



المملكة العربية السعودية  
وزارة التعليم العالي  
جامعة أم القرى  
مكة المكرمة  
كلية التربية  
قسم علم النفس

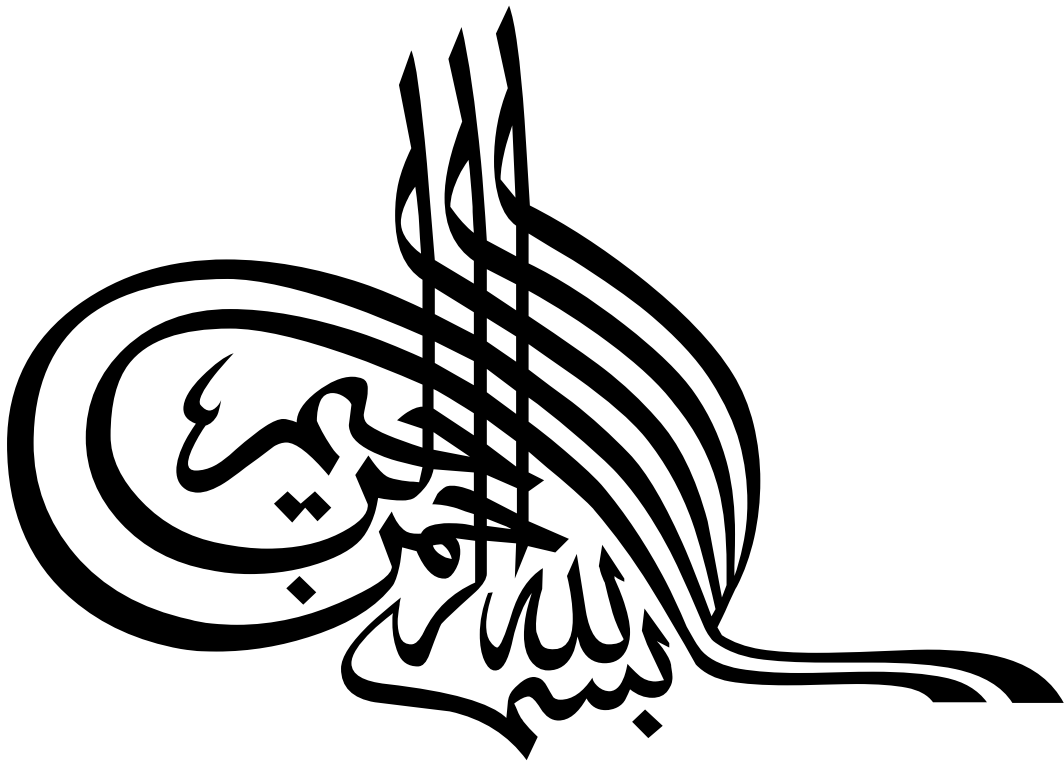
## أثر عدد البدائل في الخصائص السيكومترية للاختبار التحصيلي في الرياضيات

إعداد الطالب  
عبدالله بن أحمد آل شويل الغامدي

إشراف الأستاذ الدكتور  
ربيع بن سعيد طه  
أستاذ الإحصاء بقسم علم النفس بجامعة أم القرى

متطلب تكميلي للحصول على درجة الماجستير في علم النفس  
تخصص ( قياس وتقويم )

الفصل الدراسي الثاني ١٤٢٩هـ / ٢٠٠٨م



( - )

(Δ) ( , )

( , )

( , )

( , )

( , )



## شكر وتقدير

الحمد لله كما ينبغي لجلال وجهه وعظيم سلطانه، والشكر له على فضله وامتنانه، والصلاة والسلام على صفيه من عباده سيدنا محمد وعلى آله وصحبه وسلم، أشكر بداية سعادة الأستاذ الدكتور/ ربيع سعيد طه على تفضله بالإشراف على هذا البحث وتذليل العقبات والمصاعب التي واجهت الباحث منذ بدء الأفكار ثم اكتمال البنيان وقطف الثمار، ثقل الله بذلك موازينه وأجزل له المثوبة والأجر، وأشكر الأستاذ الدكتور/ عبدالحفيظ سعيد مقدم على قبوله مناقشة هذه الرسالة، وتحمله عناء السفر ومشقته فله مني جزيل الشكر والعرفان.

كما أتقدم بجزيل الشكر وأصدق الثناء إلى أستاذي القدير سعادة الأستاذ الدكتور/ محمد المري محمد إسماعيل والذي قام مشكوراً بمناقشة هذه الرسالة وتحكيم خطتها، فقد غمرني بعطفه وحنانه الكبير، وخصني بتوجيهاته وتعليقاته القيمة، والذي لا أنسى دعمه ووقفاته فجزاه الله عني خير الجزاء ونفع الله به وبعلمه وأدام عليه فضله ونعمته، والشكر موصول لسعادة الأستاذ الدكتور/ زايد بن عجير الحارثي على تفضله بتحكيم خطة هذا البحث وإثرائها بتوجيهاته الحكيمة، فضلاً عن تدرسه لي أهم المقررات التي استفدت منها كثيراً في هذا البحث فجزاه الله عني خير الجزاء.

كما أتقدم بالشكر الجزيل لجميع أعضاء هيئة التدريس بقسم علم النفس وأخص بالذكر الأستاذ الدكتور/ عبدالمنان ملا بار رئيس القسم، على تعاونهم الكبير مع الدارسين ومع الباحث على وجه الخصوص، والشكر لجامعة أم القرى ممثلة في عمادة كلية التربية وعمادة الدراسات العليا لما يقدمونه من عون وتسهيلات للدارسين، والشكر أيضاً لجميع زملائي في مرحلة الماجستير وأخص بالذكر رفيق دربي الأستاذ عبدالرحمن الأحمدى الذين قدم لي الكثير، والأستاذ سعيد عبدالفتاح الغامدي وعادل الهلالي ومحمد رافد و محمد عبدالرحمن الغامدي وأحمد السحيمي على ماقدموه من بذل وعطاء للباحث، كما أشكر مدارس الملك عبدالعزيز النموذجية بتبوك على تزويد الباحث بالأوراق المخصصة للإجابة، كما أتقدم بجزيل الشكر والعرفان لمنسوبي إدارة التربية والتعليم بمحافظة الليث التعليمية وعلى رأسهم سعادة الأستاذ/ مرعي بن محمد البركاتي مدير إدارة التربية والتعليم على ما بذلوه وبذلوا من مساعدة للباحث، وأخص بالشكر قسم الاختبارات المدرسية على ما قام به أعضاؤه من جهد كبير في مساعدة الباحث على تطبيق الاختبارات والتنسيق مع المدارس والإشراف عليها، وعلى رأسهم الأستاذ جزاء المطيري وأحمد الشطيبي والحسين العيايفي وعبدالله المهداوي، كما أشكر المدارس والطلاب الذي تم تطبيق أداة البحث عليهم.

كما أسطر بمداد من ذهب وعلى صفحات من نور جزيل الشكر وعظيم الامتنان لوالدي الكريمين، اللذان بذلا الغالي والنفيس في سبيل وصولي إلى هذه المرحلة، كما لا أنسى أن أشكر أخواني وأخواتي على وقوفهم بجانبني ومساعدتي في جميع مراحل دراستي، وأشكر زوجتي التي صبرت وضحت بوقتها من أجلي، كما أشكر كل من ساهم في إخراج هذا العمل بهذه الصورة.

## قائمة المحتويات

رقم  
الصفحة

الموضوع

.....	:	
.....	:	
.....	:	
( - ) .....	:	
.....	:	
.....	:	
.....	:	
.....	:	
.....	:	
.....	:	
( - ) .....( )	:	
.....	:	
.....	:	
.....	:	
.....	:	
.....	:	
.....	:	

( - ) ..... ( )  
.....

## تابع قائمة المحتويات

رقم الصفحة

الموضوع

.....	
.....	
.....	
.....	
.....	
( - )	.....( : ):
.....	
.....	
( - )	.....( ):
.....	
.....	
.....	
.....	
( - )	.....
.....	( )
.....	( )
.....	( )
.....	( )



## قائمة المعادلات

الصفحة	المعادلة	رقم المعادلة
٢٥	الوزن النسبي للأهداف السلوكية	(١)
٢٦	الوزن النسبي لموضوعات المادة الدراسية	(٢)
٢٦	إيجاد عدد الأسئلة في كل موضوع	(٣)
٢٦	إيجاد درجة كل سؤال	(٤)
٣٥	معامل سهولة المفردة في حالة النسب المئوية	(٥)
٣٦	معامل سهولة المفردة في حالة الترتيب	(٦)
٣٧	معامل تمييز المفردة في حالة المجموعتين الطرفيتين	(٧)
٣٨	معامل الارتباط الثنائي الأصيل	(٨)
٣٩	معامل جاذبية المموه	(٩)
٤١	معامل الثبات	(١٠)
٤٢	معادلة سبيرمان براون	(١١)
٤٢	معادلة رولون	(١٢)
٤٣	معادلة جتمان	(١٣)
٤٣	معادلة كيودر - ريتشاردسون ٢٠	(١٤)
٤٤	معادلة كيودر - ريتشاردسون ٢١	(١٥)
٤٤	معادلة ألفا كرونباخ	(١٦)
٧٩	الإحصائي M (هاكستين وولين)	(١٧)
٧٩	الإحصائي V (هيز)	(١٨)

## قائمة الجداول

الصفحة	الجدول	رقم الجدول
٧٠	توزيع مدارس العينة	(١)
٧٢	جدول مواصفات الاختبار التحصيلي	(٢)
٧٥	معاملات السهولة والتمييز وفعالية المشتتات للعينة الاستطلاعية	(٣)
٧٧	قيم معاملات الثبات للعينة الاستطلاعية	(٤)
٨٥	تكرارات معاملات السهولة للنماذج الثلاثة	(٥)
٨٦	تكرارات معاملات التمييز للنماذج الثلاثة	(٦)
٨٦	القدرة التمييزية للبدائل الصحيحة	(٧)
٨٧	القدرة التمييزية للمشتتات	(٨)
٨٨	تحليل التباين الأحادي لمتوسطات معاملات السهولة	(٩)
٩٠	اختبار (ت) عند نقطة القطع المثالية لمعامل السهولة	(١٠)
٩١	تحليل التباين الأحادي لمتوسطات معاملات التمييز	(١١)
٩٢	معاملات ثبات وصدق المحك للنماذج الثلاثة	(١٢)
٩٣	معاملات الثبات للنماذج الثلاثة واختبار مربع كاي	(١٣)
٩٤	معاملات صدق المحك التلازمي واختبار مربع كاي	(١٤)
٩٥	معاملات صدق المحك التنبؤي واختبار مربع كاي	(١٥)
١٣٧	معاملات تمييز وسهولة فقرات النموذج ذي الثلاثة بدائل	(١٦)
١٣٨	نسبة اختيار بدائل النموذج ذي الثلاثة بدائل	(١٧)
١٣٩	معاملات تمييز وسهولة فقرات النموذج ذي الأربعة بدائل	(١٨)
١٤٠	نسبة اختيار بدائل النموذج ذي الأربعة بدائل	(١٩)
١٤١	معاملات تمييز وسهولة فقرات النموذج ذي الخمسة بدائل	(٢٠)
١٤٢	نسبة اختيار بدائل النموذج ذي الخمسة بدائل	(٢١)
١٤٣	تكرار درجات الطلاب في النماذج الثلاثة	(٢٢)
١٤٣	مقاييس النزعة المركزية والتشتت لدرجات الطلاب على نماذج الاختبار الثلاثة	(٢٣)

- 
- 
- 
- 
- 
-

## مقدمة

يعتبر التقويم التربوي عنصراً أساسياً في منظومة العملية التعليمية والتربوية، لما له من دور فاعل في تطويرها وتقدمها، ولمعرفة مدى التقدم في تحقيق أهداف العملية التعليمية يقوم المعلم بملاحظة سلوك طلابه باستمرار، ويجمع المعلومات عن هذا السلوك بأدوات وطرق مختلفة، ولا بد من استخدام أدوات دقيقة لقياس هذا التطور والتقدم، وعلى ضوء هذه النتائج تتخذ القرارات المناسبة على أساسها من ترفيع للطلاب لمراحل أعلى وتصنيفهم وما إلى ذلك، وتعد الاختبارات التحصيلية من الوسائل المستخدمة في هذا المجال حيث يذكر سليمان (٢٠٠١م: ١١) "أن الاختبارات التحصيلية من أهم وسائل التقويم التربوي التي يستخدمها المعلمون في المدارس".

وتوفر هذه الاختبارات معلومات يمكن أن تستخدم في قياس مستوى تحصيل الطلاب ومدى نجاح عملية التعليم وتشخيص نواحي ضعف الطلاب ووضع العلاج المناسب وتصنيف الطلبة والتمييز بينهم. (أبو سل، ٢٠٠٢م).

وذكر سماره وآخرون (١٩٨٩م: ٦٥) "أن الاختبارات الموضوعية أحد أنواع الاختبارات التحصيلية واشتهرت بالموضوعية لما تمتاز به من دقة وموثوقية وعدم تأثر تصحيحها بالعوامل الذاتية للمصحح".

ويكثر المعلمون من استخدام اختبارات الاختيار من متعدد في اختباراتهم لسهولة تصحيحها ومرونة استخدامها حيث يذكر علام(٢٠٠٦م : ١٤٧) "أن اختبارات الاختيار من متعدد أحد أنواع الاختبارات الموضوعية المستخدمة حيث يتميز هذا النوع من الاختبارات بالمرونة في طريقة استخدامها في قياس المستويات المعرفية المختلفة عن غيرها من أنواع المفردات الاختبارية كما يمكن استخدامها في مختلف المجالات الدراسية".

ويلاحظ المتمعن أن أسئلة الاختيار من متعدد بأربعة بدائل هو النمط الأكثر شيوعاً في اختبارات الاختيار من متعدد، ويرجع ذلك إلى عوامل كثيرة منها اعتقاد معدي هذه الاختبارات بأنه كلما زاد عدد بدائل الفقرة عن ثلاثة قلت فرصة تخمين الإجابة الصحيحة في ظل غياب الخصائص السيكومترية الأخرى للاختبار وفقراته.

ولقد اهتمت بعض الدراسات الأجنبية في بحث وتحديد مدى تأثير عدد البدائل في اختبارات الاختيار من متعدد على الخصائص السيكومترية، وحاولت معرفة العدد الأمثل والمناسب من عدد البدائل التي ينبغي استخدامها في هذا النوع من الاختبارات، فنجد دراسة جرير(١٩٧٥، Grier) تشير إلى ارتفاع معامل ثبات النموذج ذي الثلاثة بدائل ويليها نموذج البديلين ثم نموذج الأربعة بدائل، بينما نجد دراسة ستراتون وكاتس (Straton & Catts, ١٩٨٠) ودراسة بيدسكو ونيفو (Budescu & Nevo, ١٩٨٥) تشير هاتين الدراستين إلى أن معامل الثبات يزداد بزيادة عدد البدائل.

أما على الصعيد العربي فنجد بعض هذه الدراسات كدراسة فقوسة (١٩٩٤م) التي بحثت في أثر عدد البدائل على تجانس الاختبار الداخلي وعلى معاملات الصعوبة والتمييز

وأوصت الدراسة باستخدام اختبار الاختيار من متعدد ذي الثلاثة بدائل في اختبارات الرياضيات.

وفي دراسة الموسوي (١٩٩٧م) التي ركزت على تحديد العدد الأمثل لبدايل اختبارات الاختيار من متعدد ، و أشارت نتائج هذه الدراسة أن العدد الأمثل لبدايل الإجابة هو ثلاثة بدائل في اختبارات الاختيار من متعدد.

بينما نلاحظ شيوع فكرة أنه كلما زاد عدد البدائل في اختبارات الاختيار من متعدد قلت فرصة التخمين فقد أوحى هذه الفكرة إلى كثير من معدي أسئلة اختبارات الاختيار من متعدد إلى وضع أربعة بدائل أو خمسة بدائل ، متجاهلين في ذلك بقية الخصائص السيكومترية للاختبار وفقراته ، علما بأن أغلب الدراسات تشير إلى عدم وجود فروق ذات دلالة إحصائية لأغلب الخصائص السيكومترية بين نماذج الاختبارات ذات الثلاثة بدائل والأربعة بدائل والخمسة بدائل ، وهذا كان دافعاً للباحث لإجراء هذا البحث والوقوف على الأثر الناتج عن تغيير عدد البدائل في اختبارات الاختيار من متعدد على الخصائص السيكومترية للاختبار التحصيلي في مادة الرياضيات.

## **مشكلة البحث وتساؤلاته**

يكثر المعلمون من استخدام الاختبارات الموضوعية في قياسهم لتحصيل الطلبة ، ويتوقف اختيار نوع الأسئلة على عدة عوامل ، منها طبيعة المحتوى المراد قياسه ، والهدف من القياس ، وعدد الطلاب ، والإمكانات المدرسية ، ومهارة المعلم في كتابة الأسئلة ،

حيث تتطلب الأسئلة الموضوعية مهارات معينة وخصوصاً اختبارات الاختيار من متعدد، فبعضهم يستخدم مفردات تحتوي على خمسة بدائل حتى يتم القضاء على فرصة التخمين، وبالرغم من هذه الميزة إلا أنها تزيد من الوقت الذي يحتاجه المفحوص في الإجابة على مثل هذه المفردات، حيث توصلت دراسة ستراتون وكاتس Straton & Catts (١٩٨٠م) إلى أن الوقت الذي يحتاجه المفحوص في قراءة البديل يساوي تقريباً نصف الوقت الذي يحتاجه المفحوص في قراءة متن السؤال، وبالتالي يمكن زيادة عدد مفردات الاختبار مع نقص عدد البدائل؛ مما يؤدي إلى زيادة التمثيل العيني وبالتالي زيادة صدق المحتوى الذي يعد من أهم طرق الصدق في الاختبارات التحصيلية، كما أن زيادة عدد البدائل للمفردات يقلل من إمكانية وضع مموهات أو بدائل جذابة للمفحوصين؛ فيميل بعض المعلمين إلى استخدام مفردات تحتوي على أقل من خمسة بدائل لعجزه عن إيجاد بديل خامس جذاب ومتجانس مع بقية البدائل، ومنهم من يكفي بأقل من ذلك.

واختيار العدد المناسب من البدائل عندما يتساوى تأثير العوامل السابقة في اختيار نوع الأسئلة؛ يتوقف على الخصائص السيكومترية التي تجعل مفردات الاختبار أفضل ما يمكن من حيث الصعوبة والتمييز؛ وتجعل من الاختبار أفضل ما يمكن من حيث الثبات والصدق، وبالتالي يقدم الاختبار نتائج دقيقة تعكس المستوى الحقيقي للطالب.

والمتتبع للدراسات والبحوث التي أجريت للكشف عن أثر عدد البدائل على بعض الخصائص السيكومترية للاختبار، يجد بعضها كدراسة جريير (Grier, ١٩٧٥) أثبتت أن نموذج الاختبار ذو الثلاثة بدائل يتمتع بأعلى ثبات، ويليه الاختبار ذو البديلين ثم نموذج

الاختبار ذو الأربعة بدائل، و بعضها الآخر أثبتت أن معامل الثبات يزداد بزيادة عدد البدائل مثل دراسة بيدسكو ونيفو (Budescu & Nevo, ١٩٨٥) بينما نجد أن أغلب الدراسات أثبتت عدم وجود فروق ذات دلالة بين معاملات الثبات كدراسة كوستين (Costin, ١٩٧٢) ودراسة الشبلي (١٩٩٥م) ودراسة الموسوي (١٩٩٧م).

أما الدراسات التي بحثت في أثر عدد البدائل على الخصائص السيكومترية للفقرة، نجد بعضها أثبتت أن معاملي الصعوبة والتمييز في النموذج ذو الثلاثة بدائل أعلى منه في النموذج ذو الأربعة بدائل كدراسة كوستين (Costin, ١٩٧٠)، بينما دراسة لورد (Lord, ١٩٧٧) ودراسة ستراتون وكاتس (Straton & Catts, ١٩٨٠) أثبتت أن صعوبة الفقرة وتمييزها يتناسبان طرديا مع عدد البدائل .

نلاحظ مما سبق عدم وجود اتساق في نتائج الدراسات والبحوث التي أجريت في أثر عدد البدائل على الخصائص السيكومترية للاختبار وفقراته، مما تسبب في حيرة لدى المعلمين في تحديد أفضل عدد للبدائل عند إعدادهم لاختبارات الاختيار من متعدد، وعدم التأكد من مدى فعالية عدد البدائل وأثرها في هذه الاختبارات، فتارة يستخدم المعلم ثلاثة بدائل وتارة أربعة أو خمسة، هذا ما دفع الباحث إلى التأكد من مدى الأثر الذي يحدثه عدد البدائل (ثلاثة، أربعة، خمسة) بدائل في اختبارات الاختيار من متعدد على الخصائص السيكومترية للاختبار التحصيلي للرياضيات.

ويمكن للباحث تلخيص مشكلة البحث في التساؤل الرئيس التالي:



مأثر عدد البدائل على الخصائص السيكومترية للاختبار التحصيلي في مادة

## الرياضيات ٩.

ويتفرع عن هذا التساؤل الأسئلة التالية:

- هل توجد فروق ذات دلالة إحصائية عند مستوى (٠,٠٥) بين متوسطات قيم معاملات السهولة ( $\Delta$ ) في نماذج الاختبار التحصيلي لمادة الرياضيات من نوع الاختيار من متعدد ذات البدائل الثلاثة و الأربعة و الخمسة؟.
- هل توجد فروق ذات دلالة إحصائية عند مستوى (٠,٠٥) بين متوسطات قيم الدرجات الزائفة المقابلة لمعاملات التمييز في نماذج الاختبار التحصيلي لمادة الرياضيات من نوع الاختيار من متعدد ذات البدائل الثلاثة و الأربعة و الخمسة؟.
- هل توجد فروق ذات دلالة إحصائية عند مستوى (٠,٠٥) بين قيم معاملات ثبات الإتساق الداخلي في نتائج نماذج الاختبار التحصيلي لمادة الرياضيات من نوع الاختيار من متعدد ذات البدائل الثلاثة و الأربعة و الخمسة؟.
- هل توجد فروق ذات دلالة إحصائية عند مستوى (٠,٠٥) بين قيم معاملات صدق المحك (التلازمي – والتتبؤي) في نتائج نماذج الاختبار التحصيلي لمادة الرياضيات من نوع الاختيار من متعدد ذات البدائل الثلاثة و الأربعة و الخمسة؟.

## أهداف البحث

يهدف هذا البحث بصورة عامة إلى التعرف على الأثر الذي يحدثه اختلاف عدد البدائل في اختبار الاختيار من متعدد على الخصائص السيكومترية لمفردات الاختبار (معامل السهولة والتمييز)، وعلى الخصائص السيكومترية للاختبار (الثبات و الصدق) ويمكن حصر هذه الأهداف فيما يلي:

١. الكشف عن أثر اختلاف عدد البدائل في أسئلة الاختيار من متعدد على الخصائص السيكومترية للاختبار ومفرداته في الاختبار التحصيلي لمادة الرياضيات.

٢. تحديد العدد الأمثل من البدائل المستخدمة في اختبارات الاختيار من متعدد والتي يكون صدقها وثباتها عالياً في الاختبارات التحصيلية لمادة الرياضيات.

## أهمية البحث

تكمن أهمية هذا البحث في الجوانب التالية :

١. مساعدة القائمين على مشروع تطوير الاختبارات المدرسية بوزارة التربية والتعليم في بناء الاختبارات التحصيلية للمواد الدراسية على أكمل وجه وبأقل جهد.

٢. إسهام علمي متواضع في التراكم العلمي المعرفي القائم في هذا الجانب وإضافة لبنة علمية قائمة عن تجربة فعلية إلى المكتبة العربية.

٣. تقديم اختبار تحصيلي صالح للاستخدام مستقبلا لغرض تقويم تحصيل

طلاب الصف الثاني متوسط في الوجدتين الأولى والثانية من مقرر مادة

الرياضيات في مدارس التعليم العام بالمملكة العربية السعودية.

٤. قد تسهم نتائج هذه الدراسة وتكون دافعا للباحثين للقيام بالمزيد من

البحوث المشابهة على نطاق أكبر وفي مواد دراسية أخرى.

## مصطلحات البحث

### • 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 وعددا من الإجابات

الخاطئة التي يطلق عليها بالمشتتات أو الموهات. (عوده، ١٩٩٣م: ١٥٥).

### التعريف الإجرائي لاختبار الاختيار من متعدد:

هو اختبار تحصيلي في الوجدتين الأولى والثانية من مقرر مادة الرياضيات للفصل

الدراسي الأول من العام ١٤٢٨هـ - ١٤٢٩هـ لطلاب الصف الثاني المتوسط ويتألف من

ثلاثة نماذج دون تغيير في متون مفرداتها وهي:

النموذج الأول: عبارة عن ٣٠ مفردة من نوع الاختيار من متعدد، لكل مفردة من مفرداته خمسة بدائل، ويطلب من المفحوص أن يختار الإجابة الصحيحة الوحيدة من بين خمسة بدائل لكل مفردة من مفردات الاختبار.

النموذج الثاني: يكون باحتواء كل مفردة من مفردات النموذج الأول على أربعة بدائل، وذلك بحذف أحد الموهات عشوائياً من النموذج الأول.

النموذج الثالث: يكون باحتواء كل مفردة من مفردات النموذج الثاني على ثلاثة بدائل، وذلك بحذف أحد الموهات عشوائياً من النموذج الثاني.

وبالتالي فإن النماذج الثلاثة ثابتة المتن وأن الاختلاف في عدد البدائل فالأول خمسة بدائل والثاني أربعة بدائل وذلك بحذف أحد الموهات عشوائياً من النموذج الأول والنموذج الثالث ثلاثة بدائل بحذف أحد الموهات عشوائياً من النموذج الثاني.

#### • الخصائص السيكومترية:

يقصد بها الباحث معاملات البنود (سهولة المفردة وتمييزها)، ومعاملات الثبات والصدق:

- سهولة المفردة **Item Difficulty**: عرف الدوسري (٢٠٠١م: ١٨٧) "معامل سهولة المفردة بأنه نسبة عدد الإجابات الصحيحة إلى كل الإجابات عليه".

التعريف الإجرائي لسهولة المفردة: يتم حساب معامل سهولة المفردة من خلال إيجاد نسبة الطلبة الذين أجابوا على المفردة إجابة صحيحة من بين الطلاب الذين حاولوا الإجابة

على هذه المفردة، ويحسب معامل الصعوبة بطرح معامل سهولة المفردة من الواحد الصحيح.

وتطلق كلمة السهولة أو الصعوبة في هذا البحث على أنها معامل الصعوبة، ويلاحظ أنه كلما ارتفع مؤشر الصعوبة دل ذلك على سهولة المفردة، وإذا انخفض مؤشر الصعوبة دل ذلك على صعوبة المفردة.

- تمييز المفردة **Item Discrimination**: ذكر علام (٢٠٠٣م: ٤٨٤) "بأن معامل تمييز المفردة يعبر عن مدى إمكانية قياس الفروق الفردية بواسطة مفردات الاختبار".  
التعريف الإجرائي لتمييز المفردة: وهو قيمة معامل الارتباط بين الأداء على الفقرة والدرجة الكلية على الاختبار .

- الثبات **Reliability**: ذكر إسماعيل (٢٠٠٧م: ١٦) " ثبات الاختبار بأنه الحصول على نفس النتائج تقريباً عندما يتم إعادة نفس الاختبار على نفس المجموعة وفي نفس الظروف تقريباً".  
التعريف الإجرائي للثبات : نقصد بهذا المصطلح في هذا البحث الحالي معامل ثبات الاتساق الداخلي للاختبار المحسوب باستخدام معادلة ألفا كرونباخ.

- الصدق **Validity**: عرف أبو لبدة (١٩٨٥م: ٢٤٢) الصدق " بأن يقيس الاختبار بالفعل ما وضع لقياسه".

التعريف الإجرائي للصدق : هو الصدق المرتبط بمحك ( صدق المحك التلازمي - وصدق المحك التنبؤي)، وهو عبارة عن قيمة معامل ارتباط بيرسون المحسوب بين درجات

الطلاب على الاختبار الذي أعده الباحث ، ودرجاتهم في الاختبار النصفى والنهائى لمادة الرياضيات من الفصل الدراسى الأول.

