أثر كل من عدد البدائل وترتيبها لمفردات اختبار تحصيل دراسي من نوع الاختيار من متعددعلى الخصائص السيكومترية للاختبار ومفرداته.

د. مايسة فاضل أبو مسلم *

الملخص

تعتبر الاختبارات من أهم الأدوات في المجال التربوي حيث تزوّد ببيانات يُستند إليها في اتخاذ العديد من القرارات الهامة التي تخص الفرد والمجتمع.وتتحدد مشكلة الدراسة في محاولة الإجابة عن الأسئلة التالية :هل يختلف صدق و ثبات الاختبار ككل (من نوع الاختيار من متعدد) طبقا لعدد بدائل كل مفردة من مفرداته؟ هل توجد فروق دالة إحصائيا بين متوسطات معاملات الصعوبة والتمييز لمفردات الاختبار (من نوع الاختيار من متعدد) تعزى إلى اختلاف عدد البدائل التي تضّمها كل مفردة من هذه المفردات ؟

هل يختلف صدق و ثبات الاختبار ككل (من نوع الاختيار من متعدد) طبقا لموقع (ترتيب) البديل الصحيح بين بدائل الاختبار ؟.هل توجد فروق دالة إحصائيا بين متوسطات معاملات الصعوبة والتمييز لمفردات الاختبار (من نوع الاختيار من متعدد) تعزى إلى اختلاف موقع (ترتيب) البديل الصحيح بين بدائل كل مفردة ؟وتقسم فرضيات الدراسة إلى مجموعتين: المجموعة الأولى (عدد البدائل) :توجد فروق دالة إحصائيا بين متوسطات معاملات الصعوبة والتمييز لمفردات الاختبار (من نوع الاختيار من متعدد) تعزى إلى اختلاف عدد بدائل كل مفردة من مفردات الاختبار. المجموعة الثانية (ترتيب البدائل) :توجد فروق دالة إحصائيا بين متوسطات معاملات الصعوبة والتمييز لمفردات الاختبار(من نوع الاختيار من متعدد) تعزى إلى اختلاف موقع (ترتيب) البديل الصحيح بين بدائل كل مفردة في الاختبار.

وفي ضوء ما سبق عرضه من نتائج أسفرت عنها الدراسة الحالية ، يمكن إيجاز أهم هذه النتائج في الآتي :– يزداد صدق وثبات الاختبار (من نوع الاختيار من متعدد) ككل عندما يكون الاختبار ثلاثي البدائل عما لو كان أكثر من ذلك (رباعي أو خماسي البدائل .... الخ) .–يزداد صدق وثبات الاختبار (من نوع الاختيار من متعدد) ككل عندما يتوزع البديل الصحيح عشوائيا بين بدائل كل مفردة من مفردات الاختبار .– تكون مفردات الاختبار (من نوع الاختيار من متعدد) أكثر صعوبة كلما ازداد عدد البدائل التي تضّمها المفردة .–تكون مفردات الاختبار (من نوع الاختيار من متعدد) أعلى من حيث مستوى قدرتها التمييزية كلما ازداد عدد البدائل التي تضّمها المفردة .– تكون مفردات الاختبار (من نوع الاختيار من متعدد) أكثر صعوبة عندما يتوزع البديل الصحيح عشوائيا بين بدائل مفردات الاختبار .– تكون مفردات الاختبار (من نوع الاختيار من متعدد) أعلى من حيث مستوى قدرتها التمييزية عندما يتوزع البديل الصحيح عشوائيا بين بدائل مفردات الاختبار .

---

* المركز القومي للامتحانات والتقويم التربوي، أستاذ مساعد، الأردن.

١ – مقدمة

تعتبر الاختبارات من أهم الأدوات في المجال التربوي حيث تزوّد ببيانات يُستند إليها في اتخاذ العديد من القرارات الهامة التي تخص الفرد والمجتمع ، فالاختبار بأبسط صورة هو إجراء منظم لقياس عينة من سلوك الأفراد ، وقد انتشر استخدام الاختبارات انتشاراً واسعاً في العديد من المجالات بوجه عام، فهي تحقق العديد من الأهداف ، حيث تصَمم 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 الصفية وغير الصفية (الشريفين، وطعامنة ، ٢٠٠٩ ، ص ٣٠٩).

وتكتسب عملية تقويم الأفراد أهمية كبيرة بقدر أهمية القرارات المبنية عليها ، وبقدر خطورة القرارات التي يمكن أن تترتب على ذلك في المواقف والمجالات المتعددة على مستوى الفرد والمجتمع إلى حد قد يصعب معالجته ، أو قد يحتاج علاجه إلى وقت طويل مما يعيق عملية التنمية ومواكبة تطور المجتمعات الأخرى المتقدمة ، وكذلك لاتخاذ قرارات أكثر دقة يجب توافر معلومات صادقة ودقيقة من خلال التخطيط الجيد والإعداد الدقيق للاختبار, وهناك العديد من الأشكال للاختبارات لعلها تنقسم إلى نوعين رئيسيين: هما تلك الاختبارات التي تستلزم إنتاج الاستجابة وتُعْرَف بالاختبارات المقالية، وتلك التي تستلزم إنتقاء الاستجابة وتُعْرَف بالاختبارات الموضوعية، والاختبارات الموضوعية تتعدد أشكالها ، فمنها فقرات الإكمال (completion)، وفقرات الصواب أو الخطأ (true or false)، وفقرات المزاوجة أو المقابلة (matching)، وفقرات الاختيار من متعدد (multiple choice) ، وتعد الاختبارات الموضوعية من أكثر أنواع الاختبارات استخداماً في النظام التربوي في السنوات الأخيرة حيث أصبحت أكثر انتشارا عن ذي قبل ، لما تتميز به من سهولة التصحيح ودقته وثباته وموضوعيته، وإمكانية استخدامها في قياس المستويات المعرفية العليا بشكل أوسع عن ذي قبل (الأمير ، ٢٠١٣).

ومما يجدر ذكره أن بناء مفردات الاختبار من نوع الاختيار من متعدد قد حظي باهتمام اختصاصي القياس منذ بدايات القرن الماضي ، إلا أنها أخضعت للتجريب في بداية عام ١٩٢٠ إذ كثر الحديث عن قواعد لصياغتها ، وبات من المهم تطوير قواعد لصياغتها تستند إلى التجريب أكثر من كونها تمثل عُرفاً ، ويعتبرها بعض الباحثين فناً أو عملاً إبداعياً، وأنه من الواجب مراعاة قواعد الصياغة للمفردات و"البدائل" أو

"المموهات" أو "الاختيارات" أو "المشتتات" distracters بحيث يجب مراجعة وفحص كل مفردة وكل بديل أو مشتت من مشتتات المفردة، والتأكد من كل منها، بحيث تكون صياغة المفردة ببدائلها أو مشتتاتها لا تتضمن انتهاكا لقواعد الصياغة. ولاشك أن انتهاك قواعد صياغة وبناء المفردات ومشتتاتها سيكون له أثره على خصائص المفردة، ومن ثم فإنَّ مستوي الأداء على المفردة وعلى الاختبار ككل يتوقف على خصائص المفردة، والاختبار، ومستوى قدرة المستجيب, مما يعني أن أي انتهاك لهذه القواعد من شأنه التأثير السلبي في الخصائص الإحصائية أو السيكومترية للفرد، والذي يؤثر بدوره في أداء الفرد على المفردة والاختبار ككل (ظاظا ، ٢٠١٢).

وفي هذا الاتجاه أشار بعض الباحثين (العضايلة ، ٢٠١١ ، الغامدي ، ٢٠٠٩ ، عيلبوني ، ٢٠٠٧، الشبلي ، ١٩٩٥ ، Trevisan, et al., 1994 &Trevisan& Sax, 1991) إلى أنه عند بناء مثل هذا النوع من الاختبارات سواء كانت اختبارات تحصيل دراسي، أو اختبارات استعدادات وغيرها، قد يثار تساؤل هام عن عدد البدائل التي يجب أن تتضمنها المفردة ، حيث أن زيادة عدد البدائل يؤدي إلى صعوبة المفردة وثبات الاختبار، كما تقل فرصة التخمين، إلا أن زيادة عدد هذه البدائل أكثر من اللازم قد يؤتي بنتيجة عكسية على حد قولهم ، ومن هنا كان لابد من التعرف على أثر عدد البدائل أو الاختيارات في المفردة على خصائصها الإحصائية أو السيكومترية، ومن ناحية أخرى لاحظ بعض الباحثين (الرباعي، ٢٠١٢، العريفى، ٢٠٠٧، الغامدي، ٢٠٠٠) أن ترتيب عرض هذه البدائل قد يكون له أثره أيضاً على هذه الخصائص، وأشارت دراسات أخرى (Gross, 1994 & Fray, 1991) إلى الآثار التي تنتج عندما تتضمن البدائل أو المشتتات بديل "كل ما سبق صحيحاً"، أو "ولا واحد مما سبق صحيح" في حالة ما إذا كان هذا البديل هو اختياراً صحيحاً، وفي حالة ما إذا كان اختياراً خاطئاً.

وانطلاقاً مما سبق جاءت الدراسة الحالية للتعرف على **أثر عدد البدائل** أو المشتتات التي تضمها كل مفردة من مفردات الاختبار على الخصائص الإحصائية أو السيكومترية للاختبار ككل، وعلى معاملات الصعوبة والتمييز لكل مفردة من مفرداته ، وكذلك التعرف على أثر ترتيب عرض هذه البدائل أو المشتتات وموقع (ترتيب) البديل الصحيح بينها، سواء على الخصائص السيكومترية للاختبار ككل، أو على مستويات معاملي الصعوبة والتمييز لكل مفردة من مفرداته.

٢- أهداف الدراسة

تهدف الدراسة بصورة عامة إلى التعرف على الأثر الذي يمكن أن يُحدثه كل من:

٢-١- اختلاف عدد البدائل التي تحويها كل مفردة في الاختبار من نوع الاختيار من متعدد.

