أثر موقع المشتت القوي باستخدام تقنية إخفاء البدائل الخاطئة في بعض الخصائص السيكومترية لأسئلة الاختيار من متعدد في مادة الفيزياء

م.م. جاسم حسان حميدي
تربية القادسية / طرائق تدريس الفيزياء

مستخلص البحث:

هدف البحث الحالي الى معرفة العلاقة بين قوة تشتيت البدائل في أسئلة الاختيار من متعدد في مادة الفيزياء وصدق الاختبار وثباته وتأتي أهمية هذا البحث وسط اغفال متزايد من الكثيرين للضوابط الفنية والمنهجية في اعداد وتقنين الاختبارات التحصيلية وعدم التقيد بضرورة اجراء المعالجات التحليلية لفقرات اختبارات الاختيار من متعدد. شمل البحث الحالي (105) طالباً من طلاب الصف الاول المتوسط في ثانوية المتميزين للعام الدراسي 2017-2018 ،وتمثلت أدوات البحث في اختبار الاختيار من متعدد الذي بلغ عدد فقراته (18) فقرة، وقام الباحث بتطبيق البحث الحالي على عينة البحث خلال الكورس الاول والثاني للعام الدراسي 2017-2018وعلى النحو الاتي: استخراج معاملات الصعوبة وتمييز الفقرات ، وقوة تشتيت البدائل ، والانحراف المعياري للدرجات ، ومعامل ثبات الاختبار. وفي ضوء نتائج البحث الحالي توصل البحث الحالي الى النتائج الاتية :

ان صياغة الفقرات لاسئلة الاختيار من متعدد له تأثير على معاملات تمييزها وصعوبتها, وفي ضوء نتائج البحث الحالي يوصي الباحث بضرورة التقيد بالتعليمات المتوافرة في الادب التربوي التي تتضمن الشروط العلمية لتصميم أسئلة الاختيار من متعدد لكي يتم تجنب الكثير من نقاط الضعف وأهمها التخمين.

مشكلة البحث:

تتكون أسئلة الاختيار من متعدد من سؤال أو مشكلة رئيسة تمثل الجذر stem وعدد من البدائل المطروحة لاستثارة تفكير الطالب بحيث تشمل هؤلاء البدائل على الاجابة الصحيحة وثلاثة أو أربعة مشتتات distracters أو بدائل مشتتة ، وان تصميم هذه البدائل وصياغتها يؤثر مباشرةً على إختيار الاجابة الصحيحة ومن ثم على معامل تمييز الفقرة وصدقها وبالتالي على صدق وثبات الاختبار ككل, وان المشكلة الرئيسة في إعداد فقرات أسئلة الاختيار من متعدد تتمثل في إختيار أو إنتقاء وصياغة عدد من البدائل التي تستثير حيرة الطالب وتفكيره بحيث تبدو هذه البدائل المشتتة مرتبطة منطقياً بالجذر لكنها لاتمثل الاجابة الصحيحة فيمكن لمن يعرف أن يميز بينها وبين الاجابة الصحيحة بينما تختلط الاجابة على من لا يميز بينها, والواقع أنه يصعب على المدرس أن يتنبأ بقوة المفردة أو الفقرة على التشتيت ما لم يقوم بإجراء تحليل للمفردات أو الفقرات ، ويتضمن تحليل مفردات الاختبار (فقراته) دراسة لنمط الاستجابة لكل مفردة لتقدير فاعليتها، مما يسمح بتفسير أدق للنتائج ومن ثم إصدار أحكام تقويمية أكثر فعالية فضلاً عن أن تحليل المفردات على هذا النحو يعد تغذية راجعة لواضع الاختبار مما يفيد في تطويره مستقبلاً, ويتم تحليل الفقرات في أسئلة الاختيار من متعدد من ثلاثة جوانب هي (مدى صعوبة الفقرة, مدى قوتها التمييزية أو معامل تمييزها, مدى قوة تشتيت بدائلها), وبالتأكيد تختلف إجراءات تحليل الفقرات في الاختبارات معيارية المرجع عنها في الاختبارات محكية المرجع، نظراً لاختلاف أسس إختيار الفقرات في كل منها وترتبط هذه الجوانب الثلاثة ببعضها البعض في أنها تؤثر على صدق الفقرة على أننا نرى أكثر هذه الجوانب تأثيراً هو الجانب المتعلق بقوة تشتيت البدائل اذ يترتب على ضعف تشتيت البدائل ما يلي:

- وضوح الاجابة الصحيحة مما يضعف حساسيتها لاظهار الفروق الفردية
- زيادة قابلية الفقرة للتخمين عندما تشتمل على البدائل الضعيفة
- إنخفاض قوة التمييز للفقرة ، ومعامل صعوبتها ، ومن ثم صدقها
- إنخفاض معامل ثبات الفقرة بارتفاع درجة قابليتها للتخمين
- إنخفاض صدق وثبات الاختبار ككل

بينما يفترض أن يكون معامل تمييز الاجابة الصحيحة للفقرة موجباً أي ان يكون عدد من يختارها من أفراد الارباعي الاعلى على الاختبار أكبر من عدد من يختارها من أفراد الارباعي الادني على ذات الاختبار فانه يفترض أن يكون قوة تشتيت كل من البدائل المصاحبة للاجابة من أفراد الارباعي الادنى أي ان المشتت الجيد في أسئلة الاختيار من متعدد هو المشتت الذي يختاره عدد أكبر من طلاب الارباعي الادنى اذا ما قورن بعدد من يختاره من طلاب الارباعي الاعلى, وهناك العديد من الاختبارات التي يشيع استعمالها في بيئتنا العربية لاغراض مختلفة دون التقيد باجراء معالجات تحليلية لفقراتها، تتناول معاملات التمييز ، والصعوبة ، وقوة تشتيت البدائل في أسئلة الاختيار من متعدد. ونظراً لان هذا النوع أكثر أنواع الاسئلة الموضوعية شيوعاً واستعمالها في الاختبارات النفسية والتحصيلية المقننة وتلك التي يعدها المدرس ربما لانها اقل قابلية للتخمين واكثر اشتثارة لدافعية الطالب واكثر قابلية لقياس المستويات المعرفية التي قد تعجز عن قياسها الانواع الاخرى من الاسئلة الموضوعية, لكل ذلك فان صياغة اسئلة الاختيار من متعدد تتطلب درجة عالية من الدقة ، واهم محاور هذه الدقة مدى قوة تشتيت البدائل المصاحبة للاجابة الصحيحة ، اذ يرتبط هذا المدى بفاعلية تأثير هذه البدائل على معاملات تمييز الفقرات ومعاملات صعوبتها ومن ثم صدقها. والواقع أن تأثير قوة تشتيت البدائل في أسئلة الاختيار من متعدد لم يحظ بالاهتمام الكافي في الدراسات والبحوث العربية سيان من منظور البحث أو منظور تقنية الاختبارات.

وفي ضوء ما تقدم يطرح البحث الحالي الاسئلة الاتية :

1- ما واقع معاملات صعوبة وتمييز فقرات الاختبار .

2- وما قوة تشتيت البدائل المصاحبة لهذه الفقرات ؟

3- ما مدى تأثير قوة تشتيت البدائل في أسئلة الاختيار من متعدد على معامل صعوبة فقراتها ومعامل تمييزها؟

أهمية البحث :

يكتسب البحث الحالي أهميته على النحو الاتي :

1- يتناول هذا البحث بالتحليل اهم اسس صياغة أسئلة الاختيار من متعدد وهي فاعلية تأثير قوة تشتيت البدائل والتي ترتبط محورياً بمعاملات تمييز وصعوبة وصدق هذا النوع من الاسئلة

2- يكتسب هذا البحث أهميته من قلة الدارسات والبحوث العربية في هذا المجال ، لذا فان الباحث يأمل ان يلفت أنظار الباحثين والمتخصصين في القياس والتقويم الى أهمية الموضوع مدار البحث.

