جامعـة المنصــورة
كليـــة التربية
–

أثر البديل "جميع ما سبق" على الخصائص السيكومترية للاختبار التحصيلي ذي الاختيار من متعدد

إعـداد
د. بندر بن زيد سالم الأمير
الإدارة العامة للتعليم بمنطقة مكة المكرمة

اشراف

د/ عبد الله عبد الغنى صيرفى

مجـــلة كلـــية التربية – جامعـــة المنصـــورة
العدد ١١٢ – أكتوبر ٢٠٢٠

أثر البديل "جميع ما سبق" على الخصائص السيكومترية للاختبار التحصيلي ذي الاختيار من متعدد

د. بندر بن زيد سالم الأمير

المستخلص :

هدفت الدراسة إلى الكشف عن أثر استخدام البديل "جميع ما سبق" على الخصائص السيكومترية (الصدق والثبات) لاختبار تحصيلي من نوع اختيار من متعدد في مادة الرياضيات وعلى خصائص فقراته (الصعوبة والتمييز) ، وكذلك معرفة الفروق بين معاملات (صعوبة ، وتمييز) الفقرات لكل من نموذجي الاختبار ، والكشف عن تأثير البديل "جميع ما سبق" على معاملات كل صدق المحك التلازمي وثبات الاتساق الداخلي ، وتكون مجتمع الدراسة من طلاب الصف الثالث متوسط لعام ١٤٣٩-١٤٤٠هـ في مدارس إدارة التعليم بمدينة مكة المكرمة والذي بلغ عددهم (٤٠٧٥٦) طالباً موزعين على (١٦١) مدرسة أما عينة الدراسة فتكونت من (٥٠٠) طالب موزعين على (١٢) مدرسة وقد اختيرت العينة بالطريقة العشوائية الطبقية ، وتم بناء اختبار تحصيلي لطلبة الصف الثالث متوسط في ثلاث فصول دراسية من مادة الرياضيات ، وتكون الاختبار في صورته النهائية من (٣٤) فقرة ذات أربع بدائل في نموذجين من الاختبار النموذج الأول لم تتضمن بدائل فقراته البديل "جميع ما سبق"، أما النموذج الثاني فتضمنت بدائل فقراته البديل "جميع ما سبق"، وتم تطبيق النموذجين على عينة الدراسة بواقع (٢٥٠) طالباً لكل نموذج ، وبعد المعالجة الإحصائية للبيانات توصلت الدراسة إلى عدة نتائج : وجود فروق دالة إحصائياً بين قيم معاملات صعوبة فقرات الاختيار من متعدد تعزى إلى نموذج الاختبار المتضمن البديل "جميع ما سبق"، ، وهذه الفروق بين النموذجين تعود إلى أن استخدام البديل "جميع ما سبق"، كإجابة صحيحة تزيد من صعوبة الاختبار ، كما توصلت إلى عدم وجود فروق دالة إحصائياً بين قيم معاملات تمييز فقرات الاختيار من متعدد تعزى إلى نموذجي الاختبار ، وعدم وجود فروق دالة إحصائياً بين معاملي ثبات الاتساق الداخلي للاختبارين تُعزى إلى نموذجي الاختبار ، وعدم وجود فروق ذات دلالة إحصائية بين معاملي صدق المحك التلازمي للاختبارين تعزى إلى نموذجي الاختبار ، وعلى ضوء هذه النتائج توصلت الدراسة إلى مجموعة من التوصيات منها: أنه يُفضل عدم استخدام البديل "جميع ما سبق" عند كتابة فقرات من نوع الاختيار من متعدد وذلك لأنه يزيد من صعوبة الاختبار ، كما أوصت الدراسة ببناء اختبارات تحصيلية للمقارنة بين الخصائص السيكومترية للاختبار (الصدق والثبات) وفقراته (الصعوبة والتمييز) والتي تتناول استخدام البديل "جميع ما سبق" في فقراته لمواد أخرى ولدى مراحل تعليمية مختلفة.

الكلمات المفتاحية : جميع ما سبق ، الخصائص السيكومترية ، الاختبار التحصيلي ذي الاختيار من متعدد ، معامل الصعوبة ، معامل التمييز

Abstract

The aim of this study was to investigate the effect of the use of the alternative "All the above" on the characteristics of the psychometric (Validity and Reliability) to the multiple-choice achievement test in the field of Mathematics and the characteristics of its items (difficulty and discrimination) and also distinguish the differences of both test models, as well to detecting the effect of the alternative "All the above" on the indexes of both stability of internal consistency and the validity of the collaterally test. The study population consisted of intermediate students in the third grade (1439/ 1440) in the schools of the Education Administration in Makkah city which their number reached (40756) students distributed over(161) school. The sample of the study consisted of (500) students distributed over (12) schools. The sample was selected by the random stratified method. Achievement test was constructed for the third grade intermediate students in three semesters of mathematics. The final test consisted of (34) item of four alternatives in two models of the test. The first model did not include alternatives of its alternative paragraphs "All the above", but the second model included alternatives of its alternative paragraphs "All the above". The two models were applied to a sample of study with (250) students per model. After statistical analysis treatment of the data, the results of the study were as follows : There are statistically significant differences between the indexes of the difficulty of multiple choice items due to the test model which include the alternative "All the above". And there are no statistically significant differences between the Discrimination indexes of the multiple-choice items attributed to the two test models. And there are no statistically significant differences between the two indexes of the Reliability of internal consistency for both tests attributed to the two test models. And there are no statistically significant differences between the two indexes of the validity of the collaterality test for both tests attributed to the two test models. On the basis of these results, this study came out with a set of recommendations, including : Prefer not to use when writing items of multiple choice type because it makes the test more difficult. Constructing diagnostic tests to compare the psychometric characteristics of the test (Validity and Reliability) and its items (difficulty and discrimination), which deal with the use of the achievment "All the above" in the multiple choice paragraphs of other materials.

Keywords : *"All the above", Psychometric Characteristics, Multiple-choice achievment Test*

مقدمة

تعتبرالاختبارات التحصيلية من أهم وسائل التقويم التي يستخدمها المعلمون في المدارس ، وذلك لمعرفة مدى التقدم في تحقيق الأهداف داخل غرفة الصف ، والاختبارات المستخدمة في المدارس لتقويم الطلاب هي على الأغلب اختبارات من إعداد المعلمين أنفسهم ، وهي شائعة الاستخدام ويعود انتشارها إلى أسباب عديدة .

وقد ذكر كاري(Carey,1994) بعضاً من هذه الأسباب: منها أن المعلم يستطيع باستخدامها قياس أنواع مختلفة من التعلم وعلى اختلاف المستويات .

ويمكن تصنيف الاختبارات التحصيلية وفق معايير مختلفة ، أحدها أن تصنف حسب طبيعة الإجابة المطلوبة فتصنف إلى: الاختبارات ذات الإجابة المصاغة والاختبارات ذات الإجابة المنتقاة . ويتوقف استخدام هذه الأنواع من الفقرات على مجموعة من العوامل منها : مستوى الهدف المراد قياسه ، وعمر المفحوص ، وغرض الاختبار ، وعدد الطلبة ، وظروف التطبيق ، ومهارة المعلم في صياغة الأسئلة (عودة ، ٢٠١٠).

كما تُعد الاختبارات الموضوعية من أكثر أنواع الاختبارات استخداماً في النظام التربوي في السنوات الأخيرة حيث أصبحت أكثر انتشارا عن ذي قبل ، لما تتميز به من سهولة التصحيح ودقته وثباته وموضوعيته ، وإمكانية استخدامها في قياس المستويات المعرفية العليا بشكل أوسع من ذي قبل (الأمير ،٢٠١٣)

ولعل أهم الاختبارات الموضوعية اختبار الاختيار من متعدد والذي يعتبر من أهم أدوات تقويم الطالب وأكثرها شيوعاً واستخداماً ، حيث يذكر (Aiken , 2003) أنها من أكثر الاختبارات انتشاراً في التربية ، والمجالات التنظيمية الأخرى ، وهذا النوع من الاختبارات قد تفوقت على كل أشكال الفقرات الموضوعية الأخرى ، وذلك بسبب كفاءتها وتعدد استعمالاتها ،إذ يمكن بواسطتها قياس أهداف بسيطة وأخرى مركبة في مختلف المواضيع الدراسية.

وتتكون جميع فقرات الاختيار من متعدد من العناصر نفسها ، وأول هذه العناصر هو المتن أو الأرومة Stem)) ، الذي يمكن أن يكون على شكل سؤال ، أو جملة إخبارية تتضمن قضية معينة تحتاج إلى إجابة ، أما العناصر الأخرى في فقرة الاختيار من متعدد فهي الخيارات أو البدائل Option or alternatives)) ، وهي إجابات محتملة للسؤال الوارد في المتن ، وأحد هذه البدائل هو الإجابة الصحيحة ، أما بقية البدائل فتعرف بالمموهات أو المشتتات (Distracters or Foils) (Haladyna & Roid, ١٩٨٢)

وعند تطوير اختبار من نوع من الاختيار من متعدد يجب أخذ عدة أمور تؤثر على الخصائص السيكومترية للاختبار بعين الاعتبار ، مثل (الصدق الثبات) ، وفي الخصائص السيكومترية للفقرة (الصعوبة والتمييز) ، ومن هذه الأمور شكل البدائل ، حيث أوضح (Haladyan & Crehan, ١٩٩١) بأنها تختلف من اختبار لآخر فهناك البدائل البسيطة (التقليدية) والتي تحوي إجابة واحدة صحيحة ، ومجموعة من المموهات ، وهناك بدائل مركبة تستخدم بزيادة صعوبة وتمييز فقرات الاختبار ، بحيث يكون أحد هذه المموهات البديل "جميع ما سبق" أو "لا شيء مما ذكر" أو "جمع بديلين" مثل (ب+ج).