٢-٢- اختلاف موقع (ترتيب) البديل الصحيح بين بدائل كل مفردة في هذا الاختبار.

وذلك على كل من:

- الخصائص السيكومترية العامة للاختبار ككل (صدق الاختبار / ثبات الاختبار).
- الخصائص السيكومترية لمفردات الاختبار (معامل الصعوبة/ معامل التمييز للمفردة).

٣- مشكلة الدراسة

تتحدد مشكلة الدراسة في محاولة الإجابة عن الأسئلة التالية:

٣-١- هل يختلف صدق الاختبار ككل (من نوع الاختيار من متعدد) طبقاً لعدد بدائل كل مفردة من مفرداته؟

٣-٢- هل يختلف ثبات الاختبار ككل (من نوع الاختيار من متعدد) طبقاً لعدد بدائل كل مفردة من مفرداته؟

٣-٣- هل توجد فروق بين متوسطات معاملات الصعوبة لمفردات الاختبار (من نوع الاختيار من متعدد) تعزى إلى اختلاف عدد البدائل التي تضّمها كل مفردة من هذه المفردات؟

٣-٤- هل توجد فروق بين متوسطات معاملات التمييز لمفردات الاختبار (من نوع الاختيار من متعدد) تعزى إلى اختلاف عدد البدائل التي تضّمها كل مفردة من هذه المفردات؟.

٣-٥- هل يختلف صدق الاختبار ككل (من نوع الاختيار من متعدد) طبقاً لموقع (ترتيب) البديل الصحيح بين بدائل الاختبار؟

٣-٦- هل يختلف ثبات الاختبار ككل (من نوع الاختيار من متعدد) طبقاً لموقع (ترتيب) البديل الصحيح بين بدائل الاختبار؟

٣-٧- هل توجد فروق بين متوسطات معاملات الصعوبة لمفردات الاختبار (من نوع الاختيار من متعدد) تعزى إلى اختلاف موقع (ترتيب) البديل الصحيح بين بدائل كل مفردة؟

٣-٨- هل توجد فروق بين متوسطات معاملات التمييز لمفردات الاختبار (من نوع الاختيار من متعدد) تعزى إلى اختلاف موقع (ترتيب) البديل الصحيح بين بدائل كل مفردة؟

٤- أهمية الدراسة

تكتسب الدراسة الحالية أهميتها مما لاحظته الباحثة أثناء عملها بالتدريس في الجامعة، والاشتراك مع بعض الزملاء في وضع أسئلة اختبارات للطلاب في مقررات الدراسة، أنه يشيع لدى بعض الزملاء فكرة أنه كلما ازداد عدد البدائل في اختبارات من نوع الاختيار من متعدد، كلما قلّت فرصة التخمين، مما يجعلهم يميلون إلى زيادة عدد البدائل إلى أربعة أو خمسة بدائل بدلاً من ثلاثة، متجاهلين في ذلك أثر ما يحدثه هذا الأمر على الخصائص السيكومترية الأخرى سواء للاختبار ككل، أو لكل مفردة من مفرداته، وإن كان هناك من الدراسات ما قد أشارت نتائجه إلى عدم وجود فروق ذات دلالة إحصائية بين نماذج الاختبارات المختلفة في عدد بدائلها (ثلاثة أو أربعة أو خمسة بدائل) في الخصائص السيكومترية سواء على مستوى الاختبار ككل، أو على مستوى كل مفردة من مفرداته، مما يعكس أهمية الدراسة الحالية، ومن ثم تتجلى هذه الأهمية في مساعدة المختصين والمنوط بهم صياغة وإعداد مفردات اختبارات من هذا النوع (اختيار من متعدد) لتحسين وتطوير إعداد هذه الاختبارات بالتعرف على أفضل عدد من البدائل، والموقع (الترتيب) الأفضل للبديل الصحيح بين هذه البدائل، ولعل ذلك يساهم من ناحية أخرى في أثراء المكتبة العربية بأنماط ونماذج جديدة من اختبارات هذا النوع وفق العدد الأنسب للبدائل، والموقع (الترتيب) الأنسب للبديل الصحيح بينها، في اختبارات من هذا النوع بحيث يُرَاعَى في إعدادها شروط وقواعد الاختبار الجيد.

٥- فرضيات الدراسة

تحقيقاً لأهداف الدراسة يتم اختبار صحة مجموعتين من الفرضيات، تختص الأولى بأثر اختلاف عدد البدائل التي تضمها كل مفردة من مفردات الاختبار، بينما تختص الثانية بأثر اختلاف موقع (ترتيب) البديل الصحيح بين بدائل المفردة، وذلك كما يلي:

٥-١- المجموعة الأولى (عدد البدائل):

٥-١-١- توجد فروق بين متوسطات معاملات الصعوبة لمفردات الاختبار (من نوع الاختيار من متعدد) تعزى إلى اختلاف عدد بدائل كل مفردة من مفردات الاختبار.

٥-١-٢- توجد فروق بين متوسطات معاملات التمييز لمفردات الاختبار (من نوع الاختيار من متعدد) تعزى إلى اختلاف عدد بدائل كل مفردة من مفردات الاختبار.

٥-٢- **المجموعة الثانية (ترتيب البدائل) :**

٥-٢-١- توجد فروق بين متوسطات معاملات الصعوبة لمفردات الاختبار (من نوع الاختيار من متعدد) تعزى إلى اختلاف موقع (ترتيب) البديل الصحيح بين بدائل كل مفردة في الاختبار.

٥-٢-٢- توجد فروق بين متوسطات معاملات التمييز لمفردات الاختبار (من نوع الاختيار من متعدد) تعزى إلى اختلاف موقع (ترتيب) البديل الصحيح بين بدائل كل مفردة في الاختبار.

٦- مصطلحات الدراسة

فيما يلي تعريف لأهم المصطلحات المستخدمة خلال الدراسة الحالية :

٦-١- **الاختبار من نوع الاختيار من متعدد** : هو اختبار تتكون المفردة فيه من جزأين رئيسيين يسمى الأول المتن أو الجذع (stem) وقد يكون عبارة تقريرية (statement) أو سؤال، ويسمى الجزء الثاني بالمموهات أو المشتتات أو البدائل (alternatives) وهي بمثابة إجابات محتملة للسؤال الوارد بالمتن (الجذع) أو عبارات مكملة للعبارة الواردة به ، وتتضمن البدائل إجابة واحدة صحيحة وجميع الإجابات الأخرى تكون خاطئة (عايصرة ،١٩٩٥)، والتعريف الإجرائي لاختبار الاختيار من متعدد هو ذلك الاختبار المستخدم خلال الدراسة الحالية والمطبق على أفراد العينة ، وهو من نموذجين ولكل نموذج منهما ثلاث صور على النحو التالي:

٦-١-١- **النموذج الأول للاختبار:** وله ثلاث صور – وهو نموذج اختبار تكون عدد بدائل كل مفردة فيه إما (٥) ، أو(٤) ، أو (٣) بدائل [صورة ١ ، وصورة ٢ ، وصورة ٣ على الترتيب].

٦-١-٢- **النموذج الثاني للاختبار:** وله ثلاث صور، وهو نموذج اختبار رباعي البدائل يختلف موقع البديل الصحيح من صورة لأخرى : صورة (١): اختبار يكون فيه ترتيب البديل

الصحيح ترتيباً متقدماً (الأول أو الثاني)، صورة (٢): اختبار يكون فيه ترتيب البديل الصحيح ترتيباً متأخراً (الثالث أو الرابع)، صورة(٣): اختبار فيه يكون ترتيب البديل الصحيح موزعاً عشوائياً بين البدائل الأربعة .

٦-٢- **الخصائص السيكومترية:** يقصد بها خلال الدراسة الحالية بالنسبة للاختبار ككل معامل صدق وثبات الاختبار، أما بالنسبة للمفردة فيقصد بها معامل صعوبة المفردة، ومعامل القدرة التمييزية لها ، وتحسب هذه المعاملات (الصدق – الثبات – السهولة – التمييز) بالطرق الإحصائية المتعارف عليها.

٦-٣- **النموذج الأول للاختبار:** ويقصد به اختبار التحصيل الدراسي المستخدم خلال الدراسة الحالية، وهو من ثلاث صور: الصورة الأولى ولكل مفردة فيه خمسة بدائل والصورة الثانية ولكل مفردة فيها أربعة بدائل فقط بعد حذف أحد البدائل الخمسة من الصورة الأولى، والصورة الثالثة ولكل مفردة فيها ثلاثة بدائل فقط بعد حذف أحد البدائل الأربعة من الصورة الثانية.

٦-٤- **النموذج الثاني للاختبار** : ويقصد به نفس الاختبار الخاص بالصورة الثانية من النموذج الأول (ذات الأربعة بدائل)، وهو من ثلاث صور أيضاً : الصورة الأولى وفيها يقع البديل الصحيح في ترتيب متقدم (الأول أو الثاني)، وفي الصورة الثانية يقع البديل الصحيح في ترتيب متأخر (الثالث أو الرابع)، وفي الصورة الثالثة يكون ترتيب البديل الصحيح ترتيبا عشوائياً.

٧- حدود الدراسة

يتم الالتزام بالحدود التالية لدى تعميم النتائج التي تتوصل إليها الدراسة الحالية:

٧-١- حدود أفراد العينة المطبق عليهم الدراسة، وهم الطلاب المسجلون للدراسة بالدبلوم العامة للتربية بمعهد البحوث والدراسات التربوية بجامعة القاهرة خلال العام الدراسي ٢٠١٢/٢٠١٣.

٧-٢- حدود الاختبار المطبق خلال الدراسة بنموذجيه والصور الثلاث لكل نموذج منهما.

٧-٣- حدود أثر كل من عدد البدائل وترتيبها فقط على الخصائص السيكومترية للاختبار ككل ولكل مفردة من مفرداته.

٧-٤- حدود الاعتماد على النظرية التقليدية في القياس دون النظرية المعاصرة المعروفة باسم نظرية الاستجابة للمفردة (IRT) لحساب الخصائص السيكومترية للاختبار بصوره المختلفة.

