



كلية التربية

كلية معتمدة من الهيئة القومية لضمان جودة التعليم

إدارة: البحوث والنشر العلمي (المجلة العلمية)

=====

**الأخطاء الشائعة في بناء الاختبارات الموضوعية وتأثيرها على
الخصائص السيكومترية للاختبارات التحصيلية لدى طلاب
كلية التربية بجامعة أسيوط**

إعداد

د/ علي صلاح عبد المحسن حسن

مدرس علم النفس التربوي بقسم علم النفس التربوي

كلية التربية - جامعة أسيوط

« المجلد الرابع والثلاثون - العدد الثاني - فبراير ٢٠١٨ م »

http://www.aun.edu.eg/faculty_education/arabic

أولاً - مقدمة البحث:

تعد الاختبارات من أهم الأدوات في المجال التربوي حيث تزودنا ببيانات يُستند إليها في اتخاذ العديد من القرارات المهمة التي تخص الفرد والمجتمع، فالاختبار بأبسط صوره هو إجراء منظم لقياس عينة من سلوك الأفراد، وقد انتشر استخدام الاختبارات لاختيار شخص أو عدة أشخاص لشغل وظيفة ما من بين العديد من المتقدمين لشغل هذه الوظيفة، كما تستخدم الاختبارات لأغراض التصنيف كتحديد مسارات الطلاب وتخصصاتهم الدراسية التي ينخرطون فيها بما يتناسب مع قدراتهم ومهاراتهم، وكذلك يتم استخدامها لتقويم أداء الطلاب وتحصيلهم الدراسي من خلال ما يحصلون عليه من درجات فيما يقدم إليهم من هذه الاختبارات سواء أكانت صافية أم غير صافية (نضال الشريفين، إيمان طعمانة، ٢٠٠٩، ٣٠٩).

وتعد الاختبارات التحصيلية من أهم الوسائل التي يستخدمها أعضاء هيئة التدريس بالجامعات، وذلك لمعرفة مدى التقدم في تحقيق الأهداف داخل غرفة الصف، والاختبارات المستخدمة في الكليات لتقييم الطلاب هي على الأرجح اختبارات من إعداد أعضاء هيئة التدريس أنفسهم دون الرجوع إلى متخصصي القياس والتقويم (بندر الأمير، ٢٠١١، ٢).

ولكي تحقق الاختبارات - بمختلف أنواعها- أهداف القياس النفسي والتقويم التربوي يجب أن تتمتع بخصائص ومواصفات معينة، منها: الموضوعية وسهولة التصحيح والتفسير، إضافة إلى تمتع هذه الاختبارات بالخصائص السيكومترية الجيدة من صدق وثبات.

وهناك العديد من الأشكال للاختبارات لعلها تنقسم إلى نوعين رئيسيين هما: تلك الاختبارات التي تستلزم إنتاج الاستجابة وتُعرف بالاختبارات المقالية، وتلك التي تستلزم إنتقاء الاستجابة وتُعرف بالاختبارات الموضوعية. وتتعدد أشكال الاختبارات الموضوعية فمنها فقرات الاكمال، والصواب والخطأ، والمزوجة، والاختيار من متعدد، وتُعد الاختبارات الموضوعية من أكثر أنواع الاختبارات استخداماً وانتشاراً في الفترة الأخيرة؛ وذلك للعديد من مميزاتا عن الاختبارات المقالية (مايسة أبو مسلم، ٢٠١٦، ١٥٥).

ومن خلال عمل الباحث في مجال التدريس بالجامعة، لوحظ أن هناك كثيراً من الأخطاء الشائعة في بناء مفردات الاختبارات الموضوعية - بمختلف فروعها- مما دفع الباحث لإجراء هذا البحث لتحديد أهم الأخطاء الشائعة في بناء هذه الاختبارات والوقوف على الأثر الناتج لهذه الأخطاء على الخصائص السيكومترية للاختبارات.

ولكي يحقق البحث ذلك فسيستقصي الأدبيات التي كتبت في هذا المجال، وسيتم اختيار كلية التربية كمؤسسة رائدة ومتميزة في شتى المجالات، وتشمل هذه الأدبيات كل ما توافر عن الاختبارات الموضوعية، بالإضافة إلى عينات من الاختبارات الموضوعية التي توافرت من هذه المؤسسة لتكون قيد البحث والدراسة.

ثانياً - مشكلة البحث:

هناك الكثير من المشكلات التي تعانيتها نظم التعليم في الكثير من دول العالم، ومن أهم هذه المشكلات عملية بناء الامتحانات الفصلية، والامتحانات الملحقه للشهادات العامة بالمرحلتين الإعدادية والثانوية، إذ يقوم المتخصصون في إدارة المناهج ببناء الكثير من الامتحانات لنهاية كل فصل دراسي، ويتطلب ذلك وقتاً كبيراً، وجهداً أكبر، ويضطرون إلى بناء اختبارات ملحقه للطلبة الذين لم يتمكنوا من اجتياز الامتحانات الفصلية بنجاح، ولا يوجد دليل علمي على قدرة هذه الامتحانات على تمثيل المحتوى المطلوب لتأكيد جانب الصدق فيها، كما لا يمكن اعتبارها متكافئة فيما بينها بسبب عدم تحليلها للحصول على الخصائص السيكومترية الضرورية لعملية التكافؤ والمعادلة، كالوسط الحسابي، والانحراف المعياري، ومعامل الصعوبة، ومعامل التمييز، وكثيراً ما يشتكى الطلبة في كل عام دراسي من تفاوت الامتحانات من امتحان إلى آخر، وعدم العدل في قياس سمة التحصيل قياساً علمياً سليماً ومما يزيد من خطورة المشكلة أن نتائج الطلبة في الامتحانات العامة تترتب عليها قرارات متعلقة بالانتقال إلى مراحل دراسية أعلى، أو الالتحاق بالجامعات، أو بسوق العمل (راشد الدوسري، ٢٠٠١، ١٠).

وتستخدم درجات الاختبارات غالباً كإحدى المعلومات في صنع القرارات المهمة، وتركز بعض هذه القرارات على المستوى الفردي ومثال على هذا عندما يقرر التلميذ بأية كلية سوف يلتحق أو بأي برنامج سوف يسجل اسمه، وبعض القرارات الأخرى يكون التركيز أكثر على المستوى المؤسسي، ومن أمثلة ذلك عندما تحتاج المؤسسة أو الوكالة أن تقرر أية درجة على الاختبار تكون مطلوبة للأفراد للالتحاق بمهنة أو للسماح لهم بدخول الكلية أو الجامعة أو الخدمة العسكرية، وهناك بعض القرارات تكون مطلوبة على مستوى السياسة العامة مثل تحسين التعليم (Improve education) أو كيفية تقويم الممارسات التربوية عندما يتم تعبيرها (Kolen, M. & Brennan, R., 2004, 1).

وقد أشارت العديد من الدراسات إلى أن هناك قصوراً لدى بعض المعلمين عند بناء الاختبارات بصفة عامة والاختبارات الموضوعية بصفة خاصة، ومن هذه الدراسات دراسة عماد رمضان (٢٠٠٧) والتي استهدفت المقارنة بين الخصائص السيكومترية للاختبارات الناتجة من تطبيق طرق مختلفة لاختيار الفقرات في ضوء النظرية الكلاسيكية ونظرية الاستجابة للمفردة، وقد أظهرت النتائج أن أفضل طريقتين يمكن لوضع الاختبار استخدامهما هما: طريقة الاختيار المتذبذب، والاختيار حسب مؤشرات صدق وثبات الفقرة. ودراسا محمد عبد الخالق(٢٠٠٧، و٢٠١١) والتي أجريتا على برامج تعليم العربية للناطقين بلغات أخرى، وتوصلتا إلى أن معظم أنواع الاختبارات الموضوعية التي كانت قيد الدراسة لا تحقق النواتج التعليمية المفروضة كما أن هذه الاختبارات تقتصر على قياس الأهداف المعرفية في أدنى سلم المجال المعرفي. واستهدفت دراسة عبد الله الغامدي (٢٠٠٨) الكشف عن أثر اختلاف عدد البدائل في أسئلة الاختيار من متعدد على الخصائص السيكومترية للاختبار وفقراته وتوصلت الدراسة إلى عدم وجود فروق دالة بين معاملات التمييز والصدق والثبات للنماذج الاختبارية الثلاثة. وأجرى ماجد بخيت(٢٠١١) دراسة للكشف عن أثر عدد المموهات الجذابة لفقرات الاختيار من متعدد في الخصائص السيكومترية للاختبار وفقراته، وبينت الدراسة وجود فروق دالة إحصائياً في معاملات التمييز وصعوبة الفقرات تعزى لعدد المموهات الجذابة لفقراتها. أما دراسة سناء أحمد(٢٠١١) فقد هدفت إلى تحليل أسئلة الاختبارات الموضوعية بمرحلة البكالوريوس بكلية التربية لإعداد معلمات الابتدائي بالرياض، وكان من أبرز نتائجها تواضع مستوى تمثيل الاختبارات الموضوعية للأهداف السلوكية واقتارها إلى الشمولية والدقة. بينما كشفت دراسة بسام محمود (٢٠١١) عن وجود فروق دالة إحصائياً بين متوسطات معاملات التمييز والثبات لصالح الاختبار ذي الأربعة بدائل أولاً ثم الاختبار ذي الخمسة بدائل ثانياً. وأجرى بندر الأمير(٢٠١١) دراسته على معلمي الرياضيات بالمرحلة المتوسطة بمدينة مكة المكرمة وتوصلت الدراسة إلى أن نماذج الاختبارات التي يضعها المعلمون تفتقر إلى الموضوعية وعدم الدقة ويشوبها الأخطاء ومنها عدم وضوح مدخل الفقرة، عدم ترتيب البدائل من الأعلى إلى الأسفل، وعدم توزيع الاجابة الصحيحة توزيعاً عشوائياً. وقد قام حيدر ظاظا(٢٠١٢) بدراسة هدفت إلى الكشف عن مدى مخالفة قواعد صياغة فقرة الاختيار من متعدد في امتحان شهادة الثانوية العامة بالأردن، وتوصلت الدراسة إلى أن نسبة ١٤.٣ % من الفقرات تضمنت مخالفة واحدة، و ١٢.٦ % تضمنت أكثر من مخالفة أي إجمالاً ٢٦.٩ % فقرة مخالفة للقواعد. وفي نفس السياق أجرى كل من إبراهيم محمد وباسل خميس (٢٠١٢) دراسة هدفت إلى معرفة أثر مخالفة قواعد صياغة فقرات الاختيار من متعدد في الخصائص السيكومترية للاختبار وتوصلت الدراسة إلى دلالة الفروق بين نموذجي الاختبار

لصالح الاختبار الملتزم بقواعد صياغة الفقرة في كل من الصدق والثبات والاتساق الداخلي. أما دراسة رانية بنت بكر (٢٠١٢) فقد هدفت إلى تحديد الأخطاء الشائعة لمعلمات علم النفس في بناء مفردات الصواب والخطأ، وتحديد العلاقة بين الاختبار الجيد والاختبار غير الجيد في الخصائص السيكومترية للاختبار و فقراته وتوصلت الدراسة إلى وجود أربعة أخطاء شائعة في بناء مفردات الصواب والخطأ هم خطأ خاص ببناء مفردات أسئلة الصواب والخطأ وهو نقل العبارة نصاً من الكتاب، وثلاثة أخطاء مرتبطة بإخراج أسئلة الصواب والخطأ هي عبارات الصواب والخطأ غير متقاربين في الطول، التوزيع العشوائي للعبارات، وأن تكون نسبة عبارات الصواب أقل من ٤٠%. وفي نفس السياق أجرت سجي أحمد (٢٠١٣) دراسة عن اختلاف نمط الاختبار وأثر ذلك على الخصائص السيكومترية للاختبار وتوصلت الدراسة إلى وجود فروق دالة إحصائياً بين معاملات الصدق والثبات للأنماط الثلاثة (الإكمال، الاختيار من متعدد، الصواب والخطأ) وذلك لصالح نمط الإكمال. أما دراسة مایسة أبو مسلم (٢٠١٦) فقد توصلت أنه يزداد صدق وثبات الاختبارات الموضوعية - من نوع الاختيار من متعدد- وتزداد القدرة التمييزية للمفردة عندما يتوزع البديل الصحيح عشوائياً بين بدائل الاختبار.