ويشير (Downing,2002) إلى أنه يمكن وصف خصائص فقرة الاختيار من متعدد "الفعالة" أو "الجيدة" من حيث الفقرة بشكل عام ، والمتن ، والبدائل والفقرة "المخالفة" أو "غير الجيدة" تتعارض مع ترجمة وتفسير علامات الاختبار بشكل صحيح ، وتؤثر سلباً على معدلات نجاح الطلاب ، لذلك ومن أجل تطوير اختبارات عالية الصدق والثبات لا بدّ أن يتم بناء وصياغة فقرات خالية من مخالفات قواعد كتابة فقرات الاختيار من متعدد.

وفي دراسة مسحية قام بها هالادينا وداوننج ورودوجيوز (Downing, ,Haladyna Rodriguez,2002) صنفوا فيها إحدى وثلاثين إرشاداً يتعلق بصياغة فقرات اختبار الاختيار من متعدد ، نتيجة مراجعتهم سبعة وعشرين مرجعاً متخصصاً في القياس والتقويم ، وأيضاً من خلال سبع وعشرين دراسة تجريبية في القياس والتقويم منذ عام (١٩٩٠) وجدوا أن الإرشاد الخامس والعشرين الذي يتعلق بإرشادات كتابة بدائل الفقرة والذي ينص على أنه (يمكن استخدام البديل " جميع ما سبق " ولكن مع أخذ الحيطة والحذر) لم يجد اتفاق بين الباحثين حول استخدامه من عدمه في اختبارات الاختيار من متعدد ، فبينما يصر البعض ومنهم ثورندايك وهيجن (١٩٨٩) ورويد وهالادينا (Haladyna & Roid, ١٩٨٢) وأوسترهوف (Osterhof, ١٩٩٤) و(الصبح،٢٠٠٨) ، إلى أنه يمكن استعمال هذا البديل في حالات معينة وهي إما أن يكون للفقرة إجابة مقطوعاً بصحتها بشكل مطلق كما في اختبارات الرياضيات مثلاً ، وإما أن يكون الغرض هو إيجاد عدد كاف من البدائل المناسبة ، أو عند عدم وجود بديل آخر متجانس مع بقية البدائل ، أو بديل غير مقبول ظاهرياً لدى المفحوصين ، وإما أمكانية استخدام البديل "جميع ما سبق" بفعالية مرة كل أربعة مرات استخدام ، بالمقابل هناك وجهة نظر أخرى يتبناها مجموعة من الباحثين ومنهم بوفام (Popham, ٢٠٠٠) وماكميلان (McMillan, 2004) وهما يؤكدان على ضرورة تجنب هذا البديل ، ويعللون ذلك

بأنه : يقيس القدرات لدى المفحوص في اختياره أو عدم اختياره ، ولا يقيس معرفة المفحوص بمحتوى الفقرة ، وأن المعلمين يميلون لوضعه كإجابة خاطئة عند عدم مقدرتهم إيجاد مموه منطقي ، وأن استخدام هذا البديل يجعل الاختبار أكثر صعوبة ، كذلك يأخذ وقتاً أطول من المفحوص للإجابة .

يتضح مما سبق أن استخدام البديل "جميع ما سبق" في فقرات الاختيار من متعدد يدور حوله الكثير من الجدل والخلاف ؛ حيث أن هناك انقسام بين مؤيد ومعارض نتيجة لتأثيره على مدى صعوبة وتمييز وثبات الاختبار ، لذا فإن هذه الدراسة تحاول دراسة أثر استخدام البديل "جميع ما سبق" على الخصائص السيكومترية للاختبار التحصيلي ذي الاختيار من متعدد والذي يعطينا القدرة في الحكم على جودة فقرات الاختبار من متعدد لكي تساعدنا في بناء اختبارات التحصيل بصورة فعالة.

مشكلة الدراسة وتساؤلاتها

لوحظ من خلال عمل الباحث في الميدان التربوي ، أن بعض المعلمين عند بنائهم لفقرات الاختيار من متعدد في اختبارات التحصيل المدرسية في حيرة من أمرهم حيال الالتزام ببعض قواعد صياغة الفقرات ، ومن بينها قاعدة استخدام البديل "جميع ما سبق" كأحد بدائل فقرات الاختيار من متعدد ، ويمكن تفسير ذلك نتيجة لما نصت عليه هذه القاعدة من حيث أخذ الحيطة والحذر عند استخدام هذا النوع من البدائل تارة ، والإلزام بعدم استخدامه تارةً أخرى ، ومن ناحية أخرى قد يكون استخدام البديل "جميع ما سبق" له انعكاس سلبي من خلال خفض أو زيادة مستوى صعوبة الاختبار.

وهنا تجلت مشكلة الدراسة واقتضت تقصي أثر البديل "جميع ما سبق" على الخصائص السيكومترية للاختبار وفقراته. ومن خلال إطلاع الباحث على عدد من أدبيات الدراسة في هذا المجال لاحظ أن معظم الدراسات التي أجريت لمعرفة أثر البديل "جميع ما سبق" على الخصائص السيكومترية للاختبار وفقراته ، ذكرت أن معامل الصعوبة في الفقرات التي تحتوي البديل "جميع ما سبق" أعلى من الفقرات التي لا تحتوي هذا البديل كدراسة فورسيث وسبرات (Spratt,1980 & Forsyth) ، بالإضافة إلى دراسة فراري (Frary,1991) والتي أشارت إلى أنه تزداد صعوبة الفقرات عند استخدام البديل "جميع ما سبق" كجواب صحيح فضلاً عن كونه أحد المموهات ، كما أن القدرة التميزية للفقرات عند استخدام هذا البديل من عدمه كانت متساوية.

ووجد الباحث أيضاً بعض الدراسات التي عارضت الدراسات السابقة ، والتي أثبتت أنه ليس هناك فروق دالة احصائياً في صعوبة الفقرات التي تحتوي البديل "جميع ما سبق" وكان تمييزها أقل عند استخدام هذا البديل كجواب صحيح كدراسة تولفسون وتريب Tollefson & Tripp عام ١٩٨٣، وكذلك دراسة ويسمان وبنيت Bennett & Wesman عام ١٩٨٥ والتي أشارت إليها (Tollefson,1987) والتي أثبتت أن استخدام البديل "جميع ما سبق" ليس له تأثيراً على ثبات الاختبار وصعوبته مقارنة بالبدائل العادية.

وبالنظر لما سبق يمكن ملاحظة أن هناك عدم وجود اتساق في نتائج الدراسات التي أجريت في أثر البديل "جميع ما سبق" على الخصائص السيكومترية للاختبار وفقراته ، وهذا ما يشكل غموض وتردد حول استخدامه من قبل المعلمين عند بناء الاختبارات وكذلك في مدى فعالية هذا البديل على الاختبارات ، كما يلاحظ ندرة البحوث العربية التي تطرقت لهذا الموضوع رغم أهميته ، لذا جاءت هذه الدراسة امتداداً واستكمالاً لجهود الدراسات والبحوث السابقة واستجابة لبعض توصياتها ، إذ لم يجد الباحث دراسات في البيئة السعودية تناولت هذا الموضوع ، كما أن هذه الدراسة تحاول معرفة مدى اختلاف الخصائص السيكومترية للاختبار وفقراته إثر وجود البديل "جميع ما سبق" ضمن فقرات الاختيار من متعدد في مادة الرياضيات لدى طلاب الصف الثالث متوسط.

ومن ثم تحدد مشكلة الدراسة في محاولة الإجابة على التساؤلات التالية:

١- هل توجد فروق ذات دلالة إحصائية بين متوسطات قيم معاملات الصعوبة في نموذجي الاختبار التحصيلي لمادة الرياضيات من نوع الاختيار من متعدد تعزى إلى وجود البديل "جميع ما سبق" أو عدم وجوده ؟

٢- هل توجد فروق ذات دلالة إحصائية بين متوسطات قيم معاملات التمييز في نموذجي الاختبار التحصيلي لمادة الرياضيات من نوع الاختيار من متعدد تعزى إلى وجود البديل "جميع ما سبق" أو عدم وجوده ؟

٣- هل توجد فروق ذات دلالة إحصائية بين قيم معاملات صدق المحك التلازمي في نتائج نموذجي الاختبار التحصيلي لمادة الرياضيات من نوع الاختيار من متعدد تعزى إلى وجود البديل "جميع ما سبق "أو عدم وجوده ؟

٤- هل توجد فروق ذات دلالة إحصائية بين قيم معاملات ثبات الاتساق الداخلي في نتائج نموذجي الاختبار التحصيلي لمادة الرياضيات من نوع الاختيار من متعدد تعزى إلى وجود البديل "جميع ما سبق "أو عدم وجوده ؟

أهداف الدراسة

هدفت هذه الدراسة بشكلٍ عام إلى المحاولة لمعرفة أثر البديل "جميع ما سبق" على الخصائص السيكومترية للاختبار وفقراته ، وتتلخص أهدافها فيما يلي :

١- التعرف على الفروق بين قيم معاملات الصعوبة في نموذجي الاختبار التحصيلي لمادة الرياضيات من نوع الاختيار من متعدد والتي تعزى إلى وجود البديل "جميع ما سبق" أو عدم وجوده.

