أثر كل من ترتيب فقرات الاختبار ومستوى القلق على أداء الطلاب في اختبار من نوع الاختيار من متعدد

د. فريال محمود الحاج محمود
قسم التربية والآداب
الكلية الجامعية بالقنفذة – جامعة أم القرى
Ferial_alhaj@hotmail.com

أثر كل من ترتيب فقرات الاختبار ومستوى القلق على أداء الطلاب في اختبار من نوع الاختيار من متعدد

د. فريال محمود الحاج محمود
قسم التربية والآداب
الكلية الجامعية بالقنفذة – جامعة أم القرى

الملخص

هدفت هذه الدراسة لفحص أثر كل من ترتيب فقرات الاختبار، ومستوى القلق لدى الطلاب، على أدائهم في اختبار من نوع الاختيار من متعدد. تم تطبيق: مقياس قلق الامتحان لتشارلز سبيلبيرجر، وستة نماذج مختلفة من اختبار نهاية الفصل الدراسي في مقرر مدخل إلى علم النفس، على عينة مكونة من (٤٥٤) طالبة من طالبات جامعة أم القرى. لم تظهر نتائج الدراسة فروقاً ذات دلالة إحصائية (α=٠,٠٥) في أداء الطالبات على الاختبار تعزى لتفاعل نموذج الاختبار مع مستوى القلق، بينما أظهرت فروقا دالة إحصائيا (α=٠,٠٥) في أداء الطالبات تعزى لنموذج الاختبار، فقد كان أفضل أداء للطالبات على نموذجي الاختبار:الرابع و الخامس، حيث الفقرات مرتبة وفق تسلسل المحتوى، مما يؤكد أن هذا الترتيب هو الأنسب للطلاب بصرف النظر عن مستوى القلق لديهم.

الكلمات المفتاحية: ترتيب فقرات الاختبار، مستوى القلق، اختبار اختيار من متعدد، مقياس قلق الامتحان.

* تاريخ تسلم البحث: ٢٠١٣/١٠/٨م
* تاريخ قبوله للنشر: ٢٠١٤/٢/٦م

The Effect of Test Items Order and the Level of Anxiety, on Students Performance in a Multiple Choice Test

Dr. Ferial M. Alhaj-Moh'd
University College of Qunfudah
Umm Al-Qura University

Abstract

The purpose of this study was to examine the effect of test items order, and the level of anxiety on students performance in a multiple choice test. Charles D. Spielberger test anxiety, and six different forms of Introduction to Psychology final test were applied to a sample of (454) female students from Umm Al-Qura University. The results showed no statistically significant differences (œ=0.05) in students performance on the test, due to the interaction between test form, and the level of anxiety. On the other hand there were statistically significant differences (œ=0.05) in students performance in favor of test forms where the items were arranged in ascending, or descending order, according to the content sequence, which confirms that this arrangement of multiple choice test items is the appropriate arrangement for students regardless of their test anxiety level.

Keywords: test items order, anxiety level, multiple choice test, test anxiety level.

أثر كل من ترتيب فقرات الاختبار ومستوى القلق على أداء الطلاب في اختبار من نوع الاختيار من متعدد

د. فريال محمود الحاج محمود
قسم التربية والآداب
الكلية الجامعية بالقنفذة – جامعة أم القرى

المقدمة:

على الرغم من تطور أدوات التقويم فلازالت الاختبارات التحصيلية من أهم هذه الأدوات، وهي تستخدم في قياس تحصيل الطلبة، بهدف الوقوف على مدى تحقق الأهداف التعليمية، و لمقارنة أداء الطلاب بعضهم ببعض (عودة، ٢٠٠٥). وللوقوف على المستوى الفعلي للطلاب يسعى المعلم لمنع الغش في قاعات الاختبار، من خلال إعداده لعدة نماذج من الاختبار، غالبا ما يكون الفرق بينها هو ترتيب الفقرات، مما يجعل من الصعب على الطلاب نقل الإجابات عن زملائهم. وتمتاز هذه الطريقة في تصميم نماذج الاختبار: بمحافظتها على درجة صعوبة الاختبار، فمن المعروف إن صعوبة الاختبار تتحدد بصعوبة الفقرات المكونة له، وليس بترتيبها (Sue, 2009). ولكن وعلى الرغم من تعرض الطلاب لنفس الفقرات، فهل يكون المعلم في محاولته لمنع الغش – وعلى غير قصد منه- قد وقع في مشكلة التحيز، بأن تسبب في حصول بعض الطلاب الذين أجابوا عن نموذج ما، على درجات إضافية لم يحصل عليها أقرانهم الذين أجابوا عن نماذج أخرى من نفس الاختبار؟ (Blach, 1989).

بناءً على ماسبق: ما هي أفضل طريقة لترتيب فقرات الاختبار؟ في الحقيقة لا يوجد اتفاق بين الباحثين على طريقة محددة، فبينما يصر البعض ومنهم بلاتش (Blach, 1989) على أن الطلاب يحصلون على درجات أعلى، إن كانت فقرات الاختبار مرتبة بنفس تسلسل شرح المحتوى (المادة العلمية) خلال عملية التدريس، مما سيوفر-تبعا لوجهة نظرهم- إشارات تساعد الذاكرة على استرجاع المعلومات، بطريقة تتوافق مع طريقة ترميزها، وهذا يجعل هؤلاء الطلاب يستفيدون أكثر من أقرانهم الذين عرضت عليهم الفقرات بترتيب آخر (Doerner & Calhoun, 2009). بالمقابل هناك وجهة نظر أخرى يتبناها مجموعة من الباحثين ومنهم كرونباخ (Cronbach, 1950) وهم يؤكدون على ضرورة ترتيب فقرات الاختبار تصاعديا تبعا لدرجة صعوبتها، ويعللون ذلك بأن: الطالب عندما يجيب عن الفقرات السهلة في بداية الاختبار فإن معدل القلق لديه سينخفض، وستزداد دافعيته، مما يجعله على درجة من الثقة

بالنفس تساعده على الإجابة عن الفقرات التالية، وهي الأكثر صعوبة. أما إن ظهرت الفقرات الصعبة في بداية الاختبار فهذا -تبعا لوجهة نظر كرونباخ ومؤيديه- سيرفع من درجة القلق لدى الطالب، وتجعله أكثر ارتباكا، مما يؤثر في إجاباته عن الفقرات التالية. لذلك فهم يقترحون ترتيب فقرات الاختبار من السهل إلى الصعب (Gerow, 1980).

