعون بدرجة مرتفعة من الصفة أو السمة، وبين من يتمتعون بدرجة منخفضة من نفس السمة، حيث يتم تقسيم الأفراد إلى مجموعتين: مجموعة تحصلت على درجات مرتفعة على مقياس هذه السمة، والأخرى على درجات منخفضة على مقياس السمة، ثم تتم المقارنة بين المجموعتين باي اختبار إحصائي مناسب لدلالة الفروق بين المجموعتين (محمد عبد الحفيظ و حسين باهي، 2002)، ويسمى هذا النوع من الصدق في الأدب النظري بصدق المقارنة الطرفية (Comparaison of extrem groups). تجدر الإشارة هنا أنه ما يقصد بالمقارنة الطرفية هو استخدام محك آخر يصنف الأفراد إلى مستويات في السمة المقصودة بالمقياس، وبحث الفروق بين درجات الاختبار لمجموعات المستويات المختلفة، وتدلل الفروق حينها على الصدق التجريبي، مثال ذلك صلاحية اختبار الذكاء على التمييز بين الفروقات الفردية وفق محك خارجي. (مراد و أمين، 2005).

أما الصدق التقاربي فيقصد به ارتباط المقياس بشكل جيد مع مقاييس أخرى يعتقد أنها تقيس نفس المفهوم، حيث نحصل في هذه الحالة على الدليل التقاربي للصدق. (بركات، 2008).

يمكن التأكد من الصدق التقاربي أيضا في التحليل العاملي التوكيدي من خلال اختبار تشيع فقرات المقياس على عواملها الكامنة، إذ يعتبر مقدار تشيع المؤشر المقاس على عامله معامل صدق، ويمكن تقدير الصدق على مستوى المؤشرات التي تقيس العامل الواحد من منظور الصدق التقاربي والصدق التمايزي، حيث أن القيم المرتفعة لتشيعات الفقرات على عاملها تعتبر مقياسا للصدق التقاربي، بينما معاملات الارتباط غير المرتفعة أو المعتدلة بين العوامل الكامنة أو الأبعاد دليل على الصدق التمايزي، بمعنى تفرد كل عامل بتفسير تباين مؤشراتته المقاسة. (تيغزة، 2012).

من الجدير بالذكر أنه يوجد تشابه بين الصدق المرتبط بالمفهوم الذي يهتم بالصدق التقاربي والصدق التمايزي إذا بحثنا كدليل لمثل هذا النوع من الصدق في ارتباطات الاختبار الأول باختبار آخر يشبهه، وبين الصدق المرتبط بالمحك، فكلاهما يشتركان في وجود محك آخر ثبت صدقه وثباته، وعلى أساسه يمكن مقارنة نتائجه بنتائج الاختبار الأول، لهذا هناك اتجاه جديد يعرف الصدق على أنه دليل واحد، والفئات الأخرى كالارتباط بالمحك أو الصدق التقاربي، ماهي إلا فئات فرعية لدليل الصدق. (بركات، 2008).

في هذا السياق ذكرت (أنستازي و أرينا، 1997) أن الصدق مفهوم كامل يضم عدة إجراءات تندرج تحت مفهوم شامل هو "صدق المفهوم"، أما أدلته فهي ثلاث (بركات، 2008):

- إجراءات وصف المحتوى (Content-Description Procedures).
- إجراءات التنبؤ بالمحك (Criterion-Prediction Procedures).
- إجراءات تحقيق المفهوم (Construct-Identification Procedures).

ث - صدق الاتساق الداخلي (Internal Consistency):

يتم تقدير هذا النوع من الصدق من خلال حساب معامل الارتباط بين درجة كل فقرة أو بند من بنود الاختبار مع الدرجة الكلية للبعد، والدرجة الكلية للاختبار. (محمد عبد الحفيظ و حسين باهي، 2002). تدل معاملات الارتباط على أن مكونات البنود أو أن المؤشرات تقيس شيئا مشتركا مما يدل على صدق البناء الداخلي. (مراد و أمين، 2005).

3- العوامل التي تؤثر في صدق الاختبار: حسب ما ورد عن (مجيد، 2014):

3-1- عوامل تتعلق بالاختبار:

* صعوبة الأسئلة أو سهولتها (الأُن تكون فوق أو دون مستوى أفراد عينة الدراسة).

* صياغة الأسئلة: الأسئلة المطروحة أو بنود الاختبار توحى بالإجابة.

* غموض الأسئلة أو البنود حيث يمكن أن تفسر أو تفهم بطريقة مختلفة.

* طول الاختبار.

* ثبات الاختبار، فالثبات شرط أساسي في الصدق، والاختبار الغير ثابت لا يمكن أن يكون صادقا.

3-2- عوامل تتعلق بأدلة الصدق المستعملة (مراد و أمين، 2005):

* طريقة حساب معامل الصدق، ومدى ثبات المحك المستخدم في حساب الصدق، فكلما كان ثابتا وصادقا، كلما زاد ذلك من صدق الاختبار.

*تباين الدرجات: زيادة تباين درجات الاختبار المطبق على عينة التقنين دالة على اتساع مجال السمة المقاسة وتغطية شاملة لها مما يؤدي إلى ارتفاع الصدق والثبات.

خلاصة:

البحث في صدق الاختبار هو دليل على ثقتنا في أن الاختبار يقيس فعلا ما وضع لقياسه ويزخر الأدب النظري بطرق شتى لتقدير صدق الاختبار، ولكن يمكن جمعها في ثلاث أدلة أساسية للصدق هي: الصدق المرتبط بالمحتوى أو الصدق المنطقي أو صدق المضمون، ويتم من خلال الفحص فقرات الاختبار ومدى تفسيرها للسمة المقاسة خارجيا دون القيام بإجراءات أو تطبيقات إحصائية، وهو يدل على ما يبدو أن الاختبار يقيسه، وليس ما يقيسه فعلا، ورغم أن اللجنة المشتركة الأمريكية سنة 1985 لم تقر بهذا النوع من الصدق إلا أنه لا يمكن الاستغناء عنه كمرحلة أولية في بناء الاختبار وتكوين مفرداته، يمكن الاستدلال عليه عن طريق استطلاع رأي مجموعة خبراء أو ما يسمى عادة بصدق المحكمين الذين يفحصون الاختبار من حيث صياغة مؤشرات و مدى تفسيرها للخاصية، يمكن حساب نسبة اتفاق المحكمين على الفقرات من خلال النسب المئوية أو معادلة لوشي لحساب معامل صدق المحكمين.

النوع الثاني لأدلة الصدق هو الصدق المرتبط بالمحك الذي يشترط وجود محك خارجي ثبت صدقة وثباته للمقارنة بين نتائجه ونتائج الاختبار الأول، أو يمكن حساب معاملات الارتباط بين الاختبارين، فكلما كانت مرتفعة، كلما دل ذلك على صدق الاختبار، وهو ينقسم إلى نوعين: الصدق التلازمي والصدق التنبؤي.

الدليل الثالث للصدق هو الصدق المرتبط بالمفهوم أو صدق التكوين الفرضي، وينطوي على عدة أنواع صدق هي صدق الاتساق الداخلي، الصدق العاملي، الصدق التقاربي والصدق التمايزي.

هناك اتجاه جديد يعرف الصدق على أنه دليل واحد، والفئات الأخرى كالارتباط بالمحك أو الصدق التقاربي ما هي إلا فئات فرعية لدليل الصدق.

الأهداف

في نهاية المحاضرة يجب على الطالب أن يكون قادرا على:

- 1- التعرف على ماهية الثبات في الاختبارات النفسية
- 2- يوضح الاختلافات القائمة بين طرق تقدير ثبات الاختبارات النفسية
- 3- إدراك العلاقة بين الصدق والثبات
- 4- يوظف طرق تقدير الثبات حسب طرق تقدير الصدق المعتمدة وحسب طبيعة الاختبار وخصائصه.
- 5- يتمكن من طرق حساب الثبات على اختلافها
- 6- يحرص على تطبيق أمثل للاختبار بمراعاة العوامل المؤثرة في ثباته.

المحتويات

تمهيد

- 1- تعريف الثبات
 - 2- طرق قياس ثبات الاختبار
 - 1-2 طريقة إعادة الاختبار
 - 2-2 طريقة التجزئة النصفية
 - 3-2 طريقة حساب معامل الفا كرونبيخ
 - 4-2 طريقة حساب معادلة كودر ريتشاردسون
 - 5-2 طريقة الصور المتكافئة
 - 3- العوامل المؤثرة في ثبات الاختبار
- خلاصة

المحاضرة 04: ثبات الاختبار

تمهيد:

لا يمكن استعمال مقياس نفسي أو اختبار نفسي بعد التأكد من صدقه والوثوق بقدرته على قياس ما صمم لقياسه فقط، وإنما نحتاج على غرار ذلك إلى التأكد من دقته واستقرار نتائجه وبالتالي ثباتها.

1-تعريف الثبات (Reliability):

ورد في موسوعة علم النفس والتحليل النفسي (1993) أن الثبات هو مدى إعطاء الاختبار نفس الدرجات لنفس الأفراد عند إعادة تطبيقه عليهم، مما يدل على استقرار نتائجه (غانم، 2007، ص.101).
وبالتالي إذا كان صدق الاختبار يدعو إلى الثقة في قياس الاختبار لجميع جوانب السمة المقصودة بالقياس، فإن الثبات دليل على استقرار النتائج المتحصل عليها مع إعادة التطبيق، وبالتالي فهو دليل على الدقة.
أشار علام (2000) أنه متى كانت أخطاء القياس خالية من الأخطاء العشوائية وقادرة على قياس المقدار الحقيقي للسمة في ظروف مختلفة ومتباينة كان المقياس عندئذ ثابتاً، وبالتالي الثبات هو الاتساق والدقة في قياس ما وضع المقياس لقياسه، والثبات إحصائياً يمثل النسبة من التباين الصحيح في الدرجات المستخلصة من الاختبار حسب ما تبينه المعادلة التالية التي وردت في (مجيد، 2014، ص.130):

ثبات الاختبار = التباين الحقيقي / التباين الكلي

حيث التباين الحقيقي = التباين الكلي - تباين الخطأ

حيث التباين الحقيقي (True variance) هو التفاوت بين درجات المفحوصين الذي يرجع إلى السمات الحقيقية والخصائص المستقرة لديهم.

أما تباين الخطأ (Error variance) فهو التفاوت بين درجات المفحوصين الراجع إلى عوامل الصدفة، والظروف المحيطة بموقف تطبيق الاختبار وما يحيط به من شروط الفيزيائية، أو بالمفحوص وحالته الصحية أو المزاجية أثناء التطبيق. (عبد الخالق، 2000، ص.124).

تجدر الإشارة إلى أنه كلما ارتفع التباين الحقيقي كلما ارتفع ثبات الاختبار، والعكس كلما كثر تباين الخطأ كلما كان الثبات أقل والاختبار غير دقيق.

2-طرق قياس ثبات الاختبار النفسي:

يحسب معامل الثبات إما عن طريق إعادة تطبيق الاختبار مرة ثانية وحساب الارتباطات بين درجات التطبيق الأول والثاني، أو إما عن طريق حساب العلاقة بين مجموعتين من درجات الاختبار على نفس العينة، أو عن طريق استخدام صور متكافئة أي تقيس نفس الخاصية والبحث في استقرار نتائج الاختبارين، أو عن طريق حساب مدى اتساق إجابات الاختبار من تطبيقه مرة واحدة. (مراد و أمين، 2005).

ويمكن أن نذكر جميع هذه الطرق بشيء من التفصيل:

2-1- طريقة إعادة الاختبار (Test- Retest):

تقوم هذه الطريقة على إجراء الاختبار النفسي على مجموعة أفراد عينة ما، ثم إعادة تطبيقه على نفس العينة بعد فترة زمنية من أسبوعين إلى ستة أسابيع، ثم حساب معامل الارتباط بين درجات الاختبار في مرتي التطبيق. (مراد و أمين، 2005).

غير أن هذه الطريقة قدمت لها بعض الانتقادات التي مفادها أن إعادة تطبيق الاختبار بعد مدة زمنية قصيرة تتأثر بأثر التدريب أو التعلم، بمعنى أن الفرد المختبر قد يتذكر إجاباته في المرة الأولى ويكررها، أما إذا تم تفادي عامل الذاكرة من خلال تطويل المدة الفاصلة بين التطبيقين، نفع في هذه الحالة في مشكلة عامل النمو، بمعنى أنه قد تطرأ تغييرات على السمة المقاسة بحكم طول المدة الفاصلة بين الاختبارين، لذا تبعا لبعض البحوث التجريبية تم تحديد المدة الفاصلة بين مرتي التطبيق بأسابيع لصغار السن، وألا تتجاوز ستة أشهر للبالغين، وتبقى طبيعة الظاهرة المقصودة بالدراسة أمر فاصل في تحديد هذه المدة الزمنية (غانم، 2007)، كما لا يمكن الحصول على نسبة ثبات دقيقة مع السمات المعروفة تغيرها بسرعة كنتلك التي تخص التعلم مثلا.