٢- التعرف على الفروق بين قيم معاملات التمييز في نموذجي الاختبار التحصيلي لمادة الرياضيات من نوع الاختيار من متعدد والتي تعزى إلى وجود البديل "جميع ما سبق" أو عدم وجوده.

٣- التعرف على الفروق بين قيم معاملات صدق المحك التلازمي في نتائج نموذجي الاختبار التحصيلي لمادة الرياضيات من نوع الاختيار من متعدد والتي تعزى إلى وجود البديل "جميع ما سبق" أو عدم وجوده.

٤- التعرف على الفروق بين قيم معاملات ثبات الاتساق الداخلي في نتائج نموذجي الاختبار التحصيلي لمادة الرياضيات من نوع الاختيار من متعدد والتي تعزى إلى وجود البديل "جميع ما سبق" أو عدم وجوده.

أهمية الدراسة

١- تتضح أهمية هذه الدراسة من أهميتها الواقعية من خلال دراستها للخصائص السيكومترية للاختبار وفقراته، ومحاولتها التوصل إلى معطيات تُزيل الغموض حول استخدام البديل "جميع ما سبق"، فهي تضع لبنة علمية تراكمية قد تساعد المهتمين بالقياس والتقويم في تجويد بناء اختبارات التحصيل من نوع الاختيار من متعدد بشكل عام ، ومعلمي مادة الرياضيات بشكل خاص.

٢- تسهم هذه الدراسة في تقديم أداة قياس جديدة لمعلمي الرياضيات صالحة للاستخدام مستقبلاً لغرض تقويم طلاب الصف الثالث متوسط، لما تتمتع به من درجة عالية من الصدق والثبات والقدرة التميزية.

٣- كونها الدراسة الأولى – على حد علم الباحث – التي تناولت أثر البديل " جميع ما سبق " على الخصائص السيكومترية للاختبار التحصيلي ذي الاختيار من متعدد في المملكة العربية السعودية .

مصطلحات الدراسة الإجرائية

١- الاختبارات التحصيلية Achievement Test :

يعرف (عودة ، ٢٠١٠) الاختبارات التحصيلية بأنها عينة من الأسئلة تمثل المحتوى العلمي للمادة الدراسية ، يقوم الطالب بالإجابة عليها لتحديد مستوى تحصيله لمعلومات ومهارات تلك المادة.

ويعرفها الباحث إجرائياً بأنها : الأداة المستخدمة من إعداد الباحث والتي تشتمل على اختبار ذي اختيار من متعدد في الثلاث فصول من مادة الرياضيات للفصل الدراسي الأول لعام ١٤٣٩ هـ – ١٤٤٠هـ ، ويتكون من نموذجين دون تغيير في متن مفرداتهما ، ويتكون النموذج الأول من مجموعة من الفقرات ، وكل فقرة تحتوي على أربع بدائل أحدها صحيح ، وفي النموذج الثاني يتكون من نفس فقرات النموذج الأول مع إضافة البديل "جميع ما سبق" تارة كبديل صحيح بعد حذف الإجابة الصحيحة ، وتارة كبديل خاطئ في البدائل الأخرى للفقرات بعد حذف المموهات الأقل تمييزاً .

٢- اختبار الاختيار من متعدد The Multiple–Choice Tests :

تتكون المفردة في هذا النوع من الاختبارات من جزأين رئيسيين :الجزء الأول يسمى المتن ، وقد يكون على شكل سؤال أو عبارة تتضمن قضية معينة تحتاج إلى إجابة ، ويسمى الجزء الثاني بالبدائل والتي هي بمثابة حلول أو إجابات محتملة للسؤال أو القضية الواردة في المتن وتتضمن البدائل إجابة واحدة صحيحة وعدداً من الإجابات الخاطئة التي يطلق عليها بالمشتتات أو المموهات (عودة ، ٢٠١٠م) .

ويعرفه الباحث إجرائياً بأنه : اختبار تحصيلي مكون من ٣٤ مفردة ، من نوع الاختيار من متعدد في نموذجين ، النموذج الأول : عبارة عن مجموعة من الفقرات ، تتكون كل فقرة من المتن وأربعة بدائل واحدٌ منها صحيح ، ويطلب من المفحوص أن يختار الإجابة الصحيحة من بين أربعة بدائل لكل فقرة من فقرات الاختبار ، والنموذج الثاني : وهو النموذج الأول نفسه لكن الاختلاف في إضافة البديل "جميع ما سبق" كبديل صحيح بعد حذف الجواب الصحيح من بعض البدائل ، وكبديل خاطئ بعد حذف المموهات الأقل تمييزاً (فاعلية) ، في الثلاث

فصول (المعادلات الخطية ، العلاقات ،الدوال الخطية) من مادة الرياضيات للفصل الدراسي الأول من العام ١٤٣٩ – ١٤٤٠ هـ لطلاب الصف الثالث المتوسط.

٣- صعوبة المفردة Item Difficulty:

هي نسبة عدد الأفراد الذين أجابوا إجابة صحيحة على فقرة ما بالنسبة للعدد الكلي (علام ،٢٠٠٦)

التعريف الإجرائي: يقصد بصعوبة الفقرة وفق النظرية الكلاسيكية هي نسبة الطلاب الذين أجابوا إجابة صحيحة عن كل فقرة من فقرات الاختبار من بين الطلاب الذين حاولوا الإجابة على تلك الفقرات في نموذجية .

٤- تمييز المفردة Item Discrimination :

يشير معامل التمييز إلى قدرة الفقرة على التمييز بين من يمتلكون المعرفة ومن لا يمتلكونها في اختيار البديل الصحيح .(Algina & Crocker, ١٩٨٦)

التعريف الإجرائي : يقصد بتمييز الفقرة وفق النظرية الكلاسيكية هو قيمة معامل الارتباط بين درجة الفرد على الفقرة مع درجته الكلية على الاختبار مقاساً بمعامل الارتباط الثنائي النقطي (Biserial Point) ، (Algina & Crocker, ١٩٨٦).

٥- الثبات Reliability :

هو حصول الفرد على نفس الدرجات إذا طبق عليه نفس الأداة وتحت نفس الظروف. (مراد،سليمان ، ٢٠٠٥) .

التعريف الإجرائي للثبات : يقصد به في هذه الدراسة معامل ثبات الاتساق الداخلي للاختبار محسوب باستخدام طريقة كودر – ريتشاردسون (٢٠) (KR-20).

٦- الصدق Validity :

يعرف مراد و سليمان (٢٠٠٥) بأن صدق الاختبار مدى صلاحيته لقياس ما وضع لقياسه ، بمعنى أن الاختبار الصادق يقيس ما وضع لقياسه فقط ، فاختبار الذكاء يقيس الذكاء ولا يقيس أي شيء آخر.

التعريف الإجرائي للصدق : يقصد به في هذه الدراسة صدق المحك التلازمي ، وهو عبارة عن قيمة معامل ارتباط بيرسون المحسوب بين درجات الطلاب على الاختبار التحصيلي المستخدم في الدراسة ، ودرجاتهم في الاختبار النهائي لمادة الرياضيات في الفصل الأول من العام ١٤٣٩/١٤٤٠هـ.

حدود الدراسة

حدود موضوعية : اقتصار الدراسة على أثر البديل "جميع ما سبق" على الخصائص السيكومترية للاختبار (الصدق والثبات) وفقراته (معامل الصعوبة ومعامل التمييز) بتطبيق نموذجين من الاختبارات من نوع الاختيار من متعدد (لا يتضمن البديل "جميع ما سبق" ، يتضمن البديل "جميع ما سبق").

حدود مكانية : الإدارة العامة للتعليم بمدينة مكة المكرمة.

حدود بشرية : طلاب الصف الثالث متوسط من المدارس الحكومية

حدود زمانية : الفصل الدراسي الأول من العام الدراسي ١٤٣٩-١٤٤٠هـ

منهج الدراسة :

لتحقيق أهداف هذه الدراسة اعتمدت هذا الدراسة على المنهج الوصفي المقارن الذي يعتبر عن الأسلوب الأمثل في دراسة مجالات الظواهر الإنسانية والطبيعية المختلفة .

مجتمع الدراسة :

يتكون مجتمع الدراسة من جميع طلاب الصف الثالث متوسط بالمدارس الحكومية بمدينة مكة المكرمة خلال الفصل الدراسي الأول للعام الدراسي ١٤٣٩/١٤٤٠هـ ، والبالغ عددهم بحسب الرجوع إلى التقرير الإحصائي التربوي السنوي من إدارة تقنية المعلومات بالإدارة العامة للتعليم بمكة المكرمة (٤٠٧٥٦) طالباً يتوزعون في (١٦١) مدرسة متوسطة .عينة الدراسة

أ- عينة الدراسة الاستطلاعية:

تكونت من (١٠٠) طالب من طلاب الصف الثالث متوسط بمدينة مكة المكرمة ، وذلك بغرض التحليل الأولي لفقرات الاختبار ، وشكلت نسبة ٢٠٪ من أفراد العينة النهائية

ب- عينة الدراسة الأساسية :

تم اختيار عينة الدراسة بطريقة العينة الطبقية العشوائية حتى تكون العينة ممثلة للمجتمع الأصلي للدراسة ، حيث تكونت من طلاب (١٠) مدارس بمدينة مكة المكرمة بواقع (٢٥٠) طالباً لكل نموذج .