3- يأتي هذا البحث وسط اغفال متزايد من الكثيرين للضوابط الفنية والمنهجية في اعداد وتقنين الاختبارات التحصيلية وعدم التقيد بضرورة اجراء المعالجات التحليلية لفقرات اختبارات الاختيار من متعدد.

أهداف البحث :

يهدف البحث الحالي الى الكشف عن أثر قوة تشتيت البديل القوي في أسئلة الاختيار من متعدد على:

1- معاملات صعوبة الفقرات

2- معاملات تمييز الفقرات

3- الانحراف المعياري لها وللاختبار ككل

مصطلحات البحث:

1- المشتتDisracterهو اي من البدائل غير الصحيحة المصاحبة للاجابة الصحيحة للفقرة في اسئلة الاختيار من متعدد وله قوة تمييز سالبة

2- قوة تشتيت البدائل يقصد بقوة تشتيت البدائل في اسئلة الاختيار من متعدد مدى قدرة البديل على ان يبدو مقبولاً او معقولاً plausible بالنسبة لمن لايعرف الاجابة الصحيحة

3- قوة تشتيت البديل Distractibility : وتقاس بالفرق الجبري بين عدد اختياراته من طلاب الارباعي الادنى وعدد اختياراته من طلاب الارباعي الاعلى (Gay,1981) ، ويعد المشتت جيداً اذا كان عدد من يختاره من طلاب الارباعي الادنى أكبر من عدد من يختاره من طلاب الارباعي الاعلى (Hills,1981,Ebel,1979,Gay,1981).

4- تحليل المشتت Distracter analysis : هو إجراء يقوم على مقارنة عدد اختيارات طلاب الارباعي الادنى بعدد اختيارات طلاب الارباعي الاعلى للمشتت في اسئلة الاختيار من متعدد

5- مؤشر او معامل صعوبة الفقرةDifficulty Index أو :Item Difficulty ويقصد به Ebel (1979) النسبة المئوية للاستجابات غير الصحيحة على الفقرة ، ويقاس بالنسبة المئوية لعدد من أجابوا على الفقرة إجابة خاطئة على العدد الكلي للمستجيبين، بينما يقصد به Gay (1981) النسبة المئوية لعدد الطلاب الذين اجابوا اجابة صحيحة على الفقرة الى العدد الكلي للطلاب الذين استجابوا عليها ، ويتبنى الباحث تعريف Ebel (1979) لكونه يعبر تعبيراً منطقياً عن مؤشر الصعوبة او معاملها.

6- اسئلة الاختيار من متعدد Multiple choice items: وهي نوع من الاسئلة الموضوعية يتكون من سؤال او عبارة غير كاملة تسمى الجذر stem وعدد من البدائل (اربعة الى خمسة) احدها يشكل الاجابة الصحيحة ، والبدائل الاخرى تعمل كمشتتات Distracters ، ويطلب الى الطالب ان يختار الاجابة الصحيحة أو الاجابة الافضل .

ويرى Ebel,1979 ان استخدام مفهوم صعوبة الفقرة على هذا النحو يؤدي الى نوع من الخلط او التداخل في التفسير وهذا التداخل او الخلط ينشأ من عدم منطقية هذه التسمية اذ يترتب على تناوله على هذا النحو ان يصبح مؤشراً للسهولة أكثر منه مؤشراً للصعوبة ويستطرد Ebel مشيراً الى امكانية تجنب هذا الخلط اذا نظرنا الى المفهوم على ان النسبة المئوية لعدد الطلاب الذين اجابوا على الفقرة اجابة خاطئة الى العدد الكلي للطلاب[1].

تشير الدراسات والبحوث الى اهمية تحليل فقرات الاختبار في تدعيم أو عدم تدعيم انماط الصدق المبدئية التي تقوم على صدق المحتوى والصدق الظاهري كما ان هذا التحليل يكشف عن نقاط القوة والضعف في الاختبار ويضع الاساس لكيفية تحسين فقراته فضلاً عن التفسيرات الاصدق للناتج

[1] **اخذ الباحث بما توصل اليه Ebel,1979 في تحديده لمعامل الصعوبة في ضوء الاجابات الخاطئة لا الصحيحة كما ذهب معظم الباحثين**

المتحصل عليه. ويرى Hills,1981 ان الحجم النسبي للانحراف المعياري يرتبط بمعاملات تمييز فقرات الاختبار اذ يتوقف تباين الفقرات المكونة له ، كما يؤثر حجم الانحراف المعياري للاختبار على معامل ثباته من خلال العلاقة بين عدد فقرات الاختبار والانحراف المعياري له ، اذ يرتفع ثبات الاختبار بارتفاع الانحراف المعياري لدرجاته ، وتتوقف مقدار الانحراف المعياري لدرجات الاختبار على معاملات تمييز فقراته .

وتأسيساً على ذلك فان معاملات تمييز فقرات الاختبار تؤثر على معامل ثباته ومن ثم فان تحليل فقرات الاختبار من حيث معاملات الصعوبة والتمييز يمكننا من تحسين صدق الاختبار وثباته وهذا يتفق مع ماذكره Ebel,1979 في دراسته التي استهدفت الكشف عن العلاقة بين معامل تمييز فقرات الاختبار ومعامل ثباته، وقد توصلت الدراسة الى النتائج الاتية: هناك علاقة بين نمط توزيع معاملات صعوبة فقرات الاختبار وانحرافه المعياري ومعامل ثباته فكلما كان مدى معاملات الصعوبة اصغر كان الانحراف المعياري للاختبار ومعامل ثباته اكبر.

خلفية نظرية

يحتل التقويم التربوي موقعاً مهماً في النظام التعليمي، فهو أحد عناصر المنهج الرئيسة الى جانب الاهداف، والمحتوى، وطرائق التدريس ، والانشطة ، والوسائل التعليمية (الوكيل وبشير، 2001).

وللتقويم وظائف مهمة أهمها خدمة عمليتي التعليم والتعلم ، فخدمة التقويم لهما بشكل جيد سيؤدي حتماً الى توجيه تللك العمليات الوجهة الصحيحة ، ومن ثم توجيه السياسة التعليمية بشكل صحيح (Black,1998). والعلاقة بين التقويم وطرائق التدريس وأساليبه واضحة وقوية ، فلن تستطيع طرائق التدريس أن تؤدي الغرض منها مهما كانت فاعليتها وتمركزها حول المتعلم ، إلا إذا كانت هناك معلومات مستمرة وكافية عن المتعلم من حيث تقدمه ، أو فشله في تحقيق الاهداف المرجوة (Wisker,1997). وتأسيساً على ذلك فلابد أن تكون عملية التقويم متكاملة ومتوازية مع عملية التدريس ، ويجب أن يخطط لها بصورة جيدة كما ينبغي أن يكون هناك نوع من التواصل والاستمرارية بين نوع التقويم وطرائق التدريس المستخدمة في أي برنامج دراسي (Burton & Haines,1997). وفي السياق نفسه يشير أمبوسعيدي والصارمي (2002) أن إستثارة تفكير المتعلم من خلال التركيز على مشكلات تتطلب إعمال الفكر ولايمكن تحقيق ذلك إلا اذا رافقها نظام تقويم يدفع المتعلم الى إستخدام قدراته العقلية العليا ويحثه على إعتماد أسلوب حل المشكلات في التعامل مع الخبرات الجديدة. وتعد الاختبارات بأنواعها أكثر الادوات شهرة واستخداماً في مدارسنا ، ويعزى ذلك الى سهولة إستخدامها في ضوء الاعداد المتزايدة للطلبة في الصفوف المدرسية ، ومعرفة المدرسين بها ، وسرعة تصحيحها.

