

برنامج تدريبي لصياغة مفردات الاختيار من متعدد وأثره في الاتجاه نحو بناء الاختبارات التحصيلية

د. محمد محمود عبدالوهاب

كلية التربية - جامعة المنيا - جمهورية مصر العربية
كلية التربية - جامعة أم القرى - المملكة العربية السعودية

الملخص

نظراً لما أشارت إليه نتائج الدراسات السابقة من تعدد الأخطاء التي يقع بها معدو الاختبارات التحصيلية في صياغة مفردات الاختيار من متعدد، وكذلك ما أشارت إليه نتائج الدراسة الاستطلاعية على طلاب الفرقة الرابعة في تخصص علم النفس بكلية التربية بجامعة المنيا من تعدد هذه الأخطاء، فقد هدف البحث إلى إعداد برنامج لتدريب هؤلاء الطلاب على مهارات وقواعد صياغة مفردات الاختيار من متعدد، ومعرفة أثر هذا البرنامج في الاتجاه نحو بناء الاختبارات التحصيلية. ولتحقيق هذا الغرض تم بناء قائمة لقواعد صياغة مفردات الاختيار من متعدد، ومقياس للاتجاه نحو بناء الاختبارات التحصيلية بالإضافة إلى البرنامج التدريبي، وتم الاعتماد على المنهج شبه التجريبي القائم على تصميم المجموعة الواحدة، وأسفرت النتائج عن انخفاض مستوى قدرة الطلاب على صياغة أسئلة الاختيار من متعدد قبل التدريب على قواعد الصياغة الصحيحة حيث بلغ هذا المستوى ٥٣,٩١٤٪، كما كان مستوى الاتجاه نحو بناء الاختبارات التحصيلية قبل التدريب محايداً، حيث بلغ هذا المستوى ٦٤,٦١٣٪، كما وجدت علاقة ارتباطية طردية بين قدرات الطلاب على صياغة مفردات الاختيار من متعدد واتجاهاتهم نحو بناء الاختبارات التحصيلية، كما أشارت النتائج إلى وجود فرق ذي دلالة إحصائية بين متوسطي قدرات الطلاب في صياغة الاختيار من متعدد قبل وبعد التدريب وكان هذا الفرق في اتجاه متوسط قدراتهم بعد التدريب، كما كان هناك فرق دال إحصائياً بين متوسطي اتجاهات الطلاب نحو بناء الاختبارات التحصيلية قبل وبعد التدريب وكان هذا الفرق في اتجاه متوسط اتجاهاتهم بعد التدريب، وقد أوصى البحث بضرورة تخصيص وقت أكبر لتدريب الطلاب على قواعد صياغة مفردات الاختيار من متعدد في أثناء الدراسة بكلية التربية، وضرورة العمل على تنمية اتجاهاتهم نحو بناء الاختبارات التحصيلية.

مقدمة

يهتم علماء التربية عموماً وعلماء القياس النفسي خصوصاً بضرورة تحري الدقة والعدالة والموضوعية في الاختبارات التحصيلية؛ لكونها أكثر أدوات التقويم استخداماً في التعليم العام والجامعي؛ وهذا الإهتمام المتزايد بهدف الوصول إلى أفضل قياس ممكن للأهداف التعليمية التي تسعى البرامج التعليمية إلى تحقيقها. وقد وضع علماء القياس النفسي عدة شروط أو معايير يجب أن يتسم بها الاختبار التحصيلي الجيد، مثل: الصدق والثبات والموضوعية وما يلزمهم من تحقق الشمولية ووضوح التعليمات ومناسبة زمن التطبيق... (حيمودة، ٢٠١٠؛ علام، ٢٠٠٠)، كما تناولوا فنيات وقواعد صياغة كل نوع من أنواع المفردات الاختبارية، والأخطاء التي تحدث في هذه الصياغة وكيفية تلافيها (مراد وسليمان، ٢٠١٢)، بل يكاد لا يخلو أي كتاب في القياس والتقويم - يستخدم في برامج تأهيل المعلمين في الجامعات والكليات المتوسطة - من إرشادات تتعلق بصياغة مفردات الاختبارات التحصيلية (عودة، ١٩٨٩).

ورغم ذلك فكثير من الدراسات وتقارير الندوات والاجتماعات المتعددة للجان جمعيات التقويم في العالم العربي تشير إلى وجود ضعف واضح في إعداد الاختبارات التحصيلية، وافتقار للجوانب الفنية عند بنائها وتطبيقه، فقد أكد الزهراني (٢٠٠٠) تدني درجة امتلاك معلمي الرياضيات بالمرحلة الابتدائية في المدينة المنورة للكفايات المعرفية الخاصة ببناء الاختبارات التحصيلية، بغض النظر عن عدد سنوات الخبرة التي يقضيها المعلمين في مهنة التدريس. كما توصلت الشيباب (٢٠٠٣) إلى تدني مستوى معرفة معلمات الاقتصاد المنزلي الأردنيات بأسس بناء الاختبارات وبالمفاهيم العلمية الخاصة بذلك، معللة ذلك بنقص الخبرة والتدريب الكافيين في مجال بناء الاختبارات وأنواع المفردات المختلفة وبخاصة تلك التي تقيس قدرات عقلية عليا، مثل: الاختيار من متعدد، وأكدت عباس (٢٠٠٨) قلة اهتمام واضعي امتحانات مادة التاريخ في كليات التربية بالعراق في الفترة من ١٩٩٨ وحتى ٢٠٠١ بمعايير

ومواصفات المفردات الجيدة، مع اكتفائهم بمفردات المقال دون وجود أية أنواع أخرى من المفردات. وأسفرت نتائج دراسة عبد المنعم (٢٠٠٩) عن وجود عديد من الأخطاء اللغوية والنحوية والمطبعية والفنية والعلمية في صياغة كثير من المفردات، وعدم وضوح التعليمات الخاصة بها، وعدم التحديد الدقيق للمطلوب منها. كما توصل مختار (٢٠٠٩) إلى أن أساتذة اللغتين الفرنسية والإنجليزية في الجامعات السودانية لا يلتزمون بشروط وقواعد بناء الاختبارات التحصيلية، وأن مفردات هذه الاختبارات تركز على المستويات المعرفية الدنيا في ضوء تصنيف بلوم، كما أكد أبو حطب وصادق (٢٠١٠) والحموري (٢٠١٠) وجود نقاط ضعف عديدة في إعداد الاختبارات التحصيلية، أهمها: انخفاض المستويات المعرفية للأهداف التي تقيسها المفردات، وعدم الاهتمام بإعداد جداول مواصفات للاختبارات، وكذلك عدم القدرة على صياغة مفردات التعرف، مثل: الصواب والخطأ، والاختيار من متعدد، والمزاوجة، والتصنيف؛ مما يدفعهم للجوء إلى اختبارات المقال، كما وجد أبو جراد (٢٠١١) أن نتائج كثير من البحوث أسفرت عن وجود قصور واضح في الاختبارات التي يعدها المعلمون بمديرية التربية والتعليم بغزة. وتوصل الثبتي (٢٠١٤) إلى انخفاض اتقان المعلمين للمهارات الأساسية اللازمة لإعداد الاختبارات التحصيلية لمقرر الحاسب الآلي في مدارس المرحلتين المتوسطة والثانوية بإدارة التربية والتعليم بالطائف؛ حيث بلغت نسبة المعلمين الذين أتقنوا هذه المهارات ٤,٢٪ فقط من المعلمين. وأكدت التويجري (٢٠١٥) عدم التزام معدي الامتحانات النهائية لشهادة الثانوية العامة في الكويت بمعايير المفردات الجيدة ووجود أخطاء في صياغة هذه المفردات، مثل: عدم تحديد المشكلة في متن المفردة بوضوح، وعدم تجانس البدائل.

وسعى الرويلي (٢٠٠٧) إلى دراسة مدى فعالية الإعداد التربوي بكليات المعلمين بالمملكة العربية السعودية في إكساب الطلاب مهارات القياس والتقويم المتعلقة ببناء الاختبارات التحصيلية، ووجد تدنياً ملحوظاً في درجة اكتساب الطلاب المعلمين لهذه المهارات بما تشمله من مهارات التخطيط للاختبار،

وكتابة المفردات، وتحليلها، وإحصائيات درجات الاختبار، كما توصل إلى أن درجة التركيز على هذه المهارات أثناء تدريس مادة التقويم التربوي كانت متوسطة من وجهة نظر الطلاب.

واختلفت نتائج بعض الدراسات عن ذلك، مثل: دراسة غليون (٢٠٠٩) التي أشارت إلى أن نسب تحقق معايير الصياغة الجيدة لمفردات الاختيار من متعدد يتراوح بين ٩٤ - ١٠٠٪ مما يدل على كفاءة أعضاء هيئة التدريس في التزامهم بمعايير الصياغة، ودراسة رضوان (٢٠١١) التي توصلت بتحليلها لاختبارات مادة الرياضيات بالتعليم الأساسي بدمشق إلى أن معظم المفردات تتمتع بصياغة جيدة.

وتعد مفردات الاختيار من متعدد من أكثر أنواع المفردات شيوعاً في الاختبارات (Amo-Salas, Arroyo-Jimenez, Bustos-Escribano, Fairn-Jimnez, 2014; López-Fidalgo, 2014؛ أبو حطب وصادق، ٢٠١٠)، ويعود ذلك إلى ما يحمله هذا النوع من مزايا، منها: تغطيته الواسعة للمحتوى المراد قياسه، وقصر الزمن المستغرق في الإجابة عنه، وسهولة تصحيحه (Meng Er, Ramamurthy & Pook, 2014; Rocha, Marche & Biere, 2013). كما أن مفردات الاختيار من متعدد المصاغة بعناية يمكنها أن تعكس نفس الأداء الذي تطلبه مفردات الاستدعاء الحر للإجابة في حالة تدريج أوزان بدائل الإجابة (أو ما يعرف بالتقدير الجزئي Partial Credit)، مع أفضلية مفردات الاختيار من متعدد عند التحليل الكمي والكيفي لاستجابات الطلاب (Lin, Singh, Rebello, Engelhardt & Singh, 2012)، كما أشار باغايي ودوراخشان (Baghaei Dourakhshan, 2016) أن مفردات الاختيار من متعدد المصاغة جيداً أكثر صدقاً وثباتاً في قياس المستويات المعرفية العليا، وأكد بني عطا والرابعي (٢٠١٣) وأبو فودة (٢٠١٤) وثوابيه (Thawabieh, 2016) أن هذه المفردات يمكنها شمول المحتوى بشكل جيد مع سهولة تطبيقها وموضوعية تصحيحها وتمتعها بدرجة عالية من الصدق والثبات عند قياس قدرات الأفراد.

وقد وضع الباحثون عدة قواعد أو شروط أو معايير لصياغة مفردات الاختيار من متعدد، حيث وضع هالادينا وداونينج ورودريجز (Haladyna, Downing & Rodriguez, 2012) (٣١) قاعدة لابد من اتباعها لكتابة مفردات الاختيار من متعدد، وذلك بعد مراجعة (٢٧) مرجعاً متخصصاً في القياس والتقويم، بينما ركز شيونج وبوكات (Cheung & Bucat, 2002) على (٨) قواعد فقط وقاما بوضع أمثلة للصياغات الخطأ التي تنتهك كل قاعدة من هذه القواعد الثمانية، إلا أن كولنيز (Collins, 2006) أشار إلى (٢٠) قاعدة في صياغة هذا النوع من المفردات صنفها إلى إرشادات متعلقة بالخصائص العامة للمفردة، وإرشادات متعلقة بمتن المفردة، وثالثة متعلقة ببدائل الإجابة، واختلف هال ومارشال (Hall & Marshall, 2013) عن كل ذلك، ورأوا أن عدد هذه القواعد (١٢) قاعدة فقط. وفي العالم العربي أشار حماد (٢٠١١) إلى (١٧) شرطاً يجب توافرها في مفردات الاختيار من متعدد، وأشار ظاظا (٢٠١٢) إلى (١٨) قاعدة من قواعد صياغة مفردات الاختيار من متعدد في بحثه الذي سعى فيه للكشف عن مدى انتهاك هذه القواعد في امتحانات الشهادة الثانوية بالأردن. ونظرا لتعدد هذه الرؤى حول هذه القواعد أو الشروط أو المعايير فثمة حاجة لتحليلها، للوقوف على قائمة بهذه القواعد التي يجب اتباعها عند صياغة هذه المفردات.