### • الاختبار التحصيلى Achievement Test:

ذكر أبو حطب (١٩٨٧م:٣٨٧) أن الاختبار التحصيلى هو "الأداة التى تستخدم فى قياس المعرفة والفهم والمهارات فى مادة دراسية أو تدريبية معينة أو مجموعة من المواد " بينما يعرف الكبيسى وربيىع (٢٠٠٨م:١٧) الاختبارات التحصيلية بأنها "مقاييس للكشف عن أثر تعلم أو تدريب خاص ويطلق هذا المصطلح على كل الصور والأنواع التى يقوم المعلم بإعدادها من واقع المواد التحصيلية التى درسها الطالب بالفعل".

### التعريف الإجرائى للاختبار التحصيلى :

هو اختبار من إعداد الباحث للصف الثانى المتوسط ، فى الوجدتين الأولى والثانية فى مقرر مادة الرياضيات للفصل الدراسى الأول لعام ١٤٢٨هـ - ١٤٢٩هـ ، ويتألف من ثلاثة نماذج دون تغيير فى متون مفرداتها ، عدد بدائل النموذج الأول خمسة ، والثانى أربعة بدائل ، والثالث ثلاثة بدائل.

## حدود البحث

يتحدد البحث بالحدود التالية :

### الحدود الموضوعية

• اقتصار البحث على الوجدتين الأولى والثانية من مقرر مادة الرياضيات للصف

الثانى المتوسط من العام ١٤٢٨هـ - ١٤٢٩هـ.

- اقتصار البحث على دراسة أثر عدد البدائل على الخصائص السيكومترية للاختبار ومفرداته.

## **الحدود الزمانية**

تحدد بتاريخ تطبيقها في الفصل الدراسي الأول للعام الدراسي ١٤٢٨ هـ -

١٤٢٩ هـ .

## **الحدود المكانية**

محافظة الليث التعليمية التابعة لمنطقة مكة المكرمة، وذلك على عينة من طلاب

الصف الثاني المتوسط من المدارس الحكومية.

•  
:  
:  
•  
:  
•  
:  
•  
:  
:  
:  
:  
:

**أولاً : الإطار النظري**

**أ: الاختبارات التحصيلية:**



تسعى السياسات التعليمية في أي مجتمع لإكساب الطلاب أكبر قدر ممكن من مهارات معينة كالقراءة والكتابة والحساب والقدرة على حل المشكلات والنقد وغيرها، ولمعرفة مدى تحقق هذا الهدف من عدمه كان لابد من اللجوء إلى وسائل لتقيس ما حصله الطالب من هذه المهارات، مما أدى إلى ظهور ما يسمى باختبارات التحصيل الدراسي لتبين حجم ما استوعبه أو حصله كل طالب.

ويشير مصطلح الاختبار في اللغة عند الرازي (١٥٤١هـ: ٧١) إلى "خَبَرَهُ وَاخْتَبَرَهُ إِذَا بَلَاهُ"، أما في المعنى الاصطلاحي فكما ذكر علام (٢٠٠٦م: ٣٠)، "أن براون يرى الاختبار بأنه إجراء منظم لقياس عينة من سلوك الفرد، بينما يرى شيس الاختبار بأنه أسلوب منظم لمقارنة أداء الفرد بمعيار أو مستوى أداء محدد"، ويعني الاختبار عند الكبيسي (٢٠٠٧م: ١٠٧) بأنه "قياس لأداء الفرد في مجالات التعليم وعلم النفس".

أما التحصيل في اللغة فهو تمييز ما يحصل، والحاصل من كل شيء ما بقى وثبت وذهب ماسواه (ابن منظور، ٤١٣هـ)، أما المعنى الاصطلاحي فنجده عند أبوزينة (١٩٩٨م: ١٩) أنه "المعرفة والفهم والمهارات التي اكتسبها المتعلم نتيجة خبرات تربوية محددة"، وعند حمادين (٢٠٠٥م: ١٦٩) فالتحصيل هو "مدى ما تحقق من أهداف التعلم في موضوع أو مساق أو عدة مساقات سبق للطالب دراستها أو التدريب عليها من خلال المشاركة في الأعمال المبرمجة".

أما الاختبار التحصيلي فيعرفه بامشموس وآخرون (١٤٠٥هـ: ١٥٥) بأنه "الأداة التي تستخدم في قياس المعرفة والفهم والمهارات في تعلم مادة دراسية معينة أو مجموعة من ا

لمواد الدراسية بوضع أسئلة تراعي في صياغتها شروط معينة ، بحيث تمر في خطوات تجريبية تصل بها إلى تحقيق أهداف الاختبار" ، ونجد أبو حطب ( ١٩٨٧ م: ٣٨٧) يعرف الاختبار التّحصيلي بأنه " الأداة التي تستخدم في قياس المعرفة والفهم والمهارات في مادة دراسية أو تدريبية معينة أو مجموعة من المواد " ، ويعرف القرني (١٤٢١هـ: ٣) الاختبار التّحصيلي بأنه " عبارة عن أداة من أدوات القياس التي يستخدمها المعلم بطريقة منظمة ، لتحديد مستوى تحصيل الطلاب في مادة دراسية معينة " ، بينما نجد العساف (١٤٢٤هـ: ٤٣٣) يعرف الاختبار التّحصيلي بأنه " الأداة التي صممت لتقدير ما حصل عليه التلميذ من المعلومات التي تعلمها أو المهارات التي تدرب عليها " .

ويتفق الباحث مع الكبيسي وربيح (٢٠٠٨م: ١٧) على تعريفهما للاختبار التّحصيلي بأنه "الأداة للحكم على ماتم تدريسه للتلاميذ من موضوعات تتعلق بمادة دراسية معينة بذاتها وليس ينتظر أو يتوقع تدريسيها لهم " .

### أهمية الاختبارات التّحصيلية:

تكمن أهمية الاختبارات التّحصيلية فيما تزودنا به من معلومات عن الأفراد ، ويذكر علام (٢٠٠٠م: ٣٠٦) في هذا الصدد بأن "الاختبارات التّحصيلية تقيس مدى استيعاب الطلاب لبعض المعارف والمفاهيم والمهارات المتعلقة بالمادة الدراسية ، كما يدل التّحصيل الدراسي على الوضع الراهن لأداء الفرد أو تعلمه أو ما اكتسبه بالفعل في برنامج تعليمي" .

كما تهدف الاختبارات التَّحصيلية عند غنيم (٢٠٠٣: ٧) إلى "قياس مدى تحقيق الأهداف التَّعليمية، وهي إما اختبارات عادية تعد بواسطة المعلمين أو اختبارات عامة تعدها الوزارة في نهاية كل مرحلة دراسية".

### أغراض الاختبارات التَّحصيلية:

ذكر كل من (القرني، ١٤٢٠هـ) و(سماهر، ٢٠٠٥م) عدداً من أغراض الاختبارات التَّحصيلية ، ومنها:

١. التشخيص : أي محاولة تعرف جوانب القوة والضعف لدى الطالب في جانب من جوانب التَّحصيل للاستفادة من النتائج في تدعيم جوانب القوة ومعالجة جوانب الضعف مع ما يستدعيه ذلك من تقويم لأسلوب التدريس أو المناهج أو المرافق التَّعليمية المختلفة ومصادر التعلم ، مع الإشارة إلى أن هناك اختبارات خاصة بالتشخيص ، ولكن هذا لا ينفي عن الاختبارات التَّحصيلية الصفية العادية وظيفة التشخيص.

٢. التصنيف : أي تصنيف الطلاب إلى تخصصات مختلفة : أكاديمي - تجاري - صناعي - زراعي ، وما إلى ذلك ، أو تصنيفهم إلى مجموعات اعتماداً على قدراتهم العقلية أو ميولهم . وبطبيعة الحال لا يكون هذا التصنيف ممكناً إلا بالاعتماد على نتائج الطلاب في اختبارات تحصيلية ، أو اختبارات خاصة ، أو وسائل قياس أخرى من مقابلات شخصية و استبانات وقوائم وما إلى ذلك .

٣. قياس مستوى التَّحصيل : مدى تحقق الأهداف التَّعليمية لدى المتعلم في مادة دراسية بعينها، وفي المواد الدراسية جميعها، و الاختبارات التَّحصيلية في معظمها إنما تنصب

على تحقيق هذا الهدف بقصد الأخذ بنتائجه في تحسين مستوى التعليم وترفيه الطلاب إلى صفوف أعلى وما إلى ذلك.

٤. تنشيط دافعية التعلم ، والنقل من صف إلى آخر ومنح الدرجات والشهادات .

٥. التعرف على مجالات التطوير للمناهج والبرامج والمقررات الدراسية .

### تصنيف الاختبارات التحصيلية :

هناك أكثر من أساس لتصنيف الاختبارات التحصيلية ، ذكرها كل من:

الدوسري، (٢٠٠١م) (مراد، وسليمان ، ٢٠٠٢م) ، ( علام، ٢٠٠٦م) ولعل أهمها ما يلي :

١. التصنيف وفق طريقة أداء الطلبة ( طريقة الاستجابة ) :

#### أ- اختبار شفوي :

وهو اختبار غير مكتوب يقدم للمفحوصين في صورة أسئلة شفوية ويجب عليها

بنفس الطريقة ، وهي من أقدم أنواع الاختبارات وتستخدم في عدد من المجالات كالقراءة

الجهرية وإلقاء الشعر وتلاوة القرآن الكريم .

#### ب- اختبار تحريري :

وهو اختبار مكتوب يقدم للمفحوص في صورة أسئلة تحريرية ويجب عليها كتابة

، والغرض منها معرفة قدرة الطالب على الربط والتنظيم وإدراك العلاقات .

#### ج- اختبار الأداء العملي : ( اختبار المهارات الحركية ):

وهو اختبار يقدم للمفحوص في صورة أسئلة شفوية أو تحريرية ويطلب منه الإجابة عنها في صورة أداءات ( أفعال سلوكية ) يمكن ملاحظتها وإصدار الحكم عليها ، والغرض منها قياس المهارات العملية كإجراء التجارب في المعامل والرسم والموسيقى.

## ٢. التصنيف وفق تفسير درجات الاختبار:

### أ- الاختبارات مرجعية المعيار (NRT) Norm-Referenced Test

وهي تلك الاختبارات التي تقيس أداء الفرد بالنسبة لمتوسط أداء الجماعة التي ينتمي إليها.

وتهدف هذه الاختبارات المرجعة إلى المعيار إلى تحديد مقدار المعلومات التي حصل عليها الطالب في موضوع معين وتدل هذه الدرجة على ترتيبه بالنسبة لباقي أفراد جماعته ، ويكون للدرجة تفسيراً فقط في ضوء معيار الجماعة أي متوسط أداء الجماعة في الاختبار ، وتهتم هذه الاختبارات بإبراز الفروق الفردية بين الأفراد في الصفة المقاسة ، وتستخدم لأغراض التصنيف والقبول في الجامعات والمعاهد واختبارات المهارات الأساسية.

### ب- الاختبارات مرجعية المحك : (CRT) Criterion-Referenced Test

وهي تلك الاختبارات التي تستخدم في تقييم أداء الفرد بالنسبة إلى محك أو أداء مطلق دون الحاجة أو الرجوع إلى أداء الآخرين وتهدف هذه الاختبارات مرجعية المحك إلى التعرف على مستوى التمكن أو الإتقان لدى الأفراد في المعلومات أو المهارات أو القدرات المحددة لهم مسبقاً ، وتستخدم لتحديد جوانب القوة والضعف في تحصيل كل طالب لمجموعة محددة في المعارف والمهارات .

### ٣- التصنيف لبعء محتوى الاختبار :

بعض هذه الاختبارات يكون محدود المحتوى كا اختبار حالة الطقس وبعضها يكون منسقة وعريضة فى محتواها مثل مهارات القراءة وحل المشكلات، وبعضها على شكل بطاريات تقيس عدداً من المواد كا اختبار أيوا للمهارات الأساسية (ITBS) ، واختبارات كاليفورنيا التحصيلية (CAT) .

### ٤- التصنيف وفق موعء التطبيق:

اختبارات أولية ( قبول ) ، اختبارات دورية ( بنائية ) ، واختبارات نهائية.

### ٥- التصنيف وفق طريقة تصميم وبناء الاختبار:

يمكن تصنيف الاختبارات التحصيلية إلى :

#### أ- اختبارات مقننة:

وهي الاختبارات التي تعد بواسطة خبراء القياس ومؤسسات نشر الاختبارات وتستخدم فى كافة المدارس وتتفق مع الأهداف العريضة للمادة الدراسية، وبذلك فهي تقيس معلومات أتفق خبراء التربية على أنها ضرورة لمتوسط الأداء فى مادة دراسية معينة، ويطبق على عينة كبيرة نسبياً من الطلاب لها خصائص معلومة وتتقى بطريقة عشوائية

لإيجاد معايير الاختبار استناداً إلى متوسط أداء هذه العينة وتصبح هذه المعايير الوسيلة التي يقارن في ضوءها أداء كل طالب بمتوسط أداء هذه العينة.

ب- الاختبارات التي يكتبها المعلم لطلابه:

وهي لا تتطلب معايير لأنها تتعلق بالطلاب الذين يقوم المعلم بتدريسهم، وتقيس الأهداف التعليمية المباشرة لما تم تعلمه، لذلك فإن محتواها يكون أكثر تحديداً. (علام ، ٢٠٠٦ م).

٦- التصنيف حسب نمط الإجابة :

أ- الإجابة المختارة ( انتقائية ) **Selection Type** .

ب- الإجابة المصاغة ( منشأة ) **Supply Type** .

أ- الإجابة المختارة ( الإنتقائية ) ومن أكثر هذه الاختبارات

شيوياً:

أولاً : اختبارات الاختيار من متعدد **The Multiple - Choice Tests**

ذكر عوده (١٩٩٣م: ١٥٥) "بأن الفقرة في اختبارات الاختيار من متعدد تتكون 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 وعددا من الإجابات الخاطئة التي يطلق عليها بالمشتمات أو الموهات".

وتعتبر فقرات الاختيار من متعدد من أفضل أنواع الفقرات الموضوعية وأكثرها شيوعاً وانتشاراً؛ خاصة في اختبارات التحصيل المقننة، فهي قادرة على قياس نواتج تعليم في المستويات العقلية العليا من المجال المعرفي بدرجة تفوق فقرات المطابقة، وفقرات الصواب والخطأ، وفقرات التكميل والإجابة القصيرة. (عودة، ١٩٩٣ م).

### أنماط اختبار الاختيار من متعدد:

تتضمن البدائل التي هي بمثابة حلول أو إجابات محتملة للقضية الواردة في المتن إجابة واحدة صحيحة وبقية البدائل تمثل الموهات أو المشتتات، وهناك أنماط وأشكال عديدة لفقرات اختبار الاختيار من متعدد أشهرها:

#### ١. نمط الإجابة الصحيحة:

يعتبر هذا النمط من أبسط الأنماط في اختبارات الاختيار من متعدد، حيث يحتوي السؤال على إجابة واحدة صحيحة تماماً، أما باقي البدائل فتكون خاطئة تماماً، ويرى أبو علام (١٩٨٧م:١٩٥) أن مثل هذه الأسئلة تتعلق بحقائق لاشك فيها مثل: الأسماء والتواريخ والحقائق العلمية والنتائج الحسابية.

#### ٢. نمط الإجابة الخاطئة:

في هذا النمط يحتوي السؤال على إجابة واحدة خاطئة وبقية البدائل صحيحة، ويرى جفري (Jeffrey) "أن هذا النمط يكون مناسباً حينما تكون معرفة الاستثناء لها نفس أهمية معرفة القاعدة". (العريفي، ١٩٤٢٧هـ: ١٩).



### ٣. نمط أفضل الأجوبة:

ذكر أبو لبدة (١٩٨٥م: ٣١٤) أن هذا النوع من الأسئلة قد تكون الأجوبة جميعها صحيحة بصورة جزئية ولكن أحدهما أكثر صحة من البقية أو أفضل من غيره من الأجوبة، وتؤكد ميري (Marie) "على أنه حينما يستخدم هذا النمط لا بد وأن يتضح أن الخيار الصحيح هو الخيار الأفضل" (العريفي، ١٩٨٧هـ: ١٩)، ويمكن أن تستخدم هذه الصورة عند أبو علام (١٩٨٧م: ١٩٦) "لقياس الفهم والتطبيق"، وذكر (علام، ٢٠٠٧م: ٩٩) "أن هذا النوع من المفردات يكون أكثر صعوبة في إجابته من النوع السابق وذلك لأنه يتطلب إجراء تمييزات دقيقة بين البدائل المعطاة".

### ٤. نمط أسوأ الأجوبة:

ويشبه أبو لبدة (١٩٨٥م: ٣١٥) "هذا النمط بالنمط السابق إلا أنه في هذه الحالة يطلب من المفحوصين اختيار أسوأ الأجوبة".

### ٥. نمط يتطلب إكمال فراغات باستخدام بدائل معطاة:

يستخدم هذا النوع من مفردات الاختيار من متعدد بكثرة في اللغات والحساب.

### مزايا اختبارات الاختيار من متعدد :

هناك العديد من مزايا اختبارات الاختيار من متعدد، ونذكر منها :

- ١- القدرة على تمثيل المحتوى بشكل جيد .
- ٢- إمكانية تقدير الإجابة بموضوعية تامة.

٣- سهولة التحكم في صعوبة الاختبار وقدرته على التمييز عن طريق التحكم في مستوى تجانس الخيارات وجاذبيتها.

٤- يمكن استخدامها في قياس مستويات متعددة من التعلم، ومجالات دراسية متنوعة المحتوى.

٥- تعود الطالب على الحكم الصائب، والموازنة وتمييز الأفضل.

٦- تساعد في تشخيص أخطاء الطلبة أو سوء الفهم عندهم من خلال استجاباتهم للمموهات.

٧- تستوجب على الطالب أن يراجع أكبر كمية من المادة المطلوبة .

٨- يمكن تحليل نتائجها إحصائياً بسهولة .

٩- سهولة تصحيحها يدوياً أو باستخدام آلات التصحيح.

١٠- عنصر التخمين فيها أقل من اختبارات الصواب والخطأ. (الدوسري ، ٢٠٠١م)،

( عبد الرزاق ، ٢٠٠٣م)، (الكبيسي، وربيح، ٢٠٠٨م).

**عيوب اختبارات الاختيار من متعدد:**

هناك عدد من عيوب اختبارات الاختيار من متعدد و منها :

١- تستهلك وقتاً وجهداً كبيرين من المعلم في الإعداد للاختبار.

٢- ارتفاع تكاليف الطباعة والتصوير.

٣- لا تصلح لقياس مخرجات التعلم التي تتعلق بالتأليف والتنظيم والابتكار والتي تتعلق بالتعبير الكتابي.

٤- ندرة خلو هذا النوع من الأسئلة من الأخطاء.

٥- يسهل على الطالب فيها الغش والتخمين. (أبو علام، ١٩٨٧م)، (عبد الرزاق، ٢٠٠٣م).

**قواعد صياغة اختبارات الاختيار من متعدد:**

عند كتابة أسئلة الاختيار من متعدد ينبغي مراعاة عدد من القواعد والإرشادات، منها ما يتعلق بأصل السؤال (المتن) ومنها ما يتعلق بالبدائل (المشتتات) ذكرها كل من (أبوعلام، ١٩٨٧)، (الدوسري، ٢٠٠١م)، (القرني، ١٤٢١هـ)، (عبد الرزاق، ٢٠٠٣م)، (المطيري، ١٤٢٤هـ) وهي:

**- قواعد متعلقة بأصل السؤال:**

١- يجب أن يحتوي أصل السؤال على مشكلة محددة وواضحة ومفهومة لدى المفحوص.

٢- يفضل تجنب استخدام صيغ النفي، لأن ذلك يربك المفحوص وإلا يجب وضع خط تحت أداة النفي أو إبرازها حتى يتنبه المفحوص.

٣- استقلالية الأسئلة وعدم اعتماد سؤال على آخر.

٤- يراعى عدم وجود أي تلميح أو إشارة في المتن يمكن أن تشير إلى الجواب أو تدل عليه.

٥- إذا كان السؤال متعلق بتعريف مصطلح معين فمن الأفضل وضع المصطلح في أصل السؤال.

#### - قواعد متعلقة بالبدائل :

١. يفضل عدم استخدام البديل ( كل ما ذكر صحيح ، جميع ما ذكر ) .
٢. أن تكون بدائل السؤال الواحد متجانسة في محتواها ومرتبطة بمجال المشكلة.
٣. يجب أن تكون هناك إجابة واحدة فقط صحيحة.
٤. يجب ألا تكون الإجابة الصحيحة أطول أو أقصر من البدائل الخاطئة بشكل مستمر.
٥. يجب تجنب الارتباطات اللفظية بين أصل السؤال والإجابة الصحيحة .
٦. يجب أن تكون لغة البدائل ومصطلحاتها سهلة وواضحة ومعروفة كلها لدى المتعلمين.
٧. ينوع في ترتيب الإجابات الصحيحة بين البدائل؛ بحيث لا تساعد على التخمين.
٨. يفضل أن يكتب رقم الفقرة بالعدد الحسابي، أما البدائل فبالحروف الأبجدية.
٩. يفضل ترتيب البدائل ترتيباً منطقياً.

ويذكر عودة (١٩٩٣م:١٦٦) " أنه كلما زاد عدد البدائل قلت فرصة التخمين، ففي حالة وجود خمسة بدائل للفقرة تكون نسبة التخمين ٢٠٪، بينما تكون نسبة التخمين ٢٥٪ في حالة وجود أربعة بدائل، و ٣٣٪ في حالة وجود ثلاثة بدائل".

ويذكر دالي (Dale) "أن للتخمين أثراً على ثبات الاختيار، فكلما زاد عدد البدائل في فقرات اختبار الاختيار من متعدد قل تخمين الإجابة الصحيحة، وبالتالي زاد ثبات الاختبار، وانخفض في المقابل مؤشر الصعوبة، وبالتالي يقل متوسط الصعوبة للاختبار". (فقوسة، ١٩٩٤م: ١٥).

### ثانياً - اختبارات الصواب والخطأ True –Fales Tests

يعرف أبو علام (١٩٨٧م:١٨٦) اختبارات الصواب والخطأ بأنها " تلك الاختبارات التي تكون فقراتها مكونة من عبارات أو أسئلة، وتكون استجابة الطالب بالحكم على هذه الاختبارات بالصحة أو الخطأ، ولا يوجد لها غير إجابتين محتملتين فقط".

### ثالثاً: اختبارات المزاوجة (المطابقة) Matching Tests :

ذكر أبو علام (١٩٨٧م: ٢١٦) أن "اختبار المزاوجة يتكون من مجموعة من الأسئلة أو البنود، التي يقابلها ويوازيها مجموعة من البدائل أو الإجابات المحتملة لهذه البنود أو الأسئلة".

ب - الإجابة المصاغة ( المنشأة ):

يذكر عودة(١٩٩٣م:١٥٨) بأن الفقرة ذات الإجابة المصاغة " يطلب من الطالب في هذه الفئة أن يصوغ الإجابة بكلماته وبأسلوبه الخاص" ، وبالتالي فإن الطالب في هذا النوع من الفقرات يقوم بصياغة وكتابة الإجابة بنفسه ، ومن أكثر أنواع الإجابة المصاغة شيوعاً :

#### أولاً: الأسئلة الإنشائية المحددة ( المقيدة ) :

أي أن حرية الطالب في الإجابة حرية مقيدة بطبيعة السؤال نفسه وبالتالي يسهل تحديد زمن الإجابة وحجم الفراغ المتروك لها.

#### ثانياً: الأسئلة الإنشائية المفتوحة :

وهي الأسئلة التي لا يعرف الطالب متى يتوقف عن الإجابة عنها وله الحرية في تقديمها مما يسمح له بترتيب أفكاره والتعبير عنها بحرية أكثر. ( عودة ، ١٩٩٣م )  
الدوسري ، ٢٠٠١م) .

### **ب - خطوات بناء الاختبار التحصيلي الجيد:**

اتفق معظم المختصين بالقياس النفسي والتربوي بأن الاختبارات التحصيلية تمر بعدة خطوات ، وذلك لتحقيق أكبر قدر ممكن من الأهداف التربوية المرجوة ، وتكون عاملاً فاعلاً ومساعداً في قياس النواتج التعليمية ، نورد هذه الخطوات فيما يلي :

أولا : تحديد الغرض من الاختبار:

على المعلم تحديد الغرض من الاختبار التحصيلي قبل البدء في إعداده، فمنها ما يكون للقبول والانتقاء والتصنيف في مجموعات، ومنها ما يكون لتشخيص الصعوبات التي تواجه المتعلمين ودعم خطط التوجيه والإرشاد، ومنها ما يستخدم في البحوث التربوية والتطوير التربوي، ومنها ما يتعلق بقياس مستوى التحصيل ودرجة النمو العلمي للمتعلم واتخاذ القرارات الإدارية المختلفة في ضوء ذلك، ويساعد تحديد الغرض من الاختبار في توجيه الخطوات اللاحقة كتحديد نوع الفقرات وتوقيت الاختبار وما إلى ذلك (سنان، ٢٠٠٦م)، و(الكبيسي، ٢٠٠٧م)، و(الحريري، ٢٠٠٧م).

وحدد الباحث غرضه من بناء اختبار تحصيلي، وهو تطوير أدوات قياس التحصيل الدراسي والوصول إلى العدد الأمثل من البدائل المستخدمة في اختبارات الاختيار من متعدد والتي يكون صدقها وثباتها عالياً في مادة الرياضيات.

#### ثانياً: تحديد الأهداف التعليمية :

الأهداف التعليمية هي مجموعة من العبارات أو الصياغات التي توضح ما سوف يكون عليه سلوك التلميذ بعد اكتسابه للخبرة التعليمية داخل وخارج جدران المؤسسة التربوية (خضر، ٢٠٠٤م)، وتحديد نواتج التعلم المخطط لها قبل البدء بعملية التدريس يعد من أكثر الخطوات أهمية بالنسبة لمعد الاختبار التحصيلي، ويكون ذلك من خلال صياغة الأهداف السلوكية صياغة إجرائية محددة وواضحة بعيداً عن الغموض والعمومية.

وقد صنف بلوم عام ١٩٥٦م الأهداف السلوكية إلى ثلاثة مجالات هي :

## أ- المجال المعرفي:

ويتضمن الأهداف المتعلقة باكتساب المعارف وتطبيقها وفهمها ، وتقع الغالبية العظمى من الأهداف التربوية والتدريسية ضمن هذا المجال ، وقسمها بلوم إلى ستة مستويات فرعية ، وهي :

١. التذكر: وتتمثل في قدرة الطالب على تذكر المعارف أو المعلومات السابقة

واسترجاعها ، وهي أدنى مستويات المجال المعرفي ، ومن الأفعال السلوكية التي

تستخدم في صياغة هذا المستوى: يذكر، يسمي، يصف، يعدد.

٢. الفهم: وهو القدرة على إدراك واستيعاب معنى المادة أو النص الذي يدرسه

الطالب ، ومن أفعاله السلوكية التي تستخدم في صياغة هذا المستوى: يفسر،

يترجم، يوضح، يعطي، يستخرج.

٣. التطبيق : ويعني استخدام ما تم تعلمه من قوانين ومفاهيم أو قواعد أو مبادئ أو

تعليمات أو إجراءات وغيرها في مواقف جديدة أو مواقف واقعية ، ومن الأفعال

السلوكية التي تستخدم في صياغة هذا المستوى: يطبق، يحل، يربط، يشكل،

يصحح.

٤. التحليل: ويعني القدرة على تجزئة أو تحليل المادة إلى عناصرها أو مكوناتها،

ومن الأفعال السلوكية التي تستخدم في صياغة هذا المستوى: يحلل، يصنف،

يستنبط، يميز.



٥. التركيب : ويعني القدرة على جمع عناصر أو أجزاء لتكوين كل متكامل أو

نمط أو تركيب غير موجود أصلاً، ومن الأفعال السلوكية التي تستخدم في

صياغة هذا المستوى: يركب، يؤلف، يصمم، يكون، يبني.

٦. التقويم : وهو أعلى مستويات المجال المعرفي، ويعني القدرة على إصدار أحكام

حول قيمة الأفكار أو اتخاذ قرارات مناسبة، ومن الأفعال السلوكية التي

تستخدم في صياغة هذا المستوى: يقوم، يحكم، يقارن، يعطي الطالب

حكماً، يصدر الطالب حكماً، ينتقد. (فتح الله، ٢٠٠٠م)، (الحريري، ٢٠٠٧).

#### ب - المجال الوجداني :

ويهتم هذا المجال بكل ما يرتبط بالعاطفة الإنسانية من قيم واتجاهات

وميل وتقبل واستجابة. (زيتون، وعبدالحميد، ١٩٩٥م).

#### ج - المجال النفس حركي ( المهاري ):

يتضمن هذا المجال الأهداف التي تؤكد على نتائج التعلم ذات العلاقة

بالمهارات الحركية، ويشتمل هذا المجال على عدة مستويات: الملاحظة

والتقليد، التجريب والممارسة، الإتقان والإبداع.

