الجمهورية الجزائرية الديمقراطية الشعبية
Peopl's Democratic Republic Of Alegria
وزارة التعليم العالي و البحث العلمي
Minestry Of higher Education And Scientific Research

University DjilaliBounaama of Khemis Miliana
Faculty Of Social and Human Sciences
Vice Deanship for Post Graduation
Scientific Research and External Relations

جامعة الجيلالي بونعامة خميس مليانة
كلية العلوم الاجتماعية والإنسانية
نيابة العمادة لما بعد التدرج والبحث العلمي
والعلاقات الخارجية

التاريخ 08 أفريل 2024
الرقم 347/ج ج ب خ م/ك ع ج ن/ن ع د ب ع ع خ/2024

مستخرج من محضر اجتماع المجلس العلمي للكلية رقم 2024/01
المنعقد بتاريخ : 2024/03/05

بالجلسة المنعقدة بتاريخ : 2024/03/05، صادق المجلس العلمي لكلية العلوم الإجتماعية و الإنسانية ، جامعة الجيلالي بونعامة خميس مليانة على مطبوعة دروس خاصة بالأستاذ(ة) : رابح هوادف موسومة بـ:" مطبوعة دروس القياس التربوي وبناء الإختبارات المدرسية "والموجهة لطلبة قسم : علم النفس وعلوم التربية ، المستوى : السنة الثانية ليسانس ، تخصص : علوم التربية .

وذلك بناء على التقارير الإيجابية للأستاذين :

1- د. عباس عبد الرحمان ، جامعة الجيلالي بونعامة - خميس مليانة .

2- د. صابر البحري ، المركز الجامعي بريكة .

رئيس المجلس العلمي

رئيس المجلس العلمي
كلية العلوم الاجتماعية و الإنسانية
أ.د شعباني نور الدين

Scanné avec CamScanner

وزارة التعليم العـــــالي والبحث العلمي
جامعة الجيلالي بونعامة – خميس مليانة
كلية الاجتماعيـة والإنسانية
قسم: علم النفس وعلوم التربية

مطبوعة مـــــــادة:

القياس التربوي وبناء الاختبارات المدرسية

موجهة لطلبة السنة الثانية علوم التربية (ارشاد وتوجيه)

السداسي الثالث/ الرابع

من اعداد
د. رابح هوادف

السنة الجـــــامعية
2025/2024

فهرس مطبوعة القياس التربوي وبناء الاختبارات المدرسية

مقدمة 06

الفصل الأول: مدخل إلى القياس النفسي والتربوي 07

1. مفهوم القياس النفسي والتربوي 07

2. التطور التاريخي للقياس النفسي 09

- الأحكام المبنية على الفراسة والمعتقدات الشخصية 09

- حركة التجريب في ألمانيا 09

- حركة دراسة ضعاف العقول في فرنسا 10

- الدراسة الإحصائية للفروق الفردية والقياس العقلي في إنجلترا 10

- حركة القياس العقلي في الولايات المتحدة الأمريكية 11

3. خصائص القياس النفسي والتربوي 12

- القياس غير المباشر 12

- القياس النسبي 12

- القياس غير التام 13

4. مستويات القياس 13

- المستوى الاسمي 14

- المستوى الرتبي 15

- المستوى الفتري 17

- مستوى النسبة 18

5. مسلمات القياس النفسي والتربوي 19

6. نظريات القياس النفسي 20

- النظرية التقليدية في القياس 20

- القياس معياري المرجع 21

- القياس محكي المرجع 22
- افتراضات النظرية التقليدية 23
- الانتقادات الموجهة للنظرية التقليدية 24
- النظرية المعاصرة في القياس 25

الفصل الثاني: الاختبارات النفسية والتربوية 26

- ماهية الاختبارات النفسية والتربوية واستخداماتها 26
- مفهوم الاختبارات 26
- تصنيفات الاختبارات 27
- مجالات استخدام الاختبارات النفسية والتربوية 29
- خطوات ومراحل بناء الاختبارات النفسية والتربوية 30
- الخصائص السيكومترية لأدوات القياس 34
- الصدق وأنواعه 42
- طرق تقدير الثبات 49

الفصل الثالث: بناء الاختبارات التحصيلية المدرسية 54

- خطوات إعداد وبناء الاختبارات التحصيلية 54
- مفهوم الاختبار التحصيلي 54
- مراحل بناء الاختبار التحصيلي 55
- أهمية الاختبارات التحصيلية 57
- وظائف الاختبارات التحصيلية 58
- تصنيف بلوم للأهداف التعليمية في المجال المعرفي 59

- العلاقة بين الأهداف التعليمية وأساليب التقويم التربوي 64
- جدول المواصفات 68
- تحليل مفردات الاختبارات الصفية 78
- أسئلة تقويمية وتطبيقات عملية 89

قائمة المراجع 97

الأهداف الإجرائية للمقياس

بنهاية دراسة هذا المقياس، يُنتظر من الطالب أن يكون قادرًا على:

1. تعريف مفهوم القياس النفسي والتربوي وبيان خصائصه الأساسية .
2. تفسير العلاقة التكاملية بين المنهجية العلمية والقياس والإحصاء في البحث العلمي .
3. التمييز بين مستويات القياس المختلفة وبيان خصائص كل مستوى ومجالات استخدامه .
4. التعرف على المسلمات الأساسية التي يقوم عليها القياس النفسي والتربوي .
5. تطبيق الخطوات المنهجية اللازمة لتصنيف المتغيرات وفق طبيعتها ومستويات قياسها .
6. اختيار الأساليب الإحصائية المناسبة في ضوء طبيعة الفرضيات والمتغيرات المدروسة .
7. تعريف الاختبارات النفسية والتربوية وتصنيفها وفق معايير علمية متنوعة .
8. شرح مراحل بناء الاختبارات النفسية والتربوية وفق الأسس العلمية المعتمدة .
9. إدراك أهمية التكوين الفرضي ودوره في إعداد أدوات القياس وتقنينها .
10. تحويل الأهداف العامة إلى أهداف إجرائية قابلة للملاحظة والقياس .
11. التعرف على جدول المواصفات من حيث مفهومه وأهميته ومكوناته وشروط بنائه ومجالات توظيفه .
12. بناء اختبار تحصيلي موضوعي وفق المعايير العلمية المعتمدة في القياس والتقويم .
13. التحقق من صلاحية أدوات القياس من خلال توظيف أساليب الصدق والثبات المناسبة .
14. توظيف المبادئ النظرية والتطبيقية للقياس النفسي والتربوي في المواقف البحثية والتربوية المختلفة.

مخطط الدرس –مقياس-

يهدف مقياس **القياس النفسي والتربوي** إلى تزويد الطلبة بالمعارف النظرية والمهارات التطبيقية المتعلقة بأسس القياس والتقويم، من خلال دراسة المفاهيم الأساسية للقياس، وتطوره التاريخي، ومكوناته،

ومستوياته، وعلاقته بالعلوم ذات الصلة، إضافة إلى التدريب على بناء أدوات القياس والتحقق من خصائصها السيكومترية.

السداسي الثالث: الجوانب النظرية للقياس النفسي والتربوي

يتناول هذا السداسي المحاور الآتية:

- مفهوم القياس النفسي والتربوي وأهميته .
- التطور التاريخي للقياس وأبرز رواده .
- العلاقة بين القياس والمنهجية والإحصاء .
- مسلمات القياس النفسي والتربوي .
- مستويات القياس وخصائصها .
- تصنيف المتغيرات واختيار الأساليب الإحصائية المناسبة .
- الاختبارات النفسية والتربوية: مفهومها، خصائصها، وأنواعها .

السداسي الرابع: بناء الاختبارات والتطبيقات العملية

يركز هذا السداسي على الجوانب التطبيقية لبناء أدوات القياس وفق مراحل منهجية متسلسلة:

المرحلة الأولى: مدخل إلى الاختبارات

- مفهوم الاختبارات النفسية والتربوية .
- وظائف الاختبارات ومجالات استخدامها .
- تصنيفات الاختبارات المختلفة .

المرحلة الثانية: خطوات بناء الاختبارات

- تحديد الهدف من الاختبار .

- تحديد المحتوى المراد قياسه .
- صياغة الفرضيات والأهداف الإجرائية .
- إعداد فقرات الاختبار وتنظيمها .
- تطبيق الاختبار وتجريبه الأولي .

المرحلة الثالثة: جدول المواصفات وبناء الاختبارات التحصيلية

- مفهوم جدول المواصفات وأهميته .
- مكونات جدول المواصفات ومعايير بنائه .
- توظيف جدول المواصفات في إعداد الاختبارات التحصيلية .
- بناء اختبار تحصيلي موضوعي وفق المعايير العلمية .

المرحلة الرابعة: الخصائص السيكومترية لأدوات القياس

- مفهوم الصدق وأنواعه وأساليب التحقق منه .
- مفهوم الثبات وطرائق حسابه .
- تحليل الفقرات ومعاملات السهولة والصعوبة والتمييز .
- تقويم أداة القياس والتحقق من صلاحيتها للاستخدام البحثي والتربوي .

المخرجات التعليمية المتوقعة

في نهاية المقياس يصبح الطالب قادرًا على تصميم وبناء أدوات قياس نفسية وتربوية وفق أسس علمية سليمة، وتحليل خصائصها السيكومترية، واختيار الأساليب الإحصائية الملائمة لمعالجة البيانات وتفسير نتائجها في البحوث النفسية والتربوية.

مقدمة:

يشكل القياس والتقويم أحد المرتكزات الأساسية للعملية التربوية، إذ يسهم في تشخيص خصائص المتعلمين، وتحديد مستويات أدائهم، والحكم على مدى تحقق الأهداف التعليمية والتكوينية. وقد شهد مجال القياس النفسي والتربوي خلال العقود الأخيرة تطورًا متسارعًا نتيجة التقدم في النظريات السيكومترية والأساليب الإحصائية وتطبيقات التكنولوجيا الرقمية والذكاء الاصطناعي.

وتهدف هذه المطبوعة إلى تمكين طلبة علوم التربية من اكتساب المعارف والمهارات الأساسية المرتبطة بمفاهيم القياس والتقويم وبناء الاختبارات النفسية والتربوية، مع التركيز على الأسس العلمية الحديثة التي تقوم عليها عملية بناء الأدوات وتطويرها وتقنينها وتحليل خصائصها السيكومترية.

وقد روعي في إعداد هذه الطبعة اعتماد أحدث الأدبيات العلمية في مجال القياس النفسي والتربوي، وتوحيد المصطلحات العلمية، وإثراء المحتوى بالأمثلة التطبيقية والأنشطة التقويمية، بما ينسجم مع متطلبات التكوين الجامعي وفق نظام التعليم العالي المعتمد.

وتتضمن المطبوعة مجموعة من المحاضرات المنظمة بصورة متسلسلة تبدأ بالمفاهيم الأساسية للقياس والتقويم، ثم مستويات القياس والنظريات السيكومترية، وصولًا إلى خطوات بناء الاختبارات والمؤشرات الإحصائية المستخدمة في تحليلها وتقويمها.

المحاضرة الاولى: مدخل إلى القياس النفسي والتربوي

أهداف التعلم: في نهاية هذا الفصل يكون الطالب قادراً على:

- التعرف على مفهوم القياس النفسي والتربوي.
- التمييز بين مفاهيم القياس والتقويم والتقييم والاختبار.
- إدراك أهمية القياس في البحث العلمي والعمل التربوي.
- التعرف على خصائص القياس النفسي والتربوي.
- استيعاب مجالات استخدام القياس في التربية وعلم النفس.

تمهيد:

شهدت العلوم الاجتماعية خلال العقود الأخيرة تطوراً ملحوظاً في استخدام الأساليب الكمية لقياس الظواهر النفسية والتربوية، وأصبح القياس أداة أساسية في اتخاذ القرارات التربوية والتوجيهية والعلاجية. فالعملية التعليمية الحديثة لم تعد تعتمد على الأحكام الانطباعية أو التقديرات الشخصية، بل أصبحت تستند إلى بيانات ومؤشرات موضوعية يتم الحصول عليها بواسطة أدوات قياس مقننة تتمتع بدرجات مقبولة من الصدق والثبات.

أولاً: مفهوم القياس النفسي والتربوي:

يُعرف القياس بأنه عملية إعطاء أرقام أو رموز لخصائص الأشخاص أو الأشياء أو الظواهر وفق قواعد محددة تسمح بوصف تلك الخصائص بصورة كمية قابلة للمقارنة والتحليل.

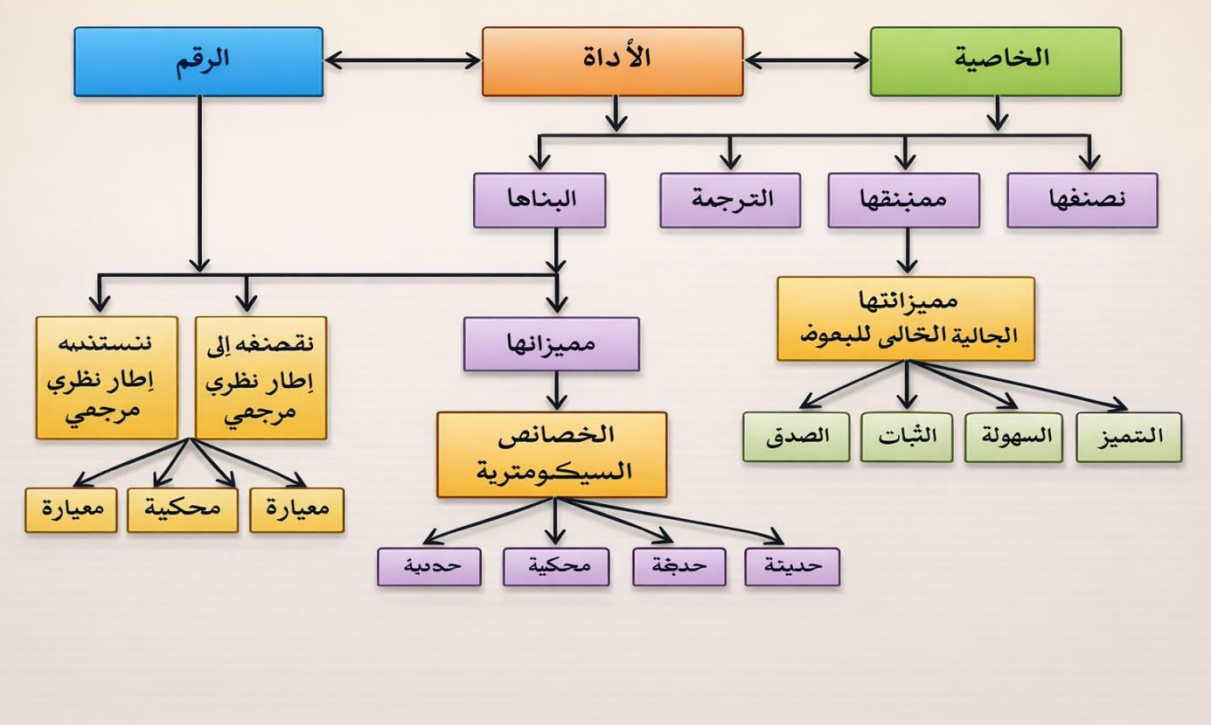
أما القياس النفسي والتربوي فهو عملية منظمة تهدف إلى تقدير مستوى السمات النفسية أو القدرات العقلية أو التحصيل الدراسي أو الاتجاهات أو الميول باستخدام أدوات مقننة تعتمد على أسس علمية وإحصائية دقيقة.

يُعدّ القياس في علم النفس عملية علمية منهجية تهدف إلى التقدير الكمي للخصائص والسمات النفسية لدى الأفراد، من خلال تحويل الظواهر النفسية غير المباشرة إلى مؤشرات رقمية قابلة للقياس والتحليل والتفسير. ويعتمد هذا المجال على مجموعة من الأدوات المقننة، مثل الاختبارات والمقاييس النفسية، إلى جانب الأساليب الإحصائية التي تضمن دقة النتائج وموضوعية تفسيرها. ويُعتبر القياس النفسي (Psychometrics) أحد الفروع الأساسية لعلم النفس، إذ يهتم ببناء وقياس السمات النفسية والسلوكات الإنسانية بطريقة معيارية، بما يسمح بفهم الفروق الفردية والتنبؤ بالسلوك المستقبلي (Borsboom, 2005).

كما يشير القياس النفسي إلى تحويل السمات النفسية إلى درجات كمية وفق قواعد علمية محددة، بما يعزز الانتقال من الوصف الكيفي إلى التفسير الكمي الموضوعي، وهو ما ينعكس إيجابًا على اتخاذ القرار في السياقات التربوية والإكلينيكية. ويتفرع هذا المجال إلى اتجاهات متعددة بحسب طبيعة الظاهرة المقاسة، حيث يُستخدم القياس السيكومتري في دراسة السمات السوية، بينما يُوظف القياس السيكوباثومتري في دراسة الاضطرابات النفسية، في حين يُعنى القياس السيكوسوسيومتري بالسمات ذات البعد الاجتماعي. (الشجيري والزهيري، 2022).

وتبرز أهمية القياس النفسي في كونه أداة تطبيقية واسعة الاستخدام في ميادين متعددة مثل التعليم، والتوجيه المهني، والصحة النفسية، والبحث العلمي، حيث يوفر بيانات كمية تساعد على الفهم الموضوعي للسلوك الإنساني والتنبؤ به في سياقات مختلفة (الشجيري والزهيري، 2022).

الشكل رقم (01) : العلاقة بين مكونات القياس (صليحة لعزالي،2023)



ثانياً: الفرق بين القياس والتقويم والتقييم والاختبار

01. القياس: عملية جمع البيانات الكمية المتعلقة بصفة أو سلوك معين.
02. التقييم: عملية تفسير نتائج القياس وإصدار أحكام حول مستوى الأداء.
03. التقويم: عملية شاملة تهدف إلى تحسين الأداء واتخاذ القرارات المناسبة بناءً على نتائج التقييم.
04. الاختبار: أداة أو وسيلة تستخدم لجمع البيانات اللازمة لعملية القياس.

ثالثاً: أهمية القياس النفسي والتربوي

تتجلى أهمية القياس النفسي والتربوي فيما يأتي:

- تشخيص الفروق الفردية بين المتعلمين.
- توجيه الطلبة نحو التخصصات المناسبة لقدراتهم.

- الكشف عن صعوبات التعلم والمشكلات النفسية.
- تحسين العملية التعليمية والتكوينية.
- دعم اتخاذ القرارات التربوية والإدارية.
- توفير بيانات علمية للبحث التربوي والنفسي.

رابعاً: خصائص القياس النفسي والتربوي

يتميز القياس النفسي والتربوي بعدد من الخصائص أهمها:

- الموضوعية.
- الدقة.
- الصدق.
- الثبات.
- الحساسية للفروق الفردية.
- القابلية للتفسير.

خامساً: مجالات استخدام القياس النفسي والتربوي

يستخدم القياس النفسي والتربوي في مجالات متعددة منها:

أ- المجال التربوي:

- قياس التحصيل الدراسي.
- تصنيف الطلبة.
- التوجيه والإرشاد التربوي.
- تقويم البرامج التعليمية.

ب- المجال النفسي:

- تشخيص الاضطرابات النفسية.
- قياس الشخصية والذكاء.
- الإرشاد والعلاج النفسي.

ج- المجال المهني:

- اختيار العاملين.
- التوجيه المهني.
- تقييم الأداء الوظيفي.

د- المجال البحثي:

- اختبار الفرضيات العلمية.
- دراسة العلاقات بين المتغيرات.
- بناء المقاييس والاختبارات.

