التحقق من الخصائص السيكومترية في اختبار اختيار من متعدد باستخدام حذف المشتت الخاطئ وفقا لنظرية السمات الكامنة

م.م علاء حسين عبيس الطائي

التحقق من الخصائص السيكومترية في اختبار اختيار من متعدد باستخدام حذف المشتت الخاطئ وفقا لنظرية السمات الكامنة

اعداد

م.م علاء حسين عبيس الطائي

قسم العلوم التربوية والنفسية - كلية التربية

للعلوم الانسانية - جامعة بابل

Laa202431@gmail.com

هاتف (07822944800)

المستخلص .

التعقيدات في القياس ضاعفت التعقيد الحسابي على الباحثين عندما حاول الباحثون الجمع بين النماذج التقليدية للقياس لكي تتناول الهياكل القياسية، مما زاد من صعوبة الاعتماد الشديد على تفسيرات الأساليب الإحصائية التكرارية والتي تحد من الاستفادة من التفسيرات والأساليب الإحصائية الحديثة والتقنيات الحاسوبية المتاحة وهنا تبرز مشكلة البحث الحالي.

ولقد اهتم البحث السيكومتري المتعلق بنظرية الاستجابة للمفردة بتطوير النماذج الاحتمالية لتحديد العلاقة بين أداء الفرد على مفردات الاختبار وبين السمات أو القدرات الكامنة وراء هذا الأداء، ولم يقتصر البحث السيكومتري بالبحث عن النماذج الرياضية بل تعدى ذلك إلى إيجاد طرق التقدير العددي التي تستخدم في تقدير معالم المفردات وقدرة الفرد للوصول إلى أفضل التقديرات، من هنا تبرز الحاجة إلى نظريات تمكن من تحويل عينات السلوك المقاسة إلى درجات تكشف عن الفروق بين الافراد في مختلف السمات.

وبعد العرض لصدق الاختبار التحصيلي بأستعمال تقنية أستبعاد بديلين والأختبار التحصيلي نفسه بدون أستعمال تقنية أستبعاد بديلين ومقارنتهما، تبين انه لم يكن فرقا واضحة بينهما ، حيث ان الأختبار التحصيلي بأستعمال تقنية أستبعاد بديلين أو بدون أستعمال تقنية أستبعاد بديلين يتمتعان بصدق وظيفي عالي وبمستوى دلالة (0.01)، فالإجراءات المتبعة تبين ان صورة الأختبار بأستعمال تقنية أستبعاد بديلين يتمتع بقدرة على التمييز بين الأفراد الذين يمتلكون مستوى عالي من القدرة والذين يمتلكون مستوى متدني من القدرة ، وكان

الاختبار التحصيلي بدون أستعمال تقنية أستبعاد بديلين الذي تمتع بقدرة على التمييز أعلى بين العينة مرتفعة المستوى والعينة منخفضة المستوى وبذلك فكلا الصورتين يمكن أستعمالها لأتخاذ مستوى دقة أعلى في القرار.

اما بخصوص تجانس الاختبار حيث لم تظهر نتائج تحليل الاختبار فرقا ذو دلالة أحصائية بين معاملي الثبات لأختبار التحصيلي لمادة علم نفس الخواص بأستعمال تقنية أستبعاد بديلين وبدون أستعمال تقنية أستبعاد بديلين عند مستوى دلالة أحصائية (0.05)، لأن القيمة الفائية المحسوبة للفرق بين معاملي الثبات أصغر من القيمة الفائية الجدولية البالغة (3.850)، وقد يعود عدم الفرق بين الأختبار بأستعمال التقنية وبدون أستعمالها في معامل الثبات إلى اختيار الفقرات ذات المؤشرات الدقيقة في ملاءمتها للأنموذج المعتمد.

وقد توصل الباحث إلى مجموعة من الاستنتاجات والتوصيات.

Abstract:

The complexities of measurement have compounded the computational complexity for researchers when they attempt to combine traditional measurement models to address standard structures. This has increased the difficulty of relying heavily on interpretations of frequentist statistical methods, which limits the benefit of modern statistical interpretations and methods and available computational techniques. This is where the problem of the current research emerges.

Psychometric research related to item response theory has been concerned with developing probabilistic models to determine the relationship between an individual's performance on test items and the traits or abilities underlying this performance. Psychometric research has not been limited to searching for mathematical models, but has extended beyond that to finding numerical estimation methods that are used to estimate the parameters of items and the individual's ability to reach the best estimates, Hence the need for theories that enable the transformation of measured behavioral samples into scores that reveal differences between individuals in various traits.

After presenting the validity of the achievement test using the two-alternative exclusion technique and the same achievement test without using the two-alternative exclusion technique and comparing them, it became clear that there was no clear difference between them, as the achievement test using the

التحقق من الخصائص السيكومترية في اختبار اختيار من متعدد باستخدام حذف المشتت الخاطئ وفقا لنظرية السمات الكامنة

م.م علاء حسين عبيس الطائي

two-alternative exclusion technique or without using the two-alternative exclusion technique enjoyed high functional validity and a significance level of (0.01), The procedures followed show that the test format using the two-alternative elimination technique has the ability to discriminate between individuals with a high level of ability and those with a low level of ability, while the achievement test without the two-alternative elimination technique has the ability to discriminate more between the high-level sample and the low-level sample. Therefore, both formats can be used to make a higher level of accuracy in decision-making.

As for the homogeneity of the test, the test analysis results did not show a statistically significant difference between the stability coefficients of the achievement test for the subject of properties psychology using the technique of excluding two alternatives and without using the technique of excluding two alternatives at a statistical significance level of (0.05), because the calculated p-value of the difference between the stability coefficients is smaller than the tabular p-value of (3.850), The lack of difference in reliability between the tests using and without the technology may be due to the selection of items with precise indicators of their suitability to the adopted model.

The researcher reached a set of conclusions and recommendations.

الكلمات المفتاحية : الخصائص السيكومترية ، اختبار اختيار من متعدد، حذف المشتت الخاطئ، نظرية السمات الكامنة .

1. مشكلة البحث :

أدوات القياس ونتائج استخدامها وتوظيفها تعتمد بالدرجة الأولى على إمكانية تحقيقها

للأهداف التي وضعت من أجلها، وقدرة تلك الأدوات ودقتها في تحديد مقدار السمة المعنية بعملية القياس، وهذا ما يسمى بصدق أداة القياس، وبغض النظر عن الخصائص الأخرى في أي مقياس فإنه إذا لم يتمتع بدرجة صدق مقبولة فإن المعلومات التي يزودنا بها تكون عديمة الفائدة، وهذا المفهوم يتطلب عدة أنواع من الشواهد الامبريقية (الأدلة) المستمدة من درجات الاختبار، فعمليات تأويل درجات الاختبار وتفسيرها التي تقتضيها أنواع الاستعمالات المقترحة هي التي يتم تقييمها لتقدير الصدق، وليس الاختبار في حد ذاته، وهو مفهوم يمكن من تعميم توظيف الاختبار واستخداماته، ويسهم في اتخاذ القرارات التربوية والمهنية الفعالة (الزيات، المحرزي، 2011: 44).