لقد تطرقت العديد من الدراسات لموضوع ترتيب فقرات الاختبار، ومن أقدمها دراسة جيرو (Gerow, 1980) الذي طبق نموذجين من اختبار مكون من ٦٠ فقرة (٥٠ فقرة متوسطة الصعوبة، و٥ فقرات صعبة جدا، و٥ فقرات سهلة جدا)، على مجموعة من طلاب السنة الجامعية الأولى. وضع الباحث الفقرات متوسطة الصعوبة بنفس الترتيب في نموذجي الاختبار، ووزع باقي الفقرات بحيث تظهر الفقرات السهلة قبل الصعبة في النموذج الأول، والصعبة قبل السهلة في النموذج الثاني. وعلى الرغم من أن نتائج الدراسة أظهرت: عدم وجود فروق دالة إحصائيا في أداء الطلاب على الاختبار تعزى لنموذج الإجابة. فقد كانت نسبة الطلاب الذين أجابوا إجابة صحيحة عن الفقرات متوسطة الصعوبة، في النموذج الأول، أعلى من نسبتهم في النموذج الثاني.

وفي نفس الفترة ظهرت دراسة دامبروت (Dambrot, 1980) حيث طبق خمسة اختبارات، يتكون كل منها من نموذجين، على (٨٤٧) طالبا من طلبة السنة الأولى المسجلين في مقرر علم النفس، في جامعة اكرون. رتب الباحث الفقرات في النموذج الأول تصاعديا حسب درجة الصعوبة، ورتبها عشوائيا في النموذج الثاني. وقد أظهرت نتائج الدراسة: عدم وجود فروق ذات دلالة إحصائية في تحصيل الطلاب، على أي من هذه الاختبارات، يمكن عزوها لترتيب الفقرات في الاختبار.

وتوصل لافيتي (Laffittee, 1984) في دراسته حول أثر ترتيب فقرات الاختبار على درجات الطلاب، وعلى إدراكهم لصعوبة الفقرة، إلى نتيجة مماثلة لما توصل إليه كل من جيرو ودامبورت، حيث صمم أربعة نماذج من اختبار تحصيلي مكون من (٥٠) فقرة من نوع الاختيار من متعدد، ورتب الفقرات تبعا لتسلسل الوحدات الدراسية، ومن السهل إلى الصعب في النموذج الأول، وعشوائيا في النموذج الثالث، وأعاد ترتيب الفقرات وبصرف النظر عن الوحدة الدراسية، ومن السهل إلى الصعب في النموذج الثاني، وعشوائيا في النموذج الرابع. طبق الباحث النماذج الأربعة على عينة عشوائية، مكونة من (٨٢) طالبا من المسجلين في مقرر علم النفس، وأظهرت النتائج عدم وجود فروق دالة إحصائيا في درجات الطلاب تعزى لطريقة ترتيب الفقرات،أو تقديرهم لصعوبة الاختبار.

وعمل كل من نيومان وكودرت (Newman & Kudert, 1988) على تطبيق أربعة نماذج من اختبار مكون من ٤٠ فقرة (معروفة الصعوبة مسبقا)، على عينة عشوائية من(١٢٠) طالباً من المسجلين في مقرر علم النفس التربوي. ورتب الباحث الفقرات تبعا لمستويات هرم بلوم، وصنفها بتسلسل من السهل إلى الصعب ضمن كل مستوى في النموذج الأول، ومن الصعب إلى السهل أيضا ضمن كل مستوى في النموذج الثاني. وأعاد ترتيب الفقرات تنازليا تبعا لمستويات هرم بلوم وبتسلسل من السهل إلى الصعب ضمن المستوى الواحد في النموذج الثالث، ومن الصعب إلى السهل في كل مستوى في النموذج الرابع. أظهرت النتائج عدم وجود فروق دالة إحصائياً في تحصيل الطلاب على الاختبار، تعزى لترتيب الفقرات تبعا لمستوى الهدف المعرفي، أو لدرجة الصعوبة. وفي النهاية أوصى الباحثان: بإجراء مزيد من الدراسات حول أثر ترتيب فقرات الاختبار على الدرجات الفرعية للاختبار، وعدم الاكتفاء بدراسة أثره على الدرجة الكلية فقط .

ومن أهم الدراسات التي تطرقت لترتيب الفقرات دراسة بلاتش (Blach, 1989)، حيث طبق ثلاثة نماذج من اختبار مكون من (٧٥) فقرة، من نوع الاختيار من متعدد، على عينة من (٤٠٤) طلاب، ورتب الباحث الفقرات تبعا للوحدات الدراسية وتسلسليا ضمن كل وحدة في النموذج الأول، وعشوائيا ضمن الوحدة في النموذج الثاني. أما في النموذج الثالث فرتبت الفقرات عشوائيا وبصرف النظر عن الوحدة الدراسية. أظهرت النتائج فروقا ذات دلالة إحصائية (α= ٠,٠٥) في درجات الطلاب تعزى لنموذج الاختبار وذلك لصالح النموذج الأول، مما يدل على أن الترتيب التسلسلي للفقرات وفق الوحدات الدراسية هو الأفضل.

وفي دراسة لفحص أثر ترتيب فقرات اختبار من نوع الاختيار من متعدد على أداء الطلاب، طبق كل من كارلسن واوستروسكي (Carlson & Ostrosky, 1992) أربعة اختبارات يتكون كل منها من نموذجين، على الطلبة المسجلين في مقرر مبادئ الإدارة، بحيث رتبت الفقرات في النموذج الأول تسلسليا تبعا للمحتوى بينما رتبت عشوائيا في النموذج الثاني. أظهرت النتائج فروقا ذات دلالة إحصائية في أداء الطلاب لصالح النموذج الأول، وذلك في ثلاثة من الاختبارات الأربعة التي تم تطبيقها. وبناء على ذلك أوصى الباحثان اعتماد الترتيب التسلسلي وفق المحتوى عند ترتيب فقرات الاختبار، وأن يتم استخراج النماذج بإعادة ترتيب بدائل الإجابة.

