الجمهورية الجزائرية الديمقراطية الشعبية
وزارة التعليم العالي والبحث العلمي
جامعة د. الطاهر مولاي سعيدة
كلية العلوم الاجتماعية والإنسانية
قسم العلوم الاجتماعية

مذكرة تخرج لنيل شهادة ماستر L.M.D
تخصص إرشاد وتوجيه

بعنـــوان :

الخصائص السيكومترية لاختبار مادة اللغة العربية في ضوء مؤشرات القياس الكلاسيكي ونموذج راش

(دراسة ميدانية بثانوية بوعمامة لمستوى تلاميذ الثالثة ثانوي شعبة آداب وفلسفة)

من إعداد الطالب :
هدي عبدالقادر

إشراف الدكتور :
د.شريفي علي

السنة الدراسية الجامعية : 2014 / 2015

كلمة شكر

أتقدم بالشكر الجزيل للأستاذ المشرف:**الدكتور شريفي علي** ،الذي تتبع هذا العمل وأحاطه بكل اهتمام وحثه المتواصل لحب المطالعة والتعلم ، وبالمناسبة أتمنى له حجا مبرورا وذنبا مغفورا وسعيا مشكورا وتجارة لن تبور ، كما أشكر الأساتذة أعضاء اللجنة على تفضلهم بقبول مناقشة و إثراء هذه المذكرة الأستاذة محصر و الأستاذ جغوبي ولمن كان له الفضل في المساعدة على انجاز هذه المذكرة من قريب أو من بعيد ، وكل أساتذة علم النفس بجامعة الدكتور مولاي الطاهر،كما أتقدم بالشكر الجزيل إلى دفعة التخرج 2015/2014ماستر إرشاد وتوجيه .

أهدي هذا العمل المتواضع إلى روح والدي – رحمه الله وأسكنه فسيح جنانه - ، وإلى والدتي أطال الله في عمرها ، وأدام الله صحتها وعافيتها ، وإلى كل أفراد أسرتي ، صغيرا وكبيرا .

كما أهدي هذا العمل المتواضع ، إلى الشهيد الدكتور مولاي الطاهر – رحمه الله– الذي تشرفت جامعة ولاية سعيدة باسمه ، وإلى كل العاملين بالجامعة وأخص بالذكر أساتذة العلوم الاجتماعية والإنسانية، كما أتقاسم فرحة التخرج مع زملائي مستشاري التوجيه والارشاد المدرسي والمهني لولاية سعيدة ،وباقي الزملاء في أنحاط الوطن ، فإلى هؤلاء وأولئك أهدي إليهم ثمرة جهدي.

هـــــــــدي عبدا لقادر.

ملخص البحث

هدفت الدراسة لتعرف عن التعرف على الخصائص السيكومترية لاختبار مادة اللغة العربية وفق مؤشرات القياس الكلاسيكي ونموذج راش، لمستوى تلاميذ السنة

الثالثة ثانوي شعبة آداب وفلسفة ، ولتحقيق أهداف البحث والإجابة عن تساؤلات الدراسة قام الطالب بدراسة الخصائص السيكومترية وفق مؤشرات النظرية التقليدية (الصدق الثبات ، معامل السهولة والصعوبة ومعامل التمييز للفقرات) ومعرفة مدى مطابقة بيانات الاختبار التحصيلي لافتراضات نموذج راش ،وكذا معرفة دلالة الفروق بين مرتفعي التحصيل ومنخفضي التحصيل في التحصيل الدراسي،وبين الجنسين في التحصيل الدراسي ،وقد تمثلت أداة الدراسة في اختبار تحصيلي لمادة اللغة العربية والذي تم إجراؤه على عينة متكونة من (30) تلميذ خلال الفصل الأول من السنة الدراسية 2015/2014 بثانوية بوعمامة ، ثم بعد ذلك قام الطالب بالإجراءات المنهجية للدراسة ،المعالجة الإحصائية باستخدام برنامج الرزمة الإحصائية للعلوم الاجتماعية (SPSS) ، وكانت فرضيات الدراسة على النحو الأتي:

- إن الاختبار ألتحصيلي لمادة اللغة العربية لا يتمتع بخصائص سيكومترية وفق مؤشرات القياس الكلاسيكي.
- إن مؤشرات الصعوبة لفقرات اختبار مادة اللغة العربية عالية وفق القياس الكلاسيكي
- إن بيانات اختبار مادة اللغة العربية مناسبة لافتراضات نموذج راش
- لا توجد فروق دالة إحصائيا بين مرتفعي التحصيل ومنخفضي التحصيل في التحصيل الدراسي.
- توجد فروق دالة إحصائية بين الذكور والإناث في التحصيل الدراسي عند مستوى الدلالة 0.05

وقد أسفرت نتائج الدراسة الميدانية على ما يلي:

- فيما يخص دراسة الخصائص السيكومترية وفق القياس الكلاسيكي، فيما يتعلق بالفرضية الأولى فقد بينت الدراسة الميدانية على أن الاختبار ألتحصيلي في مادة اللغة العربية لشعبة آداب وفلسفة لمستوى السنة الثالثة ثانوي غير صادق، فق تم حساب الصدق (الصدق الذاتي وصدق المحتوى) الذي يساوي (0.34)،والمعيار الدي يقبل به الصدق هو (1-0.50) وبالتالي فهو غير صادق والثبات(الجذر ألتربيعي معادلة ألفا كرونباخ)الذي يساوي 0.11 ، وبالتالي **تتحقق الفرضية الأولى.**

- فيما يخص الفرضية الثانية ، فإن أسئلة الاختبار ألتحصيلي وفقراته في مستوى التلاميذ إلا أنها لا تعبر عن مستواهم الحقيقي استنادا إلى عدم الترابط في فقرات الاختبار ألتحصيلي من خلال معامل التمييز الذي يقل عن (0.30) لأغلب فقرات أسئلة الاختبار وكذا التركيز على جانب واحد فقط يقيس المهارات العليا حسب تحليل أسئلة الاختبار ألتحصيلي ،فقد دل مؤشر الصعوبة على (0.34) كما دل معامل تمييز الفقرات على (0.11)وبالتالي **تتحقق الفرضية الثانية .**
- أما فيما يخص مطابقة بيانات اختبار مادة اللغة العربية لافتراضات نموذج راش فقد أسفرت نتائج الدراسة الميدانية على أن بيانات الاختبار التحصيلي غير مناسبة لافتراضات نموذج راش بعد تطبيق اختبار كا 2 ، وبالتالي لا **تتحقق الفرضية الثالثة .**
- وعن وجود الفروق الفردية في التحصيل الدراسي بين التلاميذ مرتفعي التحصيل ومنخفضي التحصيل ، بعد تطبيق اختبار (ت) تم التأكد أن هناك فروق فردية بين العينتين وبالتالي فروق بين مرتفعي ومنخفضي التحصيل في التحصيل الدراسي وبالتالي **تتحقق الفرضية الرابعة** .
- وفيما يخص الفرضية الأخيرة القائلة بأنه توجد فروق دالة إحصائية بين الذكور والإناث في التحصيل الدراسي عند مستوى الدلالة 0.05 ، فإن هذه الفرضية لم تتحقق

وقد اختتم الطالب بعدة توصيات أهمها إجراء دراسات مماثلة خاصة في مجال القياس النفسي باستخدام البرمجيات الحديثة لنموذج راش لفئة المربين والأساتذة ومستشاري التوجيه والإرشاد المدرسي بصفة خاصة .

فهرس المحتويات

الموضوع	الصفحة
❖ كلمة شكر	أ
❖ إهداء	ب

❖ ملخص البحث	ج – د
❖ فهرس المحتويات	ه
❖ فهرس الجداول	و
❖ فهرس الأشكال	ز
❖ فهرس الملاحق	ز
❖ مقدمة عامة	1 2
• الفصل الأول: مدخل إلى الدراسة	
➢ الإشكالية	3
➢ الفرضيات	4-3
➢ أهداف البحث	4-3
➢ أهمية الدراسة	4
➢ أسباب اختيار الموضوع	5
➢ التعاريف الإجرائية	5
➢ عرض الدراسات السابقة والتعليق عليها وموقع الدراسة منها	14-6
• الفصل الثاني: القياس الكلاسيكي (نظرية القياس التقليدية)	
▪ تمهيد	15
1- نشأة القياس الكلاسيكي.	16-15
2- مفهوم القياس الكلاسيكي .	17-16
3- أنواع القياس الكلاسيكي .	17-16
4- افتراضات القياس الكلاسيكي .	17-16
5- أوجه القصور في القياس الكلاسيكي	19-18
▪ خلاصة الفصل	19
❖ الفصل الثالث: نظرية القياس الحديثة (نظرية الاستجابة للفقرة)	
▪ تمهيد	20
1- مفهوم نظرية الاستجابة للفقرة .	21-20
2- مسلمات نظرية الاستجابة للفقرة .	21
3- افتراضات نظرية الاستجابة للفقرة .	22
4- نماذج نظرية الاستجابة للفقرة	24
5- أنواع نماذج نظرية الاستجابة للفقرة .	31-25
6- ميزات نماذج نظرية الاستجابة للفقرة .	32-31
7- عيوب نماذج نظرية الاستجابة للفقرة .	32
▪ خلاصة الفصل	32
• الفصل الرابع: الخصائص السيكومترية للاختبارات التحصيلية	
أولا :الاختبارات التحصيلية	33
▪ مفهوم الاختبارات التحصيلية	34-33
▪ أنواع الاختبارات التحصيلية	34-33
▪ خطوات الاختبارات التحصيلية	39-35
ثانيا :الخصائص السيكومترية	50-36
▪ تعريف الخصائص السيكومترية	51
▪ تصنيف الخصائص السيكومترية	52-51
▪ الخصائص السيكومترية لفقرات الاختبار في القياس الكلاسيكي	52

▪ الخصائص السيكومترية لفقرات الاختبار في إطار نموذج راش	53
خلاصة الفصل	53
• الفصل الخامس:عرض وتفسير نتائج الدراسة	
▪ الاجراءات المنهجية	54
▪ عرض ومناقشة نتائج الفرضيات	70
▪ خلاصة (تفسير نتائج الدراسة)	71
▪ توصيات و اقتراحات	71
قائمة المراجع	
الملاحق	

مقدمة عامة

إذا كان من بين أهداف العملية التعليمية اليوم هو العمل على خلق الإنسان المتعلم، المبدع، الواثق من نفسه وقدراته، فإن ذلك يتطلب العمل على خلق اتجاهات نفسية إيجابية لدى المتعلم، نحو ذاته أولاً كصورة من صور التوافق النفسي، ونحو الموضوعات الخارجية المحيطة به ثانياً، وهذه لا يمكن التعرف عليها بشكل واضح وفعال إلا من خلال عملية تقويمية ،قياسية و فعالة لما تحصل عليه المتعلم من أرصدة فكرية ،ثقافية ومعرفية لكل مرحلة من مراحل تعلميه، ونظراً لأهمية التقويم والقياس التربوي في الفعل التعليمي من أجل نجاح العملية التعليمية ،وتحقيق أهدافها المرجوة منها خصوصاً في هذا العصر فإننا سوف نتناوله بشيء من التفصيل .

لقد حاول الطالب الإجابة عن تساؤلات الدراسة الحالية من خلال تقسيمها إلى خمسة فصول:

حيث يتضمن **الفصل الأول**:أدبيات البحث بداية من مدخل الدراسة بتقديم الإشكالية و الإجابة عنها بفرضيات ثم التطرق إلى أهمية وأهداف الدراسة، أسباب اختيار الموضوع، التعاريف الإجرائية للدراسة، ثم قام الطالب بعرض الدراسات السابقة منها العربية والأجنبية والتعقيب على كل منها، وموقع الدراسة الحالية منها.

وبالنسبة للفصل الثاني بدأ الطالب بتمهيد للفصل الخاص بالقياس الكلاسيكي (نظرية القياس التقليدية) ،من حيث النشأة،المفهوم ، الافتراضات ، وعرض لأوجه القصور في القياس الكلاسيكي ثم اختتم بخلاصة الفصل.

وبالنسبة للفصل الثالث بدأ الطالب بتمهيد للفصل الخاص بالقياس الحديث (نظرية الاستجابة للفقرة) ،وتقديم النظرية –مفهومها-وذكر مسلمات وافتراضات النظرية ثم التطرق إلى نماذج النظرية والتركيز على نموذج راش، ليخلص إلى عرض مميزات وعيوب القياس الحديث، ثم اختتم الطالب بخلاصة الفصل الثالث.

أما الفصل الرابع فقد تناول الطالب فيه الخصائص السيكومترية للاختبارات التحصيلية– بدءا بتمهيد ، ثم مفهوم الاختبارات التحصيلية، طرق تقويم وقياس التحصيل ، أغراض الاختبارات التحصيلية وانواعها، خطوات بناءها-، والخصائص السيكومترية للاختبار

التحصـيلي- تعريفهـا ، تصـنيفها، الثبـات، الموضـوعية – صـياغة أوبنـاء اختبـارات تحصـيلية- إجـراءات مـا بعـد التجريـب –كيفيـة حسـاب معامـل الصـعوبة،ومعامل السهولة،ومعامل التمييز- ثم تعديل الاختبار ثم - خلاصة الفصل .

الفصل الخامس: الاجراءت المنهجية للدراسة تضمنت – منهج البحث المتبع- بدأ بتحديد مكان وزمن الدراسة- ومواصفات العينة- أدوات البحث المستخدمة- وصولا إلى المعايير الاحصائية المستخدمة.

الفصل السادس والأخير: كان عبارة عن عرض ومناقشة النتائج لفرضيات البحث الأربعة – ثم خاتمة عامة.

كما تضمن الفصل الرابع الإجراءات المنهجية للدراسة من خلال وصف أداة الدراسة والتحقق من صدقها وثباتها ثم بعد ذلك التطرق إلى منهج الدراسة وعينة الدراسة ، والمعالجة الإحصائية المستخدمة في الدراسة .

أما الفصل الخامس فقد قام الطالب بعرض نتائج الدراسة ومناقشة الفرضيات من خلال المعالجة الإحصائية لبرنامج الرزمة الإحصائية للعلوم الاجتماعية spss واختتم الدراسة بتوصيات واقتراحات .

1- إشكالية البحث:

تقديم الإشكالية:

يعد بناء الاختبارات جزءا أساسيا في العملية التعليمية وأحد المهام الرئيسية للمعلم نظرا لما تكتسبه من أهمية في تحديد مقدار ما يتحقق من الكفايات التعليمية المبرمجة والغايات التربوية المنشودة التي رسمتها الدولة والتي تنتظر من المدرسة أن تنعكس ايجابيا على الفرد المتعلم وعلى العملية التربوية ككل،كما يعد إجادة بناء الاختبارات ألتحصيلية بالطريقة الصحيحة نوع من تقويم للمعلم.