مشكلة البحث

نظراً لكون مهارة صياغة الأنواع المختلفة من مفردات الاختبارات التحصيلية من المهارات التي يجب إكسابها لطلاب كلية التربية؛ فإن البرامج الأكاديمية لإعداد معلمي المواد الدراسية المختلفة بكليات التربية قد تضمنت عدة معايير أكاديمية تتعلق بضرورة إكساب الطلاب هذه المهارات، ولعل المقررات في مجال القياس النفسي والتقويم التربوي أكثر المقررات التي تسهم في تحقيق هذا المعيار ومؤشراته المختلفة، وإذا كان طلاب برنامج إعداد معلم علم النفس أكثر طلاب كلية التربية دراسةً لمقررات في هذا المجال، مثل مقررات

القياس والتقويم، والاختبارات والمقاييس، والقياس النفسي، التي يدرسها الطلاب في الفرقتين الثالثة والرابعة، فإنه من المتوقع أن يكونوا أكثر الطلاب قدرة على بناء الاختبارات بشكل عام والتحصيلية بشكل خاص، كما يتوقع أن يكونوا أكثرهم قدرة على صياغة الأنواع المختلفة من مفردات الاختبارات، إلا أن الباحث لاحظ - من خلال تكليفه طلاب الفرقة الرابعة (١٥ طالباً وطالبة) بأن يصيغ كل منهم (١٥) مفردة من نوع الاختيار من متعدد في أي مقرر من المقررات التي يقومون بتدريسها في التربية العملية أو التي يدرسونها في الكلية - وجود قصور كبير في صياغة هذا النوع من المفردات؛ حيث تبين له كثرة وضع البديل "جميع ما سبق" كإجابة صحيحة في عدد كبير من المفردات بلغ عددها ١٢٧ مفردة من إجمالي عدد المفردات، التي بلغت ٢٢٥ مفردة أي بنسبة ٥٦,٤٤٤٪، كما وجد (٩٨) مفردة تحتوي على بدائل غير معقولة ظاهرياً أي بنسبة ٤٣,٥٥٦٪، وظهر النقل النصي من الكتاب في (٨٦) مفردة أي بنسبة ٣٨,٢٢٢٪، كما احتوت (٧٦) مفردة على بدائل متشابهين في المعنى أي بنسبة ٣٣,٧٧٨٪، وظهر اختلاف في طول بدائل الإجابة في (٦٧) مفردة أي بنسبة ٢٩,٧٧٨٪، وإمكانية الإجابة عن (٤٩) مفردة من خلال النمط اللغوي أي بنسبة ٢١,٧٧٨٪، كما وجدت (٣٨) مفردة تقيس معلومات غير جوهرية في المقرر بنسبة ١٦,٨٨٩٪، كما لوحظ وجود اختلاف في عدد بدائل المفردات بين ثلاثية ورباعية البدائل، وقد كان إجمالي عدد المفردات التي تحتوي على مخالفة واحدة أو أكثر (١٦٦) مفردة أي بنسبة ٧٣,٧٧٨٪، كما لم يوجد طالب واحد من بين الطلاب استطاع صياغة المفردات بدون أية مخالفة، مما يدل على ضعف قدرات الطلاب على الصياغة.

وقد تناول كثير من الباحثين القواعد التي يتم انتهاكها عند صياغة مفردات الاختيار من متعدد، فقد سعى تارنت وكنايريم وهايز ووير (Tarrant, Knierim, Hayes & Ware, 2006) إلى تحليل ٢٧٧٠ مفردة من مفردات الاختيار من متعدد طبقت على طلاب التمريض في الفترة من ٢٠٠١ وحتى ٢٠٠٥، وذلك من خلال ١٩ قاعدة، وأظهرت النتائج أن ٤٦,٢ من المفردات تقريباً تنتهك

القواعد، وأن أكثر من ٩٠ من المفردات كتبت لقياس مستوى معرّف متدن. كما قام ظاظا (٢٠١٢) بالكشف عن مدى مخالفة هذه القواعد في امتحان شهادة الثانوية العامة في الأردن من خلال مراجعة ٢٩١٠ مفردة في الأعوام من ٢٠٠٧ وحتى ٢٠١٠م، من خلال ١٨ قاعدة، وأظهرت النتائج أن ٤١٦ مفردة (أي ما نسبته ١٤,٣٪) تضمنت مخالفة واحدة على الأقل، وكذلك ٣٦٦ مفردة (أي ما نسبته ١٢,٦٪) تضمنت أكثر من مخالفة واحدة، كما تبين أن أكثر قاعدتين انتهاكاً كانتا ترتيب البدائل عمودياً، وترتيب البدائل منطقياً أو عددياً أو أبجدياً، وأوصى بتغيير مستوى تعليمات كتابة مفردات الاختيار من متعدد التي تقرها وزارة التعليم من مستوى الاسترشاد إلى مستوى الإلزام. وتوصل تارنت ووير ومحمد (Tarrant, Ware & Mohammed, 2009) إلى أن نسب مفردات الاختيار من متعدد التي تحتوي على عدد (صفر - ١ - ٢ - ٣) مشتتات فعالة في المفردات رباعية البدائل هي (١٢,٣ - ٣٤,٨ - ٣٩,١ - ١٣,٨) على الترتيب، واستنتجوا من ذلك أن المعلمين لديهم صعوبة في وضع مشتتات فعالة، مما جعلهم يوصون بضرورة مراجعة مفردات الاختبارات المعدة مسبقاً. كما لاحظ المهيدب (Al Muhaidib, 2010) عدة أخطاء في صياغة مفردات الاختيار من متعدد عند تحليله لعدد ٣٧٩ مفردة، فقد وجد ٨٠٪ من المفردات تحتوي على تفاصيل إضافية في الإجابة الصحيحة، ونسبة ٧٥٪ من المفردات تحتوي على البديل "كل ما سبق" كإجابة صحيحة، كما وجدت إجابات سلبية بنسبة (٣٥٪)، وإجابات غير مقنعة بنسبة (٢٨٪) وإجابات غير محددة بنسبة (٣٠٪)، وقد أثرت كل هذه الأخطاء بشكل سلبي في نتائج الطلاب. كما أشار الأمير (٢٠١١) إلى شيوع خمسة أخطاء في صياغة مفردات الاختيار من متعدد لدى معلمي الرياضيات بالمرحلة المتوسطة بمكة المكرمة، وهي عدم وضوح صياغة متن المفردة، والنقل النصي الحرفي لمتن المفردة من الكتاب، وعدم ترتيب البدائل ترتيباً تصاعدياً أو تنازلياً، وتكرار الكلمات في بدائل الإجابة، وعدم توزيع موضع البديل الصحيح توزيعاً عشوائياً. وتوصل نيدو كايو ولافلين ورس وهال (Nedeau-Cayo, Laughlin, Rus & Hall, 2013) إلى أن غالبية مفردات الاختيار

من متعدد تحتوي على أخطاء، وتركز على مستويي المعرفة والفهم، كما أشاروا إلى أن المزايا التي تتمتع بها مفردات الاختيار من متعدد قد تفقد إذا لم يتم صياغتها بشكل جيد.

ونادى جليسون والي وبيكر (Gleason, Alley & Baker, 2010) بضرورة فحص كل مفردات الاختيار من متعدد لتحديد ما إذا كان بها انتهاكاً أم لا؛ حتى لا يؤثر ذلك في تقدير أداء المفحوصين على هذه المفردات؛ إذ تؤدي الأخطاء في صياغة مفردات الاختيار من متعدد إلى كثير من المشكلات السيكومترية، مثل: استفادة بعض الطلاب من وجود مفاتيح أو منبهات للإجابة الصحيحة عن المفردة فيما يعرف بالقدرة على التخمين الذكي (حماد، ٢٠١٠)، وكذلك ظهور ما يعرف بتحيز المفردة، والتأثير في تقدير قدرات المفحوصين وفي التقديرات الخاصة بالأخطاء المعيارية لتقديرات القدرة (يعقوب وأبو فودة، ٢٠١٠). وكذلك عدم تحقق أحادية البعد في السمة المقيسة إذ أن القدرة على التخمين والاستفادة من أخطاء الصياغة ستؤثر أيضاً في التباين بين قدرات الأفراد (Hambleton, Swaminathan & Rogers, 1991)، كما يترتب عليها أيضاً التأثير سلباً في عدالة الاختبار، وموضوعيته، وجودته (ظاظا، ٢٠١٢)، وتؤثر بنية مفردات الاختيار من متعدد (محكمة البناء - تحتوي على مخالفات في الصياغة) على دقة تقديرات بارامترات الأفراد والمفردات وفق نظرية الاستجابة للمفردة؛ حيث تبين أن المفردات محكمة البناء قدمت تقديرات أفضل لقدرات الأفراد عند المستوى المتوسط من القدرة وفق النموذج الثنائي البارامتر (الشرفين والصبح، ٢٠١١)، وتعطي تقديرات أفضل عند المستوى المتوسط والمرتفع من القدرة وفق نموذج الاستجابة للمفردة ثلاثي البارامتر، ويؤدي تصحيح الأخطاء في صياغة المفردات واستبدال المشتتات غير الفعالة وزيادة المستوى المعرفي الذي تقيسه مفردات الاختيار من متعدد يؤدي إلى زيادة القدرة التمييزية لها (Ali & Ruit, 2015).

وبالتالي فإن أغلب الدراسات أثبتت أن وجود أخطاء في صياغة مفردات الاختيار من متعدد يؤثر سلباً في عملية القياس، ولم يشذ عن ذلك سوى نتائج

قليل من الدراسات، مثل: دراسة داونينج (Downing, 2005) التي توصلت إلى أن الصياغات الخطأ لمفردات الاختيار من متعدد لها تأثير قليل على ثبات درجات الاختبار وعلى جودة مؤشرات الخصائص السيكومترية الأخرى للاختبار.

وقد واجه الباحثون في مجال القياس النفسي كثرة هذه الأخطاء بتزويد المصادر المختلفة بالإرشادات الخاصة بكتابة هذا النوع من المفردات، إلا أنه وجد قليل من الجهد التجريبي في تطوير هذه المفردات (Neddeau-Cayo et al., 2013)، كما اهتم علماء التربية في العالم العربي عامة وعلماء القياس النفسي على وجه الخصوص بدراسة كثير من العوامل التي قد تؤثر في دقة النتائج المترتبة على استخدام هذا النمط من المفردات، مثل دراسات: (أبو جراد، ٢٠١٢؛ وأبو فودة، ٢٠١٤؛ وبني عطا والرباعي، ٢٠١٣؛ والزيون، ٢٠١٣؛ والشريفين والصبح، ٢٠١١؛ وعبد الوهاب، ٢٠١٣؛ والكحلوت، ٢٠٠٢؛ ومحمود، ٢٠١٥؛ ويعقوب وأبو فودة، ٢٠١٠) وغيرها من الدراسات، كما تناولت المراجع والمصادر المختلفة في القياس والتقويم مفردات الاختيار من متعدد من حيث أشكالها ومميزاتها وعيوبها وقواعد صياغتها والأخطاء الشائعة في هذه الصياغة، وعلى الرغم من ذلك فما تزال هناك أخطاء في صياغة المعلمين لهذه المفردات.

وأكد دي بالم (DePalma, 1990) أهمية التدريب على كتابة مفردات التعرف بشكل عام ومفردات الاختيار من متعدد بشكل خاص، وهذا ما دعا إليه عسيري (٢٠٠٢) عندما أكد ضرورة إعداد حقيبة تدريبية في مجال بناء واستخدام الاختبارات التحصيلية وإعداد بنوك أسئلة موضوعية ومقالية على درجة عالية من الصياغة الجيدة، كما نادى الصرايرة (٢٠١١) بضرورة تكثيف الدورات للمعلمين الجدد في إعداد وصياغة المفردات الممثلة لكافة المستويات العقلية خاصة العليا منها. كما أشار الزغلوط وطلافة والمحاميد (٢٠١٢) إلى ضرورة تدريب الطلاب على كيفية صياغة المفردات بشكل عام، بالإضافة إلى تعريفهم بالأهداف السلوكية وكيفية قياسها ونوع المفردات الملائمة لكل مستوى منها. وأكد شافيدا وميسرا ودوتاروي (Chavda, Misra & Duttaroy, 2015) أن

هناك أهمية كبيرة لتدريب المعلمين على إعداد مفردات الاختيار من متعدد التي تتمتع بصياغة جيدة، وكذلك ضرورة الاهتمام بتحليلها كميًا وكيفيًا. كما أشار محمود (Mahmood, 2015) إلى أن معدي المفردات الاختبارية بحاجة إلى تدريب وتصحيح لصياغة المفردات من أجل قياس السمات المختلفة بأكبر درجة من الدقة. وأوصت التوجيهي (٢٠١٥) بضرورة تكثيف الدورات التدريبية والتأهيلية في مجال التقويم التربوي مع ضرورة مساعدة المختصين في تحسين صياغتهم للمفردات بما تحقق معايير المفردات الجيدة، وإعادة النظر في مضمون مفردات الامتحانات السابقة.