خلاصة الفصل

يعد القياس النفسي والتربوي من الركائز الأساسية للعلوم السلوكية والتربوية، إذ يوفر أدوات علمية تساعد على وصف الظواهر النفسية والتربوية وتفسيرها والتنبؤ بها. كما يمثل أساساً لاتخاذ القرارات التربوية والبحثية المبنية على الأدلة.

أسئلة للمراجعة:

1. عرف القياس النفسي والتربوي.
2. بين الفرق بين القياس والتقويم والتقييم.
3. ما أهمية القياس في العملية التعليمية.

4. اذكر أهم خصائص القياس النفسي والتربوي.
5. وضح مجالات استخدام القياس النفسي والتربوي.

الفصل الثاني

التطور التاريخي للقياس النفسي والتربوي ونشأة الاختبارات النفسية

أهداف التعلم: في نهاية هذا الفصل يكون الطالب قادراً على:

- التعرف على المراحل التاريخية لتطور القياس النفسي والتربوي.
- فهم الجذور الفلسفية والعلمية للقياس النفسي.
- التعرف على إسهامات العلماء في تطوير الاختبارات النفسية.
- إدراك دور الإحصاء في تطور القياس النفسي والتربوي.
- تفسير العلاقة بين القياس النفسي والنظريات السيكومترية الحديثة.

تمهيد:

يرتبط ظهور القياس النفسي والتربوي بالتطور الذي شهدته العلوم الإنسانية والطبيعية خلال القرنين التاسع عشر والعشرين، حيث سعى الباحثون إلى إخضاع الظواهر النفسية للدراسة العلمية الموضوعية. وقد أسهمت التطورات الفلسفية والتجريبية والإحصائية في ظهور الاختبارات النفسية والتربوية التي أصبحت اليوم من أهم أدوات البحث والتشخيص والتوجيه.

أولاً: البدايات الأولى للقياس

ترجع جذور القياس إلى الحضارات القديمة التي اعتمدت أساليب مختلفة لتقييم الأفراد واختيارهم للمهام العسكرية والإدارية. غير أن هذه المحاولات كانت تعتمد على الملاحظة والخبرة الشخصية أكثر من اعتمادها على أساليب علمية منظمة.

وقد ظهرت في الحضارة الصينية القديمة بعض أشكال الاختبارات التي استعملت لاختيار الموظفين الحكوميين، حيث خضع المترشحون لاختبارات كتابية شملت المعرفة والأدب والقانون والإدارة.

ثانياً: ظهور المنهج العلمي في دراسة السلوك الإنساني

مع تطور العلوم التجريبية خلال القرن التاسع عشر، بدأ الاهتمام بقياس الظواهر النفسية بصورة موضوعية. وقد ساهم علماء الفسيولوجيا وعلم النفس التجريبي في تطوير أساليب جديدة لدراسة العمليات العقلية والحسية.

ويعتبر العالم الألماني فيلهلم فونت مؤسس أول مخبر لعلم النفس التجريبي سنة 1879، حيث ساعدت أبحاثه في ترسيخ فكرة إمكانية قياس الظواهر النفسية باستخدام أساليب علمية دقيقة.

ثالثاً: مساهمة فرانسيس غالتون في القياس النفسي

يعد فرانسيس غالتون من أبرز رواد القياس النفسي، حيث اهتم بدراسة الفروق الفردية بين الأشخاص، وسعى إلى قياس القدرات العقلية والحسية باستخدام أدوات كمية.

وقد ركز غالتون على:

- قياس زمن الرجع.
- قياس القدرات الحسية والحركية.
- دراسة الوراثة والفروق الفردية.
- إنشاء مختبرات للقياس النفسي.

كما مهدت أعماله الطريق أمام استخدام الأساليب الإحصائية في دراسة السلوك الإنساني.

رابعاً: كارل بيرسون وتطوير الأساليب الإحصائية

ساهم كارل بيرسون في تطوير مجموعة من الأساليب الإحصائية التي أصبحت أساساً للقياس النفسي والتربوي، ومن أهمها:

- معامل الارتباط.
- الانحراف المعياري.

- تحليل التوزيعات الإحصائية.

وقد سمحت هذه الأدوات بدراسة العلاقات بين المتغيرات النفسية والتربوية بصورة أكثر دقة وموضوعية.

خامساً: ألفريد بينيه وبداية الاختبارات العقلية الحديثة

يعتبر ألفريد بينيه بالتعاون مع تيودور سيمون المؤسس الحقيقي للاختبارات النفسية الحديثة.

ففي سنة 1905 طوّر الباحثان أول اختبار منظم للذكاء بهدف التعرف على الأطفال الذين يحتاجون إلى برامج تعليمية خاصة.

وقد تميز مقياس بينيه – سيمون بما يلي:

- التركيز على الوظائف العقلية العليا.
- ترتيب الفقرات حسب العمر العقلي.
- الاعتماد على الأداء الفعلي للفرد.
- إمكانية المقارنة بين الأفراد.

سادساً: اختبار ستانفورد – بينيه

قام لويس تيرمان في جامعة ستانفورد الأمريكية بتطوير اختبار بينيه – سيمون سنة 1916، حيث ظهرت النسخة المعروفة باسم اختبار ستانفورد – بينيه.

ومن أهم الإضافات التي تضمنها:

- توسيع العينة المعيارية.
- تحسين الخصائص السيكومترية.
- إدخال مفهوم نسبة الذكاء.(IQ)

وقد أصبح هذا الاختبار أحد أشهر اختبارات الذكاء في العالم.

سابعاً: الاختبارات النفسية خلال الحربين العالميتين

أدت الحربان العالميتان الأولى والثانية إلى زيادة الحاجة إلى تصنيف الأفراد وفق قدراتهم واستعداداتهم.

ولهذا الغرض تم تطوير:

- اختبار الجيش ألفا.(Army Alpha)
- اختبار الجيش بيتا.(Army Beta)

وقد استُخدمت هذه الاختبارات لتصنيف الجنود وتوزيعهم على المهام العسكرية المناسبة.

ثامناً: ظهور النظريات السيكومترية الحديثة

شهد النصف الثاني من القرن العشرين تطوراً كبيراً في مجال القياس النفسي من خلال ظهور:

1 -النظرية الكلاسيكية في القياس.(CTT)

وتقوم على دراسة العلاقة بين قدرة الفرد وخصائص الفقرة الاختبارية، وتعد من أهم الاتجاهات الحديثة في القياس النفسي.

3-الاختبارات المحوسبة التكيفية.(CAT)

تعتمد على الذكاء الاصطناعي والخوارزميات الإحصائية لتكييف الاختبار وفق مستوى أداء المفحوص.

تاسعاً: الاتجاهات المعاصرة في القياس النفسي والتربوي

يتجه القياس الحديث نحو:

- الاختبارات الإلكترونية.
- التحليل السيكومتري المتقدم.
- الذكاء الاصطناعي في التقويم.
- الاختبارات التكيفية.

- تحليل البيانات الضخمة.
- التقويم القائم على الكفاءات.

خلاصة الفصل

مرّ القياس النفسي والتربوي بمراحل متعددة بدأت بمحاولات بدائية لتقييم الأفراد، ثم تطور مع ظهور علم النفس التجريبي والإحصاء، وصولاً إلى الاختبارات النفسية الحديثة والنظريات السيكومترية المعاصرة. وقد أسهمت جهود علماء أمثال غالتون وبيرسون وبينيه وسيمون وتيرمان في بناء الأسس العلمية التي يقوم عليها القياس النفسي والتربوي في الوقت الحاضر.

أسئلة للمراجعة

1. ما المقصود بالتطور التاريخي للقياس النفسي؟
2. ما أبرز إسهامات فرانسيس غالتون في القياس النفسي؟
3. كيف ساهم كارل بيرسون في تطوير القياس النفسي؟
4. ما أهمية مقياس بينيه – سيمون؟
5. ما الفرق بين النظرية الكلاسيكية في القياس ونظرية الاستجابة للمفردة؟
6. ما أبرز الاتجاهات الحديثة في القياس النفسي والتربوي؟

الفصل الثالث: مفاهيم القياس النفسي والتربوي وخصائصه ومسلماته الأساسية

تمهيد:

يُعد القياس النفسي والتربوي أحد الأسس العلمية التي تعتمد عليها البحوث والدراسات النفسية والتربوية، حيث يهدف إلى وصف الظواهر السلوكية والسمات النفسية وصفاً كمياً يساعد على فهمها وتفسيرها والتنبؤ بها. وقد تطور مفهوم القياس تطوراً ملحوظاً مع تطور العلوم السلوكية وظهور النماذج السيكومترية الحديثة.(Cohen & Swerdlik, 2022) .

أولاً: مفهوم القياس النفسي والتربوي

يعرف القياس بأنه عملية إعطاء أرقام أو رموز لخصائص الأشخاص أو الأشياء وفق قواعد محددة تسمح بوصف تلك الخصائص وصفاً كمياً.(Stevens, 1951)

أما القياس النفسي والتربوي فيُعرف بأنه عملية منظمة تهدف إلى تحديد مقدار ما يمتلكه الفرد من سمة أو قدرة أو خاصية نفسية أو تربوية باستخدام أدوات مقننة تتمتع بدرجات مقبولة من الصدق والثبات (Urbina, 2024).

ويرى ملحم (2015) أن القياس النفسي يمثل الأساس العلمي الذي تعتمد عليه عمليات التقييم والتقويم والتوجيه والإرشاد، لأنه يوفر معلومات كمية يمكن الاستناد إليها في اتخاذ القرارات التربوية.

ثانياً: خصائص القياس النفسي والتربوي

1- القياس النفسي غير مباشر

تتميز معظم الظواهر النفسية بأنها لا تُلاحظ بصورة مباشرة، وإنما يُستدل عليها من خلال السلوك الظاهر للفرد. فالذكاء والقلق والدافعية مفاهيم افتراضية يتم قياسها عن طريق مؤشرات سلوكية يمكن ملاحظتها وتحليلها.(Anastasi & Urbina, 2019)

2- القياس النفسي احتمالي

لا يصل القياس النفسي إلى درجة الدقة المطلقة الموجودة في القياس الفيزيائي بسبب تعدد العوامل المؤثرة في السلوك الإنساني، ولذلك فإن نتائجه تبقى تقديرات احتمالية قابلة للخطأ بدرجات متفاوتة (DeVellis & Thorpe, 2021).

3- القياس النفسي يعتمد على العينات السلوكية

تعتمد الاختبارات النفسية على عينة من السلوك تمثل الخاصية المراد قياسها، إذ يصعب قياس جميع المظاهر المرتبطة بالسمة النفسية بصورة مباشرة.(Cohen & Swerdlik, 2022)

ثالثاً: مسلمات القياس النفسي والتربوي

يقوم القياس النفسي على مجموعة من المسلمات الأساسية أهمها:

أ- وجود الفروق الفردية بين الأفراد في القدرات والسمات والاستعدادات.(Cronbach, 1988)

ب- إمكانية التعبير عن السمات النفسية بصورة كمية من خلال مؤشرات سلوكية قابلة للملاحظة (Nunnally & Bernstein, 1994).

ج- تمتع السمات النفسية بدرجة من الاستقرار النسبي تسمح بقياسها والتنبؤ بها.(Urbina, 2024)

د- إمكانية ترتيب الأفراد وفق مقدار ما يمتلكونه من السمة المقاسة.(Embretson & Reise, 2020)

رابعاً: أخطاء القياس

يقصد بخطأ القياس الفرق بين الدرجة الحقيقية للفرد والدرجة المشاهدة التي يحصل عليها في الاختبار (Crocker & Algina, 2008).

وتُفسر النظرية الكلاسيكية للقياس العلاقة بين هذه المكونات وفق المعادلة:

$$X = T + E$$

حيث تمثل:

- X الدرجة المشاهدة.
- T الدرجة الحقيقية.
- E خطأ القياس.

وقد شكلت هذه المعادلة الأساس النظري لمعظم تطبيقات القياس النفسي خلال القرن العشرين (Lord & Novick, 1968).

خلاصة الفصل

يتسم القياس النفسي والتربوي بخصائص تميزه عن القياس الفيزيائي، ويستند إلى مجموعة من المسلمات النظرية التي تبرر إمكانية قياس السمات النفسية. كما أن إدراك مصادر أخطاء القياس يساعد الباحث على تحسين جودة أدواته وتحقيق مستويات أعلى من الصدق والثبات.

الفصل الرابع: مستويات القياس في العلوم النفسية والتربوية

تمهيد:

تُعد مستويات القياس من المفاهيم الأساسية في الإحصاء والقياس النفسي والتربوي، إذ يترتب على تحديد مستوى القياس اختيار الأسلوب الإحصائي المناسب وتفسير النتائج بصورة صحيحة. ويُعد تصنيف ستيفنز (Stevens, 1951)المرجع الأساسي في هذا المجال.

أولاً: مفهوم مستويات القياس

يشير مستوى القياس إلى طبيعة العلاقة الرياضية بين القيم الناتجة عن عملية القياس، وما تسمح به هذه القيم من عمليات حسابية وإحصائية.(Gravetter & Wallnau, 2017)

وقد صنف ستيفنز مستويات القياس إلى أربعة مستويات هي:

- القياس الاسمي.
- القياس الرتبي.
- القياس الفئوي.
- القياس النسبي.

01.مستوى القياس الاسمي(Nominal Scale)

يعد أبسط مستويات القياس، حيث يقتصر على تصنيف الأفراد أو الأشياء في فئات متمايزة دون ترتيب بينها.(Stevens, 1951)

ومن أمثلته:

- الجنس.
- التخصص الدراسي.
- الحالة الاجتماعية.

ويقتصر التحليل الإحصائي في هذا المستوى على التكرارات والنسب المئوية والمنوال واختبار كاي مربع (Field, 2024).

02. مستوى القياس الرتبي(Ordinal Scale)

يسمح هذا المستوى بترتيب الأفراد أو الأشياء وفق خاصية معينة دون معرفة مقدار الفروق الحقيقية بين الرتب.(Howell, 2013)

ومن أمثلته:

- الترتيب الدراسي.
- الرتب العسكرية.
- التقديرات الوصفية.

وتستخدم معه إحصاءات مثل الوسيط والرتب المئينية ومعامل ارتباط سبيرمان (Gravetter & Wallnau, 2017).

03.مستوى القياس الفئوي(Interval Scale)

يتميز بوجود فواصل متساوية بين القيم مع غياب الصفر الحقيقي.(Stevens, 1951)

ومن أمثلته:

- درجات الحرارة المئوية.
- بعض اختبارات الذكاء.
- الدرجات المعيارية.

ويتيح هذا المستوى استخدام المتوسط الحسابي والانحراف المعياري ومعامل ارتباط بيرسون واختبار (T) وتحليل التباين.(Field, 2024)

04.مستوى القياس النسبي(Ratio Scale)

يمثل أعلى مستويات القياس وأكثرها دقة لأنه يجمع خصائص المستويات السابقة ويتميز بوجود صفر حقيقي يدل على انعدام الخاصية المقاسة.(DeVellis & Thorpe, 2021)

ومن أمثلته:

- العمر.
- الطول.
- الوزن.
- الزمن.

ويتيح هذا المستوى إجراء جميع العمليات الرياضية والإحصائية الممكنة.

سادساً: أهمية مستويات القياس في البحث التربوي

تساعد مستويات القياس الباحث على:

- تحديد نوع المتغيرات.
- اختيار الاختبار الإحصائي المناسب.
- تفسير النتائج بصورة صحيحة.
- تجنب الأخطاء المنهجية في التحليل الإحصائي.(Cohen et al., 2018)

مستويات القياس في ضوء النماذج السيكومترية الحديثة:

تشير الدراسات الحديثة إلى أن نظرية الاستجابة للمفردة (IRT) ساهمت في تجاوز بعض القيود المرتبطة بمستويات القياس التقليدية من خلال بناء تدريجات أكثر دقة للسمات النفسية والقدرات العقلية (Embretson & Reise, 2020).

كما ساعدت الاختبارات المحوسبة التكيفية في تحسين دقة القياس وتقليل أخطاء التقدير المرتبطة بالاختبارات التقليدية.(Van der Linden, 2018)

خلاصة الفصل:

تشكل مستويات القياس الأساس الذي تُبنى عليه المعالجات الإحصائية في البحوث النفسية والتربوية، ويؤدي تحديدها بدقة إلى اختيار الأساليب التحليلية المناسبة وتحقيق نتائج أكثر صدقاً وموضوعية.

وتُعد مستويات القياس من المفاهيم الأساسية في القياس النفسي والتربوي والإحصاء، إذ تحدد طبيعة البيانات التي يجمعها الباحث ونوع العمليات الرياضية والإحصائية التي يمكن إجراؤها عليها. وقد صنّف ستانلي سميث ستيفنز مستويات القياس إلى أربعة مستويات رئيسية هي: الاسمي، والرتبي، والفئوي (الفاصل)، والنسبي.

ويُستخدم **المستوى الاسمي** لتصنيف الأفراد أو الظواهر في فئات متميزة دون وجود ترتيب بينها، بينما يسمح **المستوى الرتبي** بترتيب الأفراد أو الظواهر وفق درجة امتلاكهم للصفة المقاسة. أما **المستوى الفئوي** فيتميز بوجود فواصل متساوية بين القيم مع غياب الصفر الحقيقي، في حين يُعد **المستوى النسبي** أعلى مستويات القياس لأنه يجمع خصائص المستويات السابقة جميعها ويتميز بوجود صفر حقيقي يسمح بإجراء جميع العمليات الحسابية والإحصائية.

وتكمن أهمية مستويات القياس في أنها توجه الباحث إلى اختيار الأساليب الإحصائية المناسبة لتحليل البيانات وتفسير النتائج بصورة صحيحة. فكلما ارتفع مستوى القياس زادت دقة المعلومات المتاحة واتسعت إمكانات التحليل الإحصائي. لذلك فإن التحديد الدقيق لمستوى القياس يُعد خطوة أساسية لضمان سلامة البحث العلمي وصدق نتائجه وموضوعية الاستنتاجات المستخلصة من

الفصل الخامس: الإحصاء الوصفي في القياس النفسي والتربوي

أهداف التعلم:

في نهاية هذا الفصل يكون الطالب قادراً على:

- التعرف على مفهوم الإحصاء الوصفي وأهميته في القياس النفسي والتربوي .
- حساب وتفسير مقاييس النزعة المركزية .
- حساب وتفسير مقاييس التشتت .
- التعرف على الدرجات المعيارية والرتب المئينية .
- تفسير معاملات الارتباط في الدراسات النفسية والتربوية .
- توظيف الأساليب الإحصائية الوصفية في تحليل نتائج الاختبارات .

تمهيد:

يُعد الإحصاء الوصفي من الأدوات الأساسية التي يعتمد عليها الباحث في تنظيم البيانات وعرضها وتحليلها وتفسيرها. وتزداد أهمية الإحصاء الوصفي في مجال القياس النفسي والتربوي لكونه يساعد على وصف خصائص الأفراد والمجموعات وتلخيص كميات كبيرة من البيانات في مؤشرات رقمية سهلة الفهم والتفسير (Field, 2024).

ويشكل الإحصاء الوصفي الخطوة الأولى في تحليل البيانات قبل الانتقال إلى الإحصاء الاستدلالي واختبار الفروض العلمية.(Gravetter & Wallnau, 2017)

أولاً: مفهوم الإحصاء الوصفي

الإحصاء الوصفي هو مجموعة من الأساليب الرياضية والإحصائية المستخدمة في تنظيم البيانات وتبويبها وعرضها وتلخيصها بهدف وصف الظاهرة المدروسة دون تعميم النتائج على المجتمع الأصلي (Howell, 2013).