(زيتون، وعبدالحميد، ١٩٩٥م)، و(الحريري، ٢٠٠٧م).

#### ثالثاً - تحليل المحتوى :

يقصد بتحليل المحتوى الوصول إلى مفردات المقرر الدراسي، أو إحصاء المعلومات

الأساسية في المقرر الدراسي أو هو إحصاء المعارف والمهارات الأساسية المتضمنة في

الدروس، وذلك بهدف ضمان تمثيل فقرات الاختبار لمفردات المحتوى تمثيلاً متكافئاً. (الكبيسي، ٢٠٠٧م)، و(الحريري، ٢٠٠٧م).

#### رابعاً - إعداد جدول المواصفات:

جدول المواصفات عبارة عن: مخطط ثنائي الطريقة يتحدد فيه عدد الأسئلة في كل خلية بناءً على المحتوى والهدف (سيد، وسالم، ٢٠٠٥م)، ويميل الباحث إلى التعريف الذي أورده الكبيسي (٢٠٠٧م: ١٤١) بأن جدول المواصفات هو "مخطط تفصيلي يمثل محتوى الاختبار، ويربط محتوى المادة الدراسية بالأهداف التعليمية السلوكية، وبين الوزن النسبي الذي يعطيه المعلم لكل موضوع من الموضوعات المختلفة، والأوزان النسبية للأهداف المعرفية السلوكية في مستوياتها المختلفة".

ويتكون هذا الجدول من بعدين الأول منها أفقي ويمثل الأهداف التعليمية السلوكية، والآخر رأسي ويمثل موضوعات المادة الدراسية، كما تشتمل حقول الجدول على أوزان الأهمية النسبية لكل من الموضوعات والأهداف، وكذلك عدد الأسئلة التي يضعها المعلم في كل موضوع على ضوء تلك الأوزان، وبإمكان المعلم أن يضع الدرجة المستحقة لأسئلة كل موضوع في الجدول نفسه.

ويذكر الرادادي (٢٠٠٦م: ٤٦) "أن بعض حقول جدول المواصفات قد تبقى فارغة، فليس بالضرورة أن كل موضوعات المادة تشتمل على أهداف سلوكية في جميع المستويات، وخصوصاً المستويات العليا".

وهناك عدد من الخطوات اللازمة لبناء جدول المواصفات

ذكرها (الحريري، ٢٠٠٧م)، و(الكبيسي، ٢٠٠٧م) تتلخص فيما يلي:

أ. تحديد عناصر المحتوى الذي سيجرى قياسه.

ب. تحديد الأهداف التعليمية للمادة الدراسية لمعرفة مدى تحققها.

ج. تحديد الوزن النسبي للأهداف السلوكية بمستوياتها المختلفة من خلال المعادلة

رقم (١):

عدد الأهداف في ذلك المستوى

وزن الهدف في مستوى معين =  $\frac{\text{عدد الأهداف في ذلك المستوى}}{\text{مجموع أهداف المادة كاملة}} \times 100$  - - - (١)

مجموع أهداف المادة كاملة

د - تحديد عدد الحصص اللازمة لتدريس كل موضوع

هـ - تحديد الوزن النسبي لموضوعات المادة الدراسية من خلال المعادلة رقم (٢):

عدد حصص تدريس الموضوع

وزن الموضوع =  $\frac{\text{عدد حصص تدريس الموضوع}}{\text{عدد الحصص الكلية للمادة الدراسية}} \times 100$  - - - (٢)

عدد الحصص الكلية للمادة الدراسية

و - تحديد العدد الكلي لأسئلة الاختبار في ضوء الزمن المتاح للإجابة، ونوع الفقرة

المستخدمة في الاختبار، وأعمار الطلبة ومستواهم التعليمي، وطول الفقرة ومستوى

صعوبتها، ومقدار ما تتطلبه الفقرات من حساب كمية ومستوى العمليات العقلية

التي تقيسها الفقرة.

ز - حساب عدد الأسئلة في كل موضوع لكل مستوى من مستويات الأهداف من خلال المعادلة رقم (٣):

$$\text{عدد أسئلة الموضوع} : \frac{\text{وزن الموضوع}}{100} \times \frac{\text{وزن الهدف}}{100} \times \text{العدد الكلي للأسئلة} - - - (3)$$

- تحديد درجة أسئلة كل موضوع في كل مستوى من مستويات الأهداف من خلال المعادلة رقم (٤) :

$$\text{درجة أسئلة الموضوع} = \frac{\text{وزن الموضوع}}{100} \times \frac{\text{وزن الهدف}}{100} \times \text{الدرجة النهائية للاختبار} - - - (4)$$

ولجدول المواصفات عدد من الفوائد منها:

أ. إعطاء كل جزء من المقرر حقه في الأسئلة بما يتناسب مع حجم الجهود المبذولة في تدريس كل موضوع.

ب. تحقيق صدق المحتوى بنسبة كبيرة.

ج. إكساب الطالب ثقة كبيرة وشعور إيجابي بعدالة الاختبار وشموليته لجميع أجزاء

المقرر الدراسي، مما يساعده في تنظيم وقته وتوزيع جهده أثناء المذاكرة لجميع

أجزاء المنهج. (كاظم، ٢٠٠١م)، و(الردادي، ٢٠٠٦م).

خامساً - اختيار نمط أسئلة الاختبار:

اختيار النمط المناسب للاختبار يتوقف على مجموعة من العوامل، ومنها:

١. الهدف المراد قياسه: إذا كان الهدف الكشف عن الحقائق والنظريات فالأفضل أن يختار المعلم أسئلة الخطأ والصواب، ولقياس الأهداف التربوية المتقدمة كالفهم والاستيعاب فيفضل في هذه الحالة الأسئلة ذات الاختيار من متعدد، بينما تستخدم الأسئلة المقالية لقياس الحقائق والأفكار المترابطة.
٢. طبيعة المحتوى المراد قياسه: من الممكن استخدام جميع أنماط الفقرات في أي مادة دراسية إلا أن بعض تلك الأنماط تكون أكثر ملائمة لمحتوى مادة معينة، كمادة النصوص الأدبية التي تناسبها الأسئلة المقالية، ومادة الرياضيات التي تناسبها الأسئلة الموضوعية وهكذا.
٣. مهارة المعلم في كتابة الأنواع المختلفة من الأسئلة: يتباين المعلمون في نوعية المهارات التي يمتلكها كل منهم في وضع الأسئلة، فمن يمتلك مهارة عالية يضع خليطاً من الأسئلة المقالية والموضوعية، أما من يفتقد إلى هذه المهارة ويتعاس في بذل الجهد والوقت لإعداد الأسئلة الموضوعية الجيدة فإنه يضطر إلى وضع أسئلة مقالية.
٤. عمر التلميذ ومستواه التعليمي: يقل عدد الأسئلة لدى طلاب المراحل الدنيا، وذلك لأنهم يحتاجون إلى وقت أطول في الاستجابة لفقرة الواحدة مما يحتاجه طلاب المراحل العليا.

٥. الوقت المخصص لتحضيرها: يفضل وضع أسئلة الاختيار من متعدد عندما يكون

الوقت كافياً، وعندما يكون الوقت غير متاح تستخدم الأسئلة المقالية وأسئلة

التكميل والإجابة القصيرة.

٦. توفر الإمكانيات المادية: ارتفاع تكاليف طباعة وتصوير الأسئلة الموضوعية

بعكس الأسئلة المقالية، وبالتالي يتوقف استخدام الأسئلة الموضوعية على توفر

الإمكانيات المادية. (الزيود، وعليان، ٢٠٠٢م)، و(الحريري، ٢٠٠٧م).

جميع العوامل السابقة ينظر إليها بصورة كلية ، ثم يتخذ بعد ذلك القرار باختيار

النمط المناسب للاختبار، أما في حالة تساوي تأثير هذه العوامل على أنماط مختلفة من

الاختبارات فإن اختيار أنسبها يعتمد على الخصائص السيكومترية التي تتعلق بكل نمط

من الاختبارات، وعلى معد الاختبار أن يختار النمط الأكثر صدقاً وثباتاً وتمييزاً مما

يزيد من تحقيق الاختبار لأهداف القياس التربوي.

وهذا الاختيار يجب أن يكون مبنياً على أساس علمي من خلال الدراسات في هذا

المجال ومنها هذه الدراسة، حيث قارن الباحث بين ثلاث نماذج من نماذج الاختيار من

متعدد (خمسة بدائل، وأربعة بدائل، وثلاثة بدائل) من حيث أثر كل منها على الخصائص

السيكومترية للاختبار وفقراته.

سادساً - مراجعة أسئلة الاختبار:

اقترح تتكلمان (Tinkleman, ١٩٧١) أن تتم مراجعة الأسئلة في ضوء ثلاثة

معايير، وهي:

- ١- المحتوى المتعلق بالمادة الدراسية ومدى صحته وصحة نموذج الإجابة .
- ٢- الجانب التحريري وسلامة الصياغة ، ومدى التناسق في الأسلوب والشكل مع بقية الأسئلة .
- ٣- الناحية الفنية المتعلقة بالجوانب القياسية في السؤال وخصائص نوعه .

ومراجعة الأسئلة في ضوء هذه المعايير تتطلب توفر الخبرة في مجال القياس والتقييم والاختصاص في المادة الدراسية، وذلك لتحكيم الأسئلة ومدى ملائمتها لمواصفات الاختبار.(الدوسري، ٢٠٠١م).

وقام الباحث بعرض أداة البحث وهي اختبار تحصيلي من نوع الاختيار من متعدد على عدد من المتخصصين في القياس والتقييم، ومن مشرفي ومعلمي مادة الرياضيات، وأيضاً تمت مراجعتها لغوياً من أحد المتخصصين في اللغة العربية.

سابعاً- ترتيب أسئلة الاختبار :

عند صياغة فقرات الاختبار بصورتها النهائية يتم ترتيبها، ولا يوجد نظام محدد للترتيب وعلى المعلم أن يختار التنظيم المناسب لتحقيق أهداف الاختبار، وهناك أمور عدة يمكن أخذها بعين الاعتبار عند ترتيب أسئلة الاختبار، وهي :

- ١- نمط المفردة: إذا احتوى الاختبار على أكثر من نمط من أنماط الفقرات فمن المفضل ترتيبها حسب النمط والفقرات ذات النمط الواحد مع بعضها البعض.
- ٢- صعوبة المفردة: ويتم ذلك بأن تتدرج الأسئلة في المجموعة الواحدة من السهل إلى الصعب حيث يوفر هذا الترتيب الدافعية للطالب للاستمرار في محاولة الإجابة عن الأسئلة الأولى السهلة .
- ٣- المحتوى: ويقصد بهذا الترتيب تسلسل الفقرات في السؤال الواحد تسلسلاً منطقياً لمحتوى المادة الدراسية.
- ٤- مستوى الأهداف: وذلك بوضع الفقرات التي تقيس مستوى معيناً من الأهداف بشكل متسلسل، مع إمكانية ملاحظة أن مستوى الأهداف قد يستخدم لتدرج الفقرات بناء على صعوبتها. (الشيخ وآخرون ، ٢٠٠٤م)، (الردادي، ٢٠٠٦م).

#### ثامناً - صياغة تعليمات الاختبار:

تمثل التعليمات إرشادات هامة توجه التلميذ وترشده إلى كيفية الاستجابة لفقرات الاختبار، وتلعب التعليمات الواضحة والمفهومة دوراً كبيراً في تقليل أسئلة التلاميذ واستفساراتهم داخل قاعة الاختبار وتنقسم التعليمات إلى نوعين:



١. التعليمات الخارجية : وهي ورقة منفصلة توضع فوق أوراق أسئلة الاختبار، وعدد الأسئلة وأنواعها، وزمن الاجابة المخصص للاختبار، وطريقة ومكان الإجابة، وبعض الأمثلة التوضيحية وإمكانية الإجابة عن بعض الأسئلة وترك البعض الآخر، وبيان عقوبة الغش.

٢. التعليمات الداخلية: وتوضع قبل كل سؤال، لتوضيح المطلوب من السؤال مع ذكر مثال لكيفية الإجابة. (الصمادي، والزرايع، ٢٠٠٤ م)، و(سيد، وسالم، ٢٠٠٥ م).

وذكر كاظم (٢٠٠١م: ٩٢) في كتابة التعليمات مايلي:

- أ- أن تكون كلمات التعليمات واضحة ومفهومة وبجمل قصيرة.
- ب- أن تكون متسلسلة على شكل نقاط، توضح ماهو مهم.
- ج- عدم وضع افتراض على أن التلاميذ يعرفون طريقة الاستجابة استناداً إلى خبراتهم السابقة، بل يفترض أنهم يؤدون الاختبار لأول مرة.

تاسعاً - إخراج الاختبار وطباعته:

يتكون الاختبار عادة من مجموعة من الأوراق (كراسة الاختبار)، والتي يجب

على المعلم عند إخراجها ملاحظة أمور كثيرة، أهمها:

- ١- أن تحتوي كراسة الاختبار على تعليمات الإجابة عليه .

- ٢- أن تكون طباعة الأسئلة واضحة خالية من الأخطاء المطبعية والإملائية.
- ٣- أن يراعي الفصل بين التعليمات والأسئلة.
- ٤- أن يراعي الفصل بين كل سؤال والذي يليه بمسافة معقولة.
- ٥- أن لا يجزئ السؤال على صفحتين متتاليتين .
- ٦- أن يفصل بين كل نوع أو شكل من أشكال الأسئلة والشكل الآخر بخط.
- ٧- أن يراعي الفصل بين مقدمة السؤال وبدائله بمسافة معقولة .
- ٨- يراعي الترتيب المنطقي لأوراق الاختبار، وتشبيتها بشكل يسهل تناول الاختبار وتداوله. (الردادي، ٢٠٠٦م).

## عاشراً - تجريب الاختبار :

ذكر كاظم (٢٠٠١م:٩٧) بأنه يطلق على التجربة الأولية للاختبار التجريب الاستطلاعي، ويتم بتطبيق الاختبار على عينة صغيرة من التلاميذ بهدف التعرف على مدى وضوح التعليمات والأسئلة، ومدى مناسبتها لمستوى التلاميذ بشكل عام، والزمن اللازم للإجابة عليه، وبعد ذلك تجرى التعديلات اللازمة.

ويشير الدوسري (٢٠٠١م:٢٠٧) إلى أن الاعتماد على نتائج اختبارات لم يتم تجريب أسئلتها يؤدي إلى ضعف مستواها من حيث الصدق والثبات ، وبالتالي عدم مقدرة تلك الاختبارات على تحقيق أغراضها، ويذكر أن هناك ثلاثة طرق لتجريب الاختبار ، وهي :

- ١- تجريب الاختبار على أساس الاختبار الفعلي .

٢- تجريب الاختبار على أساس اختبار تجريبي .

٣- دمج الأسئلة ضمن اختبار فعلي .

وقد قام الباحث بتجربة استطلاعية للاختبار على عينة عشوائية من مجتمع الدراسة ، مع ملاحظة استبعاد هذه العينة من التطبيق الفعلي للاختبار، وفي ضوء هذه التجربة أجريت التعديلات اللازمة.

#### أحد عشر - تطبيق الاختبار الفعلي :

إن الظروف التي يتم فيها إجراء الاختبار - سواء أكانت نفسية أم بيئية - تؤثر في نتائج الطلاب تأثيراً كبيراً ، ولا بد من توفير ظروف ملائمة عند إجراء الاختبار، فإذا تم ضبط ظروف التطبيق فإننا نضبط واحداً من مصادر أخطاء القياس؛ حتى تكون الفروق بين درجات الطلاب فروقاً حقيقية ولا تعزى إلى ظروف إجراء الاختبار ومدى تفاعل الطلاب مع هذه الظروف. (الردادي، ٢٠٠٦م).

والأفضل أن يعقد الاختبار في نفس المكان الذي كان يتلقى التلاميذ فيه دروسهم، مع ضرورة بعده عن الضوضاء والضجيج، وأن تتوفر فيها الإضاءة والتهوية الكافيتين، مع ضرورة تواجد معلم المادة الدراسية يوم الاختبار لتوفير نوع من الراحة النفسية والطمأنينة لدى التلاميذ، ويجب عدم تشتيت انتباه التلاميذ وتركيزهم ويستحسن تذكيرهم بالوقت المتبقي للاختبار بين الحين والآخر، وينبغي عدم التهاون بالغش بجميع

أشكاله، لأنه يؤثر على مصداقية النتائج، ولذلك يجب وضع مسافات كبيرة بين التلاميذ عند تطبيق الاختبار ومع ضرورة المراقبة المشددة. (كاظم، ٢٠٠١م)، و(الصمادي، والزرايع، ٢٠٠٤م).

## الثاني عشر - تصحيح الاختبار :

تعتمد طريقة التصحيح على نمط الفقرات، فعندما تكون الفقرات ذات استجابة مختارة أمكن تصحيحها يدوياً عن طريق عمل ورقة إجابة؛ وقد يكون التصحيح آلياً، حيث يستخدم في هذا النوع من التصحيح آلات إلكترونية مبرمجة لهذا الغرض وفي هذه الحالة يمكن استخدام نماذج إجابة تكون مصممة مسبقاً وذات مقاسات معينة تتناسب مع أجهزة التصحيح.(الحريري، ٢٠٠٧م).

أما إذا كانت الفقرات ذات استجابة منشأة فهناك طريقتان للتصحيح، هما: الطريقة التحليلية، والطريقة الكلية، واستخدام أي منهما يعتمد على عوامل كثيرة، كالمهدف من الاختبار، والوقت المتوفر لدى المعلم للقراءة، ونوع السؤال فيما إذا كان محدداً أو مفتوحاً.

وفي الطريقة التحليلية تجزأ الإجابة على عناصر، وتكون درجة السؤال هي مجموع درجات تلك العناصر؛ إذ أن المصحح أثناء قراءته يحدد الدرجة التي يستحقها كل عنصر من عناصر الإجابة .

أما الطريقة الكلية فيتم إعطاء الدرجة للتلميذ بناء على الانطباع العام الذي تشكله إجابة التلميذ لدى المصحح، وتكون درجة الثبات متدنية لأن الدرجة تمنح حسب ذاتية المصحح والعوامل النفسية أو المزاجية الخاصة به. (كاظم، ٢٠٠١م)، و(الدوسري، ٢٠٠١م).

### الثالث عشر- تحليل فقرات الإختبار :

عرف النبهان (٢٠٠٤م: ١٩١) تحليل الفقرات بأنه "عملية تتعلق باستقصاء الخصائص الإحصائية لاستجابات الطلبة على كل فقرة من فقرات الاختبار، بينما نجد الكبيسي (٢٠٠٧م: ١٦٨) يقصد بتحليل فقرات الاختبار بأنه "استخراج معاملي الصعوبة والتمييز، وتحديد فاعلية البدائل (الموهات)".

وبذلك تكون أهم الخصائص الإحصائية لمفردات الاختبار :

(١) صعوبة المفردة Item Difficulty ، ويرمز لها بالرمز (P) .

(٢) تمييز المفردة Item Discrimination ، ويرمز لها بالرمز (D) .

(٣) جاذبية البدائل (الموهات) Distractor Attractivness ، ويرمز لها بالرمز

(DA) .

فوائد تحليل المفردات:

أورد علام(٢٠٠٦م: ١١٢)، والكبيسي (٢٠٠٧م: ١٦٨) فوائد عدة لتحليل مفردات الاختبار، ومنها:

١- تحديد مواطن الضعف والقوة عند المتعلمين وتشخيص الأسباب سواء كانت فنية في الاختبار أم تربوية في طريقة التدريس.

٢- تنمية مهارات المعلمين والقائمين بإعداد الاختبارات، نتيجة التدريب والفحص المستمر للمفردات، والتخلص من الأخطاء والعيوب التي تصيب الفقرة من حيث الصياغة أو الأخطاء الفنية الأخرى .

٣- الحصول على مفردات جيدة من حيث الصعوبة والتمييز والاحتفاظ بها، وإعادة استخدامها في الاختبارات القادمة.

٤- التعرف على درجات صعوبة الفقرات يحدد مواقع الفقرات في الاختبار ووضع كل فقرة في المكان المناسب لها .

٥- تحديد فعالية البدائل وإعادة النظر في الموهبات التي يختارها أفراد المجموعة العليا أكثر من الدنيا أولم يختارها أحد من الطلبة.

٦- يمكن أن تكون خبرة تعليمية جيدة للطلبة إذا أسهم الطلبة في بحث وتحليل الفقرات؛ لأنها تخلق لديهم نوعاً من الدقة والانتباه في فحص البدائل قبل اختيار الجواب.

أولاً: معامـل السهولة:

أورد علام(٢٠٠٧م: ٢٥١) بأن "سهولة المفردة يقصد بها النسبة المئوية لعدد الطلاب الذين أجابوا على المفردة إجابة صحيحة، وتكون المفردة صعبة جداً إذا أجاب عليها جميع الطلاب بإجابة خاطئة وتكون هذه النسبة صفراً، وعند إجابة جميع الطلاب بإجابة صحيحة تكون هذه المفردة سهلة جداً عنده تكون النسبة المئوية ١٠٠٪، وفي الحالتين لا تفيد في الكشف عن الفروق الفردية بين الطلبة في امتلاكهم للصفة أو السمة التي يتوقع من الاختبار قياسها".

وبمعنى آخر يتم حساب معامل سهولة المفردة بحاصل قسمة عدد الطلاب الذين أجابوا عن المفردة إجابة صحيحة على العدد الكلي للأفراد الذين حاولوا الإجابة عن تلك المفردة.

ويتفق الباحث مع ما ذكره نيتكو (Nitko) "بأن مدى معاملات السهولة المحصورة بين (٠,٢٠) و (٠,٨٠)، بمتوسط قدره (٠,٥٠) يمكن أن يكون مقبولاً". (النبهان ، ٢٠٠٤م: ١٩٢).

بينما ذكر علام(٢٠٠٧م: ٢٥٣) "بأن مدى معاملات السهولة المحصورة بين (٠,٣٠) و (٠,٧٠) تكون مقبولة، ويفضل أن تكون قيمته (٠,٥٠) حيث أن هذه القيمة تجعل التمييز بين المجموعتين العليا والدنيا من الطلاب أكبر ما يمكن".

علماً بأن المفردات التي تقع خارج هذا المدى قد تكون مفردات مناسبة من الناحية العملية على الرغم من أنها غير مناسبة من الناحية الإحصائية؛ فقد يلجأ واضع الاختبار

إلى وضع مفردات سهلة لأغراض تشجيع الطلبة أو اكتشاف قدراتهم المتقدمة على مفردات أصعب نسبياً. (علام، ٢٠٠٧م).

وهناك عدد من الشروط الواجب توافرها عند حساب معامل السهولة ذكرها النبهان (٢٠٠٤م: ١٩٢) وهي:

- ١- أن يكون تصحيح المفردة بواقع (٠ و ١).
- ٢- تحديد المجموعة المرجعية التي حسبت قيمة معامل سهولة المفردة بالاعتماد عليها؛ إذ أن سهولة المفردة تعتمد على خصائص المجموعة، وليس صفة مطلقة من خصائص المفردة، إذ قد تكون المفردة صعبة بالنسبة لطلبة الصف الخامس، وقد تكون المفردة سهلة بالنسبة لطلبة الصف السادس.

طرق حساب معامل السهولة:

أ- في حالة النسب المئوية نستخدم المعادلة رقم (٥):

عدد الإجابات الصحيحة

ص

معامل السهولة =  $\frac{\text{عدد الإجابات الصحيحة}}{\text{عدد الإجابات الصحيحة} + \text{عدد الإجابات الخاطئة}}$  =  $\frac{\text{ص}}{\text{ص} + \text{خ}}$  (٥)

عدد الإجابات الصحيحة + عدد الإجابات الخاطئة

ص + خ

(إسماعيل، ٢٠٠٧م)

ب- في حالة الترتيب:



نستخدم المجموعتين الطرفيتين وفيها يتم استخراج معامل سهولة المفردة بالخطوات التالية :

١- ترتيب درجات الطلبة الكلية ترتيباً تنازلياً.

٢- تحديد الفئة العليا (٢٧٪)؛ وهم من حصلوا على أعلى الدرجات، والفئة الدنيا

(٢٧٪)؛ وهم من حصلوا على أدنى الدرجات، تطبيق المعادلة التالية معادلة رقم (٦):

$$\text{معامل السهولة} = \frac{ن_١ + ن_٢}{ن_٢} - - - (٦)$$

حيث أن :

ن<sub>١</sub> = عدد الطلبة الذين أجابوا عن المفردة إجابة صحيحة من الفئة العليا.

ن<sub>٢</sub> = عدد الطلبة الذين أجابوا عن المفردة إجابة صحيحة من الفئة الدنيا .

ن = عدد الطلبة في إحدى المجموعتين .(إسماعيل، ٢٠٠٧م)

وحساب معامل السهولة بهذه الطريقة يتجاهل درجات المجموعة الوسطى والتي

تشكل حوالي (٤٦٪) من مجمل الطلبة .

ثانياً: معامل التمييز:

أورد علام(٢٠٠٧م: ٢٥٤) تعريفاً لمعامل تمييز المفردة بأنه " يشير إلى درجة تمييز

المفردة بين مرتفعي التحصيل ومنخفضي التحصيل من الطلاب في الاختبار".

وبمعنى آخر تكون الفقرة مميزة جداً إذا كان أداء الطلبة عليها يرتبط ارتباطاً

موجباً بأدائهم على الاختبار.

ويذكر(إسماعيل،٢٠٠٧م: ١١) أن حدود معامل التمييز يتراوح بين - ١ و + ١ ،

والحد الأدنى لمؤشر التمييز هو ٠,٣٠ على الأقل ، وعندما تصل إلى ٠,٤٠ فأكثر فإنها

تدل على قوة تمييز بين المجموعتين، ومن ٠,٢٠ إلى ٠,٣٠ تعتبر قوة تمييز منخفضة ولكن

لا بأس بها، ويعتبر التمييز ضعيفاً وتحتاج الفقرة إلى تعديل إذا تراوح مؤشر التمييز بين

٠,١٠ إلى ٠,٢٠ ، بينما تحذف الفقرة أو تعدل جذرياً إذا كان مؤشر التمييز أقل من ٠,١٠.

طرق حساب معامل التمييز :

أ- طريقة المجموعتين الطرفيتين:

عند استخدام هذه الطريقة يتطلب أولاً ترتيب الطلبة ترتيباً تنازلياً حسب درجاتهم

الكلية ، ومن ثم تحديد أعلى ٥٠% وأقل ٥٠% إذا كان عدد الطلبة قليلاً، أو أعلى ٢٧%

وأقل ٢٧% إذا كان العدد كبيراً .

ويتم حساب معامل التمييز بهذه الطريقة باستخدام المعادلة التالية، معادلة رقم (٧):

$$\text{معامل التمييز} = \frac{\text{ص ع} - \text{ص د}}{\text{٠,٢٧} \times \text{العدد الكلي}}$$

(٧) - - -

حيث :

ص ع : هي عدد الإجابات الصحيحة في المجموعة العليا.

ص د : هي عدد الإجابات الصحيحة في المجموعة الدنيا. (إسماعيل، ٢٠٠٧م).

ب - الطريقة الارتباطية (ويسمى في هذه الحالة صدق المفردة) :

وتتلخص هذه الطريقة في حساب معامل الارتباط بين الدرجة على الفقرة من جهة والدرجة الكلية على الاختبار من جهة أخرى ؛ حيث إن معامل التمييز الناتج بهذه الطريق يأخذ بعين الاعتبار استجابات جميع الطلبة ، وذلك باستخدام معامل الارتباط الثنائي الأصيل، أو باستخدام معامل الارتباط الثنائي المتسلسل الحقيقي (Point Biserial Correlation)، وذلك عندما تشكل الدرجات على المفردات متغيراً ثنائياً (٠ و ١) وتشكل الدرجات ككل متغيراً متصلًا، وعند استخدام معامل الارتباط الثنائي الأصيل نقوم بتطبيق المعادلة رقم (٨):

$$r_{\text{ب}} = \frac{(M_1 - M_2) \times A}{\sigma} \quad \text{ب} \times \text{أ} \quad \text{ث} \quad \text{ع}$$

حيث :

م<sub>١</sub> : هي متوسط درجات الأفراد الذين أجابوا إجابة صحيحة على المفردة.

م<sub>٢</sub> : هي متوسط درجات الأفراد الذين أجابوا إجابة خاطئة على المفردة.

ب : معامل السهولة.

أ ث : معامل الصعوبة.

ع : الانحراف المعياري. (إسماعيل، ٢٠٠٧م).