وتضاربت نتائج التدريب على مهارات القياس والتقويم عامة وعلى مهارات بناء الاختبارات التحصيلية بشكل خاص في إكساب المعلمين أو الطلاب المعلمين لمهارات صياغة الأنواع المختلفة من المفردات الاختبارية؛ فقد سعى كامل (١٩٩٩) إلى استخدام استراتيجيات التعلم حتى يتمكن في تدريس مقرر "التقويم التربوي" لإكساب الطلاب المعلمين مهارات بناء الاختبارات التحصيلية، وأسفرت نتائج دراسته عن وصول أكثر من ٨٠٪ من الطلاب المعلمين إلى درجة الاتقان في أربع مهارات أساسية في بناء الاختبارات التحصيلية متمثلة في صياغة الأهداف التعليمية، وتصميم خطة الاختبار، وبناء وتصحيح المفردات، وتحليل البيانات المستمدة من المفردات. كما قام عبيد (٢٠٠٨) بدراسة أثر الدورات التدريبية التي تعقدها وزارة التربية والتعليم في حقل القياس والتقويم لمعلمي الرياضيات أثناء الخدمة على كفاياتهم في بناء الاختبارات التحصيلية، وأسفرت الدراسة عن وجود فاعلية لهذه الدورات في معرفة المعلمين بمعايير بناء الاختبارات، وأن مستوى المعلمين الذين تلقوا تدريباً في ممارسة تصميم وإخراج الاختبارات كان متوسطاً، كما تكونت لديهم اتجاهات إيجابية نحو الاختبارات التحصيلية. وتوصل المحمد (٢٠٠٩) إلى فعالية برنامج تدريبي لمعلمي العلوم الاجتماعية بدولة الكويت في تحسين مهارات بناء الاختبارات التحصيلية. وأكد الزهراني (٢٠٠٠) أن التدريب في مجال القياس والتقويم يحدث فرقاً دالاً إحصائياً بين معلمي الرياضيات بالمرحلة الابتدائية في المدينة

المنورة في درجة امتلاكهم للكفايات المعرفية الخاصة ببناء الاختبارات التحصيلية.

بينما اختلفت نتائج دراسة ملياني (٢٠٠٣) مع نتائج تلك الدراسات؛ فقد توصل إلى أن الدورات التي حضرها المعلمون في مجال القياس والتقييم لم تحدث فرقاً دالاً إحصائياً بين معلمي الرياضيات في مكة المكرمة في مستوى معرفتهم بقواعد بناء الاختبارات التحصيلية، وأوصى بضرورة إعادة النظر في هذه الدورات وتحسينها والتركيز فيها على بناء الاختبارات، واتفقت دراسة سنان (٢٠٠٦) مع دراسة ملياني (٢٠٠٣) في نفس هذه النتائج لمعلمات الرياضيات في مكة المكرمة أيضاً. وأشار كحلوت (٢٠٠٠) (في: بركات وصباح، ٢٠٠٧) إلى أن جهود المجتمعات العربية والعالمية في تطوير عملية القياس تتجه عادة نحو تطوير الجوانب الإدارية والتنظيمية للاختبارات من لوائح وقوانين وشروط تطبيق الاختبارات أكثر مما تتجه نحو تطوير الجوانب المنهجية والفنية والتقنية في تصحيح الاختبارات وإعداد المفردات.

ويلاحظ مما سبق أن البرامج التدريبية المقدمة للمعلمين كانت موجهة لتنمية مهارات القياس والتقييم بشكل عام أو مهارات بناء الاختبارات التحصيلية بشكل خاص، إلا أنه لم توجد برامج تدريبية أكثر خصوصية في إكساب المعلمين مهارات صياغة مفردات الاختيار من متعدد بشكل أكثر تفصيلاً، ولعل الضعف الواضح في بناء هذه المفردات وتعدد الأخطاء في صياغتها يرجع إلى قلة التدريب على ذلك، وقلة الأمثلة التوضيحية المقدمة للمعلمين أو للطلاب المعلمين عن الصياغات الخطأ لهذه المفردات أو الإجراءات الخطأ التي يقع فيها المعلمون أثناء إعدادهم إياها، كما قد يكون ناتجاً عن قلة ممارسة بعض المعلمين لهذا الإجراء؛ وبذلك فثمة حاجة إلى عقد دورات تدريبية أكثر تخصصية في صياغة مفردات الاختيار من متعدد؛ ولهذا يسعى الباحث إلى إعداد برنامج تدريبي لتنمية قدرة طلاب علم النفس بكلية التربية في جامعة المنيا على مهارات صياغة مفردات الاختيار من متعدد، ومعرفة أثر هذا التدريب في درجة اكتساب هذه المهارات.

ومن ناحية أخرى فقد وجد اختلاف ملحوظ في استخدام مصطلح الاتجاهات نحو الاختبارات، فيستخدم مع الطلاب بمعنى قلق الاختبار وهو شعور الطلاب بالتوتر والإضطراب أثناء أدائهم في هذه الاختبارات، كما في دراسة عبد الحافظ (١٩٨٦)، والهواري والشناوي (١٩٨٧)، وسايحي (٢٠١٢). كما استخدم بن لادن (٢٠٠٢) هذا المصطلح أحيانا لبيان قبول الطلاب أو رفضهم لأنواع الاختبارات المختلفة (الموضوعية والمقالية)، ويستخدم أحيانا أخرى مع الطلاب والمعلمين بمعنى مجموعة المشاعر والأفكار والآراء والإدراكات للمظاهر والعناصر المتعلقة بالامتحانات المدرسية من حيث إعدادها والجهد والتكاليف التي تتطلبها، ومدى استخدامها في اتخاذ القرارات التربوية، ومدى الاهتمام بها وتقدير دورها في العملية التربوية، كما في دراسة عودة (١٩٩٠) ودراسة أبو جراد (٢٠١١)، ولم توجد سوى دراسة واحدة تتناول مصطلح "الاتجاهات نحو بناء الامتحانات المدرسية" بإضافة كلمة "بناء" وهي دراسة أبو الريش (٢٠٠٩) ورغم ذلك فقد أشار إليه بأهمية الامتحانات وأثرها في العملية التعليمية ومدى تقبل الطلاب لها، وتقييم كفاءة هذه الاختبارات في قياس نواتج التعلم؛ أي بنفس رؤية عودة (١٩٩٠) وأبو جراد (٢٠١١).

ولكن يقصد بالاتجاه نحو بناء الاختبارات التحصيلية في هذا البحث مدى قبول أو رفض المعلم أو الطالب المعلم لبناء وإعداد الاختبارات التحصيلية بنفسه، ومقدار استعداده وتهيئه النفسي لإعداد الاختبارات التحصيلية الجيدة في مادة تخصصه لتقويم أداء طلابه.

ونظراً لما أشار إليه عودة (١٩٩٠) من أن اتجاهات المعلمين نحو الاختبارات لها أثر قوي في فعالية ممارساتهم نحو تحقيق الأغراض المختلفة للاختبارات، وما أكدته جرين وستيجر (Green & Stager, 1986) بأن تحقيق المعلم للأغراض المختلفة للاختبارات التحصيلية يقتضي رفع مستوى الاتجاه الإيجابي نحوها، فإن اتجاه المعلمين نحو بناء الاختبارات التحصيلية قد يؤثر في اتقانهم لعملية بناء هذه الاختبارات والعكس، وهذا يستدعي دراسة اتجاهات المعلمين والطلاب المعلمين نحو بناء الاختبارات التحصيلية.

وبذلك فإن هناك حاجة إلى تدريب عينة من الطلاب على صياغة مفردات الاختيار من متعدد، ومعرفة أثر هذا التدريب في اتجاهاتهم نحو بناء الاختبارات التحصيلية، ومن ثم فإن مشكلة البحث الحالي تثير التساؤلين الرئيسين: "ما أثر البرنامج التدريبي المقترح في تنمية قدرة طلاب علم النفس على صياغة مفردات الاختيار من متعدد؟ وما أثره في اتجاهاتهم نحو بناء الاختبارات التحصيلية؟"، ويتفرع من هذين التساؤلين التساؤلات الآتية:

- ١ - ما مستوى قدرة طلاب علم النفس على صياغة مفردات الاختيار من متعدد قبل التدريب على صياغة مفردات الاختيار من متعدد؟
- ٢ - ما أكثر أخطاء صياغة مفردات الاختيار من متعدد شيوعاً لدى طلاب علم النفس قبل التدريب على صياغة مفردات الاختيار من متعدد؟
- ٣ - ما مستوى اتجاهات طلاب علم النفس نحو بناء الاختبارات التحصيلية قبل التدريب على صياغة مفردات الاختيار من متعدد؟
- ٤ - ما الدلالة الإحصائية للعلاقة الارتباطية بين القدرة على صياغة مفردات الاختيار من متعدد واتجاهاتهم نحو بناء الاختبارات التحصيلية قبل التدريب؟
- ٥ - ما الدلالة الإحصائية للفروق بين متوسطي درجات طلاب علم النفس في القدرة على صياغة مفردات الاختيار من متعدد قبل وبعد التدريب على صياغتها؟
- ٦ - ما الدلالة الإحصائية للفروق بين متوسطي درجات طلاب علم النفس في اتجاهاتهم نحو بناء الاختبارات التحصيلية قبل وبعد التدريب على صياغة مفردات الاختيار من متعدد؟

أهداف البحث

يسعى البحث الحالي إلى:

- ١ - إعداد برنامج تدريبي لصياغة مفردات الاختيار من متعدد، والتحقق من فعاليته.

- ٢ - تعرف مستوى قدرة طلاب علم النفس على صياغة مفردات الاختيار من متعدد قبل التدريب على صياغة مفردات الاختيار من متعدد.
- ٣ - تحديد أكثر أخطاء صياغة مفردات الاختيار من متعدد شيوعاً لدى طلاب علم النفس قبل التدريب على صياغة مفردات الاختيار من متعدد.
- ٤ - تعرف مستوى اتجاه الطلاب نحو بناء الاختبارات التحصيلية قبل التدريب على صياغة مفردات الاختيار من متعدد.
- ٥ - تحديد قيمة واتجاه العلاقة الارتباطية بين القدرة على صياغة مفردات الاختيار من متعدد واتجاهاتهم نحو بناء الاختبارات التحصيلية قبل التدريب.
- ٦ - معرفة أثر البرنامج التدريبي على قواعد صياغة مفردات الاختيار من متعدد في الاتجاه نحو بناء الاختبارات التحصيلية.

أهمية البحث

تتمثل أهمية البحث الحالي فيما يلي:

- ١ - قد يكون هذا البحث بداية لمزيد من البرامج التدريبية المتخصصة في صياغة الأنواع المختلفة من المفردات، ومعرفة أثرها في متغيرات سيكولوجية أخرى تؤثر في العملية التربوية.
- ٢ - يقدم هذا البحث برنامجاً تدريبياً قد يساهم في تنمية قدرة الطلاب المعلمين في صياغة مفردات الاختيار من متعدد، ومن ثم فقد يكون جزءاً من أجزاء محتوى مقرر القياس والتقويم أو الاختبارات والمقاييس أو القياس النفسي الذي يقدم لطلاب كلية التربية.
- ٣ - قد يثري هذا البحث مراكز تطوير نظم الامتحانات وتقويم الطلاب بالجامعات ببرنامج تدريبي يمكن استخدامه لتنمية قدرة أعضاء هيئة التدريس على صياغة مفردات الاختيار من متعدد.

٤ - قد يغير هذا البحث من وجهة نظر لجان توصيف مقررات كلية التربية في الموضوعات التي يجب تناولها في المقررات المختلفة في مجال القياس والتقويم.

حدود البحث

تتحدد نتائج البحث الحالي بما يلي:

- ١ - عينة قدرها ٣٦ طالباً من طلاب الفرقة الرابعة بتخصص علم النفس بكلية التربية جامعة المنيا، في العام الجامعي ٢٠١٤ - ٢٠١٥م.
- ٢ - مقياس الاتجاه نحو بناء الاختبارات التحصيلية (من إعداد الباحث).
- ٣ - قائمة تقييم مفردات الاختيار من متعدد (من إعداد الباحث).
- ٤ - برنامج تدريبي على صياغة مفردات الاختيار من متعدد (من إعداد الباحث).

مصطلحا البحث

القدرة على صياغة مفردات الاختيار من متعدد: تعني قدرة طلاب علم النفس على صياغة مفردات اختيار من متعدد خالية من أخطاء الصياغة ومراعية للقواعد والمعايير، وتقدر من خلال متوسط عدد القواعد التي تم مراعاتها في صياغة كل مفردة من المفردات التي يكلف ببنائها في أي مقرر من مقررات علم النفس، ويتم قياسه من خلال قائمة تقييم صياغة مفردات الاختيار من متعدد (من إعداد الباحث).

الاتجاه نحو بناء الاختبارات التحصيلية: ويعني مدى قبول أو رفض الطلاب واستعدادهم وتهيؤهم النفسي لبناء الاختبارات التحصيلية الجيدة، ويقدر هذا الاتجاه من خلال المقياس الذي أعده الباحث لهذا الغرض.

أدوات البحث

تتمثل أدوات البحث الحالي فيما يلي:

أولاً - قائمة تقييم صياغة مفردات الاختيار من متعدد: وتم إعدادها والتأكد من صلاحيتها للقياس من خلال الخطوات الآتية:

١ - الإطلاع على قواعد صياغة مفردات الاختيار من متعدد في الأدبيات والمراجع المختلفة مثل: (Haladyna, Downing & Rodriguez, 2002; Cheung & Bucat, 2002; Collins, 2006; حماد، ٢٠١١؛ ظاظا، ٢٠١٢).

٢ - وضع قائمة لقواعد صياغة مفردات الاختيار من متعدد مكونة من (٢٢) قاعدة، تتوزع على ثلاثة أبعاد أساسية هي قواعد عامة لصياغة المفردات، وعددها (٦) قواعد، وقواعد خاصة بمثن المفردة وعددها (٥) قواعد، وقواعد خاصة ببدائل الإجابة، وعددها (١١) قاعدة.