أداة الدراسة :

ولتحقيق أهداف الدراسة وهي اختبار أثر البديل "جميع ما سبق" على صدق وثبات اختبار تحصيلي ذي اختيار من متعدد في مادة الرياضيات وعلى معاملي الصعوبة والتمييز

لمفرداته،قام الباحث ببناء اختبارتحصيلي في مادة الرياضيات للصف الثالث متوسط الفصل الدراسي الأول ١٤٣٩/١٤٤٠هـ ، وفيما يلي المراحل والإجراءات التي اتبعها الباحث :

أولاً : تحديد الموضوعات المراد قياسها :

اختار الباحث ثلاث فصول دراسية من مادة الرياضيات لطلاب الصف الثالث متوسط الفصل الدراسي الأول ١٤٣٩/١٤٤٠هـ ، و هي كما يلي :

الفصل الأول : المعادلات الخطية.

الفصل الثاني: العلاقات.

الفصل الثالث: الدوال الخطية .

ويرجع السبب في اختيار الباحث لهذا المنهج الدراسي ؛ لأنه تخصص الباحث في المجال التدريسي.

ثانياً : تحديد الأهداف السلوكية :

قام الباحث بتحديد الأهداف السلوكية التي تقيس محتوى الفصل الأول والثاني والثالث من مادة الرياضيات لطلاب الصف الثالث متوسط بالفصل الدراسي الاول ، وقد استفاد الباحث من كتاب المعلم لمادة الرياضيات للفصل الدراسي الأول طبعة عام ١٤٣٩/١٤٤٠هـ ، للحصول على الأهداف الخاصة في الفصول التي سيبنى عليها الاختبار ، حيث قام عدد من المتخصصين بتحليل محتوى المادة إلى الحقائق والمفاهيم والمهارات المتضمنة في كتاب الطالب ، وكتبت على هذا الأساس الأهداف السلوكية للمادة ، وقد عرضت هذه الأهداف على عدد من المحكمين ، لتقييمها والتحقق من انطباق شروط ومعايير مواصفات الهدف السلوكي الجيد عليها ، وكذلك التأكد من مناسبة تحديد المستوى المعرفي للهدف وفقاً لتصنيف بلوم للمجال المعرفي ، وبالتالي تكون هذه الأهداف قد حكمت بصورة جيدة وموثوق بها ، ويمكن للباحث الاعتماد عليها في بناء دراسته.

ثالثاً : إعداد جدول المواصفات :

لإعداد جدول الموصفات قام الباحث بالخطوات التالية :

أ- تحديد الوزن النسبي للأهداف السلوكية بمستوياتها المختلفة:

ب- تحديد عدد الحصص اللازمة لتدريس كل موضوع ، حيث قام الباحث بتصميم جدول يتضمن العمود الأول موضوعات الفصل الأول والفصل الثاني والفصل الثالث من مادة الرياضيات ، والعمود الثاني عدد الحصص المقترحة لكل موضوع ، وتم توزيعها على

عدد من المشرفين التربويين وعدد من معلمي الرياضيات للصف الثالث متوسط، ومن ثم استخرج الباحث متوسطات مقترحات المشرفين والمعلمين واعتمدها في بحثه .

ج- تحديد الوزن النسبي لموضوعات المادة الدراسية.

د- تحديد العدد الكلي لأسئلة الاختبار في ضوء الزمن المتاح للإجابة ، ونوع الفقرة المستخدمة في الاختبار ، وأعمار الطلاب ومستواهم التعليمي ، وطول الفقرة ومستوى صعوبتها ، ومقدار ما تتطلبه الفقرات من حساب كمي ، ومستوى العمليات العقلية التي تقيسها الفقرة.

هـ- حساب عدد الأسئلة في كل موضوع لكل مستوى من مستويات الاهداف .

و- تحديد درجة أسئلة كل موضوع في كل مستوى من مستويات الأهداف .

وتوصل الباحث إلى جدول المواصفات التالي جدول (١) :

جدول (١):

مواصفات الاختبار التحصيلي للوحدة الأولى والثانية والثالثة للفصل الدراسي الأول من مادة الرياضيات لطلاب الصف الثالث المتوسط للعام ١٤٣٩هـ - ١٤٤٠هـ

م	الوحدات	الأسئلة والدرجات	الأهداف						مجموع الأسئلة	مجموع الدرجات	الأوزان النسبية للموضوعات
			تذكر	فهم	تطبيق	تحليل	تركيب	تقويم			
١	الفصل الأول المعادلات الخطية	الأسئلة	٤	٣	٠	١	٠	٠	٨	٨	%٢٠
		الدرجة	٤	٣	٠	١	٠	٠			
٢	الفصل الثاني العلاقات	الأسئلة	٢	٢	٠	٠	٠	٠	٤	٤	%١٠
		الدرجة	٢	٢	٠	٠	٠	٠			
٣	الفصل الثالث الدوال الخطية	الأسئلة	٢	٢	٠	٠	٠	٠	٤	٤	%١٠
		الدرجة	٢	٢	٠	٠	٠	٠			
		الدرجة	٢	٢	٠	٠	٠	٠			
		الدرجة	٠	٠	٠	٠	٠	٠			
	مجموع الأسئلة		٢٠	١٧	٠	٣	٠	٠	٤٠		
	مجموع الدرجات		٢٠	١٧	٠	٣	٠	٠		٤٠	
	الأوزان النسبية للأهداف		%٥٠	%٤١	%٠	%٩	%٠	%٠			%١٠٠

رابعاً : كتابة وصياغة الأسئلة :

قام الباحث بكتابة وصياغة اختباراً تحصيلياً مكون من (٤٥) فقرة من نوع الاختيار من متعدد بأربعة بدائل ، واحد منها فقط صحيح ، أي أن الإجابة تأخذ إحدى القيميتين (١،٠) وذلك في مادة الرياضيات لطلاب الصف الثالث متوسط، في الفصل الأول (المعادلات الخطية) والفصل الثاني (العلاقات) والفصل الثالث (الدول الخطية) ، وذلك لقياس الأهداف الواردة في جدول المواصفات ، مرتبة وفقاً لتسلسل عرض المحتوى .

خامساً : تحكيم أداة الدراسة :

للتأكد من صدق محتوى فقرات الاختبار قام الباحث بتحليل فقرات الاختبار بغرض تعيين صدق فقرات الاختبار والتحقق من أن الفقرات تقييس ما وضع لقياسه ، واعتمد في بحثه على الأسلوب المنطقي (طريقة التمايز اللغوي) للتحقق من صدق فقرات الاختبار التحصيلي.

ويرجع اختيار الباحث لهذه الطريقة لعدة اعتبارات منها :

اعتماد نتائجها على أكثر من معيار احصائي (الوسط الحسابي ، الوسيط ، المدى)

سهولة الحصول على استجابات المحكمين

عدم احتياجها إلى عمليات إحصائية معقدة لتحليل نتائجها.

وفيما يلي بيان تفصيلي بالخطوات التي اتبعها الباحث في تنفيذ اجراءات الطريقة :

١- قام الباحث ببناء استبيان يشتمل على ثلاثة أقسام : (القسم الاول : يضم أربع فقرات ترتبط بجدول المواصفات ، القسم الثاني : يضم سبع فقرات ترتبط بأسئلة الاختبار التحصيلي المعد ، القسم الثالث : يضم فقرة واحدة ترتبط بالاختبار التحصيلي المعد ، ومقياس تدريجي لكل قسم يمتد من (١) إلى (٥) حيث يشير العدد (١) إلى أدنى ارتباط لمدى مطابقة كل قسم على ما جاء في كل فقرة من فقراته بينما (٥) إلى أعلى ارتباط في كل قسم على ما جاء في كل فقرة من فقراته.

٢- قام الباحث بتوزيع الاستبيان المعد مسبقاً على ثمانية محكمين من الاساتذة والمشرفين التربويين لمادة الرياضيات ثم طلب من كل منهم أن يجيب من وجهة نظره على فقرات الاستبيان بأقسامه الثلاث.

٣- قام الباحث بعد الحصول على استجابات المحكمين بتفريغها في جدول خاص ، لحساب المتوسطات الحسابية لكل فقرة يشتمل عليها كل قسم من أقسام الاستبيان.

٤- قام الباحث باعتبار أن كل فقرة يزيد وسطها الحسابي عن (٤) ، فقرة ممتازة لمدى مطابقتها لكل قسم من أقسام الاستبيان الثلاثة أما الفقرة التي يقل وسطها عن ٤ ، ومن خلال الخطوات السابقة تم حذف ٥ عبارات لم تتوافر فيها الشروط السابقة .ومن ثم اصبح عدد عبارات الاختبار حتى هذه المرحلة مكون من ٤٠ فقرة .

سادساً : التجربة الاستطلاعية :

قام الباحث بتطبيق الاختبار المكون من ٤٠ فقرة على عينة استطلاعية تكونت من ١٠٠ طالب ، تم اختيارهم عشوائياً من إحدى المدارس من خارج عينة الدراسة ، موزعين على

أربع مدارس من أربع مكاتب تعليم مختلفة ، وتم إبلاغ إدارة المدارس بموعد الاختبار ، والمواضيع المحددة ، وطريقة استخدام نماذج الإجابة ، وكان الهدف من هذه التجربة التأكد من وضوح تعليمات الاختبار ، وتحديد الزمن المناسب ، والتحليل الأولي لفقرات الاختبار.