كما يغلب على الاختبارات والمقاييس المستخدمة في مجال العلوم الإنسانية في البيئة العربية اعتمادها عند بنائها وتقنينها على نظرية القياس التقليدية ، على الرغم من ظهور بعض العيوب المرتبطة بهذه النظرية ، لذا جاءت فكرة استخدام أحد النماذج الحديثة في القياس وهو نموذج راش من أجل معرفة أهم الخصائص السيكومترية التي يحققها لأحد المقاييس الذي تم بنائه في ضوء النظرية التقليدية ، كما أن كثرة استخدام هذا المقياس في البيئة الأجنبية وإخضاع بياناته لنموذج راش كان دافعاً للطالب في إجراء مثل هذه الدراسة محاولاً الإجابة على الأسئلة الآتية :

1) ما قيم مؤشرات صدق وثبات اختبار مادة اللغة العربية وفق القياس الكلاسيكي ؟
2) ما قيم مؤشرات السهولة و الصعوبة لفقرات اختبار مادة اللغة العربية وفق القياس الكلاسيكي ؟
3) ما درجة تحقيق بيانات اختبار مادة اللغة العربية لافتراضات نموذج راش ؟ .
4) ما مدى قدرة الاختبار ألتحصيلي على اكتشاف إمكانية التلميذ على الإجابة وفق افتراضات نموذج راش ؟.
5) هل توجد فروق دالة إحصائيا بين مرتفعي الذكاء ومنخفضي الذكاء في التحصيل الدراسي ؟
6) هل توجد فروق دالة إحصائيا بين الذكور والإناث في التحصيل الدراسي ؟

2 - فرضيات الدراسة:

-1 إن الاختبار ألتحصيلي لمادة اللغة العربية يتمتع بالصدق والثبات وفق القياس الكلاسيكي.
-2 إن مؤشرات السهولة و الصعوبة لفقرات اختبار مادة اللغة العربية عالية وفق القياس الكلاسيكي
-3 إن بيانات اختبار مادة اللغة العربية غير مناسبة لافتراضات نموذج راش.
-4 يمكن التنبؤ بإمكانية التلميذ على الإجابة وفق افتراضات نموذج راش.
-5 لا توجد فروق دالة إحصائيا بين مرتفعي الذكاء ومنخفضي الذكاء في التحصيل الدراسي.
-6 توجد فروق دالة إحصائية بين الذكور والإناث في التحصيل الدراسي عند مستوى الدلالة 0.05

3 -أهداف الدراسة :

❖ التعرف على الخصائص السيكومترية لاختبار مادة اللغة العربية وفق القياس الكلاسيكي ومنها

- التعرف على صدق وثبات الاختبار التحصيلي في مادة اللغة العربية.
- التعرف على مؤشرات السهولة و الصعوبة لفقرات اختبار مادة اللغة العربية.
- التعرف على مؤشرات التمييز لفقرات الاختبار التحصيلي.

❖معرفة مدى تحقيق بيانات الاختبار ألتحصيلي لمادة اللغة العربية لافتراضات نموذج راش أو مدى قدرته على اكتشاف إمكانيات التلاميذ في الإجابة على أسئلة الامتحان وفقراته .

❖معرفة مدى قدرة الاختبار ألتحصيلي على التمييز بين التلاميذ مرتفعي الذكاء ومنخفضي الذكاء في التحصيل الدراسي.

❖معرفة مدى قدرة الاختبار ألتحصيلي على الكشف عن الفروق الفردية بين الجنسين في التحصيل الدراسي.

4 -أهمية البحث:

- تفتح هذه الدراسة أمام المهتمين بالقياس النفسي والتربوي لاستخدام الطرق الحديثة لبناء الاختبارات وتقويم نتائجها.
- أنها توضح مدى مطابقة الاختبار ألتحصيلي لافتراضات نموذج راش ، وأن استخدام نموذج راش يوفر المتطلبات الموضوعية للقياس.
- مراعاة مدى تحقق الخصائص السيكومترية في بناء الاختبارات ألتحصيلية وفق القياس الكلاسيكي،
- تفتح مجال للمقارنة بين القياس الكلاسيكي والحديث في القياس النفسي والتربوي .
- مساعدة أساتذة اللغة العربية في اختيار الفقرات المناسبة لإعداد اختبار تحصيلي يعتمد على الدقة في تقدير قدرات التلاميذ بكل موضوعية.

5-أسباب اختيار الموضوع:

أولا :نظرا لظهور نظرية حديثة في القياس النفسي والتربوي -نظرية السمات الكامنة- والتي تعتبر أكثر مصداقية من النظرية الكلاسيكية في بناء الاختبار ألتحصيلي وتفسير نتائجه .

ثانيا:إلحاح الدكتور المؤطر إلى التطرق لموضوع الدراسة خصوصا تحليل النتائج المدرسية وفق الطريقة الحديثة لمواكبة التطور الحاصل في مجال القياس النفسي والتربوي وبحكم خصوصيات المهنة-مستشار التوجيه والإرشاد المدرسي والمهني- ومتطلباتها الوظيفية الرامية الى التخصص أيضا في هذا المجال.

6-التعاريف والمفاهيم الإجرائية

- **النظرية التقليدية : classical theory** : وهي واحدة من نظريات القياس التي تستخدم في تحديد العوامل التي تؤثر على الدرجة التي يحصل عليها الفرد في الاختبار . وترتكز على مفهوم الدرجة الحقيقية والدرجة الخطأ والتي يفترض انه لو أمكن أن نجري الاختبار في عدة مرات على الفرد بعناصر جديدة وتحت ظروف مختلفة ، نحصل على درجات ملحوظة مختلفة متوسطها اقرب تقدير غير متحيز لقدرة الفرد أو درجته الحقيقية(راندل 1998،Randall ،ص 4).
- **اختبار مادة اللغة العربية :** طريقة منظمة لتحديد مستوى تحصيل الطالب في المعلومات والمهارات التي تعلمها في خلال الفصل الدراسي الأول لشعبة العلوم آداب وفلسفة من خلال إجابته على عينة من الفقرات تمثل محتوى التعلم .
- **تلاميذ السنة الثالثة ثانوي** : تلاميذ متمدرسين في ثانوية بوعمامة في السنة الثالثة ثانوي تخصص آداب وفلسفة يتراوح أعمارهم ما بين 18 سنة و 20 سنة من الذكور والإناث .
- **نموذج راش :**أحد نماذج نظرية السمات الكامنة في القياس الحديث ويطلق عليه النموذج الأحادي المعلمي(one parameter model) وقد اقترحه جورج راش ويهتم بتحديد موقع الفقرة الإختبارية على سلم الصعوبة لجميع الفقرات التي تشكل الاختبار ، كما يهتم بتدريج مستويات قدرة الفرد باختبار معين على نفس متصل الفقرات (حلاصنيدلا ملاء ،2000).

7-عرض الدراسات السابقة والتعقيب عليها وموقع الدراسة منها.

1-7 - عرض الدراسات السابقة باللغة الأجنبية :

كانت مشكلة انتقاء فقرات الاختبارات استنادا إلى نموذج راش محور بحث العديد من الباحثين منهم :

الدراسة الأولى :دراسة (barbara etal ، 2006) والتي هدفت لفحص الصدق والثبات لاختبار القدرة الذاتية (see) واختبار التوقعات للمخرجات التعليمية(oee) تكونت عينة الدراسة من 166 فردا من

الإناث ، وتم تحليل النتائج باستخدام نموذج راش والتحليل العاملي ، وأظهرت نتائج التحليل أن بعض الفقرات تتميز بثبات مرتفع .(د. ماجد محمد الخياط -2012).

الدراسة الثانية : وهدفت دراسة (pallant،et. al،2007)إلى إثبات وبرهنة استخدام التحليل الإحصائي لنموذج راش عن طريق تطبيق اختبار (HADS)لقياس القلق والإحباط وطبق الاختبار على 296 مريض وتم تحليل نتائج الاختبار باستخدام نموذج راش عن طريق برنامج (RUMM2020)لمطابقة النموذج مع البيانات ، حيث أظهرت النتائج دعما واضحا لاستخدام اختبار (HADS-14)لقياس الضغوط النفسية ، وأظهرت النتائج أيضا مطابقة البيانات مع النموذج باستثناء بعض الفقرات والتي كانت غير مطابقة ، وأظهرت البيانات زوجين من الاختبارات الفرعية لاختبار الكلي هما اختبار القلق وأخر الإحباط .(د. ماجد محمد الخياط -2012) .

الدراسة الثالثة : دراسة (WILSONETAL،2007)والتي هدفت لقياس المهارات الرياضية باستخدام اختبار موضوعي (اختيار من متعدد) تكونت عينة الدراسة من 566 طالب معظم الطلبة تم إعطائهم مهارات في الجبر والحساب تعتمد على المهارات الرياضية ، تم استخدام نموذج راش عند تحليل البيانات لاعتماد الاختبار إضافة إلى وصف مستويات الطلبة في المهارات الرياضية ،وأظهرت النتائج أن درجات الطلبة على المقياس تعتمد على الأكبر سنا ومعرفة بالمهارات الرياضية ، إضافة إلى مستوياتهم السابقة بالقدرات الرياضية ، مفهوم الذات والعمر ، واظهر النموذج تطابق في البيانات المجمعة من عينة الدراسة (د. ماجد محمد الخياط -2012) .

الدراسة الرابعة : دراسة (PETER، ET، AL،2007) والتي هدفت لتقييم فقرات (ESS) الخاص بالمرضى ،البيانات من هذه الدراسة تم أخذها من دراسات أخرى سابقة أجريت على المرضى وعددهم 180 مريضا ، والبيانات المجمعة كانت جيدة ، وتتمتع بثبات عالي قدره 0.8 ،واظهر نموذج راش المستخدم في التحليل تطابق في البيانات مع النموذج ، وتم تحديد درجة في ضوئها يتم التحكم على أداء الفرد الخاضع للاختبار سواء أكانت درجة عليا آو دنيا باستخدام نموذج راش .(د. ماجد محمد الخياط -2012) .

الدراسة الخامسة: دراسة ادواردز والكوك (alcock et Edwards،2010)والتي هدفت إلى تحليل اختبار الرياضيات تكون من 11 فقرة في بريطانيا ، وتكونت عينة البحث من 164 طالب وطالبة من طلبة السنة الأولى في المرحلة الجامعية ، طلب منهم ترتيب فقرات الاختبار وفق مستوى صعوبة كل فقرة ، تم تحليل نتائج الدراسة وفق نموذج راش لتقييم مدى مطابقة الفقرات لنموذج راش أحادي المعلمة ، دلت النتائج على مطابقة فقرات الاختبار لستة أشخاص فقط من عينة الدراسة بعد إعادة ترتيب فقرات الاختبار وفق نموذج راش .(د. ماجد محمد الخياط -2012) .

الدراسة السادسة : دراسة فرانك جوني واخرون (FRANCHignoni، 2011،et al) والتي هدفت إلى تحليل فقرات اختبار داش (DASH) في ايطاليا وفق النظرية الكلاسيكية في القياس ونموذج راش أحادي المعلمة ، تكونت عينة الدراسة من 283مريضا ، أظهرت نتائج الدراسة مطابقة عشرة فقرات لنموذج راش ، ودلت النتائج على فعالية نموذج راش في مطابقة الفقرات بشكل أفضل من التحليل العملي وفق النظرية الكلاسيكية في القياس .(د. ماجد محمد الخياط -2012) .

الدراسة السابعة :دراسة هالدينا وريد haladyna 1983 ،تهدف هذه الدراسة إلى مقارنة طريقين لاختبار محكي المرجع ، الأولى بالاختيار العشوائي لفقرات الاختبار من تجمع الفقرات التي تمثل السلوك المراد قياسه ، وتعتمد الثانية الطريقة الحديثة في القياس ، وأطلق عليها الطريقة التكيفية (ADAPTIVE METHOD) والمنطق وراء الطريقة الثانية هو اختيار الفقرات بمستويات صعوبة تتناسب مع مستوى تحصيل الطالب وكانت هذه الطريقة بثلاث مستويات للصعوبة : المستوى الأول: اختيار الفقرات ذات الصعوبة ملائمة لمستوى قدرة المفحوص (ONE LEVEL)المستوى الثاني:اختيار فقرات ذات صعوبة قريبة من مستوى قدرة المفحوص (NEAR-LEVEL) المستوى الثالث : اختيار فقرات ذات صعوبة بعيدة عن مستوى قدرة المفحوص (OFF-LEVE) وقد أوضحت نتائج الدراسة انه عند إعطاء المفحوصين فقرات بصعوبة اعلي من مستوى قدراتهم فإنهم يحصلون على علامة اقل من علامة المجال ، وعند إعطاء المفحوصين فقرات بصعوبة اقل من مستوى قدرتهم يحصلون على علامة اكبر من علامة المجال ، وفي كلتا الحالتين يزداد خطا القياس . (علي زكريا- 1320 ه) .

الدراسة الثامنة : دراسة وارفل (WARFEL،1984) تهدف هذه الدراسة إلى مقارنة بعض نماذج نظرية الاستجابة للفقرة والطرق الكلاسيكية في دقة ثبات تقييم قدرة الطلاب في بعض المجالات الدراسية وقد بلغت عينة الدراسة 4900 طالبا ، وقد تم استخدام ثلاثة نماذج للسمات الكامنة وهي : النموذج اللوغاريتمي ذو المتغير الواحد (راش) ،النموذج اللوغاريتمي ذو المتغيرين (لورد) ،النموذج اللوغاريتمي ذو الثلاث متغيرات (بيرنبوم) .وقد أوضحت نتائج الدراسة أن نموذج راش هو أفضل هذه النماذج ، ويرجع ذلك إلى سهولة تطبيقه واستخدامه بالإضافة إلى كفاية الدقة الإحصائية التي يتميز بها النموذج لتقدير متغير قدرة الفرد ، وكذلك توافر برامج الكمبيوتر الخاصة بهذا النموذج .(علي زكريا- 1320 ه)

الدراسة التاسعة: دراسة هابلتون وجونز (HAMBLETON & JONES 1993، تهدف الدراسة إلى مقارنة غير تجريبية بين النظرية الكلاسيكية والحديثة في القياس ، وقد بيت الدراسة جوانب القصور في النظرية الكلاسيكية وكيف عالجت النظرية الحديثة هذه الجوانب ، كما بينت الدراسة خصائص

إحصائيات كلتا النظريتين وأشارت إلى التفوق الواضح الذي تظهره النظرية الحديثة في مجال حل المشكلات التي تعجز عن حلها النظرية الكلاسيكية مثل القياس التكيفي وبنوك الأسئلة بمفهومها الجديد وتقديرات القدرة المستقلة عن عينة الفقرات للمفحوصين . (علي زكريا-1320 ه)

الدراسة العاشرة: دراسة ألان (ALAN،1994)هدفت الدراسة إلى التعرف على العلاقة بين دالة المعلومات وفقا لنظرية الاستجابة للفقرة ومعامل الثبات وفقا لنظرية الاختبارات التقليدية ، وقد أوضحت نتائج الدراسة ما يلي :نظرية الاستجابة للفقرة ذات كفاءة عالية في بناء الاختبارات وعدم وجود فرق في دوال المعلومات بين نماذج الاستجابة للفقرة،ويرجع ذلك للأسباب منها ،انخفاض مستويات القدرة لمجموعة الأفراد الذين أدو الاختبار،وجود عدد كبير من الفقرات ذات الصعوبة المنخفضة ،انخفاض التمييز وصعوبة الحصول على معلم التخمين بصورة دقيقة بسبب وجود عدد كبير من الأفراد منخفضي القدرة (علي زكريا-1320).

تعقيب على الدراسات الأجنبية : من خلال الدراسات السابقة يتضح أن:

معظم نتائج الدراسات أوضحت أن نموذج راش هو أفضل هذه النماذج، ويرجع ذلك إلى سهولة تطبيقه واستخدامه بالإضافة إلى كفاية الدقة الإحصائية التي يتميز بها النموذج لتقدير متغير قدرة الفرد، وكذلك توافر برامج الكمبيوتر الخاصة بهذا النموذج. ودراستنا تتشابه مع الدراسات السابقة في عدة جوانب منها المنهج أي الوصفي ، الأدوات وهي الاختبارات التحصيلية . أهداف البحث وتختلف في بعض النقاط منها : مجتمع البحث ، عينة الدراسة ، المادة الدراسية التي يجرى عليها البحث، المستوى الدراسي مما يعطي خاصية لبحثنا هذا .