٣ - التأكد من صدق قائمة تقييم مفردات الاختيار من متعدد من خلال:

أ - عرض القائمة على (٧) محكمين من أعضاء هيئة التدريس بقسم علم النفس بكلية التربية بجامعة المنيا؛ لمعرفة آرائهم في هذه القائمة، وقد وافق المحكمون على مفردات القائمة مع بعض التعديلات في الصياغة وتمت مراعاتها.

ب - تم حساب الاتساق الداخلي لمفردات القائمة، على النحو الآتي:

- تكليف عدد (١٥) طالباً من طلاب الفرقة الرابعة تخصص علم النفس في العام الجامعي ٢٠١٣ - ٢٠١٤م بوضع (١٥) مفردة من نوع الاختيار من متعدد في أحد مقررات علم النفس التي يدرسونها أو التي يقومون بتدريسها للطلاب في التربية العملية؛ ليصبح عدد المفردات (٢٢٥ = ١٥ × ١٥) مفردة، ثم تحليل هذه المفردات باستخدام قائمة تقييم صياغة مفردات الاختيار من متعدد التي تم وضعها، مع الاستعانة بزميل آخر^(١) يقوم بتحليل هذه المفردات في نفس الوقت، وتم التأكد من ثبات عملية التحليل بحساب عدد مرات الاتفاق وعدد مرات الاختلاف بين الباحث والمحلل الآخر على كل

(١) د. محمد عبدالعزيز نور الدين: مدرس علم النفس بكلية التربية، جامعة المنيا.

مفردة من مفردات قائمة تقييم صياغة مفردات الاختيار من متعدد وحساب معامل الاتفاق (بين المحللين) باستخدام معادلة كوبر، التي فيها: نسبة الاتفاق = عدد مرات الاتفاق / (عدد مرات الاتفاق + عدد مرات الاختلاف)، وتراوحت نسب الاتفاق في كل قاعدة من القواعد بين (٠,٨٦٧ : ١)، مما يدل على ثبات عملية التحليل باستخدام القائمة.

- تم حساب الاتساق الداخلي لمفردات القائمة من خلال حساب معامل الارتباط المصحح^(٢) بين كل مفردة من مفردات القائمة والبعد الذي تقيسه، كما هو موضح في جدول رقم (١).

جدول رقم (١)

معاملات الارتباط المصححة بين المفردة والبعد الذي تقيسه في قائمة تقييم صياغة مفردات الاختيار من متعدد (عدد المفردات موضع التقييم = ٢٢٥)

رقم المفردة	معامل الارتباط	رقم المفردة	معامل الارتباط	رقم المفردة	معامل الارتباط	رقم المفردة	معامل الارتباط
١	٠,٥٣٦	٧	٠,٤٢٥	١٣	٠,٥٤١	١٩	٠,٦٤٥
٢	٠,٨٢٧	٨	٠,٦٨٦	١٤	٠,٤٢٥	٢٠	٠,٣٣٥
٣	٠,٦٦٩	٩	٠,٣٣٩	١٥	٠,٥٥٦	٢١	٠,٥٥٦
٤	٠,٤٢٥	١٠	٠,٥٤١	١٦	٠,٣٨٨	٢٢	٠,٣٩٦
٥	٠,٥٥٨	١١	٠,٦٣٣	١٧	٠,٧٢١		
٦	٠,٦٤٦	١٢	٠,٨٢٤	١٨	٠,٤٤٢		

(♦♦) دال إحصائياً عند مستوى دلالة (٠,٠١).

ويتضح من جدول رقم (١) أن معاملات ارتباط جميع المفردات بالبعد الذي تقيسه دالة إحصائياً عند مستوى (٠,٠١)، وتزيد عن (٠,٣) وفق محك جيلفورد؛ مما يدل على الاتساق الداخلي لمفردات القائمة.

(٢) أي بعد حذف درجة المفردة من الدرجة الكلية للبعد.

- ٤ - تم التحقق من ثبات قائمة صياغة مفردات الاختيار من متعدد من خلال حساب معامل كيوذر ريتشاردسون وكانت قيمته (٠,٧٣٦) وهي قيمة تدل على تمتع القائمة بمعامل ثبات مقبول.
- ثانياً - مقياس الاتجاه نحو بناء الاختبارات التحصيلية: وتم بناؤه والتأكد من صلاحيته للقياس من خلال الخطوات الآتية:
- ١ - الإطلاع على الأدبيات التي تناولت قياس الاتجاهات بشكل عام، والتي تناولت متغيرات الاتجاه نحو الاختبارات التحصيلية أو الاتجاه نحو بناء الاختبارات التحصيلية بشكل خاص، والتي سبقت الإشارة إليها، والاستفادة منها في بناء المقياس الحالي.
- ٢ - تم تصنيف الاتجاه نحو بناء الاختبارات التحصيلية إلى ثلاثة أبعاد، هي:
- أ - المعرفة بأسس بناء الاختبارات التحصيلية: ويمثل البعد المعرفي في الاتجاه ويتعلق بمعرفة الأسس والخطوات التي تبني في ضوءها الاختبارات التحصيلية وأهميتها والقواعد العامة في بنائها.
- ب - مهارة صياغة مفردات الاختبارات وتعليماتها: ويمثل البعد المهاري ويتعلق بمهارة وضع مفردات متنوعة في مستوى صعوبتها، وقادرة على التمييز بين الممتحنين، ومصاغة بطريقة جيدة، وتناسب الأهداف التعليمية المراد قياسها، وكذلك القدرة على صياغة تعليمات واضحة للأسئلة المختلفة وللأختبار بشكل عام.
- ج - الاستمتاع والرغبة في بناء الاختبارات التحصيلية: ويمثل البعد الوجداني ويتعلق بميول الطالب لبناء الاختبارات واستمتاعه بذلك، وثقته بنفسه وبقدرته على تحسين الاختبارات التي بها أخطاء.
- ٣ - تم وضع عدد (٨) مفردات لقياس كل بعد من الأبعاد الثلاثة السابقة.
- ٤ - تم التأكد من صدق مقياس الاتجاه نحو الاختبارات التحصيلية من خلال:
- أ - عرضه على (٧) محكمين من أعضاء هيئة التدريس بقسم علم النفس بكلية التربية بالمنيا، وقد أبدى أربعة منهم اعتراضهم على

مفردة واحدة في بعد المعرفة بأسس بناء الاختبارات التحصيلية، وهي: "أطلع على التصنيفات المختلفة لمستويات الأهداف التعليمية"، وبالتالي أصبح عدد مفردات الأبعاد المختلفة (٧، ٨، ٨) على الترتيب.

ب - تم تطبيق المقياس على عينة مكونة من (٣٩) طالباً في الفرقة الرابعة بتخصص علم النفس في العام الجامعي ٢٠١٣ - ٢٠١٤م، وحساب الاتساق الداخلي لمفردات المقياس من خلال حساب معامل الارتباط المصحح بين مفردات كل بعد من أبعاد المقياس والدرجة الكلية للبعد بعد حذف درجة المفردة من البعد، وكانت قيم معاملات الارتباط كما في جدول رقم (٢).

جدول رقم (٢)

معامل الارتباط بين المفردة والبعد الذي تقيسه في مقياس الاتجاه نحو بناء الاختبارات التحصيلية (ن = ٣٩)

معامل الارتباط بالمعزل	رقم المفردة	معامل الارتباط بالمعزل	رقم المفردة	معامل الارتباط بالمعزل	رقم المفردة	معامل الارتباط بالمعزل	رقم المفردة
٠,٣٨١	١٩	٠,٣٦٥	١٣	٠,٣٥١	٧	٠,٥٢٢	١
٠,٣٦٥	٢٠	٠,٤٤٥	١٤	٠,٣٧٨	٨	٠,٤٥٥	٢
٠,٣٥١	٢١	٠,٥٢٢	١٥	٠,٦٥٣	٩	٠,٣٦٨	٣
٠,٤٨٥	٢٢	٠,٤٥٥	١٦	٠,٥٤٥	١٠	٠,٣٦٥	٤
٠,٥٦٤	٢٣	٠,٥٤٥	١٧	٠,٤٥٥	١١	٠,٤٨٥	٥
		٠,٦٥٣	١٨	٠,٣٥٥	١٢	٠,٥٢٢	٦

(♦♦) دال إحصائياً عند مستوى دلالة (٠,٠١)، (♦) دال عند مستوى (٠,٠٥).

ويتضح من جدول رقم (٢) أن معاملات الارتباط المصححة دالة إحصائياً عند مستوى الدلالة (٠,٠٥، ٠,٠١). وتزيد عن (٠,٣) وفق محك جيلفورد؛ مما يدل على الاتساق الداخلي لمفردات مقياس الاتجاه، كما تم حساب

معاملات الارتباط المصححة بين الأبعاد الثلاثة والدرجة الكلية للمقياس، وكانت مساوية (٠,٨٧٦، ٠,٧٧٩، ٠,٨٣٢) على الترتيب، وهي جميعها مرتفعة ودالة إحصائياً عند مستوى الدلالة (٠,٠١).

ج - صدق تمييز المفردات: حيث تم تحديد مجموعتين من الطلاب إحداهما تمثل أعلى ٢٧٪ والأخرى تمثل أدنى ٢٧٪ من الطلاب في الدرجة الكلية للاتجاه نحو بناء الاختبارات التحصيلية، وقد بلغ عدد كل فئة من الفئتين (١١) طالباً، وتمت المقارنة بين استجاباتهما على كل مفردة من المفردات، وذلك باستخدام اختبار مان ويتني، ووجدت فروق ذات دلالة إحصائية بين المجموعتين في كل مفردة من المفردات؛ حيث كانت نتائج اختبار مان ويتني للمقارنة بين المجموعتين كما في جدول رقم (٣).

جدول رقم (٣)

نتائج اختبار مان ويتني بين المجموعتين العليا والدنيا في مفردات مقياس الاتجاه نحو الاختبارات التحصيلية

رقم المفردة	المجموعة العليا (ن = ١١)		المجموعة الدنيا (ن = ١١)		قيمة ي	مستوى الدلالة
	متوسط الرتب	مجموع الرتب	متوسط الرتب	مجموع الرتب		
١	١٦,٨٦	١٨٥,٥٠	٦,١٤	٦٧,٥٠	١,٥٠٠	٠,٠٠١
٢	١٧,٠٠	١٨٧,٠٠	٦,٠٠	٦٦,٠٠	٠,٠٠١	٠,٠٠١
٣	١٧,٠٠	١٨٧,٠٠	٦,٠٠	٦٦,٠٠	٠,٠٠١	٠,٠٠١
٤	١٦,٨٢	١٨٥,٠٠	٦,١٨	٦٨,٠٠	٢,٠٠٠	٠,٠٠١
٥	١٦,٠٩	١٧٧,٠٠	٦,٩١	٧٦,٠٠	١٠,٠٠١	٠,٠٠١
٦	١٧,٠٠	١٨٧,٠٠	٦,٠٠	٦٦,٠٠	٠,٠٠١	٠,٠٠١
٧	١٦,٧٣	١٨٤,٠٠	٦,٢٧	٦٩,٠٠	٣,٠٠٠	٠,٠٠١
٨	١٦,٦٨	١٨٣,٥٠	٦,٣٢	٦٩,٥٠	٣,٥٠٠	٠,٠٠١
٩	١٦,٠٦	١٧٧,٠٠	٦,٩١	٧٦,٠٠	١٠,٠٠١	٠,٠٠١
١٠	١٦,٦٤	١٨٣,٠٠	٦,٣٦	٧٠,٠٠	٤,٠٠٠	٠,٠٠١
١١	١٦,٨٢	١٨٥,٠٠	٦,١٨	٦٨,٠٠	٢,٠٠٠	٠,٠٠١

تابع/ جدول رقم (٣)

نتائج اختبار مان ويتني بين المجموعتين العليا والدنيا في مفردات مقياس الاتجاه نحو الاختبارات التحصيلية

رقم المفردة	المجموعة العليا (ن = ١١)		المجموعة الدنيا (ن = ١١)		قيمة ي	مستوى الدلالة
	متوسط الرتب	مجموع الرتب	متوسط الرتب	مجموع الرتب		
١٢	١٥,٥٥	١٧١,٠٠	٧,٤٥	٨٢,٠٠	١٦,٠٠٠	٠,٠٠٣
١٣	١٦,٠٩	١٧٧,٠٠	٦,٩١	٧٦,٠٠	١٠,٠٠١	٠,٠٠١
١٤	١٦,٥٠	١٨١,٥٠	٦,٥٠	٧١,٥٠	٥,٥٠٠	٠,٠٠١
١٥	١٦,٨١	١٧٨,٠٠	٦,٨٢	٧٥,٠٠	٩,٠٠٠	٠,٠٠١
١٦	١٧,٠٠	١٨٧,٠٠	٦,٠٠	٦٦,٠٠	٠,٠٠١	٠,٠٠١
١٧	١٥,١٤	١٦٦,٥٠	٧,٨٦	٨٦,٥٠	٢٠,٥٠٠	٠,٠٠٦
١٨	١٦,٧٧	١٨٤,٥٠	٦,٢٣	٦٨,٥٠	٢,٥٠٠	٠,٠٠١
١٩	١٦,٥٩	١٨٢,٥٠	٦,٤١	٧٠,٥٠	٤,٥٠٠	٠,٠٠١
٢٠	١٦,٧٣	١٨٤,٠٠	٦,٢٧	٦٩,٠٠	٣,٠٠٠	٠,٠٠١
٢١	١٥,٤١	١٦٩,٥٠	٧,٥٩	٨٣,٥٠	١٧,٥٠٠	٠,٠٠٣
٢٢	١٦,٣٦	١٨٠,٠٠	٦,٦٤	٧٣,٠٠	٧,٠٠٠	٠,٠٠١
٢٣	١٦,٦٨	١٨٣,٥٠	٦,٣٢	٦٩,٥٠	٣,٥٠٠٠	٠,٠٠١

ويتضح من جدول رقم (٣) أنه توجد فروق دالة إحصائية بين استجابات المجموعتين العليا والدنيا على كل مفردة من مفردات المقياس، وهذا يدل على قدرة المفردات على التمييز؛ مما يعد مؤشراً لصدق المقياس.