أسئلة الاختيار من متعدد Multiple Choice Questions

تصنف أسئلة الاختيار من متعدد ضمن الأسئلة الموضوعية Objective Questionsوذلك لموضوعية التصحيح فيها ، وبعدها عن تأثير أهواء المصحح وتقلب مزاجه، وقلما تجد إختباراً لا تشكل أسئلة الاختيار من متعدد فيه وزناً كبيراً من أسئلة الاختبار حتى على المستوى الجامعي وذلك لتميزها بالثبات العالي في التصحيح (White & Gunstone,1992). ومن الجدير بالذكر أن من إيجابيات هذا النوع من الاسئلة (Bahar&Hansell,2000):

1. يمكن قياس مجموعة كبيرة من مواضيع المقرر الدراسي في إختبار واحد وفي وقت واحد
2. يمكن إختبار عدد كبير من الطلبة في وقت قصير نسبياً

3. تتميز بثبات مقبول لما يعرف بثبات المحكمين (inter-marker reliability) اذ تتميز بعدم الذاتية في التصحيح
4. يمكن تصحيحها بسرعة وهناك مجال لاستخدام أدوات مساعدة مثل الكمبيوتر
5. تعد من الاسئلة الجيدة لاختبار أولئك الطلبة الذين يعرفون الاجابة الصحيحة ولكنهم لايستطيعون التعبير عنها بسبب ضعف القدرة الكتابية لديهم

أما عن سلبيات أسئلة الاختيار من متعدد فهي كثيرة غير اننا سنركز على نقطتين فقط نرى من الضروري توضيحهما النقطة الاولى : تناقش مشكلة التخمين في أسئلة الاختيار من متعدد فما يزال هذا العامل التي تهدد صدق هذا النوع من الاسئلة وثباته على الرغم من المحاولات الجادة والمتكررة لضبطه بالوسائل التربوية المختلفة (الكحلوت،2002) و(Schawirth et al.,1996).

اما النقطة الثانية : وهي ما يعرف في الادب التربوي بـ " تحيز موقع الاستجابة " (Positional Response Bias) ، وهو يعني إختيار الطلبة لإجابة معينة بسبب موقعها في ترتيب البدائل المعطاة بغض النظر عن محتوى السؤال (Cronbach,1950) فقد يختار بعض الطلبة البديل (a) أكثر من البديل (d) في حالة أن السؤال يتكون من أربعة بدائل، مع إحتمال أن تكون الاجابة الصحيحة هي البديل (d) وهذا يرجع الى أن بعض الطلبة قد لا يقرؤون كل البدائل وإنما يكتفون بقراءة البديلين الاول والثاني فقط خاصة عندما يكتنف الغموض جذر السؤال، أو يصعب على الطالب السؤال عند قراءته للجذر. وهناك نوعان من تحيز موقع الاستجابة : النوع الاول مرتبط بموقع البديل الصحيح بينما النوع الثاني مرتبط بموقع المشتت القوي. فبعض الدراسات السابقة أثبتت فاعلية النوع الاول (موقع البديل الصحيح) على التأثير في إجابات الطلبة، مثل دراسة (Ambusaidi,2000) ، ودراسة (Cizek,1994) ، ودراسة (Bluch,1984) ، أما دراسة (Johnstone& Ambusaidi,2000) (Friel&Johnstone,1979) فقد أثبتت فاعلية موقع المشتت القوي على التأثير في إستجابات الطلبة. ولهذا ظهرت موجة من الانتقادات وجهت الى أسئلة الاختيار من متعدد، إذ يرى منتقدوها أنها تركز على تعليم الطلبة وسائل وحيّل إختيار الجواب الصحيح أكثر من تركيزها على قياس تحصيلهم الدراسي فضلاً عن إفتقارها على إختبار مجموعة من المهارات العقلية العليا من التفكير(Black,1998).

تركيب سؤال الاختيار من متعدد:

يتكون كل سؤال في أسئلة الاختيار من متعدد من جزءين رئيسين هما :

1- الأرومة أو الجذر (Stem): ويكون على شكل سؤال أو عبارة إنشائية تطرح سؤالاً (أحياناً على شكل أكمل...

2- البدائل أو الخيارات (Options or Alternatives): وهي أين تجد الإجابة الصحيحة. والبدائل تقسم الى :

a. الإجابة الصحيحة (the key answer)
b. المموهات أو المشتتات (the distracters)

Example: When a piece of sponge is pressed, which one of the following quantity does not change: ⟸ stem

a. Volume
b. Mass } alternatives or options
c. Density
d. Shape

في المثال أعلاه يعد البديل (b) الاجابة الصحيحة ، في حين البدائل (a,c and d) مموهات أو مشتتات ومن ضمن تلك المموهات أو المشتتات يوجد ما يطلق عليه المشتت القوي (strong distracter) وهو ذلك البديل الذي يظهر في إختيارات الطلبة بقوة إذ يختاره الطلبة أكثر من إختيارهم لبديل الاجابة الصحيحة .

الدراسات السابقة:

هناك العديد من الدراسات السابقة التي تطرقت لموضوع المشتت القوي أو البديل الفعال (strong distracter) وتناولته من زوايا مختلفة ، وقد تنوعت هذه الدراسات بين الدراسات العربية والدراسات الاجنبية، وسنستعرض عديد من هذه الدراسات التي تم الاستفادة منها مع الاشارة الى أبرز ملامحها. مع تقديم تعليقاً عايها يتضمن جوانب الاتفاق والاختلاف مع البحث الحالي ويود الباحث أن يشير الى ان الدراسات السابقة التي سيتم استعراضها جاءت في الفترة الزمنية بين 1983-2012.

هذا وتم تصنيف هذه الدراسات حسب المتغيرات الرئيسة وحسب كونها دراسات عربية أو أجنبية وفيما يلي نقدم عرضاً لهذه الدراسات وجوانب الاتفاق والاختلاف بينها :

الدراسات العربية:

دراسة الزيات(1990):

هدفت هذه الدراسة بالتحليل أهم أسس صياغة أسئلة الاختيار من متعدد وهي فاعلية تأثير قوة تشتيت البدائل والتي ترتبط محورياً بمعاملات تمييز وصعوبة وصدق هذا النوع من الاسئلة. وتكونت عينة الدراسة (414) من طلبة الصفوف الاول والثاني والثالث الثانوي بمدرستي الفلاح الثانوية للبنين بمكة المكرمة وجدة الثانوية للبنات بجدة ممن تتراوح أعمارهم بين (15-20) عاماً. وقد أظهرت نتائج الدراسة أن صياغة مشتتات الاجابة في اسئلة الاختيار من متعدد له تأثير دال على معاملات صعوبة الفقرات وقوة تمييزها وأن فاعلية الفقرة تتوقف على فاعلية مشتتاتها فالمشتتات الضعيفة تؤدي الى وضوح الاجابة الصحيحة ومن ثم تصبح الفقرة غير مميزة بين الذين يعرفون والذين لا يعرفون فتميل الفقرة الى السهولة وتصبح اقل حساسية لقياس الفروق الفردية. (الزيات ،1990: 25)

دراسة أمبوسعيدي والعفيفي (2004):

هدفت هذه الدراسة الى الكشف عن أثر تغيير موقع المشتت القوي في معاملات الصعوبة ومعاملات التمييز لاسئلة الاختيار من متعدد في مادة الفيزياء . تكون مجتمع الدراسة من (3525) طالب وطالبة المنتظمين في الصف الاول الثانوي في مدارس قطاع ولاية السيب في محافظة مسقط بسلطنة عمان في العام الاكاديمي 2002-2003 اما عينة الدراسة فقد بلغت (197) طالب وطالبة تم اختيارهم بطريقة عشوائية من ثلاث مدارس ثانوية. وقد أجاب (98) طالباً وطالبة بنسبة (50%) عن أسئلة النموذج الاول للاختبار ، بينما أجاب (99) طالبا وطالبة بنسبة (50%) عن أسئلة النموذج الثاني.