٨- الإطار النظري والدراسات السابقة

تصنف الاختبارات من حيث نمط الإجابة إلى اختبارات **مقالية** يمكن أن تسمي اختبارات مصاغة أو (أو منشأة) (supply type) ، واختبارات **موضوعية** أو مختارة أو انتقائية (selection type) ، والاختبارات الموضوعية عدة أنواع أهمها (الأمير ، ٢٠١٣ ، سليمان ، ٢٠٠١ ، Cristina, et al. :2011; &Desrochers, et al., 2007)

٨-١- **اختبارات الصواب أو الخطأ**، وتكون كل مفردة من مفردات هذا الاختبار تتكون من عبارات أو أسئلة وتكون استجابة الطالب عليها إما بالصواب أو الخطأ ، ولا يوجد غير إجابتين محتملتين فقط ، إما صواب وإما خطأ.

٨-٢- **اختبارات المقابلة أو المزاوجة**، والاختبار من هذا النوع يتكون من مجموعة من البنود أو الأسئلة التي يقابلها ويوازيها مجموعة من البدائل أو الإجابات المحتملة لهذه الأسئلة.

٨-٣- **اختبارات التكملة**، وهو اختبار تكون كل مفردة من مفرداته عبارة ينقصها كلمة أو أكثر وعلى الطالب أن يكمل العبارة بكتابة الكلمة أو الكلمات الناقصة لتصبح العبارة صحيحة. وغالباً ما يكون الإكمال بكلمات محدودة فقط.

٨-٤- **اختبارات الاختيار من متعدد** ، وفي هذا النوع يتكون الاختبار من عدد من المفردات تتكون كل مفردة من جزأين رئيسيين ، الجزء الأول, ويسمى المتن أو الجذع (stem), ويكون على شكل سؤال أو عبارة تتناول قضية معينة تحتاج إلى إجابة ، ويسمي الجزء الثاني البدائل أو الاختيارات أو المموهات أو المشتتات، وهي بمثابة حلول أو إجابات محتملة للسؤال أو القضية الواردة في الجزء الأول (المتن أو الجذع)، ويكون من بين هذه البدائل بديل واحد فقط هو الصحيح، أما البدائل الأخرى فهي غير صحيحة ويطلق عليها المشتتات أو المموهات تمييزا لها عن البديل الصحيح.

إلا أن الاختبارات من نوع الاختيار من متعدد ليست نوعاً واحداً فهناك عدة أنواع أو أنماط منها ما يلي كما أشارت إلى ذلك بعض الدراسات (الأمير ، ٢٠١٣ ، الرباعي ، ٢٠١٢ ، العطوي ، ٢٠٠١ ، الغامدي ، ٢٠٠٩ ، Schroeder, et al., 2012 & Kehoe, 1995):

٨-٤-١- **نمط الإجابة الصحيحة :**

ويعتبر هذا النمط من أبسط أنماط الاختبارات من نوع الاختيار من متعدد، حيث تحتوى المفردة على إجابة واحدة صحيحة تماماً، أما باقي البدائل فتكون خاطئة تماماً، ومثل هذا النوع من الاختبارات من هذا النمط مناسب لما يتعلق بحقائق لا شك فيها مثل الأسماء والتواريخ والحقائق العلمية والنتائج الحسابية (لا تحتمل الرأي والرأي الآخر).

٨-٤-٢- **نمط الإجابة الخاطئة:**

في هذا النمط تحتوي المفردة على إجابة واحدة خاطئة وبقية البدائل صحيحة، وهذا النمط يكون مناسباً حينما تكون معرفة الاستثناء لها نفس أهمية معرفة القاعدة.

٨-٤-٣- **نمط أفضل الإجابات:**

في هذا النمط تكون جميع الإجابات أو البدائل صحيحة بصورة جزئية، ولكن واحداً منها فقط هو الأكثر صحة وأفضل من غيره من البدائل أو الإجابات الأخرى، ويلاحظ أنه عند استخدام هذا النمط فلابد أن يكون واضحا أن الخيار الصحيح هو الخيار الأفضل بشكل دقيق، وهذا النمط يعّد أكثر صعوبة من النمط السابق لأنه يتطلب إجراء تمييزات دقيقة بين البدائل المعطاة ليتسنى التوصل إلى أفضلها، ويرى البعض أن هذا النمط مناسب لقياس النواتج المعرفية من مستوى الفهم والتطبيق.

٨-٤-٤- **نمط أسوأ إجابة:**

وهذا النمط يشبه النمط السابق إلا أنه في هذه الحالة يُطلب من المستجيبين اختيار أسوأ الإجابات.

وبالنسبة للاختبارات من نوع الاختيار من متعدد بأنماطه المختلفة – وهو النوع الذي تركز عليه الدراسة الحالية – فتشير العديد من الدراسات إلى أن هناك العديد من المزايا للاختبارات من نوع الاختيار من متعدد هي على النحو التالي كما أشارت إليها هذه الدراسات (العطوي ، ٢٠١١ ، فقوسة، ١٩٩٤، Cristina, et al., 2011 &Desrochers, et al., 2007)الشمولية في قياس مخرجات التعلم بمستوياته المعرفية المختلفة (من مستوى التذكر إلى مستوى التقويم) في وقت قصير نسبياً .

(١) الفاعلية في تغطية المحتوى الدراسي بشكل أكبر.

(٢) السهولة والدقة في التصحيح وتقدير الإجابة بموضوعية تامة.

(٣) سهولة التحكم في صعوبة الاختبار وقدرته التمييزية عن طريق التحكم في مستوى تجانس الاختيارات أو البدائل وجاذبيتها.

(٤) إمكانية استخدامها في قياس مستويات التعلم المختلفة في مجالات دراسية مختلفة المحتوى.

(٥) القابلية للتعديل بسهولة من خلال طرق التحليل الإحصائي.

(٦) تعويد الطالب على اتخاذ القرار، والموازنة والتمييز بين الأفضل وإصدار أحكام صائبة.

(٧) عدم تأثرها بقدرة الطالب على التعبير والكتابة الإنشائية.

(٨) إمكانية استخدامها لإعداد بنوك الأسئلة، فهي الأنسب في أسئلة هذه البنوك.

(٩) المساعدة في تشخيص أخطاء الطلاب أو سوء الفهم لديهم من خلال استجاباتهم للمشتتات أو الاختيارات أو الممّوهات.

(١٠) تستوجب من الطالب أن يراجع أغلب محتوى المادة الدراسية المطلوبة منه.

(١١) إمكانية تحليل نتائجها الإحصائية بسهولة واستخراج الدلالات اللازمة منها.

(١٢) إمكانية التصحيح يدوياً أو آلياً باستخدام آلات التصحيح.

(١٣) تقليل احتمالات التخمين نسبياً بالمقارنة مع أسئلة الصواب أو الخطأ.

إلا أنه من ناحية أخرى، فان لهذا النوع من الاختبارات عيوباً أشارت إليها بعض الدراسات أيضاً، ومن هذه العيوب ما يلي كما أشارت إليها هذه الدراسات (فقوسة ، ١٩٩٤ ، Dibattista, et al., 2011 & Kehoe, 1995):

١- تحتاج إلى مهارة عالية وتدريب كاف على كيفية الإعداد الجيد لها.

٢- تستلزم وقتاً طويلاً وجهداً كبيراً في إعدادها.

٣- ارتفاع كلفة الطباعة والتصوير لازدياد عدد صفحات الورقة الإمتحانية.

٤- قلة مناسبتها لقياس مخرجات التعلم التي تتعلق بالتأليف والإبداع والابتكار، وعرض وتنظيم وتتابع أو تسلسل الأفكار، وغيرها من المخرجات والمهارات التي تعتمد على التعبير الكتابي.

٥- تتأثر بالقدرة القرائية للطالب.

٦- شيوع وقوع أخطاء في صياغتها وإعدادها.

٧- صعوبة إيجاد بدائل أو مشتتات مقنعة أو جاذبة.

٨- سهولة الغش بين الطلاب عند الإجابة.

٩- وجود احتمالات التخمين بنسبة عالية حسب عدد البدائل أو المشتتات أو المموهات.

أما بالنسبة لقواعد صياغة مفردات الاختبار من نوع الاختيار من متعدد وما تتضمنه هذه المفردات من مشتتات أو بدائل أو اختيارات، فقد أشارت العديد من الدراسات إلى أنه عند إعداد هذا النوع من الاختبارات فإنه يجب مراعاة عدد من القواعد والإرشادات، منها ما يتعلق بأصل المفردة أو السؤال (ويسمى متن أو جذع السؤال)، ومنها ما يتعلق بالبدائل أو الاختيارات أو المشتتات أو المموهات، وفيما يلي هذه القواعد كما ذكرتها هذه الدراسات (الأمير ، ٢٠١٣ ، الرباعي ، ٢٠١٢)

(Cristina, et al., 2011 & Kehoe, 1995):

أولاً – القواعد المتعلقة بجذع (stem) المفردة:

١- يجب أن يحتوى جذع المفردة على مشكلة محددة وواضحة ومفهومة لدى المستجيب.

٢- أن تقيس المفردة ناتجاً تعليمياً مهماً (له قيمته).

٣- صياغة المفردة بشكل دقيق واضح الصياغة في صيغة سؤال مباشر أو جملة خبرية ناقصة.

٤- عدم تضمين الاختبار مفردة تحتاج إجابتها إلى معرفة لسنا بصدد قياسها.

٥- يفضل تجنب استخدام صيغة النفي، لأن ذلك يربك المستجيب ، وإلا وجب وضع خط تحت أداة النفي لكي ينتبه المستجيب.

٦- مراعاة استقلالية المفردات وعدم اعتماد مفردة على أخرى في الإجابة.

٧- عدم وجود أي تلميح أو إشارة في الجذع يمكن أن يشير إلى الإجابة الصحيحة أو البديل الصحيح كأن تستخدم نفس الألفاظ تماما بين الاثنين.