لذلك بنود الاختبار التي تقيس في الواقع عينة من السلوك تعكس بالحقيقة السلوك مدار اهتمام الدراسة، من ذلك تأتي الحاجة إلى النظريات التي نتمكن من خلالها حل المشكلات التي تواجه الباحثين عند تحويل خصائص السلوكيات المقاسة للأفراد إلى درجات تكشف عن الفروق الحقيقية بين الافراد في خواص السمات المختلفة (Brooks, Gelman, Jones, & Meng, 2011: 27)

وفي هذا السياق تأتي هذه الدراسة ضمن الجهود المبذولة في الكشف عن مدى الاختلاف في المعنى الذي يمكن إعطاؤه للسمة الكامنة من خلال العلامة التي يحصل عليها الفرد باستخدام النظرية الحديثة في القياس، والبحث عن التبريرات لاستخدام التدريجات وفق نظرية السمات الكامنة في القياس وما تتضمنه هذه النظرية من افتراضات تختلف عن نظيرتها (النظرية الكلاسيكية)، حيث تسعى للمساهمة في حل نقاط الضعف في الاختبارات أو أنظمة القياس التربوي والنفسي لكي تزودنا بأساس علمي لقياس يتصف بالتعميم، وهذا لا يتأتى إلا إذا استند تصميم هذه الأنظمة إلى نظرية أو نموذج إحصائي سيكومتري يفسر كيفية الاستجابة للمفردات أو الأسئلة الاختبارية، فالطابع الكمي المتزايد الذي أصبحت تتسم به هذه النظم جعل تطبيقات الحاسوب والاحصاء لها دور كبير في معالجة مواطن الضعف ليس فقط في تحليل الاختبارات والبيانات ومعالجتها، وإنما في نمذجة وتبسيط كثير من الظواهر النفسية والتربوية، فأن النماذج الاحصائية الرياضية صارت لها استخدامات كبيرة وتستخدم بكثرة في وقتنا الخاضر لتفسير وتطوير الاختبارات والمقاييس وتفسير نتائجها. (علام، 1987: 112)

فهذه التعقيدات في القياس ضاعفت التعقيد الحسابي على الباحثين عندما حاول الباحثون الجمع بين النماذج التقليدية للقياس لكي تتناول الهياكل القياسية، مما زاد من صعوبة الاعتماد الشديد على تفسيرات الأساليب الإحصائية التكرارية والتي تحد من الاستفادة من التفسيرات والأساليب الإحصائية الحديثة والتقنيات الحاسوبية المتاحة (Brooks, Gelman, Jones, & Meng, 2011: 102).

التحقق من الخصائص السيكومترية في اختبار اختيار من متعدد باستخدام حذف المشتت الخاطئ وفقاً لنظرية السمات الكامنة

م.م علاء حسين عبيس الطائي

تعد الاختبارات في المجال التربوي من إحدى وسائل التقويم المتنوعة، وهي وسيلة رئيسة تعمل على قياس مستوى تحصيل الطلاب ، والتعرف على مدى تحقيق المنهج الدراسي للأهداف المرسومة له ، والكشف عن مواطن القوة والضعف في ذلك ، ومدى التقدم الذي أحرزته المؤسسة التعليمية، وبذلك يمكن على ضوءه العمل على تحسين وتطوير العملية التعليمية والتربوية إلى الأفضل. وتعد الاختبارات وسيلة من وسائل التقويم التربوي الشائعة فهو عملية مرتبطة بأهداف محددة تتم في أحوال معينة وفق قواعد خاصة يترتب عليها النجاح أو الرسوب وما يعقب ذلك من انتقال من صف إلى آخر ومن المتوقع أن الاختبارات التي يمكن الحصول عليها من خلال الكثير من المعلومات حول كيفية ومدى عمق قدرات الطلاب في مجال التدريس (معروف، 1998: 249)

يشهد التوجه العام للنظم المعرفية المعاصرة الدولية والمحلية، اهتماماً متزايداً في تقويم نظمها وضمان جودتها، وبناء المعايير والمؤشرات الأساسية لتقويم أدائها واعتمادها. والباحث المتتبع لحراك المشهد التنموي يلحظ تجسد ذلك الاهتمام في تحولات مهمة على مختلف النظم عموماً والتعليمية منها خصوصاً، رغبة في تعزيز جودة مخرجاتها، وتحقيقها لمعايير الاعتماد المحلية والعالمية، ووضع المؤشرات اللازمة، وقياس نواتجها، رافق ذلك عمليات التطوير التي أدت إلى المزيد من الاهتمام بالقياس والتقويم لضبط الجودة وتحقيق كفاية عالية لأنظمتها ورصد الأداء وربطه بالمؤشرات العالمية؛ باستخدام أدوات للقياس والتقويم على مستوى عالٍ من الموضوعية والمصداقية، تستند إلى معايير مبنية على أحدث التوجهات العالمية. ومن ثم فإن بنود الاختبار تقيس في الواقع عينة من السلوك تعكس السمة موضع الاهتمام. من هنا تبرز الحاجة إلى نظريات تمكن من تحويل عينات السلوك المقاسة إلى درجات تكشف عن الفروق بين الأفراد في هذه السمات (الزيات، المحرزي، 2011:67)

ولقد اهتم البحث السيكومتري المتعلق بنظرية الاستجابة للمفردة بتطوير النماذج الاحتمالية لتحديد العلاقة بين أداء الفرد على مفردات الاختبار وبين السمات أو القدرات الكامنة وراء هذا الأداء، ولم يقتصر البحث السيكومتري بالبحث عن النماذج الرياضية بل تعدى ذلك إلى إيجاد طرق التقدير العددي التي تستخدم في تقدير معالم المفردات وقدرة الفرد للوصول إلى أفضل التقديرات (بني عطا، الشريفين، 2012: 73).

3. هدف البحث .

يهدف البحث الحالي الى:

هدف البحث الحالي هو التحقق من الخصائص السيكومترية عند استخدام اختبار محكي المرجع) تحصيلي اختيار من متعدد(M C Q) بطريقة ابعاد المشتتات الخاطئه وفق نظرية السمات الكامنه.

4. حدود البحث :

يتحدد البحث الحالي في :

- 1- كليات التربية للعلوم الانسانية ا قسم العلوم التربوية والنفسية ا جامعات الفرات الاوسط (جامعة القادسية، جامعة الكوفة، جامعة كربلاء) المراحل الرابعة.
- 2- المادة الدراسية المقررة من قبل وزارة التعليم العالي والبحث العلمي في العراق للمراحل الرابعة من كليات التربية للعلوم الانسانية قسم العلوم التربوية والنفسية علم نفس الخواص (التربية الخاصة).
- 3- تبنى الباحث اختبار شاكر (2012) المبني لمادة علم نفس الخواص وهو اختبار (MCQ).

5. تحديد المصطلحات:

الخصائص السيكومترية : عرفه (علام 2011):

يقصد بالخصائص السيكومترية (القياسية) في تخصص القياس النفسي هي مؤشرات احصائية تعبر عن دقة وجودة وكفاءة ادوات القياس ومفردات الاختبار المستخدمة في تقييم السمات والظواهر النفسية، وهذه المؤشرات هي الصدق، الثبات، الصعوبة، التمييز، الاتساق الداخلي (ميخائيل، 1997: 100).

اختبار الاختيار من متعدد : عرفه (عودة 2010):-

يتكون اختبار الاختيار من متعدد من المتن أو النص أو الجذر حيث يشرح فيه المعلم المشكلة، ويتبعه ثلاثة بدائل أو أكثر، أحد هذه البدائل هو الإجابة الصحيحة أو الأصح أحياناً وباقي البدائل عبارة عن موهات يطلب من الطالب أن يبين الإجابة الصحيحة، وقد يظهر المطلوب على شكل سؤال أو على شكل جملة مفتوحة تكتمل بالإجابة الصحيحة (عودة، 2010 : 159).

حذف المشتت الخاطئ عرفه كل من:-

• (Henrysson.1991):-

التحقق من الخصائص السيكومترية في اختبار اختيار من متعدد باستخدام حذف المشتت الخاطئ وفقاً لنظرية السمات الكامنة

م.م علاء حسين عبيس الطائي

هو اجراء وليس خاصية قياسية للاختبار، بمعنى اجراء اساسي في تحليل الفقرات عندما نستخدم اختبارات من نوع الاختيار من متعدد ، ويهدف هذا الاجراء الى ايجاد بدائل اجابة فعالة أي جذابة للمفحوصين على الرغم من انها بدائل مخطئة (Henrysson , 1991 : 138).

• كناني وفيضي (2012):-

تعتبر طريقة يتم من خلالها الاجابة عن أسئلة اختبارات البدائل المتعددة بعد الاطلاع على السؤال وقراءته بشكل جيد عن طريق ابعاد وحذف البديلين الخطأ حسب ما يضمن المختبر وليس من ضمن الاجابة الصحيحة من وجهة نظر المختبرين ويشطبهم من ورقة الاسئلة، وبعد ذلك يحاول الطلاب التفكير العميق بين البدائل المتبقية ويضع اشارة حول الاجابة الصحيحة. (الكناني وفيضي، 2012: 212) .

نظرية السمات الكامنة : عرفها كل من :-

• امينة كاظم (1996):-

تعد من الاتجاهات الحديثة التي يسعى إلى تحقيق الموضوعية وتوفير الاستقلالية والخطية في القياس، وكذلك تكامل عملية وأهداف القياس، ومن افتراضاتها هي وجود واحدة أو أكثر من الخصائص الرئيسية غير الملاحظة التي تحدد استجابات الفرد الملاحظة لبنود اختبار ما (كاظم، 1996: 54).