وقد أظهرت نتائج الدراسة ان تعديلاً بسيطاً من حيث تغيير موقع المشتت القوي أوجد فروقاً دالة إحصائياً للمتوسط الحسابي للاختبار وكذلك في معاملات الصعوبة والتمييز والمتوسطات الحسابية لها. وقد أوصت الدراسة بعدم الاعتماد على أسئلة الاختيار من متعدد بشكل أساس وضرورة البحث عن طرق تقويم أخرى فعالة. (أمبوسعيدي والعفيفي ،2004 : 12)

دراسة أمبوسعيدي(2005):

هدفت هذه الدراسة الى تقصي أثر تغيير موقع المشتت القوي(strong distractor عند استخدام تقنية (هكذا ورد!) إبعاد البدئل الخطأ في أسئلة الاختيار من متعدد على التحصيل الدراسي (المتوسط الحسابي) وعلى بعض الخصائص السايكومترية (معاملات الصعوبة ومعاملات التمييز) لفقرات الاختبار. ولتحقيق أهداف الدراسة تم استخدام اختبار في مادة الفيزياء للصف العاشر من التعليم العام مكوناً من (24) فقرة من نوع الاختيار من متعدد. وتم تصميم نموذجين متشابهين في كل شيئ من هذا الاختبار عدا موقع المشتت القوي بالنسبة الى البديل الصحيح . كما تم استخدام تقنية إبعاد البدائل الخطأ من الورقة الامتحانية في النموذجين. وقد تم التحقق من صدقه بعرضه على مجموعة من المحكمين في تدريس الفيزياء ، كما تم حساب ثباته باستخدام معادلة كيودر-ريتشاردسون(KR-20) التي أعطت قيمة مقدارها (0.76) والتي تعد مناسبة لغرض الدراسة. تكونت عينة الدراسة من (133) طالبة من طالبات الصف العاشر من التعليم العام ، تم اختياهن عشوائياً من مدرستين من مدارس التعليم العام في قطاع ولاية السيب بمحافظة مسقط بسلطنة عمان أواخرالعام الدراسي 2003-2004.

أظهرت نتائج الدراسة عدم وجود فروق ذات دلالة إحصائية بين المتوسطات الحسابية لدرجات الطالبات في نموذجي الاختبار. كما أشارت نتائج الدراسة الى عدم وجود فروق ذات دلالة إحصائية بين المتوسطات الحسابية لمعاملات الصعوبة ، ومعاملات التمييز لنموذجي الاختبار.

وقد أوصت الدراسة بتدريب الطلبة على استخدام تقنية إبعاد البدائل الخطأ عند التعرض لاختبار مكون من أسئلة الاختيار من متعدد ، ويمكن تطبيق ذلك من خلال استخدام الطالب ورقة أوشيئ ما يستطيع به تغطية البديلين الخاطئين تماماً (بحسب الطالب) ، ومن ثم إختيار أحد البدائل المتبقية.

(أمبوسعيدي ،2005 : 67)

دراسة الكناني وفيضي (2012):

هدفت الدراسة الحالية الى بناء مقياس البديل الفعال في أسئلة الاختيار من متعدد لمادة طرائق تدريس الرياضية، ومن خلاله سعى الباحثان الى الكشف عن أثر تغيير موقع البديل الفعال عند استخدام تقنية إبعاد البدائل الخطأ على تحصيل الطلبة في أسئلة الاختيار من متعدد في مادة طرائق تدريس التربية الرياضية. تكون مجتمع الدراسة من طلبة المرحلة الثانية (لم يذكر عدد الطلاب) في كلية التربية الرياضية جامعة المثنى للعام الدراسي (2011-2010) أما عينة الدراسة فاشتملت على (60) طالباً من طلبة المرحلة الثالثة بكلية التربية الرياضية – جامعة المثنى حيث تم تقسيم عينة البحث الى مجموعتين متساويتين حيث أجاب (30) طالباً وطالبة عن أسئلة النموذج الاول كما أجاب (30) طالباً وطالبة عن أسئلة النموذج الثاني. أظهرت نتائج الدراسة عدم وجود فروق معنوية بين نموذجي الاختبار من حيث صعوبة وتمييز الفقرات ، كما أظهرت أن وجود البدائل الاخرى غير البديل الصحيح أمام الطلبة في ورقة الاختبار يؤثر بصرياً عليهم مما يؤثر سلباً في اختيار البديل الصحيح. (الكناني وفيضي ، 2012 : 32)

الدراسات الاجنبية:

دراسة Tollefson and Tripp,1983:

بعنوان" اثر صياغة الفقرات على معاملي صعوبتها وتمييزها " اذ استخدم الباحثان ثلاث صيغ مختلفة لاسئلة الاختيار من متعدد تتناول المفاهيم الاحصائية لمحتوى دراسي ، والصيغ المستخدمة هي :

- لاشئ مما تقدم كاجابة صحيحة
- لاشئ مما تقدم كبديل صحيح
- اجابة واحدة صحيحة

وقد طبقت هذه الصيغ الثلاث على عينة مكونة من (104) طالب تم اختيارهم عشوائياً ، وتشير نتائج تحليل التباين الى وجود فروق ذات دلالة احصائية في متوسطات معاملات تمييز الفقرات لهذه الصيغ الثلاث لمصلحة الصيغة الاولى ، ولم تكن الفروقات في متوسطات الصعوبة ذات دلالة احصائية. (Tollefson and Tripp,1983 214 :)

دراسة Tollefson and Chein,1986:

بعنوان" مقارنة صعوبة الفقرة ومعامل تمييزها عند استخدام لاشئ مما تقدم كبديل في اسئلة الاختيار من متعدد"، وقد استخدم الباحثان نفس الصيغ الثلاث المشار اليها في الدراسة السابقة ، اذ توصلت الى النتائج الاتية:

- انخفاض معاملات صعوبة فقرات الصيغة الاولى في المتوسط عن كل من الصيغتين الثانية والثالثة ، على حين لم تختلف معاملات التمييز اختلافا ملحوظا في الصيغ الثلاث
- ان معاملات صعوبة فقرات الصيغة الثانية كانت اقل في المتوسط بفروق ذات دلالة احصائية من الصيغة الاولى عندما كان المحتوى الدراسي المقاس هو المفاهيم ومعاني المفردات.

وتوصلت الدراسة الى النتائج الاتية :

- ان صياغة الفقرات له تأثير على معاملات تمييزها وصعوبتها
- ان استخدام صيغة " لاشئ مما تقدم " يرفع من معامل صعوبة الفقرة في ظل الصيغة الاولى ويخفض هذا المعامل في ظل الصيغة الثانية.