7-2-الدراسات السابقة باللغة العربية

الدراسة الأولى :دراسة(سيف 2008) التي هدفت إلى التعرف إلى مستوى معرفة المشرفيين التربويين بمبادئ بناء الاختبار بمحافظة تعز ، ولتحقيق أهداف الدراسة تم تطبيق اختبار تحصيلي مكون من 85 فقرة من نوع الاختيار من متعدد على عينة مكونة من 296 مشرفا ومشرفة ،وبينت نتائج الدراسة أن هناك تدني في معرفة المشرفيين التربويين بمبادئ بناء الاختبارات التحصيلية ، وعدم وجود فروق دالة إحصائيا في مستوى المعرفة بمبادئ الاختبارات التحصيلية تعزى الخبرة أو التخصص أو مكان العمل .(حمدي يونس ابو جراد-2011)

الدراسة الثانية : دراسة دعنا (2005) ،هدفت لبناء اختبار مكيف هرمي في الرياضيات للصف الثامن الأساسي وفقا لنموذج راش ، وتم بناء اختبارين من نوع متعدد وبأربعة بدائل لكل فقرة تقيس

فقراتها المستويات الخمسة الأولى من تصنيف بلوم في المجال المعرفي الأول يتعلق بمحتوى الفصل الأول والثاني بمحتوى الفصل الثاني وطبق الاختباران على عينة الدراسة المتكونة من 376 طالب تم اختيارهم بطريقة عشوائية وحللت الفقرات باستخدام برنامج rascal وتم إجراء تكافؤ لصعوبة الفقرات لتكون متدرجة تدريجا موحدا مشتركا 'بلغ عدد فقراته (62) اختير منها (55) فقرة استخدمت في بناء الاختبار المكيف، الذي تألف من (10) مستويات اختبارية مرتبة ترتيبا هرميا وفق صعوبته وبعدد فقرات مساوي لرقم المستوى الاختباري .(د. ماجد محمد الخياط -2012

الدراسة الثالثة : دراسة سنان (2006) ،هدفت إلى معرفة درجة ممارسة معلمات الرياضيات في المرحلة المتوسطة لمهارات بناء الاختبارات التحصيلية،واختلافها باختلاف المؤهل العلمي وسنوات الخبرة والدورات التدريبية في مجال القياس ، ولتحقيق أهداف الدراسة تم استخدام الاختبار الذي أعده جرادات (1988) المكون من 90 مفردة من نوع الاختيار من تعدد لكل مفردة 4 بدائل واستبانة من إعداد جرادات (1988) لقياس درجة الممارسة مكون من 25 مفردة ، وتكونت عينة الدراسة من 93 معلمة وبينت نتائج الدراسة ان هناك تدني ملحوظ في معرفة وممارسة معلمات الرياضيات بالمرحلة المتوسطة بمدينة مكة المكرمة لقواعد بناء الاختبارات التحصيلية .(حمدي يونس ابو جراد-2011)

الدراسة الرابعة : دراسة علوان (2006) هدفت إلى تحليل الأسئلة المتضمنة في كتابي المطالعة والأدب والنقد للصف الحادي عشر في المنهاج الجديد حسب تصنيف بلوم للأهداف المعرفية ، وتحديد أنماط التفكير التي تتطلبها هذه الأسئلة (تقاربي – تباعدي) وقد أظهرت نتائج الدراسة أن أسئلة الكتاب تتناول المستويات المعرفية بشكل غير متواز ، حيث ركزت على بعض المستويات مثل الفهم والتذكر والتحليل بينما قل التركيز بالنسبة لبعض المستويات (التطبيق ،التركيب ،والتقويم) إما بالنسبة لأنماط التفكير فكان فكان التفكير التقاربي هو الأكثر انتشارا ، اذ وصلت نسبته 72.5 % بينما نسبة أسئلة التفكير التباعدي كانت حوالي 27.5 % ، أما تنوع الأسئلة بين المقاربة ، والموضوعية فقد أشارت النتائج إلى أن النوع السائد هو الأسئلة المقاربة .(حمدي يونس ابو جراد-2011

الدراسة الخامسة: دراسة الثبيتي (2005) هدفت الى معرفة مدى توفر وممارسة الكفايات التعليمية في إعداد الاختبار ، وتصحيحه وتحليل نتائجه لدى معلمي الرياضيات بالمرحلة الثانوية بمدينة الطائف التعليمية ولتحقيق أهداف الدراسة تم تطبيق اختبار تحصيلي يقيس مدى توفر الكفايات التعليمية في إعداد وتصحيح وتحليل نتائج الاختبار والاستبانة لقياس مدى ممارسة الكفايات التعليمية في إعداد وتصحيح وتحليل نتائج الاختبارات على عينة مكونة من 129 معلما تم اختيارهم بطريقة الطبقية العشوائية ، وتوصلت نتائج الدراسة إلى أن هناك انخفاضا في توافر وممارسات المعلمين في إعداد الاختبار ،

وتصحيحه وتحليل نتائجه ،ووجود علاقة دالة وموجبة بين مدى توافر وممارسة الكفايات التعليمية في إعداد الاختبار وتصحيحه وتحليل نتائجه .(حمدي يونس ابو جراد-2011

الدراسة السادسة : دراسة الكحلوت (2004) ،هدفت الى تقويم اختبارات اللغة العربية للصف الثالث الأساسي من خلال معرفة مدى توافر معايير الاختبار الجيد في هذه الاختبارات ، حيث أجرت الباحثة تقويما شاملا للاختبارات المدرسية تناولت شكل الاختبار ، مضمونه بالإضافة الى تقويم كل نوع من أنواع الأسئلة الموضوعية والمقالية ، وتوصلت نتائج الدراسة إلى أن هناك اهتماما من قبل المعلمين بمعايير الشكل أكثر من اهتمامهم بمعايير المضمون ، وخلو جميع الاختبارات من مفاتيح التصحيح ، وعدم توافر عدد لا باس به من المعايير الخاصة بأنماط الأسئلة ، إضافة إلى خلو الاختبارات من أسئلة الصواب والخطأ وأسئلة المقابلة .(حمدي يونس ابو جراد-2011)

الدراسة السابعة: دراسة الزهراني (2002) إلى معرفة واقع امتلاك معلمي الرياضيات بالمرحل الابتدائية للكفايات المعرفية لأعداد الاختبارات التحصيلية ولتحقيق أهداف الدراسة تم بناء استبانة ضمت 107 فقرة موزعة على ثماني مجالات يمثل كل مجال منها إحدى كنفايات إعداد الاختبارات التحصيلية،وقد تم تطبيق الاستبانة على معلمي الرياضيات بالمدينة المنورة وتوصلت الدراسة إلى عدة نتائج أبرزها أن درجة امتلاك معلمي الرياضيات بالمرحلة الابتدائية للكفايات المعرفية لأعداد الاختبارات التحصيلية بشكل عام كانت متدنية حيث بلغت 52.1 %.(حمدي يونس ابو جراد-2011) .

الدراسة الثامنة:دراسة الأحمد 1992، تهدف هذه الدراسة إلى دراسة اثر عدد اختيار فقرات الاختبار المحكي المرجع على خصائصه السيكومترية و وقد عينة الدراسة 244 طالبا وطالبة تم اختيارهم عشوائيا من مجتمع الدراسة المؤلف من طلبة الصف الرابع الأساسي للعام الدراسي 1991- 1992 في مديرية التربية والتعليم لعمان الكبرى الأولى ،وقد تم بناء اختبار مؤلف من 85 فقرة تمثل مجال السلوك المراد قياسه ، وهو العمليات الحسابية الأربع للأعداد الصحيحة ، وقد تم التحقق من صدق الاختبار من خلال عمليات تحليل المحتوى والأهداف ، وكذلك إيجاد معامل الصدق المرتبط بالمحك بإيجاد معامل الارتباط بين العلامات على الاختبار وعلامات الطلبة في التحصيل الدراسي في الرياضيات لعينة الدراسة ، حيث بلغت قيمته 0.84 ، كما تم التحقق من ثباته باستخدام معامل الاتساق الداخلي كرونباخ الف حيث بلغت قيمته 0.96 وقد تم تشكيل ستة اختبارات بواقع 30 فقرة لكل اختبار ، كما تم اختيار نقطتي قطع 60 % لتصنيف المفحوصين عند تلك النقطتين وقد كان ابرز نتائج الدراسة ،أن خطا القياس لاختبار الذي اختيرت فقراته بالطريقة العشوائية كان الاقل منه للطرق الأخرى عند مستوى الدلالة 0.05 ،في حين كان خطا القياس اكبر ما يمكن لطريقتي نموذج راش ومعامل فاي عند نقطة قطع 80 %

فيما يتعلق بفاعلية ثبات الاتساق الداخلي (KR20) ودلالات صدق المحك لاختبارات الستة ، فقد جاءت قيمتها متقاربة وبدرجة يصعب الفصل بينها وكانت قيم (KR20) عالية وكذلك معاملات الارتباط بين كل من الاختبارات الستة واختبار المجال . (علي زكريا-1430 ه)

الدراسة التاسعة :دراسة عبد الحافظ 1999 : تهدف هذه الدراسة إلى تقويم بناء الاختبارات المرجعة إلى المحك / المعيار في ضوء نظرية الاستجابة للفقرة والنظرية التقليدية،تم جمع البيانات باستخدام أسلوب المحاكاة بالكومبيوتر ، حيث استخدم احد برامج الحاسب الآلي في توليد فقرات اختبارية ، كون منها عدد 200 اختبار ، ثم استخدم برنامجا أخر في توليد البيانات المترتبة على محاكاة التطبيق الحقيقي لهذه الاختبارات ، ومن ثم فلم يعتمد على عينة محددة لتطبيق تلك الأدوات عليها . وقد توصلت هذه الدراسة إلى النتائج التالية :

- أفضلية نماذج الاستجابة للفقرة عن الطريقة التقليدية في بناء الاختبارات بنوعيها
- أدت نماذج الاستجابة للفقرة إلى دقة اكبر في تقدير القدرة ، وكذلك اقل مستوى من الخطأ المعياري في تقديرها .
- أدت نماذج الاستجابة للفقرة إلى درجة أعلى من الثبات ودقة القياس
- النموذج الثلاثي البارامتري كان أفضل النماذج في دقة بناء وتصميم الاختبارات بنوعيها وكذلك في دقة القياس . (علي زكريا-1320 ه)

الدراسة العاشرة :دراسة جمحاوي 2000 :تهدف هذه الدراسة إلى المقارنة بين النظرية الكلاسيكية والنظرية الحديثة في القياس في اختبار القدرة الرياضية ،وقد بلغ عدد أفراد عينة الدراسة 1061 طالبا وطالبة من طلبة الصف التاسع في مديرتي اربد الأولى والثانية ، وقد تم تحليل البيانات باستخدام برنامجي ASCAL و RESCAL لنماذج النظرية الحديثة ، وبرنامجي ITEMAN و spss للتحليل وفق النظرية الكلاسيكية ، وذلك للحصول على الإحصائيات الخاصة بالفقرات والأفراد . وقد أشارت نتائج الدراسة إلى وجود اتفاق عالي بين النظريتين في تقدير صعوبة وتمييز الفقرات ، كما ظهر اتفاق عالي بين النظريتين في توزيع القدرة للمفحوصين ،

كما بينت النتائج أن عدد الفقرات التي تم اختيارها وفق النظرية الكلاسيكية بلغ 33 فقرة ، في حين طابقت 20 فقرة نموذج راش ، بينما طابقت 35 فقرة نموذج المعلمتين ، وطابقت 38 فقرة نموذج الثلاثة معالم ، مما يؤكد أن عدد الفقرات المطابقة للنموذج في النظرية الحديثة يزداد بزيادة عدد معالم النموذج ، كما أشارت النتائج إلى عدم وجود فروق ذات دلالة إحصائية بين معاملي المقياسين الناتجين عن المؤشرات الكلاسيكية والنموذج راش والنموذج ثلاثي المعلمة . (علي زكريا-1320 ه)

7-2-1- تعقيب على الدراسات العربية:

من خلال استعراض الدراسات السابقة يتضح ما يلي :

دقة تقدير نماذج نظرية الاستجابة للفقرة . حيث أدت نماذج الاستجابة للفقرة إلى دقة اكبر في تقدير القدرة ، وكذلك اقل مستوى من الخطأ المعياري في تقديرها .

- المؤشرات الإحصائية للفقرات الإختبارية المقدرة بنموذج راش أكثر استقرارا من تلك المقدرة باستخدام النظرية الكلاسيكية .
- أدت نماذج الاستجابة للفقرة إلى درجة أعلى من الثبات ودقة القياس .
- خلو جميع الاختبارات من مفاتيح التصحيح ، وعدم توافر عدد لا باس به من المعايير الخاصة بأنماط الأسئلة ، إضافة إلى خلو الاختبارات من أسئلة الصواب والخطأ وأسئلة المقابلة .
- هناك تدني في معرفة المشرفيين والمدرسين بمبادئ بناء الاختبارات التحصيلية .
- فعالية نموذج راش في مطابقة الفقرات بشكل أفضل من التحليل العملي وفق النظرية الكلاسيكية في القياس.
- نظرية الاستجابة للفقرة ذات كفاءة عالية في بناء الاختبارات

7-3-3- موقع الدراسة الحالية من الدراسات السابقة :

إن أوجه الاستفادة من الدراسات السابقة ، هو أنها بمثابة المصدر الأساسي لكثير من المعلومات المهمة والتي توجهنا في إعداد المذكرة الحالية، وتحديد أسئلة الدراسة و إعداد الإجراءات الملائمة لتحقيق أو عدم تحقيق فرضيات الدراسة، استنادا إلى تلك الدراسات السابقة ، وهذا بالإضافة إلى أن تلك الدراسات وجهت الطالب الباحث نحو العديد من البحوث والدراسات والمراجع المناسبة ومكنته من تكوين تصور شامل عن الأطر النظرية التي ينبغي أن تشملها الدراسة الحالية، وقد خلص الطالب إلى ما يلي :

- اختلاف أدوات الدراسة المستخدمة في تلك الدراسات، والأساليب الإحصائية المستعملة ، أدى هذا إلى استخدام أساليب إحصائية مناسبة لغرض البحث (اختبار ت ،كا2، المتوسط الحسابي ...) .
- تباينت العينات في الدراسات السابقة من حيث المرحلة الدراسية والعمرية ، فبعضها كانت من المرحلة الابتدائية.، ومن هنا كانت عينة البحث قصدية – تلاميذ التعليم الثانوي-
- اختلاف الدراسات السابقة في البرامج الحاسوبية المستخدمة في تحليل فقرات ،،مما يستدعي ضرورة استخدام spss للمعالجة الإحصائية للبيانات.

- أجمعت معظم الدراسات السابقة على أن نموذج راش يستند إلى افتراضات قوية تستند بدورها إلى أساس إحصائي سليم ، مما يثبت فعالية النموذج لتدريج فقرات الاختبارات والمقاييس النفسية والتربوية ، وعليه يمكن التحقق من الاختبار التحصيلي في مادة اللغة العربية ومدى مطابقته لافتراضات نموذج راش .

- عدم توفر دراسات محلية – حسب علم الطالب الباحث - توضح الخصائص السيكومترية لاختبار تحصيلي وفق القياس الكلاسيكي ونموذج راش، وعليه يمكن التحقق من ذلك في الدراسة الحالية.

القياس الكلاسيكي(نظرية القياس التقليدية):

تمهيد:

إن أغلب المقاييس العقلية المعاصرة قد تم إعدادها وتقنينها وفقًا لنظرية القياس الكلاسيكية ، في حين يبذل المعنيون جهودًا حثيثة لإعادة تدريجها ، وفقًا لنظرية الاستجابة للفقرة نظرًا لما قدمته من إسهامات في تحقيق دقة وموضوعية القياس.ولذا سيتطرق الطالب الباحث لكل من النظريتين بشيء من التفصيل على النحو التالي:

نشأة القياس الكلاسيكي:

نشأت نظرية القياس الكلاسيكية بصورة محددة من جهود سبيرمان 1927(Spearman) في تقديم الأساس الرياضي لنظريته في الذكاء (نظرية العاملين) ، والكثير من المصطلحات المعروفة ،مثل الدرجة الحقيقية "True Score" ،الصدق والثبات.

وفي حين كانت النظرية الإحصائية السائدة في ذلك الوقت تعتمد على المؤشرات الإحصائية للعالم كارل بيرسون ، (1955 C.Pearson,)فقد بنيت النظرية الكلاسيكية بشكل مكثف على مفاهيم الارتباط، وقد أدى التفاعل والاندماج بين مناهج القياس والنظريات الرياضية الإحصائية من جهة ونظريات علم النفس من جهة أخرى نتيجة الاهتمام باختبارات الذكاء إلى ظهور نظريات القياس ،ومنها نظرية القياس الكلاسيكية ، حيث كانت بداياتها في عام 1930 حين ظهرت مجموعة من الطرق الإحصائية المتناثرة في عدة مجالات ، فكان الهدف الاستفادة من هذه الطرق ووضعها في قالب علمي واسع الانتشار ، فجاءت مجلة من أهم المجلات في نظرية القياس في عام 1935 تسمى " Psychometrical" وتتعرض هذه المجلة لجميع المعادلات الإحصائية في مجال القياس ، وفي بريطانيا ظهرت مجلة أخرى تسمى "Education and Psychology"، القياس النفسي والتربوي ثم مجلة ثالثة عن الإحصاء النفسي

وقد تبلورت نظرية القياس الكلاسيكية على يد جاليكسون (Gulliksen) سنة1950، وذلك في كتابه : "نظرية الاختبارات العقلية (Theory Of Mental Tests)الذي يعتبر من أمهات الكتب بل أولها في مجال التنظير للقياس ، فقد حوى عرضًا شاملا وتفصيلا كاملا بشكل رياضي لنظرية القياس الكلاسيكية.