ولفحص إن كان قلق الاختبار يتفاعل مع ترتيب فقرات الاختبار، طبق نيلي وزملاؤه (Neely, Springston, & McCann, 1994) ثلاثة اختبارات يتكون كل منها من نموذجين،

على عينة عشوائية مكونة من (٢٥٣) طالباً جامعياً، بحيث رتبت الفقرات من السهل إلى الصعب في النموذج الأول من كل اختبار، وعشوائيا في النموذج الثاني. أظهرت نتائج الدراستين الأولى والثانية عدم وجود فروق في تحصيل الطلاب تعزى لترتيب الفقرات، بينما أظهرت نتائج الدراسة الثالثة أثراً ذا دلالة إحصائية للتفاعل بين ترتيب الفقرات ومستوى القلق، فقد كان أداء الطلاب ذو القلق المرتفع أعلى على النموذج الأول، مما يؤكد أن ترتيب الفقرات من السهل إلى الصعب، هو الملائم لذوي القلق المرتفع.

وفي دراسة حول أثر ترتيب فقرات الاختبار على كل من أداء الطلاب، وإدراكهم لدرجة صعوبة الاختبار، ومستوى القلق لديهم. طبق كل من بيتيجون وساكو (Pettijohn & Sacco, 2007) ثلاثة نماذج من نفس الاختبار على مجموعة من الطلاب. ورتبت الفقرات في النماذج الثلاثة وفق تسلسل المحتوى (تصاعديا، تنازليا، عشوائيا) على الترتيب. إضافة لذلك أجاب كل طالب بعد إنهائه للاختبار على مقياس لدرجة القلق. كما حسب الباحثان الزمن الذي استغرقه كل طالب في الإجابة. أظهرت نتائج التحليل الإحصائي: عدم وجود فروق دالة إحصائيا في درجات الطلاب تعزى لنموذج الاختبار، أو لزمن الإجابة. ولكن ظهرت فروق في درجة إدراك الطلاب لصعوبة الاختبار، حيث أشار الطلاب إلى أن أصعب نموذج هو النموذج ذو الترتيب العشوائي، بينما كان النموذجان الآخران بنفس درجة الصعوبة.

ظهرت بعد ذلك دراسات ناقضت ما ورد في هذه الدراسة، ومنها دراسة سو (sue, 2009)، والتي فحصت أثر ترتيب فقرات الاختبار على أداء الطلاب، في مقررات مبادئ الإدارة، ذات عدد الملتحقين الصغير (أقل من ٣٠ طالب). وأعدت الباحثة نموذجين من الاختبار، رتبت الفقرات في أولهما تبعا للمحتوى، بينما رتبتها عشوائياً في النموذج الثاني. حاولت الباحثة دعم رأيها بأن عدد الملتحقين الصغير في المقرر سيجعل المناقشات خلال الشرح بين المعلم والطلاب قليلة، مما يتوقع أن يقلل من قدرة الطلاب على استرجاع المعلومات، عندما تعرض فقرات الاختبار مرتبة تبعا للمحتوى. وفعلا فقد أظهر تحليل مربع كاي: عدم وجود فروق دالة إحصائيا في أداء الطلاب تعزى لترتيب الفقرات، مما يدل حسب رأيها لإمكانية إعداد المعلم عدة نماذج من الاختبار بإعادة ترتيب الفقرات، دون أن يخشى من وقوعه في التحيز.

في دراستهما حول أثر ترتيب الفقرات على تحصيل الطلاب، قام دورنر وكالهون (Doerner & Calhoun, 2009) بتطبيق: ثلاثة نماذج من اختبار نهائي يتكون من (٧٥) فقرة، من نوع الاختيار من متعدد، على (١٢٩٦) طالباً من المسجلين في مقرر الاقتصاد. ورتب الباحثان الفقرات في النماذج الثلاثة تبعا للمحتوى (تصاعديا، تنازليا، عشوائيا) على الترتيب، كما

عملا على ضبط متغيرات مثل: الجنس، و نسبة حضور المحاضرات، والتخصص، ودرجات الطلاب على اختبار القبول. وقد أظهرت النتائج فروقاً ذات دلالة إحصائية في درجات الطلاب على الاختبار، تعزى لترتيب الفقرات. فقد كانت أعلى الدرجات على النموذج الأول، ثم الثاني، ثم الثالث. مما يشير لاستفادة الطلاب الذين كان نموذج اختبارهم هو النموذج الأول أكثر من أقرانهم الذين أجابوا عن النموذجين الثاني والثالث.

بعد الاطلاع على الدراسات السابقة لوحظ ندرة الدراسات العربية التي اهتمت بربط قلق الامتحان بطريقة ترتيب فقرات الاختبار، لذلك تأتي هذه الدراسة لسد هذه الثغرة للبحث في علاقة قلق الامتحان مع طريقة ترتيب الفقرات.

مشكلة الدراسة:

على الرغم من وجود العديد من الدراسات التي تناولت موضوع طريقة ترتيب فقرات الاختبار من نوع الاختيار من متعدد فإنه يلاحظ تعارض نتائجها، كذلك يلاحظ ندرة الدراسات العربية التي تطرقت لهذا الموضوع رغم أهميته، ففي الدول العربية وبسبب العدد الكبير للطلاب في القاعات الدراسية، فإن عمل عدة نماذج من الاختبار، هو ما يلجأ إليه المعلمون. لذلك جاءت هذه الدراسة للتحقق من توفير النماذج المختلفة من الاختبار للعدالة بين الطلاب، وللكشف عن تفاعل نموذج الإجابة مع مستوى القلق لدى الطالب.