٥- تم التحقق من ثبات مقياس الاتجاه نحو بناء الاختبارات التحصيلية من خلال حساب معامل ألفا - كرونباك وكانت قيمته (٠,٨١٤) وهي قيمة تدل على تمتع القائمة بمعامل ثبات مقبول.

ثالثاً - البرنامج التدريبي على صياغة مفردات الاختيار من متعدد: تم بناء البرنامج التدريبي وفق الخطوات الآتية:

- ١ - تحديد قواعد صياغة الاختيار من متعدد، كما في الخطوات الأولى في بناء قائمة قواعد صياغة الاختيار من متعدد التي سبقت الإشارة إليها كأداة أولى للبحث.
- ٢ - تحويل قواعد صياغة الاختيار من متعدد إلى أهداف عامة ثم إلى أهداف إجرائية، كما يلي:
الهدف العام الأول - يراعي الطالب المعلم القواعد العامة لصياغة مفردات الاختيار من متعدد.
ويتفرع عنه الأهداف الإجرائية الآتية:
أ - يصيغ مفردات تتناول معلومة ضرورية، وليست شيئاً تافهاً.
ب - يصحح الصياغة اللغوية لبعض المفردات التي بها أخطاء.
ج - يعيد صياغة مفردات تحتوي على كلمات زائدة غير ضرورية.
د - يعيد صياغة مفردات منقولة نقلاً نصياً من الكتاب.
هـ - يذكر بعض المفردات التي تتوقف إجابة إحداها على إجابة مفردات أخرى.
و - يستخرج بعض المفردات التي تحتوي على كلمات غامضة.
الهدف العام الثاني - يراعي الطالب المعلم قواعد صياغة متن مفردة الاختيار من متعدد.
ويتفرع عنه الأهداف الإجرائية الآتية:
أ - يصحح صياغة مفردات يحتوي متنها على أدوات نفي.
ب - يحدد الكلمات الكاشفة عن الإجابة الصحيحة في مفردات مصاغة بطريقة خطأ.
ج - يعيد صياغة مفردات بها الفراغ في بداية المتن.
د - يحدد بعض المفردات التي تحتوي على عدة أفكار في المتن.
هـ - يذكر بعض المفردات التي بها كلمات محددة للإجابة، مثل: دائماً، أبداً، مطلقاً ... الخ.

الهدف العام الثالث - يراعي الطالب المعلم قواعد صياغة بدائل مفردة الاختيار من متعدد.

ويتفرع عنه الأهداف الإجرائية الآتية:

- أ - يعيد صياغة مفردات تحتوي بدائلها على كلمات مكررة.
- ب - يصحح بعض المفردات التي تحتوي على بديلين صحيحين.
- ج - يستخرج المشتت المتوقع ضعف فاعليته من بين بدائل بعض المفردات.
- د - يحول مفردات ذات بدائل ثلاثية إلى مفردات رباعية وخماسية البدائل.
- هـ - يعيد صياغة بدائل مفردات مختلفة في الطول.
- و - يتعرف البدائل غير المنطقية في بعض المفردات.
- ز - يتعرف الإجابة الصحيحة لبعض المفردات من خلال النمط اللغوي.
- ح - يذكر بعض المفردات التي تحتوي على بدائل غير مرتبة منطقياً.
- ط - يوزع البديل الصحيح في مواضع مختلفة (أ، ب، ج، د) بشكل عشوائي.
- ي - يعيد صياغة مفردات تحتوي على البديل "جميع ما سبق".
- ك - يستخرج المفردات التي تحتوي على بدائل متشابهة أو متضادة في معناها.

٣ - وضع الإطار العام للبرنامج التدريبي:

- تم وضع تصور مقترح للإطار العام للبرنامج يشمل الأهداف الثلاثة العامة والأهداف الخاصة الإجرائية المتفرعة من كل هدف من الأهداف العامة، وعناصر المحتوى التي يسعى من خلالها الباحث لتحقيق كل هدف من هذه الأهداف، والأنشطة التدريبية الملحقه بكل

محتوى، وأساليب التعليم والتعلم، وأدوات التقويم البنائي لقياس مدى تحقق كل هدف، والفترة الزمنية اللازمة لتحقيق كل هدف.

- تم عرض هذا التصور المقترح على (٥) من السادة المحكمين اثنين منهم من أعضاء هيئة التدريس بقسم علم النفس التربوي بكلية التربية بالمنيا، والثلاثة الآخرون من أعضاء هيئة التدريس بقسم المناهج وطرق التدريس بالكلية ذاتها.

- أشار بعض المحكمين لبعض التعديلات على هذا التصور المقترح، وتم مراعاتها، وبذلك تم الحصول على الإطار العام للبرنامج التدريبي.

٤ - تم بناء البرنامج التدريبي في ضوء الإطار العام السابق وضعه، وتضمن كتيب البرنامج (الموديول) الأهداف العامة والإجرائية المراد تحقيقها والمحتوى والأنشطة التدريبية ومفردات وأنشطة للتقويم البنائي لكل هدف من الأهداف، كما وضعت قائمة تقييم صياغة مفردات الاختيار من متعدد في بداية كتيب البرنامج وفي نهايته؛ لمساعدة المتدربين على التقييم الذاتي لقدراتهم، كما تم وضع مجموعة من المصادر العلمية التي يمكن الرجوع إليها لمزيد من القراءة والاطلاع في نهاية الكتيب.

٥ - تم تحكيم البرنامج التدريبي من قبل السادة المحكمين الذين قاموا بتحكيم الإطار العام للبرنامج، كما تم إجراء ما أشاروا إليه من تعديلات.

٦ - تم تقديم البرنامج التدريبي لعدد (٧) من طلاب علم النفس بكلية التربية بالمنيا في العام الجامعي ٢٠١٣ - ٢٠١٤م؛ للتأكد من وضوحه ومناسبته للطلاب، وبناءً على تساؤلات وملاحظات بعض الطلاب تم تزويد البرنامج بأنشطة إضافية، كما تم إجراء بعض التعديلات على الوقت المخصص لتحقيق بعض الأهداف.

عينة البحث

تكونت عينة البحث من (٣٦) طالباً من طلاب الفرقة الرابعة في تخصص علم النفس بكلية التربية - جامعة المنيا في العام الجامعي ٢٠١٤ - ٢٠١٥م،

وهم يمثلون جميع طلاب الفرقة، ما عدا الطلاب الذين تغيبوا في أي يوم من أيام التطبيق. وكانت عينة البحث مكونة من (طالبين، و ٣٤ طالبة)، وتم اختيار طلاب تخصص علم النفس خصيصاً لأنهم أكثر طلاب كلية التربية دراسة لمقررات في تخصص القياس والتقويم، كما تم اختيارهم في الفرقة الرابعة لأنهم على وشك التخرج والانخراط في العمل التدريسي.

إجراءات البحث

اعتمد المنهج شبه التجريبي وذلك وفق الخطوات الآتية:

- ١ - طُلب من عينة البحث وضع (٢٠) مفردة من نوع الاختيار من متعدد في أي مقرر من مقررات علم النفس التي يدرسونها أو التي يقومون بتدريسها في التربية العملية، ثم تم تحليل هذه المفردات التي بلغ عددها (٧٢٠) مفردة لكل الطلاب باستخدام قائمة تقييم صياغة مفردات الاختيار من متعدد.
- ٢ - تم تطبيق مقياس الاتجاه نحو بناء الاختبارات التحصيلية على عينة البحث.
- ٣ - تم تطبيق البرنامج التدريبي على عينة البحث في الفصل الدراسي الثاني من العام الجامعي ٢٠١٤ / ٢٠١٥م، واستغرق تطبيقه (٥) أسابيع بواقع ثلاثة ساعات أسبوعياً، وطُلب من عينة البحث مرة أخرى بعد انتهاء البرنامج وضع (٢٠) مفردة من مفردات الاختيار من متعدد مختلفة عن المفردات السابقة، مع تحليل هذه المفردات باستخدام القائمة المشار إليها سابقاً.
- ٤ - تم تطبيق مقياس الاتجاه نحو بناء الاختبارات التحصيلية بعد انتهاء البرنامج.
- ٥ - تمت معالجة البيانات التي تم الحصول عليها إحصائياً.

نتائج البحث وتفسيرها

الإجابة عن السؤال الأول:

ينص السؤال الأول: "ما مستوى قدرة طلاب علم النفس على صياغة مفردات الاختيار من متعدد قبل التدريب على صياغة مفردات الاختيار من متعددة؟"

- وللإجابة عن هذا السؤال تم حساب قدرة كل طالب من الطلاب على صياغة مفردات الاختيار من متعدد قبل التدريب، وذلك من خلال حساب متوسط عدد القواعد التي تم مراعاتها وفق قائمة تقييم صياغة مفردات الاختيار من متعدد، ويساوي مجموع عدد القواعد التي تمت مراعاتها مقسوماً على عدد المفردات التي صاغها، وبالتالي تتراوح قدرة الطالب في صياغة مفردات الاختيار من متعدد بين الدرجتين (صفر، ٢٢)، ثم تمت مقارنة متوسط قدرات الطلاب بقيمة اختبارية تعادل (٨٠٪) من القدرة الكلية (٢٢ درجة) التي تعادل ١٧,٦ درجة، التي تم تحديدها (كدرجة فاصلة معبرة عن الاتقان) من خلال متوسط آراء محكمي قائمة تقييم صياغة مفردات الاختيار من متعدد (بعد سؤالهم عن مستوى الاتقان اللازم لصياغة مفردات الاختيار من متعدد وفقاً للقائمة)، وذلك باستخدام اختبار (ت) لعينة واحدة فقط، وكانت النتائج كما في جدول رقم (٤).

جدول رقم (٤)

نتائج اختبار (ت) للمقارنة بين متوسط قدرات العينة على صياغة المفردات
والقيمة الاختبارية

العدد	متوسط الدرجات	الانحراف المعياري	المتوسط الاختباري	قيمة (ت)	درجة الحرية	مستوى الدلالة
٣٦	١١,٨٦١	٣,٤١٦	١٧,٦	١٠,٠٨١	٣٥	٠,٠٠١

ويتضح من جدول رقم (٤) وجود فرق دال إحصائياً بين القيمة الاختبارية التي تعادل (١٧,٦) درجة، وقيمة متوسط درجات الطلاب

المفحوصين (١١,٨٦١ درجة بنسبة ٥٣,٩١٤٪)، وذلك لصالح القيمة الاختبارية؛ مما يعني أنه يوجد فرق ذو دلالة إحصائية عند مستوى (٠,٠٠١) بين متوسط درجات الطلاب في قائمة تقييم صياغة مفردات الاختيار من متعدد (الذي يمثل المستوى الفعلي قبل التدريب)، والقيمة الاختبارية (التي تمثل مستوى الاتقان في الصياغة)، وهذا الفرق في اتجاه القيمة الاختبارية، مما يعني أن مستوى قدرة الطلاب على صياغة مفردات الاختيار من متعدد أقل من المستوى الذي يعبر عن مستوى الاتقان.

- ومما يؤكد ذلك أيضاً أن عدد الطلاب الذين وصلوا إلى محك الاتقان أو تجاوزه بلغ (٥) طلاب أي بنسبة (١٣,٨٨٩٪) وهي نسبة توحي بأن أغلب الطلاب لم يصلوا إلى مستوى الإتقان في صياغة هذه المفردات.

وتتفق هذه النتائج مع نتائج دراسات تارنت وهاميس ووير (Tarrant, Knierim, Hayes & Ware, 2006)، وتارنت ووير ومحمد (Tarrant, Ware & Mohammed, 2009)، والمهيدب (Al Muhaidib, 2010)، وظاظا (٢٠١٢)، ونيديو كايو ولافلين ورس وهال (Nedeau-Cayo et al., 2013) التي أسفرت جميعها عن انخفاض قدرات معدي مفردات الاختيار من متعدد على الصياغة الجيدة لهذه المفردات، وتعدد أخطاء الصياغة لديهم.