ومن خلال ما ورد في الدراسات والبحوث السابقة يتبين قلة الدراسات التي تحدثت عن الأخطاء الشائعة في بناء الاختبارات الموضوعية وأثرها على الخصائص السيكومترية للاختبارات التحصيلية، ومعظم الدراسات والبحوث التي أجريت بالوطن العربي تناولت إحدى أنواع أسئلة الاختبارات الموضوعية؛ غير أن هذا البحث سيتناول بالدراسة العلمية الوقوف على الأخطاء الشائعة في بناء الاختبارات الموضوعية وأثر هذه الأخطاء على الخصائص السيكومترية للاختبارات التحصيلية، ولعل ذلك مما دفع الباحث إلى إجراء مثل هذا البحث.

ويمكن تلخيص مشكلة البحث في السؤالين الرئيسيين التاليين:

أولاً- ما الأخطاء الشائعة في بناء مفردات الاختبارات الموضوعية لدى أعضاء هيئة التدريس بكلية التربية جامعة أسيوط؟

ثانياً- هل هناك فروق دالة إحصائياً بين الاختبار الجيد (جميع فقراته سليمة البنوية) والاختبار غير الجيد (فقراته تحتوي على الأخطاء الشائعة في بناء الاختبارات الموضوعية) في الخصائص السيكومترية للاختبار وفقراته في مادة الاختبارات النفسية لطلاب الفرقة الرابعة شعبة علم النفس؟

ثالثاً- أهداف البحث: يهدف البحث الحالي إلى:

١- تحديد الأخطاء الشائعة لأعضاء هيئة التدريس بكلية التربية جامعة أسيوط في بناء الاختبارات الموضوعية.

٢- الكشف عن الفروق بين الاختبار الجيد (جميع فقراته سليمة البنية) والاختبار غير الجيد(فقراته تحتوي على الأخطاء الشائعة في بناء الاختبارات الموضوعية) في الخصائص السيكومترية للاختبار وفقراته في مادة الاختبارات النفسية لطلاب الفرقة الرابعة شعبة معلمي علم النفس.

رابعاً- أهمية البحث:

تتضح أهمية البحث الحالي من خلال ما يلي:

- تعريف أعضاء هيئة التدريس بأهم الأخطاء الشائعة في بناء فقرات الاختبارات الموضوعية، وتجنب ذلك مستقبلاً.
- يمكن الاستفادة من نتائج هذا البحث في وضع خطة تدريبية لما يحتاجه أعضاء هيئة التدريس من تدريب على مهارات بناء الاختبارات الموضوعية.
- يكشف عن الفروق التي تترتب على الأخطاء الشائعة في بناء الاختبارات الموضوعية وأثر ذلك على الخصائص السيكومترية للاختبارات التحصيلية.
- يعد هذا البحث- على حد معرفة الباحث- من الأبحاث القليلة التي تناولت الأخطاء الشائعة في بناء الاختبارات الموضوعية في جمهورية مصر العربية.

خامساً- محددات الدراسة:

✓ محددات مكانية:

تم تطبيق أدوات الدراسة في معامل وفصول ومدرجات كلية التربية - جامعة أسيوط.

✓ محددات زمانية:

تم تطبيق أدوات الدراسة في الفصل الدراسي الأول من العام الجامعي ٢٠١٧ / ٢٠١٨ م.

✓ محددات موضوعية:

- يقتصر البحث على عينة من نماذج أسئلة أعضاء هيئة التدريس بكلية التربية بجامعة أسيوط.
- يقتصر البحث على مقرر الاختبارات النفسية للفرقة الرابعة شعبة معلمي علم النفس للفصل الدراسي الأول من العام الجامعي ٢٠١٧ / ٢٠١٨ م.

سادساً- عينة الدراسة:

اقتصرت هذه الدراسة على عينة عشوائية من نماذج للاسئلة الموضوعية لاختبارات أعضاء هيئة التدريس بكلية التربية جامعة أسيوط للفصل الدراسي الأول بالعام الجامعي ٢٠١٧/٢٠١٨م، كما اشتملت العينة على طلاب وطالبات كلية التربية بالفرقة الرابعة شعبة معلمي علم نفس، وجدول (١) يوضح ذلك.

جدول (١)

أعداد ذكور وإناث طلاب الفرقة الرابعة شعبة معلمي علم النفس
ونماذج الاختبارات الموضوعية لأعضاء هيئة التدريس بكلية التربية جامعة أسيوط

المجموع	إناث	ذكور	النوع / الفرقة
٦٨	٥٩	٩	الرابعة
تكرار الاختبارات الموضوعية			الكلية أو القسم
٥			كلية الآداب
١			كلية التمريض
١			كلية التربية الرياضية
٢			كلية التربية النوعية
٢			كلية رياض الأطفال
١			كلية العلوم
٢			قسم المناهج وطرق التدريس
٢			قسم أصول التربية
١٦			المجموع

سابعاً - أدوات البحث:

- قائمة التدقيق للحكم على الأخطاء الشائعة في بناء الاختبارات الموضوعية (عينة البحث).
- اختبار تحصيلي في مقرر الاختبارات النفسية لطلاب الفرقة الرابعة شعبة معلمي علم النفس.

ثامناً - مصطلحات البحث:

يتناول البحث المصطلحات الإجرائية التالية:

١ - الأخطاء الشائعة في بناء الاختبارات الموضوعية:

هي مجموعة من الأخطاء الشائعة في بناء الاختبارات الموضوعية والمحددة في قائمة التدقيق التي أعدها الباحث لهذا الغرض والتي تكررت بنسبة ٢٠ % فأكثر في الاختبارات التي أعدها أعضاء هيئة التدريس بكلية التربية جامعة أسيوط للفصل الدراسي الأول من العام الجامعي ٢٠١٧/٢٠١٨م.

٢- الاختبارات التحصيلية:

هو اختبار موضوعي لمقرر الاختبارات النفسية لطلاب الفرقة الرابعة شعبة معلمي علم النفس، ويتألف من نموذجين: النموذج الأول- اختبار جيد جميع فقراته سليمة البناء، والنموذج الثاني- اختبار غير جيد من إعداد الباحث وجميع مفرداته لها نفس محتوى مفردات النموذج الأول لكنها تحتوي على بعض الأخطاء الشائعة لأعضاء هيئة التدريس بكلية التربية جامعة أسيوط.

٣- الاختبارات الموضوعية:

وهي الاختبارات التي تمكن المعلم من تكوين أحكام موضوعية على إجابة الطالب والموضوعية هنا يقصد بها البعد عن الذاتية في تقدير الدرجة على الإجابة، أي أن الدرجة التي يحصل عليها الطالب لا تختلف باختلاف المصحح، وبذلك لا تتأثر هذه الاختبارات بالعوامل الذاتية.

٤- الخصائص السيكومترية للاختبار: ويقصد بها كلاً من الصدق والثبات للاختبار.

أ- الصدق: يشير الصدق إلى مدى صلاحية استخدام درجات المقياس في قياس ما يدعي قياسه، فالصدق يحدد قيمة الاختبار وصلاحيته في قياس ما وضع لقياسه(رجاء علام، ٢٠١١، ٥١٦).

ب- الثبات: ويعد الثبات اتساق أداء الأفراد عبر الزمن إذا ما طبقت عليهم الأداة أكثر من مرة أو هو استقرار أداء الأفراد عبر صور متكافئة من الاختبار، ونستدل أيضاً على الثبات من خلال حساب النسبة بين التباين الحقيقي إلى التباين المشاهد (التباين الكلي) لدرجات الاختبار (Kimmo, 2000, 25؛ على ماهر، ٢٠٠٣، ١٦٥).

الإطار النظري والدراسات ذات الصلة

المحور الأول- مستويات القياس:

هناك الكثير من المعلومات الضرورية التي يجب توافرها قبل تفسير علامة اختبار ما ومن أهم هذه المعلومات معرفة نوع القياس المستخدم؛ إذ أن المقاييس المختلفة لها صفات مختلفة وتحمل مسؤوليات مختلفة من المعلومات وتشمل المقاييس أربعة مستويات هي: الاسمي، الرتيبي، الفترتي، النسبي(عبد الحميد محمد، ٢٠٠٩، ٢٢).

١. المقصود بمستويات القياس:

يُعرفها صلاح مراد، أمين سليمان (٢٠١٢، ٢٤) بأنها الطريقة التي تصنف بها الأشياء أو ترتب بها أو تقارن بينها تبعاً للخصائص المشتركة، بهدف معرفة الفروق بينها وأيضاً مدى هذه الفروق، والمعالجات الإحصائية الممكنة لهذا المدى.

٢. أنواع مستويات القياس:

أشار كل من صلاح علام (٢٠٠٠، ٢١)؛ على ماهر (٢٠٠١، ١٥٩)؛ عبد الهادي عبده وفاروق عثمان (٢٠٠٢، ٢١)؛ موسى النبهان (٢٠٠٤، ٢٩)؛ صلاح علام (٢٠٠٧، ٣١)؛ صلاح علام (٢٠١١، ١٨)؛ رجاء علام (٢٠١١، ٥١٥) صلاح مراد وأمين سليمان (٢٠١٢، ٢٤) إلى وجود أربعة مستويات للقياس في ميدان علم النفس والتربية مرتبطة من البسيط إلى المعقد تبعاً لدقة الصياغة الكمية (الرقمية) للمتغيرات التي ندرسها وهي:

أولاً: مستوى القياس الاسمي (التصنيفي) Nominal Scale.

يعد هذا المستوى أدنى مستويات القياس وأضعفها بل أن تسميته مقياساً تعد أحياناً تسمية مجازية وهنا تستخدم الأعداد للتسمية بغرض التصنيف، مثال: تحديد العدد ١ ليدل على أن جنس المفحوص ذكر والعدد ٢ ليدل على أن الجنس أنثى فهنا لا يعني ذلك أن ٢ أكبر من واحد وإنما الغرض تصنيفي بحت.

ثانياً: مستوى القياس الترتيبي Ordinal Scale or Ranking

وفي هذا المستوى يكون الهدف من الأعداد هو ترتيب الأفراد في ظاهرة معينة ولكن يجب ملاحظة أن الفروق بين الرتب ليست متساوية أو لا يعني أن المسافات البينية بين الأفراد متساوية.

ثالثاً: مستوى القياس الفترتي (المسافة) Interval Scale.

وفي هذا المستوى تتساوى الفروق أو المسافات بين المستويات المتتالية مثل درجات الأطفال في اختبار الذكاء أو درجاتهم في التحصيل أو درجاتهم في اختبار القلق فهنا يكون هناك وحدة قياس ثابتة متفق عليها يقاس بها الفرق بين كل درجة والتالية لها.

رابعاً: مستوى القياس النسبي Ratio Scale.

وفي هذا المستوى نلاحظ أن للدرجات صفراً مطلقاً يعنى غياب الصفة مثل مقياس الوزن ودرجة الحرارة ويسمى هذا المستوى بالمستوى النسبي لأن النسبة بين أي درجتين لا تتأثر بوحدة القياس.

المحور الثاني - الاختبارات التحصيلية:

لعل كل طالب قد ألف الاختبارات منذ الوهلة الأولى لخبرته في أخذ الفحوص العديدة في المناسبات المتنوعة ولأغراض مختلفة. فمن المحتمل أن يكون بعض الطلاب قد مارسوا اختبارات الذكاء وبعضهم قد مارس أنواع أخرى من الاختبارات مثل القدرات أو القبول في برنامج تعليمي أو تدريبي.

أنواع الاختبارات :

تتعدد أنواع الاختبارات التي تستخدم في تقويم الطالب ويمكن تقسيم هذه الاختبارات إلى ثلاثة أنواع هي:

١- الاختبارات التحريرية. ٢- الاختبارات الشفوية. ٣- الاختبارات العملية، والتطبيقية.

• الاختبارات التحريرية :

تتمثل الاختبارات التحريرية في نوعين هما :

- الاختبارات المقالية.

- الاختبارات الموضوعية.

ويتوقف اختيارنا للاختبارات المقالية أو الموضوعية على طبيعة المادة والأهداف المنشودة.

وسوف نلقى الضوء على طبيعة كل نوع من هذه الاختبارات.

أولاً: الاختبارات المقالية Essay tests

تستخدم هذه الاختبارات لقياس قدرة الطالب على التعبير ، وترتيب الأفكار وصياغتها بأسلوبه الخاص ، والتعرف على وجهات النظر ، وإذا أحسن صياغة أسئلة هذا النوع من الاختبارات ، فإنه يمكن من خلالها التعرف على مدى اكتساب الطالب لمستويات التعلم العليا مثل القدرة على التحليل ، والتفسير ، والاستدلال ، والاستنباط ، والتفكير ، والإبداع ، كما أنها تحقق - إلى حد كبير - تقويم هذه العمليات أكثر من تقويم نواتجها ، وتعد هذه الاختبارات إحدى الأساليب الشائعة للتقويم (صلاح مراد، وأمين سليمان، ٢٠١٢، ٧٧).

وكثيراً ما تبدأ أسئلتها بكلمات مثل : صف .. اذكر .. اشرح .. علل .. قارن .. اكتب ..
وضح .. ناقش .. تكلم .. عن .. إلخ.