2-2- طريقة التجزئة النصفية (Split-half Method):

في هذه الطريقة يقدم الاختبار مرة واحدة، ثم يتم تقسيم الاختبار إلى نصفين يصحح كل منهما لوحده، ثم يقدر الثبات عن طريق حساب معامل الارتباط بين نصفي الاختبار. (بركات، 2008).

يقسم الاختبار بعدة طرق منها ما يتم بناء على درجات المفردات الفردية ودرجات المفردات الزوجية، ثم يحسب الارتباط بين المجموعتين، ومنها ما يتم بناء على تقسيم فقرات الاختبار إلى نصفين.

يجب إدراك أن الثبات الناتج عن الارتباط بين المجموعتين يخص فقط نصف اختبار، لذا يجب تصحيح معامل ثبات الاختبار باستخدام عدة طرق منها:

*معادلة سبيرمان براون (Spearman- Brown):

معامل الثبات بعد التصحيح = معامل الثبات الحالي $\times 2$ / معامل الثبات الحالي + 1

(غانم، 2007، ص. 109)

*معادلة رولون المختصرة (Rulon):

باستخدام هذه المعادلة، يمكن الاستغناء عن حساب معامل ثبات نصف الاختبار، بالاعتماد مباشرة على تباين فروق نصفي الاختبار حسب ما ورد في (غانم، 2007، ص. 109) :

معامل الثبات بعد التصحيح = $1 -$ تباين فروق درجات النصفين / تباين درجات الاختبار

*معادلة جتمان (L.Guttman):

تعمل هذه المعادلة إلى التغلب على مشكل عدم تساوي الانحرافات المعيارية لكل من نصفي الاختبار، وهي تصلح سواء تساوت الانحرافات المعيارية أو اختلفت حسب المعطيات الواردة في (غانم، 2007، ص. 111):

معادلة الثبات = تباين درجات الأسئلة الفردية + تباين درجات الاسئلة الزوجية / تباين درجات الاختبار

2-3- حساب الثبات عن طريق معامل الفا كرونباخ (CronbachAlpha Coefficient):

يتيح هذا المعامل الحصول على أقل قدر من الثبات يمكن توقعه، حيث بعد الحصول على قيمة مرتفعة له، يمكن افتراض أن ثبات الاختبار لن ينزل عن مستوى ألفا، ومنه يظل الثبات الفعلي مرتفعاً، ويمكن حسابه كما يلي (مجيد،،2014،ص.157):

$$r = \frac{[N] [1 - \sum C^2]}{[K] [1 - \sum S^2]}$$

حيث S^2 تشير إلى تباين البنود الفردية $S^2 =$ التباين الكلي للاختبار

يستعمل هذا النوع من تقدير الثبات في الاختبارات والمقاييس ذات بدائل الإجابة المتعددة، بينما الاختبارات والمقاييس التي تتضمن بديلين فقط (صح وخطأ أو نعم ولا) فيمكن استعمال معادلة "كودر ريتشاردسون" في تقدير الثبات.

2-4- معادلة كودر ريتشاردسون 20 (KR 20):

تسمى أيضاً طريقة تحليل التباين باعتبارها تعتمد أساساً على الدراسة التفصيلية لتباين فقرات الاختبار أو المقياس كبديل لتقسيم الاختبار إلى نصفين (غانم، 2007)، وهي معادلة تحسب معامل ثبات اختبار يضم بنوداً ثنائية التصحيح، بمعنى تكون الإجابة عليها بواحد أو صفر، أو عادة ما تكون صح وخطأ أو نعم ولا حيث صيغة المعادلة كالتالي حسب ما ورد في (علام، 2000، ص.162):

$$KR20 = r = \frac{N [1 - \sum pq]}{N-1 S^2}$$

KR20 الثبات المقدر

N عدد بنود الاختبار

S^2 تباين الدرجة الكلية للاختبار

P نسبة الذين أجابوا على البند بطريقة غير صحيحة (بالنسبة لكل بند فإن $q = P - 1$)

$\sum pq$ = حاصل جمع p مضروبة في k لكل بند في الاختبار

هناك صيغة أخرى KR21 هي: $r = \frac{N [1 - \sum (m-n)]}{N-1 S^2}$ ، حيث m = متوسط الدرجات الكلية

2-5- طريقة الصور المتوازية (الصور المتكافئة، Parallel Forms):

هذه الطريقة تتلخص في مقارنة صورتين متماثلتين من اختبار يقيس نفس الخاصية، ويقصد بالتكافؤ أن تكون البنود في كلا الاختبارين على نفس الدرجة من الصعوبة ونفس العدد والترتيب، وطريقة الإجابة والتصحيح، والزمن المخصص للإجابة. (مراد و أمين، 2005).

يتم تطبيق الاختبار في صورته الأولى على عينة من الأفراد، ثم تطبق مباشرة أو بعد حين (فترة زمنية قصيرة) الصورة الثانية، ويحسب معامل الثبات بحساب الارتباط بين الصورتين، ويسمى معامل الثبات بمعامل التكافؤ، حيث يدل على

اتساق درجات الأفراد خلال الصور المتكافئة. (مراد و أمين، 2005)، إذا تم تطبيق صورتى الاختبار في نفس اليوم، فإن مصدر التباير الوحيد يكون الخطأ العشوائى جراء الفرق بين صورتى الاختبار، كما يمكن تقديم الاختبارين بالتبادل، بمعنى يقدم الاختبار الأول ثم الصورة المكافئة له لمجموعة من المفحوصين، والعكس بالنسبة لباقي المجموعة، تقدم الصورة المكافئة للاختبار أولاً ثم الاختبار الأصلي ثانياً، قصد تجنب أثر التدريب، بينما يمكن تقديم الاختبار الأول ثم يليه الصورة المكافئة له بعد فترة زمنية (يوم أو أسبوع)، وهنا يكون الخطأ مرتبطاً بالمعاينة الزمنية. (بركات، 2008).

تركز هذه الطريقة على استقرار النتائج وحساب الاتساق سواء في عينة مادة الاختبار أو في الأداء عبر الزمن، إلا أنه يصعب إعداد صورتين متكافئتين من الاختبار الواحد، ناهيك عن تكلفة الجهد والوقت (غانم، 2007).

يجدر الإشارة أن الصيغ المتكافئة يجب أن تحتوي على العدد ذاته من البنود، ويجب أن تغطي نفس المحتوى، ويجب أن تكون بنودها على قدر متساو من مستوى صعوبة أو السهولة، مع التساوي كذلك في مدة تطبيق الاختبار، وكذا وضوح تعاليمه والشكل الذي يقدم به وظروف إنجازه وكذا تساوي الشروط الفيزيكية لمكان التطبيق، يعني باختصار شروط التكافؤ لا تقتصر فقط على محتوى الاختبار وإنما على توحيد نفس شروط التطبيق وما يحيط به من ظروف.

3-العوامل المؤثرة على ثبات الاختبار(غانم، 2007):

- ✓ طول الاختبار، فكلما كان المقياس أكثر طولاً كلما زادت عدد فقراته، وارتفع بالمقابل معامل ثباته، بمعنى كلما قيس الاختبار بمؤشرات أكثر دلالة عليه، كلما قلل هذا من نسبة تدخل عوامل الصدفة على ثباته.
- ✓ زمن تطبيق الاختبار، فكلما كان الزمن أقصر كلما قلل هذا من نطاق استكشاف الاختبار لجوانب الظاهرة وبالتالي التقليل من معامل الثبات.
- ✓ تشتت درجات الاختبار من حيث زيادة عدد الأسئلة شديدة الصعوبة وأخرى شديدة السهولة من شأنه أن يخفف معامل الثبات بعكس الأسئلة المتدرجة من الأصعب إلى الأسهل.
- ✓ وضوح صياغة الفقرات، بحيث يجب الاتمتمل المؤشرات إلا معنى واحداً غير قابل للتأويل.
- ✓ التخمين، فكلما زاد تأثير الإجابة على فقرات التخمين كلما أثر ذلك على ثباته.
- ✓ حالة الفرد، فالمرض والتعب وعدم الثبات الانفعالي يؤثرون لا محالة على ثباته.

خلاصة:

في حساب الثبات نعتمد عادة على تأكيد دقة الاختبار من خلال استقرار نتائجه بين فترتي التطبيق سواء تعلق الأمر بتطبيق وإعادة تطبيق الاختبار، أو بتطبيق الاختبار والصورة المكافئة له. يحسب الثبات كذلك من خلال البحث في اتساق الاختبار سواء عن طريق التجزئة النصفية أو عن طريق معامل ألفا في حالة بدائل إجابة تعتمد على سلم ليكارت، أو معامل كودر رتشاردسون في حالة الإجابات على بنود الاختبار ب صحت وخطأ، نعم أو لا.

مع أن صدق المفهوم المتعلق بصدق الاتساق الداخلي يبحث هو الآخر في الاتساق، لذا عند استعمال هذا النوع من الصدق، يفضل بالضرورة التأكد من الثبات عن طريق البحث في استقرار النتائج، حتى ننوع في أدلة الثبات فلا نبحت عن الاتساق مرتين.

كما أن استعمال طريقة التجزئة النصفية مع حساب معامل "ألفا" أو "كودر" ليس إلا تكرارا دالا على الاتساق، لذا يفضل الجمع بين الطرق الدالة على الاستقرار والأخرى الدالة على الاتساق لنتائج أفضل وأدق. الصدق إذن دليل الثقة في صلاحية الاختبار لقياس ما وضع لقياسه، والثبات هو دال على دقة الاختبار في تحصيل النتائج.

الأهداف

- في نهاية المحاضرة يجب على الطالب أن يتمكن من:
- 1- التعرف على ماهية تمييز الفقرة في الاختبارات
 - 2- القدرة على حساب كل من معامل السهولة والصعوبة
 - 3- تحليل فقرات الاختبار بحساب قدرتها التمييزية
 - 4- القدرة على اتخاذ قرار تنفيذي بشأن فقرات الاختبار بالاعتماد على مؤشرات التمييز، الصعوبة والسهولة.

المحتويات

11- تمهيد

- 1- معامل تمييز الفقرة أو البند
- 2- مؤشر صعوبة الفقرة في حالة الإجابة ببدلين
- 3- مؤشر صعوبة الفقرة في حالة الإجابة بعدة بدائل
- 4- معامل السهولة والصعوبة للاختبار ككل
- 5- مؤشر التمييز
- 12- خلاصة

المحاضرة 05: التمييز

تمهيد:

تعتبر خاصية التمييز من الخصائص الهامة التي يجب أن تتوفر في بنود اختبار لاسيما اختبارات القدرات عموماً، تستعمل عادة عند تحليل فقرات الاختبار للتأكد من قياسها للفروق الفردية.

1- المقصود بالتمييز:

ورد عن أحمد (1960) أن التمييز يقصد به مدى قياس بنود الاختبار للفروق الفردية بين الأفراد من حيث الأكثر مهارة في قدرة معينة والأقل، أو الأكثر امتلاكاً للخاصية من الأقل، ويتطلب ذلك وجود محكات التمييز بين قدرات ومهارات الأفراد منها:

✓ تحديد مستويات الصعوبة والسهولة لكل سؤال.

✓ تحديد معامل التمييز لكل سؤال.

ويتم هذا عن طريق تحليل فقرات الاختبار تحليلاً منطقياً إحصائياً وتجريبياً لوحدات الاختبار بغرض معرفة خصائصها وتحسينها سواء بالحذف والإبدال أو بالتعديل أو بالإضافة أو بإعادة الترتيب، حتى يتسنى لنا الحصول على اختبار صادق ثابت ومناسب للفتة المقصودة بالقياس.

في هذا الصدد يتم التطرق في تحليل الفقرات إلى ناحيتين (أحمد، 1960):

- تحديد مستوى صعوبة الفقرة (Difficulty level)، وتحديد الدرجة التي تميز بها الفقرة بين مستويات قدرة المفحوصين، بمعنى القوة التمييزية للفقرة والبند (Discriminative power).