ويشمل الإحصاء الوصفي:

- الجداول الإحصائية .
- الرسوم البيانية .
- مقاييس النزعة المركزية .
- مقاييس التشتت .
- الدرجات المعيارية .
- الرتب المئينية .
- معاملات الارتباط .

ثانياً: مقاييس النزعة المركزية

تشير مقاييس النزعة المركزية إلى القيم التي تمثل مركز توزيع الدرجات.

1-المتوسط الحسابي(Mean)

يُعد المتوسط الحسابي أكثر مقاييس النزعة المركزية استخداماً في البحوث النفسية والتربوية (Field, 2024).

ويتم حسابه وفق العلاقة التالية:

حيث:

- $\bar{X}$ المتوسط الحسابي .
- ΣX مجموع الدرجات .
- N عدد الأفراد .

مثال تطبيقي

درجات خمسة طلبة:

10، 12، 14، 16، 18

المتوسط الحسابي:

(10 + 12 + 14 + 16 + 18) ÷ 5 = 14

إذن متوسط المجموعة يساوي 14.

2. الوسيط(Median)

الوسيط هو الدرجة التي تقع في منتصف التوزيع بعد ترتيب الدرجات تصاعدياً أو تنازلياً (Gravetter & Wallnau, 2017).

مثال:5 ، 8، 10، 12، 15

الوسيط = 10

ويُفضل استخدامه عند وجود درجات متطرفة.

3. المنوال(Mode)

المنوال هو أكثر القيم تكراراً في التوزيع الإحصائي.(Howell, 2013)

مثال:

8، 9، 10، 10، 10، 12، 15

المنوال = 10

ويستخدم كثيراً مع البيانات الاسمية.

ثالثاً: مقاييس التشتت

تساعد مقاييس التشتت على معرفة مدى انتشار الدرجات حول المتوسط الحسابي.

01.المدى(Range) :يمثل الفرق بين أكبر وأصغر درجة في التوزيع.

مثال:

أكبر درجة = 20

أصغر درجة = 10

المدى: 20 - 10 = 10

2-التباين(Variance)

يقيس متوسط مربعات انحرافات الدرجات عن المتوسط الحسابي.(Field, 2024)

3-الانحراف المعياري(Standard Deviation)

يُعد أكثر مقاييس التشتت استخداماً في القياس النفسي والتربوي.

وهو الجذر التربيعي للتباين.

وكلما صغر الانحراف المعياري دل ذلك على تجانس الدرجات.

رابعاً: الدرجات المعيارية

نظراً لصعوبة مقارنة الدرجات الخام بين الاختبارات المختلفة، يتم تحويلها إلى درجات معيارية (Cohen & Swerdlik, 2022).

الدرجة المعياريةZ

تعبر عن عدد الانحرافات المعيارية التي تبعدها الدرجة عن المتوسط الحسابي.

تفسير الدرجة المعيارية

التفسير	قيمةZ
عند المتوسط	0
أعلى من المتوسط بانحراف معياري واحد	+1

التفسير	قيمةZ
أعلى من المتوسط بانحرافين معياريين	+2
أقل من المتوسط بانحراف معياري واحد	-1

خامساً: الرتب المئينية

تمثل الرتبة المئينية الموقع النسبي للفرد مقارنة ببقية أفراد المجموعة.(Anastasi & Urbina, 2019)

فإذا حصل طالب على الرتبة المئينية 85 فهذا يعني أنه تفوق على 85% من أفراد المجموعة المرجعية.

وتستخدم الرتب المئينية بكثرة في:

- اختبارات الذكاء .
- الاختبارات التحصيلية .
- الاختبارات النفسية المعيارية .

سادساً: الارتباط في القياس النفسي والتربوي: يشير الارتباط إلى درجة العلاقة بين متغيرين أو أكثر (Field, 2024).

معامل ارتباط بيرسون

يتراوح بين:

$-1 \leq r \leq +1$

ويُفسر كما يلي:

التفسير	قيمةr
ارتباط موجب تام	+1
ارتباط موجب قوي	0.70+إلى +0.99
ارتباط متوسط	0.30+إلى +0.69
لا توجد علاقة	0
ارتباط سالب متوسط	0.30-إلى -0.69
ارتباط سالب قوي	0.70-إلى -0.99
ارتباط سالب تام	-1

مثال تطبيقي

وجود معامل ارتباط قدره:

r = 0.82

بين التحصيل الدراسي والدافعية للإنجاز.

يدل ذلك على وجود علاقة موجبة قوية بين المتغيرين.

سابعاً: أهمية الإحصاء الوصفي في القياس النفسي والتربوي

يساعد الإحصاء الوصفي على:

- تنظيم البيانات وعرضها .
- وصف خصائص الأفراد والمجموعات .

- تفسير نتائج الاختبارات .
- بناء المعايير .
- مقارنة أداء الأفراد .
- تمهيد الطريق للتحليل الإحصائي المتقدم .

كما يمثل أساساً لدراسة الخصائص السيكومترية للاختبارات والمقاييس النفسية والتربوية (DeVellis & Thorpe, 2021).

خلاصة الفصل

يشكل الإحصاء الوصفي الركيزة الأساسية لتحليل البيانات في القياس النفسي والتربوي، حيث يسمح بتلخيص البيانات ووصفها من خلال مقاييس النزعة المركزية والتشتت والدرجات المعيارية والرتب المئينية ومعاملات الارتباط. وتُعد هذه المؤشرات ضرورية لفهم نتائج الاختبارات وتفسيرها واتخاذ القرارات التربوية المبنية على الأدلة.

أسئلة للمراجعة

1. ما المقصود بالإحصاء الوصفي؟
2. ما الفرق بين المتوسط الحسابي والوسيط والمنوال؟
3. لماذا يعد الانحراف المعياري من أهم مقاييس التشتت؟
4. ما أهمية الدرجات المعيارية؟
5. كيف تفسر الرتبة المئينية؟
6. ما المقصود بمعامل الارتباط؟
7. متى نستخدم معامل ارتباط بيرسون؟

أنشطة تطبيقية

النشاط الأول

احسب المتوسط الحسابي والوسيط والمنوال للدرجات الآتية:

12، 15، 18، 20، 20، 22، 25

النشاط الثاني

احسب المدى للبيانات السابقة.

النشاط الثالث

فسر النتائج الآتية:

- Z = +1.5
- Z = -0.75
- الرتبة المئينية = 90

النشاط الرابع

ناقش العلاقة بين التحصيل الدراسي والدافعية للتعلم في ضوء مفهوم الارتباط.

الفصل السادس: الاختبارات النفسية والتربوية: مفهومها، خصائصها وأنواعها

أهداف التعلم

في نهاية هذا الفصل يكون الطالب قادراً على:

- التعرف على مفهوم الاختبار النفسي والتربوي .
- التمييز بين الاختبار والقياس والتقويم .
- التعرف على الخصائص العلمية للاختبارات الجيدة .
- تصنيف الاختبارات النفسية والتربوية وفق أسس مختلفة .
- فهم مجالات استخدام الاختبارات في التربية وعلم النفس .
- إدراك الاتجاهات الحديثة في بناء وتطبيق الاختبارات .

تمهيد:

تُعد الاختبارات النفسية والتربوية من أهم أدوات جمع البيانات في العلوم السلوكية، إذ تُستخدم لقياس القدرات العقلية والتحصيل الدراسي والاستعدادات والميول والاتجاهات والسمات الشخصية. وقد شهدت الاختبارات تطوراً كبيراً منذ ظهور أول اختبار للذكاء على يد ألفريد بينيه سنة 1905، وصولاً إلى الاختبارات المحوسبة التكيفية المعتمدة على النماذج السيكومترية الحديثة.(Cohen & Swerdlik, 2022)

وتؤدي الاختبارات دوراً محورياً في التشخيص والتصنيف والتوجيه واتخاذ القرارات التربوية والنفسية، مما يفرض ضرورة الالتزام بالمعايير العلمية في بنائها وتطبيقها وتفسير نتائجها.

أولاً: مفهوم الاختبار النفسي والتربوي

يعرف الاختبار بأنه أداة مقننة تتكون من مجموعة من المثيرات أو الأسئلة أو المواقف المعدة لقياس خاصية أو قدرة أو سمة معينة لدى الفرد بطريقة موضوعية.(Anastasi & Urbina, 2019)

كما يعرف بأنه إجراء منظم يهدف إلى الحصول على عينة من سلوك الفرد يمكن من خلالها تقدير مستوى الصفة أو القدرة المراد قياسها.(AERA, APA, & NCME, 2014)

ثانياً: العلاقة بين الاختبار والقياس والتقويم

كثيراً ما يحدث خلط بين هذه المفاهيم، رغم أن لكل منها وظيفة محددة.

-1 الاختبار

يمثل الأداة المستخدمة لجمع البيانات.

-2 القياس

هو عملية إعطاء قيم كمية للنتائج المتحصل عليها من الاختبار.(Stevens, 1951)

-3 التقييم

يتضمن تفسير نتائج القياس وإصدار أحكام حول مستوى الأداء.

-4 التقويم

عملية شاملة تهدف إلى تحسين الأداء واتخاذ القرارات المناسبة بناءً على نتائج التقييم (ملحم، 2015.(

ثالثاً: خصائص الاختبار الجيد

-1 الصدق(Validity)

يُقصد بالصدق مدى قدرة الاختبار على قياس السمة أو الخاصية التي صُمِّم من أجل قياسها بدقة وفعالية، بحيث تكون النتائج المتحصل عليها معبرة حقيقةً عن الظاهرة المستهدفة بالقياس. ويُعد الصدق من أهم الخصائص السيكومترية للاختبارات النفسية والتربوية، إذ لا يمكن الاعتماد على نتائج أي اختبار ما لم يكن صادقًا في قياس ما وُضع لقياسه(Messick, 1995) ، ص. 741.

-2 الثبات(Reliability)

يشير الثبات إلى درجة اتساق نتائج الاختبار واستقرارها عند إعادة تطبيقه على الأفراد أنفسهم في ظروف متشابهة. فكلما أعطى الاختبار نتائج متقاربة عبر مرات التطبيق المختلفة، دلّ ذلك على ارتفاع مستوى ثباته، مما يعزز الثقة في نتائجه وإمكانية الاعتماد عليها في اتخاذ القرارات التربوية أو النفسية (Crocker & Algina, 2008، ص. 142.

3- الموضوعية(Objectivity)

تعني الموضوعية أن تكون نتائج الاختبار بعيدة عن التأثيرات الشخصية للمصحح أو الفاحص، بحيث يحصل المفحوص على النتيجة نفسها مهما اختلف المصحح أو ظروف التصحيح. وتتحقق الموضوعية من خلال وضوح التعليمات، وتحديد معايير التصحيح بدقة، واعتماد أسئلة تقلل من التحيز الذاتي (Anastasi & Urbina, 1997، ص. 113.

4- القدرة التمييزية(Discrimination)

يقصد بالقدرة التمييزية مدى قدرة الاختبار أو فقراته على التمييز بين الأفراد ذوي المستويات المختلفة في السمة أو القدرة المقاسة. فالفقرة الجيدة هي التي يجيب عنها ذوو المستوى المرتفع بصورة تختلف عن ذوي المستوى المنخفض، مما يسمح بالكشف عن الفروق الفردية بين المتعلمين(Ebel & Frisbie, 1991) ، ص. 232.

5- العملية أو القابلية للتطبيق(Practicality)

تشير العملية إلى مدى سهولة استخدام الاختبار من حيث الإعداد والتطبيق والتصحيح وتفسير النتائج، مع مراعاة الاقتصاد في الوقت والجهد والتكلفة. ويُفضل أن يكون الاختبار عمليًا بحيث يحقق أهدافه بكفاءة دون الحاجة إلى إمكانات مادية أو بشرية كبيرة(Nitko & Brookhart, 2014) ، ص. 71.

رابعاً: أنواع الاختبارات النفسية والتربوية

أ- حسب الهدف

1- الاختبارات التحصيلية

تقيس مدى اكتساب المتعلم للمعارف والمهارات التي سبق تعلمها.(Nitko & Brookhart, 2018)

أمثلة:

- اختبارات السداسي: الاختبارات التي تُجرى في نهاية السداسي الدراسي لقياس مدى تحقق مخرجات التعلم الخاصة بالمقاييس المبرمجة خلال السداسي.

- الاختبارات النهائية: الامتحانات الختامية التي تُنظم في نهاية الفترة الدراسية وتُعد أحد المكونات الأساسية للتقويم التحصيلي للطالب.
- الاختبارات الفصلية: الاختبارات التي تُجرى خلال الفصل الدراسي أو في نهايته بهدف تقويم التحصيل الدراسي ومتابعة تقدم الطلبة.

2- اختبارات الاستعداد

تهدف إلى التنبؤ بالأداء المستقبلي للفرد في مجال معين.

- **اختبارات الاستعداد الجامعي(College Aptitude Tests)**

تُعد اختبارات الاستعداد الجامعي من الأدوات المقننة التي تهدف إلى التنبؤ بقدرة الطالب على النجاح في الدراسة الجامعية، من خلال قياس مجموعة من القدرات العقلية والمعرفية المرتبطة بالأداء الأكاديمي، مثل الاستدلال اللفظي، والاستدلال الكمي، والتفكير المنطقي، والقدرة على حل المشكلات. وتُستخدم نتائج هذه الاختبارات في عمليات القبول والتوجيه الجامعي، حيث تساعد المؤسسات التعليمية على اختيار الطلبة الذين يمتلكون المقومات اللازمة لمتابعة الدراسات العليا بنجاح.(Anastasi & Urbina, 1997)

وتختلف اختبارات الاستعداد الجامعي عن اختبارات التحصيل الدراسي في كونها لا تقيس ما تعلمه الفرد سابقًا، وإنما تركز على تقدير إمكاناته المستقبلية وقدرته على اكتساب المعارف والمهارات الأكاديمية الجديدة (Aiken, 2003).

- **اختبارات الاستعداد المهني(Vocational Aptitude Tests)**

هي اختبارات تُصمم للكشف عن قدرات الفرد واستعداداته المهنية التي تؤهله للنجاح في مهنة أو مجال عمل معين. وتهدف إلى مساعدة الأفراد في اختيار المهن التي تتناسب مع قدراتهم وميولهم، كما تُستخدم في عمليات التوجيه والإرشاد المهني والانتقاء الوظيفي. وتشمل هذه الاختبارات قياس العديد من القدرات، مثل القدرة الميكانيكية، والقدرة العددية، والقدرة اللفظية، والتناسق الحركي، والإدراك المكاني، وغيرها من القدرات المرتبطة بالأداء المهني.(Cronbach, 1990)

وتُسهم اختبارات الاستعداد المهني في تحقيق التوافق بين خصائص الفرد ومتطلبات المهنة، مما يزيد من فرص النجاح والرضا المهني ويرفع من كفاءة الأداء والإنتاجية في بيئة العمل.(Brown, 2002)

ب- حسب المحتوى

-1 اختبارات القدرة العقلية العامة

تقيس الذكاء أو القدرة العقلية العامة.

- **2- اختبارات القدرات الخاصة (Specific Aptitude Tests)**

تُعرف اختبارات القدرات الخاصة بأنها أدوات قياس صُممت لتقدير مستوى قدرة محددة أو مجموعة من القدرات النوعية لدى الفرد، والتي ترتبط بالأداء الناجح في مجالات أكاديمية أو مهنية معينة. وتختلف هذه الاختبارات عن اختبارات الذكاء العام في تركيزها على جوانب محددة من الأداء العقلي أو المهاري، مما يجعلها ذات أهمية كبيرة في عمليات التوجيه الدراسي والمهني والانتقاء الوظيفي (Anastasi & Urbina, 1997).

ومن أبرز أنواع اختبارات القدرات الخاصة ما يأتي:

أ- اختبارات القدرة اللفظية (Verbal Aptitude Tests)

تهدف إلى قياس قدرة الفرد على فهم الكلمات والرموز اللغوية واستخدامها بصورة صحيحة، بالإضافة إلى قدرته على الاستيعاب اللغوي والتعبير اللفظي والاستدلال القائم على اللغة. وتُعد هذه القدرة من المؤشرات المهمة للنجاح في التخصصات التي تعتمد على التواصل اللغوي والتحليل اللفظي (Aiken, 2003).

ب- اختبارات القدرة العددية (Numerical Aptitude Tests)

تقيس قدرة الفرد على التعامل مع الأعداد وإجراء العمليات الحسابية وفهم العلاقات الكمية وتفسير البيانات الرقمية. وتستخدم هذه الاختبارات بصورة واسعة في المجالات العلمية والتجارية والمهنية التي تتطلب مهارات حسابية وتحليلية متقدمة (Cronbach, 1990).

ج- اختبارات القدرة الميكانيكية (Mechanical Aptitude Tests)

تستهدف قياس فهم الفرد للمبادئ الميكانيكية والفيزيائية الأساسية، وقدرته على إدراك كيفية عمل الآلات والأدوات الميكانيكية والعلاقات بين أجزائها. وتُستخدم هذه الاختبارات في عمليات التوجيه والاختيار للمهن الصناعية والهندسية والتقنية (Anastasi & Urbina, 1997).

د- اختبارات القدرة المكانية (Spatial Aptitude Tests)

تهدف إلى قياس قدرة الفرد على إدراك العلاقات المكانية بين الأشكال والأجسام، وتخيل حركتها أو دورانها في الفراغ، وفهم الأبعاد الهندسية المختلفة. وتُعد هذه القدرة ذات أهمية خاصة في مجالات الهندسة والعمارة والتصميم والفنون التقنية (Kaplan & Saccuzzo, 2018).

ج- حسب طريقة التطبيق

تُصنف الاختبارات وفقًا لطريقة تطبيقها إلى اختبارات فردية وأخرى جماعية، وذلك تبعًا لعدد الأفراد الذين يُطبق عليهم الاختبار في الوقت نفسه والإجراءات المتبعة أثناء عملية القياس (Anastasi & Urbina, 1997).

01.الاختبارات الفردية(Individual Tests)

هي الاختبارات التي تُطبق على فرد واحد في كل مرة، حيث يقوم الفاحص بإجراء الاختبار بصورة مباشرة مع المفحوص، مما يسمح بملاحظة سلوكه واستجاباته بدقة أثناء عملية التطبيق. وتمتاز هذه الاختبارات بإمكانية الحصول على معلومات تفصيلية حول أداء الفرد، إلا أنها تتطلب وقتًا وجهدًا أكبر مقارنة بالاختبارات الجماعية.(Kaplan & Saccuzzo, 2018)

ومن أمثلتها:

اختبار ستانفورد–بينيه للذكاء .

مقياس وكسلر لذكاء الراشدين .

مقياس وكسلر لذكاء الأطفال .

02.الاختبارات الجماعية(Group Tests)

هي الاختبارات التي تُطبق على مجموعة من الأفراد في وقت واحد وفق تعليمات موحدة وإجراءات مقننة، مما يجعلها أكثر اقتصادًا في الوقت والجهد والتكلفة. وتُستخدم هذه الاختبارات على نطاق واسع في المؤسسات التعليمية وعمليات الانتقاء والتوجيه المهني، خاصة عندما يكون عدد المفحوصين كبيرًا (Cronbach, 1990).

ومن أمثلتها:

اختبارات التحصيل الدراسي المدرسية والجامعية

اختبارات القبول الجامعي .

اختبارات القدرات والاستعدادات الأكاديمية واسعة النطاق .

الاختبارات المعيارية التي تُطبق على أعداد كبيرة من الطلبة في الوقت نفسه .