ويمكن أن نميز عدة أنواع من اختبارات المقال منها : اختبار المناقشة، والاستدعاء
الانتقائي، والاستدعاء التقويمي، والمقارنة، والتحليلي، والاستنباطي، والاستدلالي، حيث
يحدد المطلوب من الأسئلة طبيعة نوعها والهدف منها.

ثانياً: الاختبارات الموضوعية Objective tests

هي الاختبارات التي تمكن المعلم من تكوين أحكام موضوعية على إجابة
الطالب، والموضوعية هنا يقصد بها البعد عن الذاتية في تقدير الدرجة على الإجابة،
أى أن الدرجة التي يحصل عليها الطالب لا تختلف باختلاف المصحح، وبذلك لا
تتأثر هذه الاختبارات بالعوامل الذاتية.

• مزايا الاختبارات الموضوعية:

تنطوي الاختبارات الموضوعية على عدة مزايا من أهمها :

- ١- نتائجها ثابتة إلى حد كبير ولا تتأثر بذاتية المصحح.
- ٢- تغطي معظم موضوعات المقرر الدراسي.
- ٣- تحدد بدقة تباين المستويات التحصيلية للطلاب، نظراً لكثرة عدة أسئلتها والتي تغطي
معظم الأهداف المعرفية والوجدانية والمهارية.
- ٤- تكشف عن قدرة الطالب على سرعة التفكير أكثر من القدرة على الاستظهار.
- ٥- سهولة التطبيق والتصحيح.
- ٦- يمكن لغير المتخصص أن يقوم بتصحيحها بسهولة من خلال مفتاح الإجابة
وتوزيع الدرجات.
- ٧- يسهل مقارنة نتائج اختبارات بنتائج اختبارات أخرى مماثلة لها.

• عيوب الاختبارات الموضوعية :

ويؤكد علي ماهر (٢٠٠١، ٢٦٦) على وجود بعض الانتقادات التي توجه إلي الاختبارات الموضوعية من أهمها:

- ١- أنها تترك للطالب مجالاً للتخمين في اختيار بديلات الإجابة عن الأسئلة.
- ٢- عدم قدرتها على قياس جميع جوانب التحصيل ، وخاصة فيما يتعلق بقدرة الطالب على تنظيم أفكاره بأسلوبه الخاص ، وقدرته على النقد والتقويم ، وتطبيق ما تعلمه في مواقف الحياة العملية كما أنها لا تقيس الأداء الإبداعي بالنسبة للطلاب.
- ٣- يتطلب إعدادها وقتاً وجهداً كبيرين ، ومهارات معينة قد لا تتوفر لدى بعض المدرسين مما يجعل الاختبار الموضوعي ضعيف البنية، وغير قادر على قياس ما وضع من أجله.

• أنواع الاختبارات الموضوعية:

تأخذ الاختبارات الموضوعية أشكالاً متنوعة وسوف نلقى الضوء على أكثر الاختبارات شيوعاً في مجال التعليم:

١- أسئلة الصواب - الخطأ True-false questions

تقيس هذه الأسئلة النتائج التعليمية من مستوى التذكر وهذا النوع ليس مناسباً بشكل عام لقياس الأهداف ذات المستويات العليا ، وبالرغم من ذلك فلها القدرة على التمييز وتحديد علاقات السبب والنتيجة.

وتصاغ أسئلة هذه الاختبارات في عبارات تقريرية تتألف من جملة واحدة أو أكثر، ويضع الطالب علامة (√) إذا كانت العبارة صحيحة وعلامة (×) إذا كانت الإجابة غير صحيحة.

• بعض المقترحات لتحسين أسئلة الصواب - الخطأ:

- استخدام لغة سهلة واضحة ، ذات معنى محدد دقيق ، لأن معظم الاعتراضات على هذا النوع من الأسئلة هو استخدامها لمفردات غامضة وغير دقيقة.
- تجنب استخدام العبارات العامة الواسعة لأن معظم التعميمات الواسعة تكون غير سليمة إلا إذا حددت بوصفها تحديداً دقيقاً.

- تجنب استخدام العبارات المنفية ، خاصة إذا كان النفي مزدوجاً ، لأن الطلاب يميلون إلى إغفال المفردات المنفية مثل : استخدام أدوات النفي (لا ، ليس) بينما يجعل النفي المزدوج المعنى غامضاً ، مع مراعاة أن نفي النفي إثبات.
- لا تجعل العبارات الخاطئة أطول أو أقصر من العبارات الصحيحة باستمرار ويفضل أن تكون جميع العبارات متساوية في الطول (صلاح علام ، ٢٠٠٧، ١٠٨).

٢ - أسئلة الاختيار من متعدد Multiple choice questions MCQ

وتتألف فقرة الاختيار من متعدد من جزئين .

- نص السؤال .
 - قائمة البدائل أو الاختيارات. وهذه تنقسم إلى قسمين :
 - البديل أو الاختيار الصحيح.
 - المموهات وهي البدائل المقبولة ظاهرياً ولكنها ليست إجابات صحيحة.
- وإذا وضعت أسئلة هذا الاختبار بمهارة فيمكن الارتقاء بمستواها ، بحيث لا يصبح التركيز على قياس القدرة على التذكر فقط، بل تقيس القدرة على الفهم والتطبيق والتحليل والإبداع.

• أشكال أسئلة الاختيار من متعدد :

يشير صلاح علام (٢٠٠٧، ٩٨) إلى أن هناك أشكالاً متنوعة لأسئلة الاختيار من متعدد ، لكن أكثرها شيوعاً هي الأنماط التالية:

أ - الإجابة الصحيحة الواحدة :

وهذا النمط أهم أنماط أسئلة الاختيار من متعدد ، وفيه توجد إجابة صحيحة واحدة فقط، بينما تبدو البدائل الأخرى وكأنها صحيحة ولكنها بالتأكيد خاطئة.

ب - أفضل الإجابات :

وفى هذا النمط أكثر من بديل واحد صحيح ، ولكن أحدها هو أفضل الإجابات ، وعلى الطلاب اختيار هذا البديل الأكثر صحة. (وعلى واضع الاختبار أن يوضح ذلك فى تعليمات إجابة الاختبار) مثل أن يذكر فى بداية هذا النوع من الاختبارات اختر الإجابة الصحيحة أو الأكثر صحة.

وفى الغالب فإن هذا النمط من أسئلة الاختيار من متعدد أكثر صعوبة من نمط الإجابة الصحيحة الواحدة ، لأنه بحاجة إلى تمييز جيد، وهذا النوع مفيد لقياس التعلم فى مستوياته العليا ، والفهم ، والتطبيق وتفسير الحقائق المعرفية.

ج - المماثلة أو المشابهة أو التشابه الجزئى أو التناظر :

وفى هذا النمط يستنتج الطلاب العلاقة بين المفهومين الأول والثانى ، ثم يطبقون ذلك فى المفهومين الثالث والرابع.

د - النمط المعكوس للاختيار من متعدد :

هذا النمط عكس شكل النمط ذى الاستجابة الواحدة الصحيحة ، أى أن الطالب يختار الإجابة الخطأ من بين الإجابات الصحيحة، وهذا النمط يعد صعباً إلى حد ما ، وبالتالي من الأسهل وضع عدة إجابات صحيحة ، عن طريق عكس نمط السؤال والخبراء لا يشجعون استخدام هذا النمط ، أما إذا كان لابد من ذلك فيجب تنبيه الطلاب إلى اختيار الإجابة غير الصحيحة.

● بعض الاقتراحات لتحسين أسئلة الاختيار من متعدد :

فيما يلى بعض الاقتراحات التى تساعد على تلافى القصور فى هذه الاختبارات.

- أن يتضمن نص السؤال مشكلة واحدة محددة واضحة.
- يفضل تجنب تكرار المفردات فى البدائل.
- يراعى تجنب المفردات المبالغ فيها.
- يراعى عند إكمال الفقرة الناقصة بأحد البدائل أن يكون النقص فى نهاية الفقرة.
- يراعى تجنب استخدام النفى المزدوج.
- يراعى تجنب استخدام البدائل التى تحمل فى صياغتها الإيحاء بالإجابة الصحيحة.
- يفضل أن تكون كل البدائل متساوية الطول.
- يراعى أن تكون إحدى فقرات الاختيار صحيحة أو أكثر صحة.
- يراعى تغيير موضع الإجابة الصحيحة فى البدائل عشوائياً.
- تجنب استخدام بديل (كل ما ذكر سابقاً) أو (ليس مما ذكر).

٣- أسئلة المزوجة Matching items

تتكون أسئلة المزوجة من عمودين متوازيين يتضمن العمود الأول المثيرات (المقدمات). أما العمود الثانى فيتضمن الاستجابات (نتائج). وكل مثير من العمود الأول يقابله استجابة صحيحة فى العمود الثانى ، وما على الطلاب إلا أن يقوموا بالمزوجة بين المثيرات والاستجابة الصحيحة. فمثلاً يزوج بين التواريخ والأحداث التاريخية وبين المصطلحات وتعريفاتها وبين الرموز والمفاهيم.

• بعض المقترحات لتصميم أسئلة المزوجة :

يؤكد صلاح علام (٢٠٠٧، ١١٢) على بعض المقترحات لتصميم أسئلة المزوجة منها:

١- يفضل ألا يقل الفرق بين عدد المقدمات وعدد الاستجابات عن ثلاثة للإقلال من تأثير التخمين.

٢- يفضل ترتيب الاستجابات ترتيبياً منطقياً.

٤- أسئلة الإجابة القصيرة :

هذا النمط من الأسئلة يمد الطالب بالإجابة التى قد تتكون من كلمة مفردة أو رمز أو معادلة أو كلمتين أو جملة أو شبه جملة ويأخذ السؤال أحد الأشكال التالية:

- أن يكون كاملاً مثل السؤال متى ... ؟ ماذا ... ؟

- أن يكون عبارة مثل عرف ...

- أن يكون عبارة ناقصة أو ملء فراغ.

ويمكن الارتقاء بهذا النوع من الاختبارات بحيث لا يكون التركيز فيها قياس القدرة على التذكر فقط بل يمكنه قياس نواتج التعلم الأخرى، وهو يصلح فى العلوم الإنسانية والاجتماعية.

• بعض المقترحات لتحسين أسئلة الإجابة القصيرة :

- يجب أن يكون لعبارة السؤال إجابة واحدة صحيحة.

- يجب أن يكون نص عبارة السؤال مغايراً للنص المكتوب فى المذكرة أو الكتاب المقرر.

- يفضل أن تكون فقرات هذا الاختبار عبارة عن أسئلة وليس عبارات وإذا كانت عبارات يفضل أن يكون فراغ واحد فقط في نهاية العبارة.
- يفضل تجنب الأسئلة التي تتطلب في إجاباتها عدة سطور.
- يفضل أن تضع مفتاحاً للإجابة عند تصحيحك لهذا النوع من الأسئلة.
- يفضل تجنب الأسئلة الصعبة المعقدة أو الغامضة.

٥- أسئلة المشكلات :

هذا النوع من الأسئلة يستخدم على نطاق واسع في مجال العلوم والرياضيات وفيها يتم عرض الموقف للطلاب وتقدم لهم المعلومات المناسبة لهذا الموقف ثم يطلب من الطلاب تقديم الحل استناداً إلى المعطيات المعطاة.

وبذلك فإن المشكلة تتطلب عمليات عقلية متتابعة وقد تختلف مهام الحل وفقاً لطبيعة الموضوع والمشكلة وإذا أحسن صياغة هذا النوع من الأسئلة فإنها تقيس نواتج التعلم المعقدة على مستوى التطبيق والتحليل والتركيب والتقويم والإبداع وبذلك يعد هذا النوع من الأسئلة من أفضل وسائل التقويم في المراحل المتقدمة.

• بعض المقترحات لتحسين أسئلة المشكلات :

- يراعى صياغة المشكلة بصورة واضحة ومحددة بحيث تيسر فهم الطلاب لها واقتراح الحلول المناسبة.
- يراعى أن تكون نوع المعلومات التي تتضمنها المشكلة تساعد بالتنبؤ بالحلول التي يقدمها الطالب لحل المشكلة.
- يفضل تحديد الزمن المناسب لكل مشكلة.

• خطوات وأسس بناء الاختبار الجيد :

قبل أن نتحدث عن خطوات وأسس بناء الاختبار الجيد ، سوف نشير إلى بعض الأمور الشكلية التي ينبغي مراعاتها عند إخراج ورقة الأسئلة. وفيما يلي أهم هذه الأمور :

- ١- يراعى أن تكون الأسئلة مكتوبة على الآلة الكاتبة أو الكمبيوتر ويحظر كتابة الأسئلة بخط اليد ، مع مراعاة تأمين سرية الاختبار عند استخدام الكمبيوتر في كتابة أسئلة الاختبار.