وتم إجراء التحليل الأولي التالي لمفردات الاختبار :

- حساب معامل صعوبة كل مفردة من المفردات ، وهي نسبة الطلاب الذين أجابوا على المفردة إجابة صحيحة
- حساب معامل تمييز كل مفردة من المفردات وهو معامل الارتباط بين نتائج المفحوصين على كل فقرة ونتائجهم على الاختبار الكلي
- حساب فعالية المموهات (المشتتات)
- حساب معامل صدق الاختبار من خلال صدق الاتساق الداخلي
- حساب معامل الثبات بطريقة التجزئة النصفية ومعامل الاتساق الداخلي (كودر- ريتشاردسون) وذلك لمناسبته للفقرات التي إجابتها (٠ ، ١)

يتضح من الخطوات السابقة أن معاملات الصعوبة في نموذج الصورة الأولية للاختبار تراوحت بين (٠,١٥- ٠,٨٠) بمتوسط حسابي قدره (٠,٤٧) ، أما معاملات التمييز فقد تراوحت بين (٠,١٠-٠,٦٤) بمتوسط حسابي قدره (٠,٣٧) ، وبعد النظر بالفقرات التي لا تحقق الإحصائيات المتبعة في هذه الدراسة وهي الإحصائيات المقترحة من قبل (Eble,1972) والتي تتلخص بما يلي:

- الفقرات التي معامل تمييزها (سالب) تحذف ولا داعي للاحتفاظ بها.
- الفقرات التي معامل تمييزها من(٠ – ٠,١٩) تعتبر ضعيفة التمييز وينصح بحذفها.
- الفقرات التي معامل تمييزها من(٠,١٩ – ٠,٣٩) ذات تمييز مقبول وينصح بتحسينها.
- أي فقرة معامل تمييزها أعلى من(٠,٣٩) تعتبر فقرة ذات تمييز جيد ويمكن الاحتفاظ بها.
- أي فقرة معامل صعوبتها بين(٠,٣٠ – ٠,٨٠) تعتبر مقبولة ويمكن الاحتفاظ بها.
- قام الباحث في ضوء هذه المعايير بحذف الفقرات (٤ ،١١، ١٥) وذلك لتدني معاملات الصعوبة لها حيث بلغت معاملات الصعوبة لهذه الفقرات على التوالي (٠,٢٠ ، ٠,٢٣، ٠,١٥) وتم حذف الفقرة (٤٠) لأن معامل تمييزها سالب حيث بلغ معامل التمييز لهذه الفقرة (-٠,٢٢) ، كما تم حذف الفقرات (٤ ، ٢٤ ، ٣٦) وذلك لتدني معامل التمييز لهم

حيث بلغ معامل التمييز للفقرات على التوالي (٠,٠٠ ، ٠,١٨ ، ٠,٠٢) أما معاملات الصعوبة والتمييز لفقرات الاختبار المتبقية فقد كانت مقبولة ، وسيتم إجراء أي تعديل أو تحسين على هذه الفقرات في نموذج الصورة الأولية للاختبار.

كما تم تعديل بعض المموهات التي جاذبيتها موجبة أو صفر وهي : (٦-ج) ، (٩-د) ، (١٠- ب) ، (١٢- د) ، (٢٦- ج) ، (٢٨- أ) ، (٣١- ب- د) ، وفضل الباحث بقاء المموه (١٨- ب) كما هو وذلك لتناسقه مع بقي المموهات في الفقرة نفسها.

كما اعتمد الباحث في تحديد صدق الاختبار على صدق الاتساق الداخلي حيث تم حساب معامل ارتباط بيرسون بين الفقرة والدرجة الكلية للاختبار بعد حذف درجة الفقرة من الدرجة الكلية للاختبار .حيث تراوحت معاملات الارتباط من (٠,٠٥٠ ٠,٦٧٦) حيث وجد أن جميع الفقرات لها ارتباط بالدرجة الكلية للاختبار ودالة احصائياً ما عدا الفقرات (٤ ، ١١ ، ١٥ ، ٢٤ ، ٣٦ ، ٤٠) فهي غير دالة.

كما تم حساب ثبات الاختبار باستخدام معادلة كودر- ريتشاردسون (KR-20) وطريقة التجزئة النصفية ، حيث كان معامل الثبات بطريقة كودر - ريتشارسون بلغت (٠,٨٦) وهو معامل ثبات مرتفع ، كما بلغ معامل الثبات بطريقة التجزئة النصفية (٠,٨٨) وهو معامل ثبات جيد ، مما سبق يدل على أن المقياس يتصف بالثبات.

وأعيد حساب الثبات بعد حذف ستة مفردات باستخدام معادلة كودر- ريتشاردسون (KR-20) فكانت قيمة معامل الثبات لهذا النموذج (٠,٨٧) .

وفي ضوء التجربة الاستطلاعية تم تكوين نموذجين للاختبار نفسه ، يتكون كل منها من ٣٤ مفردة على النحو التالي :

النموذج الأول : احتوت كل مفردة من مفرداته على أربعة بدائل عادية.

النموذج الثاني : احتوت كل مفردة من مفرداته على أربعة بدائل (النموذج الأول نفسه) منها البديل "لا شيء مما ذكر "جواباً صحيحاً حيث حل مكان الإجابة الصحيحة في بعض فقرات الاختبار وهي (١-أ) ،(٣-ب)،(٥-د)،(٧-ج)،(٩-أ)،(١١-أ)،(١٣-ب)،(١٥-أ)،(١٧-ج)،(١٩-ب)،(٢١-أ)،(٢٣-د)،(٢٥-أ)،(٢٧-ج)،(٢٩-ب)،(٣١-ب) ،(٣٣-أ) ، وكمموه حل مكان المموهات الأقل جاذبية في فقرات الاختبار المتبقية وهي : (٢-د) ،(٤-ب) ،(٦-أ) ،(٨-ج) ،(١٠-ج) ،(١٢-أ) ،(١٤-د) ،(١٦-د) ،(١٨-أ) ،(٢٠-د) ،(٢٢-أ) ،(٢٤-ب)

،(٢٦-د) ،(٢٨-د) ،(٣٠-د) ،(٣٢-ب) ،(٣٤-د) ، ، وقد تمّ توزيع هذا البديل بصورة عشوائية في نموذج الاختبار .

المعالجة الإحصائية:

على ضوء تساؤلات الدراسة استخدم الباحث في الدراسة الحالية مجموعة من الأساليب الإحصائية لمعالجة بيانات الدراسة وهي كما يلي :

١- حساب معامل الصعوبة لكل فقرة من فقرات نموذجي الاختبار ، وذلك بإيجاد نسبة من أجاب عن الفقرة إجابة صحيحة ممن حاولوا الإجابة عليها.

٢- حساب معامل التمييز لكل فقرة من فقرات نموذجي الاختبار ، وذلك بإيجاد معامل الارتباط بين نتائج المفحوصين على هذه الفقرة ونتائجهم على الاختبار الكلي.

٣- حساب معاملي ثبات الاتساق الداخلي لنموذجي الاختبار ، وذلك باستخدام معادلة كودر- ريتشاردسون (KR-٢٠).

٤- حساب معاملي صدق المحك التلازمي ، وذلك بحساب معامل ارتباط بيرسون بين درجات المفحوصين على الاختبار ودرجاتهم في مادة الرياضيات خلال الفصل الدراسي الأول ١٤٣٩/١٤٤٠هـ.

٥- استخدام اختبار t لفحص الفروق بين قيم متوسطات معاملات الصعوبة وقيم متوسطات معاملات التمييز لفقرات نموذجي الاختبار.

٦- استخدام الإحصائي (FeldtF,1969) لفحص مدى الاختلاف في معاملي ثبات الاتساق الداخلي لعينتين مستقلتين لنموذجي الاختبار.

٧- استخدام الإحصائي (Z-Fisher) للكشف عن الفروق الثنائية بين معاملي صدق المحك التلازمي لنموذجي الاختبار.

نتائج الدراسة وتفسيرها

في هذا الفصل يتناول الباحث عرض نتائج التحليلات الإحصائية التي توصل إليها. وتفسير هذه النتائج من خلال التحقق من صحة تساؤلات الدراسة وذلك على النحو التالي :

التساؤل الأول :

هل توجد فروق ذات دلالة إحصائية بين متوسطات قيم معاملات الصعوبة في نموذجي الاختبار التحصيلي لمادة الرياضيات من نوع الاختيار من متعدد تعزى إلى وجود البديل "جميع ما سبق" أو عدم وجوده ؟

للتحقق من التساؤل الأول تمّ حساب معاملات الصعوبة للفقرات في نموذجي الاختبار وذلك بإيجاد نسبة الطلبة الذين أجابوا عن الفقرة إجابة صحيحة من بين المفحوصين الذين حاولوا الإجابة عن هذه الفقرة، ويبين الجدول الإحصائي رقم (٢) معاملات الصعوبة لكل فقرة من فقرات النموذجين :