• علام (2005):-

هي النظرية التي تهتم بالربط بين استجابة المختبرين لمفردة او فقرة اختبارية وبين معالم هذه المفردة، وتقوم على افتراض وجود خط للسمة، من خلاله يمكن تقدير احتمال الإجابة الصحيحة للممتحن عن مفردة على هذا المتصل، وأن العلاقة بين أداء المختبر على الفقرة وقدرته (θ) يمكن أن تتوضح من خلال ما يسمى منحني سمات المفردة وخصائصها (بني عطا، الشريفين، 2012:110).

6. الإطار النظري Theoretical Framework :

المحور الأول: نظريات القياس النفسي Psychological Measurement Theories :

تعد النظريات ركناً أساسياً لبناء المعرفة، من خلال كونها أداة لتفسير ونقد وتوحيد القوانين الراسخة وتوظيفها، إضافةً إلى إسهامها في اكتشاف تعميمات جديدة أكثر قوة. وأسهمت التطبيقات الأولية لتصميم التجارب في ميادين العلوم المختلفة في ظهور Bose، Box علم تصميم التجارب في شكله الحديث والتي كان من أبرز روادها علماء الانكليز (Cochran، Fisher، Kempthorne، Yates) وغيرهم، حيث كان لإسهاماتهم دور فعال في تقدم هذا الفرع الهام من فروع الإحصاء والتي أصبحت الآن تستخدم في شتى ميادين العلم وفروع المعرفة (Rekab & Shaikh, 2005, 166).

والمنتبع لمراحل التطور التاريخي لنظرية القياس وتطبيقاتها في سياق العلوم التربوية والنفسية، يلحظ الأثر الواضح لمثل هذه الجهود، تزامنا مع الإسهامات الأولى لعلماء علم الاجتماع وعلم النفس في ظهور هذه النظرية كفرع من فروع المعرفة.

ويعرف كل من (Crocker & Algina) نظرية القياس بأنها دراسة مشكلات القياس وحلولها في سياق العلوم التربوية والنفسية، كما حددا الموضوعات التي تبحثها نظرية القياس في موضوعين أساسيين هما:

١- تقدير مدى تأثير هذه المشكلات على قياس أجري في موقف معين.

٢- اقتراح وسائل أو طرائق للتغلب على هذه المشكلات أو تقليل أثرها.

كما إلى أن الهدف من نظرية القياس هو المساعدة في ادراك النماذج الرياضية والإحصائية التي تشكل الأساس في بناء الاختبارات المعيارية وتطبيقاتها، حيث يؤدي ادراك هذه النماذج وافترضاها إلى تطبيقات مطورة في اتخاذ القرارات، كما كما اشارة أيضاً إلى أنه في الوقت الذي تركز موضوعات التقييم التربوي والنفسي عادةً على إدارة اختبارات معينة وتفسيرها نجد في المقابل أن نظرية القياس تزودنا بهيكل عام يصور عملية تطوير الأداة، بحيث لا يقف عند نظرية نفسية أو تربوية معينة، بل تكون جميعها مفيدة وبدرجات متساوية في قياس العديد من الخصائص النفسية (Crocker & Algina, 2006:234).

اولاً: نظرية القياس التقليدية Classical Test Theory:

يبرز دور علماء القياس النفسي في محاولاتهم لقياس البنى النفسية والمفاهيم المجردة غير لهذه البنى Scaling Methods القابلة للقياس المباشر، حيث أسهمت طرق تدرج المقاييس والمفاهيم المجردة في تحليل البيانات الناتجة عن الاستجابات الذاتية للأفراد، كما وفرت لهم الأساليب والطرق العلمية لتحويل السمات النفسية أو المثيرات من مفاهيم مجردة إلى أرقام قد لاتصل إلى حدود المستوى الفئوي تماماً، والذي يمتاز بتحديد الاتجاه ووجود مسافات تفصل بين المستويات المختلفة المتعلقة بالمثير أو السمة على أساس تجريبي ومن واقع مشاهدات مسجلة عن الموضوع المراد قياسه وتختلف باختلاف الأبعاد للسمة المقاسة وتتعدد طرق تدرج المقاييس

حيث تسعى إلى بناء أداة قياس تقترن فيها التكوينات الفرضية أو البنى النوعية مع وحدات قياسية كمية، في محاولة لتعيين أرقام تناظر بعض خصائص الأشياء أو الأشخاص وفق قواعد تدرج محددة، ومن ثم فإن

التحقق من الخصائص السيكومترية في اختبار اختيار من متعدد باستخدام حذف المشتت الخاطئ وفقاً لنظرية السمات الكامنة

م.م علاء حسين عبيس الطائي

مفردات أداة القياس تقيس في الواقع عينة من السلوك تعكس السمة موضع الاهتمام. ومن هنا تبرز الحاجة إلى نظريات تمكن من تحويل عينات السلوك المقاسة إلى درجات تكشف عن الفروق بين الأفراد في هذه السمات. وفي هذا السياق تشكل مفاهيم القياس وفق النظرية التقليدية الركيزة الأساسية لحركة القياس النفسي والتربوي، حيث اعتمدت هذه النظرية في بناء مختلف الاختبارات والمقاييس النفسية والتربوية على معايير قياسية إحصائية خاصة بتحليل مفردات المقاييس والاختبارات، وكان الهدف فيها قياس الفروق الفردية بين الأشخاص في القدرات العقلية أو الخصائص والسمات الشخصية من خلال مقارنة درجة الفرد بدرجات مجموعته المعيارية التي ينتمي إليها بالاعتماد على التوزيع الطبيعي للظواهر والخصائص النفسية والتربوية في تمييز وإظهار الفروق الفردية في السمات والخصائص النفسية المقيسة. (Wallace, & , 2004 :97) Zhang

الدرجة الحقيقية في نظرية القياس التقليدية:

في هذا السياق تفترض نظرية القياس التقليدية أن الدرجة التي يحصل عليها الفرد في الاختبار هي ببساطة مجموع درجات المفردات التي يتألف منها الاختبار، ويشار إلى هذه الدرجة باسم الدرجة المشاهدة وهذه الدرجة هي مجموع درجتين، (Observed Score) الخام والدرجة الحقيقية التي تعكس ما يمتلكه الفرد فعلاً من السمة المقاسة موضع الاهتمام، ودرجة الخطأ التي تعود إلى عوامل عشوائية الخطأ العشوائي، وتؤدي إلى انحراف الدرجة المشاهدة عن الدرجة الحقيقية في الاتجاه السالب أحياناً، وفي الاتجاه الموجب أحياناً أخرى، فعندما يجب الفرد على مجموعة مفردات اختبارية فإن درجته تمثل فقط عينة محددة من السمة المراد قياسها وهي استجابات على مجموعة فرعية من المفردات المحتملة ضمن مجال معين، تم الحصول عليها في واحد من المواقف الممكنة، لذا فإن الدرجات التي نحصل عليها تحت الظروف الاختبارية عرضة للخطأ وخاضعة لخطأ القياس. (Crocker & Algina, 2006:164)

ويمكن تصنيف أخطاء القياس على أنها أخطاء عشوائية أو منتظمة، فالأخطاء المنتظمة هي الأخطاء التي تؤثر بانتظام على درجة الفرد بسبب خصائص معينة للفرد أو الاختبار التي لا نستطيع عمل شيء تجاهها لبناء المقياس، بالمقابل تؤثر الأخطاء العشوائية للقياس على درجة الفرد لأن حدوثها محض صدفة، واعد ستانلي قائمة تصنيفية مفصلة لمصادر التباين العشوائي الفردية يمكن الرجوع إليها والإفادة منها من قبل مطوري الاختبارات والمقاييس. ويعد كلاً من الأخطاء العشوائية والمنتظمة مصدراً مهماً في التفسير