ويرى عدد من الباحثين(1981) Ebel,1979;Gay,1979;Hills ان الطلاب عند استجابتهم على اسئلة الاختيارمن متعدد يلجئون عادة الى اسلوب استبعاد البدائل غير الصحيحة Eliminate Incorrect Alternativesللوصول الى البديل الصحيح.

ومع ان بعض بناة الاختبارات التحصيلية يرون ان شيوع هذا الاسلوب في الاستجابة على اسئلة الاختيار من متعدد يشير الى ضعف في بناء أو تركيب الفقرات ، الا ان الكثيرين منهم يرون غير ذلك Ebel,1979:150 وخاصة اذا انطوت البدائل على اسس تبدو ظاهرياً متماثلة أو متقاربة اي اذا كانت البدائل المشتتة فعالة ولها قوة تشتيت سالبة ، بمعنى أنها تجتذب عددا اكبر من الذين لا يعرفون ، وعددا اقل من الذين يعرفون الاجابة الصحيحة.

وتأسيساً على ذلك يمكن تقرير ان الاساس في كتابة فقرات أسئلة الاختيار من متعدد ليس هو كثرة عدد البدائل المشتتة وانما فاعليتها ، وقوتها التمييزية . ويؤكد هذا Ebel,1979 فيرى ان ارتفاع معاملات تمييز الفقرات لايتوقف على عدد البدائل المشتتة فقط وانما يتوقف على وجود المشتت الجيدgood distractor كما انه ليس بالضرورة ان يكون عدد البدائل متساوياً في جميع فقرات الاختبار وقد تكون الفقرة جيدة اذا قامت على مشتتين اثنين اي (ثلاث استجابات) وربما مشتت واحد

فقط اذا كان فعالاً اي ان فاعلية الفقرة تتوقف على فاعلية مشتتاتها. (Tollefson and Chein,1986 : 89)

مؤشرات ودلالات من الدراسات السابقة:

عينة الدراسة:

تنوعت عينة البحث من حيث الحجم والجنس في الدراسات السابقة فقد بلغت (414) طالباً وطالبة في دراسة الزيات(1990)، وبلغت (194) طالبة في دراسة أمبوسعيدي والعفيفي(2004) وبلغت (133) طالبة في دراسة أمبوسعيدي(2005) وقد لفت أنتباه الباحث ان كلا العينتين في الدراستين السابقتين تم اختيارهن قصدياً من نفس المدرستين ويعلل أمبوسعيدي (2005) هذا الامر لكي تكون مقارنة نتائج دراسته مع دراسة أمبوسعيدي والعفيفي (2004) منطقية وعلمية على أساس ان الطالبات في كلا الدراستين مررن بالخبرة ذاتها ، وعلى يد المعلمات انفسهن ، والمقرر الدراسي نفسه.أما دراسة الكناني وفيضي (2012) فقد بلغت عينة الدراسة (60) طالباً وطالبة من طلبة المرحلة الثانية من كلية التربية الرياضية ، أما دراسة Tollefson&tripp(1983) فقد كان حجم العينة (104) تم اختيارهم عشوائياً ، أما دراسة Tollefson&Chein(1986) فلم يذكر حجم العينة ! اما البحث الحالي فقد كان حجم العينة (105) طالباً يمثلون جميع طلاب الصف الاول المتوسط في ثانوية المتميزين.

أداة الدراسة:

تمثلت أداة الدراسة في الدراسات السابقة إعداد نموذجين من إختبار تحصيلي من نوع أسئلة إختيار من متعدد ذي الاربع بدائل (يختلفان في ترتيب البدائل)كما في دراسة أمبوسعيدي والعفيفي (2004) ودراسة أمبوسعيدي (2005) ، ودراسة الكناني وفيضي (2012) أجريت له إجراءات الصدق والثبات.

أما دراسة Tollefson& Tripp(1983) ، ودراسة Tollefson&Chein(1986) فقد استخدم الباحثان ثلاثة نماذج من اختبار تحصيلي طبقت على ثلاثة مجاميع من عينة البحث ، أما البحث الحالي فقد أعد الباحث نموذجين لاختبار تحصيلي (الكتروني) طبق على عينة البحث بعد أن أجرى له إجراءات الصدق والثبات.

الاساليب الاحصائية المتبعة:

تنوعت الاساليب الاحصائية لايجاد دلالات الصدق والثبات ، فاستخدم مربع كاي Chi- squareفي إيجاد دلالة الفروق المعنوية بين أراء المحكمين عند إيجاد الصدق الظاهري كما في دراسة أمبوسعيدي والعفيفي(2004)، وأمبوسعيدي(2005)،والكناني وفيضي(2012). أما الثبات فاستخدمت معادلة كيودور- ريتشاردسونKR-20كما في دراسة أمبوسعيدي والعفيفي(2004) أمبوسعيدي(2005) ، ولم يشر الكناني وفيضي(2012) الى طريقته في إيجاد الثبات ،

اما البحث الحالي فقد أستخدم الباحث مربع كاي chi-square لايجاد دلالة الفروق المعنوية لاراء المحكمين عند حساب الصدق الظاهري وبيان صلاحية فقرات الاختبار التحصيلي ، أما الثبات فقد استخدم الباحث معادلة الفا- كرونباخ لايجاده ، كما استخدم الاختبار t-test لايجاد دلالة الفروق بين المتوسطات.

نتائج الدراسات السابقة:

اختلفت نتائج الدراسات السابقة فيما بينها فبعضها أظهرت عدم وجود فروق ذات دلالة إحصائية في متوسطات معاملات صعوبة الفقرات ، ومعاملات تمييز الفقرات كما في دراسة Tollefson and Tripp(1983) وخالفتها في ذلك دراسة Tollefson and Chein(1986) إذ أظهرت النتائج وجود فروق ذات دلالة إحصائية في متوسطات معاملات الصعوبة ولم تشر الى وجود فروق ذات دلالة إحصائية في متوسطات معاملات التمييز. أم دراسة الزيات (1990) فقد أظهرت أن صياغة مشتتات الاجابة في أسئلة الاختيار من متعدد له دلالة إحصائية على معاملات صعوبة الفقرات وقوة تميزها وان فاعلية الفقرة تتوقف على فاعلية مشتتاتها وأتفقت دراسة أمبوسعيدي والعفيفي(2004) مع تلك النتائج إذ تبين وجود فروق ذات دلالة معنوية بين متوسطات مجموعتي الدراسة وكذلك في معاملات الصعوبة والتمييز.أما دراسة أمبوسعيدي (2005) فقد خالفت في نتائجها الدراستين السابقتين اذ أظهرت عدم وجود ذات دلالة إحصائية بين المتوسطات الحسابية لمعاملات الصعوبة ومعاملات التمييز لنموذجي الاختبار رغم اتفاقها مع دراسة امبوسعيدي والعفيفي(2004) في عينة الدراسة والخبرات التي تلقتها أفراد العينة، وأتفقت معها دراسة الكناني وفيضي(2012) إذ أظهرت نتائج الاخيرة عدم وجود فروق ذات دلالة إحصائية تعزى الى تأثير تغير موقع البديل الصحيح على معاملات الصعوبة ومعاملات التمييز للاختبار المعد.

منهجية البحث واجراءاته:

منهج البحث

استخدم الباحث المنهج الوصفي.