- تمييز الفقرة بين الدرجة العالية وبين الدرجة المنخفضة للسمة المقصودة بالدراسة، فلا تكون سهلة جداً لجميع المفحوصين ولا بالصعوبة جداً فيفضل في حلها الجميع، وبالتالي فنجاح الفقرة في تبيان الفروق الفردية بين الأفراد يتطلب تحقيقها لمستوى صعوبة يسمح بنجاح 50% من المفحوصين، حيث يمكن القول في هذه الحالة بأن الفقرة مميزة، فلا هي بشديدة الصعوبة ولا هي بالسهولة جداً.

ويمكن حساب معامل السهولة حسب (مراد وأمين، 2005، ص. 211) من خلال ما يلي:

$p = \frac{p_i}{n}$ حيث (p) مؤشر سهولة البند ، p_i عدد الأفراد الذين أجابوا إجابة صحيحة على البند

n عدد أفراد العينة

ويمكن أن يعطى معامل الصعوبة (q) كالآتي: $q = 1 - p$

2- معامل السهولة والصعوبة لأسئلة الاختبار:

معامل السهولة = $\frac{\text{مجم ص} + \text{مجم خ}}{\text{مجم ص} + \text{مجم خ}}$ حيث يشير (مجم ص + مجم خ) إلى مجموع الأفراد الذين أجابوا إجابة صحيحة والذين أجابوا إجابة خاطئة.

ويمكن التعبير عنها كالآتي أيضاً:

2-1- مؤشر صعوبة البنود في حالة الإجابة بديلين فقط:

ورد في (مراد وأمين، 2005) أنه يمكن الحصول على مؤشر سهولة بنود الاختبار في هذه الحالة بإيجاد نسبة عدد الأفراد الذين أجابوا على كل بند إجابة صحيحة إلى العدد الكلي للمشاركين، وكلما زادت هذه النسبة دل ذلك على سهولة البند وكلما قلت دل ذلك على صعوبة البند (في حالة البند يتكون من بديلين للإجابة/صح أو خطأ-نعم أو لا) حيث يتم ترتيب نتائج الطلبة تنازلياً، ثم تحدد المجموعتان الطرفيتان العليا والدنيا في حدود 27% من الأفراد وتحسب المعادلة التالية:

$$\text{معامل السهولة} = \frac{\text{مج (ع) + مج (ص د)}}{2} \quad (\text{مراد وأمين، 2015، ص. 215})$$

مج(ص ع) = عدد الإجابات الصحيحة للمجموعة العليا
مج(ص د) = عدد الإجابات الصحيحة للمجموعة الدنيا
2ن = ضعف عدد أفراد المجموعة الواحدة

تتراوح قيمة معامل الصعوبة بين 0 عندما يكون البند في غاية الصعوبة إلى 01 عندما يكون البند في غاية السهولة، لكن تتأثر درجة المفحوص عادة بأثر التخمين أي احتمال نجاح الفرد بالصدفة، لذا وجب تصحيح صعوبة البند من أثر التخمين حسب المعادلة التالية:

$$\text{الدرجة المصححة من اثر التخمين} = \text{ص} - \text{خ} \quad (\text{مراد وأمين، 2005، ص. 213})$$

وتجدر الإشارة هنا إلى أن العلاقة بين السهولة والصعوبة هي علاقة عكسية مباشرة، كما يمكن صياغة معاملي الصعوبة والسهولة في صيغة نسب مئوية أو كسور عشرية.

2-3- مؤشر صعوبة البنود في حالة الإجابة بعدة بدائل:

إذا كان الاختبار متعدد البدائل، فيمكن استعمال المعادلة التالية بعد ترتيب نتائج الإجابات تنازلياً، قم بحساب نسبة 27 بالمائة، حيث يتم تحديد عدد الإجابات من الفئة العليا وعدد الإجابات من الفئة الدنيا لتطبيق المعادلة التالية (مراد وأمين، 2005، ص. 216):

$$\text{مؤشر صعوبة البند} = \frac{\text{عدد الافراد الذين أجابوا صح من المجموعة العليا} + \text{عدد الافراد الذين أجابوا صح من الفئة الدنيا}}{\text{عدد افراد المجموعة العليا} + \text{عدد أفراد المجموعة الدنيا}}$$

2-4- حساب معامل الصعوبة والسهولة للاختبار ككل :

إذا تعلق الأمر بحساب معاملة صعوبة وسهولة الاختبار ككل، يمكن تطبيق المعادلة التي اشتقها فؤاد البهي (1978) وهي كالآتي (مراد وأمين، 2005، ص. 217):

$$\text{معامل سهولة الاختبار كله} = \frac{\text{مجموع الدرجات التي حصل عليها الأفراد}}{\text{مجموع النهاية العظمي لدرجات الأفراد}}$$

مثال تطبيقي: يفترض أن عدد الطلاب هو 40 طالبا وقد تحصلوا في مجموع درجاتهم في اختبار الرياضيات على 400، وأن أقصى درجة للاختبار هي 20

أي كل واحد يجب أن يتحصل على 20، النهاية العظمى لجميع أفراد العينة هي $40 \times 20 = 800$.

بالتطبيق في المعادلة السابقة نجد أن معامل سهولة الاختبار ككل $0.50 = 800/400$.

يمكن الحكم على الاختبار هنا أنه مميز، فلا هو بالصعب ولا هو بالسهل.

3- مؤشر التمييز (Discrimination Coefficient):

القدرة على التمييز هو قدرة الفقرة أن تميز بين درجات المفحوصين المرتفعة والمنخفضة في السمة أو الخاصية التي تقيسها فقرات الاختبار، ويمكن حساب هذا التمييز على أساس الدرجة الكلية للاختبار أو على أساس درجته على محك خارجي. (أحمد، 1960).

مؤشر تمييز الفقرة إذن يحدد لنا فعالية سؤال ما للتمييز بين الطالب ذو القدرة العليا والطالب الضعيف (خاصة باختبار القدرات)، فالقدرة التمييزية للفقرة وهو نسبة الفرق بين الذين أجابوا على الفقرة إجابة صحيحة من الفئة العليا والذين أجابوا إجابة صحيحة على الفقرة من الفئة الدنيا، تتضمن هذه الطريقة تقسيم الاختبار إلى قسمين، المجموعة التي تحصلت على أعلى الدرجات، والتي تحصلت على أقل الدرجات، حيث يأخذ كما حدد (KELLEY) نسبة 27% من الأفراد في كل مجموعتين طرفيتين، ومنه نحصل على معامل التمييز من خلال الفرق بين صعوبة البند في المجموعة العليا وصعوبة البند في المجموعة الدنيا، ويمكن حسابه حسب المعادلة التالية (مراد وأمين، 2005، ص. 218):

$$م ت = \frac{م ع - م د}{ن} \text{ حيث } م ت = \text{معامل التمييز، } م ع = \text{عدد المفحوصين الذين أجابوا إجابة صحيحة من أفراد المجموعة السفلى، } م د = \text{عدد المفحوصين الذين أجابوا إجابة صحيحة من أفراد المجموعة العليا، } ن = \text{العدد الكلي للمفحوصين في المجموعتين.}$$

يتراوح معامل التمييز ما بين $(-1 و +1)$ ، كلما كانت معاملات التمييز إيجابية كلما دل ذلك تمييز الفقرة، والاختبار الذي درجة تمييزه منخفضة لا يراعي الفروقات بين المفحوصين.

والجدول التالي يبين دلالات معامل التمييز حسب القيمة المتحصل عليها كما أشار إلى ذلك "أبييل" (Abel, 1965) نقلا عن (عدائكة وبن موسى، 2017، ص. 223):

جدول 04

دلالات معامل التمييز

دلالة معامل التمييز	قيم معامل التمييز
تدل على قوة تمييزية عالية بين أفراد المجموعتين العليا والدنيا، وبالتالي الفقرة تلي الغرض والهدف.	أكبر أو يساوي 0.40
الفقرة أو السؤال يتطلب مراجعة قليلة.	0.39- 0.30
فقرة تقع على الحد الفاصل وتحتاج إلى مراجعة.	0.29-0.20
السؤال يحذف لأنه مضلل، تحذف الفقرة وتجري مراجعة تامة.	أقل أو يساوي 0.19

خلاصة:

بعد الانتهاء من إعداد الاختبار لا بد من تحليل فقراته، التأكد من قدرته على التمييز بين الأفراد من ذوي الأداء المرتفع وذوي الأداء المنخفض، فالاختبار الجيد هو الذي يستطيع أن يقيس الفروقات الفردية. كما يمكن التأكد من صعوبة أو سهولة بنود الاختبار في اختبارات القدرات عموماً من خلال حساب معامل الصعوبة والسهولة، حيث يمكن الإبقاء على الفقرات إذا كان معامل صعوبتها 50%، في هذه الحالة يمكن الحكم أن معامل الصعوبة يساوي معامل السهولة، فالفقرة الجيدة لا تكون صعبة جداً ولا هي بالسهلة جداً. ترتب الفقرات أو البنود على أساس النتائج المتحصل عليها من الأكثر صعوبة إلى الأقل، ويتم على إثر ذلك إجراءات الحذف أو التبديل أو التعديل بغرض بناء اختبار موضوعي. تتراوح قيمة معامل الصعوبة بين 0 عندما يكون البند في غاية الصعوبة إلى 01 عندما يكون البند في غاية السهولة. بينما يتحدد مجال تمييز الاختبار من (-1 إلى +1)، فإذا كانت القيم سالبة أو أقل من 0.19، توجب حذف بنود الاختبار أو أسئلته المضللة.

- في نهاية المحاضرة يجب على الطالب أن يتمكن من:
- 1- التعرف على المعيار والمعايرة
 - 2- إدراك الاختلاف بين أنواع المعايير
 - 3- اختيار المعيار المناسب للسمة المقاسة
 - 4- حساب واستخراج معايير الاختبار بطريقة صحيحة

الأهداف

تمهيد

المحتويات

- 1- تعريف المعيار والمعايرة
 - 2- الشروط الواجب توفرها في المعيار
 - 3- أنواع المعايير:
 - 1-3 معايير الصف
 - 2-3 المعيار العمري
 - 3-3 الرتب المئينية
 - 4-4 الدرجة المعيارية أو القياسية
- خلاصة

تمهيد:

هناك صعوبة تعترض المختص النفسي في تفسيره لدرجات الاختبارات، إذ أن هذه الأخيرة لا يمكنها الحكم على الخاصية دون إسنادها إلى مرجعية معيارية يمكن من خلالها مقارنة درجة المفحوص مع بقية المفحوصين، والمرجعية المعيارية لا تأتي من فراغ وإنما لها طرق محكمة ومضبوطة، كما لها أنواع مختلفة باختلاف السمة المراد قياسها تتم ضمن عملية تسمى المعايير.

1- تعريف المعيار والمعايير:

المعيار هو الدرجة المتوسطة عند مجموعة الأفراد، تمثل نقطة الارتكاز التي يمكن من خلالها التعرف على تقدم أفراد هذه المجموعة، التي تسمى بدورها الجماعة المعيارية.

الدرجة الخام التي يتحصل عليها المفحوص في أي من الاختبارات النفسية والتربوية قد تكون مضللة ولا تحدد بالضرورة مستوى الفرد في الأداء، فحصول الفرد على صفر مثلا لا يدل على انعدام الخاصية لديه، لأن الصفر في الدراسات النفسية والتربوية هو اعتباطي، كما أن حصوله على العلامة الكاملة في اختبار تحصيلي مثلا لا يدل بالضرورة على التحكم التام، وإنما قدرته فقط على الإجابة على أسئلة الامتحان فقط وليس المقياس ككل.

كما أن أداء الفرد ونتائجه التي يتحصل عليها في اختبار معين لا تمكننا من تصنيفه ومقارنته بأداء الجماعة والتي قد تتباين من حيث الجنس والعمر والمستوى الثقافي أو المهني، ذلك لأن الدرجات المتحصل عليها هي درجات خام، لا بد من تحويلها إلى درجات معيارية توحد من خلالها معايير الحكم على الأداء. (Werner, 1965).

لذا يجب مقابلة الدرجة الخام إلى معيار محدد يفسرها ويمكن بذلك مقارنة أداء الفرد ضمن مجموعة مع أداء المفحوصين ضمن ذات المجموعة. (غانم، 2007).

لا تقتصر مقارنة المفحوص على نفس مجموعته، وإنما يقارن كذلك بمجموعة التقنين (Standradizationsample) لمعرفة معنى درجته بالنسبة لهذه العينة، حيث يقصد بهذه الأخيرة المجموعة التي طبق عليها الاختبار بغرض اشتقاق معايير واحدة يمكن على أساسها مقارنة أي شخص ينجز الاختبار بها. (عبد الخالق، 2000).