٨- تناول جذع المفردة فكرة واحدة فقط.

٩- إذا كانت المفردة تتعلق بتعريف مصطلح معين ، فمن الأفضل وضع المصطلح في جذع المفردة.

١٠- ضبط صعوبة المفردة لتكون متوسطة الصعوبة.

ثانياً – القواعد المتعلقة بالبدائل أو المشتتات:

١. يفضل عدم استخدام بدائل (كل ما ذكر صحيح – جميع ما ذكر – ليس أياً مما سبق صحيح) ضمن البدائل.
٢. تجانس البدائل في محتواها وارتباطها بمجال المشكلة.
٣. يجب أن تكون هناك إجابة واحدة صحيحة فقط.
٤. يجب مراعاة تقارب البدائل من حيث الطول، ولا تمييز للبديل الصحيح عن البدائل الأخرى بشكل قد يكشفه أو يلفت الانتباه إليه.
٥. تجنب الارتباطات اللفظية بين أصل (جذع) المفردة والبديل الصحيح.
٦. مراعاة أن تكون لغة صياغة البدائل والكلمات أو المصطلحات المستخدمة سهلة وواضحة ومعروفة للمتعلمين.
٧. يتم التنويع في ترتيب عرض البديل الصحيح بين البدائل بما لا يساعد على التخمين.
٨. يفضل كتابة رقم المفردة بالعدد الحسابي، أما البدائل فبحروف أبجدية.
٩. يفضل ترتيب البدائل أو المشتتات ترتيبا منطقياً.
١٠. استخدام العدد المناسب من البدائل (من ٣ – ٥ بدائل).
١١. أن تكون البدائل ذات جاذبية متكافئة.
١٢. استخدام البدائل التي يُعتقد أن لها استجابات متوقعة.

فمستوى الأداء على المفردة يتوقف على خصائص المفردة، أو الاختبار ككل، وخصائص الفرد المستجيب ، مما يعني أن أي انتهاك لبنية المفردة من شأنه التأثير في الخصائص الإحصائية أو السيكومترية للمفردة ، والتي تؤثر بدورها في أداء الفرد على المفردة والاختبار ككل، وهذا هو ما حاولت بعض الدراسات التحقق من صحته ، حيث هدفت بعض الدراسات إلى التحقق من أثر انتهاك قواعد صياغة مفردات الاختبار من نوع الاختيار من متعدد على الخصائص السيكومترية للاختبار ومفرداته المختلفة، ومن هذه الدراسات دراسة (الرمادى ،٢٠٠٦،سليمان ، ٢٠٠١ ، عوده، ١٩٩٣) ،(Shizukam,et al.,2006)

ومن الدراسات الأخرى التي أجريت في هذا المجال دراسة (Bruno& Dirkzwager,1995) والتي قاموا خلالها بمراجعة ٢٧٧٠ مفردة من نوع الاختيار من متعدد تمَّ تطبيقها على بعض الطلاب في مجال التمريض خلال خمس سنوات وذلك في ضوء ١٩ قاعدة من قواعد صياغة المفردات، وقد أوضحت الدراسة عدة نتائج أهمها أن ١ر٤١ % من المفردات كانت من إعداد المعلم ، بينما كان ٢ر٣٦% مسترجعة من بنوك الأسئلة ، وبلغت نسبة المفردات معروفة المصدر٤ر٤٩% ، كما أشارت النتائج إلى أن المفردات التي كانت تقيس مستويات معرفية متدنية (الحفظ والتذكر والتطبيق) كانت هي المفردات الأكثر إتبّاعاً للقواعد، أما بالنسبة لمصدر المفردة (المعلم – بنوك الأسئلة – مجهولة المصدر) فلم يكن له علاقة بانتهاك القواعد، كما أن المفردات التي من إعداد المعلم كانت هي الأكثر ميلاً لقياس المستويات المعرفية الأعلى ، وأن البدائل الصحيحة كان عرضها متوازناً حيث توزعت بشكلٍ متساوٍ على مختلف المواقع دون تحيّز لموقع على آخر.

وفي الدراسة التي أجراها (Cizek& O Day, 1994) تمَّ تحليل المفردات الواردة في عشرة كتب في المحاسبة ، ومفردات بعض الاختبارات الخاصة بالحصول على شهادة مدقق حسابات قانوني (CPA): Certified Public Accountant في ضوء ١٧ قاعدة من قواعد صياغة المفردات، وقد أشارت النتائج إلى أن ٧٥% من المفردات تنتهك واحدة أو أكثر من قواعد كتابة المفردات، وفي المقابل لم تنتهك ٧٠% منها قواعد صياغة المفردات، وقد أرجع الباحثان هذا الاختلاف إلى أن معدّي المفردات لاختبارات CPA تمَّ تدريبهم على كتابة الأسئلة ، وقد تمَّ مراجعة المفردات بشكل مكثف قبل استخدامها وتطبيقها ، بعكس الأمر بالنسبة لمؤلفي كتب المحاسبة فإنهم لم يسبق تدريبهم على إعداد مفردات اختبارات من نوع الاختيار من متعدد، كما أظهرت النتائج أيضا أن بعض الانتهاكات كانت تصحح دون إعادة صياغة المفردة ، كأن يتم حذف البديلين " جميع ما ذكر " ، و " لا شيء مما ذكر " ، أو إضافة بعض الكلمات إلى أصل (جذع) المفردة ، وإتباع قواعد الترقيم الصحيح ، وبلغت نسبة هذه المخالفات ٣٨% من أصل ٤٩٠ مخالفة تمَّ الكشف عنها.

ولعل الاهتمام بمراعاة قواعد الصياغة عند إعداد وكتابة مفردات الاختبارات من نوع الاختيار من متعدد، تزداد حسب أهمية القرارات التي تترتب على نتائج تطبيق هذه الاختبارات، مما يجعل الاختبارات ذات النتائج المهمة ، والاختبارات المقننة، واختبارات الشهادات العامة كالثانوية العامة وغيرها تحظى بالمزيد من الاهتمام ، والالتزام بدقة من جانب معدي هذه الاختبارات بالقواعد المتعارف عليها.

وتلاحظ الباحثة من خلال استقراء ما سبق عرضه من دراسات سابقة بشأن قواعد صياغة أسئلة أو مفردات الاختبارات من نوع الاختيار من متعدد تنوع أهدافها، حيث هدفت بعض هذه الدراسات إلى مسح للمخالفات التي ارتكبها معدو هذه الأسئلة أو المفردات في الاختبارات بالكتب المدرسية أو غيرها، وهدفت دراسات أخرى إلى فحص مدى الالتزام في صياغة بعض الأسئلة أو المفردات في ضوء عدد من القواعد المحددة، كما تناولت بعض الدراسات الأخرى مناقشة أثر هذه الانتهاكات سواء على الخصائص الإحصائية أو السيكومترية للمفردات وللاختبار ككل، وكذلك أثرها على التحصيل الدراسي من خلال أداء المستجيبين على هذه الأسئلة أو المفردات.

أما بالنسبة للدراسات الخاصة بعدد البدائل أو المشتتات وأثرها على الخصائص الإحصائية أو السيكومترية، فقد أجرى (Ebel, 1969) دراسته التي أكد فيها من خلال الصيغ الرياضية إلى أن التحسن في ثبات الاختبار الموضوعي يزداد بزيادة عدد البدائل من بديلين إلى ثلاثة بدائل ، ولكن يكون هذا التحسن قليلاً عند زيادة عدد البدائل إلى أربعة بدائل أو أكثر، أما الدراسات التي أجراها(Costin, 1972) فقد أشارت إلى أن الاختبار الذي تكون كل مفردة فيه تتكون من ثلاثة بدائل يكون أكثر ثباتا من الاختبارات التي تتكون المفردة فيها من بديلين أو أربعة أو خمسة بدائل.

أما الدراسة التي أجراها كل من(Dibattista& Kurzawa,2011) وقد تمَّ حساب درجة الصعوبة ودرجة التمييز للمفردات, وكذلك حساب ثبات الاختبار ككل بالاعتماد على مفاهيم النظرية التقليدية (الكلاسيكية) للقياس، ومقارنتها باختباري اختيار من متعدد ذي ثلاثة بدائل وذي أربعة بدائل ، وأوضحت النتائج أن درجة الصعوبة ودرجة التمييز للمفردات ذات الثلاثة بدائل أعلى منها بالنسبة للمفردات ذات الأربعة بدائل، كما كشفت النتائج عن أن ثبات الاختبار ذي الثلاثة بدائل أعلى منه بالنسبة للاختبار ذي الأربعة بدائل.

أما الدراسة التي أجراها (Crehan, Haladyna& Brewer, 1993) فقد أوضحت نتائجها أن مفردات الاختبار ذات البدائل الثلاثة أصعب من مفردات الاختبار ذات البدائل الأربعة، كما وجد أنه لا يوجد أثر لاختلاف عدد البدائل على معاملات القدرة التمييزية للمفردات، ويلاحظ أن الدراسة التي أجراها (Trevisan, Sax & Michael, 1994) أشارت نتائجها إلى أنه لا توجد فروق ذات دلالة إحصائية بين معاملات ثبات الاختبار تُعزَى لعدد البدائل.

أما دراسة (الشبلي ، ١٩٩٥) فقد توصلت نتائجها إلى أفضل عدد ممكن للبدائل في اختبارات الاختيار من متعدد، ومن ثم فقد أوصت الدراسة – وفقاً للنتائج التي توصلت إليها – باستخدام مفردات ثلاثية البدائل لبناء مفردات الاختبار من نوع الاختيار من متعدد.