- ٢- أن تتضمن ورقة الأسئلة بيانات أساسية تدون في أعلى الورقة وتفصل عن الأسئلة بفواصل واضح وهذه البيانات هي :
- اسم الامتحان - المقرر الدراسى - الفرقة الدراسية - زمن الامتحان - الفصل الدراسى.
- ٣- يحدد عدد الأسئلة المراد الإجابة عنها ، وإن كان هناك أسئلة إجبارية يشار إليها بخط واضح (بنط عريض) أو يوضع تحتها خط.
- ٤- أن تكون تعليمات الإجابة عن السؤال واضحة وشاملة توجه الطالب إلى الإجابة المطلوبة بدون مساعدة الآخرين.
- ٥- يراعى سلامة اللغة فى صياغة الأسئلة سواء كانت لغة الاختبار باللغة العربية أو اللغة الأجنبية.
- ٦- يفضل أن توزع درجة الاختبار على الأسئلة ويسجل أمام كل سؤال درجته.
- ٧- يراعى أن تكون ورقة الأسئلة منظمة وجذابة ويستخدم فيها أكثر من بنط فى الكتابة إذا استدعى الأمر.
- ٨- يفضل فى نهاية الأسئلة أن تكتب عبارة انتهت الأسئلة مع أطيب تمنياتنا بالتوفيق ، ويكتب اسم الممتحن.

المحور الثالث - إعداد وبناء الاختبارات التحصيلية:

لا يعد تصميم وإعداد اختبار جديد هدفا فى حد ذاته؛ فالباحث لا يلجأ إلى ذلك إلا إذا دعت حاجة ملحة، وعندما يتضح له أن الاختبارات المتوفرة فى المجال لا تفى بغرضه ودون وجود الدواعى الضرورية لتصميم اختبارا جديدا يظل استخدام الاختبارات المتوفرة وتطويرها افضل كثيرا فى الدراسات والبحوث ، وتتضمن عملية بناء اختبار جديد الخطوات الآتية كما أوردها علي ماهر (٢٠٠١، ٣١٣):

- ١- تحديد الأهداف Objectives (النواتج) التى يهدف المقرر الذى يقوم المعلم بتدريسه إلى تحقيقها.
- ٢- تحديد محتوى الاختبار وتحليله من خلال جدول المواصفات.
- ٣- تحديد مفردات (أسئلة) الاختبار وتحديد نوعها.
- ٤- كتابة تعليمات ومفردات الاختبار ومراجعتها لغوياً.

- ٥- إعداد الصورة التجريبية للاختبار (الصورة الأولية التجريبية للاختبار).
- ٦- تجهيز وإعداد مفتاح تصحيح الاختبار (نموذج الإجابة) و تطبيق الصورة الأولية التجريبية للاختبار.
- ٧- تصحيح الاختبار ورصد الدرجات على نحو يسمح بمعالجتها إحصائياً وعمل التحليلات اللازمة إحصائياً.
- ٨- عمل التعديلات اللازمة في الصورة الأولية التجريبية للاختبار.
- ٩- تطبيق الصورة التجريبية المعدلة للاختبار على مجموعات أخرى من الأفراد تمثل فصول أخرى درست نفس المقرر ونفس المحتوى الذي يمثله الاختبار، وتصحيحها.
- ١٠- عمل تحليل المفردات لدرجات الأفراد على مفردات الاختبار للتأكد من أن المفردات التي تم تعديلها والتي تم استبدالها تستوفي الخصائص السيكمترية للمفردات المرغوبة أو الجيدة، وإجراء التعديلات الإحصائية اللازمة لتقدير معاملات صدق وثبات الاختبار.

المحور الرابع - الخصائص السيكمترية للاختبارات:

يشير قاسم الصراف (٢٠٠٢، ١٨٩) إلى أن الأدوات لكي تكون مقاييس جيدة يجب أن تتصف بمجموعة من الخصائص من أهمها الصدق والثبات اللذان يعتبران من أكثر المصطلحات الفنية في القياس.

أولاً- الصدق Validity:

يوصف مفهوم الصدق من قبل مجموعة واسعة من المصطلحات النوعية، وهذا المفهوم ليس مفهوماً واحداً إنما هو بناءً مشروطاً ضرورياً لصحة البيانات، وبالرغم من أن الباحثين توصلوا إلى أن مصطلح الصدق لا ينطبق على البحوث النوعية؛ لكنهم أدركوا مدى الحاجة إلى ضرورة الفحص أو القياس المؤهل لأبحاثهم وعلى سبيل المثال أشار كلاً من (Creswell & Miller, 2000, 124) إلى أن مصطلح الصدق يتأثر بمدى إدراك الباحث له في الدراسة واختياره لافتراض النموذج، ونتيجة لذلك فإن كثيراً من الباحثين طور مفهوم الصدق، ففي كثير من الأحيان يتم الاعتماد على ما يمكن اعتباره مصطلحات أكثر ملاءمة مثل: الجودة، الثقة، الصرامة (Davies & Dodd, 2002, 279؛ Stenbacka, 2001, 551)

ويؤكد رجاء علام (٢٠١١، ٥١٦) على أن مفهوم الصدق يشير إلى الاستدلالات الخاصة التي نخرج بها درجات المقياس من حيث مناسبتها وفائدتها، وتحقيق صدق المقياس

معناه تجميع الأدلة التي تؤيد مثل هذه الاستدلالات؛ ولذلك يشير الصدق إلى مدى صلاحية استخدام درجات المقياس في قياس ما يدعي قياسه، فالصدق يحدد قيمة الاختبار وصلاحيته في قياس ما وضع لقياسه.

طرق تحديد الصدق:

تتضمن كتب مناهج البحث عدداً من طرق إيجاد الصدق، منها:

- صدق المحتوى:

يمكن حساب صدق المحتوى للاختبار عن طريق التحليل المبدئي لفقراته بواسطة عدد من المحكمين لتحديد ما إذا كانت هذه الفقرات تتعلق بالجانب المقيس، ثم يقوم الباحث بعمل تكرارات لاستجابات المحكمين ويختار المفردات التي اتفق عليها أكبر عدد من المحكمين. وتشير أدبيات البحث إلى أنه من الأفضل حساب معامل الاتفاق بين المحكمين من خلال معامل اتفاق كندال، ويكون من المناسب استخدام صدق المحتوى مع الاختبارات والاستبيانات التي تقيس مفرداتها التحصيل أو المفاهيم (سامية محمد، ٢٠٠٠، ٤٣٩؛ على ماهر، ٢٠٠٣، ١٨٥؛ Gholamreza & Fatemeh, 2008, 152؛ صلاح علام، ٢٠١١، ١٩٢).

- صدق المضمون:

يمكن حساب صدق المضمون للاختبار بحساب معاملات الارتباط بين درجات الأفراد في كل سؤال ودرجاتهم في الاختبار ككل بعد حذف درجة المفردة من المجموع الكلي للاختبار، وتشير أدبيات البحث إلى أنه من المناسب استخدام صدق المضمون مع الاختبارات واستمارات الملاحظة (محمد السيد، ٢٠٠٠، ٣١٥).

- الصدق التلازمي:

و درجاتهم في الأداء الفعلي في جوانب السلوك التي يقيسها الاختبار، وذلك بشرط أن تكون درجات الأداء الفعلية للأفراد قد تم جمعها وقت إجراء الاختبار أو قبلها، ويكون من المناسب استخدام الصدق التلازمي مع الاختبارات التي تقيس جوانب الشخصية بسماتها وأبعادها وصفاتها.

- الصدق التنبؤي:

يمكن حساب معامل الصدق في هذه الطريقة بمعامل الارتباط بين الدرجات على الاختبار ودرجات الأداء الفعلي للأفراد كما يقاس بطريقة أخرى بعد إجراء الاختبار بفترة زمنية،

وتستخدم هذه الطريقة في حساب صدق اختبارات الاستعدادات الخاصة، مثل: الاستعداد الرياضي أو الاستعداد الميكانيكي، ويطلق على الصدق التلازمي والتنبؤي معاً صدق المحكات، وتشير أدبيات البحث إلى أنه من المناسب استخدام صدق المحكات مع الاختبارات واستمارات الملاحظة والاستبيانات والمقابلات (أحمد محمد، ٢٠٠٠، ١٢٧؛ بشير صالح، ٢٠٠٠، ١٧٣).

- صدق التكوين الفرضي:

ويطلق عليه أحياناً صدق المفهوم، ويمكن حساب معامل صدق الاختبار بهذه الطريقة بتحديد معامل الارتباط بين درجات الأفراد على الاختبار، وبين مفهوم هذه الجوانب، كما تحددتها النظرية التي يتبناها الباحث في أثناء بنائه لهذا الاختبار وتشير أدبيات البحث إلى أنه من المناسب استخدام صدق المفهوم مع المقاييس التي تقيس الاضطرابات الشخصية والسلوكية والحالات المرضية (محمود عبد الحليم، ١٩٩٤، ٢٠٨).

- الصدق التطبقي:

ويطلق عليه البعض صدق المحكات ونحصل على معامل الصدق التطبقي بحساب مدى اتفاق درجات مجموعه من الأفراد في الاختبار مع درجاتهم على اختبار آخر ثبت أنه صادق في قياس نفس السمة التي يقيسها الاختبار الجديد، وتشير أدبيات البحث إلى أنه من المناسب استخدام الصدق التطبقي مع الاختبارات واستمارات الملاحظة والاستبيانات والمقابلات (أحمد محمد، ٢٠٠٠، ١٢٧؛ بشير صالح، ٢٠٠٠، ١٧٣).

- الصدق العاملي:

وتعتمد هذه الطريقة في حساب معامل صدق الاختبار على أسلوب التحليل العاملي، الذي يهدف إلى تحديد مدى قياس مجموعة اختبارات لبعض العوامل المشتركة، ولعلنا نلاحظ البعض يجمع بين الصدق التطبقي والصدق العاملي عند الحديث عن صدق المفهوم، وبالرغم من أهمية طريقة التحليل العاملي إلا أن بعض الباحثين قد يخطئ في استعمال هذه الطريقة (Martijn, G.& Jan-Bendict, E, 2010, 187). وتشير أدبيات البحث إلى أنه من المناسب استخدام الصدق العاملي مع الاختبارات والاستبيانات.

- الصدق الذاتي:

يقاس الصدق الذاتي عن طريق الجذر التربيعي لمعامل ثبات الاختبار (مصطفى باهي وفاتن النمر، ٢٠٠٩، ١٠٧) ومن أهم عيوب معامل الصدق الذاتي أنه يعتمد

على معامل الثبات، كما أن هذا الأسلوب يتجاهل تماماً المبدأ الأساسي الذي يربط بين مفهومي الصدق والثبات وهذا المبدأ الذي يرى أن كل اختبار صادق ثابت، وليس كل اختبار ثابت صادق، وخلاصة القول إن أي معامل سواء أكان للثبات أم للصدق يجب أن يؤخذ بدقة وبحذر وأن يكون المعامل المستخدم مناسباً لطبيعة البحث، فزيادة معاملات الصدق والثبات ضرورية للبحث في أي نموذج (Nahid, 2003, 602).

ثانياً- ثبات الأدوات:

الثبات من أهم الشروط السيكومترية للأدوات حيث لا غنى من حسابه مع الصدق؛ لأنه يتعلق بمدى دقة الأدوات في قياس ما تدعي قياسه، وبعد الثبات اتساق أداء الأفراد عبر الزمن إذا ما طبقت عليهم الأداة أكثر من مرة أو هو استقرار أداء الأفراد عبر صور متكافئة من الاختبار، ونستدل أيضاً على الثبات من خلال حساب النسبة بين التباين الحقيقي إلى التباين المشاهد (التباين الكلي) لدرجات الاختبار (Kimmo, 2000, 25)؛ على ماهر، (٢٠٠٣، ١٦٥).

طرق حساب الثبات:

تتعدد أساليب حساب الثبات، ويختص كل أسلوب منها بتقدير نوعية محددة من "تباين الخطأ"، وهو التباين الذي يؤثر على ثبات القياس الذي نحصل عليه كلما استخدمنا مقاييسنا المختلفة، وأشار كل من (محمود عبد الحليم (١٩٩٤، ٢٠٣)؛ على ماهر (٢٠٠٣، ١٦٧)؛ بشرى إسماعيل (٢٠٠٤، ٧٢)؛ صلاح علام (٢٠٠٧، ٢٣٤) إلى أنه توجد أكثر من طريقة لحساب معامل ثبات الاختبار وهي:

- أ- إعادة تطبيق الاختبار.
- ب- الصور المتكافئة أو البديلة.
- ج- التجزئة النصفية.
- د- التناسق الداخلي.
- هـ- تحليل التباين.