د- حسب طريقة الإجابة

تُصنَّف الاختبارات من حيث طبيعة الاستجابة المطلوبة من المفحوص إلى اختبارات مقالية واختبارات موضوعية. ويستند هذا التصنيف إلى الكيفية التي يُعبِّر بها المفحوص عن إجاباته ومدى حرية الاستجابة المتاحة له.(Nitko & Brookhart, 2018)

1-الاختبارات المقالية(Essay Tests)

هي الاختبارات التي تتطلب من المفحوص تقديم إجابات إنشائية أو وصفية مطولة باستخدام أسلوبه الخاص، حيث يُطلب منه شرح الأفكار أو تحليلها أو مناقشتها أو المقارنة بينها. وتُستخدم هذه الاختبارات بشكل واسع في المواد الإنسانية والاجتماعية التي تتطلب مستويات عليا من التفكير (Ebel & Frisbie, 1991).

مميزات الاختبارات المقالية

- تقيس القدرات العقلية العليا مثل التحليل والتركيب والتقويم والتفكير النقدي (Nitko & Brookhart, 2018).
- تسمح للمفحوص بحرية التعبير عن أفكاره وتنظيمها بأسلوبه الخاص .

- تساعد في الكشف عن قدرة المتعلم على الربط بين المعارف والمفاهيم المختلفة .
- تمكن من تقويم مهارات الكتابة والتعبير والاستدلال المنطقي .

عيوب الاختبارات المقالية

- انخفاض مستوى الموضوعية في التصحيح نتيجة تأثر التقدير بالحكم الشخصي للمصحح (Anastasi & Urbina, 1997).
- الحاجة إلى وقت طويل في التصحيح وإعطاء الدرجات .
- محدودية عدد الأسئلة التي يمكن تضمينها في الاختبار .
- تأثر النتائج أحيانًا بمهارات الكتابة والتعبير أكثر من تأثرها بالمحتوى العلمي نفسه .

2- الاختبارات الموضوعية(Objective Tests)

هي الاختبارات التي تكون إجاباتها محددة مسبقًا، بحيث لا تترك مجالًا كبيرًا للاجتهاد الشخصي أثناء التصحيح. وتتميز بإمكانية تصحيحها وفق معايير دقيقة وموحدة، مما يزيد من موضوعيتها وثباتها (Nitko & Brookhart, 2018).

وتشمل الاختبارات الموضوعية:

أ- أسئلة الاختيار من متعدد(Multiple Choice Questions)
تتكون من جذع للسؤال وعدد من البدائل، ويُطلب من المفحوص اختيار البديل الصحيح أو الأفضل.

ب- أسئلة الصواب والخطأ(True–False Questions)
يُطلب فيها من المفحوص تحديد ما إذا كانت العبارة المقدمة صحيحة أم خاطئة.

ج- أسئلة المزاوجة أو المطابقة(Matching Questions)
تتضمن قائمتين من العناصر، ويُطلب من المفحوص الربط بين العناصر المتناظرة فيهما.

د- أسئلة الإكمال(Completion Questions)
يُطلب فيها من المفحوص إكمال كلمة أو عبارة أو جملة ناقصة للوصول إلى الإجابة الصحيحة.

مميزات الاختبارات الموضوعية

- ارتفاع مستوى الموضوعية في التصحيح .
- ارتفاع معاملات الثبات مقارنة بالاختبارات المقالية .(Nitko & Brookhart, 2018)
- إمكانية تغطية جزء أكبر من محتوى المقرر الدراسي .
- سرعة التصحيح وإمكانية استخدام الوسائل الإلكترونية في رصد النتائج .
- سهولة إجراء التحليل الإحصائي للفقرات والاختبار ككل .

عيوب الاختبارات الموضوعية

- صعوبة إعداد فقرات جيدة وعالية الجودة .
- محدودية قدرتها على قياس بعض جوانب الإبداع والتعبير الحر .
- احتمال نجاح بعض المفحوصين عن طريق التخمين، خاصة في بعض أنواع الأسئلة.

هـ- حسب المرجع المستخدم في تفسير النتائج

تُصنف الاختبارات وفق المرجع المعتمد في تفسير درجات المفحوصين إلى اختبارات معيارية المرجع واختبارات محكية المرجع. ويعتمد هذا التصنيف على الأساس الذي تُفسَّر في ضوئه النتائج، سواء بالمقارنة مع أداء الآخرين أو بالرجوع إلى مستوى أداء محدد مسبقًا.(Cohen & Swerdlik, 2022)

1- الاختبارات معيارية المرجع(Norm-Referenced Tests)

هي الاختبارات التي يتم فيها تفسير أداء الفرد بمقارنته بأداء أفراد مجموعة مرجعية أو معيارية تمثل المجتمع الذي ينتمي إليه. وتهدف هذه الاختبارات إلى تحديد الموقع النسبي للفرد داخل الجماعة، أي معرفة ما إذا كان أداؤه أعلى أو أقل أو مساوياً لمتوسط أداء أقرانه.(Cohen & Swerdlik, 2022)

ويُعبر عن نتائج هذه الاختبارات غالبًا بالرتب المئينية أو الدرجات المعيارية أو الانحرافات المعيارية، مما يسمح بإجراء المقارنات بين الأفراد والجماعات. وتُستخدم الاختبارات معيارية المرجع بكثرة في اختبارات

الذكاء والاستعداد والتحصيل المقنن وعمليات الانتقاء والتصنيف التربوي والمهني (Kaplan & Saccuzzo, 2018).

خصائص الاختبارات معيارية المرجع

- تعتمد على المقارنة بين أداء الفرد وأداء المجموعة المرجعية .
- تهدف إلى ترتيب الأفراد وفق مستويات أدائهم .
- تساعد في الكشف عن الفروق الفردية بين المفحوصين .
- تُستخدم في الانتقاء والتصنيف والتوجيه التربوي والمهني .

مثال: إذا حصل طالب على درجة تضعه في الرتبة المئينية (80)، فهذا يعني أن أداءه يفوق أداء 80% من أفراد المجموعة المرجعية التي استُخدمت في تقنين الاختبار.

2- الاختبارات محكية المرجع(Criterion-Referenced Tests)

هي الاختبارات التي يتم فيها تفسير أداء المفحوص بالاعتماد على معيار أو مستوى إتقان محدد مسبقًا، دون مقارنته بأداء الآخرين. ويُركز هذا النوع من الاختبارات على تحديد مدى تمكن الفرد من المعارف أو المهارات أو الكفايات المستهدفة بغض النظر عن أداء بقية الأفراد.(Glaser, 1963)

ويُستخدم هذا النوع على نطاق واسع في المجال التربوي، خاصة في تقويم التعلم القائم على الكفاءات، حيث يتم تحديد مستوى معين من الإتقان يُعد معيارًا للحكم على نجاح المتعلم أو حاجته إلى مزيد من التعلم والتدريب.(Nitko & Brookhart, 2018)

خصائص الاختبارات محكية المرجع

- لا تعتمد على مقارنة الفرد بالآخرين .
- تركز على مدى تحقيق أهداف التعلم أو مستوى الإتقان المطلوب .
- تُستخدم للحكم على كفاءة المتعلم في مهارة أو مجال محدد .
- تساعد في تشخيص مواطن القوة والضعف واتخاذ قرارات علاجية أو تطويرية.

مثال

إذا حُددت نسبة 80% كمعيار للإتقان في اختبار الرياضيات، فإن الطالب الذي يحصل على هذه النسبة أو يتجاوزها يُعد متقنًا للمهارات المطلوبة، بغض النظر عن أداء بقية الطلبة.

الفرق بين الاختبارات معيارية المرجع ومحكية المرجع

وجه المقارنة	الاختبارات معيارية المرجع	الاختبارات محكية المرجع
أساس التفسير	مقارنة الفرد بالآخرين	مقارنة الأداء بمعيار محدد
الهدف	تحديد الموقع النسبي للفرد	تحديد مستوى الإتقان
طبيعة النتائج	ترتيب وتصنيف الأفراد	الحكم على تحقيق الأهداف
مجال الاستخدام	الانتقاء والتصنيف والتوجيه	التقويم التربوي وتشخيص التعلم

خامساً: استخدامات الاختبارات النفسية والتربوية

تستخدم الاختبارات في:

- التشخيص النفسي .
- التوجيه والإرشاد .
- الانتقاء والتوظيف .
- تصنيف الطلبة .
- الكشف عن صعوبات التعلم .
- تقويم البرامج التعليمية .

- البحث العلمي .

ويؤكد Cohen و Swerdlik (2022) أن نجاح القرارات التربوية والنفسية يعتمد بدرجة كبيرة على جودة الاختبارات المستخدمة.

سادساً: أخلاقيات استخدام الاختبارات

تتطلب الاختبارات النفسية والتربوية الالتزام بعدد من المبادئ الأخلاقية أهمها:

- احترام خصوصية المفحوص .
- سرية النتائج .
- العدالة وعدم التحيز .
- التطبيق الصحيح للاختبار .
- تفسير النتائج من قبل مختصين مؤهلين .

وقد أكدت المعايير الدولية للاختبارات النفسية والتربوية على ضرورة مراعاة هذه المبادئ في جميع مراحل القياس.(AERA et al., 2014)

سابعاً: الاتجاهات الحديثة في الاختبارات

شهدت الاختبارات خلال العقود الأخيرة تطورات مهمة من أبرزها:

1-الاختبارات الإلكترونية

أصبحت المؤسسات التعليمية تعتمد بصورة متزايدة على الاختبارات الرقمية.

2-الاختبارات المحوسبة التكيفية(CAT)

تُكيف فقرات الاختبار تلقائياً وفق مستوى أداء المفحوص.(Van der Linden, 2018)

3-توظيف الذكاء الاصطناعي

يُستخدم الذكاء الاصطناعي حالياً في:

- بناء بنوك الأسئلة .
- تحليل الاستجابات .
- الكشف عن الغش .
- تطوير الاختبارات التكيفية .

4-الاختبارات القائمة على الكفاءة

تركز على قياس الكفاءات الفعلية بدلاً من مجرد قياس المعرفة النظرية.

خلاصة الفصل

تُعد الاختبارات النفسية والتربوية من أهم أدوات القياس المستخدمة في العلوم السلوكية والتربوية، وتتنوع وفق أهدافها ومحتواها وطرق تطبيقها وتفسير نتائجها. كما أن جودة الاختبار تعتمد على توفر الصدق والثبات والموضوعية والتمييز، إضافة إلى الالتزام بالمعايير الأخلاقية والمهنية في استخدامه.

أسئلة للمراجعة

1. عرف الاختبار النفسي والتربوي .
2. ما الفرق بين الاختبار والقياس والتقويم؟
3. ما المقصود بصدق الاختبار وثباته؟
4. صنف الاختبارات وفق طريقة التطبيق .
5. ما الفرق بين الاختبارات معيارية المرجع ومحكية المرجع؟
6. ما أهم الاستخدامات التربوية للاختبارات؟
7. ما أبرز الاتجاهات الحديثة في مجال الاختبارات؟

النشاط الأول

اختر اختباراً دراسياً مستخدماً في مؤسستك التعليمية وحدد نوعه وفق التصنيفات المختلفة الواردة في الفصل.

النشاط الثاني

قارن بين الاختبارات المقالية والاختبارات الموضوعية من حيث:

- الصدق .
- الثبات .
- الموضوعية .
- سهولة التصحيح .

النشاط الثالث

ناقش أثر التحول الرقمي على مستقبل الاختبارات النفسية والتربوية.

الفصل السابع

بناء الاختبارات النفسية والتربوية وخطوات إعدادها

أهداف التعلم

في نهاية هذا الفصل يكون الطالب قادراً على:

- التعرف على مفهوم بناء الاختبارات النفسية والتربوية .
- فهم المراحل العلمية لإعداد الاختبارات .
- تحديد الأهداف التعليمية وصياغتها بصورة سليمة .
- إعداد جدول المواصفات للاختبارات التحصيلية .
- صياغة فقرات الاختبار وفق المعايير العلمية .
- التعرف على الإجراءات الأولية لتجريب الاختبار وتقنينه .

الفصل الثامن: بناء الاختبارات النفسية.

تمهيد

يمثل بناء الاختبارات النفسية والتربوية عملية علمية منظمة تهدف إلى تطوير أدوات قياس تتمتع بدرجات مرتفعة من الصدق والثبات والموضوعية. ولا يقتصر إعداد الاختبار على جمع مجموعة من الأسئلة، بل يتطلب المرور بمراحل منهجية تبدأ بتحديد الهدف من القياس وتنتهي بتحليل الخصائص السيكومترية للأداة وتفسير نتائجها.(Nitko & Brookhart, 2018)

ويُعد بناء الاختبارات من أهم الكفاءات التي ينبغي أن يمتلكها المختص في علوم التربية وعلم النفس، نظراً لما يترتب على نتائج الاختبارات من قرارات تربوية ونفسية ومهنية مهمة.(Cohen & Swerdlik, 2022)

أولاً: مفهوم بناء الاختبار

يقصد ببناء الاختبار مجموعة الإجراءات العلمية المنظمة التي تهدف إلى تصميم أداة قياس قادرة على جمع بيانات دقيقة حول سمة أو قدرة أو تحصيل معين، وفق معايير سيكومترية محددة (Anastasi & Urbina, 2019).

ويهدف بناء الاختبار إلى:

- قياس السمة أو القدرة المستهدفة .
- توفير بيانات دقيقة لاتخاذ القرار .
- التنبؤ بالأداء المستقبلي .
- تشخيص نقاط القوة والضعف .

ثانياً: مراحل بناء الاختبارات النفسية والتربوية

تشير الأدبيات الحديثة إلى أن بناء الاختبارات يمر بعدة مراحل مترابطة.(AERA et al., 2014)

المرحلة الأولى: تحديد الهدف من الاختبار

تُعد هذه المرحلة الأساس الذي تُبنى عليه جميع المراحل اللاحقة.

ويجب الإجابة عن الأسئلة الآتية:

- ماذا نريد أن نقيس؟
- من هي الفئة المستهدفة؟
- ما طبيعة القرارات التي ستبنى على النتائج؟

أمثلة:

- قياس التحصيل الدراسي .
- قياس الذكاء .
- قياس الاتجاهات .
- قياس الدافعية .

المرحلة الثانية: تحديد المحتوى المراد قياسه

بعد تحديد الهدف، يتم تحديد المجالات أو الأبعاد التي يتكون منها المحتوى.

مثال:

إذا كان الهدف بناء اختبار في مادة القياس والتقويم، فقد يشمل المحتوى:

- مفاهيم القياس .
- مستويات القياس .
- الصدق والثبات .
- بناء الاختبارات .

ويجب أن يكون المحتوى ممثلاً تمثيلاً دقيقاً لجميع جوانب المجال المراد قياسه (Nitko & Brookhart, 2018).

المرحلة الثالثة: صياغة الأهداف التعليمية

ينبغي أن تُصاغ الأهداف التعليمية بصورة سلوكية واضحة وقابلة للقياس.

ويُفضل الاعتماد على تصنيف بلوم للمجال المعرفي.(Anderson & Krathwohl, 2001)

ومن أمثلة الأهداف:

- أن يعرّف الطالب مفهوم الصدق .
- أن يميز بين أنواع الثبات .
- أن يحلل نتائج اختبار تحصيلي .
- أن يصمم جدول مواصفات .

ثالثاً: جدول المواصفات

يُعد جدول المواصفات من أهم الأدوات المستخدمة لضمان صدق المحتوى.

ويعرف بأنه مخطط يربط بين:

- محتوى المادة الدراسية .
- الأهداف التعليمية .
- عدد الأسئلة المخصصة لكل موضوع ومستوى معرفي .

ويرى عودة (2010) أن جدول المواصفات يمثل "الخريطة الهندسية" للاختبار.

نموذج مبسط لجدول المواصفات

المحتوى	المعرفة	الفهم	التطبيق	المجموع
القياس	2	1	1	4

المحتوى	المعرفة	الفهم	التطبيق	المجموع
الصدق	2	2	1	5
الثبات	1	2	2	5
المجموع	5	5	4	14

رابعاً: إعداد فقرات الاختبار

بعد الانتهاء من جدول المواصفات تبدأ عملية إعداد الفقرات.

ويجب أن تراعي الفقرة الجيدة الشروط الآتية:(Haladyna, 2016)

- الوضوح اللغوي .
- الدقة العلمية .
- خلوها من الغموض .
- مناسبة مستوى الصعوبة .
- الارتباط المباشر بالهدف .

خامساً: قواعد صياغة أسئلة الاختيار من متعدد

تُعد أسئلة الاختيار من متعدد من أكثر أنواع الفقرات استخداماً في الاختبارات التحصيلية.

شروط إعدادها

1. أن تكون صياغة السؤال واضحة .
2. أن تكون الإجابة الصحيحة واحدة فقط .
3. أن تكون البدائل متقاربة في الطول .
4. تجنب العبارات المضللة .

5. تجنب استخدام "جميع ما سبق" قدر الإمكان .

مثال

أي مما يأتي يمثل أعلى مستويات القياس؟

أ- الاسمي

ب- الرتبي

ج- الفئوي

د- النسبي ✓

سادساً: مراجعة الاختبار وتحكيمه

قبل تطبيق الاختبار يجب عرضه على مجموعة من الخبراء والمتخصصين.

ويهدف التحكيم إلى:

- التحقق من سلامة المحتوى .
- التأكد من مناسبة الفقرات .
- الكشف عن الأخطاء اللغوية والعلمية .
- تقدير صدق المحتوى .

ويُعد التحكيم أحد الإجراءات الأساسية لتعزيز صدق الاختبار.(AERA et al., 2014)

سابعاً: التجريب الأولي للاختبار

يقصد به تطبيق الاختبار على عينة استطلاعية مشابهة للفئة المستهدفة.

ويهدف إلى:

- الكشف عن الفقرات غير الواضحة .

- تقدير الزمن اللازم للإجابة .
- حساب معاملات الصعوبة والتمييز .
- تقدير الثبات والصدق الأوليين .

ثامناً: الصورة النهائية للاختبار

بعد الانتهاء من التحليل الإحصائي تُعد الصورة النهائية للاختبار وتتضمن:

- تعليمات التطبيق .
- عدد الفقرات .
- طريقة التصحيح .
- مفاتيح الإجابة .
- أساليب تفسير النتائج .

تاسعاً: أخطاء شائعة في بناء الاختبارات

من أكثر الأخطاء التي يقع فيها الباحثون:

- عدم تحديد الأهداف بدقة .
- ضعف تمثيل المحتوى .
- صياغة أسئلة غامضة .
- إهمال جدول المواصفات .
- عدم إجراء التجريب الأولي .
- الاعتماد على عدد محدود من الفقرات .

ويؤدي ذلك إلى انخفاض صدق الاختبار وثباته.(DeVellis & Thorpe, 2021)

عاشراً: الاتجاهات الحديثة في بناء الاختبارات

شهدت عملية بناء الاختبارات تطورات مهمة تمثلت في:

1-بنوك الأسئلة الإلكترونية

تسمح بتخزين أعداد كبيرة من الفقرات المصنفة وفق خصائصها السيكومترية.

2-الاختبارات التكيفية المحوسبة

تعتمد على نظرية الاستجابة للمفردة وتُكيف مستوى الأسئلة وفق أداء المفحوص (Van der Linden, 2018).

3-الذكاء الاصطناعي

أصبح يُستخدم في:

- توليد الفقرات .
- تحليل الاستجابات .
- تطوير بنوك الأسئلة .
- الكشف عن الغش الأكاديمي .

4-الاختبارات القائمة على الكفاءات

تركز على قياس الأداء الحقيقي للمتعلم بدلاً من الاقتصار على المعرفة النظرية.