ثم قام كل من لورد ونوفيك (Lord & Novick,)1968 بتطوير النظرية واستخدامها على نطاق واسع في التربية وعلم النفس ، حيث حاولا القيام بإعادة تشكيل السمات الأساسية للنظرية مستخدمين الطرق الإحصائية الرياضية الحديثة فأعتبر كتابهما" : النظريات الإحصائية لدرجات الاختبار النفسي Statistical Theories" Of Mental Test Scoresمن أهم الكتب في مجال تطوير نظرية القياس الكلاسيكية وهو آخر تجديد للنظرية القياس ثم توالت كتب لثورندايك Thorndike) سنة 1982 .وأهمها كتاب القياس التربوي النظرية والتطبيق Educational":" Measurement: Theory and Practice ولبساطة المبادئ التي ترتكز عليها هذه النظرية وسهولة فهمها، فقد حظيت بانتشار واسع لدى غالبية المعلمين والعاملين في الميدان رغم ظهور نظريات أحدث منها (عسيري، 2005) .

مفهوم القياس الكلاسيكي:

تعرف نظرية القياس الكلاسيكية بنظرية الدرجة الحقيقية"True score theory" وهي من أولى النظريات في القياس ، وبالرغم من التطور الهائل في النظريات في مجال القياس مثل : نظرية إمكانية التعميم "Generalizability theory، ونظرية الاستجابة للفقرة "Item response theory" ، " في العقدين الماضيين، إلا أن الطرق الكلاسيكية في نظرية القياس الكلاسيكي ذات أثر قوي حتى الآن ، وهناك العديد من الاختبارات التي توجد الآن تقدم دليلا على خصائص البيانات على أساس الطرق الكلاسيكية في القياس.

والهدف الأساسي لنظرية القياس الكلاسيكية هي تقدير الثبات من خلال الدرجات الملاحظة والحقيقية ، وتسمى أحيانًا هذه النظرية بنظرية الدرجة الحقيقية ؛لأنها تعتمد في اشتقاقاتها النظرية على النموذج الرياضي نموذج الدرجة الحقيقية (1990SUEN page 27)

الفكرة الأساسية للقياس الكلاسيكي:

تقوم فكرة نظرية القياس الكلاسيكية على مفهوم الدرجة الحقيقية ودرجة الخطأ والذي يفترض أنه لو أمكن أن نجري الاختبار عدة مرات على الفرد بعناصر جديدة وتحت ظروف مختلفة ، فإننا نحصل على درجات ملحوظة مختلفة متوسطها هو أقرب تقدير غير متحيز لقدرة الفرد أو درجته الحقيقية.

ولوجود أخطاء في أي نوع من القياس فإن الدرجة الملحوظة تعكس كذلك تأثير عدد من العوامل مثل : الوقت ونوع الأسئلة والظروف النفسية وغيرها، مما يجعلها مختلفة عن الدرجة الحقيقية ، إضافة إلى أننا لا نستطيع تطبيق الاختبار عددًا من المرات على الفرد.

وبمعنى آخر يوجد فرق بين الدرجة الملحوظة والدرجة الحقيقية بسبب أخطاء القياس ،ويمكن توضيح ذلك رياضيًا على النحو التالي:

$X = T + E$ ،حيث أن: الدرجة الملاحظة هي(x) والدرجة الحقيقية. (T) ودرجة الخطأ (E) وهذا يعني أن الدرجة الملاحظة (X) لكل طالب تتكون من المجموع الجبري للدرجة الحقيقية(T) التي تعكس مستوى الطالب الحقيقي، ودرجة الخطأ (E) التي مصدرها ما يسمى بالخطأ العشوائي.

لذلك يجب أن يبنى الاختبار بطريقة تمكننا من قياس الدرجة الحقيقية لكل طالب ، بمعنى آخر يجب أن نأخذ كل الاحتياطات التي تقلل من تأثير الخطأ العشوائي، بحيث تكون نسبته في الدرجة الملاحظة أقل ما يمكن (الدوسري، 2000 ص 83) .

أنواع القياس الكلاسيكي: تتضمن أساليب القياس الكلاسيكي نوعين من القياس هما:

القياس جماعي المرجع: . Norm-referenced measurement

نشأ هذا النظام مرتبطًا بالفلسفة التربوية التي سادت في أوائل قرن العشرين ، وهي تصنيف الأفراد بحسب مركزهم النسبي بين أقرانهم في القدرات المختلفة ، وكان علماء النفس أكثر من علماء التربية فاعلية في نشأة هذا النوع من القياس فقد كان بينيه "Binet"أول من صمم اختبارًا لعزل الأطفال المتخلفين عقليًا عن العاديين (الشرقاوي وآخرون83،1996)، وقد أدى ظهور ونجاح أول اختبار جمعي لتصنيف الضباط والجنود قبيل الحرب العالمية الأولى إلى تشجيع حركة القياس جماعي المرجع) .(علام ، 2001 ،21)

وهكذا تقوم المقاييس جماعية المرجع – وهي أكثر المقاييس شيوعًا وانتشارًا – على تقدير الفروق الفردية بين الأفراد والتمييز بينهم ، وفي هذه المقاييس ، لا تكون لدرجة الفرد معنى ما لم تقارن بمعيار يعتمد على مستوى جماعة الأقران التي ينتمي إليها هذا الفرد ، ويتمثل هذا المستوى بمتوسط درجات هذه الجماعة ، بينما تتمثل المقارنة بمدى انحراف درجة الفرد عن هذا المتوسط وبوساطة المعايير المحسوبة لدرجات المجموعة التي ينتمي إليها هذا الفرد (كاظم 18،1988) .

القياس محكي المرجع: Criterion- referenced measurement

مع شيوع مفهوم التعلم من أجل الإتقان، "Mastry Learning" لم يعد الهدف هو التركيز أساسًا على الفروق الفردية ، وظهرت المناداة بالابتعاد عن شكل التوزيع ألاعتدالي – الذي تعتمد عليه المقاييس جماعية المرجع – حيث إن النشاط التربوي نشاط مقصود يبذل بهدف أن يتقن الطلبة ما تعلموه ، وإذا ما تحقق الهدف

من العملية التعليمية ، فإن توزيع الأداء يختلف تمامًا عن المنحنى ألاعتدالي ، لذلك ظهر أسلوب جديد في القياس ، عرف بالقياس محكي المرجع (المرجع السابق، 18،1988) .

ويعتبر جليسر (Glaser 1963) أول من وجه الاهتمام إلى هذا الأسلوب في القياس ، ,وكان أول من أطلق عليها مصطلح الاختبارات محكية المرجع ، ثم جاء بعده كارفرCarver سنة1974 ليكون أول من استخدم مصطلح ديومتري (محكي المرجع) في مقابل سيكومتري (جماعي المرجع)) .(منصور 1987، 22) ، والهدف من القياس محكي المرجع هو تحديد المستوى الذي يصل إليه أداء الفرد وتقدير مدى تقدمه بالنسبة للأهداف الموضوعة للقياس ، بغض النظر عن مكانته بين أقرانه **(عبد السلام 1992، 34)**

- إفتراضات القياس الكلاسيكي:

تقوم نظرية القياس الكلاسيكية على مجموعة من الافتراضات ،من أهمها ما أورده جاليكسون (Gulliksen1950)في كتابه" : نظرية الاختبارات العقلية (Theory Of Mental Tests") المشار إليه فيCrocker& Algina 1986 (جمحاوي 2000،)وهي:

التوزيع الاعتدالي للدرجات على متصل القدرة التي يقيسها الاختبار ، فعدد الأفراد الواقع في مستوى معين من القدرة ، يناظر العدد المتوقع من دالة الكثافة الاحتمالية.

أن الدرجة الملاحظة_ (X) هي حاصل جمع الدرجة الحقيقية(T) زائدًا خطأ القياس(E) حيث إن (E)تأخذ قيمًا سالبة وموجبة ، $X=T+(\pm E)$ ، وأن تباين الدرجات الخام يساوي تباين الدرجات الحقيقية مضافًا إليه تباين درجات الخطأ . كما أن الارتباط بين مجموعتين من الدرجات الخام ، المستمدتين من اختبارين متوازيين ، أو ثبات الاختبارات في مجتمع معين ، يساوي تباين الدرجات الحقيقية إلى تباين الدرجات الخام ، وأن الزيادة أو النقص في طول الاختبار ، يؤثر في ثبات درجاته.

أنه لا يمكن معرفة أو قياس الدرجة الحقيقية ، بل يمكن الاستدلال عليها أو تقديرها من خلال الدرجة الملاحظة ، وذلك بحساب متوسط الدرجة الملاحظة المستقلة ، الناتجة من تطبيق الاختبار لعدد كبير من المرات ، وتزداد دقة الدرجة الحقيقية بزيادة مرات التطبيق. (Crocker& Algina1996)

أوجه القصور في القياس الكلاسيكي:

على الرغم من استخدام الباحثين للنظرية التقليدية في بناء الاختبارات وتحليل و تفسير نتائجها، إلا أن هناك بعض المشكلات التي تقلل من دقة و موضوعية هذا الاستخدام، ومن أهم جوانب القصور في هذه النظرية ما يلي :

عدم وجود وحدة قياس ثابتة : حيث لا تحدد مواضع القياس على متصل المتغير بصورة خطية، فاعتماد درجات الأفراد على مفردات الاختبار قد يؤدي إلى اختلاف المسافة بين كل درجتين متتاليتين، و يؤدي هذا إلى اختلاف المعنى الكمي لأي فرق محدد عبر مدى درجات الاختبار (أمينة كاظم، 1996 : 295).

تأثر خصائص مفردات الاختبار بقدرة الأفراد : حيث يختلف معاملات الصعوبة أو السهولة و التمييز لمفردات الاختبار باختلاف قدرة أفراد العينة، فالمفردة التي يختبر بها أفراد ذوي قدرات مرتفعة تبدو سهلة، بينما تبدو نفس المفردة صعبة لذوي القدرات المنخفضة ، و إذا كانت العينة متجانسة نسبيًا، فإنّ قيم معاملات التمييز تكون أقل من القيم التي نحصل عليها من عينة غير متجانسة (Hambelton&Swaminathan، 1989 : 4).

تأثر الدرجة الكلية للفرد في اختبار ما بمفرداته : حيث تكون درجة الفرد عندما يختبر بمفردات سهلة أعلى من المفردات الصعبة، فلا يمكن تقدير قدرته فيما تقيسه هذه المفردات تقديرًا دقيقًا، لذا تختلف نتيجة القياس باختلاف الاختبار المستخدم (صلاح الدين علام، 2000 : 203).

تقتصر الموازنة بين الأفراد في السمة ،أو القدرة التي يقيسها الاختبار على تطبيق نفس مفردات الاختبار أو مجموعة مفردات مكافئة أو موازية لها على كل فرد من الأفراد (عماد عبد السميح، 1991 : 446).

تأثر ثبات الاختبار بالموقف الاختباري : حيث يعتمد ثبات الاختبار في إطار هذه النظرية إما على تطبيق الصورة الاختبارية مرتين على أفراد العينة، أو على إعداد صور متكافئة من الاختبار و يعد هذا في الواقع أمرًا صعبًا، وبالرغم من أهمية ذلك، إلا أنه غير كاف، حيث يمكن أن يختلف الموقف الاختباري و ظروف التطبيق في هاتين المرتين، الأمر الذي يؤثر على دقة ثبات الاختبار (Hambelton&Swaminathan، 1989 : 5).

تساوي تباين أخطاء القياس لجميع أفراد العينة موضع الاختبار، و هذا بالرغم من أنه قد يكون أداء بعض الأفراد على الاختبار أكثر اتساقًا من غيرهم من الأفراد، و أن درجة هذا الاتساق تختلف باختلاف مستوى قدرة الأفراد أو بمستوى القدرة التي يقيسها الاختبار (Randall ، 1998 : 6).

لا تقدم هذه النظرية تفسيرًا سيكولوجيًا ،يوضح كيف يحاول الفرد إجابة إحدى مفردات الاختبار، على الرغم من أن هذا التفسير يعد ضروريًا و لازمًا إذا أردنا التنبؤ بخصائص الدرجات المستمدة من مجتمع معين أو مجتمعات مختلفة من الأفراد، أو إذا أردنا تصميم اختبارات تتميز بخصائص سيكومترية معينة تناسب مجتمعًا من الأفراد هذا بالإضافة إلى أن تكوين مفردات الاختبار و معناها تتغير بتغير عامل الزمن، أي بمضي الزمن بالنسبة لعينة الأفراد الذين أعد لهم الاختبار، فالظروف البيئية تتغير، والظروف الاختبارية ليست دائمًا مقننة كما أن حذف أو تغيير أي مفردة من مفردات الاختبار يؤدي إلى تغيير في درجات الأفراد، هذا التغيير يصعب التنبؤ به (صلاح الدين علام، 1985 : 102).

جميع خصائص الاختبارات التي تستند في بنائها على أسس النظرية التقليدية، مثل معاملات الصعوبة و التمييز و الثبات، تعتمد على خصائص عينة الأفراد التي يجري عليها الاختبار، و على خصائص عينة المفردات التي يشتمل علها الاختيار (صلاح الدين علام، 1987 : 18).

خاتمة الفصل:

في ضوء ما سبق، و حيث إن النظرية التقليدية للقياس تؤدي إلى بناء اختبارات غير مرنة، فقد وجه المتخصصون في القياس جهودهم لوجود نظام قياس أكثر موضوعية يركز على انتقاء المفردات الاختبارية بشكل أفضل، و يسمح بإضافة أو حذف مفردات إلى الاختبار دون أن يتأثر الاختبار ككل، و قد أدت الجهود إلى ظهور ما يطلق عليه نماذج السمات الكامنة.

تمهيد :

يعتمد الاتجاه المعاصر في القياس النفسي و التربوي على نظرية السمات الكامنة التي تفترض أنه يمكن التنبؤ بأداء الأفراد أو يمكن تفسير أدائهم في اختبار نفسي أو تربوي معين بناء على خاصية أو خصائص معينة مميزة لهذا الأداء تسمى السمات Traits، و تحاول هذه النظرية تقدير درجات الأفراد في هذه السمات، و نظرًا لصعوبة ملاحظة هذه السمات بطريقة مباشرة أو قياسها بصورة مباشرة، فإنه يجب تقديرها أو الاستدلال عليها من خلال استجابات الأفراد الملاحظة على مفردات الاختبار الذي يقيس السمة أو القدرة المطلوبة و لهذا أطلق عليها "نظرية السمات الكامنة" أو نظرية الاستجابة للمفردة (عماد عبد المسيح، 1991 : 447).

مفهوم نظرية الاستجابة للفقرة:

تعد نظرية الاستجابة للفقرة طريقة مشهورة لنمذجة البيانات، بمعنى أنها تحاول نمذجة العلاقة بين المتغير غير الملاحظ (قدرة الفرد) واحتمال استجابة الممتحن صوابا على فقرة ما (المتغير الملاحظ)، ويعتمد صدق طرق نظرية الاستجابة للفقرة في جزء كبير على البيانات المستمدة من أداء الفرد، و تقوم نظرية الاستجابة للفقرة على جمع متغيرين في علاقة رياضية وتمثل نموذج احتمالي، نظرا لأن طبيعة البيانات تسمح بالتنبؤ باحتمال النجاح على أي فقرة من الفقرات المدرجة، كما توفر تقديرات لقدرات الأفراد بوحدات النموذج المستخدم (إسماعيـــل،2007:13).