المعايير إذن هي إجراء الاختبار على عينة التقنين، وهي الدرجات المتوسطة أو النموذجية للمفحوصين في مجموعة محددة، تعبر عنها قيم كمية رقمية ترصد المستوى المتوسط للمجموعة وتحدد استنادا عليه مستويات الأداء المتنوعة على الاختبار بدءاً بأدناه وصولاً إلى أعلاها. (ميخائيل، 2016).

وهكذا لا يكون للدرجة الخام ولا للنسبة المئوية دلالة في حد ذاتها، بل تحتاج إلى معيار يعطيها معنى ودلالة، بحيث يمكن من خلاله مقارنة نتائج المفحوصين وإعطاء تفسير صحيح لها، وفي هذه الحالة نضمن المقارنة وفق معيار ثابت. (أحمد، 1960).

2- الشروط الواجب توفرها في المعيار:

ما دامت المعايير ليست إلا نتائج لإجراء الاختبار على عينات التقنين، فإن تمثيل هذه الأخيرة للمجتمع المراد بالقياس أمر ملزم لصالح المعايير والعكس صحيح، فتمثيل عينة التقنين للمجتمع تمثيلا صادقا هو شرط أساسي يضمن

صحة دلالة هذه المعايير، لذا يشترط في نقل الاختبار من بيئة إلى بيئة أخرى أن تكون عينات التقنين موازية لعينات التقنين الأصلية لاشتقاق معايير جديدة بدقة. (أحمد، 1960).

ويجب أن تتمتع المعايير بالشروط التالية (غانم، 2007):

- أن تكون للدرجة الواحدة معنى موحد حتى توفر أساسا لمقارنة درجات الافراد على الاختبار.
 - أن تكون متساوية، حتى يمكننا مقارنة أداء الفرد بفرد آخر من حيث الزيادة أو النقصان (كأن نقول درجات الفرد الأول تزيد عن الفرد الثاني بـ 10 وحدات مثلا).
 - أن تكون هناك نقطة صفر حقيقية ومطلقة تعبر عن عدم وجود الخاصية.
- ومع أن المستوى المتوسط هو الأكثر انتشارا في الحكم على الأداء، إلا أن الأمر أحيانا يتطلب حساب الانحراف المعياري والذي يعتبر بدوره شكلا من أشكال المتوسط، ويعبر عن متوسط انحراف الدرجات عن متوسطها. ويشير المعيار بذلك إلى مستوى الأداء المتوسط أو السوي أو الطبيعي كما يدل اسمه عليه.

3-أنواع المعايير:

3-1-معايير الصف:

يستخدم في تفسير درجات الأفراد في الاختبارات التحصيلية وهو يمثل متوسط تحصيل الطلبة في ذلك الصف، وبهذه الطريقة يمكن مقارنة درجات الفرد التي حصل عليها بدرجات زملائه في نفس الصف. (غانم، 2007).

3-2-المعيار العمري:

يرتبط المعيار العمري بمقياس بينيه، وهو مطبق على العديد من اختبارات الذكاء، يمكن حسابه من خلال متوسط درجات أفراد ذلك العمر في السمة السلوكية المراد قياسها: نسبة الذكاء، نسبة التعليم، نسبة التحصيل (مجيد، 2014، ص.167):

$$\text{نسبة الذكاء مثلا} = (\text{العمر العقلي} \div \text{العمر الزمني}) \times 100$$

$$\text{النسبة التعليمية} = (\text{العمر التحصيلي} \div \text{العمر الزمني}) \times 100$$

$$\text{النسبة التحصيلية} = (\text{النسبة التعليمية} \div \text{نسبة الذكاء}) \times 100$$

3-3-الرتب المئينية:

تستخرج المئينيات من عينة التقنين بتحديد أقل قيمة وأعلى قيمة على الاختبار، ثم يتم توزيع هذا المدى أو تقسم درجات المجموعة على أساس مقياس مئوي، وبالتالي يحدد المئين النسبة المئوية للحالات التي تقع بعد درجة معينة. (عبد الخالق، 2000).

وبالتالي فالمئينيات (Percentiles) تدل على النسبة المئوية من الدرجات في التوزيع التكراري التي تقع دون درجة معينة، فإذا كانت درجة قياس القلق مثلا هي 120، تقابل المئين 80، فهذا يعني أن صاحب هذه الدرجة (120) يقع دونه 80% من زملائه في درجة اكتسابهم لهذه السمة.

وإذا كانت درجة الفرد على اختبار أو مقياس معين تقابل المئين 70، فهذا دال على أن درجة المفحوص أعلى من 70% من درجات مجموعة التقنين.

إذا أردنا إذن تحديد مركز الفرد بالنسبة لجماعته أو عينة الدراسة، فإن أنسب طريقة هي المعيار المئيني، وهو مبني أساساً على نسبة الأفراد في مجموعة التقنين الذين يقعون تحت درجة معينة.

وحيث أن المئينيات تعتمد على تحديد حجم عينة التقنين، وحيث أن هذه الأخيرة قد تختلف من اختبار لآخر، تم افتراض أن المجموعة يمكن تقسيمها إلى مائة واحدة كلاً منها يسمى مئين (Percentile)، بمعنى أننا ننسب رتبة الفرد إلى حجم قياس هو مائة فرد بصرف النظر عن حجم العينة الحقيقي. (غانم، 2007).

تتكون معايير المئين من جدول من النسب المئوية تقابل درجاته كل درجة من الدرجات الخام على الاختبار، وتسمى النسب المئوية للمجموعة المعيارية الرتب المئينية للدرجة.

3-4- الدرجة المعيارية أو القياسية:

الدرجة المعيارية هي المسافة التي تبعد بها الدرجة الخام عن المتوسط الحسابي، يعبر عنها بوحدات من الانحراف المعياري، حيث الدرجة المعيارية = الدرجة الخام - المتوسط الحسابي ÷ الانحراف المعياري أي حسب ما ورد في (Cohen, 2009, P.97) :

$$Z = \frac{x - \bar{x}}{s}$$

حيث يشير (Z) إلى الدرجة المعيارية

(x) الدرجة الخام - (\bar{x}) المتوسط الحسابي - (s) الانحراف المعياري

تتراوح الدرجة المعيارية ما بين (-3 + 3)، فإذا كانت الدرجة الخام أعلى من المتوسط، أصبحت الدرجة المعيارية موجبة، أما إذا كانت الدرجة الخام أقل من المتوسط، صارت الدرجة المعيارية سالبة، أما إذا تساوت الدرجة الخام بالمتوسط ستكون الدرجة المعيارية تساوي صفر. (غانم، 2007).

أهم خاصية في الدرجات المعيارية هي أن وحدة حسابها هو الانحراف المعياري، حيث يساوي هذا الأخير واحداً صحيحاً إذا ما حسبنا الانحراف المعياري للدرجات المعيارية.

الدرجة المعيارية قد تكون قيمة موجبة كما قد تكون قيمة سالبة أو كسر عشري، ويتم تعديلها بما يسمى بالدرجة المعدلة من بينها:

أ- الدرجة التائية (T.score):

هي عبارة عن درجة معيارية متوسطها الحسابي 50 وانحرافها المعياري 10، يتكون نظام الدرجات القياسي هذا من مقياس يتراوح من 5 انحرافات معيارية أقل من المتوسط إلى 5 انحرافات معيارية فوق المتوسط، حيث إذا انخفضت النتيجة الأولية أتماماً عند 5 انحرافات معيارية أقل من المتوسط ستكون مساوية لدرجة (T=0)، والنتيجة الأولية على مستوى المتوسط تكون (T=50)، والنتيجة الأولية لـ 5 انحرافات معيارية فوق المتوسط تكون (T=100)، حيث ميزتها أن الدرجات لن تكون قيمتها سالبة على عكس الدرجات المعيارية (Cohen, 2009, P.97)، و بالتالي تحسب عن طريق المعادلة التالية:

$$T = 10 \times Z + 50$$

الدرجة التائية = 10 × الدرجة المعيارية + 50

ب-الدرجة "ج" (C.score):

فضلا على الدرجة التائية، هناك ما يسمى بالدرجة "ج" أو (C.score)، وهي تعديل للدرجات المعيارية بحيث تصبح درجات معيارية اعتدالية، إشتقها "ج Guilford) من توزيع متوسطه "5" وانحرافه المعياري "2"، بغرض إيجاد درجات معيارية اعتدالية تناسب بعض الاختبارات التي تقيس خصائص تكون فيها الفروق بين الأفراد واضحة، حيث تقاس كالتالي (أحمد، 1960، ص.322):

$$\text{الدرجة "ج"} = \text{الدرجة المعيارية} \times 2 + 5$$

$$\text{C score} = T \times 2 + 5$$

خلاصة:

الاختبار الجيد هو الاختبار القادر على توفير إجراءات تفسير مناسبة لدرجات الأفراد عليه، ويقصد بالإجراءات القدرة على تحويل الدرجات الخام إلى درجات تمكنا من تحديد مستويات الأداء المرتفعة والمتوسطة والمنخفضة، ذلك لأن الدرجة المتحصل عليها على الاختبار لا معنى لها إلا إذا نسبت إلى محك، وتصدر الإشارة أن اختبارات أخرى لا يمكن تحويل درجاتها المعيارية لأنها تتوفر على محك مسبق للتصنيف مثل الاختبارات مرجعية المحك، شأنها في ذلك شأن الاختبارات التحصيلية التي يحدد عادة المحك فيها بمتوسط درجات الاختبار، فالمحك في هذه الحالة هو معياري ينسب إلى درجات الأفراد، ففيه ننسب النقاط المتحصل عليها في امتحان دراسي إلى المتوسط (20/10) أو 10/5. أما الاختبارات التي تعتمد على معيار الجماعة في تصنيف الأداء، فهي تأخذ بعين الاعتبار أداء الفرد بالنسبة لأداء الجماعة، وهنا يمكن استعمال الرتب المئينية أو تحويل الدرجات الخام إلى درجات معيارية، المعيار العقلي أو الصفي، والمحك هنا هو مرجعي يعكس قيمة الدرجة المتحصل عليها ويقارن أداء صاحبها ببقية الأفراد.

الأهداف

في نهاية المحاضرة يجب على الطالب أن يتمكن من:

- 1- التعرف على ماهية المعاينة
- 2- إدراك شروط المعاينة
- 3- تحديد الفرق بين أسلوب العينة وأسلوب التعداد
- 4- توظيف المعارف المكتسبة في ضبط خطوات اختيار العينة.
- 5- تصنيف العينات الاحتمالية غير احتمالية
- 6- اختيار أسلوب المعاينة الأنسب لبحثه
- 7- الالتزام بخطوات دقيقة في عملية اختيار عينة دراسته.

المحتويات

تمهيد

- 1- تعريف المعاينة
 - 2- شروط المعاينة
 - 3- أسلوب العينة وأسلوب التعداد
 - 4- خطوات اختيار العينة
 - 5- أنواع العينات:
 - 1-5 العينات الاحتمالية
 - 2-5 العينات غير احتمالية
- خلاصة

لمحاضرة رقم 07: المعاينة:

تمهيد:

تحتاج عملية بناء الاختبار إلى تقنين الاختبار، عن طريق اختباره على عينة التقنين، ويختلف الأسلوب الذي يتم به اختيار عينة التقنين، حيث لا يمكن بأي حال تطبيق الاختبار على مجتمع الدراسة الأصلي، لاسيما في العينات الكبيرة جدا، لذا هناك مجموعة من الطرق المستخدمة لسحب عينة التقنين من المجتمع الأصلي بحيث تمثله، وعلى هذا الأساس قد نستطيع تعميم النتائج على جميع أفرادها، وتسمى هذه الطريقة بالمعاينة.

1-تعريف المعاينة والعينة:

تعرف العينة (Sample) بأنها جزء من مجتمع يختار بطريقة علمية محددة، يتم استخدامها في الحكم على المجتمع محل الدراسة، وعملية اختيار العينة تسمى معاينة (Sampling) (عبد الرحمن محمد، الحسيني، و محمود محمد، 1995).

والمعاينة هي: "مجموعة من الإجراءات الفنية التي تتخذ من طرف الباحث في سبيل سحب عدد جزئي من الوحدات المكونة للمجتمع المستهدف بالدراسة، وذلك بغرض استخدام هذه العينة من الوحدات في تمثيل المجتمع المسحوبة منه" (بن جندل، 2019، ص.13).