وفيما يلي يتم إلقاء الضوء على هذه الطرق:

- طريقة إعادة التطبيق:

تشير أدبيات البحث إلى أنه من المناسب استخدام هذه الطريقة مع اختبارات الاستعدادات والتحصيل والاستبيانات واستمارات الملاحظة والمقابلات (أحمد محمد، ٢٠٠٠، ١٢٥؛ صلاح علام، ٢٠١١، ١٧٣).

- طريقة الصور المتكافئة:

تشير أدبيات البحث إلى أنه من المناسب استخدام هذه الطريقة مع اختبارات التحصيل ومواقف العلاج النفسي واختبارات حل المشكلات (محمود عبد الحليم، ١٩٩٤، ٢٠٥؛ بشرى إسماعيل، ٧٤، ٢٠٠٤؛ صلاح علام، ٢٠١١، ١٤٦).

- طريقة التجزئة النصفية:

تشير أدبيات البحث إلى أنه من المناسب استخدام هذه الطريقة مع اختبارات القدرات وكثير من مقاييس الشخصية والمقابلات (عبد الحميد محمد، ٢٠٠٩، ٢٣٦؛ صلاح علام، ٢٠١١، ١٤٦).

- طريقة التناسق الداخلي:

تشير أدبيات البحث إلى أنه من المناسب استخدام هذه الطريقة مع الاختبارات والاستبيانات والمقابلات (عبد الحميد محمد، ٢٠٠٩، ٢٣٦).

- ثبات الفاحصين والمصححين:

تشير أدبيات البحث إلى أنه من المناسب استخدام هذه الطريقة مع استمارات الملاحظة والمقابلات والاختبارات المقالية والطرق الإسقاطية وموازن التقييم (صلاح علام، ٢٠١١، ١٦٧).

إجراءات البحث

أولاً- منهج البحث:

يدور البحث الحالي حول الأخطاء الشائعة في بناء الاختبارات الموضوعية وتأثيرها على الخصائص السيكومترية للاختبارات التحصيلية لدى طلاب كلية التربية، وفي سبيل ذلك تم إعداد قائمة لتحديد أهم الأخطاء التي يقع فيها أعضاء هيئة التدريس عند إعدادهم للاختبارات الموضوعية، كما تم إعداد اختبار تحصيلي في مقرر الاختبارات النفسية من صورتين؛ لذلك تم اتباع المنهج الوصفي التحليلي، وهو المنهج الملائم لأهداف البحث الحالية.

ثانياً- مجتمع وعينة البحث:

(أ) مجتمع البحث:

تكون مجتمع البحث من طلاب وطالبات كلية التربية بجامعة أسيوط للعام الجامعي ٢٠١٧/٢٠١٨م، واقتصرت العينة على طلاب الفرقة الرابعة شعبة معلمي علم النفس، حيث بلغ عددهم ٦٨ (٩ طلاب، ٥٩ طالبة) بمتوسط عمري ٢٢.٣٠ سنة وانحراف معياري ١.٦٥، وبلغ عدد الاختبارات الموضوعية للفصل الدراسي الأول ١٦ اختباراً.

(ب) العينة الاستطلاعية:

تم اشتقاق عينة عشوائية مكونة من ٣٠ طالباً وطالبة من طلاب وطالبات كلية التربية الفرقة الرابعة شعبة معلمي علم النفس للعام الجامعي ٢٠١٧/٢٠١٨ م، وتم تطبيق عليهم صورتى الاختبار للتحقق من الصدق والثبات.

(ج) العينة النهائية:

اقتصر هذا البحث على عينة عشوائية من طلاب وطالبات كلية التربية متمثلة في الفرقة الرابعة شعبة معلمي علم النفس، حيث بلغ عدد الذكور ٩ طلاب بمتوسط عمري ٢٢.٦٥ سنة وانحراف معياري ١.٧١، وعدد الإناث ٢٩ طالبة بمتوسط عمري ٢٢.٣٥ سنة وانحراف معياري ١.٦٧.

ثالثاً- أدوات الدراسة:

- قائمة تدقيق الأخطاء في بناء الاختبارات الموضوعية. إعداد الباحث

استعان الباحث بالكتابات النظرية التي كتبت من قِبل المختصين في مجال الاختبارات الموضوعية ومقترحات تحسينها، كما تمت الاستعانة ببعض الدراسات التي صممت قوائم في هذا المجال وهي: دراسة سناء أحمد (٢٠١١)، دراسة حيدر ظاظا (٢٠١٢)، ودراسة بندر الأمير (٢٠١١)، وبعد تصميم القائمة تم عرضها على مجموعة من السادة المحكمين - مدرج أسمائهم بالملاحق- وتم عمل التعديلات اللازمة، وتم حساب نسبة الاتفاق بين المحكمين من خلال معامل كندال والذي بلغ ٠.٩٣٢ مما يدل على صدق القائمة، وقد اشتملت القائمة على محورين رئيسين هما:

أ- المحور الأول- الأخطاء الشائعة في بناء مفردات الصواب والخطأ: وقد اشتمل هذا المحور على إحدى عشرة فقرة.

ب- المحور الثاني- الأخطاء الواردة في مفردات الاختيار من متعدد: وقد اشتمل هذا المحور على إحدى وثلاثون فقرة.

- الاختبار التحصيلي لمقرر الاختبارات النفسية. إعداد الباحث

استخدم الباحث الاختبار التحصيلي لمقرر الاختبارات النفسية لطلاب الفرقة الرابعة بقسم علم النفس كلية التربية، وقد تم إعداد هذا الاختبار في ضوء جدول المواصفات.

(١) الصدق Validity :

اعتمد الباحث في حساب صدق الاختبار على ما يلي :

- الصدق المنطقي (صدق المحكمين) Logical Validity

تم عرض الصورة الأولية للاختبار على مجموعة من السادة المحكمين المتخصصين في مجال علم النفس التربوي والقياس النفسي (ملحق ١) ، وقد اشتملت تلك الصورة على (٢٦) عبارة بهدف: التأكد من مناسبة العبارات للمفهوم المراد قياسه، وتحديد غموض بعض العبارات لتعديلها، وحذف بعض العبارات غير المرتبطة بالاختبار، أو غير مناسبتها لطبيعة وخصائص عينة البحث، ويوضح جدولاً (٢) ، و (٣) بعض العبارات التي تم تعديلها، والأخرى التي تم حذفها.

جدول (٢)

العبارات التي تم تعديل صياغتها لاختبار الاختبارات النفسية

م	العبارات قبل التعديل	العبارات بعد التعديل
١٠	اختبار ستانفورد بينيه واختبار رافن كلا منهما مثال على تصنيف الاختبارات النفسية من حيث الأداء.	اختبارا ستانفورد بينيه ورافن مثال على تصنيف الاختبارات النفسية من حيث الأداء.
١٢	عند حساب صدق المحكمين يجب عرض فقرات المقياس على عدد زوجي من المحكمين	يفضل عرض فقرات المقياس عند تحكيمها على عدد زوجي من المحكمين.

جدول (٣)

العبارات التي تم حذفها لاختبار الاختبارات النفسية

م	العبارات التي تم حذفها	سبب الحذف
١	عند حساب الثبات بطريقة التجزئة النصفية يفضل وضع عدد زوجي من الأسئلة.	تكرار الفقرة
٧	يصلح استخدام الصدق العاملي للعينات المعتدلة .	تكرار الفقرة

- وفي ضوء آراء السادة المحكمين تم تعديل (٢) فقرة، وحذف (٢) فقرة؛ لتكرار بعضها، ولعدم مناسبتها لطبيعة وخصائص العينة، والتي لم تحظ بنسبة اتفاق تتراوح بين (٩٠% - ١٠٠%).

- أصبح الاختبار بعد حذف العبارات التي لم تحظ بنسبة اتفاق تتراوح بين (٩٠% - ١٠٠%) من السادة المحكمين في صورته الأولية يحتوي على (٢٤) عبارة، وتم حساب نسبة الاتفاق بين المحكمين من خلال معامل كندال والذي بلغ ٠.٩٢٠ مما يدل على صدق الاختبار، وتم تطبيقه على عينة الدراسة الاستطلاعية للاستقرار على الصورة النهائية للاختبار.

(٢) الثبات Reliability :

- طريقة كيودر ريتشاردسون :

استخدم الباحث معادلة كيودر ريتشاردسون ، وهي معادلة تستخدم لإيضاح المنطق العام لثبات الاختبارات والمقاييس، ويوضح جدول(٤) معاملات كيودر ريتشاردسون للابعد والدرجة الكلية للاختبار.

جدول (٤)

معاملات كيودر ريتشاردسون لثبات اختبار الاختبارات التفسيرية وابعاده (ن = ٣٠)

معامل الثبات (كيودر ريتشاردسون)	الابعاد
٠.٨٢٤	أسئلة الصواب والخطأ
٠.٨٣٦	أسئلة الاختيار من متعدد
٠.٨٣١	المقياس ككل

يتضح من جدول (٤) ان قيم معاملات الثبات كانت مرتفعة حيث تراوحت قيم معاملات الثبات بين ٠.٨٢٤ و ٠.٨٣٦ مما يدل على ثبات الاختبار .

- التجزئة النصفية لعبارات الاختبار:

وللتأكد من ثبات الاختبار تم تجزئة فقراته إلى فقرات فردية وأخرى زوجية ثم حساب معامل ارتباط بيرسون وتصحيح ذلك من خلال معاملي سبيرمان وجتمان للتجزئة النصفية، ويوضح جدول (٥) معاملات الارتباط .

جدول (٥)

معاملات التجزئة النصفية

لثبات اختبار الاختبارات النفسية وابعاده (ن = ٣٠)

الايعاد	معامل سبيرمان	معامل جتمان	الدلالة
أسئلة الصواب والخطأ	٠.٩٢٤	٠.٩٢٠	٠.٠١
أسئلة الاختيار من متعدد	٠.٨٩٥	٠.٩١٠	٠.٠١
المقياس ككل	٠.٩١٦	٠.٩١٦	٠.٠١

يتضح من جدول (٥) أن معاملات الارتباط دالة عند مستوى ٠.٠١، وذلك يؤكد على ثبات الاختبار.

وللتأكد من اتساق الاختبار داخلياً قام الباحث بحساب معاملات الارتباط بين درجة كل عبارة من عبارات الاختبار ودرجة الاختبار الكلية بعد تطبيقه على العينة الاستطلاعية، ويوضح جدول (٦) معاملات الارتباط.

جدول (٦)

معاملات الارتباط بين درجة كل فقرة من فقرات الاختبار

لطلاب كلية التربية ودرجة الصورة الكلية (ن = ٣٠) عدد مفردات الاختبار = ٢٤

الفقرة	معامل الارتباط	الفقرة	معامل الارتباط	الفقرة	معامل الارتباط
١	**٠.٨٣٢	٩	**٠.٨٣٢	١٧	**٠.٨٩١
٢	**٠.٧٤٥	١٠	**٠.٨٢٣	١٨	**٠.٧٨٩
٣	**٠.٧٨٩	١١	**٠.٧٨٩	١٩	**٠.٧٤٨
٤	**٠.٧٨٩	١٢	**٠.٧٥٩	٢٠	**٠.٨١١
٥	**٠.٨٢٠	١٣	**٠.٧٥٣	٢١	**٠.٧٨٨
٦	**٠.٨١٥	١٤	**٠.٧٤٥	٢٢	**٠.٨١١
٧	**٠.٨٤٥	١٥	**٠.٨٩٢	٢٣	**٠.٨٨١
٨	**٠.٩١٠	١٦	**٠.٧٤٩	٢٤	**٠.٧٦٩

* * دال عند مستوى ٠.٠١

يتضح من جدول (٦) أن معاملات الارتباط بين درجة كل مفردة والدرجة الكلية للاختبار كلها دالة عند مستوى ٠.٠١، وذلك يؤكد ثبات الاختبار، وبهذا يصبح الاختبار في صورته النهائية يتكون من (٢٤) مفردة.

ج- فاعلية الفقرات:

للتأكد من فاعلية فقرات الاختبار، تم حساب معاملات السهولة والصعوبة والتمييز، لكل فقرة من فقرات الاختبار، وجدول (٧) يوضح ذلك.