ثالثاً: معامل جاذبية (فاعلية) المموهات:

المموهات أو المشتتات: هي البدائل أو الخيارات غير الصحيحة في سؤال الاختيار

من متعدد.

ويفترض في المموه أن يكون جذاباً للمفحوصين وخاصة ممن ينتمون إلى الفئة

الدنيا، وأن لا تقل نسبة الطلبة الذين يختارونه عن ٥٪ منهم ، أكثرهم من فئة مجموعة

الأداء المنخفض .

والمموه الجيد والفعال هو ذلك البديل الذي يتمتع بمعامل جاذبية سالب وكبير؛ إذ

يجب مراجعة أو تطوير أو استبدال أي بديل لا يختاره أحد من الطلبة، أو أن تكون

جاذبية موجبة أو صفراً.

طريقة حساب معامل جاذبية المموه :

يمكن حساب معامل جاذبية المموه بنفس طريقة حساب معامل تمييز الفقرة

حسب المعادلة رقم (٩):

$$\text{معامل فاعلية المموه} = \frac{\text{ن ع} - \text{ن د}}{\text{---(٩)}}$$

ن

حيث أن :

ن ع : عدد الذين اختاروا المشتت من المجموعة المرتفعة.

ن د : عدد الذين اختاروا المشتت من المجموعة الدنيا.

ن : عدد الأفراد في إحدى المجموعتين. (إسماعيل، ٢٠٠٧م).

### ج: شروط الاختبار الجيد :

يرى كل من ثورانديك وهيجن (١٩٨٦م:٥٤) وعودة (١٩٩٣م:٣٣٤) أن الاختبار

الجيد لابد أن يتصف بعدة صفات وهي :

الموضوعية والصدق والثبات ( ونعني الخصائص السيكمومترية للاختبار) والمعايير والقابلية

للاستعمال.

أما القابلية للاستعمال فتشمل مجموعة من العوامل هي:

- سهولة التطبيق.
- سهولة التصحيح.
- اقتصادية التكاليف.
- سهولة التفسير.

بينما يذكر أبو علام(١٩٨٧م٢٧٤)" بأن خاصيتي الثبات والصدق من أهم الخصائص التي يجب أن تتوفر في الاختبار كي يتم الوثوق والاعتماد عليه وعلى نتائجه التي نحصل عليها من تطبيقنا له".

## أولاً : الثبات

هناك أكثر من معنى للثبات فقد يعني الاستقرار أي تماثل نتائج أداة القياس عند تطبيقه أكثر من مرة تحت نفس الظروف ، وقد يعني الاتساق أي اتساق نتائج الاختبار مع بعضها البعض أو الانسجام بين أجزاء الاختبار (عودة ، ١٩٩٣م) (ثوراندايك ، هيجن ، ١٩٨٦م).

ويتفق الباحث مع التعريف الذي ذكره إسماعيل(٢٠٠٧م، ١٦) عن الثبات وهو "الحصول على نفس النتائج تقريباً عندما يتم إعادة نفس الاختبار على نفس المجموعة وفي نفس الظروف تقريباً".

ويعد الثبات خاصية من خصائص القياس الناتج وليس من خصائص أداة القياس وبالتالي فإن القياس الناتج عنها هو الذي يتميز بدرجة معينة من الثبات.(علام ، ٢٠٠٦م).

ونلاحظ أن درجات الاختبار التي تتمتع بقدر مرتفع من الثبات وكانت تعكس الجوانب الحقيقية للسمة أو القدرة المراد قياسها ولا تعكس عوامل الخطأ العشوائية يزداد ثبات هذه الدرجات إذا قلت هذه العوامل وتوجد مصادر متعددة تسهم في أخطاء القياس وقد تؤدي إلى انخفاض ثبات درجات الاختبارات بعضها يتعلق بأداة القياس ذاتها

وبعضها يتعلق بإجراءات وظروف تطبيقها وبعضها الآخر يتعلق بالأفراد المختبرين، وبالتالي فإن الدرجات الملاحظة التي يحصل عليها الأفراد المختبرين يشوبها قدر من الخطأ، ولو استطعنا تقديره نحصل على الدرجة الحقيقية للفرد. (علام ، ٢٠٠٦م).

أي أن الدرجة الملاحظة = الدرجة الحقيقية + درجة الخطأ .

وبالتالي فإن معامل الثبات يمثل النسبة بين تباين الدرجات الحقيقية وتباين الدرجات الملاحظة، وهذه النسبة تمثل من خلال المعادلة رقم (١٠):

$$\text{---} ( ) =$$

(إسماعيل، ٢٠٠٧م)

### طرق حساب معامل الثبات :

#### - طريقة إعادة الاختبار (Test-retest) :

يتم تطبيق الاختبار على مجموعة من الأفراد والحصول على النتائج ثم إعادة تطبيق نفس الاختبار على نفس المجموعة بعد فترة زمنية مناسبة تحت نفس الظروف تقريباً والحصول على النتائج ثم حساب معامل الارتباط بين أداء الأفراد في التطبيقين ويسمى معامل الارتباط في هذه الحالة بمعامل الاستقرار .

ومن عيوب معامل الاستقرار، أننا لا نستطيع توحيد ظروف التطبيقين. (أبو لبة ،

١٩٨٥م) و (إسماعيل، ٢٠٠٧م) .





افترض رولون أن التباين الكلي للاختبار ناتج من جزأين :

- أحدهما ناتج عن الخطأ التجريبي .

- والآخر ناتج عن الفروق الفردية .

وتأخذ معادلة رولون الصورة التالية وهي معادلة رقم (١٢) :

$$\sigma^2 = \sigma_e^2 + \sigma_b^2 \quad (12)$$

( ) .

حيث  $f$  = تباين الفرق بين درجات نصفي الاختبار

و  $k$  = تباين الاختبار ككل.

#### ٥- معادلة جتمان : Gutman Formula

يرى جتمان إمكانية إهمال شرط تساوي التباين بين نصفي الاختبار عند سبييرمان

براون وخرج بالمعادلة التالية رقم (١٣):

$$\sigma^2 = \left[ \frac{f}{k} + \frac{1}{n} \right] \sigma_e^2 \quad (13)$$

حيث : = تباين النصف الأول من الاختبار.

= تباين النصف الثاني من الاختبار.

ك = تباين الاختبار ككل. (إسماعيل، ٢٠٠٧م).

٦- معادلة كيودر- ريتشاردسون :

تستخدم هذه المعادلة لحساب التجانس الكلي للاختبار وتقوم على اعتبار كل

سؤال من الاختبار بنداً وجزءاً منه ونستخدمها لشمول الاتساق بين هذه الأجزاء (الأسئلة)

ونسقيه معامل الاتساق الداخلي (معامل التجانس)، ولها صيغتان:

الصيغة الأولى ( صيغة ٢٠):

وهي الأكثر استخداماً ويشترط لاستخدامها أن يقيس الاختبار بعد واحد ولا يكون

الاختبار اختبار سرعة ودرجة كل سؤال إما صفر أو واحد وهي المعادلة رقم (١٤):

$$k = \frac{(\sum_{i=1}^n \sum_{j=1}^n a_{ij}^2) - \frac{(\sum_{i=1}^n a_i)^2}{n}}{n(n-1)}$$

حيث ن : عدد مفردات الاختبار

ك = التباين الكلي لدرجات الاختبار.

أ بز : معامل السهولة.

أ ثث : معامل الصعوبة. (إسماعيل، ٢٠٠٧م).

الصيغة الثانية ( صيغة ٢١):

ويشترط لاستخدامها بالإضافة إلى الشروط في صيغة ٢٠ تساوي جميع مفردات  
 المقياس في درجة الصعوبة ويكون متوسط درجة الصعوبة لجميع المفردات ٠,٥، وهذا  
 الشرط أدى إلى شيوع استخدام صيغة ٢٠ أكثر من صيغة ٢١، وصيغة ٢١ هي المعادلة  
 رقم(١٥):

$$\left( \frac{\sum_{i=1}^n x_i}{n} - \frac{\sum_{i=1}^n x_i^2}{n} \right) = \left( \frac{\sum_{i=1}^n x_i}{n} \right)^2 - \frac{\sum_{i=1}^n x_i^2}{n}$$

حيث م : متوسط الدرجات الكلية، و ن : عدد مفردات الاختبار.(إسماعيل، ٢٠٠٧م)

### ٧- معادلة كرونباخ (α) Cronbach

وتستخدم لحساب ثبات الاختبار من جزأين أو أجزاء متعددة وهي المعادلة  
 رقم(١٦):

$$\left( \frac{\sum_{i=1}^n x_i}{n} - \frac{\sum_{i=1}^n x_i^2}{n} \right) = \left( \frac{\sum_{i=1}^n x_i}{n} \right)^2 - \frac{\sum_{i=1}^n x_i^2}{n}$$

ن : عدد الأجزاء التي ينقسم إليها الاختبار .

: تباين درجات الاختبار ككل .

: تباين المفردات (أي مفردة ن ) من الاختبار.

مج : مجموع تباينات الأجزاء.(إسماعيل، ٢٠٠٧م).

ملاحظة عندما ن = ٢ نحصل على معادلة جتمان .

## العوامل المؤثرة في ثبات درجات الاختبار :

١- عدد الأسئلة ( طول الاختبار ) :

معامل الثبات يزداد مع زيادة عدد الأسئلة وبالتالي فإن العلاقة بينهما طردية.

٢- زمن الاختبار :

يزداد معامل الثبات إذا زاد الزمن إلى الحد المناسب ثم يقل بعد ذلك وبالتالي فإن

العلاقة بينهما علاقة منحنية.

٣- ظروف إجراء الاختبار:

توحيد الظروف من تطبيق لآخر لأن اختلافها يؤثر في معامل الثبات وتزيد عوامل

الخطأ ومن الصعب الحصول على نفس الظروف ولكن بقدر الإمكان نصل إلى نفس

الظروف تقريباً.

٤- التخمين:

يؤدي التخمين إلى انخفاض معدل الثبات.

٥- تباين قدرات الطلاب :

كلما كانت المجموعة أقل تجانساً ومن مستويات مختلفة فإن ذلك يزيد من

معامل الثبات، والعكس لأن معامل الثبات يعتمد في حسابه على تشتت الدرجات

وتباينها، وبالتالي كلما زاد تباين الدرجات زاد ثبات الاختبار والعكس صحيح.

٦- التصحيح :

إن موضوعية التصحيح ترفع من ثبات الاختبار. (إسماعيل ، ٢٠٠٧م).

## ثانياً: الصدق: Validity

يعتبر أبو لبدة (١٩٨٥م: ٢٣٣) "الصدق أهم خاصية من خواص الاختبار ويعني ذلك

أن يقيس الاختبار بالفعل ما وضع لقياسه".

والصدق يكون نسبياً أي أنه لا يوجد اختبار صادق بشكل كامل ولا يوجد

اختبار غير صادق، أيضاً الصدق يكون نوعياً أي صادقاً للغرض الذي أعد من أجله

ضمن العينة والظروف المحيطة ( عودة ، ١٩٩٣م ).

الطرق المختلفة للتأكد من الصدق :

١- الصدق الظاهري ( صدق المحكمين ).

٢- صدق المحتوى.

٣- صدق المحك.

٤- صدق التكوين الفرضي .

أولاً : الصدق الظاهري ( صدق المحكمين ) Face Validity

ويقصد به إمكانية التعرف على ما يقيسه الاختبار من خلال العنوان وبالتالي فهو فحص سطحي ولا يعتد به في الأبحاث ويتم تحديد معامل الصدق الظاهري عن طريق صدق المحكمين. (إسماعيل ، ٢٠٠٧م).

### ثانياً : صدق المحتوى Content Validity

يدل صدق المحتوى على مدى تمثيل محتوى الاختبار للنطاق السلوكي الشامل للسمة المراد قياسها ويكون واضحاً في الاختبارات التحصيلية بأنواعها (إسماعيل، ٢٠٠٧م)، ونستطيع التأكيد من صدق المحتوى في الاختبارات النفسية عن طريق صدق المحكمين بينما في الاختبارات التحصيلية فإننا نستخدم جدول المواصفات بعد تحليل المقرر الدراسي والتعرف على أهدافه.

### ثالثاً : صدق المحك Criterion-related validity

يرتبط هذا الصدق بمدى الارتباط بين درجات الاختبار المراد معرفة صدقه ومحك له صدق وثبات موثوق بهما ومطبق على نفس العينة، وهو ينقسم إلى صدق تلازمي وصدق تنبؤي.

### أ- الصدق التلازمي Congruent Validity

ونقصد به مدى تطابق وارتباط نتائج تطبيق الاختبار ونتائج محك في نفس الوقت أي في الوقت الحاضر وذلك بإيجاد معامل الارتباط بين الاختبار وبين المحك.

وكلما كانت قيمة معامل الارتباط ( $+0,5$ ) أو أكثر دل على الصدق والأفضل أن

يكون ( $+0,7$ ) وكلما زاد عدد العينة كان أفضل (إسماعيل ، ٢٠٠٧م) .

## ب- الصدق التنبؤي (المستقبلي) Predictive Validity

يشير الصدق التنبؤي إلى مدى نجاح الاختبار وقدرته على التنبؤ بأداء الفرد في

المستقبل ويتم حسابه بإيجاد معامل الارتباط بين درجات الاختبار الحالي ودرجات محك

آخر بعد فترة زمنية مناسبة مثلاً كاختبار القدرات أو اختبارات القبول في الجامعات

والمعدل التراكمي للطالب .

ويعتبر معامل الصدق التلازمي قيمة قصوى لمعامل الصدق التنبؤي. (إسماعيل،

٢٠٠٧م).

رابعاً : صدق التكوين الافتراضي ( صدق البناء ) وينقسم إلى ثلاثة أقسام :

أ- التحليل المنطقي وينقسم إلى ثلاثة أقسام:

١- النظرية والبحوث التجريبية:

يتم تصميم الاختبار في ضوء نظرية معينة يتم التأكد من صحة فروض هذه

النظرية بالبحوث التجريبية.

٢- التحليل المتعدد:

مقارنة نتائج الاختبار الحالي بنتائج البحوث والدراسات السابقة التي تناولت نفس

الموضوع وإذا وجد اتساق فيما بينها كان ذلك دليلاً على الصدق.

### ٣- طريقة تطور الأعمار:

تناسب هذه الطريقة بعض السمات والقدرات المرتبطة في نموها مع العمر وبخاصة الذكاء ، وقد لا تتلائم مع باقي الصفات الإنسانية .

ب- طريقة الارتباطات وتنقسم إلى ثلاثة أقسام :

#### ١- الصدق العاملي : Factorial Validity

ذكر أبو لبدة (١٩٨٥م: ٢٥٢) أنه عبارة عن البحث عن عوامل مشتركة تقيسها عدة اختبارات صادقة لكي يتم تحديد اشتراك هذه الاختبارات في قياس تلك العوامل أي أن معامل الصدق العاملي هو عامل مشترك بين الاختبار وأي اختبار آخر أو مجموعة اختبارات أخرى تقيس الشيء نفسه".

#### ٢- المقارنة الطرفية ( الاتساق الداخلي ) :

يتم تطبيق الاختبار المراد حساب صدقه ثم نحسب الدرجة الكلية للاختبار، ثم نحصل على مجموعتين ٢٧٪ أعلى و ٢٧٪ أدنى، ثم نحسب الفروق بين المجموعتين المرتفعة والمنخفضة في درجات مفردات الاختبار عن طريق اختبار(ت) ويوجد احتمالين الأول وجود فروق ذات دلالة إحصائية بين المجموعتين ويكون الاختبار بذلك صادقاً لأنه استطاع



التمييز بين المجموعتين، والثاني عدم وجود فروق ذات دلالة إحصائية بين المجموعتين يكون الاختبار غير صادق لأنه لم يستطع التمييز بين المجموعتين. ( إسماعيل ، ٢٠٠٧م ).

### ٣- الصدق التقاربي والتمييزي ( التمايزي )

إذا ارتبط الاختبار ارتباطاً عالياً أو مرتفعاً بالمتغيرات التي يجب أن يرتبط بها نظرياً أو تتسق معه يسمى هذا صدق تقاربي .

أما إذا ارتبط الاختبار ارتباطاً منخفضاً بالمتغيرات التي لا يجب أن يرتبط بها نظرياً أو لا تتسق معه فهذا صدق تمييزي ( تمايزي ) ( إسماعيل ، ٢٠٠٧م ).

### ج- الصدق التجريبي :

عندما يتم اتساع نظام التطبيق القبلي للاختبار على الطلاب قبل تدريس المقرر ثم التطبيق البعدي بعد تدريس المقرر ضمن المنطقي فشل الطلاب في الاختبار القبلي ونجاحهم في الاختبار البعدي وبفحص المفردات يتم الحكم بصعوبة المفردة عند فشل معظم الطلاب في التطبيق على حلها وبسهولة المفردة عند نجاح معظم الطلاب في التطبيقين على حلها ، ويكون نجاح معظم الطلاب في الاختبار القبلي وفشلهم في الاختبار البعدي حالة شاذة وتشير إلى وجود خطأ ما في الصياغة أو خطأ التعليم أو كليهما معاً . ( إسماعيل ، ٢٠٠٧م ).

العوامل المؤثرة على الصدق:

ذكر أبو لبدة (١٩٨٥م: ٢٥٤) أن هناك عدداً من العوامل التي تؤثر على الصدق،

ومنها:

#### أ- عوامل تتعلق بالاختبار:

كأن تكون لغة الاختبار فوق مستوى الطلاب أو غموض الأسئلة مما يجعل

الطلاب يفسرونها بتفسيرات مختلفة، أو تكون الأسئلة دون مستوى الطلاب أو فوق

مستواهم فدرجاتهم لا تمثل مقدرتهم الحقيقية وبالتالي يكون الاختبار أقل صدقاً.

#### ب- عوامل تتعلق بالطالب:

اضطراب الطالب وقلقه الزائد يقللان من قدرته الحقيقية على الإجابة أيضاً الغش

والتخمين يقللان من صدق الاختبار.

#### ج- عوامل تتعلق بإدارة الاختبار:

ومن هذه العوامل البيئية التي تؤثر على أداء الطلبة كالحرارة الزائدة والضوضاء،

كما أن عدم وضوح طباعة وإخراج الاختبار وعدم وضوح تعليمات الاختبار تؤثر وتقلل من

صدق الاختبار .

أيضاً استخدام الاختبار في غير ما وضع من أجله أو لفئة غير الفئة التي وضع من

أجلها فهذه العوامل تقلل من صدق الاختبار.

## ثانياً: دراسات وبحوث سابقة

هناك العديد من الدراسات والبحوث التي تناولت أثر عدد البدائل في اختبارات الاختيار من متعدد على الخصائص السيكومترية، منها ما يبحث هذا الأثر على المفردة، ومنها ما يبحث هذا الأثر على الاختبار نفسه، ومنها ما يبحث هذا الأثر على المفردة والاختبار معاً.

وفي هذا البحث تم التركيز على الدراسات التي تناولت أثر عدد البدائل في اختبارات الاختيار من متعدد، وتم تقسيمها إلى ثلاثة أقسام، في القسم الأول الدراسات التي بحثت أثر عدد البدائل على الخصائص السيكومترية للفقرة، والقسم الثاني الدراسات التي بحثت أثر عدد البدائل على الخصائص السيكومترية للاختبار، والثالث الدراسات التي بحثت في أثر عدد البدائل على الخصائص السيكومترية للاختبار ومفرداته، وفيما يلي بيان لتلك الدراسات:

أولاً: الدراسات التي بحثت في أثر عدد البدائل في اختبارات الاختيار من متعدد على الخصائص السيكومترية للمفردة (معاملي الصعوبة والتمييز):

قام كوستين Costin (١٩٧٠م) بدراسة تناولت أثر اختلاف عدد البدائل في اختبار الاختيار من متعدد على صعوبة الفقرات وتمييزها، و استخدم في هذه الدراسة صورتين للاختبار يتكون كل منهما من ١١٠ فقرات تقيس محتوى أربعة موضوعات رئيسية من

مساق مقدمة في علم النفس، منها ٢٥ فقرة على الإدراك و ٣٠ فقرة على التعلم و ٣٠ فقرة على الدافعية و ٢٥ فقرة على الذكاء، وكان عدد بدائل فقرات الصورة الأولى أربعة بدائل، وعدد بدائل فقرات الصورة الثانية ثلاثة بدائل، تم الحصول عليها بحذف أحد بدائل فقرات الصورة الأولى عشوائياً، وبعد تطبيق هذه الاختبارات على ٢٠٧ طلاب من الطلبة المسجلين في جامعة إلينوي وينتمون إلى أربعة شعب صافية منها شعبتان من الطلبة المسجلين في الفصل الأول، وشعبتان من الطلبة المسجلين في الفصل الثاني، ولأغراض التحليل الإحصائي استبعدت منها إجابات سبعة طلاب عشوائياً، ثم حسبت معاملات صعوبة هذه الفقرات وتمييزها، وبينت نتائج هذه الدراسة أن معاملات صعوبة وتمييز الفقرات ذات الثلاثة بدائل أعلى من معاملات صعوبة وتمييز الفقرات ذات الأربعة بدائل.

وفي دراسة قام بها لورد Lord (١٩٧٧م) هدفت لمقارنة صعوبة وتمييز فقرات الاختبارات ذات البديلين، والثلاثة بدائل، والأربعة بدائل، والخمسة بدائل، ولتحقيق أهداف هذه الدراسة تم بناء أربعة اختبارات من القسم اللفظي من اختبارات القابلية المدرسية (Scholastic Aptitude Test) شكل (SAT١٣)، حيث تم بناء اختبار من (١٢٠) فقرة لكل فقرة خمسة بدائل، واختبار مكون من (١٥٠) فقرة لكل فقرة أربعة بدائل، واختبار مكون من (٢٠٠) فقرة لكل فقرة ثلاثة بدائل وآخر اختبار مكون من (٣٠٠) فقرة لكل فقرة بديلين، وكان الاختلاف في عدد الفقرات للحصول على عدد متساوٍ من البدائل وهو (٦٠٠) بديل لكل نموذج، وبعد تطبيق الاختبارات بينت النتائج أن الفقرات ذات البديلين هي الأسهل في حين كانت الفقرات ذات الخمسة بدائل أصعبها،

ولم تبين النتائج فروقاً ذات دلالة إحصائية بين معاملات التمييز للنماذج الأربعة، وقد تم استخدام منحني خاصة الفقرة في تحليل النتائج.

أما الدراسة التي قام بها روينولدز (Roynolds) (١٩٧٩م) والتي هدفت إلى تحديد شكل اختبار الاختيار من متعدد المناسب الذي يمكن استخدامه للطلاب ذوي الإعاقات العقلية البسيطة في المدارس الثانوية وهدفت أيضاً للمقارنة بين معاملات الصعوبة ومعاملات التمييز، حيث تم إعداد اختبارين في مجال الصحة والنظافة، الاختبار الأول مكون من (٣٦) فقرة من نوع الاختيار من متعدد لكل فقرة من فقراته ثلاثة بدائل ويأخذ الصورة (أ) حيث يقدم شفويًا للطلاب، والاختبار الثاني مكون من (٣٦) فقرة لكل فقرة بديلان ويأخذ الصورة (ب) ويقدم شفويًا للطلاب، وهو نفس الاختبار (أ)، إلا أن لكل فقرة بديلين، حيث تم حذف الموه الذي لا يتمتع بصدق ظاهري من كل فقرة من فقرات الشكل (أ)، وتم تطبيق الاختبارات على عينة مكونة من (١٠٠) طالب لديهم إعاقات عقلية بسيطة في مدارس ثانوية ذات تعليم خاص بالمعاقين عقلياً من كل من ولايتي أوريجن (Orgon) وواشنطن (Washington)، وتم حساب معامل ارتباط بويت بايسريال (Point-Biserial) لحساب معاملات التمييز لكل فقرة من فقرات الاختبارين، وكانت معاملات التمييز للنموذج ذي الثلاثة بدائل تتراوح بين (٠,١١ - ٠,٥٧) وللنموذج ذي البديلين بين (٠,٠٦ - ٠,٤٨)، وتم حساب معامل الصعوبة للفقرة عن طريق نسبة الذين أجابوا إجابة صحيحة عن تلك الفقرة، وكانت متوسطات معاملات الصعوبة (٠,٨١) للنموذج ذي الثلاثة بدائل و (٠,٨٧) للنموذج ذي البديلين، وتم استخدام تحليل

التباين الأحادي لمعرفة الفروق بين المتوسطات وأشارت النتائج إلى عدم وجود فروق ذات دلالة بين معاملات التمييز وكذلك بين متوسطات معاملات الصعوبة عند مستوى  $(\alpha = 0,05)$ .

وفي دراسة قام بها كرهان وآخرون (Crehan et al., 1993) هدفت لمعرفة العدد الأمثل من البدائل في اختبارات الاختيار من متعدد من حيث الصعوبة والتمييز، ولتحقيق أهداف هذه الدراسة تم إعداد اختبار في مساق مقدمة في علم النفس مكون من (40) فقرة وتم إعداد نموذجين منه الأول له ثلاثة بدائل، و الثاني له أربعة بدائل، وكانت الفقرات نفسها في النموذجين ولكن الاختلاف في عدد البدائل فقط، وتم تطبيق النموذجين على عينة مكونة من (220) طالباً تم اختيارها عشوائياً من طلاب إحدى الكليات الجامعية، وتم حساب معاملات الصعوبة عن طريق إيجاد نسبة الطلبة الذين أجابوا إجابة صحيحة عن كل فقرة، وكانت متوسطات معاملات الصعوبة للنموذج ذي الثلاثة بدائل (0,80) ومتوسط معاملات الصعوبة للنموذج ذو الأربعة بدائل (0,77)، وتم حساب معاملات التمييز عن طريق حساب معامل ارتباط (بوينت بايسيريال) وكانت متوسطات معاملات التمييز (0,35) للنموذج ذي الثلاثة بدائل، و (0,36) للنموذج ذي الأربعة بدائل.

ولمعرفة الفروق بين المتوسطات تم استخدام اختبار (t) وقد تبين وجود فروق ذات دلالة إحصائية عند مستوى  $(\alpha = 0,05)$  بين متوسطات معاملات الصعوبة لصالح النموذج ذو الثلاثة بدائل أي أن فقراته أسهل من فقرات النموذج ذو الأربعة بدائل، ولم تظهر

النتائج وجود فروق ذات دلالة بين متوسطات معاملات التمييز، وفي ضوء النتائج السابقة أوصت الدراسة باستخدام النموذج ذو الثلاثة بدائل لأنه يوفر الوقت والجهد، ومتوسط معامل الصعوبة له أكبر من متوسط معامل الصعوبة للنموذج ذي الأربعة بدائل .

وفي دراسة اليعقوب (١٩٩٦م) والتي هدفت إلى معرفة أثر اختلاف عدد البدائل والبديل لاشي مما ذكر على الخصائص السيكومترية لفقرات اختبار الاختيار من متعدد، وكانت أداة الدراسة عبارة عن اختبار تحصيلي في مادة القياس والتقويم لطلاب الدبلوم التربوي لكلية التربية بجامعة اليرموك للعام الدراسي ١٩٩٤ - ١٩٩٥م، وبلغ عدد فقراته ٧٥ فقره من نوع الاختيار المتعدد، وأعد هذا الاختبار من قبل عدد من مدرسي المادة وطبق في فصول سابقة، ولأغراض الدراسة أعدت أربع نماذج من هذا الاختبار، دون إجراء أي تغيير في متون الفقرات، وأبقيت الفقرات على حالها في النموذج الأول بأربعة بدائل وكان النموذج الثاني بثلاث بدائل وذلك بحذف احد بدائل النموذج الأول عشوائيا، أما النموذج الثالث والرابع فيتكون كل منهما من أربعة بدائل أحدهما البديل لاشيء مما ذكر، واختيرت ٣٢فقرة أي ما يشكل ٤٢,٦٪ من فقرات الاختبار ووزعت هذه الفقرات عشوائيا في النماذج الأربعة بحيث يكون النموذج الأول مكون من أربعة بدائل والثاني من ثلاثة بدائل والثالث من أربعة بدائل أحدها البديل لاشيء مما ذكر كواحد من البدائل الأربعة والرابع أربعة بدائل أحدهم البديل لاشيء مما ذكر بحيث يكون هذا البديل هو الجواب الصحيح في جميع هذه الفقرات المختارة، وتم تطبيق هذه النماذج على ١٦٥ طالبا وطالبة بواقع ٤٢ طالبا وطالبة على النموذج الأول و٤١ طالبا وطالبة لكل نموذج من النماذج

الباقية، وأشارت نتائج هذه الدراسة إلى عدم وجود فروق دالة إحصائية بين أوساط قيم معاملات الصعوبة والتمييز لفقرات اختبار الاختيار من متعدد ذو الأربعة بدائل والثلاث بدائل، كما أشارت النتائج إلى وجود فرق دال إحصائياً بين قيم معاملات الصعوبة للنموذج الرابع الذي يحتوي على البديل لاشيء مما ذكر كجواب صحيح بأنه أصعب من النموذج الثالث الذي يحتوي على البديل لاشيء مما ذكر كأحد بدائلها، بينما لم تشر النتائج إلى فرق دال إحصائياً بين وسطي قيم معاملات التمييز للنموذجين الثالث والرابع.