خلاصة الفصل

يمثل بناء الاختبارات النفسية والتربوية عملية علمية منظمة تبدأ بتحديد الهدف من القياس وتنتهي بإعداد أداة تتمتع بخصائص سيكومترية مناسبة. ويُعد الالتزام بالمراحل العلمية لبناء الاختبار شرطاً أساسياً للحصول على نتائج دقيقة يمكن الاعتماد عليها في اتخاذ القرارات التربوية والنفسية.

أسئلة للمراجعة

1. ما المقصود ببناء الاختبار؟
2. ما المراحل الأساسية لبناء الاختبارات النفسية والتربوية؟
3. ما أهمية جدول المواصفات؟
4. ما شروط صياغة الفقرة الاختبارية الجيدة؟
5. ما المقصود بالتجريب الأولي؟
6. ما دور التحكيم في بناء الاختبار؟
7. ما أبرز الاتجاهات الحديثة في بناء الاختبارات؟

أنشطة تطبيقية

النشاط الأول

قم بإعداد ثلاثة أهداف تعليمية في موضوع "القياس النفسي" وفق مستويات بلوم المعرفية.

النشاط الثاني

أنشئ جدول مواصفات لاختبار تحصيلي يتكون من 20 سؤالاً.

النشاط الثالث

صغ خمس فقرات من نوع الاختيار من متعدد وفق المعايير العلمية.

النشاط الرابع

حلل اختباراً جامعياً مستخدماً في تخصصك وحدد نقاط القوة والضعف فيه.

الفصل الثامن: صدق الاختبارات النفسية والتربوية(Validity)

أهداف التعلم

في نهاية هذا الفصل يكون الطالب قادراً على:

- تعريف مفهوم الصدق في القياس النفسي والتربوي .
- التمييز بين الصدق والثبات .
- التعرف على أنواع الصدق المختلفة .
- حساب وتفسير بعض مؤشرات الصدق .
- فهم إجراءات التحقق من صدق الاختبارات والمقاييس .
- توظيف أساليب الصدق في بناء الاختبارات النفسية والتربوية .

تمهيد:

يُعد الصدق من أهم الخصائص السيكومترية للاختبارات النفسية والتربوية، بل إن العديد من المختصين يعتبرونه الخاصية الأساسية التي تحدد قيمة الاختبار وجدواه العلمية. فالاختبار مهما بلغت درجة ثباته أو دقته لا تكون له قيمة علمية إذا كان لا يقيس السمة أو القدرة التي صُمم من أجل قياسها (Messick, 1995).

وقد شهد مفهوم الصدق تطوراً ملحوظاً خلال العقود الأخيرة، حيث انتقل من كونه مجموعة أنواع منفصلة إلى إطار نظري متكامل يعتمد على الأدلة والبراهين التي تدعم تفسير نتائج الاختبار واستخداماتها (AERA et al., 2014).

أولاً: مفهوم الصدق

يعرف الصدق بأنه مدى قدرة الاختبار على قياس الخاصية أو السمة التي وُضع لقياسها دون غيرها (Anastasi & Urbina, 2019).

كما يعرف بأنه درجة ملاءمة التفسيرات والاستنتاجات والقرارات المبنية على نتائج الاختبار للأغراض التي أُعد من أجلها

(AERA et al., 2014).

ويشير Cohen و(2022) Swerdlik إلى أن السؤال الأساسي في دراسة الصدق هو:

"هل يقيس الاختبار فعلاً ما يفترض أن يقيسه؟"

ثانياً: أهمية الصدق في القياس النفسي والتربوي

تكمن أهمية الصدق في كونه:

- يحدد جودة الاختبار .
- يسمح بتفسير النتائج بصورة سليمة .
- يدعم القرارات التربوية والنفسية .
- يزيد من موثوقية نتائج البحوث .
- يساعد في تحسين أدوات القياس .

ويرى Messick (1995) أن الصدق يمثل الأساس العلمي لجميع استخدامات الاختبار.

ثالثاً: العلاقة بين الصدق والثبات

يرتبط الصدق والثبات ارتباطاً وثيقاً، إلا أنهما ليسا مفهوماً واحداً.

العلاقة بين الثبات والصدق في الاختبارات

يُعد الثبات والصدق من أهم الخصائص السيكومترية التي ينبغي توافرها في الاختبارات النفسية والتربوية، إذ ترتبط جودة الاختبار بدرجة توافر هاتين الخاصيتين معًا.(Crocker & Algina, 2008)

1- الثبات(Reliability)

يشير الثبات إلى درجة اتساق النتائج واستقرارها عند إعادة تطبيق الاختبار على الأفراد أنفسهم في ظروف متشابهة. فكلما أعطى الاختبار نتائج متقاربة عبر مرات القياس المختلفة، دلّ ذلك على ارتفاع مستوى ثباته. ويُعبر الثبات عن مدى خلو نتائج الاختبار من الأخطاء العشوائية في القياس (Crocker & Algina, 2008).

2-الصدق(Validity)

يشير الصدق إلى مدى دقة الاختبار في قياس السمة أو الخاصية التي صُمِّم لقياسها، أي مدى تمثيل الدرجات المتحصل عليها للظاهرة المستهدفة بالقياس. ويُعد الصدق الغاية الأساسية لأي عملية قياس، لأنه يحدد مدى صلاحية النتائج للتفسير واتخاذ القرارات التربوية أو النفسية بناءً عليها.(Messick, 1995)

3-العلاقة بين الثبات والصدق

ترتبط خاصيتا الثبات والصدق بعلاقة وثيقة؛ فالثبات يُعد شرطًا أساسيًا لتحقيق الصدق، ولكنه ليس شرطًا كافيًا بمفرده. فقد يكون الاختبار ثابتًا إذا أعطى النتائج نفسها في كل مرة، لكنه قد يقيس سمة مختلفة عن السمة المستهدفة، وبالتالي يكون غير صادق. أما الاختبار الصادق فلا بد أن يكون ثابتًا بدرجة مقبولة، لأن النتائج المتذبذبة وغير المستقرة لا يمكن أن تمثل بدقة الخاصية المراد قياسها (Crocker & Algina, 2008).

وبعبارة أخرى، يمكن أن يكون الاختبار ثابتًا دون أن يكون صادقًا، لكن لا يمكن أن يكون صادقًا إذا كان غير ثابت .(Crocker & Algina, 2008)

مثال توضيحي

إذا استُخدم ميزان يعطي الوزن نفسه للشخص في كل مرة، لكنه يزيد دائمًا خمسة كيلوغرامات عن الوزن الحقيقي، فإن هذا الميزان يتمتع بدرجة عالية من الثبات لأنه يعطي نتائج متسقة، لكنه يفتقر إلى الصدق لأنه لا يقيس الوزن الحقيقي بدقة. أما إذا كان الميزان يعطي قراءات مختلفة في كل مرة للشخص نفسه، فإنه يكون غير ثابت، وبالتالي لا يمكن اعتباره صادقًا.

رابعاً: أنواع الصدق

1-الصدق الظاهري(Face Validity)

يشير إلى مدى ظهور الاختبار أمام المفحوص أو الخبراء وكأنه يقيس السمة المستهدفة.

مثال: اختبار يحتوي على مسائل رياضية لقياس القدرة الحسابية.

ورغم أهميته العملية، إلا أنه لا يُعد دليلاً علمياً كافياً على صدق الاختبار (Nitko & Brookhart, 2018).

2- صدق المحتوى(Content Validity)

يقصد به مدى تمثيل فقرات الاختبار لجميع جوانب المحتوى المراد قياسه.(Haynes et al., 1995)

ويُعد أكثر أنواع الصدق استخداماً في الاختبارات التحصيلية.

إجراءات التحقق منه

- تحديد المحتوى بدقة .
- إعداد جدول المواصفات .
- عرض الاختبار على المحكمين والخبراء .
- مراجعة الفقرات وتعديلها .

مثال: اختبار في مادة القياس والتقويم يجب أن يغطي جميع موضوعات المقرر وليس جزءاً منها فقط.

3- الصدق المرتبط بالمحك(Criterion-Related Validity)

يقيس مدى ارتباط نتائج الاختبار بمحك خارجي معتمد.

وينقسم إلى:

أ- الصدق التلازمي(Concurrent Validity)

يتم فيه مقارنة نتائج الاختبار بنتائج أداة أخرى مقننة تُطبق في الوقت نفسه.

مثال:

مقارنة نتائج اختبار ذكاء جديد بنتائج اختبار ذكاء معتمد.

ب- الصدق التنبؤي(Predictive Validity)

يقيس قدرة الاختبار على التنبؤ بالأداء المستقبلي.

مثال:

قدرة اختبار القبول الجامعي على التنبؤ بالمعدل الجامعي للطالب لاحقاً.

ويُقاس غالباً باستخدام معامل الارتباط.(Cohen & Swerdlik, 2022)

4- صدق البناء(Construct Validity)

يُعد أكثر أنواع الصدق شمولاً وأهمية في القياس النفسي الحديث.(Messick, 1995)

ويقصد به مدى نجاح الاختبار في قياس البناء النظري أو السمة النفسية المستهدفة.

أمثلة:

- الذكاء .
- القلق .
- الدافعية .
- تقدير الذات .

ويتم التحقق منه من خلال:

- التحليل العاملي .
- الارتباطات النظرية المتوقعة .
- الدراسات التجريبية .

خامساً: الأدلة الحديثة للصدق

وفق المعايير الحديثة للاختبارات النفسية والتربوية، يتم جمع أدلة الصدق من عدة مصادر (AERA et al., 2014):

1-الأدلة المستندة إلى المحتوى.

2-الأدلة المستندة إلى عمليات الاستجابة.

3-الأدلة المستندة إلى البناء الداخلي للاختبار.

4-الأدلة المستندة إلى العلاقات مع المتغيرات الأخرى.

5-الأدلة المستندة إلى نتائج استخدام الاختبار.

سادساً: معاملات الصدق

يُعبر عن الصدق غالباً بمعامل ارتباط يتراوح بين:

$$-1 \leq r \leq +1$$

وكلما اقتربت قيمة المعامل من (+1) دل ذلك على ارتفاع الصدق.

مثال

إذا بلغ معامل الصدق:

$$r = 0.82$$

فإن ذلك يشير إلى مستوى مرتفع من الصدق.

أما إذا بلغ:

$$r = 0.35$$

فإن الصدق يُعد ضعيفاً نسبياً.

سابعاً: العوامل المؤثرة في الصدق

يتأثر الصدق بعدة عوامل منها:

عوامل مرتبطة بالاختبار

- ضعف صياغة الفقرات .
- قلة عدد الأسئلة .
- عدم تمثيل المحتوى .

عوامل مرتبطة بالمفحوص

- القلق .
- التعب .
- انخفاض الدافعية .

عوامل مرتبطة بالتطبيق

- سوء التعليمات .
- الظروف البيئية غير المناسبة.

(Crocker & Algina, 2008)

ثامناً: الصدق في ضوء النظريات السيكومترية الحديثة

أصبح مفهوم الصدق في القياس الحديث أكثر شمولاً مع ظهور:

- نظرية الاستجابة للمفردة .(IRT)
- النمذجة البنائية .
- التحليل العاملي التوكيدي .

- الاختبارات المحوسبة التكيفية .

وتؤكد هذه الاتجاهات أن الصدق لا يرتبط بالاختبار نفسه، بل بتفسير النتائج واستخداماتها المختلفة.

(Embretson & Reise, 2020).

تاسعاً: تطبيقات الصدق في المجال التربوي

يُستخدم الصدق في:

- بناء الاختبارات التحصيلية .
- إعداد مقاييس الاتجاهات .
- تطوير اختبارات الذكاء .
- تصميم أدوات التقويم التربوي .
- إعداد المقاييس النفسية .

كما تعتمد الهيئات التعليمية وهيئات الاعتماد الأكاديمي على مؤشرات الصدق عند تقييم أدوات القياس المستخدمة في المؤسسات التعليمية.

خلاصة الفصل

يمثل الصدق حجر الأساس في القياس النفسي والتربوي، إذ يحدد مدى قدرة الاختبار على قياس السمة أو القدرة المستهدفة. وتتنوع أدلة الصدق بين صدق المحتوى والصدق المرتبط بالمحك وصدق البناء، بينما تؤكد الاتجاهات الحديثة أن الصدق يرتبط أساساً بصحة تفسير النتائج واستخداماتها المختلفة.

أسئلة للمراجعة

1. ما المقصود بالصدق؟
2. ما الفرق بين الصدق والثبات؟
3. ما أنواع الصدق التقليدية؟
4. كيف يتم التحقق من صدق المحتوى؟

5. ما المقصود بصدق البناء؟
6. ما الفرق بين الصدق التلازمي والصدق التنبؤي؟
7. ما مصادر أدلة الصدق وفق المعايير الحديثة؟

أنشطة تطبيقية

النشاط الأول

اختر اختباراً تحصيلياً مستخدماً في مؤسستك التعليمية وحدد نوع أدلة الصدق المتوافرة فيه.

النشاط الثاني

قم بإعداد نموذج لتحكيم محتوى اختبار تحصيلي من قبل مجموعة من الخبراء.

النشاط الثالث

ابحث في دراسة منشورة وحدد الإجراءات التي استخدمها الباحث للتحقق من صدق أداة الدراسة.

النشاط الرابع

ناقش أثر ضعف الصدق على القرارات التربوية والنفسية.

تاسعاً: تطبيقات الصدق في المجال التربوي

يُستخدم الصدق في:

- بناء الاختبارات التحصيلية .
- إعداد مقاييس الاتجاهات .
- تطوير اختبارات الذكاء .
- تصميم أدوات التقويم التربوي .
- إعداد المقاييس النفسية .

كما تعتمد الهيئات التعليمية وهيئات الاعتماد الأكاديمي على مؤشرات الصدق عند تقييم أدوات القياس المستخدمة في المؤسسات التعليمية.

. الفصل التاسع: ثبات الاختبارات النفسية والتربوية وطرق تقديره

أهداف التعلم

في نهاية هذا الفصل يكون الطالب قادراً على:

- تعريف مفهوم الثبات في القياس النفسي والتربوي .
- التمييز بين الثبات والصدق .
- التعرف على مصادر الخطأ المؤثرة في الثبات .
- حساب معاملات الثبات وتفسيرها .
- التمييز بين طرق تقدير الثبات المختلفة .
- تقييم جودة الاختبارات والمقاييس في ضوء معاملات الثبات .

تمهيد

إذا كان الصدق يجيب عن السؤال: "هل يقيس الاختبار ما وُضع لقياسه؟"، فإن الثبات يجيب عن السؤال: "هل يعطي الاختبار النتائج نفسها تقريباً إذا أعيد تطبيقه في الظروف نفسها؟ " (Cohen & Swerdlik, 2022).

ويُعد الثبات من أهم الخصائص السيكومترية للاختبارات النفسية والتربوية، لأنه يعكس درجة اتساق القياس واستقراره. فكلما ارتفع ثبات الاختبار ازدادت الثقة في نتائجه وإمكانية الاعتماد عليها في اتخاذ القرارات التربوية والنفسية.(Anastasi & Urbina, 2019)

أولاً: مفهوم الثبات

يعرف الثبات بأنه درجة الاتساق أو الاستقرار في نتائج الاختبار عند تكرار تطبيقه في ظروف متشابهة على الأفراد أنفسهم.(Crocker & Algina, 2008)

كما يُعرَّف بأنه مدى خلو الدرجات من الأخطاء العشوائية للقياس.

ويشير ملحم (2015) إلى أن الاختبار الثابت هو الاختبار الذي يعطي نتائج متقاربة كلما أعيد تطبيقه على الأفراد أنفسهم في ظروف متشابهة.

ثانياً: أهمية الثبات

تتجلى أهمية الثبات في كونه:

- يزيد من الثقة في نتائج الاختبار .
- يساعد على دقة التفسير .
- يرفع من جودة القرارات التربوية .
- يعد شرطاً أساسياً لتحقيق الصدق .
- يسهم في تحسين الخصائص السيكومترية للاختبارات .

وقد أكدت الأدبيات السيكومترية أن انخفاض الثبات يؤدي إلى زيادة أخطاء القياس وتقليل دقة النتائج (Nunnally & Bernstein, 1994).

ثالثاً: العلاقة بين الصدق والثبات

يرتبط الصدق والثبات بعلاقة تكاملية.

فالاختبار قد يكون:

- ثابتاً وغير صادق .
- لكنه لا يمكن أن يكون صادقاً إذا كان غير ثابت .

ويعبر عن هذه العلاقة بالشكل الآتي:

الثبات شرط ضروري للصدق ولكنه غير كافٍ وحده لتحقيقه .(Cohen & Swerdlik, 2022)

رابعاً: مصادر أخطاء القياس المؤثرة في الثبات

توجد مجموعة من العوامل التي تؤدي إلى انخفاض الثبات، منها:

1- عوامل مرتبطة بالمفحوص

- التعب والإرهاق .
- القلق والتوتر .
- الحالة الصحية .
- ضعف الدافعية .

2-عوامل مرتبطة بالاختبار

- غموض التعليمات .
- ضعف صياغة البنود .
- قلة عدد الفقرات .

3-عوامل مرتبطة بالبيئة

- الضوضاء .
- سوء الإضاءة .
- الظروف المناخية غير المناسبة .

4-عوامل مرتبطة بالفاحص

- اختلاف إجراءات التطبيق .
- أخطاء التصحيح والتقدير .

(Crocker & Algina, 2008)

خامساً: طرق تقدير الثبات

1- طريقة إعادة الاختبار(Test-Retest)

تعتمد هذه الطريقة على تطبيق الاختبار مرتين على المجموعة نفسها خلال فترتين زمنيتين ثم حساب معامل الارتباط بين التطبيقين.

المزايا

- تقيس الاستقرار الزمني .

العيوب

- تأثير التذكر .
- تغير خصائص المفحوصين بمرور الزمن .

(Anastasi & Urbina, 2019)

2- طريقة الصور المتكافئة(Parallel Forms)

يتم إعداد صورتين متكافئتين للاختبار ثم تطبيقهما على المجموعة نفسها.

المزايا

- تقلل من أثر التذكر .

العيوب

- صعوبة إعداد صورتين متكافئتين تماماً .

(Nitko & Brookhart, 2018)

3- طريقة إعادة الاختبار مع الصور المتكافئة

تجمع بين الطريقتين السابقتين.

وتُعد من أكثر الطرق دقة في تقدير الثبات، لكنها تتطلب جهداً ووقتاً كبيرين.

4- طريقة التجزئة النصفية(Split-Half Reliability)

يُقسم الاختبار إلى نصفين متكافئين ثم يحسب معامل الارتباط بين درجات النصفين.

ومن أشهر طرق التقسيم:

- الفردي والزوجي .
- النصف الأول والنصف الثاني .

وبعد ذلك يتم تصحيح المعامل باستخدام معادلة سبيرمان براون.

-5 طريقة الاتساق الداخلي

تعتمد على مدى تجانس فقرات الاختبار في قياس الخاصية نفسها.

ومن أشهر معاملاتها:

أ- معامل كودر ريتشاردسون(KR-20)

يستخدم مع البنود الثنائية (صحيح/خطأ).

ب- معامل ألفا كرونباخ

يُعد أكثر معاملات الثبات استخداماً في البحوث النفسية والتربوية الحديثة.(Cronbach, 1951)

سادساً: معامل ألفا كرونباخ

يقيس درجة الاتساق الداخلي بين فقرات الاختبار.

سابعاً: تفسير معاملات الثبات

يختلف تفسير الثبات باختلاف طبيعة الاختبار والغرض منه.