مجتمع البحث وعينته

تكون مجتمع البحث من طلاب الصف الاول المتوسط في ثانوية المتميزين في الديوانية البالغ عددهم (105) طالباً من طلاب الصف الاول المتوسط في ثانوية المتميزين للعام الدراسي 2017-2018 . وقد تم إختيار عينة البحث بصورة قصدية كون الباحث مدرس مادة الفيزياء في الثانوية المذكورة. وقد استعمل الباحثين صورتين لذات الاختبار التحصيلي يختلفان فقط في موقع المشتت القوي نسبة الى البديل الصحيح اذ جرى الاختبار الكترونياً في مختبر الحاسوب لثانوية المتميزين وبمساعدة من مدرس الحاسوب[1] الذي ساهم في تدريب الطلاب على تقنية إخفاء البدائل من على شاشة الحاسوب ، قبل إجراء الاختبار فعلياً صباح يوم الاحد المصادف 8\1\2017 .

أداة البحث:

تمثلت أدوات البحث في اختبار الاختيار من متعدد الالكتروني الذي بلغ عدد فقراته (18) فقرة ببدائل اربعة في نموذجين يختلفان فقط من حيث توزيع البدائل بالنسبة للبديل الصحيح , وقام الباحث بتطبيق البحث الحالي على عينة البحث خلال الكورس الثاني للعام الدراسي 2017- 2018 وقد تم التحقق من الصدق الظاهري بعرضه على ثمانية من المحكمين المتخصصين في مادة طرائق تدريس الفيزياء لبيان صلاحية الفقرات من عدمه ، وقد أظهرت نتائج إختبار مربع كاي Chi-square test صلاحية جميع فقرات الاختبار التحصيلي كما مبين في جدول (1).

[1] المدرس أحمد مجيد المرعي

جدول (1).صلاحية فقرات الاختبار التحصيلي

أرقام الفقرات	الصلاحية		قيمة x^2المحسوبة	قيمة x^2الجدولية	الدلالة الاحصائية
	تصلح	لاتصلح			
1،2،3،4،5،6،8،9،11،12،13،14،15،16،17	8	0	8	3.84	ذات دلالة
7،10،18	7	1	4.5		ذات دلالة

قيمة (x^2) الجدولية عند مستوى دلالة إحصائية (0.05) وعند درجة حرية df=1

- استخراج معاملات الصعوبة ، وتمييز الفقرات ، وقوة تشتيت البدائل ، والانحراف المعياري للدرجات ، ومعامل ثبات الاختبار .

❖ **التحليل الاحصائي لفقرات الاختبار** Statistical analysis for test

ان عملية تحليل فقرات الاختبار لها أهمية كبيرة ، اذ تساعد في قياس السمة بالشكل الدقيق وتطور فقرات الاختبار الى الحد الذي يعطي دلالة فيما يقيسه ذلك الاختبار وعملية التحليل تعد وسيلة في تحسين نوع الاختبارعن طريق معرفة سهولة أو صعوبة كل فقرة من فقرات الاختبار وقدرتها على التمييز وحذف الفقرات غير الصالحة (Seam,1999:220) بعد اكتمال تطبيق الاختبار تم احتساب النسبة المئوية وهي (27%) من المجموعتين العليا والدنيا فقد أختار الباحث (28) طالباً يمثلون المجموعة العليا و(28) طالباً يمثلون المجموعة الدنيا ثم تم احتساب الخصائص السيكومترية.

❖ **معامل صعوبة الفقرات** Coefficient of difficulty of the items

معامل صعوبة الفقرة : هو النسبة المئوية للطلاب الذين استجابوا على الفقرة استجابة صحيحة الى العدد الكلي للطلاب ، فعندما تكون النسبة عالية فهذا يدل على سهولة الفقرة وعندما تكون النسبة منخفضة فهذا يدل على صعوبة الفقرة، وان فقرات الاختبار تعد مقبولة اذا كان معامل صعوبتها يتراوح بين (0.20-0.80)(ابوجادو،واخرون،2005:210)وبعد احتساب معامل الصعوبة للاختبار وجد انها تتراوح بين (0.37 – 0.84)وهذا يدل على ان جميع الفقرات مقبولة وعليه اصبح الاختبار مؤلفاً من (18) فقرة.

❖ **معامل تمييز الفقرات** Items discrimination

معامل تمييز الفقرة : هو قدرة الفقرة على التمييز بين الطلاب ذوي المستويات العليا والمستويات الدنيا للسمة التي يقيسها الاختبار، وبعد ان طبقت معادلة معامل تمييز الفقرة تم حساب معامل تمييز كل فقرات الاختبار وظهرانها تتراوح بين (0.21-0.77) ، ويؤكد Ebelفي هذا السياق ان فقرات الاختبار تعد جيدة عندما يكون معامل تمييزها (0.20) او اكثر (Ebel,1972:404) وتأسيساً على ذلك فان جميع الفقرات مميزة وعليه فان عدد الفقرات الكلي ستكون(18)فقرة.

جدول (2). ثبات الاختبار في التجزئة النصفية

الفقرات	عدد الفقرات	الوسط الحسابي	الانحراف المعياري	قيمة t المحسوبة	قيمة t الجدولية	الدلالة الاحصائية
الفردية	9	37.21	8.28	0.34	2.1	غير دالة
الزوجية	9	35.33	9.2			

القيمة الجدولية عند مستوى دلالة إحصائية (0.05) وعند درجة حريةdf=16

❖ **ثبات الاختبار** Test reliability

يقصد بثبات الاختبار: هو الاختبار الذي يعطي نتائج متقاربة أو النتائج نفسها اذا طبق الاختبار اكثر من مرة في ظروف متماثلة (العاني ، 2010: 250).

تم احتساب ثبات الاختبار باستعمال :

❖ **معادلة الفا –كرونباخ** Cronbach's alpha

تستعمل هذه المعادلة لحساب معامل الثبات واظهرت نتيجة تطبيق المعادلة (0.78) اذ تعد قيمة مقبولة لمعامل الثبات كما في جدول (3). (الزاملي ، 2009 : 280)

جدول (3). معامل الثبات بطريقة الفا- كرونباخ

عدد الفقرات	معامل الفا- كرونباخ
18	0.78

❖ **الوسائل الاحصائية**:

1- معادلة معامل صعوبة الفقرة the formula of calculating difficulty coefficient
استعملت لاستخراج معامل صعوبة الفقرات للاختبار.(علام ، 2015: 258)

2- معادلة تمييز الفقرة equation discrimination paragraph
استعملت لايجاد القوة التمييزية لفقرات الاختبار (الهويدي، 2015 : 307)

3- معادلة الفا- كرونباخ Alpha- Cronbach equation استعملت لايجاد ثبات الاختبار

معادلة t-test لعينتين مستقلتين (غير متساويتين) لايجاد دلالة الفروق بين المتوسطات (علام ، 2012:145)

نتائج البحث:

السؤال الاول: ما أثر تغيير موقع المشتت القوي عند إستخدام تقنية إخفاء البدائل الخاطئة على المتوسطات الحسابية لاستجابات الطلاب عن أسئلة الاختيار من متعدد ؟

للإجابة عن هذا السؤال ، تم حساب المتوسطات الحسابية ، والانحرافات المعيارية، واختبار t-test لعينتين مستقلتين لايجاد دلالة الفروق عند مستوى دلالة إحصائية (0.05) كما مبين في جدول (4).