وقد تعود هذه النتيجة إلى قلة الوقت المخصص - في المقررات المتعلقة بمجال القياس النفسي والتقويم التربوي - لتدريس الأنواع المختلفة من المفردات وكيفية صياغتها وعيوبها ومزاياها، مع عدم اهتمام كثير من أعضاء هيئة التدريس بتدريب طلابهم على صياغة مفردات الاختيار من متعدد، وتكليفهم بممارسة صياغتها، وتقويم ما يقعون فيه من أخطاء وجوانب قصور، كما قد يعود السبب إلى عدم الاهتمام بقياس هذا الهدف التعليمي في أسئلة الاختبارات المتعلقة بهذه المقررات، فبمراجعة إختبارات (٥) أعوام سابقة للمقررات المتعلقة بالقياس النفسي والتقويم التربوي لم يجد الباحث سؤالا واحداً عن كيفية صياغة المفردات الاختبارية.

الإجابة عن السؤال الثاني:

ينص السؤال الثاني: "ما أكثر أخطاء صياغة مفردات الاختيار من متعدد شيوعاً لدى طلاب علم النفس قبل التدريب على صياغة مفردات الاختيار من متعدد؟"

وللإجابة عن هذا السؤال تم حساب عدد مفردات الاختيار من متعدد التي انتهكت كل قاعدة من قواعد الصياغة على حدة، وذلك لدى جميع أفراد العينة قبل التدريب (والتي سبقت الإشارة أنها ٧٢٠ مفردة)، وكذلك حساب نسبة هذه المفردات التي انتهكت كل قاعدة إلى العدد الكلي للمفردات، كما في جدول رقم (٥).

جدول رقم (٥)

ترتيب قواعد صياغة مفردات الاختيار من متعدد وفقاً لدرجة انتهاكها من قبل العينة (ن = ٧٢٠ مفردة)

م	القاعدة	عدد الانتهاكات	نسبة الانتهاكات %
١	تجنب وضع البديل (جميع / كل ما سبق) كبديل صحيح.	٥٨٨	٨١,٦٦٧
٢	تساوى بدائل الإجابة في الطول.	٥٧٩	٨٠,٤١٧
٣	تجنب وضع كلمات مكررة في بدائل الإجابة.	٥٦٧	٧٨,٧٥٠
٤	تجنب النقل النصي للمفردة من معلومة بالكتاب أو المحتوى.	٥٦١	٧٧,٩١٧
٥	إحتواء المفردة على بدائل منطقية (لا يعرف خطأها بديهاً).	٥٤٣	٧٥,٤١٧
٦	إحتواء المفردة على مشتتات جاذبة للطلاب غير المتمكنين.	٥٢٢	٧٢,٥٠٠
٧	إحتواء المفردة على كلمات ضرورية لا يمكن الاستغناء عنها.	٥٠٨	٧٠,٥٥٦
٨	صياغة المفردة صياغة لغوية صحيحة.	٤٩٣	٦٨,٤٧٢
٩	استقلال إجابة المفردة عن إجابات المفردات الأخرى ولا تعتمد عليها.	٤٧٧	٦٦,٢٥٠
١٠	إحتواء المفردة على أربعة بدائل.	٤٦٢	٦٤,١٦٧
١١	إحتواء المفردة على فكرة واحدة فقط، وليست عدة أفكار.	٤٣٢	٦٠,٠٠١

تابع/ جدول رقم (٥)
ترتيب قواعد صياغة مفردات الاختيار من متعدد وفقاً لدرجة انتهاكها من قبل
العينة (ن = ٧٢٠ مفردة)

م	القاعدة	عدد الانتهاكات	نسبة الانتهاكات %
١٢	تجنب وضع كلمات بمتن المفردة تكشف الإجابة الصحيحة.	٤١٨	٥٨,٠٥٦
١٣	وجود الفراغ في نهاية المتن وليس في بدايته أو منتصفه.	٤٠٢	٥٥,٨٣٣
١٤	تجنب وضع بدائل يمكن الاستدلال على صحتها أو خطأها من خلال النمط اللغوي.	٣٩٠	٥٤,١٦٧
١٥	تجنب وضع بدائل متشابهة أو متضادة في معناها في بدائل الإجابة.	٣٧٧	٥٢,٣٦١
١٦	تجنب وضع أدوات النفي في متن المفردة.	٣٦٢	٥٠,٢٧٨
١٧	ترتيب بدائل الإجابة ترتيباً منطقياً.	٣٤١	٤٧,٣٦١
١٨	قياس المفردة لمعلومة ضرورية، وليس شيئاً تافهاً.	٣٢٥	٤٥,١٣٩
١٩	تجنب وضع الكلمات المحددة للإجابة، مثل: دائماً، أبداً، فقط، مطلقاً ... الخ.	٣٠٤	٤٢,٢٢٢
٢٠	توزيع مكان البديل الصحيح بشكل عشوائي على مواضع البدائل (أ، ب، ج، د).	٢٨٢	٣٩,١٦٧
٢١	وجود إجابة صحيحة أو فضلى واحدة بما يتفق مع تعليمات المفردة.	٢٥٢	٣٥,٠٠٠
٢٢	تضمن المفردة كلمات واضحة في معناها.	٢١٩	٣٠,٤١٧

ويتضح من جدول رقم (٥) أن جميع قواعد الصياغة وصلت نسبة انتهاكها إلى أكثر من ٣٠٪ في المفردات المصاغة، وأن عدد (١٦) قاعدة كانت نسبة انتهاكها تزيد عن ٥٠٪ من عدد المفردات المصاغة، وهي نسب توحى بزيادة عدد انتهاكات القواعد، وعدم تمكن أفراد العينة من الالتزام بهذه القواعد عند صياغة هذه الأسئلة، وقد يكون ذلك ناتجاً عن نقص الخبرة وقلة التدريب وعدم الإلمام بالمقررات الدراسية التي وضعوا الأسئلة فيها، كما أن كثرة استخدام البديل (جميع ما سبق) قد يعود إلى رغبة الطلاب في التسهيل على أنفسهم، إذ كانت أغلب المفردات تحتوي على معلومات عامة وسطحية، وكذلك فإن اهتمام الطلاب

بالتأكيد على البديل الصحيح يزيد من التفاصيل التي يحتويها؛ مما يزيد من طوله عن باقي البدائل، كما أن عدم تمكن الطلاب من محتوى المقررات التي قاموا بوضع الأسئلة فيها أدى إلى ظهور بدائل غير منطقية (يمكن التعرف على خطئها بسهولة)، وقلة خبرتهم بصياغة الأسئلة أدت إلى تكرارهم كلمات في بدائل الإجابة بدلا من رفعها إلى متن السؤال. كما قد تكون كثرة هذه الأخطاء راجعة إلى الإهتمام بوضع البديل الصحيح في البدائل دون إعطاء أهمية كبيرة لصياغة المشتتات بطريقة فعالة، وقد يكون تعرض الطلاب إلى كثير من الاختبارات التي تهتم بقياس حفظ المعلومات دون الإرتقاء بقياس المستويات المعرفية الأعلى سببا وراء النقل النصي لكثير من المفردات من الكتاب المدرسي. كما قد يكون تعدد أخطاء الصياغة أيضاً نتيجة اكتساب بعض الطلاب الخطأ الذي يقع فيه بعض أعضاء هيئة التدريس عند صياغتهم لمفردات الاختيار من متعدد.

الإجابة عن السؤال الثالث:

ينص السؤال الثالث: "ما مستوى اتجاهات طلاب علم النفس نحو بناء الاختبارات التحصيلية قبل التدريب على صياغة مفردات الاختيار من متعدد؟"

وللإجابة عن هذا السؤال تم حساب متوسط اتجاهات الطلاب نحو بناء الاختبارات التحصيلية، ومقارنة هذا المتوسط بمتوسط افتراضي اختباري يعادل الدرجة الوسيطة للمقياس أو القيمة المحايدة التي تساوي (عدد مفردات المقياس $\times$ درجة البديل الأوسط = $23 \times 3 = 69$)، وذلك باستخدام اختبار (ت) للمقارنة بين متوسط درجات عينة بمتوسط اختباري، وكانت النتائج كما في جدول رقم (٦).

جدول رقم (٦)

نتائج اختبار (ت) للمقارنة بين متوسط اتجاهات الطلاب نحو بناء الاختبارات التحصيلية بقيمة اختبارية

العدد	متوسط الدرجات	الانحراف المعياري	المتوسط الاختباري	قيمة (ت)	درجة الحرية	مستوى الدلالة
٣٦	٧٤,٣٠٥	١٨,٣٠٩	٦٩	١,٧٣٩	٣٥	٠,٠٩١

ويتضح من جدول رقم (٦) أنه لا يوجد فرق دال إحصائياً بين المتوسط الفعلي لدرجات الطلاب في الاتجاه نحو بناء الاختبارات التحصيلية والمتوسط الاختباري الذي يمثل القيمة المحايدة، حيث بلغ متوسط درجاتهم (٧٤,٣٠٥) أي بنسبة ٦٤,٦١٣٪، وهذا يعني أن اتجاهات الطلاب نحو بناء الاختبارات التحصيلية كان محايداً، أي أن الطلاب لا يملكون اتجاهات إيجابية أو سلبية نحو بناء الاختبارات، وهذا قد يعود إلى نقص وعيهم بالدور المنوط بهم القيام به نحو تقويم أداء طلابهم مستقبلاً، مع قلة تدريبهم على بناء الاختبارات التحصيلية الجيدة في أثناء الإعداد التربوي في دراستهم السابقة، في الوقت الذي يرجون فيه القيام بهذا الدور الذي سيسند لهم مستقبلاً وما يشعرون به من تقدير لذويهم في حالة القيام به، ولعل التضارب بين ما يروجوه الطلاب من رغبة في ممارسة الدور المستقبلي والقصور الذي لحق بإعدادهم في قضية بناء الاختبارات التحصيلية هو ما قد يكون سبباً وراء ظهور اتجاه محايد لديهم نحو بناء الاختبارات التحصيلية.

الإجابة عن السؤال الرابع:

ينص السؤال الرابع: "ما الدلالة الإحصائية للعلاقة الارتباطية بين القدرة على صياغة مفردات الاختيار من متعدد واتجاهاتهم نحو بناء الاختبارات التحصيلية؟"

وللإجابة عن هذا السؤال تم استخدام معامل ارتباط بيرسون لدراسة العلاقة بين المتغيرين، وبلغت قيمته (٠,٤٨٣) وهذه القيمة دالة إحصائياً عند مستوى (٠,٠١)، وهذا يعني أنه توجد علاقة ارتباطية طردية بين المتغيرين، أي أنه كلما زادت القدرة على صياغة مفردات الاختيار من متعدد زاد الاتجاه نحو بناء الاختبارات التحصيلية، وبالتالي فإن حجم التأثير لهذه العلاقة من خلال مربع معامل الارتباط قد بلغ (٠,٢٣٣) تقريباً، وهذا يعني أن التباين في اتجاهات الطلاب نحو بناء الاختبارات التحصيلية يكون ٢٣,٣٪ تقريباً منه لتأثير قدرتهم على صياغة مفردات الاختيار من متعدد، وهذه النسبة تعد نسبة مقبولة إلى حد ما إذا ما قورنت بعدد العوامل التي يمكن أن تؤثر في الاتجاه نحو بناء الاختبارات التحصيلية، ومنها القدرة على بناء الاختبارات نفسها،

واعتقاد الطلاب بأهمية هذه الاختبارات وجدواها، والمكافآت المعنوية والمادية التي قد يحصل عليها بسبب بنائه للاختبارات، وشعوره المتوقع نتيجة ردود أفعال الطلاب والمعلمين وأولياء الأمور والإدارة نفسها تجاه هذه الاختبارات ... الخ، وكل هذه المتغيرات وغيرها مجتمعة تمثل النسبة الباقية من التباين التي تبلغ ٧٦,٧٪ تقريباً. وقد تعود هذه النتيجة إلى أن امتلاك الطلاب للقدرة على صياغة مفردات الاختيار من متعدد يزيد من ثقتهم بأنفسهم ويرفع درجة وعيهم بأهمية تحري الدقة في صياغة المفردات بشكل عام، وبأهمية بناء الاختبارات بدرجة دقة عالية، وهذا قد يزيد من رغبتهم في بناء الاختبارات التحصيلية بشكل عام، إذ أنهم قد يرون عملية بناء الاختبارات التحصيلية عملية تتمتع بفضائل عالية، وهم يمتلكون جزءاً منها.

الإجابة عن السؤال الخامس:

ينص السؤال الخامس: "ما الدلالة الإحصائية للفروق بين متوسطي درجات طلاب علم النفس في القدرة على صياغة مفردات الاختيار من متعدد قبل وبعد التدريب على صياغتها؟"

وللإجابة عن هذا السؤال تم استخدام اختبار "ت" للمقارنة بين متوسطي مجموعتين مرتبطتين، وذلك لمقارنة متوسط قدرات الطلاب في صياغة مفردات الاختيار من متعدد قبل التدريب ومتوسط قدراتهم بعد التدريب، وقد كانت النتائج كما في جدول رقم (٧).