مستوى الثبات	معامل الثبات
ضعيف	أقل من 0.60

مستوى الثبات	معامل الثبات
مقبول	0.60 – 0.69
جيد	0.70 – 0.79
جيد جداً	0.80 – 0.89
ممتاز	0.90فأكثر

(Nunnally & Bernstein, 1994)

ثامناً: العوامل المؤثرة في الثبات

-1عدد الفقرات

كلما زاد عدد الفقرات ارتفع الثبات غالباً.

-2تجانس الفقرات

يزداد الثبات عندما تقيس الفقرات البناء نفسه.

-3تباين الأفراد

كلما زادت الفروق الفردية بين المفحوصين ارتفع معامل الثبات.

-4جودة صياغة البنود

الفقرات الواضحة ترفع من استقرار القياس.

(Cohen & Swerdlik, 2022)

تاسعاً: الخطأ المعياري للقياس

يستخدم الخطأ المعياري للقياس لتقدير مقدار الخطأ المتوقع في الدرجة المشاهدة.

حيث:

- SEM = الخطأ المعياري للقياس .
- S = الانحراف المعياري .
- Rtt = معامل الثبات .

ويُعد مؤشراً مهماً في تفسير نتائج الاختبارات النفسية والتربوية.(Crocker & Algina, 2008)

عاشراً: الثبات في ضوء النظريات السيكومترية الحديثة.

أصبحت تقديرات الثبات في القياس الحديث أكثر تعقيداً مع ظهور:

- نظرية الاستجابة للمفردة .(IRT)
- النمذجة متعددة الأبعاد .
- الاختبارات المحوسبة التكيفية .

إذ لم يعد الثبات يُنظر إليه كقيمة واحدة ثابتة للاختبار، بل قد يختلف تبعاً لمستوى القدرة المقاسة لدى الأفراد.(Embretson & Reise, 2020)

خلاصة الفصل

يمثل الثبات أحد الركائز الأساسية في القياس النفسي والتربوي، ويشير إلى مدى استقرار واتساق نتائج الاختبار. وتتنوع طرق تقدير الثبات بين إعادة الاختبار والصور المتكافئة والتجزئة النصفية والاتساق الداخلي، ويُعد معامل ألفا كرونباخ من أكثر المؤشرات استخداماً في الدراسات المعاصرة. كما أن ارتفاع الثبات يسهم في تحسين جودة القياس وزيادة الثقة في القرارات المبنية على نتائجه.

أسئلة للمراجعة

1. ما المقصود بثبات الاختبار؟

2. ما الفرق بين الصدق والثبات؟
3. ما أهم مصادر الخطأ المؤثرة في الثبات؟
4. اشرح طريقة إعادة الاختبار .
5. ما المقصود بالتجزئة النصفية؟
6. ما استخدام معامل ألفا كرونباخ؟
7. كيف يفسر الخطأ المعياري للقياس؟

أنشطة تطبيقية

النشاط الأول

قارن بين طرق تقدير الثبات المختلفة من حيث المزايا والعيوب.

النشاط الثاني

اجمع نتائج استبيان مكون من 20 فقرة وحاول حساب معامل ألفا كرونباخ باستخدام برنامج.SPSS

النشاط الثالث

حلل دراسة منشورة وحدد نوع معامل الثبات المستخدم فيها.

النشاط الرابع

ناقش أثر انخفاض الثبات على نتائج البحوث النفسية والتربوية.

الفصل 11: تحليل فقرات الاختبارات النفسية والتربوية

أهداف التعلم

في نهاية هذا الفصل يكون الطالب قادراً على:

- التعرف على مفهوم تحليل الفقرات وأهميته في بناء الاختبارات .

- حساب وتفسير معامل الصعوبة ومعامل السهولة .
- حساب وتفسير معامل التمييز .
- تقييم فعالية البدائل في أسئلة الاختيار من متعدد .
- توظيف نتائج تحليل الفقرات في تحسين جودة الاختبارات .
- التعرف على تحليل البنود في ضوء النظرية الكلاسيكية ونظرية الاستجابة للمفردة .

تمهيد

يمثل تحليل الفقرات إحدى المراحل الأساسية في بناء الاختبارات النفسية والتربوية وتقنينها، إذ يهدف إلى دراسة الخصائص الإحصائية لكل فقرة من فقرات الاختبار للتأكد من ملاءمتها للغرض الذي أعدت من أجله. ويُسهم تحليل الفقرات في الكشف عن البنود الضعيفة أو غير المناسبة واستبعادها أو تعديلها بما يضمن تحسين جودة الاختبار ورفع مستوى صدقه وثباته.(Nitko & Brookhart, 2018)

وتؤكد الأدبيات السيكومترية أن الاختبار الجيد لا يُقاس بعدد فقراته فقط، وإنما بمدى جودة تلك الفقرات وقدرتها على قياس السمة المستهدفة بكفاءة.(Crocker & Algina, 2008)

أولاً: مفهوم تحليل الفقرات

يقصد بتحليل الفقرات مجموعة الإجراءات الإحصائية المستخدمة لدراسة خصائص بنود الاختبار بعد تطبيقه على عينة تجريبية من المفحوصين.(Anastasi & Urbina, 2019)

ويهدف تحليل الفقرات إلى:

- تحسين جودة الاختبار .
- زيادة الصدق والثبات .
- حذف البنود الضعيفة .
- تحسين قدرة الاختبار على التمييز بين الأفراد .

- تحديد مستوى صعوبة الفقرات .

ثانياً: أهمية تحليل الفقرات

تتجلى أهمية تحليل الفقرات في أنه يساعد على:

- اختيار أفضل البنود .
- الكشف عن الأسئلة الغامضة .
- رفع الكفاءة السيكومترية للاختبار .
- تحسين عملية التقويم .
- بناء بنوك الأسئلة .

ويرى Cohen و Swerdlik (2022)أن تحليل البنود يمثل حلقة الوصل بين بناء الاختبار وتقنينه.

ثالثاً: معامل الصعوبة(Difficulty Index)

مفهومه

يشير معامل الصعوبة إلى نسبة الأفراد الذين أجابوا عن الفقرة إجابة صحيحة.

ويعبر عن مدى سهولة أو صعوبة الفقرة بالنسبة للمفحوصين.(Ebel & Frisbie, 1991)

حيث:

- P = معامل الصعوبة .
- R = عدد الإجابات الصحيحة .
- N = عدد المفحوصين .

مثال تطبيقي

إذا أجاب 80 طالباً من أصل 100 طالب عن فقرة معينة إجابة صحيحة فإن:

P = 80 / 100 = 0.80

أي أن الفقرة سهلة نسبياً.

تفسير معامل الصعوبة

التفسير	قيمة المعامل
صعبة جداً	أقل من 0.20
صعبة	0.20 – 0.39
متوسطة	0.40 – 0.60
سهلة	0.61 – 0.80
سهلة جداً	أكثر من 0.80

(Ebel & Frisbie, 1991)

رابعاً: معامل السهولة

يعبر معامل السهولة عن نسبة الإجابات الصحيحة على الفقرة.

أو يمكن التعبير عن الصعوبة بالعلاقة:

$$P = 1 - Q$$

حيث:

- Q = معامل الصعوبة .
- P = معامل السهولة .

خامساً: معامل التمييز(Discrimination Index)

مفهومه

يقيس قدرة الفقرة على التمييز بين ذوي الأداء المرتفع وذوي الأداء المنخفض في الاختبار (Crocker & Algina, 2008).

وتُعد الفقرة الجيدة هي التي يجيب عنها أصحاب الدرجات المرتفعة أكثر من أصحاب الدرجات المنخفضة.

طريقة حساب معامل التمييز

يتم تقسيم المفحوصين إلى:

- المجموعة العليا (27 .(%
- المجموعة الدنيا (27 .(%

حيث:

- D = معامل التمييز .
- U = عدد الإجابات الصحيحة في المجموعة العليا .
- L = عدد الإجابات الصحيحة في المجموعة الدنيا .
- N = عدد الأفراد في إحدى المجموعتين .

تفسير معامل التمييز

التفسير	قيمةD
ممتاز	0.40فأكثر
جيد	0.30 – 0.39
مقبول	0.20 – 0.29

التفسير	قيمةD
ضعيف	أقل من 0.20
فقرة معيبة	قيمة سالبة

(Nitko & Brookhart, 2018)

سادساً: فعالية البدائل في أسئلة الاختيار من متعدد

تُعد البدائل الخاطئة جزءاً مهماً من جودة السؤال الموضوعي.

ويجب أن تكون:

- منطقية .
- جذابة للمفحوص غير المتمكن .
- متقاربة في الطول والصياغة .

ويُعد البديل غير الفعال هو البديل الذي لا يختاره أي مفحوص تقريباً.

(Haladyna, 2016)

سابعاً: الارتباط بين الفقرة والدرجة الكلية

من المؤشرات المهمة في تحليل الفقرات حساب معامل ارتباط كل فقرة بالدرجة الكلية للاختبار.

ويشير الارتباط المرتفع إلى أن الفقرة تقيس البناء نفسه الذي يقيسه الاختبار.

ويُفضل أن يكون معامل الارتباط:

$r \geq 0.30$

(DeVellis & Thorpe, 2021)

ثامناً: اتخاذ القرار بشأن الفقرة

بعد حساب المؤشرات السابقة يمكن اتخاذ أحد القرارات التالية:

-1الاحتفاظ بالفقرة

إذا كانت:

- متوسطة الصعوبة .
- مرتفعة التميز .
- ذات ارتباط جيد بالدرجة الكلية .

-2تعديل الفقرة

إذا ظهرت بها بعض المشكلات القابلة للتصحيح.

-3حذف الفقرة

إذا كانت:

- ضعيفة التميز .
- غامضة .
- غير مرتبطة بالبناء المقاس .

تاسعاً: تحليل الفقرات في ضوء النظرية الكلاسيكية للقياس

تركز النظرية الكلاسيكية على:

- معامل الصعوبة .
- معامل التميز .
- معامل الثبات .

- الارتباط بالدرجة الكلية .

وقد شكلت هذه المؤشرات الأساس التقليدي لتحليل البنود لعقود طويلة.(Lord & Novick, 1968)

عاشراً: تحليل الفقرات في ضوء نظرية الاستجابة للمفردة(IRT)

ظهرت نظرية الاستجابة للمفردة بهدف تجاوز بعض قيود النظرية الكلاسيكية (Embretson & Reise, 2020).

وتعتمد على تقدير معلمات البند ومنها:

-1معلمة الصعوبة(b)

تمثل مستوى القدرة اللازم للإجابة الصحيحة.

-2معلمة التمييز(a)

تعكس قدرة البند على التمييز بين المفحوصين.

-3معلمة التخمين(c)

تمثل احتمال الإجابة الصحيحة الناتجة عن التخمين.

وقد أصبحت هذه النظرية أساساً للاختبارات المحوسبة التكيفية الحديثة.

حادي عشر: استخدام البرامج الإحصائية في تحليل الفقرات

يمكن إجراء تحليل البنود باستخدام:

- SPSS.
- R.
- JASP.
- Winsteps.

- Bilog-MG.

وتساعد هذه البرامج في:

- حساب معاملات الصعوبة والتمييز .
- تقدير الثبات .
- إجراء التحليل العاملي .
- تطبيق نماذج نظرية الاستجابة للمفردة .

(Field, 2024)

خلاصة الفصل

يُعد تحليل الفقرات خطوة أساسية في بناء الاختبارات النفسية والتربوية، إذ يسمح بتحديد جودة البنود من خلال مؤشرات الصعوبة والتمييز والارتباط بالدرجة الكلية. كما أسهمت النظريات السيكومترية الحديثة، وخاصة نظرية الاستجابة للمفردة، في تطوير أساليب تحليل البنود وتحسين دقة القياس النفسي والتربوي.

أسئلة للمراجعة

1. ما المقصود بتحليل الفقرات؟
2. ما أهمية معامل الصعوبة؟
3. كيف يتم حساب معامل التمييز؟
4. ما خصائص البدائل الفعالة؟
5. متى يتم حذف الفقرة من الاختبار؟
6. ما الفرق بين النظرية الكلاسيكية ونظرية الاستجابة للمفردة في تحليل البنود؟
7. ما البرامج المستخدمة في تحليل الفقرات؟

أنشطة تطبيقية

النشاط الأول

احسب معامل الصعوبة لفقرة أجاب عنها 75 طالباً من أصل 100 إجابة صحيحة.

النشاط الثاني

احسب معامل التمييز إذا كانت المجموعة العليا تضم 25 إجابة صحيحة والمجموعة الدنيا 10 إجابات صحيحة من أصل 27 طالباً.

النشاط الثالث

حلل مجموعة من أسئلة الاختيار من متعدد وحدد البدائل غير الفعالة.

النشاط الرابع

استخدم برنامج SPSS لحساب معامل ارتباط الفقرة بالدرجة الكلية لعينة افتراضية.

معامل صعوبة ، و معامل تمييز سؤال المقال :

يمكن اجراء تحليل المفردات الذي أوضحناه في حالة أسئلة المقال . فمن المفيد أن يفحص المعلم كل سؤال في ضوء إجابات الطلاب لتحديد صدق السؤال فيما يقيسه ، ومدى ملائمته للهدف التعليمي المرجو . فإذا وجد أن الطلاب الضعاف في التحصيل كان أداؤهم أفضل من أداء الأقوياء في التحصيل في سؤال معين ، فإن السؤال بذلك لا يسهم في التمييز بينهما . إذ ربما يكون السؤال غامضا ، أو كان يتطلب تفاصيل أكثر في الإجابة لم يدركها الطالب من السؤال . و إذا أجاب جميع الطلاب تقريبا على السؤال ، أو كانت إجابتهم جميعا غير مرضية ، فإن درجة صعوبة السؤال تؤدي إلى احتمال انخفاض تمييزه .

ومن الممكن استخدام بيانات تحليل المفرداتى في حالة أسئلة المقال بطريقة مماثلة لتحليل المفردات الموضوعية . فمعامل الصعوبة في هذه الحالة سوف يكون متوسط عدد النقاط التي يحصل عليها الطالب في سؤال المقال مقسوما على عدد النقاط الممكنة ، أي أن :

$$\text{معامل صعوبة سؤال المقال} = \frac{\text{ع}_{\text{ق}} + \text{ق}_{\text{د}}}{\text{س العضمى } (\text{ن}_{\text{ع}} + \text{ن}_{\text{د}})} \quad .....$$

حيث ترمز (ع$_{ق}$) إلى مجموع عدد النقاط التي حصل عليها طلاب المجموعة العليا .

-(ق$_{د}$) إلى مجموع عدد النقاط التي حصل عليها طلاب المجموعة الدنيا .

-(س$_{العظمى}$) إلى أكبر عدد من النقاط التي يمكن أن يحصل عليها طالب في السؤال .

-(س$_{العظمى}$) إلى أكبر عدد من النقاط التي يمكن أن يحصل عليها طالب في السؤال .

-(ن$_{ع}$) إلى عدد طلاب المجموعة العليا الذين أجابوا على السؤال إجابة صحيحة .

-(ن$_{د}$) إلى عدد طلاب المجموعة الدنيا الذين أجابوا على السؤال إجابة صحيحة . و بالمثل يمكن إيجاد معامل تمييز سؤال المقال باستخدام الصيغة التالية :

$$\text{معامل تمييز سؤال المقال} = \frac{\text{ق}_{ع} + \text{ق}_{د}}{\text{س}_{العضمى}\ (\text{ن})}$$

حيث ترمز (ن) إلى عدد طلاب أي من المجموعتين العليا أو الدنيا ، نظرا لتساوي عدد كل منهما .

و ينبغي مراعاة أن إيجاد عدد هاتين المجموعتين يتطلب كما سبق ترتيب أوراق الإيجابيات من الأعلى إلى الأدنى وفقا للدرجات الكلية في الاختبار . و تقسيمها بحيث يشتمل كل منهما على (27%) من عدد هذه الدرجات . ثم يتم حساب مجموع النقاط لكل سؤال التي حصل عليها طلاب كل من المجموعتين على حدة ، و تحديد أعلى النقاط التي حصل عليها طالب في كل سؤال في كل من المجموعتين معا (س$_{العظمى}$) ، ثم تطبيق الصيغتين (7-9) ، (7-10) لإيجاد معامل صعوبة ، و معامل تمييز كل سؤال مقال .

تحليل مشتتات مفردات الاختيار من متعدد :

ينبغي ألا يكتفي المعلم بإيجاد قيمة كل من معامل صعوبة و تمييز مفردات و أسئلة الاختبار الصفي ، و إنما يوظف هذه النتائج في تحليل مشتتات المفردات الموضوعية Distractors ، و إجابات أسئلة المقال . فهذا التحليل يهدف لفحص هذه المشتتات ، و بخاصة إذا وجد أن قيمة معامل تمييز المفردة سالبة ، أو صفر أو

قريبا من الصفر . كما يمكن إجراء ذلك إذا لاحظ أن المفردة تحتاج إلى مزيد من التدقيق . فإذا تبين له مثلا أن أحد المشتتات في مفردة معينة لم يختره أحد من الطلاب ، فإن هذا ربما يكون دليلا على عدم فاعلية هذا المشتت أو وضوح خطأه . فمثل هذا المشتت لا فائدة منه لأنه يضيع وقت الطالب في قراءة مثل هذه المشتتات .

وإذا وجد أن واحدا من المشتتات اختاره معظم الطلاب و بخاصة الذين حصلوا على درجات مرتفعة في الاختبار ككل ، ولم يختاروا الإجابة الصحيحة ، فإن هذا ربما يكون دليلا على أن هذا المشتت أكثر جاذبية من الإجابة الصحيحة ، أو ربما يكون من الإجابة الصحيحة بالفعل ، و ليست الإجابة التي حددها المعلم . و هنا ينبغي أن يقوم المعلم بمراجعة الإجابة الصحيحة و التأكد من صحتها و عدم صحة هذا المشتت . فإذا تأكد أن هذا المشتت يمثل إجابة خطأ ، فإنه ينبغي الإبقاء عليه لأنه يساعد في كشف سوء الفهم ، و أنماط الأخطاء الشائعة لدى الطلاب ، و يوجه المعلم نحو أساليب العلاج المناسبة .

أما إذا كانت إجابات الطلاب موزعة توزيعا متعادلا بين المشتتات الأربعة أو الخمسة و الخمسة لمفردة الاختيار من متعدد ، ولكن المفردة غير مميزة ، فإن هذا ربما يدل على أن هذه المفردة غامضة ، أو غاية في الصعوبة ، مما جعل الطلاب يلجؤون إلى التخمين ، و هكذا . و بذلك يستفيد المعلم من تحليل المفردات والأسئلة الاختبارية في تحسينها و استخدامها في مرات تالية ، و كذلك في تشخيص الأخطاء و سوء الفهم لدى الطلاب في المادة الدراسية ، و تعديل العملية التدريسية لعلاج هذه الأخطاء .)**صلاح الدين محمود علام. ،2019).**

معامل التمييز Discrimination Index :

التمييز هو مؤشر يدل على قدرة السؤال الاختياري على الوقوف على الفروق الفردية بين مستوى الكفايات .

يتطلب تمييز الاختبار تنويع مستويات الأسئلة من حيث السهولة و الصعوبة ، و العمل على زيادة عدد الأسئلة ، و تغطية محتوى المادة الدراسية . بالإضافة إلى ذلك يجب التقليل من حدوث الغش ، و التقليل من عامل التخمين لمنع حصول الطلاب على درجات لا يستحقونها (فخري خضر 2003) .