أما بالنسبة للدراسة التي أجراها (الغامدي ، ٢٠٠٨) فقد استهدفت الكشف عن أثر عدد البدائل على الخصائص السيكومترية للاختبار التحصيلي في الرياضيات, وطبقت على طلاب الصف الثاني المتوسط بمكة ، وأظهرت النتائج عدم جود فروق دالة إحصائياً بين معاملات الصعوبة للاختبارات ذات الثلاثة أو الأربعة والخمسة بدائل، كما أوضحت النتائج أن أفضل معاملات صعوبة بشكل عام كان للنموذج ذي البدائل الثلاثة حيث لم توجد فروق دالة إحصائياً بين متوسط معاملات صعوبة مفرداته، في حين وجدت فروق دالة إحصائياً بين متوسط صعوبة مفردات النموذجين الآخرين، كذلك كانت الفروق غير دالة إحصائياً بين معاملات التمييز للاختبارات ذات البدائل الثلاثة والأربعة والخمسة بدائل، كما لم توجد فروق دالة إحصائياً بين النماذج الثلاثة من حيث معاملات ثبات الاختبار، وكذلك من حيث معاملات الصدق التلازمي والتنبؤي.

ومما يجدر ذكره أن هناك من الدراسات تلك التي قارنت أثر عدد البدائل أو ترتيبها على الخصائص السيكومترية للاختبار ومفرداته في ضوء كل من نظريتي القياس (النظرية التقليدية ونظرية الاستجابة للمفردة IRT) مثل دراسة (Revuleta,2004)ودراسة (Hwang, 2002)، وأظهرت النتائج أنه بينما لم توجد فروق من حيث الأثر في ظل النظرية التقليدية، فانه وجدت فروق دالة إحصائياً في ظل نظرية الاستجابة للمفردة، وذلك لما تقوم عليه من افتراضات في القياس تجعلها أكثر قوة ودقة لتحرر القياس من تباين مستويات قدرات أفراد العينة، وتباين مستويات صعوبة المفردات.

وباستقراء البحوث والدراسات السابقة يلاحظ ما يلي:

- هناك دراسات استهدفت التعرف على الانتهاكات أو المخالفات التي يقع فيها البعض ممن يعدوّن هذه الاختبارات من حيث قواعد وشروط الصياغة الجيدة.
- بينما تناولت بعض الدراسات دراسة العدد الأمثل للبدائل في الاختبار من نوع الاختيار من متعدد ، فقد تناولت دراسات أخرى – وإن كانت قليلة – أثر ترتيب هذه البدائل وموقع البديل الصحيح بينها على خصائص المفردات أو الاختبار ككل.

- ليس هناك اتفاق بين نتائج الدراسات حول العدد الأمثل للبدائل عند استخدام الاختبارات من نوع الاختيار من متعدد ، ولا الترتيب الأمثل لها إلا أن يكون الترتيب عشوائياً.
- بينما تناولت بعض الدراسات أثر عدد البدائل على الخصائص السيكومترية للاختبار في ضوء النظرية التقليدية للقياس، فقد تناولته بعض البحوث الأخرى في ضوء نظرية الاستجابة للمفردة (IRT).

٩- إجراءات الدراسة:

٩-١- الأدوات:

تتمثل أدوات الدراسة في اختبار تحصيل دراسي تمَّ إعداده في مقرر علم النفس التربوي الذي يتم تدريسه لطلاب الدبلوم العامة في التربية بمعهد البحوث والدراسات التربوية بجامعة القاهرة خلال العام الدراسي ٢٠١٢/٢٠١٣، وهو يتكون من (٥٠) مفردة من نوع الاختيار من متعدد، وقد أُعّد الاختبار من نموذجين ، الأول وله ثلاث صور تضم نفس المفردات والاختلاف بين الصور في عدد البدائل (المشتتات أو المموّهات):

- **الصورة الأولى**، وكل مفردة في هذه الصورة لها **خمسة بدائل**.
- **الصورة الثانية**، وكل مفردة في هذه الصورة لها **أربعة بدائل** بعد حذف أحد البدائل من الصورة الأولى عشوائيا (دون البديل الصحيح).
- **الصورة الثالثة**، وكل مفردة في هذه الصورة لها **ثلاثة بدائل** بعد حذف أحد البدائل من الصورة الثانية عشوائيا (دون البديل الصحيح).

أما النموذج الثاني فيعتمد على الصورة الثانية من النموذج الأول (أربعة بدائل)، وهذا النموذج له ثلاث صور أيضا تضّم نفس المفردات والبدائل ، والاختلاف بين الصور هو في موقع (ترتيب) البديل الصحيح بين البدائل الأربعة لكل مفردة من مفردات الاختبار:

- **الصورة الأولى**، والبديل الصحيح في كل مفردة في هذه الصورة يحتل موقعاً (ترتيباً) متقدماً بين البدائل (الترتيب الأول أو الثاني فقط).
- **الصورة الثانية** ، والبديل الصحيح في كل مفردة في هذه الصورة يحتل موقعاً (ترتيباً) متأخراً بين البدائل (الترتيب الثالث أو الرابع فقط).

- **الصورة الثالثة،** والبديل الصحيح في كل مفردة في هذه الصورة يحتل موقعاً (ترتيباً) عشوائياً بين البدائل الأربعة دون الاقتصار على ترتيب محدد.

وقد تمَّ إعداد الاختبار بنموذجيه وصوره الستة (بواقع ثلاث صور لكل نموذج كما هو موضّح أعلاه) في شكل أوّلي بواسطة ثلاثة من المحاضرين الذين يدرّسون المقرر الدراسي من بينهم الباحثة، وتم عرضه في صورته الأولية على بعض السادة المحكمين الخبراء المتخصصين في تدريس المقرر الدراسي وفي القياس والتقويم، وقد أبدوا ملاحظاتهم على الاختبار بنموذجيه وصورهما المختلفة، وقد تمَّ أخذ هذه الملاحظات بالحسبان حيث تمَّ حذف بعض المفردات، وتعديل صياغة بعض المفردات الأخرى، وإضافة مفردات أخرى جديدة ، وأصبح الاختبار بعد تعديله طبقاً لملاحظات الخبراء المحكمين يضم (٥٠) مفردة صالحة، وذلك لكل صورة من الصور الثلاث لكل نموذج من نموذجي الاختبار.

وللتحقق من ثبات الاختبار بنموذجيه وصورهما المختلفة تمَّ تطبيقه على عينة محدودة من الطلاب والطالبات (٢٠ طالباً وطالبة) هم أصلاً من أفراد عينة الدراسة الذين أعيد تطبيق الاختبار عليهم ضمن التطبيق النهائي مرة ثانية ، وقد قورنت نتائج التطبيق عليهم في المرتين (الاستطلاعية والنهائية) وحسبت معاملات الثبات بطريقة التطبيق / إعادة التطبيق (test-retest) وكانت معاملات الثبات كما يوضحها جدول (١)، وهي معاملات ثبات مقبولة لمثل هذا النوع من الاختبارات:

جدول (١)

معاملات ثبات صُوَر نموذجيْ الاختبار بطريقة التطبيق / إعادة التطبيق

النموذج / الصورة	معامل الثبات
الأول :	
صورة (١)	٠ر٨٧٥
صورة (٢)	٠ر٧٩٢
صورة (٣)	٠ر٧٥٤
الثاني :	
صورة (١)	٠ر٨١١
صورة (٢)	٠ر٧٨٩
صورة (٣)	٠ر٨٠١

٩-٢- **العينة:**

بلغ عدد أفراد عينة الدراسة (١١٣) طالبا وطالبة (٥١ طالباً ، ٦٢ طالبة) ممن يدرسون الدبلوم العامة في التربية بمعهد البحوث والدراسات التربوية / جامعة القاهرة خلال العام الدراسي ٢٠١٢ / ٢٠١٣، وقد كان يتم تقسيمهم عشوائياً إلى ثلاث مجموعات فرعية بحيث يعطى لأفراد كل مجموعة صورة من الصور الثلاث للنموذج الأول للاختبار في المرحلة الأولى، وصورة من الصور الثلاث للنموذج الثاني للاختبار في المرة الثانية حيث تمَّ التطبيق مرحلتين، وفيما يلي يوضح جدول (٢) توزيع أفراد العينة حسب المجموعات الفرعية الثلاث، والجنس، والعدد، ومتوسط العمر بالسنوات:

جدول (٢)

توزيع أفراد العينة حسب المجموعات الفرعية الثلاث والجنس والعدد ومتوسط العمر

المجموعة	الجنس	العدد	متوسط العمر بالسنوات
الأولى (ن = ٣٨)	ذكور	١٧	٢٥ر١٤
	إناث	٢١	٢٤ر٧٨
	مجموع	٣٨	٢٤ر٩٦
الثانية (ن =٣٨)	ذكور	١٨	٢٤ر٧٤
	إناث	٢٠	٢٢ر٨٦
	مجموع	٣٨	٢٣ر٨٠
الثالثة (ن = ٣٧)	ذكور	١٦	٢٥ر٧١
	إناث	٢١	٢٥ر٢٢
	مجموع	٣٧	٢٥ر٤٧

٩-٣- **جمع البيانات :**

بعد الانتهاء من إعداد وتطوير أداة الدراسة المتمثلة في الاختبار بنموذجيه والصور الثلاث لكل نموذج منهما، وتطبيق هذه الصور تطبيقاً استطلاعياً وتعديلها في ضوء ملاحظات التحكيم ونتائج التطبيق الاستطلاعي ، والاطمئنان إلى توافر قدر مناسب من درجات الصدق والثبات للاختبار بنموذجيه وصورهما المختلفة على النحو السالف الإشارة إليه، قامت الباحثة بإتباع الخطوات والإجراءات الآتية لجمع البيانات:

٩-٣-١- بعد أن تمَّ تقسيم أفراد العينة عشوائياً إلى ثلاث مجموعات فرعية على النحو السالف توضيحه في جدول (١) ، تمَّ القيام بتطبيق الاختبارات بصورها الستة على مرحلتين كالتالي:

- **المرحلة الأولى** : طبقت الصور الثلاث الخاصة بالنموذج الأول (صور مختلفة من حيث عدد البدائل التي تضمها المفردة) على المجموعات الفرعية الثلاث ، بواقع صورة لكل مجموعة ، وتمَّ ذلك خلال النصف الثاني من شهر مايو ٢٠١٣.
- **المرحلة الثانية** : طبقت الصور الثلاث الخاصة بالنموذج الثاني (صور مختلفة من حيث موقع أو ترتيب البديل الصحيح بين البدائل الأربعة) على المجموعات الفرعية الثلاث، بواقع صورة لكل مجموعة ، وتمَّ ذلك خلال النصف الأول من شهر يونيو ٢٠١٣.