أسئلة الدراسة:

تحاول هذه الدراسة الإجابة عن السؤال الرئيس التالي:

هل يتأثر أداء الطالب على الاختبار بطريقة ترتيب فقرات الاختبار وبمستوى القلق لديه؟

وينبثق عن هذا السؤال الأسئلة الفرعية التالية:

١- هل توجد فروق ذات دلالة إحصائية، عند مستوى دلالة (α = ٠,٠٥)، في أداء الطلاب على اختبار من نوع الاختيار من متعدد تعزى لمستوى القلق لديهم ؟

٢- هل توجد فروق ذات دلالة إحصائية، عند مستوى دلالة (α = ٠,٠٥)، في أداء الطلاب على اختبار من نوع الاختيار من متعدد تعزى لطريقة ترتيب فقرات الاختبار ؟

٣- هل توجد فروق ذات دلالة إحصائية، عند مستوى دلالة (α = ٠,٠٥)، في أداء الطلاب على اختبار من نوع الاختيار من متعدد تعزى لتفاعل كل من ترتيب فقرات الاختبار ودرجة القلق ؟

أهمية الدراسة:

تكتسب هذه الدراسة أهميتها من كونها: توفر بعض الأدلة التي توضح طريقة ترتيب الفقرات، التي تناسب الطلاب من ذوي مستويات القلق المختلفة. كما يمكن أن توفر نتائجها معلومات كافية لمستخدمي فقرات الاختيار من متعدد، حول أفضل الطرق لترتيب فقرات هذا النوع من الاختبارات. كذلك تقدم هذه الدراسة دعما لاستخدام اختبارات الاختيار من متعدد، كبديل لبعض أنواع الاختبارات عندما يتوقع المدرس أن لدى طلابه درجة ما من القلق.

أهداف الدراسة:

تلقي هذه الدراسة الضوء على الطريقة المثلى في ترتيب فقرات الاختبار، بحيث تحول هذه الطريقة دون حصول بعض الطلاب على درجات أعلى من زملائهم من ذوي نفس القدرة، وذلك بسبب إجابتهم عن نموذج ما في الاختبار. وهي بذلك تسعى:

- لتحديد أنسب الطرق لترتيب فقرات الاختبار.
- دراسة أثر كل من ترتيب الفقرات ومستوى القلق وانعكاسهما على أدائهم في الاختبار.

محددات الدراسة:

تتحدد نتائج هذه الدراسة بالمحددات الآتية:

- اقتصرت عينة الدراسة على عينة من طالبات الكلية الجامعية بالقنفذة – جامعة أم القرى، المسجلات في مقرر مدخل إلى علم النفس في الفصل الأول من العام الدراسي (١٤٣٣-١٤٣٤ هـ)، الموافق (٢٠١٢ – ٢٠١٣ م).
- اقتصرت الدراسة على المحتوى الدراسي الخاص بالاختبار النهائي في مقرر مدخل إلى علم النفس.
- اقتصرت الدراسة على ستة نماذج مختلفة لترتيب فقرات الاختبار تم شرحها في أداة الدراسة.

مصطلحات الدراسة:

القلق: يعرف القلق إجرائيا في هذه الدراسة بالدرجة الكلية التي تحصل عليها الطالبة على مقياس قلق الامتحان لتشارلز سبيلبيرجر (أعده وقننه للعربية د. محمد عبد الظاهر الطيب).

ترتيب الفقرات تبعا للمحتوى: يقصد به ترتيب الفقرات بشكل تسلسلي، تبعاً لتسلسل ظهور محتوى الفقرة خلال شرح المادة الدراسية.

ترتيب الفقرات تبعا لدرجة الصعوبة: يعرف معامل صعوبة الفقرة بأنه: نسبة الطلاب الذين أجابوا إجابة صحيحة عن الفقرة، وبذلك فإن قيمته تتراوح بين الصفر والواحد، وكلما كانت هذه القيمة أكبر كلما كان السؤال أسهل. وهكذا يمكن ترتيب الفقرات من السهل إلى الصعب، أو بالعكس (Nunnally, 1970).

سؤال اختيار من متعدد: يتألف سؤال الاختيار من متعدد من جزأين هما: المشكلة أو الجذر (Stem)، وقائمة من البدائل (Alternatives). والتي تضم إجابة صحيحة واحدة أو أفضل إجابة، بالإضافة إلى عدد من البدائل الخاطئة (المشتتات Distracters) (عودة، ٢٠٠٥).

متغيرات الدراسة

أولا: المتغيرات المستقلة

تتضمن هذه الدراسة متغيرين مستقلين هما:

١) **قلق الامتحان:** وله ثلاثة مستويات مرتفع ومتوسط ومنخفض، وهناك شرح مفصل في أداة الدراسة.

٢) **طريقة ترتيب الفقرات:** ولها ستة مستويات، حيث أعدت الباحثة ستة نماذج من الاختبار، وهناك شرح مفصل في أداة الدراسة، لطريقة ترتيب الفقرات في كل نموذج.

ثانياً: المتغير المضبوط

لضبط المستوى التحصيلي السابق للطالبة فقد تم ضبط متغير درجة الطالبة على اختبار منتصف الفصل الدراسي (اختبار أعمال السنة) في مقرر مدخل إلى علم النفس.

ثالثا: المتغير التابع

وهو درجة الطالبة على اختبار نهاية الفصل في مقرر مدخل إلى علم النفس.

مجتمع وعينة الدراسة:

يتكون مجتمع الدراسة من جميع الطالبات المسجلات في أقسام الكلية الجامعية بالقنفذة- جامعة أم القرى و التي تمنح الدرجة الجامعية الأولى مع إعداد تربوي. أما عينة الدراسة فتضم (٤٥٤) طالبة هن جميع الطالبات المسجلات في مقرر مدخل إلى علم النفس، خلال الفصل الدراسي الأول من العام الجامعي (١٤٣٣-١٤٣٤هـ)، والموافق(٢٠١٢-٢٠١٣م). واللواتي تقدمن للاختبار النهائي في ذلك المقرر في موعده المحدد من قبل الجامعة.