الذي يمكن إعطاؤه للسمة الكامنة من خلال العلامة التي يحصل عليها الفرد في إطار نظرية القياس التقليدية، وايضا اشار إلى أن أخطاء القياس المنتظمة لا ينتج عنها عدم تجانس في القياس، ولكنها تبقى سبباً لعدم دقة الاختبار، وبالتالي تقلل من استخداماته العملية، اما الأخطاء العشوائية فإنها تقلل من تجانس كل من درجات الاختبار ومدى الإفادة منها، وبالتالي فإنه من غير المنطقي توقع فائدة من القياسات إن لم يكن لدينا ثقة بأنها متجانسة، لذا كان من أهم مسؤوليات مطوري الاختبارات البرهنة على ثبات درجات اختباراتهم، حيث تتطلب مثل هذه البراهين دراسات تجريبية تعتمد في الأغلب على نموذج نظري يصف مدى تأثير الأخطاء العشوائية على الدرجات. لذلك تسعى نظرية القياس التقليدية إلى تقدير نسبة الخطأ الموجودة في نتائج القياس عن طريق مجموعة من التقنيات والمفاهيم وأهمها (مفهوم الصدق ومفهوم الثبات)، وبالتالي الوصول إلى تقدير مقبول للدرجة الحقيقية لأداء الفرد من خلال الدرجة المشاهدة. ويعتمد ثبات وصدق الاختبار على خصائص المفردات التي يتم اختيارها، وتقوم عملية الاختيار على ثلاث مؤشرات هي معاملات: صعوبة المفردة، والتمييز، وثبات المفردة. (Stanly, 1971: 87)

تعريف الخطأ في إطار نظرية القياس التقليدية:

إن أي نظرية في القياس التربوي والنفسي تحاول أن تقدم حلاً لمشكلات تتعلق بأخطاء القياس وفهم أثر أخطاء القياس في تقدير قدرات الافراد المفحوصين وكيفية تقليل هذه الأخطاء، وفي إطار نموذج الدرجة الحقيقية التقليدي فإن خطأ القياس هو الفرق بين درجة الفرد المشاهدة ودرجته الحقيقية. (علام، 2002:157)

ثانياً: نظرية الاستجابة للمفردة Item Response Theory :

أسهمت النماذج التقليدية للاختبارات في العقود الماضية إسهامات متعددة في بناء الاختبارات النفسية والتربوية، وتحليل وتفسير نتائجها، إلا أنه ونظراً لنواحي القصور التي عانت منها أساليب القياس التي يتم إعدادها وفقاً للنظرية التقليدية للقياس والتي بدت بصورة واضحة من خلال الانتقادات التي وجهت إليها، وأسهمت جهود كثير من علماء القياس والتي اتجهت في الآونة الأخيرة إلى تطوير نماذج سيكومترية جديدة أو نماذج الاستجابة للمفردة أطلق عليها نماذج السمات الكامنة. (علام، 1987: 213)

حيث أسفرت هذه الجهود في ظهور ما يسمى بنظرية الاستجابة للمفردة Item Response Models (IRM)، والنماذج الرياضية المرتبطة بها، وتقترض هذه النظرية وجود سمات أو خصائص يطلق عليها السمات أو القدرات تكمن خلف أداء الفرد على الاختبار، بحيث يمكن التنبؤ بقدرة الفرد من خلال هذا الأداء على الاختبار ودرجاته على الخصائص والسمات المقاسة من الاختبار وحيث أن هذه السمات لا يمكن ملاحظتها أو قياسها بشكل مباشر فيطلق عليها السمات الكامنة. (Haviland, 2005: 177)

وانبثقت عن نظرية السمات الكامنة بناءات إحصائية تم اشتقاقها إحصائياً من خلال العلاقة الملاحظة تجريبياً من الاستجابات على مفردات الاختبار، فالدرجة الكلية للاختبار تعد هي الحساب الأولي التقريبي

التحقق من الخصائص السيكومترية في اختبار اختيار من متعدد باستخدام حذف المشتت الخاطئ وفقاً لنظرية السمات الكامنة

م.م علاء حسين عبيس الطائي

للسمة الكامنة للمفردة ونماذج الاستجابة هي نماذج أو دوال رياضية على أساس رياضي احتمالي، حيث تتحدد العلاقة بين أداء الفرد على الاختبار والقدرة التي تكمن خلف هذا الأداء وفقاً لأسس نظرية الاحتمالات، حيث ان السبب في استخدام لفظة "استجابة" بدلاً من "الدرجة" يرجع إلى أن الاستجابة للمفردة يمكن أن تكون غير كمية وتتبع تدرجاً معيناً. (Anastasi & Urbina:1997)

نماذج نظرية الاستجابة للفقرة :Item Response Theory Models

أسهمت النماذج التقليدية للاختبارات في العقود الماضية إسهامات متعددة في بناء الاختبارات النفسية والتربوية، وتحليل وتفسير نتائجها، إلا أنه ونظراً لنواحي القصور التي عانت منها أساليب القياس التي يتم إعدادها وفقاً للنظرية التقليدية للقياس والتي بدت بصورة واضحة من خلال الانتقادات التي وجهت إليها، أسهمت جهود كثيرة من علماء القياس والتي اتجهت في الآونة الأخيرة إلى تطوير نماذج سيكومترية جديدة أو نماذج الاستجابة (Latent Trait Models) (LTM) أطلق عليها نماذج السمات الكامنة، حيث أسفرت هذه الجهود في ظهور ما يسمى بنظرية الاستجابة للمفردة، والنماذج الرياضية المرتبطة بها، وتقتضى هذه النظرية وجود سمات أو خصائص يطلق عليها السمات أو القدرات تكمن خلف أداء الفرد على الاختبار، بحيث يمكن التنبؤ بقدرة الفرد من خلال هذا الأداء على الاختبار ودرجاته وعلى الخصائص والسمات المقاسة من الاختبار، حيث أن هذه السمات لا يمكن ملاحظتها أو قياسها بشكل مباشر فيطلق عليها السمات الكامنة. (علام، 1987: 109)

صنفت نماذج نظرية الاستجابة للمفردة إلى نوعين: IRT Models هما النماذج التي تستخدم المنحنى الطبيعي التراكمي والنماذج اللوجستية Logistic Models وتُعد النماذج اللوجستية أوسع انتشاراً في الوقت الحاضر، وهي ملائمة ل فقرات ثنائية التدرج وتختلف هذه النماذج فيما بينها باختلاف عدد معالم المفردة، ويوجد العديد من النماذج في نظرية الاستجابة للمفردة غير أن هناك ثلاثة نماذج شائعة الاستخدام في البيانات الثنائية وهي المقدر بالدرجة (1) للاستجابة الصحيحة، و (0) للاستجابة الخاطئة، وهذه النماذج هي:

أولاً: النموذج اللوجستي أحادي المعلم: (1PLM) (One-Parameter Logistic Model)

ويُعرف بنموذج راش، وهو أكثر النماذج شهرة واستخداماً وذلك لسهولة تطبيقه، حيث يفترض هذا النموذج أن جميع الفقرات تُميز بين الافراد بنفس القدر، كما يفترض عدم تأثر الإجابات بعامل التخمين، ولكنها تتباين فقط في صعوبتها.

ثانياً: النموذج اللوجستي ثنائي المعلم: (Two-Parameter Logistic Model) (2PLM) : يفترض هذا النموذج اختلاف الفقرات في صعوبتها وتمييزها، كما يفترض عدم تأثر الإجابات بعامل التخمين.

ثالثاً: النموذج اللوجستي ثلاثي المعلم: (Model Three Parameter Logistic) (3PLM) المستخدم لاغراض تحليل بيانات هذه الدراسة ويعد اكثر النماذج أحادية البعد تشدداً لأنه يسمح بأن تختلف مفردات الاختبار في صعوبتها وتمييزها، وهو ما يلاحظ عادة في بناء الاختبارات، إذ من الصعب إيجاد مجموعة من المفردات التي تميز بدرجة واحدة من مستويات السمة، أو القدرة (One Parameter) التي يقيسها اختبار معين، وهذا الافتراض الذي اعتمده النموذج الأحادي المعلم كذلك يفترض تأثر الإجابات بعامل التخمين وهو المعلم الثالث الذي تميز به عن النموذج الثنائي للمعلمة (Logistic Model or Rasch Model) (Two- Parameter Logistic Model) أو (Lower Asymptote Line) الذي أضافه لورد، وأطلق عليه معلم الخط التقاربي الأدنى والحقيقة أن إضافة هذا المعلم الثالث يكون بمثابة معلم التخمين (Guessing Parameter). (عبابنة، 2008:115)

المحور الثاني: الاختبارات التحصيلية:

نستطيع أن نقول أن الاختبار هو وسيلة من الوسائل التقويمية الموجودة التي من خلالها يتم عملية جمع المعلومات والبيانات، تشمل فيها الامتحان لقياس الشيء ولمعرفة قدرات الطلبة حتى يستطيع المعلم أخذ القرار فيما يتعلق بنجاح الطلبة وانتقالهم إلى المرحلة التالية، وأيضاً يستطيع المعلم معرفة نجاح تعليمه وفق الأهداف المرجوة في المنهج المقرر.