وعندما يود عالم النفس قياس إحدى السمات العقلية، ولتكن الذكاء مثلا، فإنه عندما يلاحظ طفل معين لا يرى ذكاؤه، لأن الذكاء مفهوم مجرد، ولكنه يلاحظ سلوك الطفل في مواقف مختلفة تتطلب الذكاء، أي: إن السمات العقلية تعد بمثابة تكوينات يفترض أنها تشكل مجموعة مترابطة من السلوك الذي يمكن ملاحظته وتسجيله في مواقف إختبارية مقننة (علام، 2000: 21-22).

وإن الفكرة الأساسية لنماذج نظرية الاستجابة للفقرة تتلخص في ربطها خصائص الفقرات بمعلم مقدار السمة التي يمتلكها الفرد واحتمال إجابته عن فقرة محددة بمستوى معين من الإجابة حسب نوع الفقرة، ويقوم هذا الرابط على تحديد موقع الفرد على مقياس السمة بصرف النظر عن مجموعة الفقرات التي يجيب عنها الفرد (عودة، 1992).

مسلمات نظرية الاستجابة للفقرة:

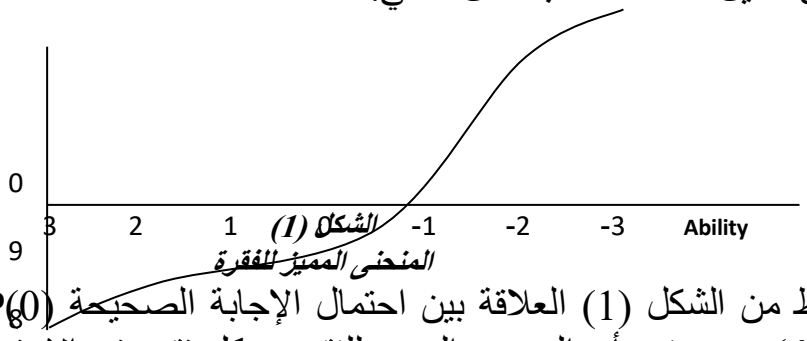
تقوم نظرية الاستجابة للفقرة على مسلمة أساسية هي أن القيمة الاحتمالية لاستجابة فرد لفقرة إختبارية ما تكون دالة لكل من السمة أو القدرة التي يفترض أن الاختبار يقيسها لدى الفرد وخصائص الفقرة التي يحاول الإجابة عنها، أي أنها تفترض أن هناك دالة احتمالية تربط بين بارامترين أحدهما يتعلق بالفرد والآخر يتعلق بالفقرة التي يختبر بها، وتهدف النظرية للتوصل إلى قيم تقديرية لكل من هذين البارامترين (Linden & Hambleton ,1997:5).

ويذكر الوليلي (2002) أن نظرية الاستجابة للفقرة تقوم على بعض المسلمات التي يجب أن يدركها كل مستخدم لهذه النظرية أو لأحد النماذج المتعلقة بها، ومن أهم هذه المسلمات مايلي:

-يمكن التنبؤ بأداء أي مختبر في أي اختبار بواسطة مجموعة من العوامل يطلق عليها سمات أو قدرات كامنة.

-يمكن وصف العلاقة بين أداء المختبر على أي فقرة إختبارية ومجموعة السمات أو القدرات الكامنة التي يفترض أنها تؤثر في أدائه على هذه الفقرة بدالة تزايدية (تزايد طردي)، ويطلق عليها دالة خصائص الفقرة، لأن هذه الدالة تحدد المختبرين الذين حققوا درجات مرتفعة في السمات التي لها توقعات احتمالية عالية للإجابة الصحيحة للفقرة من المختبرين الذين حققوا درجات منخفضة على السمات (الوليلي، 2002: 47-48).

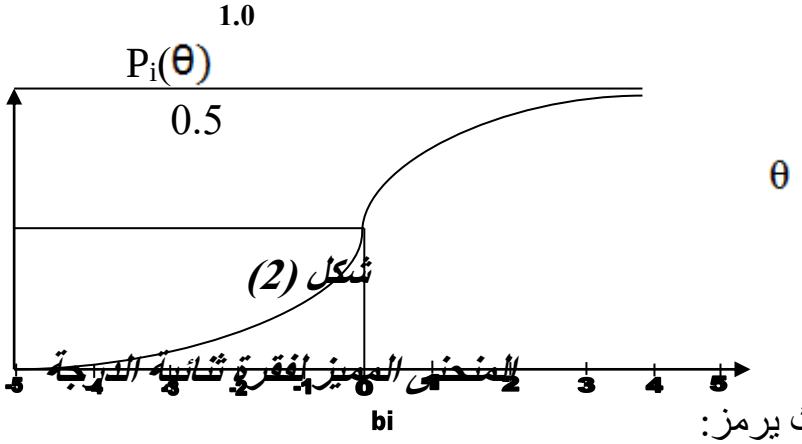
[2]ويمكن تمثيل هذه العلاقة بالشكل التالي:



الشكل (1) المنحنى المميز للفقرة

ويلاحظ من الشكل (1) العلاقة بين احتمال الإجابة الصحيحة P(0) لفقرة ما ومقياس القدرة (Ability)، ويعرف بأنه المنحنى المميز للفقرة، وكل فقرة في الاختبار يكون لها منحنى مميز خاص بها. واحتمال الإجابة الصحيحة تقترب من الصفر عند مستويات منخفضة من القدرة . ويزيد احتمال الإجابة الصحيحة عند المستويات العالية للقدرة حتى تصل احتمال الإجابة الصحيحة إلى واحد صحيح (Baker ,2001 :7).

ويأخذ المنحنى المميز للفقرة في حالة الفقرة ثنائية الدرجة الشكل التالي:



شكل (2) المنحنى المميز لفقرة ثنائية الدرجة

حيث يرمز:

θ: قدرة الفرد.

b_i: صعوبة الفقرة.

$P_i(\theta)$: احتمالية أن يجيب المفحوص الذي قدرته θ على الفقرة i بطريقة صحيحة.

ويعد مفهوم المنحنى المميز للفقرة من المفاهيم المهمة في نماذج الاستجابة للفقرة التي تتناول سمة كامنة متصلة يتم قياسها بفقرات إختبارية من النوع الثنائي أي: تتطلب إجابة صحيحة أو إجابة خطأ. ويمثل هذا المنحنى بدالة رياضية تربط بين احتمال نجاح الفرد في إجابة الفقرة، وبين السمة أو القدرة التي تقيسها مجموعة فقرات أو يقيسها اختبار معين. أي أنها دالة انحدار درجة الفقرة على السمة الكامنة التي يقيسها الاختبار، وهذه الدالة غير خطية "Non-Linear" (علام، 2000: 690).

افتراضات نظرية الاستجابة للفقرة:

تتميز نظرية الاستجابة للفقرة بعدة افتراضات التي ينبغي أن تتوفر في البيانات المتعلقة بالاختبارات، حيث يتم اختيار النموذج الملائم للبيانات وفقا لمدى تحقق هذه الافتراضات في البيانات، حيث إن ذلك يؤدي إلى دقة وإحكام القياس وما يستتبعه من أحكام وقرارات. وفيما يلي توضيح للافتراضات الأساسية التي ترتكز إليها نماذج نظرية الاستجابة للفقرة:

1/أحادية البعد Unidimensionality :

يرى ويبرغ (2004) Wiberg أن أحادية البعد تشير إلى الافتراض بأن الاختبار يقيس قدرة كامنة واحدة، وهي تعني وفق ما يراه أورلاندو (.Orlando (w.d أن التغاير بين الفقرات يمكن أن يفسر بواسطة بعد واحد يقف خلف ذلك التغاير.

والخلاصة كما يقول ويبرغ (2004) Wiberg بأنه وفقا للنظرية الحديثة والتي تفترض أحادية البعد، بمعنى أن يقيس الاختبار قدرة كامنة واحدة. وهذا يتفق مع أن المفحوص ذو القدرة العالية يفترض أن تكون احتمالية إجابته على الفقرة بشكل صحيح عالية أيضا، والعكس صحيح (Lord, 1980)، (كاظم، 1988 ب).

ويثير دودين (2004) نقطة تطبيقية في غاية الأهمية، وهي أن العديد من الاختبارات عادة ما تحتوي على عدة مجالات، والسؤال كما يثيره دودين: كيف ندعي عندئذ في مثل هذا الوضع أن الاختبار أحادي البعد في حين أنه يقيس عدة مجالات؟

ويجيب دودين نفسه قائلا: "والحقيقة أنه لا تناقض في ذلك إذا ما أخذ في الاعتبار أن هناك عاملا مشتركا أساسيا بين هذه المجالات، وهي قياسها للسمة المراد قياسها".

ثم يضيف دودين ما اقتبسه من سميث (1996) Smith حيث قال: " من المهم التفريق بين أحادية البعد النظرية والعملية. فمن الناحية النظرية فإن مسائل الحساب تتكون من أربعة مكونات أو أبعاد مثلا هي: الجمع والطرح والضرب والقسمة. والتحليل العاملي لاختبار الحساب يمكن أن يظهر هذه البناءات أو الأبعاد الأربعة. ولكننا من الناحية العملية نفترض أن هذه البناءات الأربعة تقيس فعليا موضوعا واحدا وهو القدرة الحسابية. وفي الواقع فإن من الصعب إيجاد إختبار حساب لا يجمع بينهم" (دودين، 2004: 115).

2/الاستقلال المحلي Local Independence:

يرى الدوسري (2001) أن خاصية الاستقلال المحلي تعني أن الإجابة على فقرة ما مستقلة إحصائيا عن الإجابة عن أي فقرة أخرى في الاختيار (الدوسري، 2001: 5).

والمفهوم الأدق كما عرفه هامبلتون وآخرون (1991) .Hambleton et al بأنه "عند تثبيت القدرة المسؤولة عن أداء المفحوصين على الاختيار، فإن استجابات المفحوصين على أي زوج من الفقرات تكون مستقلة إحصائيا" (Hambleton et al. 1991 :10).

ويرى أولاندو (.Orlando(w.d بأن هذا الافتراض، فنيا يعتبر، امتدادا للافتراض الأعم وهو أحادية البعد، وهو يعني أنه إذا تم إزالة أثر العامل أو البعد أو السمة الكامنة خلف المقياس، فلن يوجد أي تغاير منتظم إضافي بين الفقرات. وبذلك فإن الاستقلال المحلي يمكن أن يظهر بين المجموعات الجزئية من الفقرات التي تسير في اتساق واحد (.Orlando(w.d .

3/عامل السرعة في الإجابة Speeded ness :

هناك افتراض أساسي عام لجميع نماذج نظرية الاستجابة للفقرة المستخدمة، وهو أن الاختبار الذي يسعى النموذج لمطابقة بياناته لم يتم تطبيقه تحت ظرف السرعة. بمعنى أن المفحوصين الذين أخفقوا في الإجابة على فقرات الاختبار لم يكن ذلك بسبب إخفاقهم في السرعة الكافية لانجاز الاختبار، وإنما يعود ذلك إلى محدودية قدراتهم (Hambleton & Swaminathan, 1985)، (علام، 1986).

ويشير هامبلتون وسواميناثان Hambleton and Swaminathan (1985) إلى أنه من النادر الإشارة إلى هذا الافتراض، لأنه قد يكون متضمنا في افتراض أحادية البعد، فعندما تؤثر السرعة في الأداء على الاختبار، فإن هذا يعني أن هناك سمتين على الأقل تؤثران في الأداء هما: سرعة الأداء، والسمة الأخرى التي يسعى الاختبار لقياسها.

الافتراضات السابقة التي تمت مناقشتها هي افتراضات عامة لجميع نماذج نظرية الاستجابة للفقرة.

4/تساوي مؤشرات التمييز Equal Discrimination Indices:

يرى الدوسري والمرشد (2004) أنه يرتبط بمشكلات التحقق من فرضيات النظرية مسألة اختيار النموذج المناسب لتحليل البيانات الإختبارية، بمعنى هل يتم استخدام النموذج ذي الثلاثة معالم أو المعلمتين أو المعلمة الواحدة؟ (الدوسري والمرشد، 2004).

الخطوة الأولى لتحديد أنواع المعالم التي يمكن تضمينها في النموذج تعتمد على التحليل وفق النظرية الكلاسيكية. فعلى سبيل المثال: معلمة التمييز (الميل Slope) في نظرية الاستجابة للفقرة والتي يرمز لها بالرمز (a_i) تناظر ارتباط الفقرة بالدرجة الكلية في النظرية الكلاسيكية (Reeve,w.d.).

ويقصد بتساوي مؤشرات التمييز، أي أن تكون جميع المنحنيات المميزة لفقرات الاختبار متوازية (متساوية الميل)، وتختلف عن بعضها البعض في نقط التقائها بالمحور الأفقي الذي يمثل متصل القدرة أو السمة الكامنة.

5/الحد الأدنى من أثر التخمين Minimal Guessing:

أما معلمة التخمين (c) فيمكن فحص مدى مناسبة تضمينها في النموذج المقترح، وبذلك يمكن استخدام النموذج 3PLM، حيث إن النموذج ذا المعالم الثلاثة 3PLM له معلمة خاصة بالتخمين، وهي عبارة عن المقدار الذي يقترب فيه المنحنى من نقطة الصفر.

وهذا يعني أن معلمة التخمين يجب أن تكون جزءا من النموذج، خاصة إذا كانت الفقرات هي الاختيار من متعدد فإن إمكانية التخمين تزداد بشكل عام (Wiberg, 2004).

نماذج نظرية الاستجابة للفقرة:

لقد أدت الجهود المتواصلة للمهتمين بنماذج الاستجابة للفقرة إلى التواصل إلى مجموعة من النماذج السيكومترية الجديدة تعرف بنماذج الاستجابة للفقرة Item Response Models ، وتتخلص الفكرة الأساسية لهذه النماذج في أنها تحاول اشتقاق قيم تقديرية للسمة أو السمات التي تتضمنها الاستجابة لمجموعة من فقرات الاختبار، بمعنى أن هذه النماذج تقدم تفسيرا لاستجابة

الفرد لفقرة الاختبار التي تنطوي على السمة التي تقيسها هذه الفقرة، والسمة المقاسة قد تكون قدرة معينة أو خاصية من خصائص الفرد، بحيث تكون هناك علاقة منتظمة بين مستويات السمة المقاسة لدى أفراد مختلفين وبين احتمالات الاستجابة الصحيحة لفقرات الاختبار (علام، 2000: 686).

إن الهدف الأساسي من أي نموذج ينتج من أي نظرية إختبارية بصفة عامة هو استخدامه (أي النموذج) في التدريج Scaling، أي إعطاء قيم للأفراد على التدريج Scale values، ولن يتحقق ذلك إلا إذا حدد النموذج العلاقة بين أداء الفرد، والقدرة التي تكمن خلف هذا الأداء، وعلى هذا فإن نموذج السمة الكامنة يحدد العلاقة بين البيانات (أداء الفرد على الاختبار) وقيم التدريج أو الدرجات على تدريج السمة الكامنة، وكلما كان هناك ملائمة (Fit) بين النموذج المستخدم ومجموعة البيانات أدى ذلك إلى الحصول على تقديرات دقيقة للسمة المقاسة، حيث يجب وضع كل من الأفراد والفقرات على مقياس تدريج السمة، ويتم ذلك من خلال عمليات التقدير طالما توجد علاقة تقارب ممكنة بين الاحتمالات المتوقعة للممتحنين، والاحتمالات الواقعية (الفعلية) لأدائهم في كل مستوى من مستويات السمة المقاسة، وتتميز هذه النماذج بأنها دوال احتمالية Probabilistic وليست دوال حتمية Deterministic وهذا يعني أن العلاقة بين أداء الفرد والسمة التي يتم قياسها تسير وفقا لنظرية الاحتمالات (علام، 1982: 40).

إن استخدام الاحتمالات أقرب إلى طبيعة العلم والظواهر السلوكية بصفة خاصة ولا يتنافى ذلك مع دقة القياس، فتلك النماذج الرياضية هي وسائل قياس دقيقة يمكن الاعتماد عليها بدرجة عالية من الثقة في المجال النفسي والعقلي والتربوي، وتختلف النماذج فيما بينها في الشكل الرياضي للدالة المميزة للفقرة، كما تختلف في عدد البارامترات الواجب تحديدها في كل نموذج للوصول للعلاقة السابقة ، فكل نموذج يحتوي على واحدا أو أكثر من البارامترات التي تصف الفقرة وواحد أو أكثر من البارامترات التي تصف الفرد (علام، 2000: 698).