أدوات الدراسة:

استخدمت في هذه الدراسة أداتان رئيستان هما:

١) **اختبار نهاية الفصل الدراسي في مقرر مدخل إلى علم النفس:** قامت الباحثة بإعداد ستة نماذج من اختبار مكون من (٦٠) فقرة، من نوع الاختيار من متعدد (هذا الاختبار تم إعداده في فصل دراسي سابق). ورتبت الفقرات في النماذج الأول والثاني والثالث (من السهل إلى الصعب، من الصعب إلى السهل، عشوائيا)، مع مراعاة عدم ظهور الفقرات المتعلقة بنفس المحتوى الدراسي بشكل متتالٍ في أي من هذه النماذج الثلاثة. وفي النماذج الرابع والخامس والسادس: رتبت الفقرات تبعاً لتسلسل تدريس المحتوى العلمي الخاص بها (تسلسلي، عكسي، عشوائي). وصحح الاختبار بإعطاء درجة واحدة لكل إجابة صحيحة تقدمها الطالبة، وبالتالي تراوحت الدرجات بين (٠-٦٠).

٢) **مقياس قلق الامتحان:** استخدم في الدراسة مقياس قلق الامتحان لتشارلز سبيلبيرجر (أعده وقننه للعربية د. محمد عبد الظاهر الطيب (الملحق ١))، وهو من مقاييس التقرير الذاتي، وقد تم إعداده ليقيس الفروق الفردية في قلق الامتحان كسمة للشخصية في موقف محدد، وهو يتكون من ٢٠ عبارة، يتم الإجابة عن كل منها باختيار أحد البدائل (أبدا – أحياناً – غالباً - دائماً)، والتي يعبر المفحوصون من خلالها عن كيفية ومدى تكرار معاناتهم لأعراض القلق في مواقف الامتحان (سبيلبيرجر، ١٩٩٧). وتبعا لتعليمات تصحيح هذا المقياس، بإعطاء الإجابة بأبدا: درجة واحدة، وأحيانا: درجتان، وغالبا: ثلاث درجات، ودائما: أربع درجات. وهكذا تراوحت الدرجات على هذا المقياس بين (٢٠-٨٠). ووفقا للتعليمات يصنف من يحصل على درجة (٢٠-٣٩) على أنه مرتاح جدا، أي من ذوي قلق الامتحان المنخفض. يصنف من يحصل على درجة (٤٠-٥٩)، على أنه من ذوي قلق الامتحان المتوسط. بينما يصنف من يحصل على درجة (٦٠-٨٠)، بأنه من ذوي قلق الامتحان المرتفع.

خطوات إعداد اختبار نهاية الفصل الدراسي في مقرر مدخل إلى علم النفس:

إن هذا الاختبار التحصيلي تم إعداده في فصل دراسي سابق وفق خطوات إعداد الاختبار التحصيلي:

١) تحديد هدف الاختبار وهو قياس تحصيل الطالبات في مقرر «مدخل إلى علم النفس».

٢) استعانت الباحثة بزميلاتها اللواتي يدرسن نفس المقرر، و بالمراجع وبخبرتها في تدريس نفس المقرر في تحديد المفردات التي سيتم قياسها، ومن ثم صياغة الأهداف السلوكية، وإعداد جدول المواصفات.

٣) صياغة (١٠٠) فقرة اختبارية من نوع الاختيار من متعدد، وبعد عرض الاختبار بصورته الأولية على مجموعة من المحكمين، من تخصص القياس و التقويم، وممن يدرسون المقرر تم حذف (١٢) فقرة أجمع (٨٠٪) من المحكمين على أنها مكررة. كما تم تعديل بعض الفقرات بناء على قرار المحكمين.

٤) لإجراء التحليل العاملي (Factor Analyses) بطريقة تدوير المحاور (Varimax)، وللتأكد من أن الاختبار يقيس سمة أُحادية البعد. تم تطبيقه بفقراته (٨٨) فقرة في الفصل الأول (٢٠١١-٢٠١٢م)، على عينة استطلاعية مكونة من (١٠٠) طالبة، وبناء على نتائج التحليل العاملي تم حذف (٥) فقرات.

٥) بناء على قيم كل من معامل الصعوبة، ومعامل التمييز (معامل الارتباط بوينت بايسيريال بين الدرجة على الفقرة و الدرجة الكلية على الاختبار). تم حذف الفقرات التي وقعت قيمة معامل صعوبتها خارج المدى (٠,٣ – ٠,٨)، كما حذفت الفقرات التي يقل معامل تمييزها عن (٠,٣٩) (عودة، ٢٠٠٥). وبالتالي حذفت (٢٣) فقرة من فقرات الاختبار. وهكذا تكون الاختبار في صورته النهائية من (٦٠) فقرة. كذلك لم تشكو الطالبات من الزمن المخصص للإجابة، أو من أي غموض في صياغة الفقرات.

صدق وثبات أدوات الدراسة :

للتأكد من صدق كل من مقياس قلق الامتحان، و اختبار مقرر مدخل إلى علم النفس، تم عرض كل منهما على عدد من المختصين في علم النفس التربوي، و القياس والتقويم التربوي، وذلك في عدد من الجامعات السعودية، وبناء على آرائهم تم إجراء بعض التعديلات البسيطة على كل منهما قبل القيام بعملية التطبيق.كذلك تم حساب الاتساق الداخلي لفقرات الاختبار بحساب معامل الارتباط بين الدرجة على الفقرة و الدرجة الكلية على الاختبار، وكانت قيم جميع معاملات الارتباط دالة إحصائيا.

للتأكد من ثبات مقياس قلق الامتحان تم تطبيقه على (٦٣) طالبة من خارج عينة الدراسة، وكانت قيمة معامل ثبات كرونباخ ألفا (٠,٨٣)، وهي قيمة دالة إحصائيا (α=٠,٠٥). كذلك قامت الباحثة بحساب ثبات المقياس بناء على استجابات نفس مجموعة الطالبات ولكن بطريقة التجزئة النصفية، حيث حسبت الارتباط بين الفقرات الفردية والفقرات الزوجية وبلغت قيمة معامل الارتباط بعد تصحيح الطول وفق معادلة سبيرمان-براون (٠,٨١) وهي قيمة تشير لاتساق داخلي مرتفع للمقياس. أما فيما يتعلق بثبات الاختبار فقد طبق في فصل سابق على عينة من (٥٠) طالبة، وأعيد تطبيقه على نفس الطالبات بعد أسبوعين من التطبيق

الأول، وحسب معامل الثبات للاختبار-إعادة الاختبار (Test- Retest Reliability) وكانت قيمة معامل الارتباط بين التطبيقين (٠,٨٧) وهي قيمة دالة إحصائيا (α=٠,٠٥).