٩-٣-٢- تصحيح استجابات أفراد العينة على الاختبار بنموذجيه، وبصورهما الستة لكل صورة على حدة.

٩-٣-٣- رصد الدرجات الناتجة عن تصحيح استجابات أفراد العينة على الاختبار بصوره المختلفة في كشوف من شأنها تسهيل إدخال البيانات إلى الحاسب بشكل مناسب تمهيدا لمعالجتها إحصائياً.

٩-٤- **المعالجة الإحصائية للبيانات :**

للإجابة عن السؤالين الأول والثاني، والسؤالين الخامس والسادس، تمَّ حساب قيم معاملات الصدق والثبات للاختبار بنموذجيه (وبصوره الست بواقع ثلاث صور لكل نموذج منهما)، وإجراء المقارنة بينها طبقا لقيمها المحسوبة ، أما بالنسبة للإجابة عن الأسئلة الأربعة الأخرى (الثالث والرابع والسابع والثامن)، فقد تمَّ اختبار صحة الفرضيات الأربع التي سعت الدراسة إلى اختبارها، وذلك باستخدام تحليل التباين الأحادي الاتجاه (one-way ANOVA) للتعرف على ما إذا كانت هناك فروق دالة إحصائياً بين

متوسطات معاملات الصعوبة أو معاملات التمييز للمفردات تعزى إلى اختلاف عدد البدائل التي تضمها كل مفردة من مفردات الاختبار، وكذلك التعرف على ما إذا كانت هناك فروق دالة إحصائياً بين متوسطات معاملات الصعوبة أو معاملات التمييز للمفردات تعزى إلى اختلاف موقع (ترتيب) البديل الصحيح بين البدائل الأربعة التي تضمها كل مفردة من مفردات الاختبار.

١٠- نتائج الدراسة :

قبل عرض النتائج الخاصة باختبار صحة الفرضيات الأربع للدراسة ، وللإجابة عن أسئلتها الأربعة (الأول والثاني والخامس والسادس)، قامت الباحثة بمقارنة قيم معاملات الصدق والثبات على مستوى الاختبار ككل بنموذجيه وصوره الستة، وهو ما يوضحه جدول (٣)، أما بالنسبة لمعاملات صعوبة المفردات وتمييزها فقد قامت بحساب تكراراتها مبوبة في ثلاثة مستويات (عال – متوسط – منخفض)، وهو ما يوضحه الجدولان (٤) و (٥)، ثم تعرض الباحثة بعد ذلك النتائج الخاصة باختبار الفرضيات مبوبة في قسمين أساسيين يتعلق أولهما بالنتائج الخاصة بأثر عدد بدائل كل مفردة من مفردات الاختبار على الخصائص السيكومترية للاختبار ككل ، ويتعلق القسم الثاني بالنتائج الخاصة بموقع (ترتيب) البديل الصحيح بين البدائل الأربعة للمفردة على خصائص مفردات الاختبار (معاملات الصعوبة والتمييز):

جدول (٣)

قيم معاملات الصدق والثبات للاختبار ككل بنموذجيه لكل صورة من صورهما الثلاث

المعامل / النموذج / الصورة	معامل الصدق	معامل الثبات
النموذج الأول (عدد البدائل):		
(٥ بدائل) : صورة ١		
(٤ بدائل) : صورة ٢	٠ر٧٥٣٣	٠ر٨٧٥١
(٣ بدائل) : صورة ٣	٠ر٦٤٥٩	٠ر٧٩١٩
	٠ر٦٣٨٩	٠ر٧٩٠٥

النموذج الثاني (ترتيب البديل الصحيح):		
(ترتيب متقدم) : صورة ١		
(ترتيب متأخر) : صورة ٢	٠ر٥٩١١	٠ر٨١١٢
(ترتيب عشوائي) : صورة ٣	٠ر٥٠١٥	٠ر٧٨٨٩
	٠ر٦١١٣	٠ر٨٤١١

وكما يتضح من جدول (٣) فإن قيمة معامل صدق الاختبار ككل تكون أكبر عندما يكون عدد بدائل الاختبار خمسة بدائل، وتقلُّ قيمة معامل الصدق كلما قلَّ عدد البدائل, أما بالنسبة للثبات, فإن قيمة معامل الثبات تكون أكبر عندما يكون عدد بدائل الاختبار خمسة بدائل أيضاً، وتقل أيضا كلما قلَّ عدد بدائل كلِّ مفردة من مفردات الاختبار.

ومن ثم يمكن القول إنَّ ثبات الاختبار وصدقه يكون أفضل عندما يكون الاختبار ثلاثي البدائل، أما بالنسبة لموضوع أثر اختلاف موقع (ترتيب) البديل الصحيح بين البدائل، فيلاحَظ أن أفضل معامل للصدق عندما يكون الترتيب عشوائياً، وكذلك أيضاً فإنه بالنسبة للثبات، فإن أفضل معامل ثبات للاختبار عندما يكون الترتيب عشوائياً.

جدول (٤)

مستويات معاملات صعوبة* مفردات الاختبار بنموذجيه لكل صورة من صورهما الثلاث

النموذج	مستوى الصعوبة / الصورة	عال (صعب)		متوسط		منخفض (سهل)	
		العدد	%	العدد	%	العدد	%
الأول	صورة ١	١٠	٢٠	٣٩	٧٨	١	٢
	صورة ٢	٩	١٤	٤٠	٨٠	١	٢
	صورة ٣	٧	١٨	٣٨	٧٦	٥	١٠
الثاني	صورة ١	١٤	٢٨	٣٤	٦٨	٢	٤
	صورة ٢	١١	٢٢	٣٥	٧٠	٤	٨
	صورة ٣	١٢	٢٤	٣٢	٦٤	٦	١٢

***[عالي (صعب) : أقل من ٣٠ر] ، [متوسط : ٣٠ر٠– أقل من ٧٠ر] ، [منخفض (سهل): ٧٠ر فأكثر]**

وكما يتضح من جدول (٤) فإنَّ الاختبار ذي البدائل الخمسة تكون نسبة عدد المفردات الصعبة فيه ٢٠% من جملة عدد مفرداته، بينما هي أقل من ذلك كلما قلَّ عدد البدائل التي تتضمنها المفردة، وهذا يعني أن الاختبار يصبح أكثر صعوبة كلما زاد عدد البدائل، كما يلاحظ أيضاً أن مستوى صعوبة مفردات الاختبار قد لا يختلف طبقاً لاختلاف ترتيب البديل الصحيح بين بدائل مفرداته، حيث أنه بينما بلغت نسبة المفردات الصعبة ٢٨% عندما يكون ترتيب البديل الصحيح ترتيباً متقدماً، فإنها بلغت ٢٢% عندما يكون هذا الترتيب ترتيباً متأخراً، وتبلغ هذه النسبة ٢٤% عندما يكون هذا الترتيب عشوائياً .

جدول (٥)

مستويات معاملات التمييز* لمفردات الاختبار بنموذجيه لكل صورة من صورهما الثلاث

النموذج	مستوى التمييز / الصورة	عال		متوسط		منخفض	
		العدد	%	العدد	%	العدد	%
الأول	صورة ١	٣	٦	٤٢	٨٤	٥	١٠
	صورة ٢	٤	٨	٤٤	٨٨	٢	٨
	صورة ٣	٥	١٠	٤٠	٨٠	٥	١٠
الثاني	صورة ١	٤	٨	٤١	٨٢	٤	٨
	صورة ٢	٦	١٢	٣٩	٧٨	٥	١٠
	صورة ٣	٧	١٤	٤٠	٨٠	٣	٦

***[عالي: أقل من ٣٠ر] ، [متوسط : ٣٠ر٠– أقل من ٧٠ر] ، [منخفض: ٧٠ر فأكثر]**

وكما يتضح من جدول (٥) فإنَّ مستوى القدرة التمييزية لمفردات الاختبار تزداد بقلة عدد بدائل مفرداته، فالاختبار ذو البدائل الثلاثة تبلغ نسبة المفردات عالية القدرة التمييزية فيه ١٠% من جملة عدد مفرداته, وتقلُّ هذه النسبة كلما زاد عدد البدائل، أما بالنسبة لموقع(ترتيب) البديل الصحيح بين البدائل فيلاحظ أنه كلما كان الترتيب عشوائياً كلما زادت نسبة المفردات ذات القدرة التمييزية العالية (١٤%), وتقلُّ هذه النسبة إذا ما كان الترتيب ترتيباً متأخراً (١٢%), أو ترتيباً متقدماً (٨%) .