جدول (٢) : معاملات صعوبة الفقرات في نموذجي الاختبار

نموذج الاختبار	نموذج لا تتضمن بدائل فقراته البديل "جميع ما سبق"			نموذج تتضمن بدائل فقراته البديل "جميع ما سبق"		
رقم الفقرة	معامل الصعوبة	رقم الفقرة	معامل الصعوبة	رقم الفقرة	معامل الصعوبة	رقم الفقرة	معامل الصعوبة
١	٠,٦٦	١٨	٠,٥٠	١	٠,٣٤	١٨	٠,٤٩
٢	٠,٤٢	١٩	٠,٥٠	٢	٠,٤٩	١٩	٠,١٦
٣	٠,٤٣	٢٠	٠,٤٨	٣	٠,٥٩	٢٠	٠,٢٨
٤	٠,٧٠	٢١	٠,٣٨	٤	٠,٤٤	٢١	٠,١٦
٥	٠,٤٧	٢٢	٠,٤٥	٥	٠,٥١	٢٢	٠,٥٠
٦	٠,٥١	٢٣	٠,٣٥	٦	٠,٣١	٢٣	٠,١٨
٧	٠,٣٨	٢٤	٠,٥٠	٧	٠,٤٥	٢٤	٠,٥٠
٨	٠,٦٤	٢٥	٠,٥٨	٨	٠,٣٥	٢٥	٠,٥٦
٩	٠,٥٩	٢٦	٠,٤٨	٩	٠,٦٢	٢٦	٠,٢٧
١٠	٠,٤٠	٢٧	٠,٣٧	١٠	٠,٤٥	٢٧	٠,٦٢
١١	٠,٥٨	٢٨	٠,٤١	١١	٠,٥٥	٢٨	٠,٢٠
١٢	٠,٤٥	٢٩	٠,٥٠	١٢	٠,٢٠	٢٩	٠,٣٦
١٣	٠,٥٧	٣٠	٠,٤٠	١٣	٠,٦٣	٣٠	٠,٤٦
١٤	٠,٤٠	٣١	٠,٣٦	١٤	٠,٢١	٣١	٠,١٣
١٥	٠,٤٣	٣٢	٠,٣٦	١٥	٠,٤٥	٣٢	٠,٤٣
١٦	٠,٤٣	٣٣	٠,٦١	١٦	٠,١٩	٣٣	٠,٦٢

يتضح من الجدول السابق أن قيم معاملات الصعوبة لفقرات النموذج الذي لا تتضمن بدائل فقراته البديل "جميع ما سبق" قد تراوحت ما بين (٠,٣٠ – ٠,٧٠) ، بينما تراوحت قيم معاملات الصعوبة لفقرات النموذج الذي تضمنت بدائل فقراته البديل" لا شيء مما ذكر" ما بين (٠,١٣ – ٠,٦٣).

وللكشف عن الفروق في معامل صعوبة الفقرات تبعاً لنموذجي الاختبار ، تم إجراء الاختبار الإحصائي (T-test) لاختبار دلالة الفروق بين متوسطات صعوبة فقرات نموذج الاختبار الذي لا تتضمن بدائل فقراته البديل "جميع ما سبق" ، والنموذج الذي تتضمن بدائل فقراته البديل "جميع ما سبق" كما يتضح في جدول (٣) التالي :

جدول (٣) : نتائج اختبار (ت) لفحص الفروق بين متوسطي معاملات صعوبة فقرات نموذجي الاختبار

م	العينة	العدد	المتوسط الحسابي	الانحراف المعياري	درجة الحرية	قيمة اختبار ت	الدلالة
١	نموذج ١	٣٤	٠,٤٧	٠,١٠	٦٦	٢,٣٧	٠,٠٢
٢	نموذج ٢	٣٤	٠,٣٩	٠,١٥			

نموذج(١): لا تتضمن بدائل فقراته البديل "جميع ما سبق"

نموذج(٢): تتضمن بدائل فقراته البديل "جميع ما سبق"

من جدول (٣) السابق يتضح وجود فروق دالة إحصائياً بين قيم معاملات صعوبة فقرات الاختيار من متعدد تعزى إلى نموذجي الاختبار ، حيث بلغت قيمة " ت " (٢,٣٧) باحتمال يساوي (٠,٠٢) وهي قيمة دالة إحصائياً عند مستوى الدلالة الإحصائية (٠,٠٥) مما يدل على أنه :

توجد فروق ذات دلالة إحصائية بين متوسطات قيم معاملات الصعوبة في نماذج الاختبار التحصيلي لمادة الرياضيات من نوع الاختيار من متعدد تعزى إلى وجود البديل "جميع ما سبق" أو عدم وجوده ، لصالح النموذج الذي تتضمن بدائل فقراته البديل "جميع ما سبق".

كما يتضح من قيم متوسطات معاملات صعوبة الفقرات أن متوسط معامل صعوبة النموذج الذي لا تتضمن بدائل فقراته البديل "جميع ما سبق" (٠,٤٧) أعلى من وسط معامل صعوبة النموذج الذي تضمنت بدائل فقراته البديل "جميع ما سبق" (٠,٣٩) .

ويمكن تفسير نتيجة التساؤل الأول على النحو التالي :

أظهرت نتائج اختبار (ت) لفحص الفروق بين متوسطي صعوبة فقرات نموذجي الاختبار إلى وجود فروق ذات دلالة إحصائية عند مستوى (٠,٠٥=) تعزى إلى نموذجي الاختبار ، ويتضح ذلك من خلال تباعد توزيع قيم معاملات صعوبة الفقرات إلى حدٍ ما في هذين النموذجين.

كما لاحظ الباحث أنّ عدد الفقرات التي تقل عن معامل الصعوبة (٠,٥) في النموذج الذي لا تتضمن بدائل فقراته البديل "جميع ما سبق"(٢١) فقرة ، و (٢٤) فقرة في النموذج الآخر ، أما عدد الفقرات التي معامل صعوبتها (٠,٥) فأكثر فكانت (١٣) فقرة في النموذج

الذي لا تتضمن بدائل فقراته البديل "جميع ما سبق"، و (١٠) فقرة في النموذج الآخر ، وهذا التقارب يعود للتشابه الكبير بين بدائل هذين النموذجين من الفقرات فالاختلاف بين النموذجين فقط في تضمين البديل "جميع ما سبق" من بين بدائل الفقرة ، فالجواب الصحيح موجود تارةً كواحد من بدائل الفقرة في كلا النموذجين وأحيانا موجود على شكل "جميع ما سبق" ، كما أنّ المموهات هي نفسها والاختلاف الوحيد هو وجود البديل "جميع ما سبق" كواحد من المموهات ، الأمر الذي يجعل إمكانية إجابات المفحوصين عن الفقرات المتناظرة في النموذجين متقاربة إلى حد كبير.

واتفقت هذه النتيجة مع نتائج عدد من الدراسات السابقة كدراسة كل من أوسترهوف وكوتس (Coats,1984 & Oosterhof) ، توليفسون (Tollefson,1987) ، كرهان وآخرون (et al.,1989 Crehan) ، كريهان وهالادينا وبريور(Haladyna , Crehan & Brewer,1993) ، يعقوب (١٩٩٦) ، كوين و أودقارد (Odegard & Koen, ٢٠٠٧) ، (DiBattista, etal, ٢٠١٤) ، (Pachai , etal, ٢٠١٥) التي أشارت إلى أنّ البديل "جميع ما سبق" يزيد من صعوبة الفقرات خاصة إذا كان البديل هو الجواب الصحيح ، وأن الفرق بين صعوبة الفقرات التي فيها البديل "جميع ما سبق" كجواب صحيح والفقرات التي لا تحتوي هذا البديل كانت دالة إحصائياً عند مستوى (٠,٠٥=) ، ولكنها تعارضت مع نتائج دراسات أخرى كدراسة وليامسون وهوبكنز (Hopkins,1967 & Williamson) ، فورسيث وسبرات (Spratt,1980 & Forsyth) ، أبو فودة (٢٠٠٣).

التساؤل الثاني :

هل توجد فروق ذات دلالة إحصائية بين متوسطات قيم معاملات التمييز في نموذجي الاختبار التحصيلي لمادة الرياضيات من نوع الاختيار من متعدد تعزى إلى وجود البديل "جميع ما سبق" أو عدم وجوده ؟

للتحقق من التساؤل الثاني تم حساب معاملات التمييز لكل فقرة في نموذجي الاختبار وذلك بإيجاد معامل الارتباط بين نتائج المفحوصين على كل فقرة من فقرات الاختبار ونتائجهم على الاختبار الكلي ، كما تم حساب القيم الوسيطية لهذه المعاملات ويبين الجدول (٤) معامل التمييز لكل فقرة من فقرات نموذجي الاختبار.

جدول (٤) :معاملات تمييز الفقرات في نموذجي الاختبار

نموذج الاختبار	نموذج لا تتضمن بدائل فقراته البديل "جميع ما سبق"			نموذج تتضمن بدائل فقراته البديل "جميع ما سبق"			
رقم الفقرة	معامل التمييز	رقم الفقرة	معامل التمييز	رقم الفقرة	معامل التمييز	رقم الفقرة	معامل التمييز
١	0.45	١٨	0.50	١	0.32	١٨	٠,٤٢
٢	0.41	١٩	0.55	٢	0.40	١٩	٠,٢٩
٣	0.32	٢٠	0.27	٣	0.24	٢٠	٠,٤٩
٤	0.47	٢١	0.44	٤	0.59	٢١	٠,٢٤
٥	0.45	٢٢	0.40	٥	0.54	٢٢	٠,٤٧
٦	0.39	٢٣	0.38	٦	0.56	٢٣	٠,٣٠
٧	0.40	٢٤	0.52	٧	0.31	٢٤	٠,٤٩
٨	0.53	٢٥	0.40	٨	0.49	٢٥	٠,٤٠
٩	0.60	٢٦	0.44	٩	0.51	٢٦	٠,٤٧
١٠	0.34	٢٧	0.19	١٠	0.19	٢٧	٠,٤٢
١١	0.47	٢٨	0.47	١١	0.39	٢٨	٠,٣٧
١٢	0.33	٢٩	0.49	١٢	0.43	٢٩	٠,٥٧
١٣	0.49	٣٠	0.30	١٣	0.47	٣٠	٠,٢٨
١٤	0.42	٣١	0.29	١٤	0.42	٣١	٠,١٩
١٥	0.45	٣٢	0.30	١٥	0.19	٣٢	٠,٣١
١٦	0.27	٣٣	0.54	١٦	0.33	٣٣	٠,٤٠
١٧	0.19	٣٤	0.49	١٧	0.47	٣٤	٠,٥٠

يتضح من خلال الجدول السابق أنّ قيم معاملات التمييز لفقرات النموذج الذي لا تتضمن بدائل فقراته البديل "جميع ما سبق" قد تراوحت ما بين (٠,١٩-٠,٦٠) ، بينما تراوحت قيم معاملات التمييز لفقرات النموذج الذي تضمنت بدائل فقراته البديل "جميع ما سبق" ما بين (٠,١٩-٠,٥٩).