أنواع نماذج نظرية الاستجابة للفقرة:

توجد نماذج متعددة في إطار نظرية الاستجابة للفقرة، وذلك لاختلاف الافتراضات المتعلقة بالبيانات الإختبارية. إذ ربما يفترض أن الأداء في الاختبار ينطوي على سمة أحادية البعد، أو على سمات متعددة. أو ربما يفترض أن الفقرات تتباين في قدرتها على التمييز بين المستويات المختلفة للقدرة ، أو أن الإجابات الصحيحة على فقرات الاختيار من متعدد تتأثر بالتخمين العشوائي أو لا تتأثر كما في الأسئلة المفتوحة، أو عما إذا كانت الدرجات ثنائية أم متعددة الأقسام. لذلك يمكن تقسيم هذه النماذج إلى:

-أ- نماذج سكونية Static models:

ويهتم هذا النوع من النماذج بالقياس في مدة زمنية واحدة، وكذلك بتحديد العمليات التي ينطوي عليها الأداء في الاختبارات السيكولوجية والتربوية. وتختلف هذه النماذج في صيغها الرياضية التي تربط أداء الفرد في الاختبار بدرجات قدرته، كما تختلف في عدد الخصائص أو البارامترات المستخدمة في وصف كل فقرة من فقرات الاختبار (علام، 2005: 67).

ومن أهم النماذج السكونية ثلاثة نماذج أساسية شائعة الاستخدام وهي:

1/النموذج اللوغاريتمي أحادي المعلم One-Parameter Logistic Model:

ويسمى بنموذج "راش" Rasch Model نسبة إلى عالم الرياضيات "جورج راش" بجامعة كوبنهاجن (1960)، وهو من أبسط نماذج نظرية الاستجابة للفقرة ويفترض أن استجابة كل فرد تكون إما (1) أو (صفر) على كل فقرة من فقرات الاختبار (المرجع السابق، 2005: 69).

كما يفترض النموذج تساوي معاملات التمييز (a_i)، وانعدام التخمين (c_i)، ويقوم بتقدير صعوبة الفقرات (b_i) فقط، وتتمثل الصيغة الرياضية لهذا النموذج فيما يلي:

$$P_i(\theta) = \frac{e^{D(\theta - b_i)}}{1 + e^{D(\theta - bi)}}$$

حيث:i: هي رقم الفقرة وتساوي 1، 2، .....، n

$P_i(\theta)$: احتمالية أن يجيب المفحوص الذي قدرته θ على الفقرة i بطريقة صحيحة.

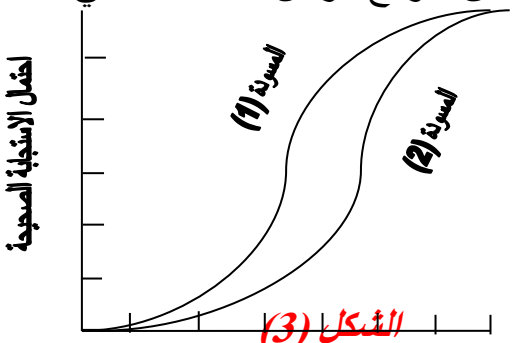
b_i: هو صعوبة الفقرة.

e: الأساس اللوغاريتمي ويساوي (2.718) تقريبا.

D: معامل القياس أو التدريج Scaling Factor وهو مقدار ثابت يساوي 1.7، ويحول هذا المعامل المنحنى اللوغاريتمي إلى تطابق تقريبي مع المنحنى الاعتدالي.

(Fan, 1998 : 359).

ويمكن تمثيل الدالة التي تعبر عن نموذج "راش" بالشكل التالي:



الشكل (3)

التمثيل البياني للنموذج اللوغاريتمي أحادي المعلم (نموذج راش)

ويلاحظ من الشكل (5) أن معلم التخمين يمثل بالخط الذي يلتقي بالطرف الأسفل للمنحنى المميز لكل فقرة في اللانهاية، وهو يسمى بالخط التقاربي. ولذلك تتقاطع المنحنيات المميزة لفقرات الاختبار، أي تختلف عن بعضها البعض في كل من الجزء الذي تلتقي فيه بالمحور الأفقي الذي يمثل متصل السمة الكامنة وهو بارامتر الصعوبة، والميل يمثل بارامتر التمييز، والخط التقاربي الأدنى هو الخط الذي يلتقي بالطرف الأسفل للمنحنى في اللانهاية ، ويمثل بارامتر التخمين (علام، 2005: 73-74).

ـب ـنماذج ديناميكية Dynamic Models:

وتهتم هذه النماذج بقياس التغير الحادث في السمات الإنسانية عبر الزمن. فالبعض يرى أن هذا التغير يعد عملية تدريجية، بينما يرى البعض الآخر أنه عملية تحول من حالة إلى أخرى، وفي الحالة الأولى يكون المتغير الذي ينطوي عليه التغير متصلا، بينما في الحالة الثانية يكون

المتغير منفصلا، غير أن الحالة الأولى هي التي نالت الاهتمام في النماذج الديناميكية (المرجع السابق، 2005: 84).

معايير اختيار النموذج المناسب للبيانات:

من أجل فحص وتقويم النماذج المختلفة وصولا للنموذج الأمثل لمطابقة البيانات، اقترح كل من هامبلتون وسواميناثان (Hambleton & Swaminathan)، وهامبلتون وآخرون (Hambleton et al., 1991) ثلاث معايير لتقييم نماذج نظرية الاستجابة للفقرة هي:

أولا- التحقق من افتراضات النموذج.

ثانيا- التحقق من الخصائص المتوقعة للنموذج.

ثالثا- التحقق من قدرة النموذج على التنبؤ بالدرجات الفعلية.

أولا: التحقق من افتراضات نموذج الاستجابة للفقرة:

1-أحادية البعد Unidimensionality:

هناك أساليب عديدة يمكن استخدامها لفحص افتراض أحادية البعد لبيانات الاختبار. أحد هذه الأساليب هو فحص معامل الثبات فمعامل الثبات ألفا كرونباخ هو مؤشر للاتساق الداخلي الذي يشير إلى أحادية البعد (الحربي، 2006: Wiberg, 2004). ومع ذلك فهناك أساليب أكثر تحديدا، تساعد في فحص تحقق افتراض أحادية البعد في البيانات المشاهدة من أهمها التحليل العاملي.

ويتم فحص خاصية أحادية البعد في التحليل العاملي من خلال تفحص قيم الجذور الكامنة Eigen values. فإذا أعطى التحليل العاملي عاملا واحدا مميزا والعديد من العوامل الصغيرة، بحيث كانت قيمة الجذر الكامن لذلك العامل تشكل نسبة واضحة ومرتفعة من التباين الكلي للدرجات، فإن ذلك يعني أن هناك عاملا واحدا مهيمن في الاختبار، مما يشير إلى تحقق افتراض أحادية البعد.

ويشير دودين (2004) نقلا عن ريكاس Reckase أن التحليل العاملي الذي يفسر العامل الأول فيه 20% من التباين الكلي للدرجات على الأقل يعتبر أحادي البعد (دودين، 2004).

لكن هامبلتون "Hambleton" أثار ملاحظة تطبيقية في غاية الأهمية ، وهي أن الحصول على عامل يفسر نسبة أقل من 10% من التباين الكلي للدرجات أمر ليس مستغربا، ولذا فإنه طالما كان ذلك العامل أكبر بشكل واضح ومميز من بقية العوامل، فإن ذلك يعني صحة افتراض أحادية البعد لذلك الاختبار (Wiberg, 2004).

كما يمكن استكشاف ما سبق من خلال فحص الرسوم Scree Plot والتي هي عبارة عن تمثيل لقيم الجذور الكامنة لجميع العوامل الممكن استخلاصها من الاختبار. فإذا اتضح من فحص تلك الرسوم أن التغير في قيمتي الجذور الكامنة بين العامل الأول والعامل الذي يليه لافت للنظر وسريع، مقارنة بالتغير التدرجي والبسيط بين العوامل التي تأتي بعد ذلك، فإن هذا يشير إلى أن هناك عامل واحد مسيطر ومهيمن على درجات الاختبار، أي أن افتراض أحادية البعد متحقق.

والطريقة الأخرى للكشف عن أحادية البعد باستخدام التحليل العاملي هي تفحص معاملات تشبع الفقرات على العامل الأول. فإذا كانت جميع الفقرات متشعبة على العامل الأول فهذا يعني تحقق افتراض أحادية البعد.

ويرى أورلاندو "Orlando" بأن وجود فقرة واحدة أو أكثر غير مشبعتين على العامل الأول لا يعني عدم توفر خاصية أحادية البعد ، ومع ذلك يجب فحص نتائج نموذج نظرية استجابة الفقرة من أي شذوذ ربما يظهر بسبب انتهاك هذا الافتراض، مثل أن يظهر بأن هناك فقرة واحدة أو أكثر لهما معالم تمييز منخفضة جدا (Orlando,w.d.).

2-الاستقلال المحلي Local Independence:

يعتبر هذا الافتراض امتدادا للافتراض الأعم وهو أحادية البعد، وهو يعني أنه إذا تم إزالة أثر العامل أو البعد أو السمة الكامنة خلف المقياس، فلن يوجد أي تغاير منتظم اضافي بين الفقرات. وبذلك فإن الاستقلال المحلي يمكن أن يظهر بين المجموعات الجزئية من الفقرات التي تسير في اتساق واحد. بحيث يكون بمقدور الباحث أن يختبر احصائيا أحد الافتراضيين نيابة عن الأخر . ويمكن التحقق من هذا الافتراض أمبريقيا عن طريق تثبيت القدرة عند مستويات مختلفة وحساب معامل (F) لكل فئة من مستويات القدرة، حيث يفترض أن معاملات الثبات (F) تقترب عند مستوى ثابت من القدرة من الصفر (Nandakumar & Ackerman, w.d.).

3-عامل السرعة في الاجابة Speeded ness:

يرى كل من هامبلتون وسواميناثان (Hambleton & Swaminathan, 1985) وعلام (1986) بأنه يمكن التحقق ما إذا كان الاختبار هو اختبار قوة أو سرعة من خلال حصر عدد المفحوصين الذين لم يتمكنوا من إكمال الاختبار موضع التطبيق.

4-تساوي مؤشرات التمييز Equal Discrimination Indices:

يمكن التحقق من تساوي مؤشرات التمييز من خلال فحص ذلك النوع من الارتباط بين الفقرة والدرجة الكلية للاختبار والمعروف بمعامل الارتباط ثنائي التسلسل الحقيقي Point-biserial correlation. فإذا كانت قيمة الانحراف المعياري لمعاملات الارتباط الثنائي التسلسل الحقيقي للفقرات صغيرا وتقترب من الصفر، فإن معنى ذلك أن مؤشرات التمييز متساوية في جميع الفقرات (Wiberg, 2004).

فإذا تحقق مثل الوضع السابق، فإن هذا يعني أن النموذج سوف يكون أحادي المعلمة أي نموذج راش Rasch Model (كاظم، 1988 أ : 42).

أما إذا اتضح من فحص قيم معاملات الارتباط الثنائي التسلسل الحقيقي بأنها ليست متساوية، فإن من الأفضل استخدام النموذج ثنائي المعلم 2PLM بدلا من النموذج أحادي المعلم 1PLM (Hambbleton et al. 1991).

5-الحد الأدنى من أثر التخمين Minimal Guessing:

يمكن التحقق ما إذا كان هناك أثر للتخمين (c) أم لا ، وذلك من خلال مؤشرات معاملات الصعوبة المحسوبة وفق النظرية الكلاسيكية (التقليدية). فإذا اتضح أن هناك فقرات صعبة، فإن المفحوصين غالبا ما سيلجئون إلى التخمين، وبذلك يمكن استخدام النموذج 3PLM، حيث أن النموذج ذا المعالم الثلاثة 3PLM له معلمة خاصة بالتخمين، والتي هي عبارة عن المقدار الذي يقترب فيه المنحنى من نقطة الصفر (Wiberg, 2004).

أما إذا كانت كل الفقرات سهلة نسبيا، أو إذا كان الاختبار يتكون من فقرات تكون الاستجابة عليها بإجابات مفتوحة وقصيرة فإن من غير المحتمل أن يكون النموذج المختار هو نموذج المعالم الثلاثة (Hambleton et al. 1991).

أما الطريقة الأخرى للتحقق من الحد الأدنى للتخمين فإنها تعتمد على فحص أداء المفحوصين ذوي القدرات المنخفضة، لأن هؤلاء المفحوصين هم الذين يلجئون لهذا السلوك عندما يواجهون الفقرات الصعبة (Leeson & Fletcher, w.d.).

ويتم ذلك التحليل من خلال اختيار 10% من المفحوصين الأقل قدرة على الدرجة الكلية للاختبار، ومن ثم دراسة أدائهم على الفقرات الأكثر صعوبة، ومقارنة نسب أولئك المفحوصين الأقل قدرة، الذين أجابوا إجابة صحيحة على تلك الفقرات الصعبة بالقيم النظرية، إذا حدث أن خمن المفحوصين بطريقة عشوائية على تلك الفقرات، مع ملاحظة أن تلك القيم النظرية للتخمين تعتمد على عدد فئات الاستجابة لكل فقرة (Hambleton & Swaminathan, 1985).

ويرى ويبرغ (2004) Wiberg بأن نسبة إجابة المفحوصين الأقل قدرة على الفقرات الصعبة إذا كانت قريبة من القيمة النظرية أو أكبر منها، فهذا يعني أن معلمة التخمين يجب أن تكون جزءا من النموذج.

ثانيا: التحقق من الخصائص المتوقعة للنماذج:

إن هذا المعيار للخصائص المتوقعة للنموذج هو معيار عام، فأيا كان نوع النموذج المستخدم، فإن هناك خاصيتان مهمتان لنماذج نظرية الاستجابة للفقرة يمكن فحصهما هما: خاصية الثباتية لمعالم القدرة Invariance of ability parameter ، وخاصية الثباتية لمعالم الفقرات Invariance of Item parameter. تعني خاصية الثباتية لمعالم القدرة، أن تقديرات القدرة للمفحوصين يفترض ألا تعتمد على خصائص الفقرات، أي على صعوبة أو سهولة الفقرات (Hambleton et al. 1991).

بينما تعني خاصية الثباتية لمعالم الفقرات أن تقديرات معالم الفقرات يفترض ألا تعتمد على طبيعة وخصائص المفحوصين، سواء كانت قدراتهم مرتفعة أم منخفضة، أو مهما كان تصنيفهم داخل العينة الكبيرة للدراسة (Wiberg, 2004).

ويتم التحقق من الخاصتين السابقتين على النحو التالي:

1-التحقق من ثباتية معالم القدرة:

يمكن التحقق من خاصية الثباتية لمعالم القدرة من خلال تقسيم فقرات الاختبار إلى فقرات سهلة وأخرى صعبة بناء على معاملات صعوبة الفقرة المستخرجة من النظرية الكلاسيكية (التقليدية)، ومن ثم يتم تقدير قدرات جميع المفحوصين في مجموعتي الفقرات السهلة والفقرات الصعبة، ثم تفحص العلاقة بين تقديرات قدرات المفحوصين في المجموعتين المختلفتين في خصائص الفقرات (معاملات الصعوبة). فإن التمثيل البياني لرسم الانتشار بين تقديرات معالم قدرة المفحوصين بناء على مجموعتي الفقرات السهلة والصعبة يمكن أن يعكس مدى تقارب أو اختلاف التقديرات . فإذا لم يتم الحصول على قيمة جيدة لمعامل الارتباط بين التقديرات، فإن ذلك يشير إلى أن خاصية ثباتية معالم القدرة لن تكون متوفرة عند نمذجة البيانات بذلك النموذج (Hambleton et al. 1991),(Wiberg, 2004).