وتوضح الجداول الأربعة (من جدول ٦ إلى جدول ٩) النتائج الخاصة باختبار صحة الفرضيات الأربع للدراسة مبوبة في قسمين يختصُّ أولهما بالتعرف على أثر اختلاف عدد بدائل المفردة، ويختصُّ الثاني بالتعرف على أثر موقع (ترتيب) البديل الصحيح بين بدائل المفردة، وذلك على النحو التالي:

القسم الأول – أثر اختلاف عدد البدائل:

الفرضية الأولى – يوضح الجدول (٦) نتائج تحليل التباين لاختبار صحة الفرضية الأولى التي تنصّ على أنه توجد فروق دالة إحصائياً بين متوسطات معاملات الصعوبة لمفردات الاختبار تعزى إلى اختلاف عدد بدائل كلَّ مفردة من مفرداته

جدول (٦)

نتائج تحليل التباين الأحادي (one–way ANOVA)

بين متوسطات معاملات الصعوبة لمفردات النموذج الأول للاختبار بصوَره الثلاث

مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	قيمة "ف"	مستوى الدلالة
بين المجموعات	١٠٢٥	٢	٥١٢ر٥٠	٦ر٤٤٧	٠ر٠٥
داخل المجموعات	٨٧٤٥	١١٠	٧٩ر٥٠		
المجموع	٩٧٧٠	١١٢			

وكما يتضح من جدول (٦) فإنَّ قيمة "ف" دالة إحصائياً عند مستوى ٠ر٠٥ مما يؤكد صحة الفرضية الأولى، بما يعني وجود فروق دالة إحصائياً بين متوسطات معاملات الصعوبة لمفردات الاختبار تعزى إلى اختلاف عدد بدائل كل مفردة من مفرداته . ولمعرفة اتجاه الفروق الدالة، تمَّ التعرف على دلالة الفروق باستخدام اختبار " توكي" وكانت نتائجها كالتالي:

	مج ١ (صورة ١)	مج ٢ (صورة ٢)	مج ٣ (صورة٣)
مج ١ (صورة ١)	–	١ر٦٥٥	٧ر٣١١***
مج ٢ (صورة ٢)		–	٢ر٣٤٤
مج ٣ (صورة ٣)			–

* = دالة عند ٠ر٠٥ **= دالة عند ٠ر٠١ ***= دالة عند ٠ر٠٠١

وكما يتضح من نتائج استخدام اختبار "توكي" على النحو الموضح أعلاه أن الفرق الوحيد الدال إحصائياً كان عند مستوى ٠ر٠٠١, وهو بين الصورتين (١) و (٣) من حيث متوسطات معاملات صعوبة المفردات ولصالح الصورة الأولى (خماسية البدائل) ، مما يعني أن مستويات صعوبة المفردات للاختبار تكون أعلى إذا ما كان الاختبار مفرداته خماسية البدائل عما إذا كانت أقل (ثلاثية أو رباعية).

الفرضية الثانية – يوضح الجدول (٧) نتائج تحليل التباين لاختبار صحة الفرضية الثانية التي تنصّ على أنه توجد فروق دالة إحصائياً بين متوسطات معاملات التمييز لمفردات الاختبار (النموذج الأول) تعزى إلى اختلاف عدد بدائل كل مفردة من مفرداته:

جدول (٧)

نتائج تحليل التباين الأحادي (one-way ANOVA)

بين متوسطات معاملات التمييز لمفردات النموذج الأول للاختبار بصوَره الثلاث

مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	قيمة "ف"	مستوى الدلالة
بين المجموعات	١١٢٢	٢	٥٦١ر٠٠	٦ر٦٧٤	٠ر٠٥
داخل المجموعات	٩٢٤٧	١١٠	٨٤ر٠٦٤		
المجموع	١٠٣٦٩	١١٢			

وكما يتضح من جدول (٧) فإنَّ قيمة "ف" دالة إحصائيا عند مستوى ٠ر٠٥ مما يؤكد صحة الفرضية الثانية، بما يعني وجود فروق دالة إحصائياً بين متوسطات معاملات التمييز لمفردات الاختبار تعزى إلى اختلاف عدد بدائل كل مفردة من مفرداته . ولمعرفة اتجاه الفروق الدالة، تمَّ التعرف على دلالة الفروق باستخدام اختبار " توكي" وكانت نتائجها كالتالي:

	مج ١	مج ٢	مج ٣
مج ١	–	٠ر٣٢١	٦ر٢٥٧***
مج ٢		–	١ر٧٢٣
مج ٣			–

* = دالة عند ٠ر٠٥ **= دالة عند ٠ر٠١ ***= دالة عند ٠ر٠٠١

وكما يتضح من نتائج استخدام اختبار "توكي" على النحو الموضح أعلاه، فإنَّ الفرق الوحيد الدال إحصائياً كان عند مستوى ٠ر٠٠١ ، وهو بين الصورتين (١)و(٣) من حيث متوسطات معاملات التمييز للمفردات ولصالح الصورة (١) " خماسية البدائل "، مما يعني أن مستويات القدرة التميزية لمفردات الاختبار تكون أعلى إذا ما كان الاختبار مفرداته خماسية البدائل، وذلك عما اذا كان عدد بدائلها أقلَّ (ثلاثة أو أربعة).

القسم الثاني – أثر اختلاف موقع (ترتيب) البديل الصحيح :

الفرضية الثالثة – يوضح الجدول (٨) نتائج تحليل التباين لاختبار صحة **الفرضية الثالثة** التي تنصّ على أنه توجد فروق دالة إحصائياً بين متوسطات **معاملات الصعوبة** لمفردات الاختبار تعزى إلى اختلاف **موقع (ترتيب) البديل** الصحيح بين بدائل كل مفردة من المفردات:

جدول (٨)

نتائج تحليل التباين الأحادي (one–way ANOVA)

بين متوسطات معاملات الصعوبة لمفردات النموذج الثاني للاختبار بصورة الثلاث

مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	قيمة "ف"	مستوى الدلالة
بين المجموعات	٣١٢٦	٢	١٥٦٣ر٠٠	١٥ر٣١٧	٠ر٠٠١
داخل المجموعات	١١٢٢٥	١١٠	١٠٢ر٠٤٦		
المجموع	١٤٣٥١	١١٢			

وكما يتضح من جدول (٨) فإنَّ قيمة "ف" دالة إحصائياً عند مستوى ٠ر٠٠١ مما يؤكد صحة الفرضية الثالثة ، بما يعني وجود فروق دالة إحصائياً بين متوسطات معاملات الصعوبة لمفردات الاختبار تعزى إلى اختلاف موقع (ترتيب) البديل الصحيح بين بدائل كل مفردة من المفردات. ولمعرفة اتجاه الفروق الدالة، تمَّ التعرف على دلالة الفروق باستخدام اختبار " توكي" وكانت نتائجها كالتالي:

	مج ١	مج٢	مج ٣
مج ١	-	١ر٧٦٨	٤ر٤٧٨*
مج ٢		-	٥ر٣٢٩**
مج ٣			-

* = دالة عند ٠ر٠٥ **= دالة عند ٠ر٠١ ***= دالة عند ٠ر٠٠١

وكما يتضح من نتائج استخدام اختبار "توكي" على النحو الموضح أعلاه ، فإنَّ الفروق الدالة إحصائياً كانت بين الصورة (٣) وكل من الصورتين (١) و(٢)، حيث كان الفرق دال إحصائياً عند مستوى ٠ر٠٥ بين الصورتين (٣) و (١)، وعند مستوى ٠ر٠١ بين الصورتين (٣) و (٢) ولصالح الصورة (٣) في الحالتين، مما يعني أنه عندما يتوزع موقع (ترتيب) البديل الصحيح بين بدائل كل مفردة من مفردات الاختبار عشوائياً، فإنَّه متوسط مستويات الصعوبة للمفردات يكون أعلى منه عما في الصورتين الأخريين.

الفرضية الرابعة – يوضح الجدول (٩) نتائج تحليل التباين لاختبار صحة الفرضية الرابعة التي تنصّ على أنه توجد فروق دالة إحصائياً بين متوسطات معاملات التمييز لمفردات الاختبار تعزى إلى اختلاف موقع (ترتيب) البديل الصحيح بين بدائل كل مفردة من مفرداته:

جدول (٩)

نتائج تحليل التباين الأحادي (one-way ANOVA)

بين متوسطات معاملات التمييز لمفردات الاختبار (النموذج الثاني) بصورة الثلاث

مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	قيمة "ف"	مستوى الدلالة
بين المجموعات	٢٦٤١	٢	١٣٢٠ر٥٠	١١ر٠٧٨	٠ر٠٠١
داخل المجموعات	١٣١١٢	١١٠	١١٩ر٢٠		
المجموع	١٥٧٥٣	١١٢			

وكما يتضح من جدول (٩) فإنَّ قيمة "ف" دالة إحصائياً عند مستوى ٠ر٠٠١ مما يؤكد صحة الفرضية الرابعة، بما يعني وجود فروق دالة إحصائياً بين متوسطات معاملات التمييز لمفردات الاختبار تعزى إلى اختلاف موقع (ترتيب) البديل الصحيح بين بدائل كل مفردة من مفرداته . ولمعرفة اتجاه الفروق الدالة، تمَّ التعرف على دلالة الفروق باستخدام اختبار " توكي" وكانت نتائجها كالتالي:

	مج ١	مج ٢	مج ٣
مج ١	–	١ر٩٣٢	٤ر٤٧٨*
مج ٢		–	٤ر٣٢٩*
مج ٣			–

* = دالة عند ٠ر٠٥ **= دالة عند ٠ر٠١ ***= دالة عند ٠ر٠٠١

وكما يتضح من نتائج استخدام اختبار "توكي" على النحو الموضح أعلاه ، فإنَّ الفروق الدالة إحصائياً كانت بين الصورة (٣) وكل من الصورتين (١) و(٢)، حيث كان الفرق دالاً إحصائياً عند مستوى ٠ر٠٥ بين الصورتين (٣) و (١) ، وعند مستوى ٠ر٠٥ أيضاً. بين الصورتين (٣) و (٢) ولصالح الصورة (٣) في الحالتين، مما يعني أنه عندما يتوزع موقع (ترتيب) البديل الصحيح بين بدائل كل مفردة من مفردات الاختبار عشوائياً، فانه متوسط مستويات القدرة التمييزية للمفردات يكون أعلى منه عماً في الصورتين الأخريين.

ملخص نتائج الدراسة:

في ضوء ما سبق عرضه من نتائج أسفرت عنها الدراسة الحالية، يمكن إيجاز أهم هذه النتائج في الآتي :

(١) يزداد صدق وثبات الاختبار (من نوع الاختيار من متعدد) ككل عندما يكون الاختبار ثلاثي البدائل عما لو كان أكثر من ذلك (رباعي أو خماسي البدائل .... الخ).

(٢) يزداد صدق وثبات الاختبار (من نوع الاختيار من متعدد) ككل عندما يتوزع البديل الصحيح عشوائياً بين بدائل كل مفردة من مفردات الاختبار .