جدول رقم (٧)

نتائج اختبار "ت" للمقارنة بين متوسطي قدرات الطلاب في صياغة مفردات الاختيار من متعدد قبل وبعد التدريب (ن = ٣٦)

القياس	المتوسط	الانحراف المعياري	قيمة "ت"	درجة الحرية	مستوى الدلالة	قيمة مربع ايتا	حجم التأثير
قبل التدريب	١١,٨٦١	٣,٤١٦	١١,٥٨٥	٣٥	٠,٠٠١	٠,٨٩١	كبير
بعد التدريب	١٨,٠٢٨	١,٧٨١					

ويتضح من جدول رقم (٧) أنه يوجد فرق ذو دلالة إحصائية بين متوسط قدرات الطلاب قبل التدريب على صياغة مفردات الاختيار من متعدد ومتوسط قدراتهم بعد التدريب، وهذا الفرق في اتجاه متوسط قدراتهم بعد التدريب، وهذا قد يعود إلى تأثير البرنامج التدريبي الذي تلقاه الطلاب في صياغة مفردات الاختيار من متعدد، إذ تم إطلاع الطلاب على كثير من المفردات رديئة الصياغة، وتم تدريبهم على كيفية إعادة صياغتها بطريقة أفضل يراعى من خلالها تطبيق قواعد الصياغة الجيدة، كما تم تكليفهم باستخراج المفردات جيدة الصياغة والمفردات رديئة الصياغة من بين مفردات مجموعة كبيرة من الاختبارات التي تعقد في كلية التربية للدبلوم التربوي، كما تم تقديم البرنامج كما لو كان جزءاً من مقرر القياس النفسي الذي يتم تدريسه للطلاب؛ مما زاد اهتمام الطلاب به وزاد من حرصهم على التعلم والإجادة في التكاليفات التي يقدمونها، وهذا قد يكون السبب وراء زيادة قدراتهم على صياغة هذه المفردات.

الإجابة عن السؤال السادس:

ينص السؤال السادس: "ما الدلالة الإحصائية للفروق بين متوسطي درجات طلاب علم النفس في اتجاهاتهم نحو بناء الاختبارات التحصيلية قبل وبعد التدريب على صياغة مفردات الاختيار من متعدد؟"

وللإجابة عن هذا السؤال تم استخدام اختبار "ت" للمقارنة بين متوسطي مجموعتين مرتبطتين، وذلك لمقارنة متوسط اتجاهات الطلاب نحو بناء الاختبارات التحصيلية قبل تطبيق البرنامج التدريبي ومتوسط اتجاهاتهم بعد هذا التطبيق، وكانت النتائج كما في جدول رقم (٨).

جدول رقم (٨)

نتائج اختبار "ت" للمقارنة بين متوسطي اتجاهات الطلاب نحو بناء الاختبارات التحصيلية قبل وبعد التدريب

القياس (ن = ٣٦)	المتوسط	الانحراف المعياري	قيمة "ت"	درجة الحرية	مستوى الدلالة	قيمة مربع ايتا	حجم التأثير
بعد التدريب	٧٩,٦٣٩	١٦,٩٩١	٤,٤٩٠	٣٥	٠,٠٠١	٠,٦٠٥	كبير
قبل التدريب	٧٤,٣٠٥	١٨,٣٠٩					

ويتضح من جدول رقم (٨) أنه يوجد فرق ذو دلالة إحصائية بين متوسط درجات الطلاب في الاتجاه نحو بناء الاختبارات التحصيلية بعد التدريب على صياغة مفردات الاختيار من متعدد ومتوسط درجاتهم قبل التدريب، وقد كان هذا الفرق في اتجاه متوسط درجاتهم بعد التدريب، وقد كان تأثير البرنامج التدريبي على اتجاهات الطلاب نحو بناء الاختبارات التحصيلية كبيراً، وهذا يعني أن البرنامج التدريبي أدى إلى نمو اتجاهات الطلاب نحو بناء الاختبارات التحصيلية، وقد يعود هذا النمو الحادث في الاتجاه نحو بناء الاختبارات التحصيلية إلى ما تلقاه الطلاب من تدريب على صياغة مفردات الاختيار من متعدد؛ إذ أن هذا التدريب قد أكسبهم معلومات عن قواعد صياغة المفردات ومهارات في صياغتها، وكذلك فإن قيامهم بنقد كثير من المفردات المصاغة في الاختبارات التحصيلية التي اعدّها بعض أعضاء هيئة التدريس بالكلية قد يكون سبباً في رفع تقديرهم لذويهم واستمتاعهم بصياغة المفردات، وبالتالي فإن البرنامج قد يكون له أثر على الجانب المعرفي والمهاري والوجداني لدى الطلاب فيما يتعلق بالعملية الاختبارية بشكل عام، مما قد يكون سبباً في تنمية اتجاهاتهم نحو عملية بناء الاختبارات التحصيلية.

توصيات البحث

يمكن تقديم التوصيات الآتية:

- ١ - ضرورة تخصيص وقت أكبر لتدريب الطلاب على مهارات صياغة المفردات بشكل عام، ومهارات صياغة مفردات الاختيار من متعدد بشكل خاص، وذلك أثناء دراستهم للمقررات ذات الصلة.
- ٢ - ضرورة اهتمام أعضاء هيئة التدريس بكلية التربية بتقييم أداء طلابهم في صياغتهم لمفردات الاختيار من متعدد، وأن يتم تمثيل ذلك في الاختبارات التي تقدم لهم.
- ٣ - أن يكون إعداد بنك أسئلة في مقرر ما جزء من مشروع التخرج الذي يقدمه طلاب تخصص علم النفس.

٤ - ضرورة الاهتمام بتنمية اتجاهات الطلاب نحو بناء الاختبارات التحصيلية.

البحوث المقترحة

١ - برنامج تدريبي لصياغة أنواع المفردات الإختبارية المختلفة وأثره في اتجاهات الطلاب المعلمين نحو مهنة التدريس.

٢ - العوامل المؤثرة في اتجاهات الطلاب المعلمين نحو بناء الاختبارات التحصيلية المختلفة.

٣ - المستويات المعرفية لمفردات الاختيار من متعدد التي يعدها الطلاب المعلمون وعلاقتها بالتحصيل الدراسي لديهم.

A Training Program in Writing Multiple Choice Items and its Effect on the Attitude Towards Achievement Test Construction

Dr. Mohammed M. Abdel-wahab

Faculty of Education - Minia University - A.R.E

Faculty of Education - Umm Al-Qura University - K.S.A

Abstract

Due to the results of previous studies which indicated that there are multiple errors made by constructors of achievement tests in writing multiple choice questions (MCQ), as well as the results indicated by the pilot study done on the fourth year students in psychology at the Faculty of Education in Minia University which indicated the multiplicity of these errors, this study was initiated. This research aims to prepare a training program for those students on the skills and rules of writing MCQ and see the impact of this program on attitude towards construction of achievement tests. For this purpose a list of rules of writing MCQ was built, and a scale of attitude towards constructing achievement tests in addition to training program. Quasi-experimental approach based on the design of a single group was used, the results indicated a low level in students' ability in construct MCQ before training, this level was 53.914%, as well as the level of attitude toward constructing achievement tests before training was neutral, It was 64.613%, and there was a positive correlation between students' abilities in writing MCQ and attitudes towards constructing achievement tests. The results also indicated that there was statistically significant difference between means of students' abilities in writing MCQ before and after training, in favor of abilities after training, in addition to a statistically significant difference between means of students' attitudes towards construction of achievement tests before and after training in direction of mean of attitudes after training, The researcher recommended that there should be more time to train students on rules of writing MCQ during their study in Faculty of Education, and there is need of work on developing attitudes towards constructing achievement tests.

المراجع

- ١ - أبو الريش، معاوية (٢٠٠٩). دراسة العلاقة بين الاتجاهات وبناء الامتحانات الصفية لمعلمي مرحلة التعليم الأساسي بوحدة كبوشيه محلية بولاية نهر النيل. رسالة ماجستير غير منشورة، جامعة أم درمان الإسلامية، كلية التربية.
- ٢ - أبو جراد، حمدي (٢٠١١). مدى إلتزام المعلمين بتحليل نتائج الاختبارات التحصيلية وعلاقته باتجاهاتهم نحوها. مجلة جامعة الازهر بغزة (سلسلة العلوم الإنسانية)، ١٣ (١٢)، ٨٩ - ١٠٦.
- ٣ - أبو حطب، فؤاد وصادق، آمال (٢٠١٠). علم النفس التربوي (ط ٧). القاهرة، مكتبة الأنجلو المصرية.
- ٤ - أبو فودة، باسل (٢٠١٤). أثر إعادة ترتيب البدائل في صعوبة فقرة الاختيار من متعدد. دراسات عربية في التربية وعلم النفس، (٥٣)، ٢٦٥-٢٨٧.
- ٥ - الأمير، بندر (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 غير منشورة، جامعة أم القرى، كلية التربية.
- ٦ - بركات، زياد وصباح، عبد الهادي (٢٠٠٧). مدى تحقيق أسئلة الامتحانات النهائية في جامعة القدس المفتوحة للأهداف التعليمية تبعاً لهرم بلوم. مجلة جامعة القدس المفتوحة للأبحاث والدراسات، ٩، ١٢٣ - ١٥٥.
- ٧ - بن لادن، سامية (٢٠٠٢). اتجاه طالبات كلية التربية نحو الاختبارات المقالية والموضوعية وعلاقة ذلك بتحصيلهن الدراسي. مجلة البحوث النفسية والتربوية - جامعة المنوفية، ١٧ (٢)، ١٥٥ - ١٩٦.
- ٨ - بني عطا، زايد والرباعي، إبراهيم (٢٠١٣). أثر عدد البدائل وتغيير موقع الموه القوي في فقرات اختبار الاختيار من متعدد على معالم الفقرات

- وقدرة الفرد ودالة المعلومات. المجلة الأردنية في العلوم التربوية، ٩ (٣)، ٣١٩ - ٣٣٣.
- ٩ - التويجري، نوال (٢٠١٥). دراسة تحليلية لأسئلة امتحانات اللغة العربية لشهادة إتمام الثانوية العامة في الكويت في ضوء مهارات التفكير والإبداع اللغوي. مجلة العلوم التربوية والنفسية، ١٦ (٤)، ٧٩-١١٣.
- ١٠ - الثبتي، سالم (٢٠١٤). بناء اختبار تشخيصي مرجعي المحك لقياس مهارات المعلمين في بناء الاختبارات التحصيلية لمقرر الحاسب الآلي. رسالة ماجستير غير منشورة، جامعة أم القرى، كلية التربية.
- ١١ - حماد، شريف (٢٠١١). جودة أسئلة الامتحانات النصفية لبرنامج التربية في جامعة القدس المفتوحة في ضوء معايير جودة الامتحان. ورقة علمية مقدمة إلى اليوم الدراسي "جودة الامتحانات الجامعية: الواقع والمأمول، في ٢٨ / ٤ / ٢٠١١، جامعة القدس المفتوحة، غزة.
- ١٢ - الحموري، هند (٢٠١٠). دراسة استكشافية لملاءمة نماذج نظرية الاستجابة للمفردة في بناء اختبار تحصيلي من إعداد المعلم. مجلة العلوم التربوية والنفسية، ١٢ (٢)، ٤٨ - ٨١.
- ١٣ - حيمودة، جمعة (٢٠١٠). كيفية بناء اختبار تحصيلي (اختبار مادة مدخل إلى علم النفس نموذجاً). مجلة الواحات للدراسات والبحوث، (٨)، ٢٩٥-٣٠٤.
- ١٤ - رضوان، رانيا (٢٠١١). دراسة تحليلية لاختبارات مادة الرياضيات المدرسية في ضوء مستوياتها المعرفية وخصائصها السيكمومترية بالتعليم الأساسي. مجلة العلوم التربوية، ٣، ٢١٩ - ٢٤٩.
- ١٥ - الرويلي، مشعل (٢٠٠٧). مدى فاعلية الإعداد التربوي في كليات المعلمين بالمملكة العربية السعودية في تزويد الطالب المعلم بمهارات قياس وتقويم التحصيل الطلابي. رسالة ماجستير غير منشورة، جامعة أم القرى، كلية التربية.