المعاينة هي إذن عملية اختيار عينة البحث بطريقة تضمن تمثيلها للمجتمع الأصلي حتى تتمكن من تعميم نتائج الدراسة، حيث ليس من السهل على الباحث أن يقوم بتطبيق بحثه على جميع أفراد المجتمع الأصلي واختيار عينة ممثلة للمجتمع الأصلي من شأنه توفير الجهد والوقت والمال.

2-شروط المعاينة:

نقلا عن (عبد الرحمن محمد، الحسيني، ومحمود محمد، 1995) تشترط طرق المعاينة المختلفة تحقيق ما يلي:

- أن تكون العينة المختارة ممثلة للمجتمع أصدق تمثيل.
- أن تكون التقديرات للمعالم الإحصائية لخصائص المجتمع مثل المتوسط، التباين، النسبة، ذات دقة ويمكن قياس درجة الثقة لها.

➤ أن تكون تكاليف اختيار العينة المطلوبة أقل ما يمكن، يراعى في ذلك الإمكانيات المتوفرة. الشائع في البحوث الكمية أن ترتبط بتقدير القواعد التقنية لعملية اختيار العينة لتحقيق تمثيل أمثل للمجتمع الذي استقيت منه، وحتى يمكن تعميم نتائج الدراسة لاحقا؛ بينما في البحوث الكيفية، فيشترط في المعاينة على مبادئ نوعية مثل الكفاية والملائمة بين نوع العينة وطبيعتها وأهداف البحث وغنى النتائج وقابليتها للإفادة في بحوث أخرى متشابهة.

3-أسلوب العينة وأسلوب التعداد:

هناك حالات معينة مثل البحوث المسحية يتعين على الباحث فيها استعمال أسلوب التعداد بدل اللجوء إلى المعاينة، في مثل حالات الحصر الشامل، أو الحاجة إلى إدراج فئات صغيرة جدا من المجتمع في الدراسة، أو في حالة عدم تجانس المجتمع. (دانييل، 2015).

والجدول التالي يشرح الفرق بين استعمال التعداد واستعمال العينة والحاجة إلى ذلك (دانييل، 2015، ص.93):

جدول 05

جوانب القوة والضعف في استعمال كل من التعداد والعينة

العينة	التعداد	الظروف والحالات
قوة	ضعف	البحث ذو هدف استطلاعي
ضعف	قوة	الحصر الشامل المطلوب إجراؤه
قوة	ضعف	الحاجة إلى اتخاذ قرار سريع
قوة	ضعف	الحاجة إلى جمع معلومات حديثة بانتظام
قوة	ضعف	الحاجة إلى قدر كبير من المعلومات المتعمقة والتفصيلية
قوة	ضعف	الحاجة إلى دراسة أو فئات معينة مستهدفة في المجتمع
قوة	ضعف	حجم المجتمع كبير
ضعف	قوة	تباين المجتمع
ضعف	قوة	صعوبة في الوصول إلى مكان عناصر المجتمع
قوة	ضعف	انتشار المجتمع على نطاق واسع
قوة	ضعف	الإنتاج الصناعي المستمر للمجتمع
قوة	ضعف	محدودية الوقت والمال والموظفين
ضعف	قوة	أهمية تقليل خطأ التحيز
ضعف	قوة	احتمال حدوث معدل مرتفع جدا لعدم الإجابة
قوة	ضعف	أهمية استخدام إجراءات تنفيذية سهلة
ضعف	قوة	الحاجة لتحليل البيانات بصورة تفصيلية
ضعف	قوة	القيود الأخلاقية والقانونية لاستخدام المجموعات الضابطة

وبالتالي، فإن أسلوب التعداد قد يقتصر على مجتمعات الدراسة صغيرة الحجم أو المتباينة، وقد يفرض هذا الأسلوب بسبب الحاجة إلى التفصيل في تحليل البيانات، أو حصر شامل للمجتمع وصعوبة الوصول إلى عناصره. في المقابل، فإن أسلوب العينة جد ضروري في حالة كبر حجم العينة وتجانس مجتمع الدراسة، والأنسب لتوفير المال والوقت والجهد.

وقد ورد عن (دانييل، 2015) دواعي اختيار أسلوب العينة نذكر منها ما يلي:

- أن يكون البحث ذا هدف استطلاعي.
- أن تكون هناك حاجة لاتخاذ قرار سريع.
- أن تكون هناك حاجة لجمع معلومات حديثة بالانتظام.
- أن تكون هناك حاجة إلى دراسة عناصر وفئات معينة مستهدفة من المجتمع.

- أن يكون حجم المجتمع كبيرا.
- أن يكون المجتمع منتشرا على نطاق جغرافي واسع.
- الإنتاج والتصنيع المستمر للمجتمع.
- أن تكون الموارد محدودة.
- أن تكون هناك أهمية لتقليل تحيز الإجابة.

4-خطوات اختيار العينة:

عند تقرير اختيار أسلوب العينة بدل أسلوب التعداد بناء على الاعتبارات التي تم ذكرها سلفا، يلجأ الباحث إلى اختيار تصميم معين للعينة سواء باختيار عينة احتمالية، أو عينة غير احتمالية، وهذا يتطلب مسبقا تحديد ما يلي:

4-1-الهدف من المعاينة:

يحدد الهدف من المعاينة والمشكلة المراد دراستها تحديدا دقيقا وواضحا، وتوضع بعدها التصميمات المختلفة للحصول على إجابات للأسئلة وكذا المصادر المناسبة لتحقيق الهدف البحثي. (عبد الرحمن محمد، الحسيني، و محمود محمد، 1995).

4-2-تحديد قائمة بمجتمع البحث:

يقصد بمجتمع البحث مجموع الوحدات التي يتعامل بها الباحث في إطار جمع بياناته، حيث يقوم الباحث بالحصول على قائمة كاملة دقيقة وحديثة لجميع مفردات المجتمع، يمكن لمجتمع البحث أن يكون أكثر ضبطا ووضوحا متوافقا تماما مع أهداف البحث يسمى مجتمع البحث المستهدف. (بن جنخل، 2019).

فإذا كان مجتمع البحث يخص المدارس مثلا، فإن مجتمع البحث المستهدف يمكن أن يشمل المجتمع المتاح مثلا والذي قد يشمل جميع مدارس مدينة معسكر مثلا.

4-3-تحديد وحدة العينة:

قد تكون فردا أو مؤسسة اجتماعية أو تربوية أو اقتصادية، فمثلا لدراسة العنف لدى تلاميذ المدارس الابتدائية، فإن مجتمع البحث هو جميع المدارس الابتدائية، بينما المدرسة الواحدة هي وحدة العينة التي تتكون من جميع التلاميذ ذكورا وإناثا الذين ينتمون إلى مجتمع البحث ووحداته التي هي المدارس الابتدائية، ويسمى هذا كذلك بالإطار الذي يعتبر قائمة مكونة من وحدات المعاينة التي يتكون منها مجتمع المعاينة.

تتطلب في هذه المرحلة تحديد كل الوحدات والعناصر التي نريد أخذ العينة منها، حيث يجب أن يكون تعريف المجتمع دقيقا ووحدات المعاينة واضحة. (عبد الرحمن محمد، الحسيني، و محمود محمد، 1995).

4-4-اختيار العينة:

اختيار العينة وتقدير أحجامها ليس بالأمر السهل، بل تعتمد هذه المرحلة على عدة اعتبارات بدء بالنظر إلى وحدة المعاينة، الأهداف البحثية، المنهج المتبع، وسائل جمع البيانات، وكذلك حجم العينة وطريقة المعاينة المعتمدة، وقد ورد عن دانيليل(2015) ضرورة التقييد بالاعتبارات التالية في عملية المعاينة:

4-4-1- تحديد حجم العينة:

يتوقف تحديد حجم العينة على عدة اعتبارات منها تجانس المجتمع الأصلي، وهنا يمكن اختيار عينة صغيرة الحجم، أما إذا كان المجتمع متباينا فمن الضروري اعتماد عينة كبيرة الحجم للتقليل من خطأ الصدفة؛ كما أن هناك اعتبارات أخرى في اختيار حجم العينة ترجع إلى الإمكانيات المادية المخصصة للبحث والوقت المحدد لجمع البيانات. يحدد العدد المناسب لأفراد العينة بناء على عدة معايير منها:

✓ أهداف الدراسة

✓ تجانس أو تباين المجتمع: فكلما زاد التجانس بين أفراد المجتمع كان العدد اللازم لتمثيل المجتمع أقل والعكس صحيح.

✓ أسلوب البحث المستخدم وطبيعة تصميمه: فالدراسات المسحية تحتاج إلى أكبر عدد ممكن لتمثيل المجتمع.

✓ درجة الدقة المطلوبة: لتحقيق الدقة المطلوبة لا بد من اختيار حجم عينة كبير لتعطي الثقة المطلوبة لتعميم النتائج.

✓ كما يعتمد تحديد حجم العينة على الإمكانيات المتوفرة للبحث من حيث المال والجهد والوقت.

فإذا أخذنا بعين الاعتبار أهداف الدراسة، فالدراسات الاستطلاعية والاستكشافية تستخدم عينة ذات حجم صغير، أما إذا كانت الدراسة وصفية للمجتمع أو تنبئية أو تفسيرية، فالحاجة تحتم استخدام عينة كبيرة نسبيا، نفس الشيء بالنسبة للدراسات التي تتطلب دقة عالية وقضايا علمية ذات نتائج دقيقة، فإن تصميم البحث هو الآخر يتطلب دقة كبيرة، وبالتالي اختيار عينة ذات حجم كبير أمر حتمي في هذه الحالة، لأن كبر حجم العينة يقلل الخطأ العشوائي.

عموما في حالة العينات غير الاحتمالية، تحديد حجم العينة يعتمد على طرق غير إحصائية مثل قواعد الحساب التقريبي المعتمد على الخبرة، حيث ورد عن (دانييل، 2015) الاحتمالات التالية:

➤ بحوث دراسة الحالة: من ثلاثة إلى خمسة مشاركين.

➤ البحوث الفينمولوجية: من ستة إلى عشرة مشاركين.

➤ البحوث التجريبية: من 15 إلى 30 مشاركا لكل مجموعة.

➤ البحوث المسحية: موضوع متعدد أو دراسة قومية من 10000 إلى 15000 مشاركا.

➤ البحوث الاستكشافية والدراسات الاستطلاعية: من 20 إلى 150 مشاركا.

وعموما صاغ الإحصائيون تحديدا للحجم الأمثل للعينة على النحو التالي:

* في الدراسات الوصفية ينصح باستخدام ما نسبته 20% من أفراد مجتمع يتكون من بضع مئات، و 10% من مجتمع يتكون من بضع آلاف، و 5% من مجتمع كبير جدا (عشرات الآلاف).

5- أنواع العينات:

اختيار أسلوب العينة يحتم على الباحث اتخاذ القرار المناسب باختيار تصميم العينات الاحتمالية أو العينات غير احتمالية بناء على عدة عوامل حددها (دانييل، 2015، ص. 115) في الجدول التالي:

جدول 06

جوانب القوة والضعف للعينات الاحتمالية وغير احتمالية

الاحتمالات	العينه غير احتمالية	العينه الاحتمالية
الأهداف		
هدف البحث استطلاعي	جانب قوة	جانب ضعف
الحاجة لاتخاذ قرار سريع	جانب قوة	جانب ضعف
الحاجة لدراسة عناصر معينه في المجتمع المستهدف	جانب قوة	جانب ضعف
الحاجة لاستخدام عينه ممثله	جانب ضعف	جانب قوة
الحاجة لإجراء استدلال إحصائي من العينه	جانب ضعف	جانب قوة
الحاجة لتقليل تحيز الاختبار	جانب ضعف	جانب قوة
أهميه الدراسة	جانب ضعف	جانب قوة
الهدف من العينه تقديم مثال توضيحي	جانب قوة	جانب ضعف
طبيعه المجتمع الإحصائي:		
تباين المجتمع	جانب قوة	جانب ضعف
صعوبه الوصول على عناصر المجتمع	جانب قوة	جانب ضعف
انتشار المجتمع جغرافيا بصورة كبيره	جانب قوة	جانب ضعف
الموارد المتاحة:		
محدودية الوقت والمال	جانب القوة	جانب ضعف
إطار العينه غير متاح	جانب قوة	جانب ضعف
اعتبارات تصميم البحث:		
تصميم البحث نوعي	جانب قوة	جانب ضعف
أهميه استخدام أساليب عمليه سهله	جانب قوة	جانب ضعف
حجم العينه المستهدف صغير جدا	جانب قوة	جانب ضعف

يتضح مما سبق أن خيار العينات الغير احتمالية هو الأفضل في حالات صغر حجم العينه وعدم تجانسها، انتشارها انتشارا على نطاق جغرافي واسع جدا، واستهداف عناصر بذاتها من المجتمع، كما تدعو نقص الموارد المتاحة من محدودية الوقت والجهد والمال إلى اختيار مثل هذا الأسلوب في المعاينه، في حين عكس ما ذكر هو يحتم اختيار العينات الاحتمالية.