وللكشف عن الفروق في معامل تمييز الفقرات تبعاً لنموذجي الاختبار ، تم إجراء الاختبار الإحصائي (T-test) لاختبار دلالة الفروق بين متوسطات تمييز فقرات نموذج الاختبار الذي لا تتضمن بدائل فقراته البديل "جميع ما سبق" ، والنموذج الذي تتضمن بدائل فقراته البديل "جميع ما سبق" كما يتضح في جدول (٥) التالي :

جدول (٥) : نتائج اختبار (ت) لفحص الفروق بين متوسطي تمييز فقرات نموذجي الاختبار

م	العينة	العدد	المتوسط الحسابي	الانحراف المعياري	درجة الحرية	قيمة اختبار ت	الدلالة
١	نموذج ١	٣٤	٠,٤١	٠,١٠	٦٦	٠,٥٥	٠,٥٨
٢	نموذج ٢	٣٤	٠,٤٠	٠,١١			

نموذج(١): لا تتضمن بدائل فقراته البديل "جميع ما سبق"

نموذج(٢): تتضمن بدائل فقراته البديل "جميع ما سبق"

من جدول (٥) السابق يتضح عدم وجود فروق دالة إحصائياً بين قيم معاملات تمييز فقرات الاختيار من متعدد تعزى إلى نموذجي الاختبار ، حيث بلغت قيمة " ت " (٠,٥٥) باحتمال يساوي (٠,٥٨)وهي قيمة غير دالة إحصائياً عند مستوى الدلالة الإحصائية (٠,٠٥) مما يدل على أنه :

لا توجد فروق ذات دلالة إحصائية بين متوسطات قيم معاملات التمييز في نموذجي الاختبار التحصيلي لمادة الرياضيات من نوع الاختيار من متعدد تعزى إلى وجود البديل "جميع ما سبق" أو عدم وجوده ؟

كما يتضح من قيم متوسطات معاملات تمييز الفقرات أنها تساوت في فقرات النموذج الذي لا تتضمن بدائل فقراته البديل "جميع ما سبق"(٠,٤١) ، وفي فقرات النموذج الذي تضمنت بدائل فقراته البديل "جميع ما سبق" (٠,٤٠) .

ويمكن تفسير نتيجة التساؤل الثاني على النحو التالي :

أظهرت نتائج اختبار (ت) لفحص الفروق بين متوسطي تمييز فقرات نموذجي الاختبار إلى عدم وجود فروق ذات دلالة إحصائية عند مستوى (٠,٠٥=) تعزى إلى نموذجي الاختبار ، ويتضح ذلك من تقارب توزيع قيم معاملات تمييز الفقرات في هذين النموذجين.

كما لاحظ الباحث أنّ عدد الفقرات التي قل معامل تمييزها عن (٠,٣) (الحد الأدنى المفضل في قيم تمييز الفقرات) كان (٥) فقرات في النموذج الذي لا تتضمن بدائل فقراته البديل "جميع ما سبق"، و(٧) فقرات في النموذج الآخر.

كذلك لاحظ الباحث أنّ عدد الفقرات التي كانت معاملات تمييزها ما بين (٠,٣) و (٠,٤٩) كان متقارباً في النموذجي ، أما عدد الفقرات التي كانت معاملات تمييزها (٠,٥) فأكثر فكانت (٦) فقرات في النموذج الذي لم تتضمن بدائل فقراته البديل "جميع ما سبق"، و

(٦) فقرات في النموذج الآخر ، وتتفق هذه النتيجة مع دراسة كل توليفسون (Tollefson,1987) ، ودراسة كرهان وآخرون (Crehan et al.,1989) ، ودراسة فراري (Frary,1991) ، ودراسة كريهان وهالادينا وبريور (Crehan,Haladyna & Brewer,1993).

التساؤل الثالث :

هل توجد فروق ذات دلالة إحصائية بين قيم معاملات صدق المحك التلازمي في نتائج نموذجي الاختبار التحصيلي لمادة الرياضيات من نوع الاختيار من متعدد تعزى إلى وجود البديل "جميع ما سبق" أو عدم وجوده ؟

للتحقق من التساؤل الثالث تمّ حساب معامل صدق المحك التلازمي لنموذجي اختبار الاختيار من متعدد وذلك بحساب معامل ارتباط بيرسون بين درجات الطلاب على النموذجين ودرجاتهم النهائية في مادة الرياضيات المقررة في الفصل الدراسي الأول من العام الدراسي ١٤٣٩هـ – ١٤٤٠هـ ، حيث بلغت قيمة معامل صدق المحك التلازمي للنموذج الذي لا تتضمن بدائل فقراته البديل "جميع ما سبق" (٠,٩٧٠) ، أما للنموذج الذي تضمنت بدائل فقراته البديل "جميع ما سبق" فكانت تساوي (٠,٩٦٢).

وللكشف عن دلالة الفروق تمّ استخدام الإحصائي (Fisher-Z) لفحص مدى اختلاف العلاقة الارتباطية لمجموعتين مستقلتين ويبين الجدول رقم (٦) نتائج التحليل:

جدول (٦) : نتائج المقارنة الثنائية بين معاملي صدق المحك التلازمي لنموذجي الاختبار

النموذج	معامل الصدق	قيمة Z المحسوبة	قيمة Z الحرجة
لا تتضمن بدائل فقراته البديل "جميع ما سبق"	٠,٩٧٠		
تتضمن بدائل فقراته البديل "جميع ما سبق"	٠,٩٦٢	١,١٠	١,١٦

من الجدول السابق يتضح أن قيمة "F" تساوي (١,١٠) ، وتشير إلى عدم وجود فروق ذات دلالة إحصائية عند مستوى دلالة (٠,٠٥) بين معاملي صدق المحك التلازمي للاختبارين تعزى إلى نموذجي الاختبار مما يدل على أنه :

لا توجد فروق ذات دلالة إحصائية بين قيم معاملات صدق المحك التلازمي في نتائج نموذجي الاختبار التحصيلي لمادة الرياضيات من نوع الاختيار من متعدد تعزى إلى وجود البديل "جميع ما سبق" أو عدم وجوده ؟

ويمكن تفسير نتيجة التساؤل الثالث على النحو التالي :

لكشف العلاقة الارتباطية بين أداء الطلبة على نموذج الاختبار الذي قام الباحث بإعداده ودرجة الطالب النهائية في مادة الرياضيات للفصل الأول ، فقد تمّ حساب معاملي الارتباط لنموذجي الاختبار ، فكانت معاملات الارتباط بين أداء الطلاب على الاختبار الذي أعده الباحث ودرجة الرياضيات النهائية عالية وموجبة ؛ ويرجع الباحث السبب إلى أنه قد يعود إلى قوة العلاقة بين الاختبار الذي أعده الباحث والمحك المستخدم ، فكليهما يقيسان نفس الهدف ونفس السمة لدى جميع الطلاب في نموذجي الاختبار.

كما أظهرت نتائج التحليل الإحصائي للكشف عن دلالة الفروق بين معاملي صدق المحك التلازمي لنموذجي الاختبار (وجود البديل "جميع ما سبق" أو عدم وجوده) باستخدام الإحصائي(Fisher-Z) إلى عدم وجود فروق دالة إحصائياً بين قيم معاملات صدق المحك التلازمي تعزى إلى نموذجي الاختبار ، ويلحظ ذلك في عدم وجود فروق بين معاملي الصدق لنموذجي الاختبار ، وهذه النتيجة تتفق مع نتائج دراسة كل من وليامسون وهوبكنز (Hopkins,1967 & Williamson) ، ودراسة فورسيث وسبرات (Forsyth & Spratt,1980) ، ودراسة يعقوب (١٩٩٦) ، ودراسة أبو فودة (٢٠٠٣).

التساؤل الرابع :

هل توجد فروق ذات دلالة إحصائية بين قيم معاملات ثبات الاتساق الداخلي في نتائج نموذجي الاختبار التحصيلي لمادة الرياضيات من نوع الاختيار من متعدد تعزى إلى وجود البديل "جميع ما سبق" أو عدم وجوده ؟

للتحقق من التساؤل الرابع تمّ حساب معامل الثبات لنموذجي اختبار الاختيار من متعدد بطريقة كودر- ريتشاردسون (KR-20) فكان للنموذج الذي لا تتضمن بدائل فقراته البديل "جميع ما سبق" (٠,٨٤٤) ، أما للنموذج الذي تضمنت بدائل فقراته البديل "جميع ما سبق" فكان (٠,٨٣٨).