إجراءات الدراسة:

بعد التأكد من الخصائص السيكومترية لكل من مقياس قلق الامتحان والاختبار التحصيلي:

- طبقت الباحثة مقياس قلق الامتحان على عينة الدراسة.
- بعد انتهاء الطالبات من الإجابة عن فقرات مقياس قلق الامتحان، طبقت عليهن الباحثة اختبار نهاية الفصل الدراسي، وذلك تحت إشرافها، وبمساعدة عدد من أعضاء هيئة التدريس.
- قبل البدء بعملية التحليل تأكدت الباحثة من خلو الإجابات من عامل التخمين، وذلك من خلال فحص صعوبة الفقرات وتحديد الفقرات الصعبة، حيث تم اختيار الفقرات التي قيمة معامل صعوبتها أقل من (٠,٦) وعددها (١٥) فقرة، وتم فحص أداء الطالبات ذوات أقل (٢٧٪) درجة كلية على الاختبار وعددهن (١٢٢) طالبة على هذه الفقرات الصعبة، ووفق ماذكره هامبيلتون وسوامينثان (Hambleton & Swaminathan, 1985) فإن كانت نسبة إجابتهن بشكل صحيح على هذه الفقرات تختلف عن (٠,٢٥)، وهي نسبة التخمين العشوائي على الفقرة ذات البدائل الأربعة كان ذلك دليلا على تحقق شرط عدم وجود التخمين. وقد تحققت الباحثة من أن نسبة إجابة الطالبات ذوات التحصيل المتدني اللواتي أجبن إجابة صحيحة، عن الفقرات الصعبة في الاختبار هي نسبة منخفضة، مقارنة بنسبة الإجابة العشوائية عن هذه الفقرات وهي (٠,٢٥)، لذلك يمكن القول بخلو الإجابات من عامل التخمين.

الأساليب الإحصائية:

استخدمت في هذه الدراسة الأساليب الإحصائية التالية:

- الوسط الحسابي، الانحراف المعياري، معاملات الارتباط ومعامل ثبات كرونباخ ألفا.
- تحليل التباين الثنائي المصاحب (Two Way ANCOVA).
- مقياس حجم الأثر (مربع ايتا η^2) والذي يحسب وفق المعادلة $\eta^2 = \frac{SS_{effect}}{SS_{effect} + SS_{error}}$ حيث SSeffect: مجموع مربعات المعالجة، وSSerror: مجموع مربعات الخطأ (cohen, 1988).

نتائج الدراسة:

أولاً: النتائج المتعلقة بالسؤال الأول

نص هذا السؤال على:» هل توجد فروق ذات دلالة إحصائية، عند مستوى دلالة (α = ٠,٠٥)، في أداء الطلاب على اختبار من نوع الاختيار من متعدد تعزى لمستوى القلق لديهم؟"

بلغ عدد أفراد عينة الدراسة الذين أجابوا عن الاختبار النهائي، في مقرر مدخل إلى علم النفس (٤٥٤) طالبة، وكان المتوسط الحسابي لدرجاتهن (٤٢,٠١)، بانحراف معياري (٩,٠٧)، وقد تراوحت الدرجات بين (٢٥-٥٩). بينما تراوحت درجاتهن على مقياس قلق الامتحان بين (٢٥-٧٠).

الجدول رقم (١)
تحليل التباين الثنائي المصاحب (Two Way ANCOVA) لأثر مستوى القلق و ترتيب الفقرات على التحصيل مع ضبط أثر اختبار منتصف الفصل

مصدر التباين	مجموع المربعات	درجة الحرية	متوسط المربعات	قيمة ف المحسوبة	الدلالة الإحصائية	الدلالة العملية (مربع إيتا η٢)
اختبار منتصف الفصل (المتغير المضبوط)	٢٧٨١,٨٨	١	٢٧٨١,٨٨	٤٩,٥١	٠,٠٠٠	٠,١٠
مستوى القلق	٣٠٠,٠٧	٢	١٥٠,٠٣	٢,٦٧	٠,٠٧٠	٠,٠١
ترتيب الفقرات	٩٣٢٩,٩٦	٥	١٨٦٥,٩٩	٣٣,٢١	٠,٠٠٠	٠,٢٨
مستوى القلق × ترتيب الفقرات	٣٣٥,٨٩	١٠	٣٣,٥٩	٠,٦٠	٠,٨١٦	٠,٠١
الخطأ	٢٤٤٤١,٥	٤٣٥	٥٦,١٩			
الكلي	٣٧١٨٩,٣	٤٥٣				

يظهر الجدول رقم (١) تحليل التباين الثنائي المصاحب (Two Way ANCOVA) لأثر كل من مستوى القلق وترتيب الفقرات على التحصيل، مع ضبط أثر اختبار منتصف الفصل. حيث يظهر بوضوح عدم وجود فروق ذات دلالة إحصائية عند مستوى دلالة (α = ٠,٠٥) في أداء الطالبات تعزى لمستوى القلق، وكانت قيمة الدلالة العملية (مربع إيتا) لهذا الأثر = ٠,٠١، أي أن ١٪ من التباين في أداء الطالبات يمكن عزوه لدرجة القلق لديهن. وهي قيمة صغيرة وفق معايير كوهن (Cohen, 1988).