5-1- العينات الاحتمالية:

العينات الاحتمالية هي أسلوب معاينة يعطي كل عنصر في المجتمع المستهدف احتمالاً معروفاً لاختياره في العينة، بمعنى أن التضمين المحتمل لأي عنصر من عناصر المجتمع البحثي المستهدف يحدث بلا تحيز استناداً إلى الاختيار العشوائي. (بن جخدل، 2019).

هناك أربع أنواع رئيسية للعينة الاحتمالية هي حسب ما ورد عن دانييل (2015):

5-1-1- العينة العشوائية البسيطة:

في هذا النوع كل فرد في المجتمع لديه نفس الاحتمال في الاختيار، كما أن اختيار أي فرد لا يؤثر في اختيار الفرد الآخر، وتستهمل عدة طرق في الاختيار العشوائي لأفراد العينة منها:

✓ كتابة جميع أفراد المجتمع الأصلي في بطاقات وتوضع في صندوق مثلاً وتقلب ثم تختار بطاقة حتى نحصل على عدد الأفراد المطلوبين في عينة البحث (طريقة القرعة).

✓ استخدام جداول الأعداد العشوائية: ترقيم مفردات المجتمع الأصلي في قائمة بأرقام متسلسلة، ثم يفتح جدول الأعداد العشرية على أي صفحة كانت، يتم اختيار أي رقم وتقرأ الأعداد بالترتيب من الأعلى إلى الأسفل، أو من الأسفل إلى الأعلى، أو من اليمين إلى اليسار، أو من اليسار على اليمين حتى يتم الحصول على حجم العينة المطلوب، ثم يعد كشفاً بالأسماء التي تقابل الأرقام التي تم اختيارها.

وكمثال على طريقة اختيار العينة باستخدام جدول الأرقام العشوائية، يمكن عرض الجدول حسب ما ورد في بن جخدل (2019، ص. 37):

جدول 07

نموذج مصغر عن جدول الأرقام العشوائية

10431	75100	26320	37674	75061	36318	11164
93205	31687	87054	58678	76831	91791	21215
12462	90870	08882	37649	66558	44482	10438
20461	97395	60881	66583	33266	26236	36792
00249	38370	82384	51414	12032	04773	73944
68714	72717	78483	93104	14063	12872	49563
83049	42314	33242	73117	41701	48237	64208
33835	80151	80749	29341	38605	72875	51486
04638	09610	48478	17971	64516	26360	99756
33827	45117	00431	72907	13015	55217	71325
16805	15838	94601	53010	12138	97198	65285
34976	38109	31381	29301	38224	57327	17264
51092	04485	68368	92558	31199	99754	95639
97438	45147	12841	11808	86210	76404	61555
84967	08153	79225	87130	04689	98768	78137
24550	14733	19177	28759	84987	99215	62490
13211	37284	92729	07044	21283	63444	24216
22250	43817	31605	55903	33226	95428	16975
74244	77904	91510	57609	71168	39542	59138
12944	87154	87912	31966	50414	59652	29478

لنفرض أن العدد الكلي لوحدات إطار المعاينة هو 700، وأنا نريد استخراج ما عدده 14 من أفراد العينة باستخدام الجدول أعلاه، فإذا أخذنا الرقم الأول عشوائياً متكوناً مثلاً من ثلاث أرقام هو (124)، فإننا نعتمد الأعداد التي تليه نزولاً والتي تتكون من ثلاث أرقام على أن لا تتجاوز العدد 700 الذي هو العدد الكلي لوحدات إطار المعاينة، وبالتالي فالأفراد المختارين تكون أرقامهم (124-204-202-687-338-046-338-168-349-510-129-222-132-245).

كما يمكننا اختيار الأرقام الواقعة على اليمين أو اليسار، صعوداً أو نزولاً بالنسبة للرقم الأول المختار عشوائياً.

كما يمكن التوليد العشوائي للأرقام باستخدام جهاز الحاسوب باستعمال برمجيات مختلفة مثل برنامج (Microsoft Excel).

تجدر الإشارة هنا أن هناك نوعين فرعيين للعينة العشوائية البسيطة هما العينة مع الإحلال والعينة بدون إحلال. ففي النوع الجزئي الأول، وعند اختيار عنصر من إطار العينة، يمكن إرجاعه إلى إطار العينة ويكون مؤهلاً للاختيار مرة ثانية في العينة.

أما في النوع الجزئي الثاني، فعند اختيار عنصر من العينة يتم استبعاده من إطارها حتى لا يتم اختياره مجدداً. (دانييل، 2015).

5-1-2- العينة العشوائية الطبقية:

تعتمد هذه الطريقة على التقسيمات الطبقية للمجتمع الأصلي الذي تختار منه العينة، حيث يقسم هذا الأخير إلى طبقات رئيسية بناء على خاصية معينة، ثم يختار بطريقة عشوائية من هذه الطبقات العدد المطلوب من المفردات بما يتناسب مع حجمها الحقيقي في المجتمع الأصلي فمثلا إذا كان مجتمع مكون من ثلاث طبقات العليا مثلا عددها 1000، الوسطى عددها 4000، الطبقة الدنيا وعددها 5000، والمطلوب اختيار عينة عشوائية عددها 100، باستخدام أسلوب التوزيع النسبي يمكن اختيار العينات كالآتي:

جدول 08

نموذج عن العينة العشوائية الطبقية

الفئات	العدد	حجم العينة المختارة
الطبقة العليا	1000	10
الطبقة الوسطى	4000	40
الطبقة الدنيا	5000	50
المجموع	10000	100

للعينة العشوائية الطبقية هي الأخرى أنواع فرعية هي العينة الطبقية النسبية والعينة الطبقية غير النسبية. أما العينة الطبقية النسبية فيقصد بها توزيع عدد العناصر على الطبقات حسب نسبة تمثيلها للمجتمع، وبهذا يتناسب حجم العينة المسحوبة مع حجم هذه الطبقة في المجتمع، مما يضمن طريقة الاختيار الاحتمالي المتساوي، وهذا يسمح لكل عنصر في المجتمع بالحصول على فرص متساوية للاختيار في العينة حسب حجم تواجده في كل طبقة. في حين العينة الطبقية غير النسبية لا تتناسب فيها عناصر العينة المختارة مع كل طبقة مع نسبة تمثيلها في المجتمع الكلي، ولا يعطى لعناصر المجتمع فرصة الاختيار المتساوي ضمن العينة ولا تطبيق كسر العينة (ناتج قسمة حجم العينة على حجم المجتمع). (دانييل، 2015).

5-1-3- العينة العشوائية المنتظمة:

تتميز هذه الطريقة بسهولة اختيار مفرداتها، من خلال مسافات متساوية على القائمة، حيث تختار الوحدة الأولى بطريقة عشوائية، أما باقي الوحدات فيتم اختيارها مراعيًا انتظام المسافات بين وحدات الاختبار وصولًا إلى حجم العينة المراد أخذه.

مثال: نفرض أنه في جامعة 1000 طالب، وأردنا اختيار عينة بنسبة 20% من الأفراد، أي بواقع طالب من كل خمسة طلاب وصولًا إلى العدد 200 طالبًا، فنأتي بقائمة تضم ألف طالب، ونضع لها أرقامًا متسلسلة، ثم نختار رقمًا عشوائيًا يقع ما بين الواحد والخمسة، فإذا اخترنا رقم 04، فسيكون صاحب هذا الرقم أول فرد في العينة، ونضيف 05 كل مرة حتى نصل إلى العدد 200، مثلًا الأرقام تكون: 04-09-14-19-26-31-36-41-46...

لعينة العشوائية المنتظمة هي الأخرى أنواع فرعية خاصة بها، حيث يمكن تصنيفها إلى ثلاث أنواع هي: العينة المنتظمة ذات الاتجاه الخطي، العينة المنتظمة ذات الاتجاه الدائري، والعينة المنتظمة المتكررة أو المضاعفة (دانييل، 2015). تختلف هذه الأنواع باختلاف المسار الخطي لانتقاء أرقام العينة، فالمثال المذكور سابقا يخص الاتجاه الخطي، ويمكن اختيار الاتجاه الدائري باختيار رقم يقع ما بين 1 وعدد العناصر في إطار العينة، وعند الوصول إلى نهاية القائمة في عملية الاختيار، يتم الرجوع إلى بداية القائمة للاستمرار في الاختيار بحيث يشكل بذلك نمط دائري في الاختيار. أما عن العينة المنتظمة المتكررة، فهي تنطوي على اختيار عينات متعددة بدءاً باستخدام بدايات عشوائية متعددة.

5-1-4- العينة العنقودية:

تستخدم إذا ما طبق البحث على عينة لها مثلاً امتداد جغرافي كبير يتعذر فيه الاختيار العشوائي للعينة، حيث لا تتوفر قائمة بعناصر المجتمع، فلو أردنا مثلاً دراسة الدخل السنوي للأسرة في مدينة معسكر، فيمكننا اختيار عينة عنقودية على مرحلتين:

المرحلة الأولى: نقسم المدينة إلى أحياء ونأخذ منها عينة بحجم مناسب مع حجم الحي ويعتبر هذا العنقود الأول. المرحلة الثانية: نقسم كل حي من الأحياء المختارة إلى عمارات ونختار من كل منها عدد مناسب من المنازل ثم نختار دخل الأسر التي تسكن تلك المنازل، وبالتالي نكون قد حصلنا على عينة عنقودية من مرحلتين. تقسم الأنواع الفرعية للعينة العنقودية على أساس عدد مراحل الاختيار، بمعنى يمكن أن تكون العينة العنقودية ذات المرحلة الواحدة، والعينة العنقودية ذات المرحلتين، والعينة العنقودية متعددة المراحل. (دانييل، 2015).

فبالاستناد إلى المثال المذكور أعلاه، تكون تلك العينة العنقودية متعددة المراحل، أما إذا تم اختيار العينة مرة واحدة، كأن نهتم مثلاً بالبحث حول أطفال دور الحضانة، فإذا افترضنا وجود ست دور للحضانة في المدينة، نختار عشوائياً دار حضانة واحدة ونطبق على أطفالها فحوى دراستنا، وبهذا نكون قد اخترنا أسلوب العينة العنقودية ذات المرحلة الواحدة.

وبالاستناد إلى نفس المثال أعلاه يمكن للباحث أن يختار عشوائياً مجموعة فرعية من الأطفال المسجلين بدار الحضانة المختارة، ليكونوا عينة دراسته، ومن هذا المنطلق يكون قد استخدم أسلوب العينة الاحتمالية ذات مرحلتين.

5-2- العينات غير الاحتمالية:

من خلال هذا الأسلوب في المعاينة يختار الباحث عينته بناء على حكمه الشخصي بدل الاختيار العشوائي، وهي تعتمد بشكل خاص على خبرة الباحث والهدف من بحثه.

والمعاينة غير الاحتمالية لا يتمتع فيها جميع أفراد مجتمع الدراسة من فرص متساوية للمشاركة، كما أنها تخص العينة بذاتها ولا يمكن بأي حال تعميم استعمالها على باقي مفردات المجتمع.

هذا النوع من المعاينة أكثر فائدة في الدراسات الاستطلاعية والاستكشافية، ويعتمده الباحثون عادة لعدة اعتبارات أخرى قد تتعلق بالوقت أو تكلفة الدراسة.

5-2-1- العينه الغرضية أو القصدية:

تسمى أيضا العينه العمديه يقوم الباحث باختيارها طبقا للغرض الذي يهدف إلى تحقيقه من خلال البحث، وفي ضوء تلك المعرفة يقوم الباحث باختيار وحدات معينه يعتقد أنها تمثل المجتمع الأصلي تمثيلا صادقا(لظفي، 2011)، كأن يريد الباحث دراسة اتجاهات طلبة السنة الثانية علم النفس حول تخصصات علم النفس، فيختار مباشرة طلبة السنة الثانية علم النفس في كليته، لأنها أولا عينه متاحة، وثانيا لأنه يعتقد أنها تملك خصائص المجتمع الذي يود دراسته.

5-2-2- العينه الحصصية:

يتم اختيارها عن طريق تقسيم مجتمع الدراسة إلى طبقات طبقا للخصائص التي ترتبط بالظاهرة محل البحث، حيث يتسم المجتمع بعدم التجانس، ثم يختار عينه من كل طبقة تتناسب مع حجم الطبقة في المجتمع(القحطاني، العامري،المذهب، و العمر، 2013).