2-التحقق من ثباتية معالم الفقرات:

ويتم ذلك من خلال تقسيم المفحوصين إلى مجموعتين أو فئتين، احداهما مرتفعة القدرة، والأخرى منخفضة القدرة، وذلك بناءً على درجاتهم الكلية المكتسبة على الاختبار (Hambleton & Swaminathan, 1985) ، وفي حالة إذا كان الاختبار محكمي المرجع، فإن تقسيم المفحوصين إلى مجموعتين مرتفعة القدرة ومنخفضة القدرة يمكن أن يعمل بناء على درجة القطع (Wiberg, 2004). بعد ذلك يتم عمل تقدير لمعالم الفقرات (الصعوبة والتمييز والتخمين) لكل مجموعة من المجموعتين السابقتين، ثم تمثل تقديرات معالم الفقرات للمجموعتين على رسم الانتشار لملاحظة ما إذا كانت العلاقة قوية أم ضعيفة بين التقديرات. إذا كانت العلاقة قوية فإن هذا يعني توفر خاصية الثباتية لمعالم الفقرات، بمعنى أن تقديرات معالم فقرات الاختبار لم تعتمد على خصائص عينة المفحوصين.

وتجدر الإشارة إلى أن الفحوصات السابقة لثباتية معالم القدرة وثباتية معالم الفقرات يتم اجراؤها في هذه الخطوة على النماذج الثلاثة المقترحة لتوفيق البيانات، بحيث يتمكن فحص الاختبار من مقارنة درجة توفر خصائص الثباتية في النماذج الثلاثة (Leeson & Fletcher, w.d.), (Wiberg, 2004).

ثالثا: التحقق من قدرة النموذج على التنبؤ بالدرجات الفعلية للاختبار:

إن جميع تطبيقات نظرية الاستجابة للفقرة تفترض ببساطة أن النموذج المستخدم هو نموذج صحيح، وفائدة وجدوى أي نموذج من نماذج نظرية الاستجابة للفقرة تعتمد على المدى الذي يعكس فيه النموذج البيانات الفعلية بدقة (Sinharay, 2005),(Orlando, w.d.).

وهناك عدة طرق للتحقق من قدرة النموذج على التنبؤ بالنتائج الفعلية للاختبار ومن ذلك ما يأتي:

1-مقارنة التوزيعات المشاهدة بالتوزيعات المتوقعة: ويمكن ذلك من خلال استخدام الرسوم البيانية لتوزيعات الدرجات المشاهدة والدرجات المتوقعة. فعلى قدر ما تكون تلك التوزيعات متقاربة أو متطابقة يكون ذلك النموذج الأفضل في مطابقة البيانات مقارنة بالنماذج الأخرى.

2-استخدام اختبار x^2 لاختبار جودة المطابقة:

يستخدم هذا الاختبار لفحص مدى مطابقة كل فقرة من فقرات الاختبار للبيانات المشاهدة للاختبار، وهو يقارن نسبة الإجابات الصحيحة على الفقرة (i) للمجموعة الجزئية التي قدرتها (h) مع النسبة المتوقعة للإجابات الصحيحة وفقا للنموذج المستخدم.

فإذا كانت قيمة (x^2) المحسوبة أكبر من القيمة الحرجة، ترفض الفرضية الصفرية، وبذلك نتوصل إلى أن منحنى خصائص الفقرة (ICC) لا يلائم الفقرة موضع الاختبار (Wiberg, 2004).

3-فحص البواقي: يعتبر فحص البواقي (البواقي المعيارية) أفضل طريقة لفحص مطابقة البيانات. والمعادلة المستخدمة لذلك كما يلي:

البواقي= القيم الملاحظة – القيم المتوقعة

$$R_{ij}=P_{ij} - E(P_{ij})$$

وتمثل P (احتمالية الإجابة بشكل صحيح على كل فقرة) في مقابل القدرات، وبالتحديد في مقابل نوعين من القدرات إحداهما القدرات المشاهدة أو الحقيقية أو ما تسمى أيضا التقليدية، والأخرى القدرات المتوقعة أو المتنبأ بها من خلال النموذج. كما يمكن أيضا تمثيل البواقي المعيارية حسب فئات القدرات لكل فقرة، حيث يظهر ما إذا كانت هناك نمطية معنية للبواقي المعيارية في أحد النماذج، كما تتم ملاحظة أي من النماذج يعطي توزيعا أكثر انتشارا وعشوائية للبواقي المعيارية (Hambleton et al. 1991).

4-التوزيع الطبيعي للقدرات:

يفترض أن تتوزع القدرة بشكل طبيعي وبوسط قيمته (صفر) وانحراف معياري قيمته (واحد) أي توزيع طبيعي معياري. ولذا فإن تقديرات القدرات للبيانات المشاهدة يتم تمثيلها بمدرج تكراري، وتقارن بالتوزيع الطبيعي المعياري بحيث ينظر هل تحقق ذلك التوقع أم لا؟ (Wiberg, 2004).

5-مقارنة إحصاءات عدم المطابقة مع معاملات الارتباط الثنائي التسلسل الحقيقي للفقرات:

وهي إيجاد علاقة بين إحصاءات عدم المطابقة للفقرة والارتباطات الثنائية التسلسل الحقيقة Point-biserial correlations. احصاءة عدم التطابق misfit يمكن أن تحدد بإيجاد معدل القيم المطلقة للبواقي المعيارية التي حصلنا عليها بعد مطابقة النموذج موضع الفحص بيانات الفقرة . ويكون الوضع المفضل هنا عندما لا يكون هناك نمط محدد لانتشار البواقي المعيارية المطلقة مع الارتباطات ثنائية التسلسل الحقيقية، وعندما تكون قيمة البواقي المعيارية المطلقة أقل ما يمكن (Hambleton et al. 1991).

مميزات نماذج نظرية الاستجابة للفقرة:

يمكن تلخيص أهم مميزات نماذج نظرية الاستجابة للفقرة فيما يلي:

-إن استخدام نماذج نظرية الاستجابة للفقرة يجعل بارامترات الفقرة مستقلة عن عينة الأفراد المستخدمة في تقدير هذه البارامترات، التي تمثل المجتمع المستهدف من تصميم الاختبار، كما أن تقديرات قدرات الأفراد مستقلة عن مجموعة فقرات الاختبار المستخدمة في الحصول على هذه التقديرات، المنتقاة من مجموعة كبيرة من الفقرات التي تم تدريجها باستخدام أحد نماذج الاستجابة للفقرة (),(Hambleton, 1989 :148) (281).

-إمكانية الحصول على مؤشر إحصائي (إحصاءات صدق المطابقة) يوضح مدى دقة تقدير قدرات الأفراد وبارامترات الفقرات (Hambleton & Swaminathan, 1985 :11).

-تحسين دقة وثبات النتائج من خلال تحديد وحذف الأفراد والفقرات غير المطابقتين للنموذج المستخدم، إضافة إلى حذف التأثيرات المتطرفة لقياسات كل من

-سهولة التمثيل البياني لبارامترات الفقرات وقدرات الأفراد.

-استخدام نماذج نظرية الاستجابة للفقرة يسهل دراسة كل من تحيز الأفراد، والفقرات، وصدق الاستجابات، والفروق الفردية، واستراتيجيات الاستجابة، وكذلك سهولة إعداد بنوك الأسئلة، والاختبارات الآلية، واستخدام الحاسب الآلي في تكرار العملية الإختبارية.

-يمكن تحديد الأفراد الذين يلجأون إلى التخمين، كما يمكن تكميم مدى استجاباتهم الشاذة، كما يمكن ترتيب هؤلاء الأفراد عن طريق جودة المطابقة لنموذج احتمالي معين.

تقليل عدد الفقرات اللازمة للقياس الدقيق، وبالتالي توفير الزمن اللازم للإجابة على فقرات الاختبار، وإعادة التحليل عند الصور البديلة للاختبار (Henning, 1984 : 132).

عيوب نماذج نظرية الاستجابة للفقرة:

يمكن تلخيص أهم عيوب نماذج نظرية الاستجابة للفقرة فيما يلي:

-تمثل هذه النظرية فرعا معقدا من فروع نظرية الاختبارات، واستيعاب البحوث السيكومترية التي تجري في هذا المجال ونشرها في الدوريات العلمية المتخصصة يحتاج إلى قدر كبير من الفهم الرياضي العالي، والإحصاء متعدد المتغيرات.

-معظم خبراء القياس الذين تناولوا هذا الموضوع وجهوا دراساتهم وبحوثهم لمنظري القياس، وليس للذين سيقومون بتطبيقه في بناء الاختبارات وتحليل وتفسير نتائجها، وفي غير ذلك من التطبيقات السيكولوجية والتربوية.

*** القوة التميزية للفقرة Item Discrimination:**

يعرف القياس الكلاسيكي مفهوم القوة التمييزية للفقرة بأنه مدى فاعلية فقرة الاختبار في التمييز بين المستويات المختلفة للسمة المراد قياسها لدى الممتحنين (علام، 2001: 216).

وتفصل مؤشرات تمييز الفقرة بين الممتحنين ذوي الدرجات المرتفعة والممتحنين ذوي الدرجات المنخفضة. وباستخدام هذه المؤشرات يستطيع الفرد أن يتعرف على الممتحنين الأقوياء ويمتلكون احتمالية أعلى للإجابة بطريقة صحيحة، والضعفاء المنخفضة درجاتهم ويمتلكون احتمالية أقل للإجابة بطريقة صحيحة (Crislip & Chin-Chance, 2001).

- **صدق القياس Measurement Validity:**

يعرف القياس الكلاسيكي مفهوم صدق القياس بأنه يتعلق بما يقيسه الاختبار وإلى أي مدى ينجح في قياسه. فالصدق يحدد قيمة الاختبار وصلاحيته في قياس ما وضع لقياسه (خطاب، 2001: 159).

- **ثبات القياس Measurement Reliability:**

يعرف القياس الكلاسيكي مفهوم ثبات القياس، بأنه مدى اتساق درجات الاختبار من قياس إلى آخر، أو القدرة على قياس الظواهر لعدة مرات، والحصول على نفس النتائج أو قريبا جدا من بعضها (El-Korashy, 1995 : 761).

الخصائص السيكومترية لفقرات الاختبار في إطار نموذج "راش" مايلي:

و من أهم الخصائص السيكومترية لفقرات الاختبار في إطار نموذج "راش" ما يلي:

- **صعوبة الفقرة Item Difficulty**

يدل مفهوم صعوبة الفقرة في ضوء نموذج "راش" بمستوى القدرة (θ) التي يكون عندها المنحنى المميز للفقرة أكثر انحدارا، وبعبارة أخرى كلما زادت صعوبة الفقرة زادت فرصة قدرة الفرد على اجتياز الفقرة بنسبة (50%) (Yen & Edwardson, 1999, 237).
ويعرف مفهوم صعوبة الفقرة أيضا بأنه "النقطة على متصل السمة الكامنة التي يتوقع عندها أن يكون احتمال الإجابة الصحيحة عن الفقرة دون تخمين تساوي (0,50)(علام، 2001: 218).
وقد أشار هامبلتون وسومیناثان (Hambleton & Swaminathan) إلى أن قيمة الصعوبة (b) تتراوح من (− ∞) إلى (+ ∞) نظريا، إلا أنه عمليا فقيمتها تتراوح بين (-3) إلى (+3).

- **قدرة الفرد: Ability**

لمرجع نجاح الفرد على الفقرات التي تعبر نقطة صفر التدريج عن صعوبتها وتقدر باللوجيت (كاظم، 1988أ: 56).
ويرى هامبلتون وسومیناثان (Hambleton & Swaminathan,1985) أن القدرة تتدرج من (− ∞) إلى (+ ∞)، إلا أنه نادرا ما لوحظ تجاوزها للمدى من (-3) إلى (+3).

- **القوة التميزية للفقرة Item Discrimination**

يعرف مفهوم القوة التمييزية للفقرة في ضوء نموذج "راش" هو معدل التغير في احتمال الاستجابة الصحيحة للأفراد على الفقرة بالنسبة لمستوى القدرة، وتقدر هذه القوة بمعامل التمييز الذي يقصد به الميل النسبي للمنحنى المميز للفقرة على محور القدرة (كاظم، 1988 أ : 74).

- **صدق القياس Measurement Validity:**

يعرف مفهوم صدق القياس في ضوء نموذج "راش" بأن تعرف الفقرات فيما بينها متغيرا واحدا، أي: إن فقرات الاختبار تتدرج من حيث صعوبتها بحيث تعرف متغيرا واحدا، كما يعني تدرج قدرات الأفراد على المتغير محددة تقديرات أدائهم على هذا الاختبار، وهذا يوضح أن كلا من صعوبات الفقرة، وقدرات الأفراد تتدرج على متصل واحد يمثل متغيرا واحدا (المرجع السابق، 1988 أ : 98).

- **ثبات القياس Measurement Reliability:**

يعرف مفهوم ثبات القياس في ضوء نموذج "راش" بتحقق ما بقي من مطالب الموضوعية في القياس، أي: عندما تستخدم أداة القياس التي أنشئت باستخدام هذا النموذج، حيث يتحقق:

-استقلال القياس عن الاختبار المستخدم.
-استقلال القياس عن مجموعة الأفراد المؤدية للاختبار.

أي أن استقلالية القياس وتحرره الذي يوفره نموذج راش تتيح الفرصة لثبات القياس بحيث لا يختلف القياس (سواء لقدرة الفرد أو صعوبة الفقرة) باختلاف عينة التدريج أو باختلاف الاختبار المستخدم لقياس قدرة ما (كاظم، 1988 أ : 99)، (إسماعيل وآخرون، 1994 : 131).

كما أن نموذج راش لا يعطي فقط معامل ثبات الاختبار الكلي، ولكنه يعطي أيضا معامل ثبات لكل فرد وفقرة (El-Korashy, 1995 : 760).

خاتمة:

تهدف نظرية الاستجابة للمفردة و النماذج الرياضية المرتبطة بها إلى تقدير جميع إحصاءات المفردة و القدرة، و كلما كان هناك ملاءمة Fit بين النموذج المستخدم و مجموعة البيانات أدى ذلك إلى الحصول على تقديرات دقيقة لهذه الإحصاءات، حيث يجب وضع كل من الأفراد و المفردات على قائمة للقدرة، و يتم ذلك من خلال عمليات التقدير طالما يوجد علاقة تقارب ممكنة بين الاحتمالات المتوقعة للممتحنين والاحتمالات الواقعية الفعلية لأدائهم في كل مستوى من مستويات القدرة، مع الأخذ في الاعتبار أن تقديرات إحصاءات المفردة و تقديرات قدرة الممتحنين يجب أن تعدل وتراجع بصفة مستمرة حتى يمكن الحصول على أقصى اتساق ممكن بين التنبؤات على أساس تقديرات القدرة و متغيرات المفردة و بين البيانات .

تمهيد:

تعد الاختبارات التحصيلية أكثر أساليب التقويم شيوعاً، بل قد تكون الوحيدة في كثير من الأحيان, ولاشك أن لها دوراً مهماً في العملية التعليمية وخاصة في تقويم تحصيل المتعلمين وفي القرارات التربوية المبنية على ذلك, فبناءً على نتائج الاختبارات يسمح للطالب بالانتقال من صف لآخر أو يبقى في صفه, وبناءً عليها يوجه الطالب لتخصص دون آخر, وعملية بناء واستخدام الاختبارات التحصيلية تتطلب مهارات متعددة من المعلم، واتباع خطوات علمية منظّمة في إعدادها وتنفيذها وتصحيحها.

1- **تعريف الاختبار التحصيلي:**

1-1- **تعريف الاختبار (test) عند علماء التربية:**حسب كرو نباخ هو طريقة منظمة لمقارنة سلوك شخصية بسلوك شخصية أخرى (احمد :2011م)،أما ثورنديك فهو طريقة لقياس الكم من الشيء على اساس ان أي شيء موجود يكون موجوداً بكمية معينة . (احمد :2011م)،ويعرفه بين بأنه مجموعة من المثيرات التي يمكن بواسطتها قياس بعض العمليات العقلية أو الخصائص النفسية بطريقة كمية أو كيفية (احمد :2011م)

1-2-تعريف الاختبار التحصيلي:يعرفه (عودة،1993) بأنه طريقة منظمة لتحديد مستوى تحصيل الطلبة لمعلومات ومهارات في مادة دراسية كان قد تعلمها مسبقا بصبغة رسمية من خلال إجابته عن على مجموعة من الفقرات تمثل محتوى المادة الدراسية.