ثانيا: النتائج المتعلقة بالسؤال الثاني

نص هذا السؤال على:» هل توجد فروق ذات دلالة إحصائية، عند مستوى دلالة (α = ٠,٠٥)، في أداء الطلاب على اختبار من نوع الاختيار من متعدد تعزى لطريقة ترتيب فقرات الاختبار؟"

يظهر الجدول رقم (١) فروقا ذات دلالة إحصائية عند مستوى دلالة (α=٠,٠٥)، في أداء الطلاب تعزى لترتيب فقرات الاختبار، حيث يلاحظ أن قيمة الدلالة العملية كانت ٠,٢٨، أي أن ٢٨٪ من التباين في درجات الطلاب يمكن عزوه لترتيب الفقرات وهذه القيمة كبيرة وفق معايير كوهن (Cohen, 1988).

الجدول رقم (٢)
المقارنات البعدية للفروق في الأوساط الحسابية (المعدلة) لدرجات الطلاب تبعا لترتيب الفقرات بعد ضبط أثر درجات اختبار منتصف الفصل (اختبار أعمال السنة)

نموذج الاختبار	المتوسط الحسابي	١	٢	٣	٤	٥	٦
١	٣٩,٩٥	–	١,٠٩	١,٦٧	*٩,١٥-	*٧,٩٥-	١,٩٤
٢	٣٨,٨٧	١,٠٩-	–	٠,٥٨	*١٠,٢٤-	*٩,٠٣-	٠,٨٥
٣	٣٨,٢٩	١,٦٧-	٠,٥٨-	–	*١٠,٨٢-	*٩,٦١-	٠,٢٧
٤	٤٩,١١	*٩,١٥	*١٠,٢٤	*١٠,٨٢	–	١,٢١	*١١,٠٩
٥	٤٧,٩٠	*٧,٩٥	*٩,٠٣	*٩,٦١	١,٢١-	–	*٩,٨٨
٦	٣٨,٠٢	١,٩٤-	٠,٨٥-	٠,٢٧-	*١١,٠٩-	*٩,٨٨-	–

* دال عند مستوى دلالة (α = ٠,٠٥)

ولمعرفة لصالح أي المجموعات كانت الفروق استخدمت المقارنات البعدية. ويظهر الجدول رقم (٢) المقارنات البعدية (Post Hoc Comparisons) حيث يتضح أن الفروق كانت لصالح المجموعتين الرابعة (الفقرات مرتبة تصاعديا وفق المحتوى)، والخامسة (الفقرات مرتبة تنازليا وفق المحتوى).

ثالثاً: النتائج المتعلقة بالسؤال الثالث

وينص على: «هل توجد فروق ذات دلالة إحصائية، عند مستوى دلالة (α = ٠,٠٥)، في أداء الطلاب على اختبار من نوع الاختيار من متعدد تعزى لتفاعل كل من ترتيب فقرات الاختبار ودرجة القلق؟"

يظهر الجدول رقم (١) عدم وجود فروق دالة إحصائيا عند مستوى دلالة (α = ٠,٠٥)، في أداء الطلاب تعزى لتفاعل مستوى القلق وترتيب الفقرات. كما يظهر الجدول رقم (٣) توزيع أفراد عينة الدراسة والمتوسط الحسابي للدرجات تبعا لمتغيري نموذج الاختبار (ترتيب الفقرات)، ومستوى القلق لديهم.

الجدول رقم (٣)
المتوسط الحسابي لدرجات مجموعات الدراسة تبعا لمستويات متغير القلق

المجموعة (نموذج الاختبار)	قلق منخفض			قلق متوسط			قلق مرتفع			العينة ككل		
	العدد	المتوسط الحسابي	الانحراف المعياري	العدد	المتوسط الحسابي	الانحراف المعياري	العدد	المتوسط الحسابي	الانحراف المعياري	العدد	المتوسط الحسابي	الانحراف المعياري
١	١٨	٤١,٢	١١,٨	٣٦	٤٠,٨	٨,٧	٢١	٣٨,٢	٨,٧	٧٥	٣٩,٩٥	٩,٤٦
٢	٢٢	٤١	١١,٢	٣٦	٣٨,٦	٨,٦	١٨	٣٥,٤	٣,٧	٧٦	٣٨,٨٧	٨,٨٠
٣	٢٤	٤١,٣	٨,٥	٣٢	٣٧,٧	٩,٦	٢٠	٣٦,٣	٧,١	٧٦	٣٨,٢٩	٨,٨٠
٤	٢١	٥٠	٥,٣	٣٦	٤٨,٧	٤,٧	٢٠	٤٨,٩	٥,٦	٧٧	٤٩,١١	٥,٠٥
٥	٢٢	٤٨,٥	٥,٦	٣١	٤٧	٧,١	٢١	٤٦,٨	٦,١	٧٤	٤٧,٩٠	٦,٣٨
٦	٢٥	٣٧,٢	٨,٣	٣٣	٣٨,٩	٩,١	١٨	٣٩,١	٦,١	٧٦	٣٨,٠٢	٨,١٢
العينة ككل	١٣٢	٤٣,١	٩,٧	٢٠٤	٤١,٩	٩,١	١١٨	٤٠,٩	٨,٢	٤٥٤	٤٢,٠١	٩,٠٧

مناقشة النتائج:

على الرغم من أن نتائج السؤال الأول أظهرت عدم وجود فروق دالة إحصائيا في أداء الطالبات تعزى لمستوى القلق لديهن، فإن الجدول رقم (٣) يظهر أن متوسط درجات ذوي القلق المنخفض على الاختبار، كانت أعلى من متوسط درجات ذوي القلق المرتفع. وهذه النتيجة وإن لم تكن دالة إحصائيا فهي تتفق مع نتائج العديد من الدراسات، والتي ربطت بين القلق والتحصيل ومنها دراسة: سوترسو (Sutarso, 1992)، وويلسون (Wilson, 1997)، وهمبري (Hembree, 1988)، والتي أكدت جميعها على أن التحصيل يتناسب عكسيا مع زيادة مستوى القلق.