(٣) تكون مفردات الاختبار (من نوع الاختيار من متعدد) أكثر صعوبة كلما ازداد عددالبدائل التي تضّمها المفردة.

(٤) تكون مفردات الاختبار (من نوع الاختيار من متعدد) أعلى من حيث مستوى قدرتهاالتمييزية كلما ازداد عدد البدائل التي تضّمها المفردة.

(٥) تكون مفردات الاختبار (من نوع الاختيار من متعدد) أكثر صعوبة عندما يتوزع البديل الصحيح عشوائيا بين بدائل مفردات الاختبار.

(٦) تكون مفردات الاختبار (من نوع الاختيار من متعدد) أعلى من حيث مستوى قدرتهاالتمييزية عندما يتوزع البديل الصحيح عشوائياً بين بدائل مفردات الاختبار.

١١- توصيات ومقترحات لدراسات مستقبلية :

تقترح الباحثة إجراء المزيد من الدراسات ذات العلاقة بقضية الدراسة الحالية حول ما يلي على سبيل المثال:

١١-١- دراسة مقارنة لأثر استخدام الأنواع الأخرى من الاختبارات الموضوعية (مثل الصواب أو الخطأ - المزاوجة / المقابلة - الإكمال) على الخصائص السيكومترية للاختبار ككل ولكل مفردة من مفرداته.

١١-٢- دراسة للصعوبات التي تواجه مُعّدي الاختبارات من نوع الاختيار من متعدد عند صياغة مفردات هذا النوع من الاختبار صياغة دقيقة محكمة باللغة العربية.

١١-٣- دراسة أثر كل من عدد البدائل وترتيبها على الخصائص السيكومترية لمفردات الاختبار وفقاً لنظرية الاستجابة للمفردة (IRT).

١١-٤- دراسة مقارنة لنوْعَي الاختبارات (المقالية والموضوعية) من حيث دقة وموضوعية القياس والتصحيح وفقاً لكل من نظريتي القياس التقليدية والمعاصرة (IRT).

١١-٥- الأخطاء الشائعة لدى صياغة مفردات الاختبارات الموضوعية وأثرها على الخصائص السيكومترية للاختبار.

١١-٦- دراسة لاقتراح أنسب الأساليب الإحصائية لتصحيح الاستجابات من نوع الاختيار من متعدد وتنقيتها من أثر التخمين.

المراجع

أولا – مراجع باللغة العربية :

(١) الأمير، بندر بن زيد. (٢٠١٣). ***الأخطاء الشائعة في بناء مفردات الاختيار من متعدد وتأثيرها على الخصائص السيكومترية للاختبارات التحصيلية لدى معلمات الرياضيات بالمرحلة المتوسطة بمدينة مكة.*** رسالة ماجستير، كلية التربية، جامعة أم القرى.

(٢) الرباعي، إبراهيم محمد متعب. (٢٠١٢). ***أثر عدد البدائل وتغيير موقع الممّوه القوي في فقرات اختبار الاختيار من متعدد على الخصائص السيكومترية للاختبار ومعالم الفقرات وقدرات الأفراد.*** رسالة ماجستير، الجامعة الأردنية.

(٣) الردادي، عبد العالي سليمان. (٢٠٠٦). ***أثر اختلاف نمط الاختبار على الخصائص السيكومترية للاختبار وفقراته*** ، رسالة ماجستير، كلية التربية، جامعة أم القرى.

(٤) سليمان، شاهر خالد. (٢٠٠١). ***أثر فنية فقرات الصح والخطأ والتكميل على الخصائص السيكومترية للاختبار وفقراته.*** رسالة دكتوراه، جامعة الخرطوم.

(٥) الشبلي، على مصطفى موسى. (١٩٩٥). ***أثر عدد البدائل في اختبار تحصيلي من نوع الاختيار من متعدد على تجانسه الداخل والخصائص السيكومترية لفقراته*** . رسالة ماجستير، الجامعة الأردنية.

(٦) الشريفين، نضال و طعامنة ، إيمان. (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، مجلد (٥)، عدد(٤)، ص ٣٠٩ – ٣٣٥.

(٧) ظاظا، حيدر ابراهيم. (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، مجلد (٨) ، عدد(١) ، ص ٨١ – ٩١ .

(٨) ظاظا، حيدر إبراهيم. (١٩٩٥). ***المقارنة بين أثر أربع طرق لحذف المموهات في فقرات اختبار الاختيار من متعدد على الخصائص السيكومترية للاختبار وفقراته،*** الجامعة الأردنية.

(٩) العريفي، هدى محمد (٢٠٠٧). ***تباين الخصائص السيكومترية لاختبار متعدد البدائل بتباين عدد مفرداته وبعض الخصائص الدافعية للمفحوصين*** . رسالة ماجستير، كلية التربية، جامعة أم القرى.

(١٠) العضايلة، بسام محمود عبد الله. (٢٠١١). ***أثر عدد البدائل في الأسئلة الموضوعية على الخصائص السيكومترية للفقرة والاختبار*** . رسالة ماجستير، الجامعة الأردنية.

(١١) العطوي، ماجد بخيت. (٢٠١١). ***أثر عدد المموّهات الجذابة لفقرات الاختيار من متعدد في الخصائص السيكومترية للفقرات والاختبار*** . رسالة ماجستير، الجامعة الأردنية .

(١٢) عياصرة ، على مصطفي الشبلي. (١٩٩٥). ***أثر عدد البدائل في اختبار تحصيلي من نوع الاختيار من متعدد على تجانسه الداخلي والخصائص السيكومترية لفقراته.*** الجامعة الأردنية.

(١٣) عيلبوني ، سمير فؤاد. (٢٠٠٧). ***أثر عدد البدائل وتمييز المموّهات في فقرات الاختيار من متعدد على التوافق مع نموذج ثلاثي المعلم.*** رسالة ماجستير، الجامعة الأردنية .

(١٤) الغامدي ، سعيد حسن. (٢٠٠٠). ***مدى اختلاف الخصائص السيكومترية لأداة القياس في ضوء تغاير بدائل الاستجابة والمرحلة الدراسية (دراسة حالة مقياس ليكرت).*** رسالة ماجستير، كلية التربية، جامعة أم القرى.

(١٥) الغامدي، عبد الله أحمد سعيد. (٢٠٠٩). ***أثر عدد البدائل في الخصائص السيكومترية للاختبار التحصيلي في الرياضيات.*** رسالة ماجستير، كلية التربية، جامعة أم القرى.

(١٦) فقوسه، محمد عبد الله. (١٩٩٤). ***أثر عدد البدائل في فقرات اختبار الاختيار من متعدد على الخصائص السيكومترية للاختبار وفقراته.*** رسالة ماجستير، الجامعة الأردنية.

ثانيا – مراجع باللغة الإنجليزية :

- Bruno, J. &Dirkzwager, A. (1995). Determining the Optimal Number of Alternatives to a Multiple-Choice Test Item: An Information Theoretic Perspective. *Educational and Psychological Measurement*, vol. 55, no. 6, p. 595 – 60.

- Cizek, G. &O'Day, D. (1994). Further Investigation of Nonfunctioning Options in Multiple-Choice Test Items.

Educational and Psychological Measurement, vol. 54, no. 4, p. 861 – 72.

- Crehan, K. and Others (1993). Use of an Inclusive Option and the Optimal Number of Options of Multiple-Choice Items. *Educational and Psychological Measurement*, vol. 53, no.1, p. 241 – 47.

- Cristina, T.; Lopes, A.; Babo, L. &Azevedo, J. (2011). *Improving Multiple- Choice Questions*. US-China Education Review B 1, p. 1 – 11.

- Desrochers, M.; Fink, H.; Thomas, A.; Kimmerling, J. & Tung, W. (2007). Student Assessment: A Comparison of Solitary, Cooperative, and Competitive Testing. *International Journal of Teaching and Learning in Higher Education*, vol. 19, no. 3, p. 289 – 96.

- Dibattista, D. &Kurzawa, L. (2011). *Examination of the Quality of Multiple- Choice Items on Classroom Tests*. The Canadian Journal for the Scholarship of Teaching and Learning, vol. 2, Iss. 2, Article 4.

- Frary, R. (1995). *More Multiple-Choice Item Writing Do's and Don'ts*. ERIC/AE Digest.

- Frary, R. (1991). The None-of-the-Above Option: An Empirical Study. *Measurement in Education,* vol.4, no.2, p. 115 – 24.

- Gross, L. (1994). *Logical Versus Empirical Guidelines for Writing Test Items: The Case of "None of the Above".*

Evaluation and the Health Professions, vol. 17, no. 1, p. 123 – 26.

- Kehoe, J. (1995). *Writing Multiple-Choice Test Items*. ERIC/AE Digest.

- Revuelta, J. (2004). *Analysis of Distractor Difficulty in Multiple-Choice Items.* Psychometrika, vol. 69, no. 2, p. 217 – 234.

- Schroeder, J.; Hurphy, K. &Holme, T. (2012). Investigating Factors that Influence Item Performance on ACS Exam. *Journal of Chemical Education*, vol. 89, no.3, p. 346 – 350.

- Shizukam T.; Takeuchi, O.; Yashinma, T. &Yoshizawa, K. (2006). *A Comparison of Three-and Four-Option English Tests for University Entrance Selection Purposes in Japan*. *Language Testing*, vol. 23, no.1, p. 35 – 57 .

- Trevisan, M. and Others (1994). Estimating the Optimum Number of Options per Item Using an Incremental Option Paradigm. *Educational and Psychological Measurement*, vol. 54, no. 1, p. 86 – 91.

- Trevisan, M. & Sax, G. (1991). *Estimating the Optimal Choice Format Using an Incremental Option Paradigm. Evaluative Report.*

<< وصل هذا البحث إلى المجلة بتاريخ ٢٠١٥/٨/١٠، وصدرت الموافقة على نشره بتاريخ ٢٠١٥/١١/٨>>