جدول (4). المتوسطات الحسابية والانحرافات المعيارية وقيمة t-test لاداء الطلاب في نموذجي الاختبار

النموذج	عدد الطلاب	المتوسط الحسابي	الانحراف المعياري	t-test	t-test الجدولية	مستوى الدلالة
الاول	53	11.90	3.99	0.594	1.98	غير دالة
الثاني	52	12.30	3.91			

قيمة t-testالجدولية عند مستوى دلالة احصائية (0.05) وتحت درجة حرية df=103

يتبين من جدول (4) ان المتوسط الحسابي لاداء الطلاب في النموذج الثاني أعلى من المتوسط الحسابي لاداء الطلاب في النموذج الاول ، اذ بلغ المتوسط الحسابي للنموذج الثاني (12.30) والانحراف المعياري (3.91) في حين بلغ المتوسط الحسابي للنموذج الاول (11.90) وبانحراف معياري (3.99) ولمعرفة ما اذا كانت تلك الفروق دالة احصائياً تم استخدام اختبار t-test لعينتين مستقلتين . وقد أشارت نتائج الاختبار الى عدم وجود فروق ذات دلالة إحصائية عند مستوى دلالة

إحصائية(0.05) بين المتوسطات الحسابية لنموذجي الاختبار. وتتفق نتائج هذا البحث مع ماتوصلت اليه نتائج دراسة أمبوسعيدي (2004)، والكناني وفيضي(2012) وتخالف ما توصلت اليه نتائج دراسة امبوسعيدي والعفيفي (2005) اذ أظهرت الاخيرة وجود فروق ذات دلالة إحصائية بين المتوسطات الحسابية لنموذجي يعزى بحسب الباحث الى تغير موقع المشتت القوي الذي وضع بجانب البديل الصحيح في النموذج الاول وعلى بعد بديل او بديلين في النموذج الثاني.

السؤال الثاني: ما أثر تغيير موقع المشتت القوي عند استخدام تقنية إخفاء البدائل الخاطئة على معاملات الصعوبة لاسئلة الاختيار من متعدد في مادة الفيزياء

للاجابة عن هذا السؤال تم اولا حساب معاملات الصعوبة لجميع فقرات الاختبار للنموذجين ثم بعد ذلك حساب المتوسطات الحسابية والانحرافات المعيارية لمعاملات الصعوبة لكل نموذج واختبار t-test **للكشف عن دلالة** الفروق في المتوسطات لمعاملات الصعوبة لنموذجي الاختبار كما موضح في جدول (5).

جدول (5). يوضح معاملات الصعوبة لفقرات نموذجي الاختبار

الفقرة	معامل الصعوبة			الفقرة	معامل الصعوبة		
	النموذج الاول	النموذج الثاني	الفرق بين النموذجين		النموذج الاول	النموذج الثاني	الفرق بين النموذجين
1	0.52	0.55	0.03	10	0.57	0.54	0.03
2	0.54	0.51	0.03	11	0.63	0.52	0.11
3	0.76	0.77	0.01	12	0.39	0.26	0.13
4	0.46	0.42	0.04	13	0.67	0.83	0.16
5	0.76	0.35	0.41	14	0.29	0.20	0.09
6	0.31	0.38	0.07	15	0.42	0.34	0.08
7	0.58	0.55	0.03	16	0.60	0.42	0.18
8	0.30	0.69	0.39	17	0.55	0.73	0.18
9	0.73	0.74	0.01	18	0.60	0.66	0.06

يتضح من الجدول اعلاه ان عدد الفقرات التي كان الفرق فيها في معاملي الصعوبة بين النموذج الاول ومعامل الصعوبة للنموذج الثاني يساوي(20%) لا يتجاوز عددها (2) فقرات هي (5،8) اي ما نسبته (11 %) من المجموع الكلي للفقرات البالغة (18) ، اما باقي الفقرات فان الفرق في معامل الصعوبة فيها بين النموذجين فأقل من ذلك. وقد اختيرت القيمة (20%) معياراً للحكم على وجود الفروق بين معاملات الصعوبة بين نموذجي الاختبار استناداً الى ماذهبت اليه الدراسات السابقة مثل دراسة امبوسعيدي (2004) ، ودراسة امبوسعيدي والعفيفي (2005). يتضح من جدول(3) ان المتوسط الحسابي لمعاملات الصعوبة لنموذجي الاختبار متقاربة جداً اذ ان الفارق بينهما لايكاد يذكر وعند اختبار ما اذا كانت تلك الفروق البسيطة ذات دلالة احصائية تم استخدام اختبار t-test لعينتين مستقلتين وقد اشارت النتائج الى عدم وجود فروق ذات دلالة احصائية عند مستوى دلالة (0.05) اي ان نموذجي الاختبار متساويين في متوسطي معاملات صعوبتيهما وانه لاتفضيل لاي من النموذجين على الاخر في مستوى الصعوبة.

جدول(3). المتوسط الحسابي والانحراف المعياري وقيمة t-test للفروق بين معاملات الصعوبة لنموذجي الاختبار

النموذج	عدد الطلاب	المتوسط الحسابي	الانحراف المعياري	قيمة t-test الحسابية	قيمة t-test الجدولية	مستوى الدلالة
الاول	53	0.54	0.15	0.82	1.98	غير دال
الثاني	52	0.52	0.18			

قيمة t-test الجدولية عند مستوى دلالة إحصائية (0.05) وتحت درجة حرية df=103

السؤال الثالث: ما أثر تغيير موقع المشتت القوي عند استخدام تقنية اخفاء البدائل الخاطئة على معاملات التمييز لاسئلة الاختيار من متعدد في مادة الفيزياء

للاجابة عن هذا السؤال تم حساب معاملات التمييز لجميع الفقرات في نموجي الاختبار كما تم حساب المتوسطات الحسابية والانحرافات المعيارية لمعاملات التمييز لكل نموذج واختبار t-test لعينتين مستقلتين للكشف عن دلالة الفروق بين المتوسطات الحسابية لمعاملات التمييز لنموذجي الاختبار كما مبين في جدول (4).

جدول (4). يوضح معاملات التمييز لفقرات نموذجي الاختبار

الفقرة	معامل الصعوبة			الفقرة	معامل الصعوبة		
	النموذج الاول	النموذج الثاني	الفرق بين النموذجين		النموذج الاول	النموذج الثاني	الفرق بين النموذجين
1	0.55	0.44	0.11	10	0.40	0.11	0.29
2	0.50	0.61	0.11	11	0.45	0.55	0.10
3	0.50	0.38	0.12	12	0.10	0.11	0.01
4	0.65	0.44	0.21	13	0.35	0.22	0.13
5	0.58	0.61	0.03	14	0.40	0.16	0.24
6	0.60	0.56	0.04	15	0.20	0.22	0.02
7	0.35	0.55	0.20	16	0.45	0.55	0.10
8	0.50	0.38	0.12	17	0.60	0.61	0.01
9	0.25	0.32	0.07	18	0.70	0.66	0.04

يظهر من جدول (4) ان هناك عدداً من الفقرات يكون فيها الاختلاف في معامل التمييز بين النموذجين يساوي (20%) فأكثر وهي (14،10،7) وتشكل ما نسبته (17%) من المجموع الكلي للفقرات والبالغة (18) فقرة اما باقي الفقرات فان الفارق في معامل التمييز فيما بينها فهو قليل لايكاد يذكر.