معامل التمييز = عدد الإجابات الصحيحة في المجموعة العليا – عدد الإجابات الصحيحة في المجموعة الدنيا

عدد الطلاب في إحدى المجموعتين

كيفية تحديد المجموعة العليا و المجموعة الدنيا :

نحدد المجموعة العليا و الدنيا عن طريق :

أ-ترتيب أوراق الطلاب حسب الدرجات ترتيبا تنازليا ، ثم تقسيمها إلى مجموعتين متساويتين هي مجموعة الأوراق ذات الدرجات العليا ، و مجموعة الأوراق ذات الدرجات الدنيا .

ب-نستخرج إحصائيات السؤال .

مثال :

مجموعة عددها (60) طالبا و المطلوب حساب معامل التمييز لها في اختبار من الاختبارات ؟

1- نرتب أوراق (60) طالبا حسب الدرجات ترتيبا تنازليا .

2- نأخذ (30) ورقة ذات الدرجات العليا (مجموعة 1) .

3- نأخذ (30) ورقة ذات الدرجات الدنيا (مجموعة 2) .

4- فإذا افترضنا أن عدد الإجابات الصحيحة في المجموعة العليا (20) ، و عدد الإجابات الصحيحة في المجموعة الدنيا كان (10) .

نوضح الحل في الجدول التالي :

معامل التمييز – لمجموعة من (60) طالبا

د	ج	ب	أ	
1	5	20	4	المجموعة العليا 30
3	9	10	8	المجموعة الدنيا 30
				ب هي الإجابات الصحيحة

$$\text{معامل التمييز} = \frac{\text{عدد الإجابات الصحيحة في المجموعة العليا} - \text{عدد الإجابات الصحيحة في المجموعة الدنيا}}{\text{عدد الطلاب في إحدى المجموعتين}}$$

$$\text{معامل التمييز} = \frac{20-10}{30} = \frac{10}{30} = \mathbf{0.3}$$

الجدول التالي يوضح قيم معامل التمييز المقبولة (مندور فتح الله 2000)

قيم معامل التمييز المقبولة

مدلولها و ما ينبغي عملها	**قيم معامل التمييز**
ليس لها قيمة ، و لا تساهم بالإيفاء بغرض الاختبار و لذلك ينبغي أن تفحص بدقة ، فإما أن تعدل أو تحذف من الاستخدام	قيمة سالبة لمعامل تمييز المفردة
ذات تمييز جيد و يمكن الاحتفاظ بها	قيمة موجبة لمعامل التمييز ما بين 0.39-0.20
ذات تمييز جيد و يمكن الاحتفاظ بها	قيمة موجبة تزيد عن 39
مفردات ذات تمييز قوي و يجب أن يستخدمها المعلم	القيمة الموجبة أكبر من 0.6

لأسئلة السهلة أو الصعبة تكون عادة ذات معامل تمييز منخفض . أسئلة كهذه ينبغي ألا تستخدم في الاختبار إلا إذا كان هناك ضرورة لتنشيط دافعية الطلاب .(أحمد يعقوب النور ،2007 ص208،،206).

مؤشر الصعوبة المفردة (أو مستوى صعوبة المفردة) :

ITEM DIFFICULTY INDEX (OR ITEM DIFFICULTY LEVEL)

عند تقييم المفردات في اختبارات الأداء الأقصى أو القدرات ،

ينبغي مراعاة مستوى صعوبة المفردات . و تعرف صعوبة المفردة بأنها النسبة المئوية أو النسبة من المختبرين الذين يجيبون إجابة صحيحة عن المفردة . و يرمز لمستوى أو مؤشر صعوبة المفردة بالرمز P ، و يتم حسابه باستخدام الصيغة الآتية :

تعرف صعوبة المفردة بأنها النسبة المئوية أو النسبة من المختبرين الذين يجيبون إجابة صحيحة عن المفردة

$$P = \frac{\text{عدد المختبرين الذين أجابوا على المفردة إجابة صحيحة}}{\text{عدد المختبرين}}$$

فمثلا ، إذا كان عدد طلبة صف معين 30 طالبا ، و إذا أجاب 20 طالبا إجابة صحيحة ، 10 أجابوا إجابة خطأ، فإن مؤشر صعوبة المفردة يساوي 0.66 ، و طريقة حساب ذلك كالآتي :

$$P = \frac{20}{30} = 0.66$$

وفي نفس الصف، إذا أجاب 10 طلبة إجابة صحيحة، 20 أجابوا إجابة خطأ ، فإن مؤشر صعوبة المفردة يساوي 0.33 . و يمكن أن يتراوح مؤشر صعوبة المفردة بين صفر ، 1.0 ، حيث تكون قيمة المؤشر بالنسبة إلى المفردات الأكثر سهولة مرتفعة ، و قيمته بالنسبة إلى المفردات الصعبة منخفضة . و تكون قيمة المؤشر بالنسبة إلى المفردة التي أجاب عليها جميع الطلبة إجابة صحيحة 1.00 ، بينما تكون قيمته بالنسبة إلى المفردة التي أجاب عنها جميع الطلبة إجابة خطأ صفرا . و المفردات التي قيمة مؤشر صعوبتها إما 1.00 أو صفر لا تقدم أية معلومات عن الفروق الفردية ، و لا تكون لها فائدة من منظور القياس . و سوف يضمن بعض القائمين ببناء الاختبارات مفردة أو مفردتين قيمة P لهما 1.0 في مستهل اختبار معين لغرس حس الثقة لدى المختبرين .

و هذه تعد من الممارسات التي يمكن تبريرها من منظور الدافعية ،

و لكن هذه المفردات – من منظور فني – لا تساهم في الخصائص السيكومترية للاختبار.

و أحد العوامل الأخرى التي ينبغي مراعاتها فيما يتعلق بتضمين مفردات غاية في السهولة أو الصعوبة هو فعالية الوقت. فالوقت الذي يقضيه المختبرون في إجابة مفردات غير فاعلة يعد وقتا مهدرا ، و كان من الأفضل الاستفادة منه في مفردات تثري الخصائص السيكومترية للاختبار.

لتعظيم التباين بين المختبرين ، يكون المستوى الأمثل لصعوبة المفردة 50 % ، و هي تدل عن أن 50 %من المختبرين أجاب عن المفردة إجابة صحيحة ، 50 %أجاب عنها إجابة خطأ .

و لتعظيم التباين و الثبات ، يكون المستوى الأمثل لصعوبة المفردة 0.50 ، و هي تدل على أن 50% من المختبرين أجاب عن المفردة إجابة صحيحة ، 50% أجاب عنها إجابة خطأ . و ربما تستنتج – استنادا إلى هذه العبارة – أنه من المرغوب فيه أن يكون مستوى صعوبة جميع مفردات الاختبار 0.50 ، ولكن ليس من الضروري أن يكون هذا صحيحا لأسباب متعددة . و أحد الأسباب أن مفردات اختبار معين ترتبط ببعضها بعضا في أغلب الأحيان . فعندما ترتبط مفردات اختبار ببعضها بعضا ، فإن عملية القياس ربما تتداخل إذا كانت قيم P لجميع المفردات 0.50 . و نتيجة لذلك ، فإنه من المرغوب فيه في أغلب الأحيان اختيار بعض المفردات التي تكون قيم P لها أقل من 0.50 ، و مفردات أخرى تكون قيم P لها أكبر من 0.50 ، و لكن يكون متوسطها جميعا 0.50 . و قد أوصى (Aiken 2000 (بأنه ينبغي أن يكون هناك مدى 0.20 تقريبا لهذه القيم p حول القيمة الأمثل . فمثلا،

ربما يختار من يقوم ببناء اختبار مفردات تتراوح مستويات صعوبتها بين 0.40 ، 0.60 بمتوسط قيمته 0.50

و أحد الأسباب الأخرى المتعلقة بأن 0.50 ليست مستوى الصعوبة الأمثل لجميع المواقف الاختبارية ، يتضمن تأثير التخمين . ففي المفردات التي تتطلب إنشاء الاستجابة)مثل ، أسئلة المقال ، و الإجابة القصيرة) ، حيث لا يكون للتخمين أثر يذكر ، تكون 0.50 هي القيمة التي تعد عادة مستوى الصعوبة الأمثل . غير أنه في حالة المفردات التي تتطلب اختيار الاستجابة (مثل ، مفردات الاختيار من

بينما متوسط P الذي قيمته 0.50 يعد مثاليا لتعظيم التباين و الثبات ، إلا أنه من المرغوب فيه وجود مستويات مختلفة في كثير من التطبيقات الاختيارية المختلفة .

متعدد ، و الصواب أو الخطأ) حيث ربما يجيب المختبرون

إجابة صحيحة على المفردة عن طريق التخمين وحده ، يتباين مستوى الصعوبة . و لكي نأخذ تأثيرات التخمين بعين الاعتبار ، فإن مستوى الصعوبة الأمثل للمفردة يحدد أعلى من المستوى الأمثل للمفردات التي تتطلب إنشاء الاستجابة . فمثلا ، بالنسبة إلى المفردات الاختيار من متعدد ذات أربعة خيارات ، ينبغي أن يكون متوسط p قيمته 0.74 تقريبا . (Lord, 1952) أي أن القائم ببناء الاختبار ربما يختار مفردات تتراوح مستويات صعوبتها بين 0.64 ، 0.84 بمتوسط قيمته 0.74 تقريبا . ويوضح جدول 7-1 معلومات عن المتوسط الأمثل لقيمة p للمفردات التي تتطلب اختيار الاستجابة و يتباين عدد بدائلها أو خياراتها.

جدول 7-1 قيم p الأمثل لمفردات يتباين عدد خياراتها :

عدد الخيارات	القيمة الامثل لمتوسط p
(2مثل ، الصواب أو الخطأ)	0.85
3	0.77
4	0.74
5	0.69
إنشاء الاستجابة (مثل ، المقال)	0.50

مواقف تقييم خاصة و صعوبة المفردة

Special Assessment Situations and Item Difficulty:

تنطبق مناقشتنا لصعوبة المفردة حتى الآن على الاختبارات مرجعية المعيار بدرجة أكبر . و بالنسبة للاختبارات مرجعية المحك – و بخاصة اختبارات الإتقان – يتم تقييم صعوبة المفردة بطريقة مختلفة.

فعادة في اختبارات الاتقان إما أن ينجح الفرد المختبر أو يرسب ، و يوجد توقع أن معظم المختبرين سوف يكونون في النهاية ناجحين . و نتيجة لذلك ، فإنه من الشائع في اختبارات الاتقان أن يصل متوسط قيمة P للمفردات إلى 0.90 . و ربما تتباين اختبارات أخرى صممت لأغراض تقييم خاصة ، من حيث ما يمثل مستويات صعوبة مرجوة للمفردات . فمثلا ، إذا تم بناء اختبار لمعاونة أصحاب الاعمال في اختيار 25 %

العليا من المتقدمين لعمل معين ، فإنه من المرغوب فيه أن تكون هناك مفردات متوسط قيم P لها حوالي 0.25 ، ومن المرغوب فيه أن يميز اختبار معين بين المختبرين ذوي الأداء الأعلى (مثلا ، عند اختبار الطلاب الموهوبين و الفائقين) ، و ربما يكون من المرغوب فيه أيضا تضمين بعض المفردات الصعبة للغاية على الأقل . و الخلاصة ، أنه بينما تعد قيمة P المتوسطة 0.50 مثالية لتعظيم التباين بين المختبرين ، إلا أن مستويات صعوبة مختلفة تكون مرغوبة في كثير من التطبيقات الاختبارية (أنظر "موضوع ذو أهمية خاصة 1.7 " ، حيث يقدم مثالا آخر.

و ينبغي ذكر أن قيمP تعتمد على العينة . أي أنه من الممكن الحصول على قيم مختلفة لمؤشر الصعوبة P للمفردة من عينات مختلفة . فمثلا ، يتوقع أن سوف تختلف قيم صعوبة المفردة وفقا لمستوى الفرقة . فسوف يحتمل أن تكون مفردة معينة أكثر صعوبة بالنسبة إلى اختبار طبق على تلاميذ الفرقة الثالثة مقارنة بتلاميذ الفرقة الرابعة(. صلاح الدين محمود علام ، 2013 ، ص 303،301)

اسئلة تقييمية خاصة بمقياس القياس النفسي وبناء الاختبارات المدرسية

أولا: من خلال دراستك لمقياس القياس النفسي تعرضت للمراحل التاريخية لتطوره(10 نقاط)

المطلوب:

1. اذكر هذه المراحل دون شرح.

-.....

-.....

-.....

2. Wundt / Benet and Simon /Galton / Person/ gosse كلهم علماء ورواد للمراحل التي ستذكرها في إجابتك عن السؤال الأول.

1.2. ضع كل رائد في المرحلة المناسبة له

.....

.....

.....

2.2.اذكر كل مرحلة في أي دولة كانت

-.....

-.....

-.....

3.2.ما الذي قدمته كل مرحلة لعلم النفس ؟

-.....

-.....

-.....

ثانيا: إليك المتغيرات التي قد تصادفك وأنت تدرس إحدى موضوعات في تخصصك (6 نقاط)

الجنس (ذكر/ أنثى) – الاكتئاب- التفاؤل - متعلم /غير متعلم – ناجح /راسب- المراحل الدراسية – المراحل العمرية - التحصيل الدراسي – الرضا الوظيفي- مستوى الطموح – الجنسية – الديانة –لون العينين-رتب الموظفين – رتب المتعلمين – الذكاء الانفعالي – الخصائص الجسمية

المستوى.....	المستوى.....	المستوى.....	المستوى.....
-.....	-.....	-.....	-.....
-.....	-.....	-.....	-.....
-.....	-.....	-.....	-.....
-.....	-.....	-.....	-.....
-.....	-.....	-.....	-.....
-.....	-.....	-.....	-.....

المطلوب:

1. أكمل الجدول .

2. صنف كل متغير حسب المستوى المناسب له.

ثالثا:

أثناء دراستك لمحاضرة مستويات القياس، تعلمت أن العدد لا معنى له إلا من خلال المعنى الذي تعطيه له أنت كباحث، ومن خلال الغرض منه

فعندما يكون الغرض من العدد هو الوصف والتصنيف فنحن في مستوى القياس.....

أما إذا كان الغرض هو كمي من خلال تحويل الخاصية من الكيف إلى الكم، ومعرفة الفرق بين فرد وأخر فنحن في مستوى قياس .....

أما إذا كان الغرض من العدد هو التصنيف والترتيب فنحن أمام مستوى القياس.....

رابعا: اثناء محاولتك بناء اختبار نفسي لتستخدمه في مذكرتك، عليك مراعاة مواصفات أداة القياس الجيدة .

س1: ماهي الخصائص التي لابد أن تتوفر في الاختبار الذي تريد بناءه حتى يوصف بأنه أداة قياس جيدة؟

س2: ماهي الطريقة التي تسمح لك بالتحقق من ثباته من دون بناء صورة مكافئة له؟

س3: لوكان اختبارك تحصيلي وأردت التحقق من ثباته، ما الطريقة التي ستستخدمها؟

خامسا :تمرين

أراد أستاذ لغة العربية لسنة الثانية متوسط، بناء إختبار تحصيلي موضوعي، في ثلاث وحدات ، وبعد أن قرر وضع 20 فقرة لهذا الإختبار، وان يقوم بتصحيحه على 40، قدم لك هذه المعطيات وطلب منك مساعدته على بناء إختبار تحصيلي موضوعي:

الوحدات	عدد الحصص	الأهداف					
		الأول	الثاني	الثالث	الرابع	الخامس	السادس
الوحدة الأولى	2	هدفين	ثلاثة أهداف	هدف واحد	/	/	/
الوحدة الثانية	3	اربع أهداف	خمس أهداف	هدف واحد	/	/	هدف واحد
الوحدة الثالثة	3	هدفين	هدفين	هدفين	/	هدف واحد	/

2-1- بماذاتنصحه؟

2-2- قم بتطبيق ما نصحته به من خلال المعطيات السالفة الذكر ؟

2-3- عرف مانصحته بتطبيقه؟

2-أكمل مايلي :

سادسا :

هذه مجموعة من الأسئلة يتكون كل سؤال من أربعة إجابات اختيارية ، توجد إجابة واحدة صحيحة لكل سؤال المطلوب وضع دائرة حول العبارة الصحيحة فعلا.

1- ادخل التجريب إلى علم النفس عن طريق :

أ-وليام فوندت . ب- فوانسيس غالتون . ج- وليام جيمس. د-سبيرمان براون.

2- من خصائص الظاهرة النفسية :

أ-متداخلة . ب-ديناميكية. ج- غير مدركة. د- كل ماسبق.

3- يمنح للعدد معنى في القياس النفسي عن طريق:

أ –قواعد رياضية . ب- معالجات إحصائية . ج- أسس علمية . د- مستويات القياس.

4- خطأ القياس الذي لا يؤثر على كمية الفروق بين الأفراد هو الذي يعود إلى :

أ- الحالة النفسية للمفحوص. ب- أداة القياس . ج- ظروف التطبيق . د- طريقة التصحيح.

5- أعلى مستوى قياس يمكن تصنيف عنده تقديرات الطلبة هو المستوى :

أ- النسبي . ب- المسافات المتساوية - ج- الرتبي . د- الاسمي.

6- المفهوم الذي ادخله "كاتل" لأول مرة في ميدان القياس النفسي هو :

أ-التغذية الراجعة. ب- الاختبار النفسي . ج- الاختبار العقلي. د- الدرجة الحقيقية.

7- طبق "فوندت" المنهج التجريبي لدراسة سلوك الفرد بهدف :

أ- قياس الاختلافات بين الأفراد في زمن الرجع .

ب- الوصول إلى تفسير علمي لإختلافات سلوك الأفراد.

ج- قياس كمي و موضوعي لاستجابات الأفراد.

د- الوصول إلى قوانين معممة للسلوك الإنساني.