أهداف الاختبار:

للاختبار أهداف كثيرة، فأهداف الاختبار عند لي جرونيج Lee Cronbach وعلى ضوء المفهوم المعاصر للاختبارات سواء أكانت نصف فصلية أو فصلية يمكن تحقيق عدد من الأهداف نجملها في التالي:

١ - قياس مستوى تحصيل الطلاب العلمي ، وتحديد نقاط القوة والضعف لديهم.

التحقق من الخصائص السيكومترية في اختبار اختيار من متعدد باستخدام حذف المشتت الخاطئ وفقاً لنظرية السمات الكامنة

م.م علاء حسين عبيس الطائي

٢ -تصنيف الطلاب في مجموعات ، وقياس مستوى تقدمهم في المادة.

٣ -التنبؤ بأدائهم في المستقبل.

٤ -الكشف عن الفروق الفردية بين الطلاب سواء المتفوقون منهم ، أم العاديون أم

بطيئو التعليم.

٥ -تنشيط واقعية التعليم ، ونقل الطلاب من صف إلى آخر ، وفتح الدرجات

والشهادات.

٦ -التعرف على مجالات التطوير للمناهج والبرامج والمقررات الدراسية. (خولي، 1998:155)

أنواع الاختبارات:

للاختبارات أنواع عديدة كل منها يناسب أهدافا مختلفه عن النوع الآخر ، والاختبارات الموضوعية هي احد انواع الاختبارات والتي سيتطرق لها الباحث في هذه الدراسة كونها اداة البحث الحالي:

الاختبار الموضوعي: (Objective Test)

تظهر الاختبارات الموضوعية ثباتاً أكثر من الاختبارات المقالية أو غير الموضوعية، ويرجع ذلك إلى ذاتية تقدير الدرجات في الاختبارات المقالية. ولكن ليس المقصود من ذلك هو عدم إعطاء اختبارات مقال حتى نحصل على اختبارات ثابتة، فهذا من ناحية أخرى يؤثر على الصدق الذي يعتبر أهم صفة من صفات الاختبار الجيد. ولكن الحل الأمثل هو استخدام وسيلة القياس الأنسب لقياس مخرجات التعلم، وبعد ذلك نحاول زيادة موضوعية الاختبار، بإتباع مجموعة من القواعد سواء في الإعداد أو التصحيح. ويعرف بأنه اختبار يتطلب كل بند فيه إجابة تختلف عن البند الاخر و كذلك تكون الإجابة مرتبطة بالموضوع ولا مجال فيها للاجتهاد الشخصي، والإجابة واحدة محددة ولا يقبل سواها ولا يختلف عليها المصححون ويتميز هذا النوع بارتفاع مستوى الصدق والثبات والتميز بين الافراد. (الفوزان، 2011: 99)

ومن أمثلة الاختبار الموضوعي ما يلي:

- 1- اختبار المزوجة (Matching-Test): هو اختبار موضوعي تعرُفي يطلب فيه من خلال التعليمات من الطالب أن يزوج بين مفردات مجموعتين أو قائمتين من البنود.
- 2- اختبار الصواب والخطأ (True-False Test): هو الاختبار الذي يكون خطأ من حيث المحتوى او من حيث المضمون او اللغة اذا كان من ضمن اختبارات اللغة، ويطلب من الطالب أن يقرأ جملة ما ويدقق فيها ويحكم من خلال معرفته بها، ويمكن أن يطلب منه تصحيح الخطأ إن وجد.
- 3- اختبار الاختيار من متعدد (Multiple Choice- Test): في هذا النوع من الاختبارات تطلق كلمة بند على المشكلة المراد اختبارها والتي تمثل وحدة كاملة ومختلفة عن الاجابة والجزء الأول الذي يحتوى على المشكلة فقد اصطلح على تسميته بالساق، وقائمة الإجابة التي تليه اصطلح على تسميتها بالبدائل أو الخيارات، وفي أحيان قليلة تسمى الخيار الصحيح، وما تبقى بعد اختيار الإجابة الصحيحة يسمى المشتت أو المموه.
- 4- اختبار ملء الفراغ (Fill in- Test): وفيه يطلب من الطالب الاجابة بكلمة واحدة مناسبة او قد تكون محتوى قصير او جملة او وظيفة حسب ما مطلوب في السؤال.(خولي، 1998: 127)

7. التحليل الإحصائي لفقرات الاختبار التحصيلي على وفق نظرية الاستجابة .

أ- الصدق :

يعد من أهم الاعتبارات في تقويم الاختبارات التربوية والنفسية والاجتماعية ، ويعرف بأنه تقييم شامل يوفر من خلاله الدليل المادي والمبرر النظري اللازمين لإثبات كفاية وملائمة اي معنى أو فعل يبني على درجة الاختبار.

ولمعرفة قدرة الاختبار التحصيلي الذي أعد بأستعمال تقنية أستبعاد بديلين على التمييز بين المجموعتين العليا والدنيا اي على التمييز بين العينتين مرتفعة المستوى ومنخفضة المستوى، استخدم الباحث الاختبار التائي لعينتين مستقلتين لمعرفة دلالة الفرق بين متوسطي درجات الأفراد الذين حصلوا على أعلى الدرجات أي (27% العليا) والأفراد الذين حصلوا على أدنى الدرجات أي (27% الدنيا) إذ كانت نتائج الاختبار التائي كما هي موضحة في الجدول (1).

جدول رقم (1)

قيم الإحصاءات الوصفية لتوزيع معاملات الارتباط لأرتباط درجة الفقرة بالدرجة الكلية

الاختبار التحصيلي	عدد الفقرات	المدى اقل ارتباط اعلى ارتباط	المتوسط الحسابي	الانحراف المعياري

التحقق من الخصائص السيكومترية في اختبار اختيار من متعدد باستخدام حذف المشتت الخاطئ وفقا لنظرية السمات الكامنة

م.م علاء حسين عبيس الطائي

0,067	0,419	0,270	89	مادة علم نفس الخواص بأستعمال تقنية أستبعاد البديلين
0,070	0,348	0,334	89	مادة علم نفس الخواص بدون أستعمال تقنية أستبعاد البديلين

يتبين من الإحصاءات الوصفية لتوزيع معاملات الارتباط أن توزيع مؤشرات التمييز لجميع فقرات اختبار مادة علم نفس الخواص بأستعمال تقنية أستبعاد البديلين وبدون أستعمال تقنية أستبعاد البديلين متباينة الى حد ما في قوتها التمييزية ، وهي بذلك تستدعي أستجابات متباينة من الأفراد، وهذا يتسق مع الانموذج الأحادي المعتمد وبالتالي استخدام الأنموذج في تحليل بيانات البحث الحالي.

جدول رقم (2)

نسبة الأفراد منخفضي القدرة وإجاباتهم الصحيحة على الفقرات الصعبة في الاختبار

بدون أستعمال تقنية أستبعاد بديلين			ت الفقرة	أستعمال تقنية أستبعاد بديلين			ت الفقرة
نسبة الأفراد الذين أجابوا إجابة صحيحة	عدد الأفراد الذين أجابوا إجابة صحيحة	صعوبة الفقرة		نسبة الأفراد الذين أجابوا إجابة صحيحة	عدد الأفراد الذين أجابوا إجابة صحيحة	صعوبة الفقرة	
0,05	2	0,121	8	0,08	3	0,224	8
0,03	1	0,216	15	0,05	2	0,312	15
0,08	3	0,166	23	0,08	3	0,271	23
0,05	2	0,114	31	0,05	2	0,221	31
0,08	3	0,234	48	0,05	2	0,326	48
0,08	3	0,189	52	0,03	1	0,314	52
0,05	2	0,109	61	0,03	1	0,451	61
0,00	0	0,254	71	0,00	0	0,325	71
0,03	1	0,218	79	0,10	4	0,367	79
0,02	3	0,162	84	0,05	2	0,325	84

ويتبين من الجدول (2) النسب المنخفضة جدا لأجابات الأفراد الصحيحة من ذوي القدرات المنخفضة على الفقرات الصعبة جدا ، وعند بمقارنتها مع النسبة النظرية للتخمين العشوائي (0,25)، نستخلص الى

نتيجة مفادها خلو الأختبار التحصيلي لمادة علم نفس الخواص بأستعمال تقنية أستبعاد بديلين وبدون أستعمال تقنية أستبعاد بديلين تقريبا من عامل التخمين وبذلك تصبح البيانات ملائمة للتحليل وفقا للنموذج الأحادي .