جدو (٧)

معاملات السهولة و الصعوبة والتمييز لفقرات الاختبار (ن = ٣٠) عدد

مفردات الاختبار = ٢٤

م	معامل السهولة	معامل الصعوبة	معامل التمييز	م	معامل السهولة	معامل الصعوبة	معامل التمييز
١	٠.٦٥	٠.٣٥	٠.٢٣	١٣	٠.٦٤	٠.٣٦	٠.٢٣
٢	٠.٥٩	٠.٤١	٠.٢٤	١٤	٠.٦٢	٠.٣٨	٠.٢٤
٣	٠.٦٤	٠.٣٦	٠.٢٣	١٥	٠.٧٠	٠.٣٠	٠.٢١
٤	٠.٥٥	٠.٤٥	٠.٢٥	١٦	٠.٧٣	٠.٢٧	٠.٢٠
٥	٠.٦٨	٠.٣٢	٠.٢٢	١٧	٠.٥٥	٠.٤٥	٠.٢٥
٦	٠.٦٤	٠.٣٦	٠.٢٣	١٨	٠.٦٤	٠.٣٦	٠.٢٣
٧	٠.٦٤	٠.٣٦	٠.٢٣	١٩	٠.٥٠	٠.٥٠	٠.٢٥
٨	٠.٧١	٠.٢٩	٠.٢١	٢٠	٠.٦٦	٠.٣٤	٠.٢٢
٩	٠.٥٩	٠.٤١	٠.٢٤	٢١	٠.٥٨	٠.٤٢	٠.٢٤
١٠	٠.٦٤	٠.٣٦	٠.٢٣	٢٢	٠.٥٧	٠.٤٣	٠.٢٥
١١	٠.٦٦	٠.٣٤	٠.٢٢	٢٣	٠.٥٩	٠.٤١	٠.٢٤
١٢	٠.٦٨	٠.٣٢	٠.٢٢	٢٤	٠.٧٣	٠.٢٧	٠.٢٠

يتضح من جدول (٧) أن معاملات الصعوبة تراوحت بين ٠.٢٧ إلى ٠.٥٠، ومعاملات التمييز تراوحت بين ٠.٢٠ إلى ٠.٢٥، وأشار كلاً من عماد حسن ومحمد رياض وعلي صلاح (٢٠١٥)، ومحمود عبد الحليم (١٩٩٤، ١٧٥) إلى أن قيمة معامل التمييز تتراوح بين ١ - : ١ وعادة ما يتم اختيار المفردات التي تزيد معاملات تمييزها عن ٠.٢٠ وجميعها قيم مقبولة وتشير إلى فاعلية جيدة لفقرات الاختبار.

سابعاً- الخطوات الإجرائية للدراسة:

١- تم تصميم اختبار لمقرر الاختبارات النفسية للفرقة الرابعة شعبة معلمي علم النفس، وتم إعداد صورتين من هذا الاختبار صورة جيدة وأخرى غير جيدة، كل صورة تتكون من ٢٤ فقرة.

٢- تم تطبيق الصورتين على طلاب وطالبات الفرقة الرابعة شعبة معلمي علم النفس بكلية التربية، وذلك بهدف تقنين الصورتين على العينة الاستطلاعية.

- ٣- تم تطبيق صورتى الاختبار النهائيتين -كل صورة تتكون من ٢٤ فقرة - على العينة الأساسية، وأصبح لكل طالب وطالبة درجة كلية على إحدى الصورتين.
- ٤- تم الحصول على الدرجات الخام بعد تصحيح الصورتين الاختباريتين.
- ٥- تم الاستعانة ببعض الاختبارات الموضوعية من أقسام وكليات مختلفة بالجامعة.
- ٦- تم تحليل الاختبارات وفقا لقائمة تحديد الأخطاء المحكمة.
- ٧- تم تفسير ومناقشة نتائج الدراسة في ضوء الإطار النظري والدراسات السابقة.
- ٨- تم إقتراح مجموعة من التوصيات المستقبلية.

ثامناً - الأساليب الإحصائية المستخدمة:

- تم استخدام عدداً من الأساليب الإحصائية، سواء أثناء التقنين بغرض حساب الصدق والثبات والاتساق الداخلي أو أثناء معالجة البيانات، وهذه الأساليب هي:
- ١- معامل كيبودر ريتشاردسون.
 - ٢- معادلة سبيرمان وبراون لحساب ثبات التجزئة النصفية.
 - ٣- معامل الصعوبة / السهولة.
 - ٤- معامل التمييز.
 - ٥- معامل بيرسون لحساب الاتساق الداخلي للاختبار.
 - ٦- تم استخدام مجموعة من البرامج الإحصائية لحساب الأساليب الإحصائية السابقة هي:
برنامج SPSS وتم استخدامه في حساب الاتساق الداخلي ومعاملات الارتباط واختبارت لحساب الفروق بين صورتى الاختبار.
برنامج Excel وتم استخدامه لحساب معاملات الصعوبة والتمييز للصورتين الاختباريتين.

نتائج البحث وتفسيرها

وينص السؤال الأول على " ما هي الأخطاء الشائعة في بناء مفردات الاختبارات الموضوعية لدى أعضاء هيئة التدريس بكلية التربية جامعة أسيوط؟

وللإجابة عن هذا السؤال تم فحص مجموعة عشوائية من الاختبارات الموضوعية لأعضاء هيئة التدريس بكليات (الآداب، التمريض، التربية الرياضية، التربية النوعية، رياض الأطفال، العلوم) وقسمي (المناهج وطرق التدريس، أصول التربية) ويوضح جدول (٨) أعداد الاختبارات التي تم اختيارها من كل كلية وقسم.

جدول (٨)

بعض نماذج الاختبارات الموضوعية لأعضاء هيئة التدريس بكلية التربية جامعة أسيوط

تكرار الاختبارات الموضوعية	الكلية أو القسم
٥	كلية الآداب
١	كلية التمريض
١	كلية التربية الرياضية
٢	كلية التربية النوعية
٢	كلية رياض الأطفال
١	كلية العلوم
٢	قسم المناهج وطرق التدريس
٢	قسم أصول التربية
١٦	المجموع

ويوضح جدولاً (٩، ١٠) أهم الأخطاء التي تكررت في الاختبارات الموضوعية التي أعدها أعضاء هيئة التدريس بكليات (الآداب، التمريض، التربية الرياضية، التربية النوعية، رياض الأطفال، العلوم) وقسمي (المناهج وطرق التدريس، أصول التربية).

جدول (٩)

المحور الأول- الأخطاء الشائعة في بناء مفردات الصواب والخطأ:

الخطأ في بناء مفردات الصواب والخطأ				لا توجد		توجد	
الرقم	الوصف	ك	%	ك	%	ك	%
٢	صياغة الفقرة غير واضحة وتؤدي إلى التخمين.	١٢	٧٥%	٤	٢٥%		
٣	تتطلب الإجابة على الفقرة أكثر من هدف.	١٤	٨٨%	٢	١٣%		
٤	غموض المفردات وعدم دقتها.	١٥	٩٤%	١	٦%		
٥	استخدام العبارات العامة الواسعة.	١٥	٩٤%	١	٦%		
٦	طول أو قصر بعض المفردات عن غيرها.	١٥	٩٤%	١	٦%		
٧	الوزن النسبي في توزيع مفردات الصواب والخطأ غير عادل.	١٢	٧٥%	٤	٢٥%		
٨	توجد أخطاء إملائية أو نحوية تغير من معنى الفقرة.	١	٦%	١٥	٩٤%		
٩	اقتصار المفردات على بعض الوحدات والفصول دون غيرها.	١٥	٩٤%	١	٦%		
١٠	استخدام صيغة السؤال لصياغة الفقرة بدلا من العبارات التقريرية.	٢	١٣%	١٤	٨٨%		
١١	ال فقرات لا تشمل على غالبية المحتوى (المقرر).	١٤	٨٨%	٢	١٣%		

يتضح من جدول (٩) أن:

- نسبة الأخطاء الإملائية واستخدام صيغة السؤال في الفقرات كانت أقل من ٢٠ %.
- باقي الأخطاء بالمفردات كانت موجودة بنسب أكبر من ٢٠ %، حيث تراوحت هذه النسب ما بين (٢٥% إلى ٩٤ %) وهي نسب كبيرة توضح مدى الأخطاء التي يقع فيها الباحثون عند صياغة مفردات الصواب والخطأ.

وأظهرت نتائج البحث أن هناك أخطاء تسعة أخطاء شائعة في بناء الاختبارات

الموضوعية - أسئلة الصواب والخطأ - وهي:

- غموض المفردات وعدم دقتها، واستخدام العبارات العامة الواسعة، وطول أو قصر بعض المفردات عن غيرها، واقتصار المفردات على بعض الوحدات والفصول دون غيرها، وهذه الأخطاء بلغت نسبة وجودها (٩٤ %).
- تتطلب الإجابة على الفقرة أكثر من هدف، والفقرات لا تشتمل على غالبية المحتوى (المقرر)، وهذه الأخطاء بلغت نسبة وجودها (٨٨ %).
- صياغة الفقرة غير واضحة وتؤدي إلى التخمين، والوزن النسبي في توزيع مفردات الصواب والخطأ غير عادل، وهذه الأخطاء بلغت نسبة وجودها (٧٥ %).
- الفقرة مصاغة بصورة النفي بلغت نسبة وجودها (٢٥ %).

وهذه النتائج تتفق مع ما توصلت إليه نتائج الدراسات التالية:

- دراسة مروان الصمادي (٢٠٠٦)، ودراسة سناء أحمد (٢٠١١)، ودراسة بندر الأمير (٢٠١١)، ودراسة محمد عبد الخالق (٢٠٠٧، ٢٠١١)، ودراسة رانية بنت بكر (٢٠١٢).

جدول (١٠)

المحور الثاني - الأخطاء الواردة في مفردات الاختيار من متعدد:

المجلة العلمية لكلية التربية - جامعة اسيوط

الخاصية	الخطأ في بناء مفردات الاختيار من متعدد					
	توجد		لا توجد			
	ك	%	ك	%		
مدخل أو مقدمة السؤال	١	صياغة مدخل الفقرة غير واضح.	١٢	٧٥%	٤	٢٥%
	٢	الفقرة تناقش أكثر من هدف.	١٦	١٠٠%	٠	٠%
	٣	غموض المفردات وعدم دقتها.	١٤	٨٨%	٢	١٣%
	٤	استخدام العبارات العامة المبهمة.	١٦	١٠٠%	٠	٠%
	٥	استعمال صيغة النفي في بناء الفقرة.	٩	٥٦%	٧	٤٤%
	٦	طول أو قصر بعض المفردات عن غيرها.	١٥	٩٤%	١	٦%
	٧	توجد أخطاء إملائية أو نحوية تغير من معنى الفقرة.	٠	٠%	١٦	١٠٠%
	٨	اقتصار المفردات على بعض الوحدات والفصول دون غيرها.	١٦	١٠٠%	٠	٠%
	٩	وجود تلميحات بالإجابة في مدخل الفقرة.	١٥	٩٤%	١	٦%
	١٠	استخدام صيغة السؤال لصياغة الفقرة بدلا من العبارات التقريرية.	٢	١٣%	١٤	٨٨%
البدائل أو الموهبات	١١	البدائل غير مرتبة منطقياً أو عددياً أو أبجدياً.	١٥	٩٤%	١	٦%
	١٢	تكرار المفردات في بدائل الإجابة.	١١	٦٩%	٥	٣١%
	١٣	عدم تجانس البدائل.	١٦	١٠٠%	٠	٠%
	١٤	عدم اتساق البدائل طولاً أو قرصاً.	١٦	١٠٠%	٠	٠%
	١٥	الفقرة تخلو من البديل " لا شئ مما سبق " كإجابة صحيحة.	١٢	٧٥%	٤	٢٥%
	١٦	الفقرة تخلو من البديل " كل ما سبق " كإجابة صحيحة.	١٠	٦٣%	٦	٣٨%
	١٧	الفقرة تخلو من البديل " لا أعرف " كإجابة صحيحة.	٠	٠%	١٦	١٠٠%
	١٨	تتضمن البدائل مؤشرات -إيماءات- نحو الإجابة الصحيحة.	١٥	٩٤%	١	٦%
	١٩	الموهبات تخلو من الالفاظ المحددة مثل " قطعاً"، " دائماً"، " فقط".	١٣	٨١%	٣	١٩%
	٢٠	تتساوى الموهبات في جاذبيتها ظاهرياً.	٢	١٣%	١٤	٨٨%
	٢١	استخدام صيغة النفي المزدوج.	٠	٠%	١٦	١٠٠%
	٢٢	تكرار بعض الكلمات في البدائل.	١٢	٧٥%	٤	٢٥%
	٢٣	عدم مناسبة عدد البدائل للفقرة.	١٦	١٠٠%	٠	٠%
	٢٤	البدائل في نفس سطر مدخل السؤال ولا تتفصل عنه.	٨	٥٠%	٨	٥٠%
٢٥	اقتصار البدائل على ثلاثة بدائل فقط أو بديلين أحياناً.	١٣	٨١%	٣	١٩%	
الإجابة الصحيحة	٢٦	الإجابة الصحيحة موجودة في بديلين أو أكثر.	١٦	١٠٠%	٠	٠%
	٢٧	عدم توزيع الإجابة الصحيحة توزيعاً عشوائياً.	١٦	١٠٠%	٠	٠%
	٢٨	التكرار اللفظي بين مدخل السؤال والبديل الصحيح.	١٣	٨١%	٣	١٩%
	٢٩	طول أو قصر الإجابة الصحيحة مما يوحي بالإجابة.	١٥	٩٤%	١	٦%
	٣٠	تجانس الإجابة الصحيحة مع باقي البدائل.	١	٦%	١٥	٩٤%
	٣١	بديل واحد فقط يمثل الإجابة الصحيحة أو الأكثر صحة.	٠	٠%	١٦	١٠٠%

يوضح من جدول (١٠) أن:

- نسبة استخدام صيغة السؤال لصياغة الفقرة بدلا من العبارات التقريرية كانت أقل من ٢٠%.