ثانياً : الدراسات التي بحثت في أثر عدد البدائل في اختبارات الاختيار من متعدد في

خصائص الاختبار ( الصدق - الثبات ) :

قام جرير Grier (١٩٧٥) بدراسة هدفت لمقارنة ثبات اختبارات الاختيار من متعدد ذات البديلين والثلاثة بدائل والأربعة بدائل، وتم حساب معاملات الثبات (التجانس الداخلي) باستعمال معادلة كيودر ريتشاردسون (٢١) (Kudder-Richardson) وأشارت النتائج أن النموذج الذي يكون لفقراته ثلاثة بدائل يتمتع بثبات أعلى من بقية الاختبارات، ويتبعه النموذج الذي لكل فقره من فقراته بديلين، ثم النموذج ذو الأربعة بدائل.

وفي دراسة قامت بها ويبر Weber (١٩٧٧م) هدفت لمعرفة أثر نماذج الاختيار من متعدد على تجانس عدداً من الاختبارات التي قام بإعدادها مجموعة من المعلمين وطبقت على مجموعتين أحدهما عالية التحصيل والأخرى منخفضة التحصيل، وقسمت الدراسة إلى قسمين الأول يهدف لمقارنة معامل الثبات (التجانس) بين نموذج له خمسة بدائل،



ونموذج له ثلاثة بدائل ، والقسم الثاني لمقارنة معامل الثبات لنموذج ذو ثلاثة بدائل ولنموذج له أربعة بدائل ، وأشارت النتائج بأنه لا تحسن على معاملات الثبات بإنقاص عدد البدائل بالنسبة لمجموعة التحصيل العالي ، ولكن هناك تحسن على معاملات الثبات بالنسبة لمجموعة التحصيل المنخفض.

أما الدراسة التي قام بها كل من بيدسكو ونيفو Budescu & Nevo (1985م) والتي هدفت إلى معرفة العدد الأمثل من البدائل في فقرات اختبارات الاختيار من متعدد ، ولتحقيق أهداف هذه الدراسة تم استخدام اختبار للرياضيات ، يتكون من أربعة نماذج تختلف فقط في عدد بدائلها ، حيث كان النموذج الأول يتكون من بديلين ، و الثاني يتكون من ثلاثة بدائل ، والنموذج الثالث يتكون من أربعة بدائل والنموذج الرابع يتكون من خمسة بدائل ، وكانت العينة مؤلفة من (1018) طالباً جامعياً ، وتم حساب معاملات الثبات باستخدام معادلة كيودر ريتشاردسون (20) (Kudder-Richardson) (20) وكانت (0,797) للنموذج ذي البديلين و (0,843) للنموذج ذي الثلاث بدائل ، (0,840) للنموذج ذي الأربع بدائل (0,850) للنموذج ذي الخمس بدائل وكانت المتوسطات الحسابية للنموذج ذي البديلين (16,79) ، و (15,52) للنموذج ذي الثلاث بدائل ، و (14,27) للنموذج ذي الأربعة بدائل ، (14,27) للنموذج ذي الخمسة بدائل ويتضح من ذلك أن معاملات الثبات تزداد بزيادة عدد البدائل وأن المتوسط الحسابي يتناقص بزيادة عدد البدائل ، وأوصت الدراسة باستخدام النموذج ذي الثلاثة بدائل لأن معامل ثباته قريب من معاملات ثبات النموذج ذي الأربعة بدائل ، و النموذج ذي الخمسة بدائل.

وفي دراسة تريفيزان وساكس وميشيل Trevisan, Sax & Michael (1991م) والتي هدفت لمعرفة أثر عدد البدائل في اختبارات الاختيار من متعدد على صدق وثبات الاختبار، ولتحقيق أهداف الدراسة تم استخدام اختبار القبول التابع لجامعة واشنطن وهو اختبار من نوع الاختيار من متعدد لكل فقرة من فقراته خمسة بدائل، ومن أجل بناء اختبار لفقراته أربعة بدائل، واختبار لفقراته ثلاثة بدائل تم حذف الموهبات الأقل تمييزاً من بطارية الاختبار الأصلية ذات الخمسة بدائل، بعد ذلك تم تطبيق نماذج الاختبار على عينة حجمها (435) طالباً تابعين للمدارس الثانوية التي تقع في منطقة بورتلاند؛ وطلب من المفحوصين وضع معدلاتهم على ورقة الإجابة، حيث تم تصنيفهم حسب هذه المعدلات إلى ثلاث مجموعات ( ذات تحصيل عال ومتوسط، ومنخفض)، وأشارت نتائج الدراسة إلى وجود فروق ذات دلالة إحصائية عند مستوى ( $\alpha = 0,05$ ) بين معاملات الثبات لمجموعة الطلبة منخفضي التحصيل ولصالح النموذج الذي لفقراته ثلاثة بدائل، ولا توجد فروق ذات دلالة إحصائية بين معاملات الثبات لمجموعة الطلبة متوسطي ومرتفعي التحصيل، وأن ثلاثة خيارات هو العدد الأمثل في اختبار الاختيار من متعدد للطلبة على اختلاف مستوياتهم التحصيلية، كما أشارت النتائج إلى وجود فروق ذات دلالة إحصائية بين معاملات الصدق لمجموعة الطلبة منخفضي التحصيل ولصالح النموذج الذي لفقراته أربعة بدائل، ولا توجد فروق ذات دلالة إحصائية بين معاملات الصدق لمجموعتي الطلبة مرتفعي ومتوسطي التحصيل .

وفي دراسة لاحقة قام بها تريفيزان وساكس وميشيل & Trevisan, Sax Michael (١٩٩٤م) هدفت إلى مقارنة معاملات ثبات اختبارات الاختيار من متعدد ذات البديلين، والثلاثة بدائل والأربعة بدائل والخمسة بدائل، ولتحقيق هذا الهدف تم استخدام اختبار مكون من (٤٥) فقرة من نوع الاختيار من متعدد لكل فقرة بديلين تم تطويره باستخدام قواعد كتابة الفقرات التي اقترحها رويد و هالدينا (١٩٨٢) ويقيس موضوعات في الموسيقى و الآداب و التاريخ و الجغرافيا، بعد ذلك تم تكوين نموذج آخر له نفس العدد من الفقرات ولكل فقرة ثلاثة بدائل وذلك بإضافة بديل لكل فقرة من فقرات النموذج الذي لفقراته بديلين، وتم تكوين نموذج لكل فقرة من فقراته أربعة بدائل وذلك بإضافة بديل لكل فقرة من فقرات النموذج الذي لفقراته ثلاثة بدائل، وتم تكوين نموذج لكل فقرة من فقراته خمسة بدائل وذلك بإضافة بديل لكل فقرة من فقرات النموذج الذي لفقراته أربعة بدائل، وطبقت نماذج الاختبار على عينة من طلبة الصف الثاني عشر في مدارس بورتلاند حجمها (١٥٤) طالباً، بواقع (٤٠) طالباً للنموذج ذي البديلين، و(٣٨) طالباً لبقية النماذج الثلاثة، وأشارت نتائج الدراسة إلى وجود فروق ذات دلالة إحصائية بين معاملات الثبات لنماذج الاختبار الأربعة قبل تنفيذ شرط تفرسكي (Tversky) الذي يفترض أن زمن الاختبار يتناسب مع طول الاختبار ولصالح النموذج الذي لفقراته أربعة بدائل، ولا توجد فروق ذات دلالة إحصائية بين معاملات الثبات بعد تنفيذ شرط تفرسكي (Tversky) حيث كانت معاملات الثبات للنماذج الأربعة متقاربة، مع أن أعلى معامل ثبات كان للنموذج الذي لفقراته أربعة بدائل، و أوصت الدراسة باستخدام النموذج الذي لفقراته ثلاثة بدائل لأنه يوفر الجهد ويعمل على زيادة صدق المحتوى أكثر من الاختبار الذي لفقراته خمسة بدائل أو أربعة بدائل.

ثالثاً : الدراسات التي بحثت في أثر عدد البدائل في اختبارات الاختيار من متعدد في

خصائص الفقرة و الاختبار :

قام كوستين Costin (١٩٧٢م) بدراسة هدفت لمقارنة خصائص اختبار لفقراته ثلاثة بدائل باختبار لفقراته أربعة بدائل، واستخدم في هذه الدراسة ١٠٠ فقرة اختيرت عشوائياً من مجموعة من الفقرات تقيس موضوعات في علم النفس وهي الذكاء، والشخصية، والتفاعل الاجتماعي، والاضطراب السلوكي، واختيرت هذه الفقرات بما يتناسب مع حجم محتوى هذه الموضوعات، ثم اختير من هذه الفقرات ٥٠ فقرة حذف أحد بدائلها عشوائياً لتشكّل الفقرات ذات الثلاثة بدائل، وأما الفقرات الخمسون الباقية فأبقي على بدائلها كما هي، ثم طبق الاختبار بكامله على ١٢٠٠ طالب كاختبار نهائي للمساق، ولأغراض التحليل اعتبر كل نوع من هذه الفقرات اختباراً منفصلاً، وحسب معامل ثبات الاتساق الداخلي لكل منهما بمعادلة (كبودر - ريشاردسون ٢٠) فكان ٠,٧٥ للفقرات ذات الثلاثة بدائل و ٠,٧٨ للفقرات ذات الأربعة بدائل وتبين أنه لا يوجد فروق دالة إحصائية بين قيم معاملات الثبات هذه، كما حسبت معاملات تمييز هذه الفقرات وهي عبارة عن معاملات ارتباط الفقرات مع الاختبار ككل، ثم حسب الوسط لمعاملات تمييز الفقرات في كلا النموذجين بعد تحويلها إلى درجات فشر الزائفة فكانت ٠,٢٩ للفقرات ذات الثلاثة بدائل و ٠,٣٠ للفقرات ذات الأربعة بدائل، كما حسب متوسط عدد الفقرات التي أجيب من كل نوع فكان للفقرات ذات الثلاثة بدائل ٣٦,٨ ومن الفقرات ذات الأربعة بدائل ٣٦,٤، وتبين أنه لا يوجد فرق دال إحصائياً بين هذه المتوسطات.

كما قام ستراتون وكاتس Straton & Catts (١٩٨٠م) بدراسة هدفت إلى المقارنة بين خصائص اختبارات الاختيار من متعدد حيث قام الباحثان ببناء ثلاثة نماذج لاختبار في مادة الاقتصاد للصف الثاني عشر، وتضم فقرات كل نموذج منها ١٢٠ بديلاً، ويتكون النموذج الأول من ٦٠ فقرة لكل منها بديلان، والنموذج الثاني من ٤٠ فقرة لكل منها ثلاثة بدائل، أما النموذج الثالث فيتكون من ٣٠ فقرة لكل منها أربعة بدائل، وطبقت الاختبارات في ظروف صافية عادية حيث وزع كل نموذج منها عشوائياً على ٦٥ تلميذاً من تلاميذ الصف الثاني عشر، وبينت نتائج هذه الدراسة أن أسهل الفقرات كانت ذات البديلين وأصعبها ذات الأربعة بدائل، حيث كانت الأوساط لمعاملات الصعوبة هي: (٠,٧٥) و (٠,٧٦) و (٠,٥٩) للفقرات ذات البديلين والثلاثة والأربعة بدائل على التوالي، أما فيما يتعلق بمعاملات تمييز الفقرات فكان أقلها تمييزاً الفقرات ذات البديلين وأعلىها تمييزاً الفقرات ذات الأربعة بدائل، حيث كانت أوساط معاملات التمييز لهذه الفقرات هي: (٠,٢٤) و (٠,٣٥) و (٠,٤١) على التوالي، وكان أدنى معامل ثبات لاختبار الفقرات ذات البديلين وأعلىها لاختبار الفقرات ذات الأربعة بدائل وكانت قيم معاملات الثبات هي (٠,٤٧) و (٠,٦٦) و (٠,٦٨) على التوالي، كما أظهرت نتائج هذه الدراسة أن الوقت الذي يحتاجه المفحوص في قراءة البديل يساوي تقريباً نصف الوقت الذي يحتاجه المفحوص في قراءة متن السؤال، وتوصلت الدراسة إلى أن تقليل عدد البدائل يؤدي إلى تقليل الوقت الذي يحتاجه المفحوص للإجابة على مثل هذه الفقرات.

أما دراسة فقوسة (١٩٩٤م) والتي هدفت لمعرفة أثر اختلاف عدد البدائل في فقرات اختبار الاختيار من متعدد على الخصائص السيكومترية للاختبار وفقراته ، ولتحقيق أهداف هذه الدراسة قام الباحث ببناء اختبارين في مادتي الرياضيات والتاريخ للصف التاسع وكان كل اختبار يتألف من أربعة نماذج ، وكانت فقرات النماذج الأربعة متشابهة في المتن ومتساوية في العدد (٤٠ فقرة لكل منها) ، ولكنها مختلفة في عدد البدائل فكان النموذج الأول يتكون من خمسة بدائل ، والنموذج الثاني يتكون من أربعة بدائل ، وثلاثة بدائل لفقرات النموذج الثالث ، وبديلان لفقرات النموذج الرابع .

تم تطبيق هذه النماذج على عينة مكونة من (٢٣٦) طالباً تم اختيارها بطريقة عشوائية من طلبة الصف التاسع بواقع (٥٩) طالباً لكل نموذج. حسب معامل ثبات الاتساق الداخلي لنماذج الاختبارين باستخدام معادلة كرونباخ الفا ، كما حسبت معاملات الصدق لنماذج الاختبارين باستخدام معامل ارتباط بيرسون بين علامات المفحوصين على الاختبار الذي طبقه الباحث في مادة الرياضيات ومادة التاريخ .

وأشارت نتائج تحليل فقرات اختبار الرياضيات إلى وجود فروق ذات دلالة إحصائية بين متوسط معاملات صعوبة الفقرات ذات البديلين والاختبار الذي تحتوي كل فقرة من فقراته على ثلاثة أو أربعة أو خمسة بدائل ولصالح الاختبار ذي البديلين ، وبين الاختبار ذي الثلاثة بدائل والاختبار الذي تحتوي كل فقرة من فقراته على أربعة أو خمسة بدائل ولصالح الاختبار ذي الثلاثة بدائل ، أما بالنسبة لمتوسط معاملات التمييز فكان هناك فرق دال إحصائياً بين متوسط معاملات تمييز الفقرات ذات الخمسة بدائل و الأربعة بدائل

لصالح الفقرات ذات الخمسة بدائل، وكان الفرق دال إحصائياً أيضاً بين متوسط معاملات التمييز للفقرات ذات الثلاثة بدائل و ذات البديلين ولصالح الفقرات ذات الثلاثة بدائل، أما بالنسبة لفقرات اختبار التاريخ فلم تكن هناك فروق ذات دلالة إحصائية بين متوسطات معاملات صعوبة وتمييز الفقرات في النماذج الأربعة، وأشارت الدراسة إلى وجود فروق ذات دلالة إحصائية بين معامل الثبات لنموذج اختبار الرياضيات الذي لفقراته خمسة بدائل والاختبار الذي لفقراته بديلين لصالح الاختبار الذي لفقراته خمسة بدائل، ولم يكن هناك فروق ذات دلالة إحصائية بين معاملات الثبات لبقية النماذج الأخرى من اختبار الرياضيات، أما اختبار التاريخ بأشكاله الأربعة فلم يكن هناك فروق ذات دلالة إحصائية بين معاملات الثبات لهذه الأشكال، وبينت النتائج وجود فروق ذات دلالة إحصائية بين معاملات الصدق لشكل اختبار الرياضيات الذي لفقراته ثلاثة بدائل والاختبار الذي تحتوي فقراته على بديلين ولصالح النموذج الذي لفقراته ثلاثة بدائل، ولم يكن هناك فروق ذات دلالة إحصائية بين معاملات الصدق لبقية النماذج الأخرى لاختبار الرياضيات، وكذلك لم يكن هناك فروق ذات دلالة إحصائية بين معاملات الصدق للنماذج الأربعة لاختبار التاريخ، وأوصت الدراسة بأن يستخدم المعلمون اختبار الاختيار من متعدد ذا الثلاثة بدائل في اختباراتهم التحصيلية لمادة الرياضيات، وأن يستخدموا الاختبار ذا الثلاثة بدائل أو الأربعة بدائل أو الخمسة بدائل لمادة التاريخ.

أما الدراسة التي قام بها الشبلي (١٩٩٥م) فكان هدفها معرفة أثر عدد البدائل في فقرات اختبار الاختيار من متعدد التحصيلي على تجانسه الداخلي، والخصائص

السيكومترية لفقراته ، ولتحقيق أهداف هذه الدراسة تم بناء اختبار تحصيلي في مادة الرياضيات لطلبة الصف العاشر، حيث تكون الاختبار من ثلاثة نماذج فقراتها متشابهة في المتن، ومتساوية في العدد، ولكنها مختلفة في عدد البدائل.

فقد كان عدد البدائل لكل فقرة من فقرات النموذج الأول خمسة ، وفي النموذج الثاني أربعة حيث تم حذف أحد الموهات عشوائياً من كل فقرة من فقرات النموذج الأول وفي النموذج الثالث ثلاثة بدائل حيث تم حذف أحد الموهات عشوائياً من كل فقرة من فقرات النموذج الثاني ، تم تطبيق هذه النماذج على عينة مكونة من (٤٥٠) طالباً تم اختياره بطريقة عشوائية من طلبة الصف العاشر في مديرية تربية جرش وبواقع (١٥٠) طالباً لكل نموذج، وأشارت نتائج هذه الدراسة إلى وجود فروق ذات دلالة إحصائية بين متوسطات معاملات الصعوبة للفقرات ذات الثلاثة بدائل والفقرات ذات الخمسة بدائل ولصالح الفقرات ذات الثلاثة بدائل. كما بينت النتائج عدم وجود فروق ذات دلالة إحصائية بين متوسطات معاملات صعوبة الفقرات ذات الأربعة بدائل من جهة ومتوسط معاملات صعوبة الفقرات ذات الثلاثة بدائل، أو الخمسة بدائل من جهة أخرى. أما بالنسبة لمتوسطات معاملات التمييز فقد أشارت النتائج إلى عدم وجود فروق ذات دلالة إحصائية عند مستوى ( $\alpha = 0,05$ ) بين أي من متوسطات التمييز لفقرات النماذج الثلاثة للاختبار . أما بالنسبة لأثر الاختلاف في عدد البدائل في فقرات اختبار الاختيار من متعدد التحصيلي على تجانسه الداخلي فقد أشارت النتائج إلى عدم وجود فروق ذات دلالة إحصائية بين قيم معاملات الثبات للنماذج الثلاثة فقد كانت معاملات الثبات للنماذج الثلاثة متقاربة،



وكان معامل الثبات للنموذج الذي فقراته خمسة بدائل (٠,٨٦) وللنموذج الذي لفقراته أربعة بدائل (٠,٨٦) ، وللنموذج الذي لفقراته ثلاثة بدائل (٠,٨٥) .

وفي دراسة قام بها الموسوي(١٩٩٧م) وكان الهدف من هذه الدراسة تحديد العدد الأمثل لخيارات الفقرة في اختبار الاختيار من متعدد أي اقل عدد من الخيارات يحقق أقصى قيمة ممكنة لبعض الخصائص السيكومترية للاختبار كالصدق والثبات والقدرة التمييزية عند تطبيقه على مفحوصين ذوي مستويات تحصيلية متباينة ، قام الباحث ببناء اختبار تحصيلي يشتمل على ٤٠ فقرة بأربعة خيارات لكل منها في مقرر القياس والتقويم التربوي، وتم تطبيقه على عينة عشوائية من ١٤٤ طالبا وطالبة من جامعة البحرين، ثم أعيد تطبيقه على نفس المجموعة بعد أسبوعين من التطبيق الأول في أعقاب تقليص عدد خيارات الفقرة إلى ثلاثة، وبعد تصحيح الاختبار قسم الطلبة إلى ثلاث فئات وفقا لدرجة الطالب النهائية في المقرر، والتي اعتمدها الباحث محكا لاحتساب معامل الصدق التنبؤي للاختبار، وقد تم استخراج الفروق بين معاملات الصدق باستخدام الإحصائي  $U$  الذي يتبع توزيع كاي تربيع، وقد تم استخراج الفروق بين معاملات الثبات باستخدام الإحصائي  $M$  ، وقد أبرزت نتائج الدراسة فروقا دالة إحصائيا عند مستوى ٠,٠٥ بين معاملي الصدق التنبؤي لصورتى الاختبار بأربعة وبثلاثة خيارات للفئة الوسطى لصالح الخيارات الثلاثة ، كما لم تبرز النتائج فروقا دالة إحصائيا عند مستوى ٠,٠٥ بين معاملي الثبات، أو بين متوسطي القدرة التمييزية ، لصورتى الاختبار لكل فئة ولجميع الفئات، وبالنتيجة فإن العدد الأمثل لخيارات الفقرة في اختبار الاختيار من متعدد هو ثلاثة، لذا

أوصى الباحث باستخدام هذا النمط من الفقرات بثلاثة خيارات فقط في الاختبارات الفصلية والنهائية لغرض تقويم التحصيل لدى الطلبة الجامعين.

وفي دراسة قام بها ظاظا (٢٠٠٠م) هدفت إلى استقصاء أثر أربع طرق لحذف الموهات في فقرات اختبار الاختيار من متعدد على الخصائص السيكومترية للاختبار وفقراته، ولتحقيق أهداف الدراسة تم استخدام اختبار الاستدلال الحسابي الصورة (ج) المستوى الأول من بطارية القدرات الفارقة والمؤلف من ٤٠ فقرة من نوع الاختيار من متعدد ولكل فقرة من فقرات الاختبار خمسة بدائل، وفي ضوء نتائج التجربة الاستطلاعية تم تكوين ثمانية نماذج من الاختبار الأصلي، ونموذجين بطريقة الحذف التجريبي أحدهما بأربع بدائل والآخر بثلاث بدائل، ونموذجين بطريقة الحذف العشوائي للموهات أحدهما بأربعة بدائل والآخر بثلاثة بدائل، ونموذجين بطريقة الحذف التتابعي للموهات أحدهما بأربعة بدائل والآخر بثلاثة بدائل، وآخر نموذجين بطريقة الحذف بالتحكيم أحدهما بأربعة بدائل والآخر بثلاثة بدائل، وطبقت النماذج الثمانية على ٤٢٦ طالباً من طلاب الصف العاشر بمحافظة عمان، وأشارت نتائج التحليل الإحصائي إلى عدم وجود فروق ذات دلالة إحصائية في معاملات الصعوبة عند مستوى ٠,٠٥ بين النماذج الأربعة ذات الأربعة بدائل، وأيضاً بين النماذج الأربعة ذات الثلاثة بدائل تعزى لطريقة الحذف، كما أشارت النتائج إلى عدم وجود فروق ذات دلالة إحصائية في معاملات التمييز بين النماذج الأربعة ذات الأربعة بدائل تعزى لطريقة الحذف، بينما أظهرت فروق ذات دلالة إحصائية بين النموذجين ذوي الثلاثة بدائل المكونين وفق طريقة الحذف التجريبي والحذف

بالتحكم لصالح الحذف بالطريقة التجريبية، أما النتائج المتعلقة بمعاملات ثبات الاتساق الداخلي ومعاملات الصدق فقد تبين عدم وجود فروق ذات دلالة إحصائية بين النماذج المكونة وفق طرق الحذف المختلفة والتي لها نفس عدد البدائل.

أما بالنسبة لأثر عدد البدائل على صعوبة الفقرة وتمييزها، فقد أظهرت النتائج أن معامل صعوبة الفقرة يزداد بزيادة عدد البدائل، ولا يوجد أثر لعدد البدائل على معاملات التمييز، كما أظهرت النتائج عدم وجود فروق ذات دلالة إحصائية بين معاملات ثبات الاتساق الداخلي بين النموذج الذي عدد بدائله ثلاثة بدائل وأربعة بدائل، ووجدت فروق ذات دلالة إحصائية بين معاملات صدق النموذج الذي عدد بدائله أربعة والنموذج الذي عدد بدائله ثلاثة لصالح النموذج ذي الأربعة بدائل ضمن طريقة الحذف التجريبي، ولم توجد فروق ذات دلالة إحصائية بين معاملات صدق المحك التلازمي تعزى لعدد البدائل للنموذجين ذوي الثلاثة بدائل والأربعة بدائل ضمن بقية طرق الحذف الأخرى.

### التعليق على الدراسات السابقة :

بعد العرض السابق لبعض الدراسات المتعلقة بأثر عدد البدائل على الخصائص السيكومترية للاختبار و فقراته قام الباحث بالتعليق عليها وفق العناصر التالية:

أولاً: الهدف من البحوث والدراسات السابقة:

كان هدف بعض البحوث والدراسات السابقة البحث في أثر عدد البدائل على الخصائص السيكومترية للاختبار ممثلة في قيم معاملات الصعوبة والتمييز لل فقرات، كدراسة كوستين (١٩٧٠م)، ودراسة لورد (١٩٧٧م)، ودراسة روينولدز (١٩٧٩م)، ودراسة كرهان وآخرون (١٩٩٣م)، ودراسة اليعقوب (١٩٩٦م).

بينما كان هدف البعض الآخر من هذه البحوث والدراسات البحث في أثر عدد البدائل على ثبات الاختبار كدراسة جرير (١٩٧٥م)، والدراسة التي قامت بها وبيير (١٩٧٧م)، ودراسة بيدسكي ونيفو (١٩٨٥م)، ودراسة تريفيزان وساكس وميشيل (١٩٩٤م)، أما دراسة تريفيزان وساكس وميشيل (١٩٩١م) فكان الهدف منها معرفة أثر عدد البدائل على صدق وثبات الاختبار.

وجمعت بعض الدراسات والبحوث السابقة في معرفة أثر عدد البدائل على الخصائص السيكمترية للاختبار وفقراته، فنجد دراسة كوستين (١٩٧٢م) بحثت في أثر عدد البدائل على معاملات التمييز لفقرات الاختبار وثباته، أما دراسة ستراتون وكاتس (١٩٨٠م) ودراسة الشبلي (١٩٩٥م) بحثنا في معاملات صعوبة وتمييز فقرات الاختبار وثباته، أما دراسة الموسوي (١٩٩٧م) فكانت تبحث في صدق وثبات الاختبار وقدرته التمييزية، أما دراسة فقوسة (١٩٩٤م) ودراسة ظاظا (٢٠٠٠م) فبحثنا في أثر عدد البدائل على معاملات صعوبة وتمييز الفقرات، وأيضاً صدق وثبات الاختبار.

#### ثانياً: عينة البحوث والدراسات السابقة:

تنوعت عينات الدراسة في البحوث والدراسات السابقة فطبق منها على الطلاب الجامعيين كدراسة كوستين (١٩٧٠م)، ودراسة كوستين (١٩٧٢م)، ودراسة بيدسكو ونيفو (١٩٨٥م)، ودراسة تريفيزان وساكس وميشيل (١٩٩١م)، ودراسة كرهان وآخرون (١٩٩٣م)، ودراسة اليعقوب (١٩٩٦م)، ودراسة الموسوي (١٩٩٧م).

أما بقية الدراسات فقد طبقت على طلاب مراحل التعليم العام ابتداءً من الصف التاسع وحتى الصف الثاني عشر وهي دراسة روينولدز (١٩٧٩م)، ودراسة وبيير (١٩٧٧م)، ودراسة تريفيزان وساكس وميشيل (١٩٩٤م)، ودراسة فقوسة (١٩٩٤م)، ودراسة الشبلي (١٩٩٥م)، ودراسة ظاظا (٢٠٠٠م).

وتراوحت أعداد عينة البحوث والدراسات السابقة بين ١٠٠ طالباً كما في دراسة روينولدز (Roynolds) (١٩٧٩م) إلى ١٢٠٠ طالباً كما في دراسة كوستين Costin (١٩٧٢م).

### ثالثاً: مدى استفادة الباحث من الدراسات السابقة:

لقد ساعدت البحوث والدراسات السابقة الباحث في بناء هذا البحث ابتداءً من تحديد المشكلة، ورسم الأهداف وصياغة الفروض وبناء الإطار النظري، ومروراً بإجراءات الدراسة من اختيار المنهج المناسب ومجتمع وعينة البحث وحجمها، وطريقة اختيار أداة البحث وكيفية تطبيقها، وأرشدت الباحث إلى الأساليب الإحصائية المناسبة لبحثه، كما كانت نتائج الدراسات السابقة إطاراً مرجعياً اعتمد عليه الباحث في تفسير ومناقشة نتائج دراسته.