وهي مهمة في استطلاع الرأي العام لأنها من جهة تتم بسرعة، ومن جهة أخرى بتكاليف أقل، حيث يتسنى للباحث اختيار عينات من فئات ذات خصائص مختلفة حسب الحجم العددي للجماعات، وهي تشترك مع العينه العشوائية الطبقيه من حيث اختلاف خصائص الفئات. (در، 2017).

الفرق بين هذا النوع من العينه والعينه العشوائية الطبقيه هو أسلوب اختيار أفراد كل طبقة إذ لا يتم بصورة عشوائية وإنما بصورة قصدية، أي أن الباحث يختار من الطبقات المفردات التي يريد بشرط أن يلتزم بالعدد المحدد لكل فئة.

5-2-3- عينه الصدفة:

يعتبر عامل الصدفة عاملا مشتركا في التصاميم غير احتمالية، تتكون هذه العينه من الأفراد الذين يقابلهم الباحث بالصدفة، فقياس الرأي العام للجمهور مثلا، قد يتم من خلال استطلاع رأي الأشخاص الذين يصادفهم الباحث سواء في الشارع أو في أماكن عامة، ويؤخذ على هذه الطريقة أنها لا تمثل المجتمع الأصلي، بل هي عينه تمثل نفسها فقط ولا يمكن بأي حال تعميم نتائجها، مع أن الحصول على مثل هذه العينه يتم بسرعة إلا أنها ترتبط بأخطاء قد تغير مسار البحث. (يونسي، شينار، و عماري، 2021).

5-2-4- عينه كرة الثلج:

يتم في هذه الطريقة اختيار فرد معين، وبناء على ما يقدمه الفرد من معلومات، يمكن للباحث أن يستغلها في اختيار شخص آخر لاستكمال المعلومات والمشاهدات المطلوبة، فالفرد الأول هو النقطة التي يبدأ من خلالها تكثيف حجم العينه والخيط الذي يجر بقية الأفراد (در، 2017).

تستخدم عادة هذه الطريقة في البحوث التي تكون موضوعاتها تشكل نوعا من الطابوهات والمحظور الحديث عنها أو صعوبة التعرف على الأفراد الذين يحملون خصائصها، كموضوع تعاطي المخدرات مثلا، أو الانحرافات بأنواعها، فالكشف شخص واحد قد يجرنا إلى التعرف على بقية الأشخاص وبالتالي تكبير العينه تدريجيا ككرة الثلج.

خلاصة:

توجد أربعة احتمالات لتصميم العينة الاحتمالية هي العينة العشوائية البسيطة، العينة العشوائية الطبقية، العينة العشوائية المنتظمة، والعينة العنقودية، من خصائصها أن لكل فرد نفس نسبة الحظوظ للظهور ضمن العينة، هي عينة ممثلة للمجتمع ويمكن تعميم نتائج الدراسة المتحصل عليها على باقي أفراد المجتمع.

أما العينات غير الاحتمالية لا يتمتع فيها جميع أفراد مجتمع الدراسة من فرص متساوية للمشاركة، كما أنها تخص العينة بذاتها ولا يمكن تعميم نتائجها على باقي مفردات المجتمع.

العينات غير احتمالية أنواع هي الأخرى، فمنها العينة القصدية، عينة الصدفة، العينة الحصصية وعينة كرة الثلج. عملية اختيار الأسلوب الأمثل للمعاينة يعتمد على عدة اعتبارات من بينها أهداف البحث، طبيعة المجتمع الإحصائي، الموارد المتاحة، واعتبارات تصميم البحث؛ فالعينة غير الاحتمالية هي الأنسب في البحوث الاستطلاعية والاستكشافية، أو تلك التي تخص خصائص مجموعة بعينها، كما تعتبر مناسبة أيضا في الأبحاث النوعية، أو في حالة محدودية الوقت والمال والحاجة إلى اتخاذ قرار سريع، أو يكون المجتمع منتشرا بطريقة كبيرة أو متجانسا، وإطار العينة غير متاح. في حين يعتبر اختيار أسلوب المعاينة الاحتمالية هو الأنسب في حالة كان هدف البحث غير استطلاعي، أو كان المجتمع متباينا أو غير منتشر جغرافيا بدرجة كبيرة يصعب حصره بسببها، أو كانت الموارد من وقت وجهد ومال متاحة ومتوفرة، كذلك عند الحاجة لاستخدام عينة ممثلة، والحاجة لإجراء استدلال إحصائي من العينة وتقليل تحيز الاختبار.

في نهاية المحاضرة يجب على الطالب أن يتمكن من:

الأهداف

- 1- إدراك ماهية التقنين وأهدافه
- 2- توظيف المكتسبات القبلية في تطبيق شروط تقنين الاختبار كاملة
- 3- التفريق بين العينة الاستطلاعية وعينة التقنين
- 4- الالتزام باستيفاء شروط دليل الاختبار

تمهيد

المحتويات

- 1- تعريف التقنين
 - 2- خطوات تقنين اختبار نفسي
 - 3- عينة التقنين ودليل الاختبار
- خلاصة

تمهيد:

بناء الاختبار بجميع بنوده ليس إلا مرحلة أولى في البناء، التي تستوجب مرحلة ثانية يتم فيها التأكد من مصداقية وموثوقية الاختبار واستقرار نتائجه وكذلك اختيار العينة المناسبة لعملية التقنين، ثم وضع معايير تكفل تصحيح الإجابات بشكل موحد لدى جميع المفحوصين، كل هذه الخطوات نقصد بها عملية التقنين.

1 - تعريف التقنين (Standarization):

الاختبار المقنن هو اختبار حددت إجراءات تطبيقه وأجهزته، وتصحيحه، بحيث يصبح من الممكن تقديم الاختبار نفسه في أوقات وأماكن مختلفة (مراد و أمين، 2005).

عندما يكون التقنين في تمام الفاعلية، فإن المفحوص يحصل على نفس الدرجة تقريبا إذا ما طبق عليه الاختبار ثانية.

يعتبر التقنين ضروري عند بناء الاختبار بغرض التحقق من مناسبه لقياس السمة وخصائصه السيكومترية. عمليات التقنين هي إذن تلك الخطوات التجريبية التي يمر بها المقياس في صورته النهائية عند إجرائه على عينات من الجماعات التي سيطبق عليها فيما بعد وذلك بغرض اشتقاق المعايير، لذا فيمكن القول أن المعايير تشتق من عينة التقنين (Standardization Sample) التي تمثل المجتمع الأصلي المدروس. (أحمد، 1960).

تتضمن فكرة التقنين تحديد شروط تطبيق الاختبار بعد بنائه تحديدا دقيقا، بحيث تستخدم طريقة واحدة في تطبيقه وفي وضع الدرجات وتفسيرها، وهذا يتطلب تثبيت كافة المتغيرات التي من شأنها التأثير على نتائج المقياس حتى تكون الفروقات بين الأفراد على هذا الاختبار أو المقياس ناتجة عن الاختلافات الحقيقية القائمة بينهم وليس إلى متغيرات أخرى. (كراجة، 1997).

التقنين يشترط التأكد أولا من صدق وثبات الاختبار، حيث يعتبران مراحل أساسية تتضمنها عملية التقنين، كما هناك مراحل أخرى هي بالترتيب كالتالي:

2- خطوات تقنين اختبار:

يمكن تقنين الاختبار باتباع الخطوات التالية:

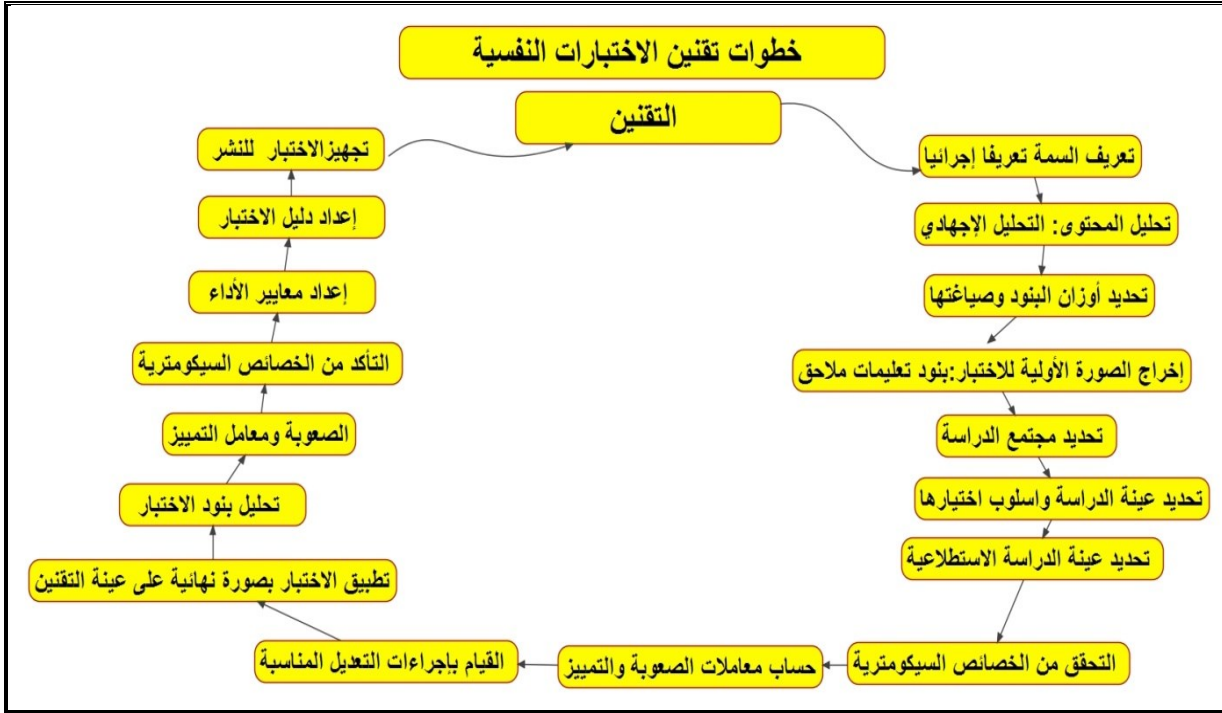
- ✓ تحديد المجتمع الذي سيقنن عليه الاختبار تحديدا إجرائيا دقيقا من تحديد سمات المجتمع وللعينة التي تمثل المجتمع تمثيلا جيدا والتي على ضوءها يتم تعميم النتائج على المجتمعات الأخرى.
- ✓ اختيار العينة الممثلة للمجتمع وتحديد أسلوب اختيارها.
- ✓ التخطيط الجيد لتطبيق الاختبار
- ✓ تطبيق الاختبار وتوحيد ظروف الإجراء لجميع أفراد العينة
- ✓ تحليل فقرات الاختبار ومدى تحقيقها للخصائص السيكومترية من صدق وثبات
- ✓ حساب معامل التمييز والسهولة والصعوبة.

✓ إيجاد معايير الأداء من خلال أداء مجموعة التقنين بغرض توفير إطار مرجعي يمكن من خلاله الحكم على أداء الفرد في الاختبار.

والشكل التالي المنقول عن (النفيعي، 2001) بتصرف يوضح خطوات التقنين:

شكل 06

خطوات تقنين اختبار



يلاحظ أن هذه الخطوات تتضمن معلومات يجب أن يتضمنها دليل الاختبار في صورته النهائية، فتقنين الاختبار لا يمس فقط الاختبار جديد البناء والتصميم، وإنما يشمل تلك الاختبارات المعدة سابقا والتي تحتاج إلى تقنينها في بيئة جديدة بغرض توحيد إجراءات التطبيق والتصحيح وضبط المعايير المناسبة.

تحديد الهدف من الاختبار يشير في محتواه على الغاية من تطبيق الاختبار، وطبيعة المجتمع المراد التطبيق عليه، والإمكانات المتاحة للبحث والتطبيق، ويتحدد من خلال كل هذا حجم العينة وطبيعة المعاينة، سواء تعلق الأمر باختبار معد مسبقا أو ببناء اختبار جديد.

في كلتا الحالتين نجد أمامنا عينتين الأولى استطلاعية بغرض التحقق من الخصائص السيكومترية للاختبار، والثانية هي عينة التقنين الغرض منها التأكد أيضا من الخصائص السيكومترية من صدق والثبات والتأكد من القوة التمييزية للاختبار وكذا اشتقاق المعايير ليصبح الاختبار جاهزا للنشر والتطبيق.