ثامنا: اختر الاجابة الصحيحة فيما يأتي

.عندما يقال لولي أمر متعلم إن تحصيل ابنك في امتحان العلوم الطبيعية هو أفضل من مستوى أداء 50% من المتعلمين صفه، يكون تصنيف هذا الامتحان على أنه:

أ. مقنن ☐ ب.موضوعي ☐ ج. محكي المرجع ☐ د. معياري المرجع☐

2.يستطيع متعلم أن يحل 90بالمئة من مسائل القدرة العددية من ضمن اختبارات القبول السنوي بنجاح، يصنف هذا الاختبار على أنه:

أ. مرجعي الفرد ☐ ب. موضوعي ☐ ج. محكي المرجع ☐ د. معياري المرجع☐

3.أنت مختص تريد التشخيص والحكم على ما إذا كان طفل ذو الأربع سنوات لديه لغة جيدة هل تحتاج إلى معلومات:

أ.معيارية المرجع ☐ ب. محكية المرجع☐ ج. ذاتية المرجع☐ د. موضوعية المرجع☐

4.لتحديد نوع ودرجة الصعوبات التي تواجهها مجموعة من المتدربين أو المتعلمين نستخدم أدوات قياس من نوع:

أ.موضوعية ☐ ب. ذاتية المرجع ☐ ج. محكي المرجع ☐ د. المعياري المرجع ☐

5. حدد أين يكمن التفسير المعياري للنتائج مما يلي:

أ. تحصل متعلم على الرتبة الميئينية 95 من بين زملائه ☐ ب.وصول المتعلم إلى درجة الإتقان لجميع الكفاءات المحددة ☐

ج. معلم حقق أكبر عدد من الأهداف المحددة قبل بدء التدريس ☐ د. متعلم حقق ما نسبته 70%من المقرر فهو ناجح. ☐

6. يستخدم القياس المرجعي المحك عادة في:

أ. اتخاذ القرارات التربوية الخاصة بإصلاح المناهج ☐ ب. تحديد الطلبة الناجحين في مسابقة ما. ☐

ج. توجيه الطلبة الناجحين إلى تخصصات معينة ☐ د. تحديد معايير اختبارات معينة لتفسير نتائجها ☐

7. عندما نقول بوجود علاقة بين قدرات الأفراد الطبيعية وأدائهم هذا يعني أن:

أ. أداء الأفراد يمكن أن يتوزع اعتداليا ☐ ب. الاستعداد المنخفض يعد عائقا أمام التعليم ☐

ج. الاستعداد المنخفض لا يعد عائقا أمام التعلم ☐ د.أداء الأفراد يمكن التنبؤ به من خلال الاستعداد ☐

8. العالم الأمريكي جليزر هو أول من استخدم مصطلح:

أ. الاختبارات المرجعية ☐ ب.الاختبارات المحكية ☐ ج. الاختبارات الموضوعية ☐ د. الاختبارات النفسية ☐

9. تستخدم الاختبارات المعيارية المرجع في أغراض:

أ. التقويم والتقييم ☐ ب.التعيين والتشخيص ☐ ج. التقويم البنائي والاختيار ☐ د. التقييم والاختيار ☐

10. التحصيل تبعا لمراعاة إيقاع التعلم يتوزع توزعا

أ. ملتويا ☐ ب. اعتداليا ☐ ج. تصاعديا ☐ د. تنازليا ☐

11. حسب النظرية الايديومترية يمكن لمتعلم منخفض التحصيل أن يتقن مهارة ما إذا ما تم معاملته:

(أ). بنفس طريقة أفراد المجموعة التي ينتمي اليها ☐ **(ب)** بطريقة تتناسب وتمكنه وقدراته ☐

(ج)بطريقة مختلفة عن أفراد عينته ☐ (د) بطريقة تتناسب طردا وقدراته ☐

12. الخاصية الأساسية في تقدير جودة البند هي قدرته على:

أ. القياس ☐ ب. التقييم ☐ ج. الثبات ☐ د. التمييز ☐

13. الفرق الأساسي بين القياس المحكي المرجع والقياس المعياري المرجع هو:

(أ) طبيعة العلاقة بين الخاصية والقدرة ☐ (ب) طريقة تفسير النتائج في الاختبار ☐

(ج) عدد البنود والبدائل في الاختبار ☐ (د) طريقة التطبيق ومنح الدرجات ☐

14.يمكننا الاستفادة من الاختبارات المعيارية المرجع في:

(أ) اختيار جميع الأشخاص المتمكنين من مهارة معينة ☐ (ب) اختيار أفضل الأشخاص في وظيفة معينة ☐

(ج) الكشف عن قدرة الشخص في إنجاز مهمة معينة ☐ (د) الكشف عن ملائمة الشخص لمهنة معينة ☐

15.حدّد أين يكمن التفسير المعياري المرجع للنتائج التالية:

(أ) حقق متعلم ما نسبته 72% من المستوى المطلوب ☐ (ب) تفوق الطالب على 73% من الطلاب ☐

(ج) تقدم المتعلم بنسبة 74% في عملية المراجعة ☐ (د) أتقن طالب حوالي 75% من الأهداف ☐

16. حدّد العبارة الصحيحة الوحيدة مما يلي:

(أ) اعتمد القياس النفسي في بدايته على القياس المحكي ☐ (ب) يعد العالم Glaser من رواد القياس المعياري ☐

(ج) تأثر القياس المحكي بظهور حركة علم النفس الفارقي ☐ (د) يتم تحديد المعيار بعد تطبيق الاختبار المعياري ☐

17.حدّد العبارة الصحيحة الوحيدة مما يلي:

(أ) يؤكد الاتجاه المعياري على وجود علاقة طردية بين الخصائص والأداءات ☐ (ب) يؤكد أصحاب الاتجاه المحكي بأن أداء الأفراد يتوزع توزيعاً اعتدالياً ☐

(ج) يفسر الاتجاه السيكومتري أداء الفرد في ضوء محك تحديده مسبقاً ☐ (د) يفسر الاتجاه الايديومتري أداء الفرد في ضوء معيار يتم تحديده مسبقاً ☐

***إليك العنوان البحثي التالي:**

- دور برنامج إرشادي قصصي قائم على نظرية العلاج العقلاني الانفعالي في خفض السلوك العدواني لدى اطفال المدارس

حدد مستوى قياس متغير السلوك العدواني :.....

إذا أردت دراسة الفروق في مستوى السلوك العدواني لدى العينة قبل وبعد تطبيق البرنامجالارشادي القصصي . الأسلوب الإحصائي المناسب هو ..... والمنهج الملائم للدراسة هو .....

*** بصفتك باحث في علم النفس ومختص في الإرشاد والتوجيه، لجأ إليك أستاذ يعمل في مؤسسة تربوية لمعرفته أنك تعرضت -في مادة القياس النفسي-لمواضيع يمكنك مساعدته من خلالها، فطرح عليك التساؤلات التالية. ساعده من خلال إجابتك**

1. في الخطوة الأولى من خطوات بناء الاختبارات النفسية والتربوية سأحدد الهدف بدقة، ثم أعرف السمة المراد قياسها بالرجوع إلى ....، والهدف من ذلك استخراج .....
2. لأحدد الأهداف بصورة إجرائية يعني أحول .....إلى أهداف يمكن قياسها .....
3. أثناء بنائي لاختبار تحصيلي أعتمد في تحديد أوزان الفقرات على ....، وأنت كمختص نفسي على ماذا تعتمد في تحديد أوزان فقرات الاختبار النفسي؟ اعتمد على .....
4. أريد أن أبني اختبار ا لأعرف مدى تمكن المتعلمين من قدرات محددة، تنصحني باستخدام .....

تاسعا :

تمّ التطرق إلى تطور القياس النفسي وذكر تطوراته من خلال مجموعة من التيارات : قارن بين هذه التيارات في نقاط محددة؟ (من حيث : الرواد ، الإنجازات ، الأماكن)

قائمة المراجع المطلع عليها لاعداد محاضرات المطبوعة:

-إسماعيل مُحَمَّد الفقي .2005. ***التقويم والقياس النفسي والتربوي***، القاهرة: دار غريب للطباعة والنشر، ب ر ط .- أحمد عودة .1999. ***القياس والتقويم في العملية التدريسية***،دار الأمل ، الطبعة الثالثة.

-أمينة مُحَمَّد كاظم . 1988. ***دراسة نظرية نقدية حول القياس الموضوعي للسلوك " نموذج راش"*** .مؤسسة الكويت لتقدم العلمي، الكويت.

-إبراهيم مبارك الدوسري.2000 . ***الإطار المرجعي للتقويم التربوي،*** مكتبة التربية العربي لدول الخليج الطبعة الثانية،الرياض.

-أحمد يعقوب النور. 2007. ***القياس و التقويم في التربية و علم النفس،*** الجنادرية للنشر و التوزيع الأردن – عمان.

-إسماعيل بشرى.2007. ***مدخل إلى علم النفس في القرن الحادي والعشرون***، الاتجاهات الثقافية للنشر و التوزيع.

-إبراهيم الشوربيجي. 2004. ***تقويم التحصيل الدراسي للطالب الجامعي في ضوء الاتجاهات العالمية المعاصرة،***جامعة الزقازيق،القاهرة.

-أبو هاشم ، السيد مُحَمَّد. 2006. ***دراسة مقارنة بين النظرية التقليدية و نموذج راش ، في اختيار فقرات مقياس مداخل الدراسة لدى طلاب الجامعة*** ، مجلة كلية التربية – العدد (08)،جامعة الزقازيق،مصر، 52 صفحة.

-أمينة مُحَمَّد كاظم . 1996. ***اتجاهات معاصرة فى القياس والتقويم النفسى والتربوي" القاهرة ، الانجلو المصرية،***

-ان اناستازي، سوزانا اوريانا، ترجمة : صلاح الدين محمود علام .2015. ***القياس النفسي***، عمان: دار الفكر ناشرون وموزعون، الطبعة الأولى.

- أنور مُحَمَّد الشرقاوي، وسليمان الخضري الشيخ، وأمينة كاظم، ونادية عبد السلام.1996 . ***اتجاهات معاصرة في القياس و التقويم النفسي و التربوي،*** مكتبة الأنجلو - المصرية ، القاهرة.

-بدر مُحَمَّد الأنصار .2000 .***قياس الشخصية*** ، دار الكتاب الحديث، كويت، الطبعة الأولى.

-بركات مُحَمَّد خليفة : بدون سنة .***الإختبارات والمقاييس العقلية*** ، مكتبة مصر ، الطبعة الثانية .

-بوسالم عبد العزيز.2008. ***توظيف نموذج راش أحادي البرامتر في بناء اختبار تحصيلي في القياس النفسي و تحقيق التفسير الموضوعي لنتائجه بالمقارنة مع النظرية الكلاسكية ، دراسة سيكومترية مقارنة بين نظرية السمات الكامنة و النظرية الكلاسكية في بناء الاختبارات و التحقيق من موضوعيتها***، رسالة دكتوراة في علوم التربية، غير منشورة.

-تيسير مفلح كوافحة. 2003. ***القياس و التقييم و أساليب القياس و التشخيص في التربية الخاصة***، دار المسيرة ، الطبعة الأولى ،عمان– الأردن.

-حسن غانم مُحَمَّد. 2006. ***القياس النفسي للشخصية*** ، المكتبة المصرية الإسكندرية، الطبعة الأولى.

- خطاب علي ماهر . 2005. ***القياس و التقويم في العلوم النفسية و التربوية الإجتماعية***، المكتبة الأنجلو مصرية، القاهرة ، الطبعة الثانية.

-خطاب علي ماهر .2000. ***القياس و التقويم في العلوم النفسية و التربوية الاجتماعية*** ، المكتبة الأنجلو – مصرية، القاهرة ، الطبعة الثانية ،.

-ديانا فهمي علي حماد . 2010..***علاقة الحكمة الإختبارية بالأداء الناتج من إختبار تحصيلي ذي إختيار متعدد مبني وفق نموذج "راش" لدى طالبات كلية التربية***، رسالة دكتوراه- تخصص قياس و التقويم، غير منشورة،

- سامي مُحَمَّد ملحم (2007) ، ***مناهج التربية في التربية وعلم النفس*** ، عمان : دار المسيرة ، الطبعة الخامسة

. موسى نبهان (2004)، ***أساسيات القياس في العلوم السلوكية*** ،عمان : دار الشروق للنشر و التوزيع،الطبعة الأولى.

. سامي مُحَمَّد ملحم، (2015)، ***القياس و التقويم في التربية*** ،عمان : دار المسيرة للنشر و التوزيع،الطبعة السابعة.

.الطريري عبد الرحمان .1997.***القياس النفسي و التربوي- نظريته،أسسه،تطبيقاته***، مكتبة الرشد،الرياض.

.سامي مُحَمَّد ملحم .2005. ***القياس والتقويم في التربية وعلم النفس***، دار المسيرة للنشر والتوزيع،الأردن ،الطبعة الثالثة.

.سبع مُحَمَّد أبو لبدة .2008.***مبادئ القياس النفسي و التقييم التربوي***، دار الفكر موزعون و ناشرون الطبعة الأولى،الأردن - عمان،.

.سعد حسن آل عبد الفتاح الغامدي.2003. ***مدى اختلاف الخصائص السيكومترية لأداة القياس في ضوء تغاير عدد بدائل الاستجابة و المرحلة الدراسة – دراسة حالة مقياس ليكرت***،رسالة ماجستيير، جامعة أم القرى ،، غير منشورة .

.سعد عبد الرحمان.1998 .**القياس النفسي (النظرية و التطبيق)**، دار الفكر العربي، القاهرة ،الطبعة الثالثة.

.صلاح احمد مراد ، فوزية عباس هادي ، هشام فتحي جاد الرب (2017) ،***الإحصاء الاستدلالي في العلوم السلوكية*** ،القاهرة : دار الكتاب الحديث ،الطبعة الأولى

.طارق بن إبراهيم جابر الجهني. 2004. ***أثر إختلاف طرق تحليل بنود الاختبار المحكي المرجع على إختيار الطبقات و الثبات*** ، رسالة ماجستير غير منشورة ، جامعة الملك سعود ، سعودية.

.عباس فيصل . 1996.***الإختبارات النفسية تقنياتها واجراءته.*** دار الفكر العربي،الطبعة الأولى، بيروت.

.عبد الهادي السيد عبده، فاروق السيد عثمان.2002. ***القياس والاختبارات النفسية الأسس والأدوات***، القاهرة:دار الفكر العربي، الطبعة الأولى.

.عبد الهادي السيد عبده،فاروق السيد عثمان. 2002. ***القياس والاختبارات النفسية الأسس والأدوات***، دار الفكر العربي القاهرة ، الطبعة الأولى.

.علام صلاح الدين محمود .2001. ***الاختبارات التشخيصية المرجعية المحك في المجالات التربوية والنفسية والتدريبية***، دار الفكر العربي، القاهرة، الطبعة الثانية.

.علام صلاح الدين محمود .2005. ***نماذج الاستجابة للمفردة الاختبارية أحادية البعد و متعددة الأبعاد و تطبيقاتها في القياس النفسي و التربوي*** ، دار الفكر العربي، القاهرة،الطبعة الاولى.

.علام صلاح الدين محمود .2000. ***القياس و التقويم التربوي و النفسي أساسياته و توجيهاته المعاصرة***، دار الفكر العربي، القاهرة.

علام صلاح الدين محمود . 2006. ***الاختبارات و المقاييس التربوية و النفسية***، دار الفكر ،الطبعة الأولى، عمان – الأردن،.

عبد العزيز بوسالم . 2008.***توظيف نموذج راش أحادي البرامتر في بناء اختبار تحصيلي في القياس النفسي و تحقيق التفسير الموضوعي لنتائجه بالمقارنة مع النظرية الكلاسكية ، دراسة سيكومترية مقارنة بين نظرية السمات الكامنة و النظرية الكلاسكية في بناء الاختبارات و التحقيق من موضوعيتها***، رسالة دكتوراة في علوم التربية ، غير منشورة ،.

علي زكري. 2009. ***الخصائص السيكومترية لإختبار (أوتيس – لينون) للقدرة العقلية مقدرة وفق القياس الكلاسيكي و نموذج راش لدى طلبة المرحلة المتوسطة بمحافظة صبيا التعليمية***، رسالة دكتوراه غير منشورة ، كلية التربية جامعة أم القرى ، السعودية،.

عماد عبد المسيح . 1994.***استخدام نموذج راش اللوغاريتمى أحادى البارامتر فى تحليل مفردات الاختبارات المعرفية مرجعية المعيار ثنائية القطب (دراسة تجريبية)***، جامعة المنيا ، كلية التربية ، مجلة البحث فى التربية وعلم النفس ، العدد(4) ابريل ، 443-475.

.علي ماهر خطاب.(2000)، ***القياس و التقويم في العلوم النفسية و التربوية الاجتماعية*** ، القاهرة : المكتبة الانجلو – مصرية، الطبعة الثانية.

.عوض محمود عباس. 1999 . ***القياس النفسي بين النظرية والتطبيق***، دار المعرفة الجامعية، الاسكندرية.

.فرج صفوت. 2000. ***القياس النفسي***، الأنجلو المصرية، القاهرة، الطبعة الرابعة ،.

.فؤاد أبو حطب و سيد أحمد عثمان . 1997. ***التقويم النفسي***، مكتبة الأنجلو مصرية، القاهرة، الطبعة الرابعة،.

.كوافحة تسير مفلح . 2005. ***القياس و التقويم و أساليب القياس و التشخيص في التربية الخاصة***، دار المسيرة للنشر و التوزيع، عمان ، الطبعة الثانية.

.ليونا,أ,تايلر، ترجمةسعد عبد الرحمان، مُحَمَّد عثمان النجاتي.ب.ت. ***الاختبارات والمقاييس***، ديوان المطبوعات الجامعية بن عكنون ، بدون طبعة .

.مجدي عبد الكريم حبيب . 1996. ***التقويم و القياس في التربية و علم النفس،*** مكتبة النهضة، مصر ،الطبعة الأولى .

ماجد مُحَمَّد الخياط (2010)، ***أساسيات البحوث الكمية والنوعية في العلوم الاجتماعية***، الأردن: دار الراية لنشر والتوزيع، الطبعة الأولى.

.محمود أبو علام رجاء. 2004. ***مناهج البحث في العلوم النفسية والتربوية***، دار النشر للجامعات ، الطبعة الرابعة، القاهرة - مصر ،

.مراد احمد صلاح، سليمان امين علي.2005. ***الإختبارات والمقاييس في العلوم النفسية والتربوية -خطوات إعدادها وخصائصها***-، دار الكتاب الحديث ، الطبعة الأولى ، الكويت،

.-مقدم عبد الحفيظ ، (2003)، ***الإحصاء و القياس النفسي و التربوي مع- نماذج من المقاييس و الاختبارات*** الجزائر : ديوان المطبوعات الجامعية ، الطبعة الثانية.

.ملحم سامي مُحَمَّد .2012. ***القياس و التقويم في التربية وعلم النفس*** ، دار المسيرة للنشر و التوزيع،الطبعة السابعة، عمان - الأردن،.

.ملحم سامي مُحَمَّد. 2002. ***القياس و التقويم في التربية*** ، دار المسيرة للنشر و التوزيع،الطبعة الثانية عمان - الأردن،

.نادر فهمي الزيود، هشام عامر عليان. 2005. ***مبادئ القياس و التقويم في التربية،*** دار الفكر ناشرون و موزعون، عمان- الأردن، الطبعة الثالثة.

.نبهان موسى . 2004. ***أساسيات القياس في العلوم السلوكية*** ، دار الشروق للنشر و التوزيع، الطبعة الأولى، عمان ،

-سامي مُحَمَّد ملحم (2005)، ***القياس والتقويم في التربية وعلم النفس***،الأردن : دار المسيرة للنشر والتوزيع، ،الطبعة الثالثة.

- Anastasi, A. and Urbina, S. Psychological Testing New jersey 07458: Upper (1997) Saddle

-Back,R.D: A brief history of item response Theory.**Educational Measurement: Issues and practice.(1997)**

-Crocker, L. & Algina, J. Introduction to Classical and Modern Test Theory.NY: Holt, Rinhart and Winston.(1986) . -
Edel, R. Frisbie, D. Essentials of Educational Measurement. New Jersey1995) .

-Engelhard, G: Historical views of invariance: Evidence from the measurement theories of Thorndike, Thurstone, and Rasch**. Educational and Psychological Measurement, N**52**,P 275 – 291 , (1992) .**

-Garland, Ron. The Mid-Point on aRating Scale: Is it Desirable?.Marketing Bulletin, 2, 66-70.(1991) . -

-Gregory, R. Psychological testing History ,Principles aplications 4 th .Ed , New York .Pearson Education Group Inc.(2004) .

Mehrens J. and Lehmann, D.R, "Are Three-Point ScalesAlways Good Enough?" Journal of Marketing Research, 9, 4(November).(1980) ,

-Hambleton, R. & Swaminathan ,H. Item response Theory Principles and applications **. Boston , Kluwer Nijhoff Publishing.(1989)**

-Randall ,S .Comparing Measurement Theories . Paper Present at the Annual Meeting of the American Educational Research Association (San Diego , CA , April 13-17).(1998). .

Randall ,S. Comparing Measurement Theories **. Paper Present at the Annual - Meeting of the American Educational Research Association (San Diego , CA , April 13-17), (1998).**

- Suen , H .Principles Of Test Theories. Louence Erlbaum Associates ,Inc (1990) .Hillsdale, New Jersey