### رابعاً: أوجه الاتفاق والاختلاف بين نتائج الدراسات السابقة:

يتضح من الدراسات السابقة أن النتائج التي توصل لها الباحثون فيما يختص بأثر عدد البدائل في فقرات اختبارات الاختيار من متعدد على صعوبة الفقرة وتميزها متباينة، فبعض هذه الدراسات أشار إلى وجود فروق ذات دلالة إحصائية بين متوسطات معاملات الصعوبة والتميز كدراسة كوستين (Costin, 1970) أشارت إلى أن معاملات الصعوبة والتميز في النموذج ذي الثلاثة بدائل أعلى منها في النموذج ذي الأربعة بدائل، أما دراسة لورد (Lord, 1977) فقد كانت فقرات النموذج ذي الأربعة بدائل أصعب النماذج وأسهلها فقرات النموذج ذي البديلين، وأثبتت دراسة ستراتون وكاتس (Straton & Catts, 1980) أن عدد البدائل يتناسب عكسياً مع صعوبة الفقرة، وتميز الفقرة يتناسب طردياً مع عدد البدائل، ودراسة كرهان وآخرون (Crehan et al., 1993) حيث كانت هناك فروقا بين متوسطات معاملات الصعوبة لصالح النموذج ذي الثلاثة بدائل، بينما أشارت بعض الدراسات بعدم وجود فروق دالة إحصائية بين متوسطات معاملات الصعوبة

والتمييز، كدراسة كوستين (Costin, ١٩٧٢)، ودراسة روينولدز (Roynolds, ١٩٧٩)، ودراسة الشبلي (١٩٩٥).

ونجد كذلك تبايناً في نتائج هذه الدراسات فيما يتعلق بمعامل ثبات الاختبار فبعضها أشار إلى وجود فروق دلالة إحصائية كدراسة كوستين (Costin, ١٩٧٠)، ودراسة جرير (Grier, ١٩٧٥) ودراسة ستراتون وكاتس (Straton & Catts, ١٩٨٠) حيث كان الاختبار ذو الثلاثة بدائل يمثل أعلى ثبات، أما دراسة بيدسكو ونيفو (Budescu & Nevo, ١٩٨٥) فقد كان معامل الثبات يزداد بزيادة عدد البدائل، بينما أشارت بعض الدراسات بأنه لا توجد فروق ذات دلالة إحصائية بين معاملات الثبات لبعض الاختبارات كدراسة كوستين (Costin, ١٩٧٢)، ودراسة تريفيزان وآخرون (Trevisan et al., ١٩٩٤)، ودراسة فقوسة (١٩٩٤م)، ودراسة الشبلي (١٩٩٨م).

أما من حيث الصدق فلم توجد فروق ذات دلالة إحصائية بين معاملات صدق المحك التلازمي كدراسة فقوسة (١٩٩٤)، ودراسة الموسوي (١٩٩٧م) ودراسة ظاظا (٢٠٠٠م).

#### **خامساً: اختلاف البحث عن الدراسات السابقة:**

البحث الحالي يعد من البحوث الأولى - على حد علم الباحث - في المجتمع السعودي، حيث لم يجد الباحث في قواعد المعلومات ومراكز الدراسات والبحوث دراسات أو بحوث تتعلق بأثر عدد البدائل على الخصائص السيكومترية للاختبار، كما

تميز هذا البحث بالجمع بين استخدام طريقتين لإيجاد الصدق المرتبط بمحك، وهما صدق المحك التلازمي وصدق المحك التنبؤي.

وهكذا يتضح من الدراسات السابقة عدم وجود اتساق في النتائج المتعلقة بأثر عدد البدائل على الخصائص السيكومترية للاختبار وفقراته، فأتى هذا البحث محاولاً تقصي الحقيقة، وللخروج بنتائج مؤيدة أو معارضة للدراسات السابقة أو إضافة متواضعة عليها، هذا ما سعى إليه هذا البحث.

### **ثالثاً: فروض البحث:**

- لا توجد فروق ذات دلالة إحصائية عند مستوى (0,05) بين متوسطات قيم معاملات السهولة ( $\Delta$ ) في نماذج الاختبار التحصيلي لمادة الرياضيات من نوع الاختيار من متعدد ذات البدائل ثلاثة وأربعة وخمسة.

- لا توجد فروق ذات دلالة إحصائية عند مستوى (٠,٠٥) بين متوسطات قيم الدرجات الزائفة المقابلة لمعاملات التمييز في نماذج الاختبار التحصيلي لمادة الرياضيات من نوع الاختيار من متعدد ذات البدائل ثلاثة و أربعة و خمسة.
- لا توجد فروق ذات دلالة إحصائية عند مستوى (٠,٠٥) بين قيم معاملات ثبات الإتساق الداخلي في نتائج نماذج الاختبار التحصيلي لمادة الرياضيات من نوع الاختيار من متعدد ذات البدائل ثلاثة و أربعة و خمسة.
- لا توجد فروق ذات دلالة إحصائية عند مستوى (٠,٠٥) بين قيم معاملات صدق المحك (التلازمي - والتنبؤي) في نتائج نماذج الاختبار التحصيلي لمادة الرياضيات من نوع الاختيار من متعدد ذات البدائل ثلاثة و أربعة و خمسة.



- 
- 
- 
- 
- 
-

## منهج البحث

بعد أن قام الباحث بتحديد مشكلة البحث، وهي معرفة أثر عدد البدائل على الخصائص السيكومترية للاختبار التحصيلي للرياضيات، وبعد الإطلاع على الدراسات والبحوث السابقة، ومراجعة العديد من المناهج البحثية، توصل الباحث إلى أن المنهج الملائم للبحث الحالي هو المنهج الوصفي، لمناسبته في الإجابة على تساؤلات البحث، والذي يقول عنه عبيدات وآخرون (١٩٩٣م: ٢١٩) بأنه "يعتمد على دراسة الظاهرة كما توجد في الواقع ويهتم بوصفها وصفاً دقيقاً، ويعبر عنها تعبيراً كيفياً أو كمياً، فالتعبير الكيفي يصف لنا الظاهرة ويبين خصائصها، بينما التعبير الكمي يعطينا وصفاً رقمياً لمقدار الظاهرة أو حجمها، ودرجات ارتباطها مع الظواهر المختلفة الأخرى".

## مجتمع البحث

يقصد بمجتمع البحث كما ذكر عبيدات وآخرون (١٩٩٣م: ١٠٩) بأنه "جميع الأفراد أو الأشخاص أو الأشياء الذين يكونون موضوع مشكلة البحث" ويتكون مجتمع البحث الحالي من جميع طلاب الصف الثاني المتوسط بالمدارس الحكومية بمحافظة الليث، التابعة لإدارة التربية والتعليم بمحافظة الليث للعام الدراسي ١٤٢٨هـ - ١٤٢٩هـ والبالغ عددهم الإجمالي (١٢١٨) طالباً يتوزعون في (٤٧) مدرسة متوسطة، ضمن أربعة قطاعات للمحافظة، قطاع أضم ويتبعه (١٥) مدرسة، قطاع الشاقة ويتبعه (١٠) مدارس، ويليه قطاع الليث ويوجد به (١١) مدرسة، وقطاع حفار وفيه (١١) مدرسة.

### **عينة البحث**

نظراً لصعوبة تطبيق أداة هذا البحث على جميع أفراد المجتمع، قام الباحث بأخذ عينة طبقية عشوائية من مدارس المجتمع الأصلي تكونت من (٩) مدارس، وذلك بعد الحصول على آخر إحصائية لأعداد المدارس المتوسطة بمحافظة الليث وأعداد طلاب الصف الثاني المتوسط من قسم الحاسب الآلي بإدارة التربية والتعليم بمحافظة الليث للعام الدراسي ١٤٢٨هـ - ١٤٢٩هـ، وتم تقسيم مجتمع البحث إلى أربعة قطاعات (أضم، الشاقة، الليث، حفار)، وتم اختيار عينة عشوائية بطريقة المعاينة العشوائية المنتظمة داخل كل قطاع، تكونت من طلاب (٩) مدارس وذلك كالتالي:

أ- تحديد طول الفترة ( المسافة الفاصلة بين رقم المدرسة التي نختارها ورقم

المدرسة التي تليها) وذلك كالتالي:

عدد المدارس بالمجتمع ٤٧

تحديد طول الفترة = \_\_\_\_\_ = ٥,٢٢ = ٥

عدد مدارس العينة ٩

ب- تحديد بداية الفترة وذلك باختيار رقم عشوائي داخل طول الفترة (١ - ٥)،  
وتم اختيار المدرسة رقم (٥).

ج - تم الحصول على المدارس الثلاثة التي تحمل الأرقام (٥ - ١٠ - ١٥) من قطاع  
أضم، وعلى مدرستين من قطاع الشاقة تحمل الأرقام (٥ - ١٠)، وعلى مدرستين  
من قطاع الليث تحمل الأرقام (٥ - ١٠)، وعلى مدرستين من قطاع حفار تحمل  
الأرقام (٥ - ١٠)، بحيث أصبح عدد مدارس العينة (٩) مدارس بنسبة ١٩٪ من  
مدارس المجتمع الأصلي، وفيها (٣٥١) طالباً وهو ما يمثل ٢٩٪ من نسبة عدد  
الطلاب في مجتمع البحث.

وتم تطبيق النماذج الثلاثة للاختبار بشكل عشوائي على طلاب الصف الثاني  
المتوسط في مدارس العينة، بحيث يقسم كل صف إلى ثلاثة أقسام عشوائياً، وطبق كل  
نموذج من النماذج الثلاثة عشوائياً على الثلاثة أقسام بشكل جماعي داخل كل فصل،  
والجدول رقم (١) يبين توزيع عينة البحث والنماذج الثلاثة على الطلاب.

جدول رقم (١)

توزيع مدارس عينة البحث ونماذج الاختبار الثلاثة على الطلاب

					( )		

## أدوات البحث

هدف هذا البحث إلى مقارنة أثر عدد البدائل على الخصائص السيكمومترية للاختبار التحصيلي للرياضيات، ولتحقيق أهداف هذا البحث قام الباحث بإعداد اختبار تحصيلي في مادة الرياضيات للصف الثاني المتوسط، وفيما يلي المراحل والإجراءات التي اتبعها الباحث في بناء أداة البحث:

أولاً: تحديد الموضوعات المراد قياسها:

اختار الباحث الوجدتين الأولى والثانية من مقرر مادة الرياضيات لطلاب الصف الثاني المتوسط (الفصل الدراسي الأول) للعام الدراسي ١٤٢٨هـ - ١٤٢٩هـ، وهاتين الوجدتين تمثلان ٦٦٪ من المقرر الدراسي:

• الوحدة الأولى: العلاقات والتطبيقات.

• الوحدة الثانية: مجموعة الأعداد النسبية.

**ثانياً: تحديد الأهداف السلوكية:**

قام الباحث بتحديد الأهداف السلوكية التي تقيس محتوى الوجدتين الأولى والثانية من مقرر الرياضيات لطلاب الصف الثاني المتوسط (الفصل الدراسي الأول)، وتم الاعتماد على الأهداف السلوكية التي وضعها نخبة من المتخصصين في مشروع الاختبارات المدرسية بوزارة التربية والتعليم حيث قامت كل من إدارتي التربية والتعليم بمنطقة مكة المكرمة والإحساء، بتحليل محتوى المادة إلى الحقائق والمفاهيم والقوانين والمهارات المتضمنة في كتاب الطالب، وكتب على هذا الأساس الأهداف السلوكية للمقرر، حيث عرضت هذه الأهداف على عدد من المحكمين لتقييمها والتحقق من انطباق شروط ومعايير مواصفات الهدف السلوكي الجيد عليها، وكذلك شمولية هذه الأهداف لمستويات بلوم المعرفية، وبالتالي تكون هذه الأهداف حكمت بصورة جيدة وموثوق فيها، ويمكن للباحث الاعتماد عليها في بناء أداة دراسته.

**ثالثاً: إعداد جدول المواصفات:**

لإعداد جدول المواصفات قام الباحث بالخطوات التالية :

أ - تحديد الوزن النسبي للأهداف السلوكية بمستوياتها المختلفة من خلال المعادلة

رقم (١) في الفصل السابق.

ب - تحديد عدد الحصص اللازمة لتدريس كل موضوع، حيث قام الباحث بتصميم

جدول يتضمن العمود الأول موضوعات الفصل الأول والثاني من مقرر الرياضيات للصف

الثاني المتوسط، والعمود الثاني عدد الحصص المقترحة لكل موضوع، وتم توزيعها على

عدد من المشرفين التربويين و وعدد من معلمي الرياضيات للصف الثاني المتوسط ملحق

رقم(١).

ج - تحديد الوزن النسبي لموضوعات المادة الدراسية من خلال المعادلة رقم (٢) في

الفصل السابق.

د - تحديد العدد الكلي لأسئلة الاختبار في ضوء الزمن المتاح للإجابة، ونوع الفقرة

المستخدمة في الاختبار، وأعمار الطلبة ومستواهم التعليمي، وطول الفقرة ومستوى

صعوبتها، ومقدار ما تتطلبه الفقرات من حساب كمية، ومستوى العمليات العقلية التي

تقيسها الفقرة.

هـ - حساب عدد الأسئلة في كل موضوع لكل مستوى من مستويات الأهداف من خلال

المعادلة رقم (٣) في الفصل السابق.

- تحديد درجة أسئلة كل موضوع في كل مستوى من مستويات الأهداف من خلال

المعادلة رقم(٤) في الفصل السابق.

وتوصل الباحث إلى جدول المواصفات التالي جدول رقم (٢) :

جدول رقم (٢)

جدول مواصفات الاختبار التحصيلي للوحدتين الأولى والثانية للفصل الدراسي الأول من مادة الرياضيات لطلاب الصف

الثاني المتوسط للعام ١٤٢٨هـ - ١٤٢٩هـ

المحتوى	الموضوعات	الأهداف التربوية				الوزن النسبي للموضوعات	
		التذكر	الفهم	التطبيق	التحليل		
الوحدة الأولى العلاقات والتطبيقات	ضرب المجموعات	١	١	١	٠	٣	١٠%
	العلاقات وتمثيلها	٠	١	١	٠	٢	١٠%
	التطبيق	١	٠	٠	١	٢	٥%
	أنواع التطبيق	٠	١	١	٠	٢	٥%
مجموعة الأعداد النسبية الوحدة الثانية	مجموعة الأعداد النسبية	١	١	١	٠	٣	١٠%
	ضرب وقسمة الأعداد النسبية	١	١	٠	١	٣	١٠%
	جمع وطرح الأعداد النسبية	٠	١	١	١	٣	١٠%
	مقارنة الأعداد النسبية	١	١	١	٠	٣	١٠%
	تبسيط العبارات العددية	١	٢	٠	٠	٣	١٠%
	الكتابة العشرية	١	١	١	٠	٣	١٠%
	قوى عدد نسبي	١	١	١	٠	٣	١٠%
	المجموع	٨	١١	٨	٣	٣٠	
	الاوزان النسبية للأهداف	٢٧%	٣٦%	٢٧%	١٠%		١٠٠%

رابعاً: كتابة وصياغة الأسئلة:

قام الباحث بكتابة وصياغة اختباراً تحصيلياً في مادة الرياضيات لطلاب الصف

الثاني المتوسط، في الوحدة الأولى (العلاقات والتطبيقات) والوحدة الثانية (مجموعة



الأعداد النسبية)، وتم إضافة سؤال إضافي في كل موضوع قبل عرض جدول المواصفات على المحكمين بحيث أصبح عدد فقرات الاختبار ٤١ فقرة من نوع الاختيار من متعدد، ولكل فقرة خمسة بدائل.

### خامساً: تحكيم أداة البحث:

تم عرض جدول المواصفات وفقرات الاختبار على عدد من المتخصصين في مادة الرياضيات، وفي القياس والتقويم، وذلك من خلال استبيان للحكم على مدى ملائمة جدول المواصفات وفقرات الاختبار التحصيلي في الملحق رقم (١) وتم الاجتماع مع المحكمين وأخذ وجهات نظرهم، حيث تم حذف ثلاث فقرات ليصبح عدد فقرات الاختبار ٣٨ فقرة، وذلك لتشابه هذه الفقرات مع فقرات أخرى في نفس النموذج، وأجريت بعض التعديلات على صياغة بعض الفقرات وبعض البدائل، والملحق رقم (١) يمثل بيان بأسماء المحكمين لأدوات البحث.

### سادساً: توفير أدوات الدراسة:

قام الباحث بتوفير عدد كاف من أوراق الإجابة الخاص بأجهزة التصحيح الآلي، وتوفير أقلام الرصاص ليستخدمها الطلاب في تظليل دائرة حرف البديل الصحيح في كل نموذج، وأيضاً توفير نسخ كافية من نماذج الاختبارات.

### سابعاً: التجربة الاستطلاعية:

قام الباحث بتطبيق الاختبار المكون من ٣٨ فقرة على عينة استطلاعية تكونت من ٦٠ طالباً، تم اختيارهم عشوائياً من إحدى المدارس من خارج عينة البحث، وتم إبلاغ إدارة المدرسة 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.

وكانت نتائج هذه التجربة تشير إلى وضوح تعليمات الاختبار من خلال أداء الطلاب على الاختبار وقلة استفساراتهم، وتم تحديد الزمن اللازم للإجابة على الاختبار وقدر بـ (٥٠) دقيقة، وذلك عن طريق المعادلة التالية:

$$\text{زمن أسرع طالب} + \text{زمن أبطأ طالب} \\ \text{زمن الاختبار} = \frac{\text{-----}}{2} - - - (١٧)$$

وتم إجراء التحليل الأولي التالي لمفردات الاختبار:

- حساب معامل سهولة كل مفردة من المفردات؛ وهي نسبة عدد الطلاب الذين أجابوا على المفردة إجابة صحيحة.
- حساب معامل تمييز كل مفردة من المفردات وهو معامل الارتباط بين الدرجة على الفقرة من جهة والدرجة الكلية على الاختبار من جهة أخرى، أي بمعامل الارتباط الثنائي الأصيل.
- حساب فعالية الموهبات (المشتتات).

• حساب معامل الثبات بطريقة التجزئة النصفية ومعامل الإتساق الداخلي.

والجدول رقم ( ٣ ) يوضح معاملات الصعوبة والتمييز للعينات الاستطلاعية.

جدول رقم ( ٣ )

معاملات السهولة والتمييز وفعالية المشتتات للعينات الاستطلاعية

- ١٠٧ -

*	. -	. -	. -	. -	.	.	
	. -	*	. -	. -	.	.	
.	. -	. -	*	. -	.	.	
. -	. -	. -	. -	*	.	.	
. -		.	*	. -	.	.	
. -	*	. -		. -	.	.	
. -	. -	. -	*	. -	.	.	
.	*	. -	. -	. -	.	.	
. -	. -	*	. -	.	.	.	
. -		. -	. -	*	.	.	
.	.	*	.	. -	.	.	
. -	. -	. -	*	.	.	.	
. -	*	. -	. -	.	.	.	
.	.	.	*	.	. -	.	
. -	.	. -	. -	*	.	.	
. -	*	. -	. -	. -	.	.	
. -	. -		*	. -	.	.	
. -	. -	*	. -	. -	.	.	
.	.	.	*	. -	.	.	
. -	. -		. -	*	.	.	

.	-	.	*	.	-	.	-	.	.	.	.	.	.
.	-	.	-	.	-	.	-	.	*	.	.	.	.
.	-	.	-	.	-	.	-	.	*	.	.	.	.
.	-	*	.	-	.	-	.	-	.	.	.	.	.
.	-	.	-	*	.	-	.	-	.	.	.	.	.
.	-	*	.	.	-	.	-	.	-	.	.	.	.
.	-	*	.	-	.	-	.	-	.	.	.	.	.
.	-	.	.	-	*	.	-	.	-	.	.	.	.
.	-	.	.	-	*	.	-	.	-	.	.	.	.
.	-	.	.	-	*	.	-	.	-	.	.	.	.
.	-	.	.	-	*	.	-	.	-	.	.	.	.
.	-	.	.	-	*	.	-	.	-	.	.	.	.
.	-	.	.	-	*	.	-	.	-	.	.	.	.
.	-	.	.	-	*	.	-	.	-	.	.	.	.
.	-	.	.	-	*	.	-	.	-	.	.	.	.
.	-	.	.	-	*	.	-	.	-	.	.	.	.
.	-	*	.	-	.	-	.	-	.	.	.	.	.

وللحكم على فقرات الاختبار اعتمد الباحث على المعايير التالية:

١. حذف المفردات التي تكون معاملات سهولتها أقل من (٠,٢٠) وأكبر من (٠,٨٠).

٢. حذف المفردات التي يكون معامل تمييزها أقل من (٠,١٠)، وتحتاج المفردة إلى تعديل إذا كان معامل تمييزها يتراوح بين (٠,١٠) و (٠,٢٠)، بينما معاملات التمييز التي تتراوح بين (٠,٢٠) و (٠,٣٠) فهي منخفضة ولكن لا بأس بها، والحد الأدنى لمعامل التمييز (٠,٣٠) على الأقل وأكثر من (٠,٤٠) تدل على قوة تمييز مرتفعة.

٣. إعادة النظر في الموهات التي تكون جاذبيتها موجبة أو صفر.

و بالاعتماد على ما سبق تم حذف المفردات (٥، ١١، ١٤، ١٥، ١٩، ٢٦، ٢٩، ٣٢)

حيث كانت معاملات التمييز لتلك المفردات على التوالي:

(٠,٠٨، ٠,٠٨ - ٠,٠٨، ٠,٠٨، صفر، ٠,٠٨، ٠,٠٨ - ٠,١٧، صفر).

ومعاملات سهولتها: (٠,٦٥، ٠,١٢، ٠,٢٣، ٠,٩٣، ٠,١٩، ٠,٢١، ٠,١٤، ٠,٢٦).

وبدراسة فعالية الموهات تم استبدال بعض الموهات أو إعادة صياغتها ومراجعتها ،

لأن الموه (أ) للفقرات (٩، ١٢، ١٣، ٢٤، ٢٥، ٣١، ٣٨) كان موجباً أو صفر، وأيضا

الموه (ب) للمفردتين (٦، ٢٢) فكانت جاذبيتها على الترتيب (صفر، ٠,٠٨)، أما الموه

(ج) للمفردات (١٧، ٢٠، ٣٥) كانت جاذبيتها صفراً، والبديل (د) للمفردتين (١٠، ٣٥)

كان صفراً، وللمفردتين (٢١، ٢٨) فجاذبيتها (٠,٠٨، ٠,٠٣)، أما الموه (هـ) كان

صفراً للمفردتين (٢، ٢٧) وكان (٠,١١، ٠,٠٨) للمفردتين (٣، ٨) على الترتيب.

وباستخدام معادلة ألفا كرونباخ تم حساب معامل الثبات وكانت قيمته ٠,٧٥٦ ،  
وأعيد حساب الثبات بعد حذف المفردات الثمانية بطريقة التجزئة النصفية، و معامل  
الاتساق الداخلي، والجدول رقم(٤) التالي يبين قيم معاملات الثبات:

جدول رقم(٤)

قيم معاملات ثبات التجزئة النصفية والاتساق الداخلي للعينات الإستطلاعية

طريقة الثبات	قيمة معامل الثبات
التجزئة النصفية (معادلة جتمان لتصحيح الثبات)	٠,٧٥٠
معادلة كرونباخ (الاتساق الداخلي)	٠,٧٨٠

ونلاحظ من الجدول السابق أن قيمة معامل الثبات بالتجزئة النصفية كانت بمعادلة  
جتمان ٠,٧٥٠ ، وبمعادلة كرونباخ ٠,٧٨٠ .

وفي ضوء التجربة الاستطلاعية تم تكوين ثلاثة نماذج للاختبار نفسه، يتكون كل  
منها من ٣٠ مفردة على النحو التالي:

النموذج الأول: احتوت كل مفردة من مفرداته على خمسة بدائل.

النموذج الثاني: احتوت كل مفردة من مفرداته على أربعة بدائل، وذلك بحذف أحد

الموهبات عشوائياً من كل مفردة من مفردات النموذج الأول.

النموذج الثالث: احتوت كل مفردة من مفرداته على ثلاثة بدائل، وذلك بحذف أحد الموهات عشوائياً من كل مفردة من مفردات النموذج الثاني.

والملاحق رقم (٢) يبين النماذج الثلاثة بالصورة النهائية.

## متغيرات البحث

المتغير المستقل: عدد البدائل في اختبار الاختيار من متعدد خمسة بدائل وأربعة بدائل وثلاثة بدائل.

المتغير التابع: الخصائص السيكومترية لمفردات الاختبار (معاملات السهولة والتمييز)، والخصائص السيكومترية للاختبار (معامل الثبات والصدق المرتبط بمعك).

## الأساليب الإحصائية

تم استخدام الحزمة الإحصائية SPSS لتحليل البيانات واستخراج وحساب كل

من:

• حساب معامل السهولة لكل مفردة وذلك بإيجاد نسبة من أجابوا على المفردة

إجابة صحيحة إلى من حاولوا الإجابة عليها وذلك باستخدام المعادلة رقم (٥).



• حساب معامل التمييز لكل مفردة من مفردات نماذج الاختبار وذلك بإيجاد معامل

الارتباط بين درجات الطلبة على هذه المفردة ودرجاتهم على الاختبار الكلي ،

وذلك باستخدام معامل الارتباط الثنائي الأصيل ، وهي المعادلة رقم (٨).

• حساب معاملات ثبات الاتساق الداخلي لكل نموذج وذلك باستخدام معادلة ألفا

كرونباخ وهي معادلة رقم (١٦).

• حساب معامل الصدق المرتبط بمحك ( التلازمي - و التتبؤي) لكل نموذج وذلك

بحساب معامل ارتباط بيرسون بين درجات الطلبة على الاختبار ودرجاتهم في

اختبار منتصف الفصل الدراسي الأول واختبار نهاية الفصل الدراسي الأول لمادة

الرياضيات من العام ١٤٢٨ - ١٤٢٩ هـ .

• استخدام تحليل التباين الأحادي (ANOVA) عند مستوي (٠,٠٥) لفحص

الفروق بين متوسطات معاملات السهولة ( $\Delta$ ) و الدرجات الزائفة المقابلة لمعاملات

التمييز بين النماذج الثلاثة.

• اختبار (ت) للعينة الواحدة عند درجة القطع المثالية لمعامل الصعوبة (٠,٥٠).

• استخدام الإحصائي (M) المقترح من قبل هاكستين وولين (Hakstain &

Whalen, ١٩٧٦:٢٢٤) والذي يتبع توزيع كاي تربيع بدرجات حرية تساوي

(عدد معاملات الثبات - ١) ، وذلك للكشف عن الاختلاف بين عدة معاملات

ثبات الإتساق الداخلي ، علما أن قيمة (M) تحسب من المعادلة رقم (١٧) التالية:

$$(17) \quad M = \frac{J-1}{18J} \left\{ \sum_{k=1}^K B_k - \frac{\left[ \sum_{k=1}^K B_k (1-r_{ak})^{-1/3} \right]^2}{\sum_{k=1}^K B_k (1-r_{ak})^{-2/3}} \right\}$$

$$B_k = \frac{(9n_k - 11)^2}{(n_k - 1)}$$

:

: n : J

.k :rk

• الاختبار الإحصائي V الذي وضعه هيز (Hays) والذي يتبع توزيع كاي تربيع

بدرجات حرية تساوي (عدد معاملات الارتباط - 1)، والخاص بمعرفة دلالة

الفروق بين عدة معاملات ارتباط بيرسون أي بين معاملات الصدق التلازمي

والصدق التنبؤي، والمعادلة رقم (18) تمثل الإحصائي V بعد تحويل معاملات

الارتباط إلى قيم Z بطريقة Fisher:

( )

$$V = \sum_1^k (n_i - 3)(z_i - u)^2, \quad u = \frac{\left[ \sum_1^k (n_i - 3) z_i \right]}{\left[ \sum_1^k (n_i - 3) \right]}, \quad z_i = \frac{1}{2} \ln \left[ \frac{(1+r_i)}{(1-r_i)} \right]$$

نقلا عن / <http://home.ubalt.edu/ntsbarsh/Business-stat/opre004.htm>

:

:

## أولاً: نتائج البحث :

هدف هذا البحث إلى معرفة أثر عدد البدائل على الخصائص السيكومترية للاختبار وفقراته، ومن أجل ذلك تم تطبيق اختبار تحصيلي في مادة الرياضيات للصف الثاني متوسط في الوجدتين الأولى والثانية من مقرر الفصل الدراسي الأول ويتألف من ثلاثة نماذج :

النموذج الأول: عبارة عن مجموعة من المفردات، تتكون كل مفردة من المتن وخمسة بدائل، أحدها فقط صحيح، ويطلب من المفحوص أن يختار الإجابة الصحيحة من بين الخمسة بدائل لكل مفردة من مفردات الاختبار.