- نسبة استخدام صيغة النفي المزدوج كانت أقل من ٢٠ %.
- باقي الأخطاء بالمفردات كانت موجودة بنسب أكبر من ٢٠ %، حيث تراوحت هذه النسب ما بين (٢٥% إلى ١٠٠%) موزعة ما بين مدخل الفقرة والبدائل والاجابات الصحيحة، وهي نسب كبيرة توضح مدى الأخطاء التي يقع فيها الباحثون عند صياغة مفردات الاختيار من متعدد.

وأظهرت نتائج البحث أن هناك ثمانية أخطاء شائعة في مدخل أو مقدمة السؤال وهي:

- الفقرة تناقش أكثر من هدف، واقتصار المفردات على بعض الوحدات والفصول دون غيرها وهذه الأخطاء بلغت نسبة وجودها (١٠٠%).
- طول أو قصر بعض المفردات عن غيرها، و وجود تلميحات بالاجابة في مدخل الفقرة، وهذه الأخطاء بلغت نسبة وجودها (٩٤%).
- غموض المفردات وعدم دقتها، وهذا الخطأ بلغت نسبة وجوده (٨٨%).
- صياغة مدخل الفقرة غير واضح، وهذا الخطأ بلغت نسبة وجوده (٧٥%).
- صياغة مدخل الفقرة غير واضح، وهذا الخطأ بلغت نسبة وجوده (٧٥%).
- استعمال صيغة النفي في بناء الفقرة، وهذا الخطأ بلغت نسبة وجوده (٥٦%).

وهناك ثلاثة عشر خطأ شائعاً في البدائل أو المموهات وهي:

- عدم تجانس البدائل، وعدم اتساق البدائل طولاً أو قصراً، وعدم مناسبة عدد البدائل للفقرة، وهذه الأخطاء بلغت نسبة وجودها (١٠٠%).
- البدائل غير مرتبة منطقياً أو عددياً أو أبجدياً، وتتضمن البدائل مؤشرات -إيماءات- نحو الاجابة الصحيحة، وهذتن الخطأن بلغت نسبتا وجودهما (٩٤%).
- تتساوى المموهات في جاذبيتها ظاهرياً، وهذا الخطأ بلغت نسبة وجوده (٨٨%).
- اقتصار البدائل على ثلاثة بدائل فقط أو بديلين أحياناً، والمموهات تخلو من الالفاظ المحددة مثل "قطعاً"، "دائماً"، "فقط"، وهذان الخطأن بلغت نسبة وجودهما (٨١%).
- الفقرة تخلو من البديل " لا شئ مما سبق " كإجابة صحيحة، وتكرار بعض الكلمات في البدائل وهذان الخطأن بلغت نسبة وجودهما (٧٥%).

- تكرار المفردات في بدائل الإجابة، وهذا الخطأ بلغت نسبة وجوده (٦٩ %).
- الفقرة تخلو من البديل " كل ما سبق " كإجابة صحيحة، وهذا الخطأ بلغت نسبة وجوده (٦٣ %).
- البدائل في نفس سطر مدخل السؤال ولا تتفصل عنه، وهذا الخطأ بلغت نسبة وجوده (٥٠ %).

وهناك ستة أخطاء شائعة في الإجابة الصحيحة وهي:

- الاجابة الصحيحة موجودة في بديلين أو أكثر، وعدم توزيع الاجابة الصحيحة توزيعاً عشوائياً، وبديل واحد فقط يمثل الاجابة الصحيحة أو الأكثر صحة، وهذه الأخطاء بلغت نسبة وجودها (١٠٠ %).
- طول أو قصر الاجابة الصحيحة مما يوحي بالاجابة، وتجانس الاجابة الصحيحة مع باقي البدائل وهذان الخطآن بلغت نسبة وجودهما (٩٤ %).
- التكرار اللفظي بين مدخل السؤال والبديل الصحيح، وهذا الخطأ بلغت نسبة وجوده (٨١ %).

وهذه النتائج تتفق مع ما توصلت إليه نتائج الدراسات التالية:

دراسة سمير فؤاد (٢٠٠٧)، ودراسة هدى العريفي (٢٠٠٨)، ودراسة عبد الله الغامدي (٢٠٠٨)، ودراسة زياد ملياني (٢٠١٠)، ودراسة ماجد بخيت (٢٠١١)، ودراسة سناء أحمد (٢٠١١)، ودراسة بسام محمود (٢٠١١)، ودراسة بندر الأمير (٢٠١١)، ودراسة حيدر ظاظا (٢٠١٢)، ودراسة كلا من إبراهيم محمد وياسل خميس (٢٠١٢)، ودراسة إبراهيم الرباعي (٢٠١٢)، ودراسة سجي أحمد (٢٠١٣)، ودراسة أيمن فريحات (٢٠١٤)، ودراسة مايسة أبو مسلم (٢٠١٦).

ويرجح الباحث أن السبب الرئيس وراء كل هذه الأخطاء في بناء الاختبارات الموضوعية هو عدم اهتمام كثير من أعضاء هيئة التدريس بالخطوات العلمية لبناء الاختبارات ومراعاة شروط وقواعد بناء الاختبارات الموضوعية، مع عدم الاهتمام ببناء جدول مواصفات لكل اختبار يساعد في عمل بنك أسئلة مما يؤثر على فقر البدائل وتكرارها.

وينص السؤال الثاني على " هل هناك فروق دالة إحصائياً بين الاختبار الجيد) جميع فقراته سليمة البنية) والاختبار غير الجيد) فقراته تحتوي على الأخطاء الشائعة في بناء الاختبارات الموضوعية) في الخصائص السيكومترية للاختبار وفقراته في مادة الاختبارات النفسية لطلاب الفرقة الرابعة شعبة علم النفس؟

وللإجابة عن هذا السؤال تم تطبيق صورتني الاختبار على طلاب الفرقة الرابعة شعبة علم النفس، ثم تم حساب درجة كل طالب وطالبة على هاتين الصورتين ومن خلال درجات الطلاب تمت معرفة الفروق بين درجات الطلاب على الصورتين من خلال التحقق من هذه الفروض:

أ- الفرض الفرعي الأول، وينص على " لا توجد فروق ذات دلالة إحصائية بين عند مستوى دلالة ٠.٠٥ بين متوسطات معاملات السهولة لمفردات الاختبار الجيد ومفردات الاختبار غير الجيد".

ب- الفرض الفرعي الثاني، وينص على " لا توجد فروق ذات دلالة إحصائية بين عند مستوى دلالة ٠.٠٥ بين متوسطات معاملات الصعوبة لمفردات الاختبار الجيد ومفردات الاختبار غير الجيد".

ج- الفرض الفرعي الثالث، وينص على " لا توجد فروق ذات دلالة إحصائية بين عند مستوى دلالة ٠.٠٥ بين متوسطات معاملات التمييز لمفردات الاختبار الجيد ومفردات الاختبار غير الجيد".

وللتحقق من صحة هذا الفروض الفرعية تم حساب الفروق بين متوسطات معاملات السهولة والصعوبة والتمييز لمفردات الصورتين الاختباريتين الجيدة وغير الجيدة باستخدام اختبار ت من خلال البرنامج الإحصائي Spss، وجدول (١١) يوضح هذه الفروق.

جدول (١١)

يوضح المتوسط الحسابي والانحراف المعياري وقيمة (ت) لنتائج تطبيق صورتني الاختبار على طلاب الفرقة الرابعة شعبة علم النفس (ن= ٣٨)

مستوى الدلالة	قيمة (ت) المحسوبة	الصورة غير الجيدة		الصورة الجيدة		المعاملات
		ع	م	ع	م	
دالة عند ٠.٠١	٤.١٩	٠.٢٨٣٠	٠.٤٦١٥	٠.٠٥٩٩	٠.٦٠٢٦	معامل السهولة
دالة عند ٠.٠١	٤.٢٠	٠.٢٨٣٠	٠.٥٣٨٥	٠.٠٥٩٩	٠.٣٩٧٤	معامل الصعوبة
دالة عند ٠.٠١	١٦.١٨	٠.٠٣٠٥	٠.١٧٢٦	٠.٠١٤٧	٠.٢٣٦٣	معامل التمييز

يتضح من جدول (١١) ما يلي:

- ١- توجد فروق ذات دلالة إحصائية عند مستوى دلالة ٠.٠١ بين متوسطات معاملات السهولة للصورتين الاختباريتين لصالح متوسط معاملات الصورة الجيدة.
 - ٢- توجد فروق ذات دلالة إحصائية عند مستوى دلالة ٠.٠١ بين متوسطات معاملات الصعوبة للصورتين الاختباريتين لصالح متوسط معاملات الصورة غير الجيدة.
 - ٣- توجد فروق ذات دلالة إحصائية عند مستوى دلالة ٠.٠١ بين متوسطات معاملات التمييز للصورتين الاختباريتين لصالح متوسط معاملات الصورة الجيدة.
 - د- الفرض الفرعي الرابع، وينص على " لا توجد فروق ذات دلالة إحصائية بين عند مستوى دلالة ٠.٠٥ بين متوسطات معاملات الثبات لمفردات الاختبار الجيد ومفردات الاختبار غير الجيد".
 - هـ- الفرض الفرعي الخامس، وينص على " لا توجد فروق ذات دلالة إحصائية بين عند مستوى دلالة ٠.٠٥ بين متوسطات معاملات الصدق لمفردات الاختبار الجيد ومفردات الاختبار غير الجيد".
- وللتحقق من صحة هذين الفرضين الفرعيين تم حساب الفروق بين متوسطات معاملات الثبات والصدق لمفردات الصورتين الاختباريتين الجيدة وغير الجيدة باستخدام اختبار ت من خلال البرنامج الإحصائي Spss، وجدول (١٢) يوضح هذه الفروق.

جدول (١٢)

يوضح المتوسط الحسابي والانحراف المعياري وقيمة (ت) لنتائج تطبيق صورتين الاختبار على طلاب الفرقة الرابعة شعبة معلمي علم النفس (ن= ٣٨)

مستوى الدلالة	قيمة (ت) المحسوبة	الصورة غير الجيدة		الصورة الجيدة		المعاملات
		ع	م	ع	م	
دالة عند ٠.٠١	٨٣.٥٦	٠.٠٤٥١	٠.١٥٩٦	٠.٠٤٩٣	٠.٨٠٨٧	معامل الثبات
دالة عند ٠.٠١	٣٣.٥١	٠.١٠٤٦	٠.٢٣٠٠	٠.٠٨٨٤	٠.٧٦٣٦	معامل الصدق

يتضح من جدول (١٢) ما يلي:

١- توجد فروق ذات دلالة إحصائية عند مستوى دلالة ٠.٠١ بين متوسطات معاملات الثبات للصورتين الاختباريتين لصالح متوسط معاملات الصورة الجيدة.

٢- توجد فروق ذات دلالة إحصائية عند مستوى دلالة ٠.٠١ بين متوسطات معاملات الصدق للصورتين الاختباريتين لصالح متوسط معاملات الصورة الجيدة.

وتتفق هذه النتائج مع نتائج الدراسات التالية:

دراسة هدى العريفي (٢٠٠٨)، ودراسة عبد الله الغامدي (٢٠٠٨)، ودراسة زياد ملياني (٢٠١٠)، ودراسة ماجد بخيت (٢٠١١)، ودراسة سناء أحمد (٢٠١١)، ودراسة بسام محمود (٢٠١١)، ودراسة بندر الأمير (٢٠١١)، ودراسة حيدر ظاظا (٢٠١٢)، ودراسة كلا من إبراهيم محمد وباسل خميس (٢٠١٢)، ودراسة إبراهيم الرباعي (٢٠١٢)، ودراسة سجي أحمد (٢٠١٣)، ودراسة ماييسة أبو مسلم (٢٠١٦).