تجدر الإشارة هنا أن تقييم وتشخيص الظاهرة النفسية قد يكون في شكل اختبار أو مقياس أو الاستبيان، أو سلسلة من المواقف التي تضع المفحوص في وضع امتحان، فقد يتطلب الأمر حل مشكلات بسيطة في حالات الاختبارات المعرفية كاختبار الفهم القرائي مثلا، حيث كل عنصر منها يسمى بندا أو فقرة، وكل العناصر تخضع للتقييم حسب طبيعة الاختبار والهدف منه (كالمتوى، طرق الاستجابة أو مدة الإجابة). (Chartier, 2018).

3- خصائص عينة التقنين ودليل الاختبار:

من خصائص عينة التقنين أنها يجب أن تكون ممثلة للمجتمع الأصلي من حيث تركيبته وحجمه وصفاته، ويشترط فيها أن تحمل نفس خصائص عينة المجتمع الذي أعد فيه الاختبار (إن كان الاختبار معد مسبقا في بيئة أخرى)، وتخضع لجميع شروط المعاينة التي تم التطرق إليها سابقا.

أما دليل الاختبار فيجب أن يتضمن تعليمة موحدة وواضحة، الهدف من الاختبار، العينة الموجه إليها، شروط الأداء، خصائصه السيكومترية، وكذا مفتاح التصحيح وكذا معايير أو جداول المستويات التي تتضمن مستويات الأداء والدرجات المقابلة لها.

خلاصة:

بناء الاختبار عملية منهجية تتطلب خطوات صحيحة ومتتالية حيث لا يمكن الاستغناء عن أي واحدة لحساب الأخرى، وكلما تحرى الباحث الدقة في استكمال متطلبات كل مرحلة بعناية كلما قلل من الوقوع في أخطاء البناء والتقنين.

كما أن للتقنين شروطه بدء بتحديد الهدف من البناء والعينة المستهدفة إلى تحليل إجهادي يبحث في أدبيات الدراسة بطريقة علمية، حيث على الباحث تبني التزام نظري بعينه في اشتقاقه للبنود التي يخضعها لاحقا إلى التحقق من مصداقيتها وموثوقيتها وقوتها التمييزية، فهي تشترط البساطة والوضوح والسهولة، والموضوعية، ويتم تعديلها لاحقا بعد القيام بحساب النتائج المتحصل عليها، ليحملها إلى مرحلة أخرى ملزمة وهي التقنين من أجل اشتقاق معايير وتوحيد إجراءات التطبيق والتصحيح حتى يتم إخراج الاختبار في حلته النهائية الجاهزة للتطبيق بكل ثقة.

خاتمة:

تم تقديم محتويات مقياس بناء الاختبار، وهو مادة منهجية موجهة إلى طلبة السنة الثانية علم النفس، تساعدهم في إرساء قواعد البحث العلمي بمنهجية علمية صحيحة ودقيقة للوصول إلى نتائج ذات مصداقية تحقق الغرض من انتهاج المنهج العلمي من وصف، وتنبؤ ثم تحكم في الظاهرة المقاسة.

الاختبار النفسي يختلف باختلاف السمة المقاسة، لذا فإن طبيعة الاختبار تفرض طرق التحقق من مصداقيته، وتفرض كذلك أساليب معاينة ومعايرة تخص اختبارا بعينه وفئة بعينها، فاختبارات الشخصية تختلف عن اختبارات القدرات، وهذه الأخيرة تختلف عن باقي الاختبارات الموضوعية الأخرى.

بناء الاختبارات التحصيلية مثلا يتطلب مراعاة تحقيق الأهداف السلوكية للتعلم وشموليته للمحتويات من خلال بناء جدول مواصفات، وتحليل الفقرات فيه يعتمد أساسا على تحقيق صدق المحتوى والقدرة التمييزية لفقرات ومعاملات الصعوبة والسهولة، كما يتطلب اختيار طرق قياس الثبات تناسب خصوصيته، فالثبات عن طريق تطبيق الاختبار وإعادة تطبيقه مثلا تطرح إشكال وقوع اثر التعلم والتدريب حتى لو كانت المدة الفاصلة بين تطبيق الاختبارين بعيدة نسبيا، كما أننا في الحكم على الأداء قد نستعين بمعايير الصف والعمر بدل معايير أخرى.

كما أن اختبارات الشخصية تشترط أسلوب التعداد مثلا بدل العينات الكبيرة لأننا نحتاج إلى تحليل للبيانات بصورة تفصيلية وأكثر عمقا، في حين الاختبارات الموضوعية يمكن أن تشترط حجم عينة كبير في ظل شروط متعلقة بالإمكانات المتاحة والهدف من الدراسة، وطبيعة تصميم المجتمع، ويمكن كذلك التحقق من خصائصه السيكومترية بشتى الطرق لكن بمراعاة شروط معينة كذلك تتعلق بطبيعة الاختبار وحجم العينة.

التمكن من محتويات هذا المقياس أمر في غاية الأهمية، لأنه يتعلق بعصب البحث في الظاهرة النفسية ألا وهو القياس وأدواته وطرق التأكد من خصائصها السيكومترية وتقنينها، فنحن بحاجة دائمة إلى بناء مقاييس جديدة، وتقنين مقاييس معدة في غير بيئتنا حتى نلبي احتياجاتنا في البحث في الظاهرة النفسية باتباع منهج علمي صحيح واعتماد أساليب علمية حديثة.

قائمة المراجع:

المراجع العربية:

- أحمد، محمد عبد السلام. (1960). *القياس النفسي والتربوي*. القاهرة: مكتبة النهضة المصرية.
- أبو حويج، مروان، الخطيب، إبراهيم، وابو مغلي، سمير. (2002). *القياس والتقويم في التربية وعلم النفس*. عمان: الدار العلمية الدولية للنشر والتوزيع ودار الثقافة للنشر والتوزيع.
- أحمد الخطيب، محمد، و حامد الخطيب، أحمد. (2010). *الاختبارات والمقاييس النفسية*. عمان: دار الحامد للنشر والتوزيع.
- أحمد، محسن لطفي. (2011). *مقدمة في الإحصاء الاجتماعي*. جامعة الملك سعود: النشر العلمي والمطابع.
- بركات، حمزة حسن. (2008). *مبادئ القياس النفسي*. القاهرة، مصر: الدار الدولية للاستثمارات الثقافية .
- بن جخدل، سعد الحاج ب. (2019). *العينة والمعاينة، مقدمة منهجية قصيرة جدا*. عمان: دار البداية.
- دانييل، جوني. (2015). *أساسيات اختيار العينة في البحوث العلمية*. (طارق عطية عبد الرحمن، ترجمة) المملكة العربية السعودية : مركز البحوث.
- در، محمد. (2017). *أهم المناهج وعينات وأدوات البحث العلمي*. مجلة الحكمة للدراسات التربوية
- تيغزة، أحمد بوزيان. (2012). *التحليل العاملي الاستكشافي والتوكيدي مفاهيمهما ومنهجيتهما باستعمال حزمة SPSS و ليزرل LISREL*. عمان: دار المسيرة.
- سليمان، أمين علي محمد. (2010). *القياس والتقويم في العلوم الإنسانية*. الطبعة الثانية. القاهرة: دار الكتاب الحديث.
- السيد، فؤاد بهي. (1994). *الذكاء*. مصر: دار الفكر العربي.
- سي موسى، عبد الرحمان، و بن خليفة، محمود (2010). *علم النفس المرضي التحليلي والإسقاطي*. الجزائر: ديوان المطبوعات الجامعية.
- عباس، فيصل. (1996). *الاختبارات النفسية، تقنياتها وإجراءاتها*. بيروت: دار الفكر العربي.
- عبد الرحمن، أحمد محمد. (2011). *تصميم الاختبارات: أسس نظرية وتطبيقات عملية*. عمان: دار أسامة للنشر والتوزيع
- عبد الرحمان، سعد. (2008). *القياس النفسي النظرية والتطبيق*. الجزيرة: هبة النيل العربية للنشر والتوزيع.
- عبد الرحمن محمد الحسيني، راضي ، و محمود محمد، إبراهيم هندي. (1995). *مقدمة في المعاينة الإحصائية*. الرياض: دار المريخ.
- عبد الخالق، محمد أحمد. (2000). *استخبارات الشخصية (الإصدار الطبعة الثالثة)*. الإسكندرية: دار المعرفة الجامعية.

- عبد العظيم، حمدي عبد الله. (2013). موسوعة الاختبارات والمقاييس. مصر: مكتبة أولاد الشيخ للتراث.
- عدائكة، دنيا ، و بن موسى، يمينة. (2017). معيير الجودة في بناء الاختبارات التحصيلية الجيدة. مجلة السراج في التربية وقضايا المجتمع (01)، 216-278.
- عطية، نعيم. (1982). ذكاء الأطفال من خلال الرسوم. لبنان: دار الطليعة.
- علام، صلاح الدين محمود. (2001). القياس والتقويم النفسي والتربوي، أساسياته وتطبيقه وتوجهاته المعاصرة. القاهرة: دار الفكر العربي.
- غانم، حامد محمد. (2007). القياس النفسي للشخصية. الإسكندرية: المكتبة المصرية.
- الفتحي، إسماعيل محمدي. (2005). التقويم والقياس النفسي والتربوي. القاهرة: دار غريب.
- القحطاني، سالم سعيد ، العامري، أحمد سليمان ، آل مذهب، معدي محمد ، و العمر، بدران عبد الرحمن. (2013). منهج البحث في العلوم السلوكية . الرياض: مكتبة العبيكان.
- كراجة، عبد القادر. (1997). القياس والتقويم في علم النفس، رؤية جديدة. عمان: دار اليازوري العلمية للنشر والتوزيع.
- مجيد، سوسن شاكر. (2014). أسس بناء الاختبارات النفسية والتربوية. الأردن: مركز ديونو لتعليم التفكير.
- محمد عبد الحفيظ، إخلاص و حسين باهي، مصطفى. (2002). طرق البحث العلمي والتحليل الإحصائي في المجالات التربوية والنفسية والرياضية . القاهرة: مركز الكتاب للنشر.
- محمد، محمد إبراهيم. (2016). صدق المحتوى في البحوث التربوية: الواقع والتطبيق. المجلة المصرية للدراسات النفسية، 26(92)، 215-247 .
- مراد، صلاح أحمد ، و أمين، علي سليمان ن. (2005). الاختبارات والمقاييس في العلوم النفسية والتربوية. القاهرة: دار الكتاب الحديث.
- معمرية، بشير. (2013). مداخلات الملتقى الوطني الأول حول إشكالية القياس في علم النفس. الجزائر: مكتبة الرشاد للطباعة والنشر والتوزيع.
- مقدم، عبد الحفيظ. (2010). الإحصاء والقياس النفسي والتربوي مع نماذج من المقاييس والاختبارات. الجزائر: ديوان المطبوعات الجامعية.
- مقدم، عبد الحفيظ. (2011). الإحصاء والقياس النفسي والتربوي مع نماذج من المقاييس والاختبارات (الإصدار 03). الجزائر: ديوان المطبوعات الجامعية.
- ميخائيل، إسماعيل، إمطانيوس نايف. (2016). بناء الاختبارات والمقاييس النفسية والتربوية وتقنياتها. عمان: دار الإعصار العلمي

للنشر والتوزيع.

النفيعي، عبد الرحمان بن عبد الله بن أحمد. (2001). تقنين رافن للمصفوفات المتتابعة المتقدم على طلاب المرحلتين المتوسطة والثانوية بمنطقة مكة المكرمة. رسالة ماجستير تخصص اختبارات ومقاييس . كلية التربية جامعة أم القرى، مكة المكرمة.

يونسى، عيسى ، شينار، سامية ، و عماري، عائشة. (2021). العينة وأسس المعاينة في البحوث الاجتماعية. مجلة الرواق للدراسات الاجتماعية والإنسانية، 07 (02)، 539-528.

المراجع الأجنبية:

Anastasi.A ; Urbuna.S .(2006)*psychological testing*.India: Pearson .

Cambon.J .(1967) .La méthode des tests .*Bulletin de Psychologie* ،(256) 20 ،549-554.

Chabert. C.(2018). *Psychanalyse et es méthodes projectives*.Paris :Dunod

Cohen.S.(2009).*Psychological testing and assessment : An introduction to test and measurement*(7th ed).USA :McGraw-Hill.

Freeman.F.S .(1962) .*Theory and practice of psychological testing*.New york: Holt.

Werner.I.(1965).Investigation psuichométrique des aptitudes humaines .*bulettin de psychologie*,02,144-184.

Philipe Chartier .(2018) .*l'évaluation en psychologie: Tests et questionnaires*.Paris: In press.