وللكشف عن دلالة الفروق تم استخدام الإحصائي (F) (Feldet,١٩٦٩) للكشف عن دلالة الفروق بين معاملي الثبات لعينتين مستقلين، ويبين الجدول رقم (٧) نتائج التحليل :

النموذج	(KR-20)	قيمة (F) المحسوبة	قيمة (F) الحرجة
لا تتضمن بدائل فقراته البديل "جميع ما سبق"	٠,٨٤٣	١,٠١٨	١,٢١
تتضمن بدائل فقراته البديل "جميع ما سبق"	٠,٨٣٩		

جدول (٧) : نتائج الإحصائي(F) لفحص مدى الاختلاف في معاملي ثبات الاتساق الداخلي لنموذجي الاختبار

من الجدول السابق يتضح أن قيمة "F" تساوي (١,٠١٨) ، وتشير إلى عدم وجود فروق دالة إحصائياً عند مستوى دلالة (٠,٠٥) بين معاملي ثبات الاتساق الداخلي للاختبارين تعزى إلى نموذجي الاختبار مما يدل على أنه :

لا توجد فروق ذات دلالة إحصائية بين قيم معاملات ثبات الإتساق الداخلي في نتائج نموذجي الاختبار التحصيلي لمادة الرياضيات من نوع الاختيار من متعدد تعزى إلى وجود البديل "جميع ما سبق" أو عدم وجوده ؟

ويمكن تفسير نتيجة التساؤل الرابع على النحو التالي :

أظهرت نتائج التحليل الإحصائي للكشف عن دلالة الفروق في معاملي ثبات الاتساق الداخلي بين نموذجي الاختبار (وجود البديل "جميع ما سبق" أو عدم وجوده) إلى عدم وجود فروق جوهرية بين معاملي ثبات الاتساق الداخلي تعزى إلى نموذجي الاختبار.

ويُرجع الباحث السبب في عدم وجود الفروق بين معاملي الثبات إلى تطابق نموذجي الاختبار مع الاختلاف في تضمين البديل "جميع ما سبق" الأمر الذي جعل تباين الأداء على فقرات النموذجين متقارباً إلى حد كبير.

كما أن الاختبارين يتمتعان بمعاملات تمييز عالية إذ أنّ الاختبار الأكثر تمييزاً أكثر ثباتاً ، كما يلاحظ من الجدول (٦) تجمع عدد كبير من الفقرات الأعلى تمييزاً والتي تسهم بصورة فعالة من الفقرات الأقل تمييزاً في زيادة التباين الحقيقي في نموذجي الاختبار ، وهذا بدوره يرفع من قيمة معامل ثبات الاتساق الداخلي في النموذجين الأمر الذي أدى إلى تقارب معاملي ثبات الاتساق الداخلي في النموذجين بشكل غير دال إحصائياً ، وهذه النتيجة تتفق مع دراسة كريهان وهالادينا وبريور (Brewer,1993 & Crehan,Haladyna) ، ودراسة

فورسيث وسبرات (Spratt,1980 & Forsyth) ، ودراسة وليامسون وهوبكنز (Hopkins,1967 & Williamson)، وتناقضت مع نتائج دراسة كل من أوسترهوف وكوتس (Coats,1984 & Oosterhof) ، ودراسة يعقوب (١٩٩٦).

ثانياً : التوصيات :

توصي الدراسة بما يلي :

١- استخدام الاختبار التحصيلي لمقرر الرياضيات للصف الثالث متوسط كأداة لقياس تحصيل الطلاب على مستوى إدارات التعليم بالمناطق الأخرى .

٢- يُفضل عدم استخدام البديل "جميع ما سبق" عند كتابة فقرات من نوع الاختيار من متعدد وذلك لأنه يزيد من صعوبة الاختبار ، على الرغم من أنه يمكن استخدام هذا البديل في حين كون مصمم الاختبار يُريد أن يزيد من صعوبة الاختبار.

٣- استخدام البديل "جميع ما سبق" له دور في دقة الاجابة في اختبارات الرياضيات ، حيث يفصل بين من يعرف الاجابة الصحيحة ومن لا يعرفها.

ثالثاً : الدراسات والبحوث المقترحة :

١- بناء اختبارات تحصيلية للمقارنة بين الخصائص السيكومترية للاختبار (الصدق والثبات) وفقراته (الصعوبة والتمييز) والتي تتناول استخدام البديل "جميع ما سبق" في فقرات الاختيار من متعدد لمواد أخرى ولدى مراحل تعليمية مختلفة.

٢-إجراء دراسات تتناول العوامل الاخرى المؤثرة على الخصائص السيكومترية للاختبار وفقراته باختلاف شكل البديل كالمستويات المعرفية التي تقيسها الفقرة وفق تصنيف بلوم.

٣- إجراء دراسات للمقارنة بين اختبار الاختيار من متعدد الذي يستخدم البديل "جميع ما سبق" مع أشكال أخرى للاختبارات من حيث الخصائص السيكومترية للاختبار وفقراته.

المراجع

أبو فودة،باسل خميس سالم(٢٠٠٣.(*أثر تضمين البديل*" جميع ما سبق "*الفقرة الاختبار لمفحوصين من مستويات تحصيلية مختلفة على خصائص الفقرات والاختبار*، رسالة ماجستير غير منشورة، جامعة اليرموك، الأردن.

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

عودة،أحمد(٢٠١٠).القياس والتقويم في العملية التدريسية.عمان، اربد:دار الأمل للنشر والتوزيع.

مراد،صلاح احمد؛سليمان،أمين احمد(٢٠٠٥).الاختبارات والمقاييس في العلوم النفسية والتربوية.ط٢.القاهرة:دار الكتاب الحديث.

يعقوب،إبراهيم محمد(١٩٩٦). أثر اختلاف عدد البدائل والبديل "لا شيء مما ذكر" على الخصائص السيكومترية لفقرات الاختيار من متعدد ، المجلة العربية للتربية ، ٤ (١٦)،٩١-١١٧.

Aiken, L. (2003 .(*Psychological Testing and Assessment* .Boston: Pearson Education Group.

Carey, L (1994) Measuring and Evaluating School Learning . Bouston : Allyn and Bacon .

Crehan,K.& Haladyna,Th.(1991).The Validity of Two Items Writing Rules. *Journal of experimental education*.59, 183-192.

Crehan,K., & Haladyna,Th., & Brewer.B.(1993).Use of an Inclusive Option and the Optimal Number of Options for Multiple Choice Items .*Educational and psychological measurement*.53, 241-247.

Crehan,Kevin & Others.(1989).*An investigation of the Optimal Number of Options and the Inclusive " None of the These "* Option.ED:314500.

Crocker, L. & Algina, J. (1986). Introduction and Modern Test Theory. Holt Rinehart and Winston, New York: *Education Group*.

DiBattista, D., Sinnige-Egger, J.-A., & Fortuna, G. (2014). The "None of the Above" Option in Multiple-Choice Testing: An Experimental Study. *The Journal of Experimental Education*, 82, 168–183.

Downing,S.M, (2002).Construct – Irrelevant Variance and Flawed Test Questions: Do Multiple-Choice Item Writing Principles Make Any Difference? *Academic Medicine*, 77(10), 103-104.

Eble,R.(1972).*Essentials of Educational Measurement* .New jersey: prentice-Hall,inc.

Feldet,l,.S.(1969).A Test of the Hypothesis that Cronbachs Alpha or Kuder-Richardson Coefficient twenty is the Same for Two test .*Psychometrika*. 34,363-373.

Forsyth, R.A. & Spratt, K.F. (1980). Measuring Problem Solving Ability in Mathematics with Multiple Choice Items: The Effect of Item Format on Selected Item Test Characteristics.*Journal of education measurement* .1, 31 – 43.
Forsyth, R.A. & Spratt, K.F. (1980). Measuring Problem Solving Ability in Mathematics with Multiple Choice Items: The Effect of Item Format on Selected Item Test Characteristics.*Journal of education measurement* .1, 31 – 43.
Frary,R.B.(1991).The None of the Above Option: an Empirical Study. *Applied measurement in education*,2, 115-124.

McMillan, J. H. (2004). *Classroom Assessment: principles and for Effective Instruction* (3rd ed.). United states Of america.

Odegard, T. N., & Koen, J. D. (2007). *"None of the Above" as a Correct and Incorrect Alternative on a Multiple-Choice test: Implications for the Testing Effect*. Memory, 15, 873–885.

Osterhof,A.C., & Coats, P.K.(1984).Comparison of Difficulties and Reliabilities of Quantitative Word Problems in Completion Multiple-Choice Items Format. Applied psychological measurement. 8,287 – 294.

Pachai, M. V., DiBattista, D., & Kim, J. A. (2015). A Systematic Assessment of 'None of the Above' on Multiple Choice Tests in a First Year Psychology Classroom. *The Canadian Journal for the Scholarship of Teaching and Learning*, 6 (3).

Popham, W.j. (2000). *Modern Educational Measuerment: Practical Guidelines for Educational Leader* (3rd ed.). U.S. of Amarica.

Roid,G.,& Haladyna,T.(1982).*Atechnology for Test Item Writing*.New york:Academic press,49-55.

Tollefson,N.(1987).Acomparison of Item Difficult and Item Discrimination of Multiple - Choice Items Using The " None of the Above"and One Correct Response Options.*Educational and psychological measurement*.47, 377-384.

Williamson,M.L.& Hopkins,k. D.(1967).The use of "none of these" versus homogeneous alternatives on multiple-choice tests: experimental reliability and validity comparisons.*Journal of education measurements* ,4, 53-58.