جدول رقم (3)

نتائج الاختبار التائي لدلالة الفرق بين متوسطي درجات الأختبار التحصيلي بتقنية أستبعاد بديلين المجموعتين العليا والدنيا

مستوى الدلالة عند مستوى 0.001	القيمة التائية		الأنحراف المعياري	المتوسط الحسابي	المجموعة	الأختبار التحصيلي
	الجدولية	المحسوبة				
دالة	3,373	33,640	2,981	62,407	العليا 27%	بأستعمال تقنية أستبعاد بديلين
			2,625	44,222	الدنيا 27%	

*القيمة التائية الجدولية عند مستوى دلالة (0.05) ودرجة حرية (106) تساوي 1.980

** القيمة التائية الجدولية عند مستوى دلالة (0.01) ودرجة حرية (106) تساوي 2.617

*** القيمة التائية الجدولية عند مستوى دلالة (0.001) ودرجة حرية (106) تساوي 3.373

ويتضح من الجدول (3) أن القيمة التائية المحسوبة أكبر من القيم التائية الجدولية عند مستوى دلالة (0.001) وبدرجة حرية (106)، مما يشير ذلك إلى وجود فروق ذات دلالة إحصائية بين متوسطي المجموعتين العليا والدنيا، مما يعني قدرة الاختبار التحصيلي لمادة علم نفس الخواص بأستعمال تقنية أستبعاد بديلين اي على التمييز بين المجموعتين العليا والدنيا ،وبذلك فإن الاختبار التحصيلي يتميز بصدق وظيفي .

ولمعرفة قدرة الاختبار التحصيلي الذي أعد بدون أستعمال تقنية أستبعاد بديلين على التمييز بين المجموعتين العليا والدنيا على التمييز بين العينتين مرتفعة المستوى ومنخفضة المستوى استخدم الباحث الاختبار التائي لعينتين مستقلتين لمعرفة دلالة الفرق بين متوسطي درجات العينتين (مرتفعة المستوى ومنخفضة المستوى) إذ كانت نتائج الاختبار التائي كما هي موضحة في الجدول (4).

الجدول (4)

نتائج الاختبار التائي لدلالة الفرق بين متوسطي درجات الأختبار التحصيلي بدون تقنية أستبعاد بديلين المجموعتين العليا والدنيا

مستوى الدلالة عند مستوى 0.001	القيمة التائية		الأنحراف المعياري	المتوسط الحسابي	المجموعة	الأختبار التحصيلي
	الجدولية	المحسوبة				
	3,373	25,801	3,346	6,574	العليا 27%	بدون أستعمال تقنية

التحقق من الخصائص السيكومترية في اختبار اختيار من متعدد باستخدام حذف المشتت الخاطئ وفقاً لنظرية السمات الكامنة

م.م علاء حسين عبيس الطائي

دالة		4,173	44,796	الدنيا 27%	أستبعاد بديلين
------	--	-------	--------	------------	----------------

*القيمة التائية الجدولية عند مستوى دلالة (0.05) ودرجة حرية (106) تساوي 1.980

** القيمة التائية الجدولية عند مستوى دلالة (0.01) ودرجة حرية (106) تساوي 2.617

*** القيمة التائية الجدولية عند مستوى دلالة (0.001) ودرجة حرية (106) تساوي 3.373

ويتضح من الجدول (4) أن هناك فروقاً ذات دلالة إحصائية بين متوسطي درجات العينتين مرتفعة المستوى ومنخفضة المستوى . إذ كانت القيمة التائية المحسوبة البالغة (3.373) وهي أكبر من القيمة التائية الجدولية البالغة (3.373) عند مستوى (0.001)، مما يعني ذلك ، أن الاختبار التحصيلي لمادة علم نفس الخواص بدون أستعمال تقنية أستبعاد بديلين له القدرة على التمييز بين العينتين مرتفعة المستوى ومنخفضة المستوى او أن الاختبار يتميز بصدق وظيفي.

وبعد العرض السابق لصدق الاختبار التحصيلي لمادة علم نفس الخواص بأستعمال تقنية أستبعاد بديلين والأختبار التحصيلي نفسه بدون أستعمال تقنية أستبعاد بديلين ومقارنتهما، تبين انه لم يكن فرقا واضحة بينهما ، حيث ان الأختبار التحصيلي بأستعمال تقنية أستبعاد بديلين أو بدون أستعمال تقنية أستبعاد بديلين يتمتعان بصدق وظيفي عالي وبمستوى دلالة (0.01) ، فالإجراءات المتبعة بين ان صورة الأختبار بأستعمال تقنية أستبعاد بديلين يتمتع بقدرة على التمييز بين الأفراد الذين يمتلكون مستوى عالي من القدرة والذين يمتلكون مستوى متدني من القدرة ، وكان الاختبار التحصيلي بدون أستعمال تقنية أستبعاد بديلين الذي تمتع بقدرة على التمييز أعلى بين العينة مرتفعة المستوى والعينة منخفضة المستوى وبذلك فكلا الصورتين يمكن أستعمالها لأخذ مستوى دقة أعلى في القرار.

ب- الثبات :

يختص بمدى الوثوق بالدرجات التي نحصل عليها من تطبيق الاختبار بمعنى أن هذه الدرجات أو النتائج يجب ألا تتأثر بالعوامل التي تعود إلى أخطاء الصدفة ، فهو يعني دقة الاختبار أو اتساقه ، فإذا حصل نفس

الفرد على نفس الدرجة أو ما يقرب منها في نفس الاختبار أو في مجموعات من الأسئلة المتكافئة أو المتماثلة في مناسبات مختلفة فإننا نصف الاختبار أو المقياس في هذه الحالة بأنه على درجة عالية من الثبات • وبهذا المعنى يرتبط مفهوم الثبات بما يسمى إحصائياً أخطاء القياس المتضمن في كل درجة من درجات الاختبار •

الجدول (5)

القيمة الفائية لدلالة الفرق بين معاملي الثبات الذي يشير للثبات المحسوب بكلا الأسلوبين (القياس معياري المرجع والقياس محكي المرجع)

مستوى الدلالة 0.05	القيمة الفائية		معامل الثبات	الأختبار التحصيلي
	الدرجة	المحسوبة		
غير دالة	3,850	1,195	0,843	بأستعمال تقنية أستبعاد بديلين
			0,922	بأستعمال الطريقة الاعتيادية التقليدية

ويتضح من الجدول (5) لم يظهر فرقا ذو دلالة إحصائية بين معاملي الثبات للأختبار التحصيلي لمادة علم نفس الخواص بأستعمال تقنية أستبعاد بديلين وبدون أستعمال تقنية أستبعاد بديلين عند مستوى دلالة إحصائية (0.05)، لأن القيمة الفائية المحسوبة للفرق بين معاملي الثبات أصغر من القيمة الفائية الجدولية البالغة (3.850)، وقد يعود عدم الفرق بين الأختبار بأستعمال التقنية وبدون أستعمالها في معامل الثبات إلى اختيار الفقرات ذات المؤشرات الدقيقة في ملاءمتها للنموذج المعتمد . ويعزو الباحث سبب ذلك إلى الدور المهم للقياس وفقا للنظرية الحديثة وما انبثق عنها من نماذج نظرية ، إذ تهتم بتقدير مستوى القدرة للأفراد من خلال فقرات اختبارية تتناسب في صعوبتها مع مستوى قدرة هؤلاء الأفراد، وعليه يضمن الباحث التفاعل بين قدرة الفرد وصعوبة الفقرة، ومن ثم ينقص معلم التخمين الذي يظهر من خلال الأداء على أسئلة الاختيار من متعدد ، إذ يلجأ الطلاب ذو القدرة المنخفضة إلى التخمين عن بعض فقرات الاختبار وبخاصة الفقرات ذات مستوى الصعوبة المرتفع إلى حد ما عن قدرتهم ، مما يزيد من متوسط الأخطاء المعيارية لقيمتي معامل الثبات ، وقد استخدم الباحث هذه الطريقة للمقارنة بوصفها المعتمد عليها في معظم الدراسات التي قارنت بالثبات.