جدول (5). المتوسط الحسابي والانحراف المعياري وقيمة t-testلعينتين مستقلتين لايجاد دلالة الفروق بين معاملات التمييز لنموذجي الاختبار

النموذج	عدد الطلاب	المتوسط الحسابي	الانحراف المعياري	قيمة t-test الحسابية	قيمة t-test الجدولية	مستوى الدلالة
الاول	53	0.45	0.16	0.53	1.98	غير دالة
الثاني	52	0.41	0.18			

قيمة t-test الجدولية (1.98) عند مستوى دلالة أحصائية (0.05) وعند درجة حرية df=103

يظهر من جدول (5) ان معامل التمييز للنموذج الاول للاختبار مقارب لمعامل التمييز للنموذج الثاني لذات الاختبار اذ بلغ للنموذج الاول (0.45) وللنموذج الثاني (0.41) ولذا لاتوجد فروق ذات دلالة احصائية بين نموذجي الاختبار اي ان لكلا النموذجين قدرة تمييزية متساوية بين الطلاب .

مناقشة نتائج البحث:

اظهرت نتائج البحث الحالي عدم وجود فروق ذات دلالة احصائية عند مستوى دلالة (0.05) في المتوسطات الحسابية ، ومعاملات الصعوبة ، ومعاملات التمييز لاداء الطلاب على نموذجي الاختبار وهذا ما دفع الباحث الى وضع التساؤل الاتي : هل هناك تأثير بصري لوجود البدائل مجتمعة على شاشة الاختبار؟ والذي دفع الباحث الى هذا التساؤل ان المتوسط الحسابي للاختبار باجمعه انخفض حين اتيح للطالب اخفاء البديلين المتأكد من خطأهما (بحسب الطالب) ومن ثم اختيار البديل الصحيح (بحسب الطالب) من بين البديلين المتبقيين. فعند ظهور بديلين على الشاشة وامام الطالب عندها يستطيع عمل مقارنة بينهما بشكل جيد دون تشتيت (اي ان التركيز سيكون مرتفعاً هذه المرة) وهنا يمكن للباحث القول ان ظهور جميع البدائل على الشاشة ربما يؤثر بصرياً (برأي الباحث) على الطلاب ويشتت تركيزهم في اختيار البديل الصحيح .

التوصيات والمقترحات:

في ضوء النتائج أعلاه يوصي الباحث بمجموعة من التوصيات ويقترح عدداً من الافكار:

1- عدم الاعتماد على أسئلة الاختيار من متعدد بشكل أساس في إختبارات الطلبة

2- يجب التقيد بالتعليمات المتوافرة في الادب التربوي حول تصميم الاختيار من متعدد لكي يتم تجنب الكثير من نقاط الضعف لهذه الاسئلة مثل التخمين

3- ضرورة البحث عن طرائق تقويم أخرى فعالة بجانب اسئلة الاختيار من متعدد

4- القيام بمزيد من الدراسات والبحوث في موضوع أسئلة الاختيار من متعدد تناقش التأثير البصري للبدائل بصورة معمقة .

المصادر:

▪ أمبوسعيدي ،عبد الله بن خميس والعفيفي ، منى بنت محمد (2004). أثر تغيير المشتت في بعض الخصائص السايكومترية لاسئلة الاختيار من متعدد في مادة الفيزياء ، **مجلة العلوم التربوية والنفسية** ، المجلد (5) ، العدد (2) ،

▪ أمبوسعيدي،عبد الله بن خميس(2005). أثر تغيير موقع المشتت القوي عند استخدام تقنية البدائل الخطأ في صعوبة الفقرات لاسئلة الاختيار من متعدد في الفيزياء وتمييزها ، **مجلة العلوم التربوية والنفسية** ، المجلد (6) ، العدد (3)

▪ الزيات ، فتحي مصطفى(1990).أثر قوة تشتيت البدائل في أسئلة الاختيار من متعدد على صدق الاختبار وثباته **مجلة كلية التربية** ، جامعة المنصورة ، العدد (11) .
▪ علام ، صلاح الدين محمود (2003). **تحليل بيانات البحوث النفسية والتربوية والاجتماعية** ، ط3، كلية التربية ، جامعة الأزهر ، دار الفكر للنشر والتوزيع.
▪ عودة ، احمد سليمان وملكاوي ، فتحي حسن (1998) . **القياس والتقويم في العملية التدريسية** ، ط2، دار الأمل ، عمان ، الأردن.
▪ عودة ، أحمد سليمان(1999). **القياس والتقويم في العملية التدريسية** ، ط4 ، دار الأمل ، عمان ، الأردن.
▪ فهمي ، فاروق ، وعبد الصبور، منى (2002). **المدخل المنظومي في مواجهة التحديات التربوية المعاصرة والمستقبلية** ، ط1 ، دار المعارف ، القاهرة .
▪ الكناني ، عايد كريم وفيضي ، فاضل عبد (2012). فاعلية استخدام موقع البديل الخاطئ في صعوبة فقرات اسئلة الاختيار من متعدد وتمييزها ، **مجلة القادسية لعلوم التربية الرياضية** ، المجلد (12) ، العدد(1)
▪ الهويدي ، زيد (2005). **الأساليب الحديثة في تدريس العلوم** ، ط1 ، دار الكتاب الجامعي ، الأمارات العربية المتحدة.
▪ الوكيل، حلمي احمد وبشير،محمود حسين (2001). **الاتجاهات الحديثة في تخطيط وتطوير مناهج المرحلة الاولى** ، القاهرة ، دار الفكر العربي.

Refrence:

▪ Anastasi,A(1976).**Psychological Testing**, New York ,McMillan company,4th ed.
▪ Buros,O.N.(1972).**The seventh mental measurement year book**, high and park,N.J.,The Gryphon Press.
▪ Ebel,R.L.(1979).**Essential of Educational Measurement**,Prentice-Hall,INC.,Englewood cliffs, New Jersey.
▪ Ebel,R.L.(1979).The relation of item discrimination to test reliability, **journal of education measurement**,(fall 1676),4,p.125-128.
▪ Ebel,R.L.,Willams,B.J.(1979).**The effect of varying the number of alternatives per item on multiple** – choice vocabulary test item.
▪ Gary,R.B. & Lowery,S.R.(1976). Misinformation reliability and item discrimination indices on multiple choice tests, **the paper presented at the annual meeting of the A.E.R.A.**,San Francisco,California,April,Eric127529.
▪ Gay,L.R.(1980).**Educational Evaluation Measurement Competencies for Analysis and Application**, Charles E.,Merill Publshing company,A.Bell &Howell company,London.
▪ Hills,J.(1981).**Educational Evaluation & Measurement** , Prentice – Hall Inc.,Englewood Cliffs, New Jersey.

▪ Tollefson,N.&Chen(1986).A comparison of item difficulty and item discrimination of multiple choice items using, none of the above options, paper presented at the annual meeting of the Midwest,E.R.A.,Chicago,11october,ERIC,ED278704.
▪ Tollefson,N.&Tripp,A.(1983).The effect of item format on item difficulty and item discrimination **paper presented at the annual meeting of the Amer**.Ed. Res.Assoc., April11-15,F.D.230582.
▪ West,D(1993).**Introduction in Teaching learning and Assessment in Science Education** ,London, Paul Chapman Publishing Ltd.
▪ Wisker,G.(1997).**Assessing for learning in English studies**, Some innovative practices, Teaching in higher education,2(2),123-139.

Abstract:
The aim of this research was to investigate the power of distractor upon difficulty index and discrimination index of the test items. To achieve the aim above, a multiple choice test in physics consist of (18) items was constructed. the sample consist of (105) male first grade student was chosen from distinguished school in Al-diwaniya city(2017-2018).the results showed that there was statistical significant difference in mean scores of the difficulty index and crimination index.A number of recommended and suggestions were proposed in the light of the above results.