- ١٦ - الزبون، حابس (٢٠١٣). التلاغير في تقديرات معالم فقرات في الاختيار من متعدد عندما تطبق في سياقات مختلفة باستخدام نماذج النظرية الحديثة في القياس. *المجلة التربوية بالكويت*، ١٠٧ (٢٧)، ٣٠٩-٣٣٨.
- ١٧ - الزغلول، عماد وطلافة، فؤاد والمحاميد، شاكر (٢٠١٢). أثر استخدام استراتيجية التعريف بالأهداف السلوكية والتدريب على كيفية وضع المفردات في خفض قلق الاختبار لدى المتعلمين. *مجلة جامعة دمشق*، ٢٨ (١)، ٣٩٣ - ٤١٧.
- ١٨ - الزهراني، محمد (٢٠٠٠). واقع امتلاك معلمي الرياضيات بالمرحلة الابتدائية للكفايات المعرفية لإعداد الاختبارات التحصيلية. رسالة ماجستير غير منشورة، جامعة أم القرى، كلية التربية.
- ١٩ - سايحي، سليمة (٢٠١٢). قلق الامتحان وبعض العوامل المساعدة لظهوره لدى التلاميذ. *مجلة العلوم الإنسانية والاجتماعية*، ٧، ٧٤ - ٨٩.
- ٢٠ - سنان، إيناس (٢٠٠٦). درجة معرفة وممارسة معلمات الرياضيات في المرحلة المتوسطة لمهارات بناء الاختبارات التحصيلية بمدينة مكة المكرمة. رسالة ماجستير غير منشورة، كلية التربية، جامعة أم القرى.
- ٢١ - الشريفين، نضال والصبح، رانيا (٢٠١١). أثر بنية فقرات الاختيار من متعدد ومستوى القدرة لدى الأفراد على دقة التقديرات لمعالم الفقرات والأفراد وفق نظرية الاستجابة للفقرة. *مجلة جامعة أم القرى للعلوم التربوية والنفسية*، ٣ (٢)، ٤٦ - ١١٠.
- ٢٢ - الشياب، سوسن (٢٠٠٣). مدى كفاءة معلمات الإقتصاد المنزلي في بناء الاختبارات التحصيلية في مدارس وزارة التربية والتعليم بالأردن وفق معايير الاختبار الجيد. رسالة ماجستير غير منشورة، جامعة عمان العربية للدراسات العليا، كلية الدراسات التربوية العليا.
- ٢٣ - الصرايرة، آيات (٢٠١١). دراسة تحليلية لأنماط مفردات الاختبارات التحصيلية لدى معلمي التربية الاجتماعية والوطنية للصفوف الرابع

- والخامس الأساسيين في لواء المزار الجنوبي. رسالة ماجستير غير منشورة، جامعة مؤتة، عمادة الدراسات العليا.
- ٢٤ - ظاظا، حيدر (٢٠١٢). الكشف عن مدى انتهاك قواعد صياغة فقرة الاختيار من متعدد في مفردات امتحان شهادة الدراسة الثانوية العامة في الأردن. *المجلة الأردنية في العلوم التربوية*، ٨ (١)، ٨١ - ٩١.
- ٢٥ - عباس، مها (٢٠٠٨). مستويات التقويم والتصانيف التي تضمنتها في مفردات امتحان مادة التاريخ. *مجلة مركز البحوث التربوية والنفسية*، (١٦)، ٢٨٧ - ٣٢٨.
- ٢٦ - عبد الحافظ، ليلي (١٩٨٦). مقياس قلق الامتحان (الاتجاه نحو الامتحان). *المجلة التربوية لكلية التربية بسوهاج*، ١، ٤٠ - ٨٢.
- ٢٧ - عبد المنعم، محمد (٢٠٠٩). تقويم أسئلة الاختبارات التحصيلية بكلية المعلمين جامعة الملك فيصل في ضوء معايير جودة التقويم. *مجلة البحوث النفسية والتربوية*، (١)، ٢٠٧ - ٢٣٦.
- ٢٨ - عبد الوهاب، محمد (٢٠١٣). المقارنة بين النظرية الكلاسيكية والنموذج ثلاثي البارامتر في معالجة أثر التخمين عند تقدير إحصائيات الأفراد والمفردات في اختبار اختيار من متعدد، *المجلة التربوية بالكويت*، ١٢٩ (٢٨)، ١٦٩-٢١١.
- ٢٩ - عبيد، صديق (٢٠٠٨). فاعلية الدورات التدريبية أثناء الخدمة في تطوير كفايات معلمي الرياضيات في بناء الاختبارات التحصيلية للمرحلة الأساسية في الأردن. رسالة ماجستير غير منشورة - كلية الدراسات التربوية والنفسية العليا بجامعة عمان العربية للدراسات العليا.
- ٣٠ - عسيري، علي (٢٠٠٢). واقع الاختبارات التحصيلية في التعليم العام والجامعي وأثره على المخرجات التعليمية. بحث مقدم في اللقاء السنوي العاشر للجمعية السعودية للعلوم التربوية والنفسية. (القياس والتقويم التربوي والنفسي). جامعة الملك سعود: الرياض، ١٦ - ١٨ أبريل.

- ٣١ - علام، صلاح الدين (٢٠٠٠). القياس والتقويم التربوي والنفسي: أساسياته وتطبيقاته وتوجهاته المعاصرة. القاهرة، دار الفكر العربي.
- ٣٢ - عودة، أحمد (١٩٨٩). اتجاهات المعلمين نحو الامتحانات المدرسة والعوامل المؤثرة فيها. أبحاث اليرموك، ٥(٤)، ١٣٣-١٤٩.
- ٣٣ - عودة، أحمد (١٩٩٠). تطوير مقياس اتجاهات المعلمين نحو الامتحانات المدرسية: نموذج مقترح. مجلة كلية التربية، جامعة الإمارات، ٥(٥)، ٥١-٧٢.
- ٣٤ - غليون، أزهار (٢٠٠٩). تقويم الاختبارات النهائية للدراسات العليا بكلية التربية صنعاء في ضوء معايير الاختبار الجيد. المجلة العلمية لكلية التربية، ١(٨)، ١٧٦ - ٢٠٠.
- ٣٥ - كامل، مصطفى (١٩٩٩). استخدام استراتيجية التعلم حتى يتمكن في تدريس مقرر في التقويم التربوي لإكساب الطلاب المعلمين مهارات بناء الاختبارات التحصيلية تجربة ميدانية. مجلة علم النفس، ٥١، ٦-٣١.
- ٣٦ - الكلحوت، أحمد (٢٠٠٢). مقارنة بين الخصائص السيكومترية لكل من اختبارات الاختيار من متعدد واختبارات التكميل. مجلة مركز البحوث التربوية بجامعة قطر، ٢٢، ١٢٧ - ١٥٣.
- ٣٧ - المحمد، فيصل (٢٠٠٩). فاعلية برنامج تدريبي في تحسين مهارات بناء الاختبارات التحصيلية لدى معلمي العلوم الاجتماعية في دولة الكويت. رسالة ماجستير غير منشورة، كلية الدراسات التربوية والنفسية العليا، جامعة عمان العربية للدراسات العليا.
- ٣٨ - محمود، فريال (٢٠١٥). أثر كل من ترتيب فقرات الاختبار ومستوى القلق على أداء الطلاب في اختبار من نوع الاختيار من متعدد. مجلة العلوم التربوية والنفسية، ١٦(٢)، ٥٨٧ - ٦٠٥.
- ٣٩ - مختار، ضياء الدين (٢٠٠٩). تحليل وتقويم الاختبارات التحصيلية لطلاب اللغتين الفرنسية والإنجليزية ومدى إلمام أساتذتهما بمفاهيم وأسس

- إعدادها بالجامعات السودانية. مجلة العلوم والثقافة، ١٠ (٥)، ١٤٢ - ١٨٣.
- ٤٠ - مراد، صلاح وسليمان، أمين (٢٠١٢). الاختبارات والمقاييس في العلوم النفسية والتربوية (خطوات إعدادها وخصائصها). القاهرة، دار الكتاب الحديث.
- ٤١ - ملياني، زياد (٢٠٠٣). درجة معرفة معلمي الرياضيات في مدينة مكة المكرمة بقواعد بناء الاختبارات التحصيلية. رسالة ماجستير غير منشورة، كلية التربية، جامعة أم القرى.
- ٤٢ - الهواري، ماهر والشناوي، محمد (١٩٨٧). مقياس الاتجاه نحو الاختبارات (قلق الاختبارات): معايير ودراسات ارتباطية. مجلة رسالة الخليج العربي، ٢٢، ١٧١ - ١٩٦.
- ٤٣ - يعقوب، إبراهيم؛ أبو فودة، باسل (٢٠١٠). أثر مخالفة قواعد صياغة فقرات الاختيار من متعدد على التقديرات المختلفة لنظرية استجابة الفقرة. مجلة كلية التربية بجامعة الإسكندرية، ٢٠ (٢)، ٥٣ - ٨٩.
- 44 - Ali, S.H., & Ruit, K.G. (2015). The Impact of item flaws, testing at low cognitive level, and low distractor functioning on multiple-choice question quality. **Perspectives on Medical Education**, 4(5), 244-251.
- 45 - Al Muhaidib, N. S. (2010). Types of Item-Writing Flaws in Multiple Choice Question Pattern - A Comparative Study. **Umm Al-Qura University Journal of Educational & Psychological Science**, 2(2), 9 - 44.
- 46 - Amo-Salas, M., Arroyo-Jimenez, M. d. M., Bustos-Escribano, D., Fairn-Jimnez, E., & Lopez-Fidalgo, J. (2014). New Indices for Refining Multiple Choice Questions. **Journal of Probability and Statistics**, 2014, 1-8. doi: 10.1155/2014/240263
- 47 - Baghaei, P. & Dourakhshan, A. (2016). Properties of Single Response and Double Response Multiple Choice Grammar Items. **International Journal of Language Testing**, 6(1), 33-48.
- 48 - Chavda, P., Misra, S., & Duttaroy, B. (2015). Item Analysis of Multiple

- Choice Questions based Undergraduate Assessment in Community Medicine. **South East Asian Journal of Medical Education**, 9(1), 66-68.
- 49 - Cheung, D. & Bucat, R. (2002). How can we construct good Multiple-choice items?. Paper presented at the Science and Technology Education Conference, Hong Kong, June 20-21, 2002.
- 50 - Collins, J. (2006). Education techniques for lifelong learning: writing multiple-choice questions for continuing medical education activities and self-assessment modules. **Radiographics**. 26(2), 543-551.
- 51 - DePalma, A. (1990, November 1). Revisions adopted in college entrance tests. The New York Times. Retrieved from <http://www.nytimes.com/1990/11/01/us/revisions-adopted-in-college-entrance-tests.html?page-wanted=all>
- 52 - Downing, S.M. (2005). The effects of violating standard item-writing principles on tests and students: the consequences of using flawed test items on achievement examinations in medical education. **Advances in Health Sciences Education: Theory and Practice**. 133, 10-43.
- 53 - Green, K. & Stager, S. (1986). Measuring attitudes of teachers toward testing. **Measurement and Evaluation in Counseling and Development**, 19(3), 141-150.
- 54 - Gleason, J., Alley, A., & Baker, S. (2010). The effect of item-writing rules on the reliability of an instrument to measure mathematical knowledge for teaching. **Journal of Mathematical Sciences and Mathematics Education**, 5(2), 21-27
- 55 - Haladyna, T. M., Downing, S. M., & Rodriguez, M. C. (2002). A Review of Multiple-Choice Item-Writing Guidelines for Classroom Assessment. **Applied Measurement in Education**, 15(3), 309-333. doi: 10.1207/S15324818AME1503_5
- 56 - Hall, C. & Marshall. S. (2013). **A guide for developing multiple choice and other objective style questions**. Centre for Academic Development, Victoria University of Wellington, New Zealand.
- 57 - Hambleton, R., Swaminathan, H. & Rogers, H. (1991). **Fundamentals of item response theory**, International Educational and Professional. Publisher Newbury Park.

- 58 - Lin, S.-Y., Singh, C., Rebello, N. S., Engelhardt, P. V., & Singh, C. (2012). Can multiple-choice questions simulate free-response questions?. **AIP Conference Proceedings**, 47-50. doi: 10.1063/1.3679990
- 59 - Mahmood, H. (2015). Correlation of MCQ and SEQ Scores in Written Undergraduate Ophthalmology Assessment. **Journal of the College of Physicians and Surgeons Pakistan**, 25(3), 185-188
- 60 - Meng Er, H., Ramamurthy, S., & Pook, P. C. (2014). Can learning outcomes in cognitive domain be assessed effectively using multiple choice questions? A study in an undergraduate pharmacy curriculum. **International E-Journal Of Science, Medicine & Education**, 8(3), 9-18.
- 61 - Nedeau-Cayo, R., Laughlin, D., Rus, L., & Hall, J. (2013). Assessment of Item-Writing Flaws in Multiple-Choice Questions. **Journal for Nurses in Professional Development**, 29 (2), 52-57.
- 62 - Rocha, E. M., Marche, T. A., & Briere, J. L. (2013). The effect of forced-choice questions on children's suggestibility: A comparison of multiple-choice and yes/no questions Canadian. **Journal of Behavioural Science**, 45 (1), 1-11. doi: 10.1037/a0028507
- 63 - Tarrant, M., Knierim, A., Hayes, S. K., & Ware, J. (2006). The frequency of item writing flaws in multiple-choice questions used in high stakes nursing assessments. **Nurse Education Today**, 26(8), 662-671. doi: 10.1016/j.nedt.2006.07.006/c
- 64 - Tarrant, M., Ware, J., & Mohammed, A. M. (2009). An assessment of functioning and non-functioning distractors in multiple-choice questions: a descriptive analysis. **BMC Med Educ**, 9, 40. doi: 10.1186/1472-6920-9-40.
- 65 - Thawabieh, A. M., (2016). A Comparison between two test item formats Multiple-Choice Items and Completion Items. **British Journal of Education**, 4(8), 63-74.