جدول رقم (6)

قيم معاملات ارتباط درجة الفقرة بالدرجة الكلية لفقرات الأختبار التحصيلي بأستعمال تقنية أبعاد البديلين

ت الفقرة	معامل الصدق	ت الفقرة	معامل الصدق	ت الفقرة	معامل الصدق
----------	-------------	----------	-------------	----------	-------------

التحقق من الخصائص السيكومترية في اختبار اختيار من متعدد باستخدام حذف المشتت الخاطئ وفقاً لنظرية السمات الكامنة

م.م علاء حسين عبيس الطائي

0,497	61	0,426	31	0,422	1
0,481	62	0,429	32	0,451	2
0,475	63	0,507	33	0,272	3
0,491	64	0,301	34	0,338	4
0,494	65	0,371	35	0,399	5
0,496	66	0,360	36	0,317	6
0,479	67	0,387	37	0,407	7
0,481	68	0,301	38	0,382	8
0,494	69	0,336	39	0,413	9
0,496	70	0,389	40	0,472	10
0,481	71	0,345	41	0,383	11
0,489	72	0,411	42	0,399	12
0,500	73	0,365	43	0,333	13
0,468	74	0,447	44	0,473	14
0,485	75	0,382	45	0,378	15
0,477	76	0,339	46	0,360	16
0,356	77	0,345	47	0,326	17
0,326	78	0,332	48	0,332	18
0,468	79	0,491	49	0,487	19
0,496	80	0,490	50	0,310	20
0,481	81	0,484	51	0,455	21
0,326	82	0,478	52	0,447	22
0,441	83	0,495	53	0,403	23
0,365	84	0,485	54	0,448	24
0,419	85	0,498	55	0,354	25
0,531	86	0,475	56	0,366	26
0,412	87	0,495	57	0,347	27
0,332	88	0,488	58	0,441	28
0,356	89	0,415	59	0,392	29
		0,546	60	0,333	30

*القيمة الحرجة لمعامل الارتباط عند مستوى دلالة (0,05) ودرجة حرية (198) = (0,117)

يتضح من الجدول (6) أن القيم المحسوبة لمعاملات ارتباط درجة الفقرة بالدرجة الكلية كانت أكبر من القيمة الحرجة لمعامل الارتباط ، ويدل ذلك على التجانس الداخلي ل فقرات اختبار مادة علم نفس الخواص باستعمال تقنية أبعاد البديلين في قياس السمة الكامنة المقاسة ، وهو مؤشر إضافي على تحقيق افتراض أحادية البعد.

كما قام الباحث بحساب معامل ارتباط درجة كل فقرة بالدرجة الكلية للأختبار التحصيلي باستعمال الاسلوب التقليدي وكانت النتائج كما مبينة في الجدول(7).

الجدول (7)

قيم معاملات ارتباط درجة الفقرة بالدرجة الكلية لفقرات الأختبار التحصيلي باستعمال الاسلوب الاعتيادي (التقليدي)

معامل الصدق	ت الفقرة	معامل الصدق	ت الفقرة	معامل الصدق	ت الفقرة
0,441	61	0,355	31	0,336	1
0,329	62	0,456	32	0,404	2
0,336	63	0,309	33	0,247	3
0,412	64	0,400	34	0,305	4
0,396	65	0,441	35	0,226	5
0,412	66	0,392	36	0,251	6
0,325	67	0,207	37	0,396	7
0,412	68	0,337	38	0,331	8
0,333	69	0,375	39	0,466	9
0,329	70	0,345	40	0,313	10
0,411	71	0,237	41	0,270	11
0,315	72	0,307	42	0,284	12
0,322	73	0,373	43	0,258	13
0,364	74	0,375	44	0,270	14
0,541	75	0,345	45	0,372	15
0,511	76	0,237	46	0,269	16
0,418	77	0,339	47	0,327	17
0,339	78	0,217	48	0,301	18
0,374	79	0,368	49	0,232	19
0,360	80	0,496	50	0,265	20
0,229	81	0,307	51	0,276	21
0,364	82	0,323	52	0,349	22
0,521	83	0,355	53	0,333	23
0,472	84	0,395	54	0,355	24

التحقق من الخصائص السيكومترية في اختبار اختيار من متعدد باستخدام حذف المشتت الخاطئ وفقا لنظرية السمات الكامنة

م.م علاء حسين عبيس الطائي

0,364	85	0,352	55	0,369	25
0,423	86	0,437	56	0,326	26
0,299	87	0,356	57	0,365	27
0,362	88	0,388	58	0,284	28
0,357	89	0,410	59	0,241	29
		0,370	60	0,306	30

*القيمة الحرجة لمعامل الارتباط عند مستوى دلالة (0,05) ودرجة حرية (198) = (0,117)

يتضح من الجدول (7) أن القيم المحسوبة لمعاملات ارتباط درجة الفقرة بالدرجة الكلية كانت أكبر من القيمة الحرجة لمعامل الارتباط ، ويدل ذلك على التجانس الداخلي ل فقرات اختبار مادة علم نفس الخواص باستعمال الاسلوب التقليدي في قياس السمة الكامنة المقاسة ، وهو مؤشر إضافي على تحقيق أفترض أحادية البعد.

الاستنتاجات :

في ضوء النتائج التي توصل إليها البحث الحالي ، يمكن للباحث أن يستنتج الآتي :

1. وجود اتفاق بين الأسلوبين في القدرة على التمييز بين المجموعات المتطرفة.
2. دقة الإجراءات المتبعة لأعداد الاختبار التحصيلي سواء بأستعمال تقنية أستبعاد بديلين أو بدون أستعمالها ، وذلك لما يمتاز به من انخفاض في قيمة الخطأ المعياري وارتفاع في معامل الثبات .
3. ليس هناك تباين في المعايير المستخدمة لحذف الفقرات الضعيفة في بالتحليل المعتمد للاختبار التحصيلي بأستعمال تقنية أستبعاد البديلين أو بدونها.

التوصيات :

1. استخدام نماذج النظرية الحديثة في تحليل فقرات الاختبارات التحصيلية وبنائها
2. تطوير طرق وأنواع الاختبارات المقدمة الى الطلبة ومقارنتها بطرق واساليب حديثة للتأكد من خصائصها القياسية.
3. التأكيد على استعمال الاختبارات الموضوعية المبنية بشكل صحيح ومحبوكة، لاستخدامها في قياس المهارات العليا للطلبة.

15. المصادر العربية:

- 1) أبو حطب, فؤاد, وعثمان, سيد أحمد, وصادق, أمال (2008): التقويم النفسي. ط4, القاهرة, مكتبة الأنجلو المصرية.
- 2) التقي, احمد محمد (2013): النظرية الحديثة في القياس, عمان, ط2, دار المسيرة للنشر والتوزيع.
- 3) التقي, احمد محمد (2013). النظرية الحديثة في القياس, ط2, عمان, دار المسيرة للنشر والتوزيع.
- 4) الزيات, السيد محمد (1997): الإحصاء النفسي, القاهرة, دار الفكر العربي.
- 5) الدليمي, خالد جمال جاسم (2009): بناء بنك اسئلة في مادة القياس والتقويم التربوي والنفسي, اطروحة دكتوراه غير منشورة, جامعة بغداد, كلية التربية - ابن رشد.
- 6) زكري, علي بن محمد عبد الله (2009). الخصائص السيكمترية لاختبار أوتيس - لينون للقدرة العقلية مقدرة وفق القياس الكلاسيكي ونموذج راش لدى طلبة المرحلة المتوسطة بمحافظة صبيا التعليمية, أطروحة دكتوراه غير منشورة, جامعة أم القرى, كلية التربية.
- 7) سليمان, امين محمد علوان, وابو علام, رجاء محمود (2012): القياس والتقويم في العلوم الانسانية: ط2, دار الكتاب الحديث.
- 8) الشافعي, محمد منصور محمد, ونور الدين, أمين محمد صبري (2007): استخدام نموذج التقدير الجزئي الرياضي اللوغاريتمي في تطوير اختبارات القدرات العقلية الأولية على عينة من البيئة السعودية, مجلة كلية التربية, جامعة الزقازيق, العدد56.
- 9) الشرقاوي, أنور محمد والشيخ, سليمان الخضري, و كاظم, أمينة محمد وعبد السلام, نادية محمد. (1996). اتجاهات معاصرة في القياس والتقويم النفسي والتربوي, القاهرة, مكتبة الانجلو المصرية.