النموذج الثاني: يكون باحتواء كل مفردة من مفردات النموذج الأول على أربعة بدائل، وذلك بحذف أحد الموهات عشوائياً من النموذج الأول.

النموذج الثالث: يكون باحتواء كل مفردة من مفردات النموذج الثاني على ثلاثة بدائل، وذلك بحذف أحد الموهات عشوائياً من النموذج الثاني.

وبالتالي فإن النماذج الثلاثة ثابتة المتن وأن الاختلاف في عدد البدائل، فالأول خمسة بدائل والثاني أربعة وذلك بحذف أحد الموهات عشوائياً من النموذج الأول، والنموذج الثالث ثلاثة بدائل بحذف أحد الموهات عشوائياً من النموذج الثاني، وقد تم عرض النتائج وفقاً لترتيب فروض البحث:

أ: النتائج المتعلقة بأثر اختلاف عدد البدائل (ثلاثة، أربعة، خمسة) على الخصائص السيكومترية للمفردات (الصعوبة و التمييز):

بالنسبة لمعاملات صعوبة الفقرات فقد حسبت في كل نموذج من النماذج الثلاثة بإيجاد نسبة الطلاب الذين أجابوا عن المفردة إجابة صحيحة من بين الطلاب الذين حاولوا الإجابة عن هذه المفردة، ولأن هذه النسبة تقع ضمن المستوى الترتيبي، لذا تم تحويلها إلى قيم دلتا ( $\Delta$ ) المناظرة حيث  $\Delta = 13 + 4 \times (Z)$ ، حيث ( $Z$ ) الدرجة المعيارية المناظرة لمعامل الصعوبة و تستخرج من الجداول الخاصة بالتوزيع الاعتنالي، والجدول رقم (٦) يوضح ملخصاً لجدول الملحق رقم (٤) لتكرارات معاملات صعوبة المفردات للنماذج الثلاثة للاختبار:

جدول ( ٦ )  
تكرارات معاملات الصعوبة للنماذج الثلاثة للاختبار

متوسط قيم دلتا المناظرة لمعاملات الصعوبة	٠,٧١ - ١		٠,٣٠ - ٠,٧٠		٠ - ٠,٢٩		النموذج
	سهل		متوسط		صعب		
	نسبتها	عدد الفقرات	نسبتها	عدد الفقرات	نسبتها	عدد الفقرات	
١٣,٤٢	٠	٠	%٩٠	٢٧	%١٠	٣	نموذج ٣ بدائل
١٢,٩٥	%٣	١	%٧٧	٢٣	%٢٠	٦	نموذج ٤ بدائل
١٢,٥٩	٠	٠	%٨٠	٢٤	%٢٠	٦	نموذج ٥ بدائل

ومن خلال الجدول السابق رقم (٦) يمكن ملاحظة التالي:

١. بلغت نسبة المفردات الصعبة ١٠٪ من مجمل مفردات النموذج ذي الثلاثة بدائل،

و ٢٠٪ من مجمل مفردات النموذجين ذي الأربعة بدائل والخمسة بدائل.

٢. بلغت نسبة المفردات متوسطة الصعوبة ٧٧٪ من مجمل مفردات النموذج ذي الأربعة

بدائل، و ٨٠٪ من مجمل مفردات النموذج ذي الخمسة بدائل، و ٩٠٪ من مجمل

مفردات النموذج ذي الثلاثة بدائل.

٣. بلغت نسبة المفردات السهلة ٣٪ في النموذج ذي الأربعة بدائل، وصفاً في النموذجين

ذي الثلاثة بدائل والخمسة بدائل.

٤. بلغ متوسط معاملات الصعوبة بعد تحويلها إلى قيم دلتا المناظرة للنموذج ذي

الثلاثة بدائل ١٣,٤٢ وهي أعلى قيمة، بينما نجد قيمتها للنموذج ذي الأربعة بدائل

١٢,٩٥، أما النموذج ذي الخمسة بدائل فقد بلغت قيمتها ١٢,٥٩ وهي الأدنى بين

قيم معاملات الصعوبة وبالتالي فهي تتناسب عكسياً مع عدد البدائل.

أما بالنسبة لمعاملات تمييز الفقرات فقد حسبت لكل نموذج من النماذج الثلاثة بإيجاد

معامل الارتباط بين درجات الطلاب على هذه الفقرة ودرجاتهم على الاختبار الكلي

باستخدام معامل الارتباط الثنائي المتسلسل الحقيقي، والجدول رقم (٧) يوضح ملخصاً

لجداول الملحق رقم (٤) لتكرارات المفردات الاختبارية بحسب قدرتها التمييزية في النماذج الثلاثة للاختبار:

جدول (٧)

تكرارات المفردات الاختبارية مصنفة بحسب معاملات التمييز

متوسط قيم Z بطريقة فشر	أكبر من ٠,٤٠		٠,٣١ - ٠,٤٠		٠,٢١ - ٠,٣٠		٠,٢٠ - ٠		أقل من ٠	النموذج
	تمييز مرتفع		تمييز متوسط		تمييز مقبول تحتاج إلى تعديل طفيف		تمييز ضعيف تحتاج إلى تعديل كبير		تحذف الفقرة	
	نسبتها	عدد الفقرات	نسبتها	عدد الفقرات	نسبتها	عدد الفقرات	نسبتها	عدد الفقرات	عدد الفقرات	
٠,٣٥	%٢٧	٨	%٤٣	١٣	%٢٠	٦	%١٠	٣	٠	نموذج ٣ بدائل
٠,٣٥	%٢٣	٧	%٤٧	١٤	%٢٣	٧	%٧	٢	٠	نموذج ٤ بدائل
٠,٣٩	%٣٣	١٠	%٤٧	١٤	%١٣	٤	%٧	٢	٠	نموذج ٥ بدائل

ومن خلال الجدول السابق رقم (٧) يمكن ملاحظة التالي:

١. لا توجد أي مفردة تمييزها أقل من صفر في النماذج الثلاثة.
٢. قلة عدد المفردات التي معامل تمييزها يقع بين ٠ و ٠,٢٠ في جميع النماذج.
٣. ارتفاع عدد المفردات التي معامل تمييزها أكبر من ٠,٣ في النماذج الثلاثة، حيث تمثل هذه المفردات أكثر من ٧٠٪ من مفردات الاختبار.
٤. تقارب متوسطات معاملات التمييز المحولة إلى قيم Z بطريقة فشر، فكان أعلاها في النموذج ذي الخمسة بدائل، وأدناها في النموذجين ذي الثلاثة بدائل والأربعة بدائل.

أما بالنسبة لمعاملات تمييز البدائل الصحيحة التي عددها ٣٠ بديلاً في كل نموذج والتي يجب أن تكون موجبة التمييز، وكذلك معاملات التمييز للمشتتات أو البدائل الخاطئة والتي يجب أن تكون سالبة التمييز، ويلخصها الجدولين (٨) و (٩) التاليين:

جدول (٨)

القدرة التمييزية للبدائل الصحيحة لكل نموذج

البدائل الصحيحة وعددها ٣٠ بديل في كل نموذج				النموذج
نسبتها	الموجبة	نسبتها	السالبة أو الصفرية	
%١٠٠	٣٠	٠	٠	٣ بدائل
%١٠٠	٣٠	٠	٠	٤ بدائل
%١٠٠	٣٠	٠	٠	٥ بدائل

جدول (٩)

القدرة التمييزية للبدائل الخاطئة (المشتتات أو الموهبات)

النموذج	عدد المشتتات	المشتتات غير الفعالة (الموجبة)	نسبتها	المشتتات الفعالة وغير الجذابة	نسبتها	المشتتات الفعالة والجذابة	نسبتها
٣ بدائل	٦٠	٥	%٨	١	%٢	٥٤	%٩٠
٤ بدائل	٩٠	١٢	%١٣	١	%١	٧٧	%٨٦
٥ بدائل	١٢٠	٢٥	%٢٠	٦	%٥	٨٩	%٧٥

ومن خلال الجدولين السابقين (٨) و (٩) يمكن ملاحظة التالي:

١. جميع قيم معاملات التمييز للبدائل الصحيحة موجبة التمييز.
٢. قلة المشتتات غير الفعالة والموجبة في جميع النماذج فكانت نسبتها في النموذج ذي الثلاثة بدائل ٨٪، و ١٣٪ في النموذج ذي الأربعة بدائل، و ٢٠٪ في النموذج ذي الخمسة بدائل، وبالتالي فإن نسبة البدائل غير الفعالة تزداد بزيادة عدد البدائل وهذا تناسب طردي.
٣. زيادة نسبة المشتتات الفعالة والجذابة في جميع النماذج ونعني بالمشتت الجذاب أنه جذب أكثر من ٥٪ من عدد طلاب عينة الدراسة، فكانت نسبتها في النموذج ذي الثلاثة بدائل ٩٠٪، بينما نجدها في النموذج ذي الأربعة بدائل بلغت



٨٦٪، و٧٥٪ للنموذج ذي الخمسة بدائل، وبالتالي فإن نسبة البدائل الفعالة

الجذابة تقل بزيادة عدد البدائل للفقرة وهذا تناسب عكسي.

٤. أما البدائل الفعالة وغير الجذابة أي جذبت أقل من ٥٪ من عدد طلاب عينة

الدراسة، فكانت فقرة واحدة في النموذجين ذي الثلاثة بدائل والأربعة بدائل،

وخمسة فقرات في النموذج ذي الخمسة بدائل.

نتيجة الفرض الأول وتفسيرها :

ينص الفرض الأول على أنه لا توجد فروق ذات دلالة إحصائية عند مستوى (٠,٠٥)

بين متوسطات قيم معاملات الصعوبة في نماذج الاختبار التحصيلي لمادة الرياضيات من

نوع الاختيار من متعدد ذات البدائل ثلاثة أو أربعة أو خمسة.

وللتحقق من صحة الفرض السابق قام الباحث بتحويل معاملات صعوبة الفقرات إلى

قيم دلتا المناظرة ( $\Delta$ ) والتي هي درجات معيارية محولة بوسط مقداره (١٣)، وانحراف

معياري (٤)، وذلك لان قيم دلتا ( $\Delta$ ) المناظرة تقع على مقياس فنوي وحسب المعادلة

التالية:

ولمعرفة مدى الاختلاف في متوسطات معاملات الصعوبة بين النماذج المختلفة (النموذج

ذو الثلاثة بدائل والنموذج ذو الأربعة بدائل والنموذج ذو الخمسة بدائل) تم استخدام تحليل

التباين الأحادي للكشف عن الفروق في متوسطات معاملات الصعوبة بين النماذج الثلاثة، وذلك بعد التأكد من توفر شروط استخدام تحليل التباين الأحادي على البيانات،

والجدول رقم ( ١٠ ) يبين نتائج هذا التحليل:

جدول رقم ( ١٠ )

نتائج تحليل التباين الأحادي لمتوسطات معاملات الصعوبة في النماذج الثلاثة (٣ ، ٤ ، ٥) بدائل

الدلالة	(ف) المحسوبة	متوسط المربعات	درجات الحرية	مجموع المربعات	مصدر التباين
٠,٠٠٠	٥٤,٦٩٤	٥,١٩٦	٢	١٠,٣٩٢	بين المجموعات
		٠,٠٩٥	٨٧	٨,٢٤٠	داخل المجموعات
			٨٩	١٨,٦٣٣	المجموع

وبدراسة نتائج الجدول رقم (١٠) نجد قيمة (ف) المحسوبة بلغت (٥٤,٦٩٤) عند درجة

حرية (٢) وهذه قيمة دالة إحصائياً عند مستوى أقل من (٠,٠٥).

ولتحديد اتجاهات الفروق ذات الدلالة الإحصائية تم استخدام طريقة شيفيه

للمقارنات البعدية، والجدول رقم (١١) يبين نتائج المقارنات الثنائية بين متوسطات معاملات

الصعوبة للنماذج الثلاثة:

جدول رقم (١١)

نتائج المقارنات الثنائية بين متوسطات معاملات الصعوبة للنماذج الثلاثة وفق طريقة شيفيه

النموذج	٣ بدائل	٤ بدائل	٥ بدائل

النموذج	المتوسط	١٣,٤٢	١٢,٩٥	١٢,٥٩
٣ بدائل	١٣,٤٢	-	*٠,٤٧	*٠,٨٢
٤ بدائل	١٢,٩٥		-	*٠,٣٥
٥ بدائل	١٢,٥٩			-

\* تعني ذات دلالة إحصائية عند مستوى (٠,٠٥).

وبدراسة الجدول رقم (١١) نجد أن النتائج أشارت إلى مايلي:

١. وجود فروق دالة إحصائية بين النموذج ذي الثلاثة بدائل والنموذج ذي الأربعة بدائل

وكانت لصالح النموذج الذي لفقراته ثلاثة بدائل، حيث بلغت قيمتها (٠,٤٧) وهي

ذات دلالة إحصائية عند مستوى أقل من (٠,٠٥).

٢. وجود فروق دالة إحصائية بين النموذج ذي الثلاثة بدائل والنموذج ذي الخمسة

بدائل وكانت لصالح النموذج الذي لفقراته ثلاثة بدائل، حيث بلغت قيمتها

(٠,٨٢) وهي ذات دلالة إحصائية عند مستوى أقل من (٠,٠٥).

٣. وجود فروق دالة إحصائية بين النموذج ذي الأربعة بدائل والنموذج ذي الخمسة

بدائل، وكانت لصالح النموذج الذي لفقراته أربعة بدائل، حيث بلغت قيمتها

(٠,٣٥) وهي ذات دلالة إحصائية عند مستوى أقل من (٠,٠٥).

واعتماداً على النتيجة السابقة للفرض الأول يقرر الباحث رفض الفرض الصفري،

بمعنى أنه توجد فروق ذات دلالة إحصائية بين النموذج ذي الأربعة بدائل والنموذج ذي

الثلاثة بدائل لصالح النموذج ذي الثلاثة بدائل، وتوجد فروق ذات دلالة إحصائية بين

النموذج ذي الخمسة بدائل والنموذج ذي الثلاثة بدائل لصالح النموذج ذي الثلاثة بدائل،

وتوجد فروق ذات دلالة إحصائية بين النموذج ذي الخمسة بدائل والنموذج ذي الأربعة بدائل لصالح النموذج ذي الأربعة بدائل.

وقد ذكر علام (٢٠٠٧م) أن أفضل قيمة لمعامل الصعوبة هي (٠,٥٠) حيث أن هذه القيمة تجعل التمييز بين المجموعتين العليا والدنيا من الطلاب أكبر ما يمكن، ولمعرفة أي النماذج كان متوسط صعوبة فقراته أقرب إلى (٠,٥٠) تم استخدام اختبار (ت) للعينة الواحدة عند نقطة القطع (١٣,٥٠) وهي قيمة ( $\Delta$ ) المناظرة لمعامل الصعوبة (٠,٥٠)، وذلك بعد التأكد من توفر شروط استخدام اختبار (ت) على البيانات، والجدول رقم (١٢) يبين ذلك:

جدول (١٢)

متوسطات معاملات الصعوبة للنماذج الثلاثة ونتائج اختبار (ت) للعينة الواحدة عند نقطة القطع ١٣,٥٠

النموذج	متوسط معامل الصعوبة	الانحراف المعياري	قيمة ت	درجات الحرية	مستوى الدلالة
٣ بدائل	١٣,٤٢	٠,٣٤	١,٢٥ _	٢٩	٠,٢٢١
٤ بدائل	١٢,٩٥	٠,٢٨	١٠,٥٩ *	٢٩	٠,٠٠٠
٥ بدائل	١٢,٥٩	٠,٣٠	١٦,٤٢ *	٢٩	٠,٠٠٠

\* تعني ذات دلالة إحصائية عند مستوى (٠,٠٥)

وبدراسة الجدول رقم (١٢) نجد أن النتائج أشارت إلى مايلي:

بأنه لا يوجد فرق دال إحصائياً بين متوسط معاملات الصعوبة لفقرات النموذج ذي

الثلاثة بدائل وقيمة دلتا المناظرة لمعامل الصعوبة المثالي (٠,٥٠)، حيث بلغت قيمة ت (-

١,٢٥)، في حين أظهر اختبار (ت) أن متوسط صعوبة فقرات النموذجين ذو الأربعة بدائل

والخمسـة بدائل قد اختلفت بدلالة إحصائية عن القيمة المثالية لمعامل الصعوبة (٠,٥٠)، حيث بلغت قيمة ت للنموذج ذو الأربعة بدائل (- ١٠,٥)، وللنموذج ذو الخمسة بدائل (- ١٦,٤٢)، وهما ذو دلالة إحصائية عند مستوى (٠,٠٥)، وقد بلغ متوسط معاملات الصعوبة للنموذج ذو الثلاثة بدائل ١٣,٤٢، وللنموذج ذو الأربعة بدائل ١٢,٩٥، وللنموذج ذو الخمسة بدائل ١٢,٥٩.

نتيجة الفرض الثاني وتفسيرها :

ينص الفرض الثاني على أنه لا توجد فروق ذات دلالة إحصائية عند مستوى (٠,٠٥) بين متوسطات قيم معاملات التمييز في نماذج الاختبار التحصيلي لمادة الرياضيات من نوع الاختيار من متعدد ذات البدائل ثلاثة أو أربعة أو خمسة.

وللتحقق من صحة الفرض السابق قام الباحث بتحويل قيم معاملات تمييز الفقرات إلى درجات فشر الزائفة، ولمعرفة مدى الاختلاف في متوسطات معاملات التمييز بين النماذج المختلفة (نموذج الثلاثة بدائل، ونموذج الأربعة بدائل، ونموذج الخمسة بدائل) تم استخدام تحليل التباين الأحادي للكشف عن الفروق في متوسطات معاملات التمييز بين النماذج الثلاثة، وذلك بعد التأكد من توفر شروط استخدام تحليل التباين الأحادي على البيانات، والجدول رقم (١٣) بين نتائج هذا التحليل:

جدول رقم (١٣)

نتائج تحليل التباين الأحادي لمتوسطات معاملات التمييز في النماذج الثلاثة (٣، ٤، ٥) بدائل

مصدر التباين	مجموع	درجات	متوسط	(ف) المحسوبة	الدلالة
--------------	-------	-------	-------	--------------	---------

		المربعات	الحرية	المربعات	
٠,٣٣٤	١,٠٦٧	٠,٠١٦	٢	٠,٠٣٣	بين المجموعات
		٠,٠١٥	٨٧	١,٢٧٦	داخل المجموعات
			٨٩	١,٣٠٩	المجموع

وبدراسة نتائج الجدول السابق رقم (١٣) نجد أن قيمة (ف) المحسوبة بلغت (١,٠٦٧) عند درجة حرية (٢) وهذه القيمة غير دالة إحصائياً عند مستوى (٠,٠٥)، وبالتالي لا يوجد فروق ذات دلالة إحصائية بين أيّاً من المتوسطات الثلاثة تعزى لإختلاف عدد البدائل. و اعتماداً على النتيجة السابقة للفرض الثاني يقرر الباحث قبول الفرض الصفري، بمعنى أنه لا توجد فروق ذات دلالة إحصائية بين متوسطات معاملات التمييز للنماذج ذات الثلاثة والأربعة والخمسة بدائل.

ب: النتائج المتعلقة بأثر اختلاف عدد البدائل (٣، ٤، ٥) على الخصائص السيكومترية للاختبار (الثبات والصدق):

تم حساب معامل الثبات لكل نموذج من نماذج الاختبار الثلاثة وذلك باستخدام معادلة ألفا كرونباخ، كما تم حساب معامل صدق المحك التلازمي وذلك بحساب معامل ارتباط بيرسون بين درجات الطلاب في كل نموذج ودرجاتهم في اختبار منتصف الفصل الدراسي الأول لمادة الرياضيات، كما تم حساب معامل صدق المحك التنبؤي وذلك بحساب معامل ارتباط بيرسون بين درجات الطلاب في كل نموذج ودرجاتهم في اختبار

نهاية الفصل الدراسي الأول لمادة الرياضيات ، والجدول رقم (١٤) يبين معاملات الثبات ومعاملات الصدق التلازمي والتبؤي للنماذج الثلاثة للاختبار:

جدول (١٤)

معاملات الثبات وصدق المحك التلازمي وصدق المحك التبؤي للنماذج (٣ ، ٤ ، ٥) بدائل

النموذج	معامل ثبات ألفا كرونباخ	معامل الصدق التلازمي	معامل الصدق التبؤي
٣ بدائل	٠,٧١٩	٠,٥٩٥	٠,٥٠٢
٤ بدائل	٠,٧٢٩	٠,٦٣٢	٠,٤١٣
٥ بدائل	٠,٧٧٤	٠,٥٥١	٠,٥٠٢

ومن خلال الجدول رقم (١٤) السابق يمكن ملاحظة التالي:

١. ارتفاع معاملات الثبات في جميع النماذج، وكان أعلى معامل ثبات للنموذج ذي الخمسة بدائل حيث بلغ معامل ألفا كرونباخ (٠,٧٧٤) ، ويليه النموذج ذي الأربعة بدائل حيث بلغ معامل ألفا كرونباخ (٠,٧٢٩) ، وأقل معامل ثبات كان للنموذج ذي الثلاثة بدائل حيث بلغ معامل ألفا كرونباخ (٠,٧١٩) وبالتالي فإن معاملات الثبات تزيد بزيادة عدد البدائل.
٢. يتمتع النموذج ذو الأربعة بدائل بأعلى قيمة بالنسبة لصدق المحك التلازمي، حيث بلغ معامل الصدق التلازمي له (٠,٦٣٢) ، ويليه النموذج ذو الثلاثة بدائل حيث بلغ معامل الصدق التلازمي له (٠,٥٩٥) ، وكان أقل معامل صدق للنموذج ذي الخمسة بدائل حيث كان (٠,٥٥١).

٣. بالنسبة لمعامل الصدق التنبؤي في النموذجين ذي الثلاثة بدائل والخمسة بدائل

(٠,٥٠٢)، وأقل قيمة كانت للنموذج ذي الأربعة بدائل حيث بلغ معامل الصدق

التنبؤي له (٠,٤١٣).

نتيجة الفرض الثالث وتفسيرها :

ينص الفرض الثالث على أنه لا توجد فروق ذات دلالة إحصائية عند مستوى (٠,٠٥)

بين قيم معاملات ثبات الإتساق الداخلي في نماذج الاختبار التحصيلي لمادة الرياضيات من

نوع الاختيار من متعدد ذات البدائل ثلاثة أو أربعة أو خمسة.

وللتحقق من صحة الفرض السابق قام الباحث باستخدام الإحصائي (M) المقترح من

قبل هاكستين و ولين (Hakstain & Whalen, ١٩٧٦) وهي معادلة (١٩)، والذي يتبع

توزيع كاي تربيع بدرجات حرية تساوي (عدد معاملات الثبات - ١)،

( )

جدول (١٥)

معاملات الثبات بطريقة ألفا كرونباخ للنماذج المختلفة واختبار مربع كاي لفحص

الفروق بينها

.		.	.	.	.

وبدراسة الجدول رقم (١٥) نجد أن قيمة مربع كاي بلغت (١,٥٣) عند درجة حرية

(٢) وهذه القيمة غير دالة إحصائياً عند مستوى (٠,٠٥)، وبالتالي لا يوجد فروق ذات دلالة

إحصائية بين معاملات الثبات بطريقة ألفا كرونباخ تعزى لإختلاف عدد البدائل.



و اعتماداً على النتيجة السابقة للفرض الثالث يقرر الباحث قبول الفرض الصفري،  
بمعنى أنه لا توجد فروق ذات دلالة إحصائية بين معاملات ثبات الإتساق الداخلي للنماذج  
ذات الثلاثة والأربعة والخمسة بدائل.

نتيجة الفرض الرابع وتفسيرها :

ينص الفرض الرابع على أنه لا توجد فروق ذات دلالة إحصائية عند مستوى (٠,٠٥) بين  
قيم معاملات صدق المحك التلازمي في نماذج الاختبار التحصيلي لمادة الرياضيات من نوع  
الاختبار من متعدد ذات البدائل ثلاثة أو أربعة أو خمسة.

وللتحقق من صحة الفرض السابق قام الباحث بحساب معامل ارتباط بيرسون بين  
درجات الطلاب على الاختبار ودرجاتهم في اختبار منتصف الفصل الدراسي الأول لعام  
١٤٢٨ - ١٤٢٩ هـ، وللكشف عن دلالة الفروق في معاملات صدق المحك التلازمي تم  
استخدام الاختبار الإحصائي  $V$  الذي وضعه هيز (Hays, ١٩٨١) والذي يتبع توزيع كاي  
تربيع، وهي المعادلة (٢٠)، والجدول التالي رقم (١٦) يبين ذلك.

جدول (١٦)

معاملات صدق المحك التلازمي للنماذج الثلاثة واختبار مربع كاي لفحص الفروق بينها

.		.	.	.	.

وبدراسة الجدول رقم (١٦) نجد أن قيمة مربع كاي بلغت (١,٠٧) عند درجة حرية (٢) وهذه القيمة غير دالة إحصائياً عند مستوى (٠,٠٥) ، وبالتالي لا يوجد فروق ذات دلالة إحصائية بين معاملات صدق المحك التلازمي تعزى لإختلاف عدد البدائل. و اعتماداً على النتيجة السابقة للفرض الرابع يقرر الباحث قبول الفرض الصفري، بمعنى أنه لا توجد فروق ذات دلالة إحصائية بين معاملات صدق المحك التلازمي للنماذج ذات الثلاثة والأربعة والخمسة بدائل. نتيجة الفرض الخامس وتفسيرها :

ينص الفرض الخامس على أنه لا توجد فروق ذات دلالة إحصائية عند مستوى (٠,٠٥) بين قيم معاملات صدق المحك التنبؤي في نماذج الاختبار التحصيلي لمادة الرياضيات من نوع الاختيار من متعدد ذات البدائل ثلاثة أو أربعة أو خمسة.

وللتحقق من صحة الفرض السابق قام الباحث بحساب معامل ارتباط بيرسون بين درجات الطلاب على الاختبار ودرجاتهم في اختبار نهاية الفصل الدراسي الأول لعام ١٤٢٨ - ١٤٢٩هـ، وللكشف عن دلالة الفروق في معاملات صدق المحك التنبؤي تم استخدام الاختبار الإحصائي V الذي وضعه هيز (Hays) والذي يتبع توزيع كاي تربيع، وهي المعادلة (٢٠)، والجدول التالي رقم (١٧) يبين ذلك.

( )

--	--	--	--	--	--

--	--	--	--	--	--

وبدراسة الجدول رقم (١٥) نجد أن قيمة مربع كاي بلغت (٠,٩٨٤) عند درجة حرية (٢) وهذه القيمة غير دالة إحصائياً عند مستوى (٠,٠٥) ، وبالتالي لا يوجد فروق ذات دلالة إحصائية بين معاملات صدق المحك التنبؤي تعزى لإختلاف عدد البدائل. و اعتماداً على النتيجة السابقة للفرض الرابع يقرر الباحث قبول الفرض الصفري، بمعنى أنه لا توجد فروق ذات دلالة إحصائية بين معاملات صدق المحك التلازمي للنماذج ذات الثلاثة والأربعة والخمسة بدائل.