تتفق نتيجة السؤال الثاني مع ما ذكره كل من: جيرو (Gerow, 1980)، ودامبروت (Dambrot, 1980)، ولافيتي (Laffittee, 1984)، من عدم وجود فروق في أداء الطلاب تعزى لترتيب الفقرات من السهل إلى الصعب، أو من الصعب إلى السهل.كما تتفق مع دراسة بلاتش (Blach, 1989)، وكارلسن واوستروسكي (Carlson & Ostrosky, 1992)، ودورنر وكالهون (Doerner & Calhoun, 2009) والتي أظهرت أن الأداء كان أفضل على نموذج الاختبار الذي يعرض الفقرات بشكل تسلسلي تبعا للوحدات الدراسية. كما تتفق هذه النتيجة جزئيا مع ماذكره نيلي وآخرون (Neely et al, 1994)، من حصول أصحاب القلق العالي على نتائج أفضل إن كانت الفقرات مرتبة من السهل إلى الصعب، فعلى الرغم من عدم دلالة الفروق بين المجموعات عندما كانت الفقرات مرتبة تبعا لمعامل الصعوبة، ولكن يلاحظ من

الجدول رقم (٣) أن متوسط أداء مجموعة القلق المرتفع على النموذج الأول كان أعلى من متوسط أداء نفس المجموعة على أي من النموذجين الثاني و الثالث.

وفيما يتعلق بدراسة كل من بيتجون وساكو (Pettijohn & Sacco, 2007) وإن لم تتفق نتائج هذه الدراسة معها من حيث وجود فروق دالة إحصائيا في أداء الطلاب، عندما رتبت الفقرات تبعا للمحتوى، فإنه يوجد من جانب آخر اتفاق بين الدراستين. فقد ذكر الباحثان أن الطلاب أدركوا أن أصعب نماذج الاختبار هو النموذج ذو الفقرات المرتبة عشوائيا، تلاه النموذجان ذوا الترتيب التصاعدي والتنازلي تبعا للمحتوى. وفي هذه الدراسة لم تظهر فروق دالة إحصائيا بين النموذجين الرابع و الخامس، وهذا يوضح جانب الاتفاق بين الدراستين. كما لا تتفق النتائج مع دراسة سو (sue, 2009) والتي أظهرت عدم وجود فروق ذات دلالة إحصائية في أداء الطلاب عندما عرضت الفقرات مرتبة وفق المحتوى، أو عرضت بشكل عشوائي. ولعل صغر حجم العينة في دراسة سو كان هو السبب وراء عدم ظهور مثل هذه الفروق.

التوصيات:

توصي هذه الدراسة بترتيب فقرات الاختبار من نوع الاختيار من متعدد تصاعديا، أو تنازليا وفق تسلسل عرض المحتوى العلمي. كذلك توصي هذه الدراسة بدراسة أثر ترتيب فقرات من نوع الصواب والخطأ، أو الاختبارات المقاليه على أداء الطلاب في هذه الاختبارات. وفي النهاية توصي الدراسة بدراسة أثر ترتيب فقرات الاختبار على الزمن الذي يقضيه الطالب في الإجابة عن ذلك الاختبار.

المراجع:

عودة، أحمد (٢٠٠٥). القياس والتقويم في العملية التدريسية. الأردن: الأمل للطباعة والنشر والتوزيع.

سبيلبيرجر، تشارلز (١٩٩٧). اختبار قلق الامتحان. (ترجمة: محمد عبد الظاهر الطيب)، القاهرة: مكتبة الأنجلو المصرية.

Balch, W. (1989). Item order affect performance on multiple-choice exams. *Teaching of Psychology, 16,* 75- 77.

Carlson, J., & Ostrosky A. (1992). Item sequence and student performance on multiple-choice exams: Further evidence, *Journal of Economic Education, 23,* 232-235.

Cohen, J. (1988). *Statistical power analysis for the behavior sciences* (2nd ed.). New Jersey: Lawrence Erlbaum Associates.

Cronbach, L. (1950). Further evidence on response sets and test design, *Educational and Psychological Measurement, 10,* 3-31.

Dambrot, F. (1980). Test item order and academic ability, or should you shuffle the test item deck?. *Teaching of Psychology, 7*(2), 95-96.

Doerner, W., & Joseph P. (2009). *The impact of the order of test questions in introductory economics*. *(working paper)*, Retrieved from, SSRN: http://ssrn.com/abstract=1321906 .

Gerow, J. (1980). Performance on achievement tests as a function of the order of item difficulty, *Teaching of Psychology, 7*(2), 93-94.

Gohmann, S., & Spector, L.(1989). Test scrambling and student performance. *Journal of Economic Education, 20*(3), 235-238.

Hambleton, R. & Swaminathan, H. (1985). *Item response theory principles and applications* (1st ed.). Boston: Kluwer, Nijhoff Publishing.

Hembree, R. (1988). Correlates, causes, effects and treatment of test anxiety. *Review of Educational Research, 58*(1), 47–77.

Laffittee, R. (1984). Effects of item order on achievement test scores and students' perceptions of test difficulty. *Teaching of Psychology, 11*(4), 212-214.

Neely, D., Springston, R., & McCann, S. (1994). Does item order affect performance on multiple-choice exams? *Teaching of Psychology, 21*(1), 44-45.

Newman, D. & Kudert, D.(1988). Effect of varying item order on multiple-choice test scores: Importance of statistical and cognitive difficulty, *Applied Measurement in Education, 1*(1), 89-97.

Nunnally, J. (1970). *Introduction to psychological measurement*. Tokyo: Kogakusha, McGraw-Hill.

Pettijohn, T. , & Sacco, M. (2007). Multiple-choice exam question order influences on student performance, Completion time, and Perceptions, *Journal of Instructional Psychology, 34*(3), 142-149.

Spielberger, C. (1977). *Test attitude inventory*. Palo Alto, CA: Consulting Psychology Press.

Sue, D. (2009). The effect of scrambling test questions on student performance in a small class setting. *Journal for Economic Educators, 9*(1), 32-41.

Sutarso, T. (1992). *Students' attitudes toward statistics (STATS)*. Paper presented at the annual meeting of the Mid-South Educational Research Association, November, Knoxville, TN.

Wilson, V. (1997). *Factors related to anxiety in the graduate statistics classroom*. Paper presented at the meeting of the Mid-south Educational Research Association, November, Memphis, TN.