ويرجح الباحث أن السبب الرئيس وراء كل هذه الأخطاء في بناء الاختبارات الموضوعية وضعف صدق وثبات الصور غير الجيدة هو عدم اهتمام كثير من أعضاء هيئة التدريس بالخطوات العلمية لبناء الاختبارات ومراعاة شروط وقواعد بناء الاختبارات الموضوعية، مع عدم الاهتمام ببناء جدول مواصفات لكل اختبار يساعد في عمل بنك أسئلة مما يؤثر على ضعف صدق وثبات الاختبارات، مما يجعلنا نشك في نتائج هذه الاختبارات، فمن المعلوم أن من الشروط الأساسية لكي نحكم على جودة الاختبار هو صدقه وثباته.

توصيات البحث:

في ضوء نتائج البحث الحالي يمكن تقديم هذه التوصيات:

- ✓ ضرورة مراعاة شروط وقواعد إعداد الاختبارات والالتزام بمواصفات الورقة الامتحانية من قبل واضعي الامتحانات.
- ✓ ضرورة مراعاة شروط وقواعد إعداد الاختبارات الموضوعية بشتى أنواعها.
- ✓ يجب أن ينوع أعضاء هيئة التدريس من أشكال الاختبارات الموضوعية وعدم الاقتصار فقط على مفردات الصواب والخطأ والاختيار من متعدد وأحيانا المزوجة، فهناك أنواع أخرى يمكن الاستعانة بها.
- ✓ عقد ورش عمل وتدريبات عملية -على أيدي متخصصين في مجال القياس والتقويم- لأعضاء هيئة التدريس عن كيفية تصميم الاختبارات الموضوعية بصورة سليمة.
- ✓ إنشاء وحدة أو مركز للقياس النفسي والتقويم التربوي بكلية التربية يختص بنشر ثقافة إعداد الاختبارات الموضوعية بمختلف أشكالها.

قائمة المراجع

أولاً : المراجع العربية :

إبراهيم محمد الرباعي(٢٠١٢). أثر عدد البدائل وتغيير موقع المموه القوي في فقرات اختبار الاختيار من متعدد على الخصائص السيكومترية للاختبار ومعالم الفقرات وقدرات الأفراد. رسالة ماجستير، كلية التربية، جامعة اليرموك، الأردن.

إبراهيم محمد يعقوب، باسل خميس أبو فودة(٢٠١٢). أثر مخالفة قواعد صياغة فقرات الاختيار من متعدد على الخصائص السيكومترية للاختبار وفقراته. مجلة جامعة دمشق، ١(٢٨)، ص ص ٤١٩ - ٤٤٣.

أحمد محمد عبد الخالق(٢٠٠٠). قياس الشخصية. الإسكندرية: دار المعرفة الجامعية.

أيمن محمد فريحات(٢٠١٤). أثر عدد البدائل في دقة تقدير معالم فقرات الاختبار وفقاً لنماذج الاستجابة للفقرة. رسالة دكتوراه، كلية التربية، جامعة اليرموك، الأردن.

بسام محمود عبد الله(٢٠١١). أثر عدد البدائل في الأسئلة الموضوعية على الخصائص السيكومترية للفقرة والاختبار. رسالة ماجستير، كلية التربية، جامعة مؤتة، الأردن.

بشرى إسماعيل(٢٠٠٤). المرجع في القياس النفسي. القاهرة: الأنجلو المصرية.

بشير صالح الرشيد(٢٠٠٠). مناهج البحث التربوي " رؤية تطبيقية مبسطة ". الكويت: دار الكتاب الحديث.

بندر سالم الأمير(٢٠١١). الأخطاء الشائعة في بناء مفردات اختبار الاختيار من متعدد وتأثيرها على الخصائص السيكومترية للاختبارات التحصيلية لدى معلمي الرياضيات بالمرحلة المتوسطة بمدينة مكة المكرمة. رسالة ماجستير، كلية التربية، جامعة أم القرى، السعودية.

حيدر ظاظا(٢٠١٢). الكشف عن مدى انتهاك قواعد صياغة فقرة الاختيار من متعدد في أسئلة امتحان شهادة الدراسة الثانوية العامة في الأردن. المجلة الأردنية في العلوم التربوية، ٨(١)، ص ص ٨١ - ٩١.

راشد حماد الدوسري (٢٠٠١). معادلة الاختبارات مفهومها وطرقها ومشكلات تطبيقها. مجلة العلوم التربوية والنفسية ، ١(٤)، ص ص ١٠٧ - ١٤١.

رانية بنت بكر بالحر(٢٠١٢). الأخطاء الشائعة في بناء مفردات الصواب والخطأ وعلاقتها ببعض الخصائص السيكومترية للاختبارات والمفردات المعدة من قبل

- بعض المعلمات في مقرر مادة علم النفس لطالبات المرحلة الثانوية بمدينة جدة . رسالة ماجستير، كلية التربية، جامعة أم القرى، السعودية.
- رجاء محمود علام (٢٠١١). مناهج البحث في العلوم النفسية والتربوية. القاهرة: دار النشر للجامعات.
- زياد حسين ملياني(٢٠١٠). مقارنة أثر اختلاف عدد البدائل في اختبار الاختيار من متعدد على صعوبة الفقرة وقدرة الفرد في ضوء كل من النظرية الكلاسيكية ونموذج راش. رسالة دكتوراه، كلية التربية، جامعة أم القرى، السعودية.
- سامية محمد جابر (٢٠٠٠). منهجيات البحث الإجتماعي والإعلامي. الإسكندرية: دار المعرفة الجامعية.
- سجى أحمد الذنبيات(٢٠١٣). أثر اختلاف نمط الاختبار على الخصائص السيكمترية للاختبار. رسالة ماجستير، كلية العلوم التربوية، جامعة مؤتة، الأردن.
- سمير فؤاد عيلبوني(٢٠٠٧). أثر عدد البدائل وتمييز الموهبات في فقرات الاختيار من متعدد على التوافق مع نموذج ثلاثي المعلم. رسالة دكتوراه، كلية التربية، جامعة اليرموك، الأردن.
- سناء أحمد المنصور(٢٠١١). الاختبارات الموضوعية لبعض مقررات مرحلة البكالوريوس بكلية التربية لاعداد معلمات الابتدائي: دراسة تحليلية. المجلة العلمية لجامعة الملك فيصل- العلوم الانسانية والادارية، ١٢(١)، ص ص ٢٦١ :٣٢٩.
- صلاح أحمد مراد، وأمينة علي سليمان(٢٠١٢). الاختبارات والمقاييس في العلوم النفسية والتربوية "خطوات إعدادها وخصائصها". القاهرة: دار الكتاب الحديث.
- صلاح الدين محمود علام(٢٠٠٠). تحليل بيانات البحوث النفسية والتربوية الاجتماعية. القاهرة: دار الفكر العربي.
- صلاح الدين محمود علام (٢٠٠٧). القياس والتقويم التربوي في العملية التدريسية. عمان: دار المسيرة.
- صلاح الدين محمود علام(٢٠١١). القياس والتقويم التربوي والنفسي أساسياته وتطبيقاته وتوجهاته المعاصرة. القاهرة: دار الفكر العربي.

عبد الحميد محمد علي(٢٠٠٩). الاتجاهات الحديثة في القياس النفسي والتقييم التربوي. القاهرة: طيبة للنشر.

عبد الله أحمد الغامدي(٢٠٠٨). أثر عدد البدائل في الخصائص السيكومترية للاختبار التحصيلي في الرياضيات. رسالة ماجستير، كلية التربية، جامعة أم القرى، السعودية.

علي ماهر خطاب(٢٠٠١). علم النفس الفارق. القاهرة: الأنجلو المصرية.

علي ماهر خطاب(٢٠٠٣). علم النفس الفارق. ط٣، القاهرة: الأنجلو المصرية.

عماد أحمد حسن علي ومحمد رياض أحمد وعلى صلاح عبد المحسن (٢٠١٥). تقييم الأساليب الإحصائية المستخدمة في معالجة الفروض في بعض الرسائل العلمية بالأقسام التربوية بكلية التربية - جامعة أسيوط. المؤتمر العلمي الثالث لشباب الباحثين بكلية التربية- جامعة أسيوط.

عماد رمضان البطاح(٢٠٠٧). أثر طريقة اختيار فقرات الاختبار في خصائصه السيكومترية. رسالة ماجستير، كلية العلوم التربوية، جامعة مؤتة، الأردن.

قاسم علي الصراف(٢٠٠٢). القياس والتقييم في التربية والتعليم. القاهرة: دار الكتاب الحديث.

ماجد بخيت العطوي(٢٠١١). أثر عدد المموهات الجذابة لفقرات الاختبار من متعدد في الخصائص السيكومترية للفقرات والاختبار. رسالة ماجستير ، كلية العلوم التربوية، جامعة مؤتة، الأردن.

مايسة فاضل أبو مسلم(٢٠١٦). أثر كل من عدد البدائل وترتيبها لمفردات اختبار تحصيل دراسي من نوع الاختيار من متعدد على الخصائص السيكومترية للاختبار ومفرداته. مجلة اتحاد الجامعات العربية للتربية وعلم النفس- سوريا، ١٤(٢)، ص ص ١٥٤ : ١٨٧.

مروان علي الصمادي(٢٠٠٦). فاعلية طرق تصحيح اختبار الصواب- الخطأ المتعدد وتأثيرها على دقة معادلة الاختبار باستخدام نماذج النظرية الحديثة للقياس. رسالة دكتوراه ، كلية التربية، جامعة اليرموك ، الأردن.

محمد السيد علي(٢٠٠٠). علم المناهج" الأسس والتنظيمات في ضوء الموديلات". القاهرة: دار الفكر العربي.

محمد عبد الخالق محمد(٢٠٠٧). الاختبارات الموضوعية في برامج تعليم اللغة العربية للناطقين بلغات أخرى. مجلة العربية للناطقين بغيرها- معهد اللغة العربية بجامعة أفريقيا العالمية- السودان، (٥)، ص ص ٢٧١ : ٢٧٤.

محمد عبد الخالق محمد(٢٠١١). أوجه القصور في الاختبارات الموضوعية. مجلة العربية للناطقين بغيرها- معهد اللغة العربية بجامعة أفريقيا العالمية- السودان، (١١)، ص ص ١ : ٤٦.

محمود عبد الحليم منسي(١٩٩٤). القياس والإحصاء النفسي والتربوي. الإسكندرية: دار المعارف.

مصطفى باهي، وفاتن النمر(٢٠٠٩). التقويم في مجال العلوم التربوية والنفسية " مبادئ، ونظريات، وتطبيقات". القاهرة: الأنجلو المصرية.

موسى النبهان(٢٠٠٤). أساسيات القياس في العلوم السلوكية. عمان: دار الشروق.

نضال كمال محمد الشريفي، إيمان صالح صلاح طعمنة(٢٠٠٩). أثر عدد البدائل في اختبار الاختيار من متعدد في تقديرات القدرة للأفراد والخصائص السيكومترية للفقرات والاختبار وفق نموذج راش في نظرية الاستجابة للفقرة. المجلة الأردنية في العلوم التربوية، جامعة اليرموك، (٤)٥، ص ص ٣٠٩ - ٣٣٥.

هدى محمد العريفي(٢٠٠٨). تباين الخصائص السيكومترية لاختبار متعدد البدائل بتباين عدد مفرداته وبعض الخصائص الدافعية للمفحوصين. رسالة ماجستير، كلية التربية، جامعة أم القرى، السعودية.

ثانياً المراجع الأجنبية:

Creswell, J. & Miller, D. (2000). Determining Validity in qualitative inquiry. Theory into Practice, 39(3), 124-131.

Davies, D. & Dodd, J (2002). Qualitative research and the question of rigor. Qualitative Health research, 12(2), 279-289.

- Gholamreza, J. & Fatemeh, S. (2008). Validity, Reliability and Difficulty indices for Instructor-Built Exam Question. Journal of applied quantitative Methods 3(2), 151-155.
- Kimmo, V. (2000). Reliability of Measurement Scales, Tarkkonen's general method supersedes Cronbach's alpha, Department of Statistics, Finland: University of Helsinki.
- Kolen, M. & Brennan, R. (2004). Test equating, scaling, and linking: methods and practices. 2nd Ed. The United States of America, Springer.
- Martijn, G. & Jan-Benedict, E. (2010). Finite Mixture Multilevel Multidimensional Ordinal IRT Models for Large Scale Cross-Cultural Research, The Psychometric Society. This article is published with open access at Springerlink.com.
- Nahid, G. (2003). Understanding Reliability and Validity in Qualitative Research. Canada: University of Toronto, Toronto, Ontario.
- Stenbacka, C. (2001). Qualitative research requires quality concepts of its own. Management Decision, 39(7), 551-555.