## ثانياً: مناقشة النتائج:

حاول البحث الحالي معرفة أثر عدد البدائل على الخصائص السيكومترية للاختبار التحصيلي في مادة الرياضيات، ومن أجل ذلك تم تطبيق اختبار تحصيلي في مادة الرياضيات للصف الثاني المتوسط له ثلاثة نماذج تتشابه مفرداتها في المتن، وتختلف في عدد بدائلها، وتم تطبيقها بشكل عشوائي على ٣٥١ طالباً، وتم حساب معاملات الصعوبة والتمييز والثبات والصدق لهذه النماذج، وتمت المقارنة بين النماذج الثلاثة بناءً على ترتيب فروض البحث، وفيما يلي مناقشة للنتائج:

أ: مناقشة النتائج المتعلقة بأثر اختلاف عدد البدائل (ثلاثة، وأربعة، وخمسة) على الخصائص السيكومترية لل فقرات (الصعوبة و التمييز):

## ١. صعوبة الفقرة:

أظهرت نتائج التحليل الإحصائي وجود فروق بين متوسطات معاملات الصعوبة للمفردات بين نماذج الاختبار الثلاثة ( ثلاثة، وأربعة، وخمسة) بدائل، وكان متوسط صعوبة مفردات النموذج ذي الثلاثة بدائل يمثل أعلى قيمة بين المتوسطات الثلاثة، ويأتي بعده النموذج ذي الأربعة بدائل ثم النموذج ذي الخمسة بدائل، وهذا يشير إلى أن الاختبار الذي له ثلاثة بدائل أسهل الاختبارات، ويليه النموذج ذي الأربعة بدائل، ثم النموذج ذي الخمسة بدائل، ويظهر ذلك جلياً في متوسط درجات الطلاب على الدرجة الكلية، فقد بلغ متوسط علامات الطلبة في الاختبار ذو الثلاثة بدائل ١٤,٦٢، في حين بلغ في النموذج ذو الأربعة بدائل ١٢,٩٦، وبلغ للطلبة الذين خضعوا للاختبار ذو الخمسة بدائل ١١,٦٧، وكانت هذه الفروق دالة بين النموذجين ( ٣ بدائل، ٤ بدائل) وكانت لصالح النموذج ذي الثلاثة بدائل، أيضاً كانت هذه الفروق دالة إحصائياً بين النموذجين (٣ بدائل، ٥ بدائل) وكانت لصالح النموذج ذي الثلاثة بدائل، أما الفروق بين النموذجين (٤ بدائل، ٥ بدائل) فكانت دالة إحصائياً ولصالح النموذج ذي الأربعة بدائل.

وهذه النتيجة تتفق مع النتائج التي توصلت لها الدراسات التالية:

دراسة كوستين Costin (١٩٧٠م)، ودراسة لورد Lord (١٩٧٧م)، ودراسة ستراتون وكاتس Straton & Catts (١٩٨٠م)، ودراسة كرهان وآخرون Grehan et al., (١٩٩٣، ودراسة فقوسة (١٩٩٤م)، ودراسة الشبلي (١٩٩٥م).

وهذه النتيجة اختلفت مع النتائج التي توصلت لها الدراسات التالية:

دراسة روينولدز (Roynolds ١٩٧٩م)، ودراسة اليعقوب (١٩٩٦م)، ودراسة  
ظاظا (٢٠٠٠م).

ويعتقد الباحث أن تفسير هذه النتيجة يرتبط بعدد البدائل في اختبارات الاختيار من  
متعدد، فكلما زاد عدد البدائل في مفردات اختبار الاختيار من متعدد قلت معاملات  
الصعوبة لتلك المفردات، وبالتالي يقل متوسط صعوبة الفقرات بزيادة عدد البدائل،  
ويظهر ذلك على متوسط درجات الطلاب في الاختبارات الثلاثة، فنلاحظ أن أعلى  
متوسط كان للنموذج ذي الثلاثة بدائل، وأقل متوسط كان للنموذج ذي الخمسة بدائل،  
وهذا يؤكد ما أشرنا إليه في الإطار النظري، كما نلاحظ أن أعلى نسبة للفقرات  
متوسطة الصعوبة كانت في النموذج ذي الثلاثة بدائل كما في الجدول رقم (٦)، وقد  
تبين أن أفضل النتائج كانت للنموذج ذي الثلاثة بدائل، حيث كان متوسط معامل  
صعوبة مفرداته قريباً جداً من درجة الصعوبة المثالية (٠,٥٠)، في حين كان متوسطي  
معاملات الصعوبة للنموذجين ذي الأربعة بدائل وذي الخمسة بدائل أقل من الدرجة المثالية.

## ٢. تمييز الفقرة:

أظهرت نتائج التحليل الإحصائي عدم وجود فروق ذات دلالة إحصائية بين متوسطات  
معاملات التمييز للفقرات بين نماذج الاختبار الثلاثة (ثلاثة، وأربعة، وخمسة) بدائل،  
وكانت متوسطات معاملات التمييز للنماذج الثلاثة متقاربة، حيث بلغ متوسط تمييز  
فقرات النموذجين ذي الثلاثة بدائل والأربعة بدائل (٠,٣٥)، أما النموذج ذي الخمسة بدائل  
فقد بلغ متوسط معامل التمييز له (٠,٣٩).

وهذه النتيجة تتفق مع النتائج التي توصلت لها الدراسات التالية:

دراسة كوستين Costin (١٩٧٢م)، ودراسة لورد Lord (١٩٧٧م)، ودراسة روينولدز  
Roynolds (١٩٧٩م)، ودراسة ستراتون وكاتس Straton & Catts (١٩٨٠م)،  
ودراسة كرهان وآخرون (Crehan et al., ١٩٩٣)، ودراسة الشبلي (١٩٩٥م)، ودراسة  
الموسوي (١٩٩٧م).

وهذه النتيجة اختلفت مع النتائج التي توصلت لها الدراسات التالية:

دراسة كوستين Costin (١٩٧٠م)، ودراسة فقوسة (١٩٩٤م).

ويعتقد الباحث أن هذه النتائج التي تم الحصول عليها من خلال مقارنة متوسطات  
معاملات التمييز للنماذج ذات البدائل ثلاثة وأربعة وخمسة بدائل، ترجع إلى أن متوسطات  
معاملات التمييز على تلك النماذج كانت متقاربة.

ب: مناقشة النتائج المتعلقة بأثر اختلاف عدد البدائل (ثلاثة، وأربعة، وخمسة) على  
الخصائص السيكمترية للاختبار (الثبات و الصدق):

#### ١. ثبات الاختبار:

أظهرت نتائج التحليل الإحصائي عدم وجود فروق ذات دلالة إحصائية بين معاملات  
الثبات للفقرات بين نماذج الاختبار الثلاثة (٣، ٤، ٥) بدائل، حيث كانت معاملات ثبات  
ألفا كرونباخ للنماذج الثلاثة متقاربة، فنجد معامل ثبات ألفا كرونباخ للنموذج ذي  
الثلاثة بدائل (٠,٧١٩)، أما النموذج ذي الأربعة بدائل (٠,٧٢٩)، أما النموذج ذي الخمسة

بدائل فكان يتمتع بأعلى معامل ثبات وهو (٠,٧٧٤)، وبالتالي نجد أن معامل الثبات يزيد بزيادة عدد البدائل، وهذا يؤكد ما ذكر في الإطار النظري.

وهذه النتيجة تتفق مع النتائج التي توصلت لها الدراسات التالية:

دراسة كوستين Costin (١٩٧٢م)، ودراسة ستراتون وكاتس Straton & Catts (١٩٨٠م)، ودراسة بيدسكو ونيفو Budescu & Nevo (١٩٨٥م)، ودراسة فقوسة (١٩٩٤م)، ودراسة الشبلي (١٩٩٥م)، ودراسة الموسوي (١٩٩٧م)، ودراسة ظاظا (٢٠٠٠م).

وهذه النتيجة اختلفت مع النتائج التي توصلت لها الدراسات التالية:

دراسة كوستين Costin (١٩٧٠م)، ودراسة جريير Grier (١٩٧٥).

ويفسر الباحث عدم وجود فروق ذات دلالة إحصائية بين معاملات الثبات للنماذج الثلاثة للاختبار بأن المفردات هي نفسها في الثلاثة نماذج مع اختلاف عدد البدائل، مما جعل تباين الأداء على فقرات كل نموذج من النماذج متقارباً إلى حد كبير، وزيادة معامل الثبات بزيادة عدد البدائل يتفق مع ما ذكره دالي (Dale) بأن معامل الثبات يزيد بزيادة عدد البدائل.

## ٢. الصدق التلازمي للاختبار:

أظهرت نتائج التحليل الإحصائي عدم وجود فروق ذات دلالة إحصائية بين قيم معاملات الصدق التلازمي للفقرات بين نماذج الاختبار الثلاثة ذات (الثلاثة، أو الأربعة، أو الخمسة) بدائل، حيث كانت هذه القيم للنماذج الثلاثة متقاربة، وكان أعلاها في

النموذج ذي الأربعة بدائل حيث بلغ (٠,٦٣٢)، ويليه النموذج ذي الثلاثة بدائل حيث بلغ (٠,٥٩٥)، وأقل معامل للنموذج ذي الخمسة بدائل حيث كان (٠,٥٥١).

وهذه النتيجة تتفق مع النتائج التي توصلت لها الدراسات التالية:

دراسة تريفزان وساكس وميشيل Trevisan, Sax & Michael (١٩٩١م)، ودراسة فقوسة (١٩٩٤م)، ودراسة ظاظا (٢٠٠٠م).

ويفسر الباحث عدم وجود فروق دالة إحصائية لتقارب متوسطات معاملات التمييز مما أدى إلى انخفاض التباين في الأداء على النماذج الثلاثة.

### ٣. الصدق التنبؤي للاختبار:

أظهرت نتائج التحليل الإحصائي عدم وجود فروق ذات دلالة إحصائية بين قيم معاملات الصدق التنبؤي للفقرات بين نماذج الاختبار الثلاثة ذات (٣، ٤، ٥) بدائل، وكانت أدنى قيمة في النموذج ذي الأربعة بدائل حيث بلغت ٠,٤١٣، ويليه النموذجين ذي الثلاثة بدائل وذي الخمسة بدائل حيث بلغتا ٠,٥٠٢.

وهذه النتيجة اختلفت مع نتيجة الدراسة التي توصلت لها دراسة الموسوي (١٩٩٧م).

ويلاحظ أن معاملات الصدق التنبؤي كانت أقل من معاملات الصدق التلازمي ويفسر الباحث ذلك بأن معامل الصدق التلازمي يمثل قيمة قصوى لمعامل الصدق التنبؤي، وهذا يتفق مع ما ذكره (إسماعيل، ٢٠٠٧م) أن معامل الصدق التلازمي يعتبر قيمة قصوى لمعامل الصدق التنبؤي.





# الفصل الخامس

**خلاصة البحث**

**النتائج والتوصيات**

**المقترحات**

**خلاصة البحث:**

هدف هذا البحث إلى التعرف على أثر عدد البدائل على الخصائص السيكومترية للاختبار التحصيلي للرياضيات، وتم بناء اختبار تحصيلي من نوع الاختيار من متعدد في

الوحدتين الأولى والثانية من مقرر الرياضيات لطلاب الصف الثاني المتوسط وفق جدول المواصفات، وتم تحكيم ومراجعة هذا الاختبار من قبل عدد من المتخصصين في مجال القياس والتقويم، و من معلمي ومشرفي مادة الرياضيات.

ويتكون هذا الاختبار من ثلاثة نماذج، الأول يتكون من خمسة بدائل، و الثاني يتكون من أربعة بدائل، والثالث يتكون من ثلاثة بدائل، وقد طبقت هذه النماذج الثلاثة بطريقة عشوائية منتظمة على ٣٥١ طالباً من طلاب الصف الثاني المتوسط، ينتمون إلى تسعة مدارس من مدارس التعليم العام الحكومية بمحافظة الليث التعليمية، وقد تم تطبيق الاختبار في الفصل الدراسي الأول من العام الدراسي ١٤٢٨هـ - ١٤٢٩هـ.

وبعد تطبيق النماذج الثلاثة للاختبار بشكل عشوائي على عينة البحث، تم تصحيح الاختبار عن طريق برنامج آلي ثم التعامل مع البيانات وتحليلها عن طريق برنامج الحزمة الإحصائية SPSS وبرنامج الإكسل، حيث تم استخدام الأساليب الإحصائية المناسبة، وهي:

١. حساب قيم معاملات الصعوبة والتمييز ومتوسطات هذه المعاملات لفقرات كل

نموذج من النماذج الثلاثة.

٢. اختبار تحليل التباين الأحادي للكشف عن وجود فروق دالة إحصائية بين متوسطات

قيم معاملات الصعوبة والتمييز ثم استخدام اختبار شيفيه لتحديد اتجاهات هذه الفروق.

٣. اختبار (ت) للعينة الواحدة لمتوسطات معاملات الصعوبة عند درجة القطع المثالية

لمعامل الصعوبة (٠,٥٠).

٤. حساب قيم معاملات ألفا كرونباخ وقيم معاملات صدق المحك التلازمي و التنبؤي

لنماذج الثلاثة للاختبار.

٥. الاختبار الإحصائي M المقترح من قبل هاكستين وولين (Hakstain &

Whalen, ) الذي يتبع توزيع كاي تربيع والخاص بمعرفة دلالة الفروق بين

عدة معاملات ثبات ألفا كرونباخ.

٦. الاختبار الإحصائي V الذي وضعه هيز (Hays) والذي يتبع توزيع كاي تربيع،

والخاص بمعرفة دلالة الفروق بين عدة معاملات ارتباط بيرسون أي بين معاملات

الصدق التلازمي والصدق التنبؤي.

### **ومن خلال المعالجة الإحصائية توصل الباحث إلى النتائج التالية:**

١. أظهرت النتائج فروقاً ذات دلالة إحصائية عند مستوى أقل من (٠,٠٥) بين متوسط قيم

معاملات الصعوبة للنموذجين ذي الخمسة بدائل وذي الثلاثة بدائل وكانت هذه

الفروق لصالح النموذج ذي الثلاثة بدائل، كما أظهرت النتائج فروقاً ذات دلالة

إحصائية عند مستوى أقل من (٠,٠٥) بين متوسط قيم معاملات الصعوبة للنموذجين

ذي الأربعة بدائل وذي الثلاثة بدائل وكانت هذه الفروق لصالح النموذج ذي الثلاثة

بدائل، كما أظهرت النتائج فروقاً ذات دلالة إحصائية عند مستوى أقل من (٠,٠٥)

بين متوسط قيم معاملات الصعوبة للنموذجين ذي الخمسة بدائل وذي الأربعة بدائل

وكانت هذه الفروق لصالح النموذج ذي الأربعة بدائل.

٢. لم تظهر النتائج فروقاً ذات دلالة إحصائية بين متوسطات معاملات التمييز للنماذج

الثلاثة (ثلاثة، وأربعة، وخمسة) بدائل.

٣. لم تظهر النتائج فروقاً ذات دلالة إحصائية بين قيم معاملات ثبات ألفا كرونباخ للنماذج الثلاثة (ثلاثة، وأربعة، وخمسة) بدائل.

٤. لم تظهر النتائج فروقاً ذات دلالة إحصائية بين قيم معاملات صدق المحك التلازمي للنماذج الثلاثة (ثلاثة، وأربعة، وخمسة) بدائل.

٥. لا توجد فروقاً ذات دلالة إحصائية بين قيم معاملات صدق المحك التنبؤي للنماذج الثلاثة (ثلاثة، وأربعة، وخمسة) بدائل.

### **التوصيات:**

جاء هذا البحث كمساهمة متواضعة في إثراء أدبيات البحث العربي من خلال تناول هذا الموضوع، و الذي له أثر كبير في اتخاذ القرارات الإدارية والتربوية المترتبة على نتائج الاختبارات المدرسية، وبناءً على نتائج هذا البحث فإن الباحث يوصي بما يلي:

١. استخدام النموذج ذي الثلاثة بدائل في اختبارات الاختيار من متعدد حيث فيه توفير للوقت والجهد المبذول في إعداده، وذلك لأن متوسط معامل صعوبة فقرات النموذج ذي الثلاثة بدائل كان ذا دلالة إحصائية، و لتقارب معاملات التمييز و الثبات والصدق بين بقية النماذج وعدم وجود فروق ذات دلالة إحصائية بينها.

٢. دعم مشروع الاختبارات التحصيلية بنتائج هذه الدراسة وما يماثلها، وذلك للتقليل من تكاليف إعداد الاختبارات التحصيلية والإسهام في إنجاز أكبر عدد ممكن من الاختبارات التحصيلية في جميع المواد الدراسية بمختلف المراحل.

### **المقترحات:**

١. إجراء بحث مماثل للبحث الحالي للتعرف على أثر عدد البدائل على الخصائص السيكومترية للاختبار التحصيلي للرياضيات على الطالبات.
٢. إجراء بحث مماثل للبحث الحالي للتعرف على أثر عدد البدائل على الخصائص السيكومترية للاختبار التحصيلي لمواد أخرى كاللغة العربية ومواد الدين.
٣. إجراء بحث عن أثر التخمين على الخصائص السيكومترية للاختبار وفقراته في اختبارات الاختيار من متعدد ذات الثلاثة بدائل والأربعة بدائل والخمسة بدائل.
٤. الاهتمام بتطوير مهارات الكوادر الفنية والعلمية لأعضاء مشروع الاختبارات التحصيلية، وتقديم الدعم التقني والتكنولوجي والتوسع في إقامة الدورات التدريبية التي تعمل على تطوير كفايات التقويم عند المعلمين.

## قائمة المراجع

١. ابن منظور، محمد بن مكرم (١٤١٣هـ): معجم لسان العرب، المجلد الثالث، دار إحياء التراث العربي، بيروت.
٢. أبو حطب، فؤاد (١٩٨٧م): التقويم النفسي ، ط٢، مكتبة الانجلو المصرية، القاهرة .
٣. أبو زينة، فريد كامل (١٩٩٨م): أساسيات القياس والتقويم في التربية، ط٢، مكتبة الفلاح – الكويت.
٤. أبو علام، رجاء محمود (١٩٨٧م): قياس وتقويم التحصيل الدراسي، دار القلم: الكويت.
٥. أبو لبده، سبع محمد (١٩٨٥م): مبادئ القياس النفسي، ط٣، كلية التربية، الجامعة الأردنية، عمان .
٦. أبوسل، محمد عبدالكريم (٢٠٠٢م): قياس وتقويم تعلم الطلبة ، دار الفرقان: عمان.
٧. إسماعيل، محمد المري محمد (٢٠٠٧م): محاضرات في النظرية التقليدية للقياس لطلاب الماجستير، جامعة أم القرى.

- ٨ . بامشموس، سعيد وخيري، السيد محمد ومهني، يحيى محمد (١٤٠٥هـ): التقويم التربوي، دار الفيصل الثقافية: الرياض.
- ٩ . ثورندايك، روبرت وآخرون (١٩٨٦): القياس والتقويم في علم النفس والتربية ترجمة عبد الله الكيلاني وعبد الرحمن عدس، مركز الكتب الأردني: الأردن.
- ١٠ . الحريري، رافدة عمر (٢٠٠٧م): التقويم التربوي الشامل للمؤسسة المدرسية، دار الفكر: عمان.
- ١١ . حمادين، فخري فريد (٢٠٠٥): قياس مستوى أداء المهارات التدريسية لدى الطلبة معلمي الدراسات الاجتماعية في مستوى الدبلوم العام وعلاقتها بالتحصيل العلمي لديهم ، المجلة التربوي، عدد ٧٤ المجلد ١٩ مارس ٢٠٠٥ ، جامعة الكويت .
- ١٢ . خضر، فخري رشيد(٢٠٠٤م): التقويم التربوي، دار القلم، دبي.
- ١٣ . الدوسري، إبراهيم بن مبارك(٢٠٠١م): الإطار المرجعي للتقويم التربوي ، ط٢ ، مكتب التربية العربي لدول الخليج ، الرياض .
- ١٤ . الرّازي، محمد بن أبي بكر: (١٤١٥هـ)، مختار الصحاح، مكتبة دار المعاجم: بيروت.
- ١٥ . الرّاداي، عبدالعالي سليمان(٢٠٠٦م): أثر اختلاف نمط الاختبار على الخصائص السيكمترية للاختبار وفقراته ، رسالة ماجستير غير منشورة بكلية التربية ، جامعة أم القرى .
- ١٦ . زيتون، حسن حسين وعبد الحميد، كمال(١٩٩٥م): تصنيف الأهداف التدريسية، دار المعارف، الإسكندرية.
- ١٧ . الزيود، نادر فهمي وعليان، هشام عامر(٢٠٠٢م): مبادئ القياس والتقويم في التربية، دار الفكر: عمان.
- ١٨ . سليمان ،شاهر خالد(٢٠٠١م): أثر بنية فقرات الصح والخطأ والتكميل على الخصائص السيكمترية للاختبار وفقراته ، رسالة دكتوراه غير منشورة :جامعة السودان .
- ١٩ . سماره، عزيز وآخرون(١٩٨٩م): مبادئ القياس والتقويم في التربية ، الطبعة الثانية ، دار الفكر :عمان.
- ٢٠ . سماره، نواف أحمد(٢٠٠٥م): الطرائق والأساليب ودور الوسائل التعليمية في تدريس العلوم، الطبعة الأولى، دار الفكر، عمان.
- ٢١ . سنان، إيناس محمد(٢٠٠٦م): درجة معرفة وممارسة معلمات الرياضيات في المرحلة المتوسطة لمهارات بناء الاختبارات التحصيلية بمدينة مكة المكرمة، رسالة ماجستير غير منشورة بكلية التربية ، جامعة أم القرى .



٢٢. سيد، علي أحمد وسالم، أحمد محمد (٢٠٠٥م): التقويم في المنظومة التربوية، مكتبة الرشد، الرياض.
٢٣. السيد، فؤاد البهي (١٩٥٨م): الجداول الإحصائية لعلم النفس والعلوم الإنسانية الأخرى، ط١، دار الفكر العربي، القاهرة.
٢٤. الشبلي، علي مصطفى (١٩٩٥م): أثر عدد البدائل في اختبار تحصيلي على تجانسه الداخلي والخصائص السيكومترية لفقراته، رسالة ماجستير غير منشورة، جامعة اليرموك، اربد.
٢٥. الشيخ، تاج السر وآخرون (٢٠٠٤م): القياس والتقويم التربوي، مكتبة الرشد: الرياض.
٢٦. الصمادي، عبدالله والزرايع، ماهر (٢٠٠٤م): القياس والتقويم النفسي و التربوي، دار وائل: عمان.
٢٧. ظاظا، حيدر إبراهيم (٢٠٠٠م): المقارنة بين اثار أربع طرق لحذف الموهات في فقرات اختبار الاختيار من متعدد على الخصائص السيكومترية للاختبار و فقراته، رسالة ماجستير غير منشورة، الجامعة الأردنية.
٢٨. عبدالرزاق، صلاح عبدالسميع (٢٠٠٣م): تنمية مهارات صياغة الأسئلة التحريرية، دار القاهرة: مصر.
٢٩. عبيدات، ذوقان وآخرون (١٩٩٣م): البحث العلمي مفهومه وأدواته وأساليبه، دار أسامة - الرياض.
٣٠. العريفي، هدى محمد (١٤٢٧هـ): تباين الخصائص السيكومترية لاختبار متعدد البدائل بتباين عدد مفرداته وبعض الخصائص الدافعية للمفحوصين، رسالة ماجستير غير منشورة بكلية التربية، جامعة أم القرى.
٣١. العساف، صالح بن حمد (١٤٢٤هـ): المدخل إلى البحث في العلوم السلوكية، ط٣، مكتبة العبيكان - الرياض.
٣٢. علام، صلاح الدين محمود (٢٠٠٠م): القياس و التقويم التربوي (أساسياته وتطبيقاته وتوجهاته المعاصرة)، ط١، دار الفكر العربي، القاهرة.
٣٣. علام، صلاح الدين محمود (٢٠٠٣م): تحليل البيانات في البحوث التربوية والنفسية والاجتماعية، دار الفكر العربي، القاهرة.
٣٤. علام، صلاح الدين محمود (٢٠٠٦م): الاختبارات والمقاييس التربوية والنفسية، الطبعة الأولى، دار الفكر: عمان.
٣٥. علام، صلاح الدين محمود (٢٠٠٧م): القياس والتقويم التربوي، الطبعة الأولى، دار المسيرة: عمان.

٣٦. عوده، أحمد (١٩٩٣م): القياس والتقويم في العملية التدريسية، دار الأمل: الأردن.
٣٧. غنيم، محمد أحمد (٢٠٠٣م): الاتجاهات الحديثة في بحوث مشكلات تقويم التحصيل الدراسي، كلية التربية، جامعة بنها.
٣٨. فتح الله، مندور عبدالسلام (٢٠٠٠م): التقويم التربوي، دار النشر الدولي، الرياض.
٣٩. فقوسة، محمد عبدالله (١٩٩٤م): أثر عدد البدائل في اختبار الاختيار من متعدد على الخصائص السيكومترية للاختبار وفقراته، رسالة ماجستير غير منشورة الجامعة الأردنية: عمان.
٤٠. القرني، علي بن عبد الخالق و آخرون (١٤٢٠هـ): دليل المعلم في بناء الاختبارات، الإدارة العامة للقياس والتقويم، وزارة التربية والتعليم: الرياض .
٤١. القرني، ناصر بن صالح (١٤٢١هـ): دليل المعلمين والمعلمات في تقويم الاختبارات التحصيلية، الطبعة الثانية، الإدارة العامة للاختبارات، الرياض.
٤٢. كاظم، علي مهدي (٢٠٠١م): القياس والتقويم في التعلم والتعليم، دار الكندري للنشر، عمان.
٤٣. الكبيسي، عبدالواحد حميد (٢٠٠٧م): القياس والتقويم، دار جرير: عمان.
٤٤. الكبيسي، عبدالواحد وربيح، هادي مشعان (٢٠٠٨م): الاختبارات التحصيلية المدرسية، مكتبة المجتمع العربي، الأردن.
٤٥. مراد، أحمد صلاح وسليمان، أمين علي (٢٠٠٢م): الاختبارات والمقاييس في العلوم النفسية والتربوية، دار الكتاب الحديث.
٤٦. المطيري، جزاء مرزوق (١٤٢٤هـ): الخطوات الأساسية في بناء الاختبارات التحصيلية، وزارة التربية والتعليم: تعليم الليث.
٤٧. الموسوي، نعمان (١٩٩٧م): تحديد العدد الأمثل لخيارات الفقرة في اختبار الاختيار من متعدد في ضوء بعض خصائصه السيكومترية، العدد ٤٤ من المجلة التربوية المجلد الحادي عشر جامعة الكويت.
٤٨. النبهان، موسى (٢٠٠٤): أساسيات القياس في العلوم السلوكية، دار الشروق للنشر والتوزيع: عمان.
٤٩. اليعقوب، إبراهيم محمد (١٩٩٦م): أثر اختلاف عدد البدائل والبديل لاشيء مما ذكر على الخصائص السيكومترية لفقرات الاختيار من متعدد، العدد الثاني من المجلة العربية للتربية المجلد السادس عشر، جامعة الدول العربية.
٥٠. Budescu, D. Nevo, B. (١٩٨٥). Optimal Number of Options: an Investigation of the assumption of

Proportionality. Journal of Education Measurement, vol. 22, No. 3, PP. 183-196.

51. Costin, Frank, (1970). "The optimal number of alternatives in multiple choice tests"m. some empirical evidence for mathematical proof. Educational and psychological measurement, Vol. 30, N. 2, pp. 230-308.

52. Costin, Frank, (1972). "Three – choice versus four – choice items implications for reliability and validity objective achievement ". Educational and psychological, vol. 32, N2, pp. 1030 – 1038.

53. Crehan, K., Haladyna, T. and Brewer, B. (1993). Use of an Inclusive Option and the Optimal Number of Option for multiple Choice items. Educational and Psychological Measurement, VOL. 53, PP 241-247.

54. Grier, J. Brown, (1970). "The number of alternatives for optimum test reliability" .Journal of Educational Measurement, Vol. 12, N. 2, pp. 109 – 113.

55. Hakstian, A.R. & Whalen, T.E. (1976). A K-sample significance test for independent alpha coefficients. PSYCHOMETRIKA, 41, PP 219-231.

56. Lord, Frederic, M., (1977). " optimal number of choice per item " comparison of four approaches. .Journal of Educational Measurement, Vol. 14, N. 1 pp. 33 – 38.

57. Straton, R.G. and Catts, R. M. (1980). A Comparison of two, three and four choices item tests given a fixed total number of choices. Educational and Psychological Measurement. Vol. 40, PP. 307-360.

58. Trevisan, Sax & Michae, W.B. (1991). The effects of the number of Option per item and student ability on test validity and reliability. Educational and Psychological Measurement, VOL. 51, PP 829-837.

59. Trevisan, Sax & Michae, W.B. (1994). Estimating the optimum number of options per item using an Incremental option Paradigm. Educational and Psychological Measurement, VOL. 53, PP 87-91.

60. Weber, M.B. (1977). The effect of choice Format On Internal Consistency. ED. 161940.

٦١. Willim, M , Roynolds.(١٩٧٩) "The utility of multiple - choice test formats with mildly retarded adolescents".Educational and Psychological Measurement,VOL. ٣٩, N.٢, ١٩٧٩, PP ٣٢٥-٣٣١.
٦٢. [ [home.ubalt.edu/ntsbarsh/Business stat/opre٥٠٤.htm](http://home.ubalt.edu/ntsbarsh/Business%20stat/opre٥٠٤.htm) ]٢٣-٥-٢٠٠٨.