التحقق من الخصائص السيكومترية في اختبار اختيار من متعدد باستخدام حذف المشتت الخاطئ وفقا لنظرية السمات الكامنة

م.م علاء حسين عبيس الطائي

- (10) عبابنة, عماد (2006): التحقق الإمبريقي من معادلات ستوكنج في تحديد مستويات القدرة المناظرة لأقصى معلومات لتقدير معالم الفقرات في نظرية الاستجابة للفقرة, المجلة الأردنية في العلوم التربوية, المجلد 2, العدد 2.
- (11) عبابنة, عماد خصاب (2009). الاختبارات محكية المرجع. ط1, عمان, دار المسيرة للنشر والتوزيع.
- (12) عبد الخالق ، أحمد محمد . (1983) . الأبعاد الأساسية للشخصية ، الإسكندرية ، دار المعرفة الجامعية .
- (13) عبد الرحمن, سعد (1998): القياس النفسي - النظرية والتطبيق, ط3, القاهرة, دار الفكر العربي.
- (14) عبد السلام ، نادية محمد (1992 أ) ، مشكلات عند بناء الاختبارات محكية المرجع – تحليل وتقويم ، مجلة علم النفس ، الهيئة المصرية العامة للكتاب ، العدد (23) .عبيدات، ذوقان وآخرين (2000): البحث العلمي مفهومه وأدواته وأساليبه، ط ٣، الرياض, دار أسامه للنشر والتوزيع.
- (15) العنبي , سالم حميد معيض (2019): التنبؤ بالتحول العقلي في ضوء ما وراء التعلم وقوة السيطرة المعرفية لدى طلبة الجامعة, اطروحة دكتوراه غير منشورة, كلية التربية, جامعة أم القرى.
- (16) علام, صلاح الدين محمود (2006): الاختبارات والمقاييس التربوية والنفسية, ط 1, دار الفكر العربي.
- (17) عيسوي ، عبد الرحمن محمد . (1983) . القياس والتجريب في علم النفس والتربية ، بيروت ، دار النهضة العربية .
- (18) فاندالين، ديوبورت. ب (1988). مناهج البحث في التربية وعلم النفس، ترجمة محمد نبيل نوفل، القاهرة، مكتبة الانجلو المصرية.

- (19) فخرو، حصة عبد الرحمن، وعبد الرحيم، انور رياض، ومحمد، محمد إبراهيم محمد (2009): فاعلية استخدام نموذج سلم التقدير في تحليل مفردات مقياس مداخل الدراسة لدى طلبة جامعتي قطر والمينا، مجلة البحث في التربية وعلم النفس كلية التربية، جامعة المينا، المجلد 22، العدد 2.
- (20) الفيل، حلمي (٢٠١٨). برنامج مقترح للتوظيف أنموذج التعلم القائم على السيناريو (SBL) في التدريس وتأثيره في تنمية مستويات عمق المعرفة وخفض التجول العقلي لدى طلاب كلية التربية النوعية جامعة الإسكندرية، مجلة كلية التربية، جامعة المنوفية، العدد الثاني.
- (21) كاظم ، أمينة محمد . (1988) . دراسة نظرية نقدية حول القياس الموضوعي للسلوك " نموذج راش " ، الكويت ، مؤسسة الكويت للتقدم العلمي .
- (22) كاظم ، أمينة محمد . (1988) . دراسة نظرية نقدية حول القياس الموضوعي للسلوك " نموذج راش " ، الكويت ، مؤسسة الكويت للتقدم العلمي .
- (23) الكبيسي ، كامل ثامر (1987) : بناء وتقنين مقياس السمات الشخصية ذات الاولوية للقبول في الكليات العسكرية لدى طلبة الصف السادس الاعدادي في العراق، اطروحة دكتوراه غير منشورة، جامعة بغداد، كلية التربية، ابن رشد .
- (24) كروكر، ليندا والجينا، جيمز (2009): مدخل إلى نظرية القياس التقليدية والمعاصرة. ترجمة: زينات يوسف دعنا، ط 1، دار الفكر .
16. المصادر الاجنبية:

- 25) Baker, Frank, B. (2001): The Basics of item responses Theory, 2ed, ERIC Clearinghouse on Assessment and Evaluation.
- 26) Brown , T.A. (2006) , Confirmatory Factor Analysis for Applied Research New York , Guilford Press.
- 27) Gregory , R. (2004) . Psychological Testing : History Principles applications . 4th ed, New York, Pearson Education group Inc.
- 28) Hambleton & Jones, R .W. (1993): Introduction model on Comparison of classical test theory and item response theory And

التحقق من الخصائص السيكومترية في اختبار اختيار من متعدد باستخدام حذف المشتت الخاطئ وفقا لنظرية السمات الكامنة

م.م علاء حسين عبيس الطائي

-
- their application to test development, Educational Measurement: Issues and Practice, vol, 12 No. 3, 38–47.**
- 29) Hambleton ,R.K. (1978):**On the Use of Cut-off Scores with Criterion Referenced Tests in Instructional Setting** ,Journal of Educational Measurement No.4,p.277-292.
- 30) Hickey, J. (2013): **Mind wandering, Meoryand mood**, master thesis, university.
- 31) Kane, M., Brown, L., Mcvay, J., Silivia, P. Myin – Germeys, J.& Mcvay, J.Silivia, P. Myin–Germeys, J.& kwapil, T. (2007): **for whom the mind wanders, and when: A experience–sampling tudy of working memory and excutive control in daily life**, psychological science 18 (7) 61– 621.
- 32) Kane, M., Brown, L., Mcvay, J., Silivia, P. Myin – Germeys, J.& Mcvay, J.Silivia, P. Myin–Germeys, J.& kwapil, T. (2012): **for whom the mind wanders, and when : A experience–sampling tudy of working memory and excutive control in daily life**, psychological science 18 (7) 61– 621.
- 33) Lustina, Michael J. (2004): **A Comparison of Andrich s Rating Scale Model and Rosts Successive Intervals Model**, Ph.D Thesis, The University of Texas at Austin.
- 34) Masters, G. (1982): **A Rasch model for partial credit scoring**. Psychometrika, 47, 149–174.

- 35) Masters, Geoff N. (2010): **Partial Credit Model, Ch, 5, in Nering, M. L., & Ostini, M. (EDT).** Handbook Polytomous Item Response Theory Models.
- 36) Ostini, R., & Nering, M. L. (2006): **Polytomous Item response theory models.** Sage University paper Series QASS.
- 37) Randall, J. (2015): **Mind Wandering and Self – Directed Learning : Testing the Efficacy of self – Regulation Interventions to Reduce Mind wandering and Enhance online Training Performance,** PhD Dissertation, Rice University.
- 38) Shepard, Lorrie A (1984) .Setting Performance Standards , In Berk (Ed): **A guide to criterion – referenced Test Construction.** Hopkins University press.
- 39) Smallwood, J., and Schooler, J. W. (2006): **“The Restless Mind,”.** Psychological Bulletin (132:6), pp. 946– 958.
- 40) Smallwood, J., D'Connor, R. (2011): **Imprisoned by the past.** Unhappy moods lead to retrospective bias to mind wandering. Cognition and Emotion, 25 (8), pp 1481 – 1490.
- 41) Smallwood, J.Schooler, J. (2015): **the science wind wandering.** Empirically navigating the stream of consciousness. Annual Review of Psychology, 66, pp 487 – 518.
- 42) Stenner , A. S. (1990) . **Specification of the Universe of Generalization under construct definition Theory** , D. A. , Vol. (45) , No. (8) , PP. 502-503.
- 43) Stenner , A. S. (1990) . **Specification of the Universe of Generalization under construct definition Theory** , D. A. , Vol. (45) , No. (8) , PP. 502-503.
- 44) Traub, E, R (1997): **Classical Test Theory in**