2 - طرق تقويم وقياس التحصيل الدراسي للمتعلمين:

لقد لجأت المدارس الحديثة إلى استخدام طرق مختلفة لتقويم وقياس تعليم أبنائها واتخذت بعضها كمقياس لقيمة المعلومات والبعض الآخر كوسيلة لتحسين عملية التعليم وهنا نشير وبإيجاز إلى الطرق التقويمية والقياسية التالية :

2-1- الاختبارات الشفوية: وفيها يوجه المعلم للمتعلم أسئلة شفوية، ويستجيب المتعلم بالطريقة نفسها، وهي من أقدم أنواع الاختبارات وتستخدم في تقويم مجالات معينة من التحصيل كالقراءة الجهرية وإلقاء الشعر وتلاوة القرآن الكريم.

2-2- الاختبارات المقالية: وهي الاختبارات ذات الإجابة الحرة، ويطلق عليها أحياناً اسم الاختبارات الإنشائية أو التقليدية، ولأن هذه الاختبارات تتيح للمتعلم فرصة إصدار جوابه الخاص به وكيفية تنظيم الإجابة وتركيبها فهي تساعد على قياس أهداف معقدة معينة كالابتكار والتنظيم والمكاملة بين الأفكار والتعبير عنها باستخدام ألفاظه الخاصة، ومن نقاط ضعف هذا النوع من الاختبارات قلة شمول أسئلتها للمادة الدراسية جميعها، وتأثر تصحيحها بالعوامل الذاتية للمصحح .

2-3- **الاختبارات الموضوعية:** ويطلق عليها اسم الاختبارات الحديثة مقارنة بالاختبارات المقالية، وقد اشتهرت باسم الموضوعية لما تمتاز به من دقة وموثوقية ولعدم تأثر تصحيحها بالعوامل الذاتية للمصحح، وهي على أنواع متعددة أشهرها: الصواب والخطأ والاختيار من متعدد، والمقابلة، والتكميل .ومع ما تتميز به الاختبارات الموضوعية من موضوعية وشمول وارتفاع في معاملي الصدق والثبات وسهولة في التطبيق والتصحيح إلا أن أعدادها صعب وتقتصر عن قياس بعض الأهداف التعليمية المعقدة كالتركيب والتقويم كما أنها تفتح مجالاً للغش والتخمين من قبل المفحوصين، ولذلك فإنه ينصح بعدم استخدامها منفردة دون الاختبارات المقالية، بل يفضل المزج بين النوعين، وهذا يعود طبعاً إلى طبيعة المادة الدراسية .مجلة العلوم الإنسانية أ.د/ يحي علوان2007 ص25

2-4 **الاختبارات الأدائية:** وهي الاختبارات التي تقيس أداء الأفراد بهدف تعرف بعض الجوانب الفنية في المادة المتعلمة وفي بعض المهارات التي لا يمكن قياسها بالاختبارات الشفهية أو الكتابية من مقالية

وموضوعية، وبذلك فهي لا تعتمد على الأداء اللغوي المعرفي للطالب، وإنما تعتمد على ما يقدمه الطالب من أداء عملي في الواقع، ولهذا النوع من الاختبارات التحصيلية عدة أغراض.

3- أغراض الاختبارات التحصيلية:

3-1-**التشخيص**: أي محاولة تعرف جوانب القوة والضعف لدى الطالب في جانب من جوانب التحصيل للاستفادة من النتائج في تدعيم جوانب القوة ومعالجة جوانب الضعف مع ما يستدعيه ذلك من تقويم لأسلوب التدريس أو المنهاج أو المرافق التعليمية المختلفة ومصادر التعلم، مع الإشارة إلى أن هناك اختبارات خاصة بالتشخيص، ولكن هذا لا ينفي عن الاختبارات التحصيلية الصفية العادية التي تتناسب مع وظيفة التشخيص .

3-2-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 على اختبارات تحصيلية، أو اختبارات خاصة، أو وسائل قياس أخرى من مقابلات شخصية واستبانات وقوائم وما إلى ذلك.

3-3-قياس مستوى التحصيل: والذي عبرنا عنه سابقاً بمدى تحقيق الأهداف التعليمية لدى المتعلم في مادة دراسية يعينها، وفي المواد الدراسية جميعها، هذا وإن الاختبارات التحصيلية في معظمها إنما تنصب على تحقيق هذا الهدف بقصد الأخذ بنتائجه في تحسين مستوى التعلم وترفيع الطلاب إلى صفوف أعلى وما إلى ذلك.

4-أنواع أسئلة الاختبارات

أولا : الفقرات ذات الإجابة المنتقاة:

يمتاز هذا النوع من الفقرات بالموضوعية حيث يتم تقدير العلامة بعيدا عن ذاتية المصحح ، وعلى الطالب اختيار الإجابة الصحيحة من بين عدد من البدائل ويقسم إلى الأنواع الآتية

1- **فقرات الصواب والخطأ:** الفقرة في هذا النوع من الأسئلة عبارة عن جملة خبرية يطلب إلى الطالب أن يجيب عنها بالصواب ،أ والخطأ ويهدف هذا النوع من الأسئلة إلى قياس قدرة الطالب على التمييز بين المعلومات الصحيحة والخاطئة،ويستخدم لقياس المستويات الدنيا من المجال المعرفي.

2- **فقرات الاختيار من متعدد:**يعتبر هذا النوع من الأسئلة من أفضل أنواع الأسئلة وأكثرها صدقا وثبات ا واستخداما في الاختبارات ويستخدم هذا النوع لقياس معرفة الحقائق العلمية واستيعابها وتطبيق المعرفة العلمية في مواقف جديدة ، ويتصف هذا النوع من الأسئلة بتغطية محتوى الكتاب المدرسي ، -سهولة التصحيح، - صعوبة الإعداد والتصميم ، صعوبة قياس القدرات التعبيرية والأدائية والإبداعية و ارتفاع تكاليف إعداده.

3- **أسئلة المطابقة أو المزاوجة :**يتألف هذا النوع من الأسئلة من مجموعتين) قائمتين من الكلمات أو العبارات وتسمى القائمة الأولى المقدمات ، وتسمى القائمة الثانية الإجابات بحيث يكون لكل مقدمة إجابة في القائمة ، ويطلب إلى الطالب أن يربط كل كلمة أو بما يناسبها ، ويستخدم هذا النوع لقياس قدرة الطالب على ربط المفاهيم والمبادئ والتعميمات العلمية.

ثانيا : الفقرات ذات الإجابة المصوغة:

يطلب إلى الطالب في هذا النوع من الفقرات أن يصوغ الإجابة بكلماته وبأسلوبه ولكن بدرجات متفاوتة من الحرية بحسب درجة تحديد الإجابة التي تختلف باختلاف نوع الفقرة .وتشتمل الفقرات ذات الإجابة المصوغة على الأنواع الآتية:

1- **فقرات التكميل:** وهي جملة خبرية غير مكتملة المعنى ، ويطلب إلى الطالب أن يكملها بوضع الكلمة المناسبة ، أو شبه الجملة أو الرمز ، أو الرقم . ويعتبر هذا الشكل من اكثر أشكال فقرات الإجابة المصوغة تقييدا لحرية الطالب في صياغة الإجابة .

2- فقرات الإجابة القصيرة: الإجابة على هذا النوع من الفقرات أطول من إجابة فقرة التكميل ، فهي تعطي درجة أعلى من الحرية للطالب في الاجابة عليها . خاصة إذا ظهرت الفقرة على شكل سؤال بدلا من جملة غير تامة المعنى.

3- الفقرات الإنشائية المحددة وحل المسائل: هذا النوع من الأسئلة يعطى الطالب حرية أكبر من أسئلة التكميل والإجابة القصيرة ، ويقسم هذا النوع إلى قسمين:

أ . الفقرات الإنشائية محددة الإجابة، وحل المسائل الحسابية أوالفقرات التي تحتمل إجابتها نقاطا محددة ، كأن يعدد أسبابا ، أو يذكر مكونات ، أو يحل مسألة حسابية.

ب . الفقرات الإنشائية مفتوحة الإجابة: الفقرات التي تعطي حرية للطالب بأن يجيب على الفقرة دون قيود على طول الإجابة أو تنظيمها أو زمن الإجابة.

ثالثا : الأسئلة التركيبية:

السؤال التركيبي ويعنى بوحدة الموضوع ، ويتألف من عدة فقرات متدرجة في صعوبتها، ومتنوعة في المهارات التي تقيسها ، ويراعي في بنائها ما يراعى في بناء الأسئلة الأخرى ويمكن تقديم السؤال فيها بأحد الأشكال الآتية:

1- **المشكلة:** يصاغ السؤال بطريقة تمكن الطالب عند الإجابة عليه الافادة من المعلومات المتوفرة في السؤال ، وما لديه من معلومات حول الموضوع ؛ أي أن يصاغ السؤال على شكل مشكلة ، ويوظف الطالب هذه المعلومات والمعرفة السابقة في حل هذه المشكلة.

2-الأشكال التوضيحية: يصاغ السؤال بتقديم المعلومات عن طريق رسم ، أو جدول ، أو صور...

3- **الموقف الحياتي:** يصاغ السؤال على شكل موقف حياتي مألوف لدى الطالب.

خطوات بناء الاختبار التحصيلي الجيد :

اتفق معظم المختصين بالقياس النفسي و التربوي بأن الاختبارات التحصيلية تمر بعدة خطوات، و ذلك لتحقق أكبر قدر ممكن من الأهداف التربوية المرجوة، و تكون عاملاً فاعلاً و مساعدًا في قياس النواتج التعليمية، نورد هذه الخطوات فيما يلي :

أوّلاً : تحديد الغرض من الاختبار

على المعلم تحديد الغرض من الاختبار التحصيلي قبل البدء في إعداده، فمنها ما يكون للقبول و الانتقاء و التصنيف في مجموعات، و منها ما يكون لتشخيص الصعوبات التي تواجه المتعلمين و دعم خطط التوجيه و الإرشاد، و منها ما يستخدم في البحوث التربوية و التطور التربوي، و منها ما يتعلق بقياس مستوى التحصيل ودرجة النمو العلمي للمتعلم و اتخاذ القرارات الإدارية المختلفة في ضوء ذلك، ويساعد تحديد الغرض من الاختبار في توجيه الخطوات اللاحقة كتحديد نوع الفقرات و توقيت الاختبار و ما إلى ذلك (سنان، 2006 م)، و (الكبيسي، 2007 م)، و (الحريري، 2007 م).

و حدد الباحث غرضه من بناء اختبار تحصيلي، و هو تطوير لأدوات قياس التحصيل الدراسي و الوصول إلى العدد الأمثل من البدائل المستخدمة في اختبارات الاختيار من متعدد و التي يكون صدقها و ثباتها عاليًا في مادة الرياضيات.

ثانيًا : تحديد الأهداف التعليمية

الأهداف التعليمية هي مجموعة من العبارات أو الصياغات التي توضح ما سوف يكون عليه سلوك التلميذ بعد اكتسابه للخبرة التعليمية داخل و خارج جدران المؤسسة التربوية (خضر، 2004 م) و تحديد نواتج العلم المخطط لها قبل البدء بعملية التدريس يعد من أكثر الخطوات أهمية بالنسبة لمعد الاختبار التحصيلي، و

يكون ذلك من خلال صياغة الأهداف السلوكية صياغة إجرائية محددة و واضحة بعيدًا عن الغموض و العمومية.

و قد صنف بلوم عام 1956 م الأهداف السلوكية إلى ثلاثة مجالات هي :

أ- **المجال المعرفي :**

و يتضمن الأهداف المتعلقة باكتساب المعارف و تطبيقها و فهمها، و تقع الغالبية العظمى من الأهداف التربوية و التدريسية ضمن هذا المجال، و قسمها بلوم إلى ستة مستويات فرعية، و هي :

1- **التذكر :**و يتمثل في قدرة الطالب على تذكر المعارف أو المعلومات السابقة واسترجاعها، ، و من الأفعال السلوكية التي تستخدم في صياغة هذا المستوى : يذكر، يسمي، يصف، يعدد.

2- **الفهم :**و هو القدرة على إدراك و استيعاب معنى المادة أو النص الذي يدرسه الطالب، و من أفعاله السلوكية التي تستخدم في صياغة هذا المستوى : يفسر، يترجم، يوضح، يعطي،

3- **التطبيق :** و يعني استخدام ما تم تعلمه من قوانين و مفاهيم أو قواعد او مبادئ ، ومن الأفعال السلوكية التي تستخدم في صياغة هذا المستوى : يطبق، يحل، يربط، ...

4- **التحليل :** و يعني القدرة على تجزئة أو تحليل المادة إلى عناصرها أو مكوناتها، و من الأفعال السلوكية التي تستخدم في صياغة هذا المستوى : يحلل، يصنف، يستنبط، يميز.

5- **التركيب :** و يعني القدرة على جمع عناصر أو أجزاء لتكوين كل متكامل ، و من الأفعال السلوكية التي تستخدم في صياغة هذا المستوى : يركب، يؤلف، يصمم، يكون، يبني.

6- **التقويم ::** و هو أعلى مستويات المجال المعرفي، و يعني القدرة على إصدار أحكام حول قيمة الأفكار أو اتخاذ قرارات مناسبة، و من أفعاله السلوكية التي تستخدم في صياغة هذا المستوى : يقوم، يحكم، يقارن، ينتقد. (فتح الله، 2000 م)، (الحريري، 2007 م).

ب- **المجال الوجداني :**و يهتم هذا المجال بكل ما يرتبط بالعاطفة الإنسانية من قيم و اتجاهات وميول و تقبل و استجابة. (زيتون، و عبد الحميد، 1995 م).

ج- **المجال النفسي حركي (المهاري) :** يتضمن هذا المجال الأهداف التي تؤكد على نتاجات التعلم ذات العلاقة بالمهارات الحركية، و يشتمل هذا المجال على عدة مستويات : الملاحظة والتقليد، التجريب و الممارسة، الإتقان ...(زيتون، و عبد الحميد، 1995 م).

ثالثًا :تحليل المحتوى :

يقصد بتحليل المحتوى الوصول إلى مفردات المقرر الدراسي، أو إحصاء المعلومات الأساسية في المقرر الدراسي أو هو إحصاء المعارف و المهارات الأساسية المتضمنة في الدروس، و ذلك بهدف ضمان تمثيل فقرات الاختبار لمفردات المحتوى تمثيلاً متكافئًا (الكبيسي، 2007 م)، و (الحريري، 2007 م).

رابعًا :إعداد جدول المواصفات

جدول المواصفات عبارة عن : مخطط ثنائي الطريقة يتحدد فيه عدد الاسئلة في كل خلية بناءً على المحتوى و الهدف (سيد و سالم، 2005 م)، و يميل الباحث إلى التعريف الذي أورده الكبيسي (2007 م : 141) بأن جدول المواصفات هو "مخطط تفصيلي يمثل محتوى الاختبار، و يربط محتوى المادة الدراسية بالأهداف التعليمية السلوكية، و بين الوزن النسبي الذي يعطيه المعلم لكل موضوع من الموضوعات المختلفة، و الأوزان النسبية للأهداف المعرفية السلوكية في مستوياتها المختلفة".

و لجدول 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 و شعور إيجابي بعدالة الاختبار و شموليته لجميع أجزاء المقرر الدراسي، مما يساعده في تنظيم وقته و توزيع جهده أثناء المذاكرة لجميع أجزاء المنهج. (كاظم، 2001 م)، و(الردادي، 2006 م).

خامسًا : اختيار نمط أسئلة الاختبار

اختيار النمط المناسب للاختبار يتوقف علة مجموعة من العوامل، و منها :

1- **الهدف المراد قياسه :** إذا كان الهدف الكشف عن الحقائق و النظريات فالأفضل أن يختار المعلم أسئلة الخطأ و الصواب، و لقياس الأهداف التربوية المتقدمة كالفهم و الاستيعاب فيفضل في هذه الحالة الأسئلة ذات الاختيار من متعدد، بينما تستخدم الأسئلة المقالية لقياس الحقائق و الأفكار المترابطة.

2- **طبيعة المحتوى المراد قياسه :** من الممكن استخدام جميع أنماط الفقرات في أي مادة دراسية إلا أن بعض تلك الأنماط تكون أكثر ملاءمة لمحتوى مادة معينة، كمادة النصوص الأدبية التي تناسبها الأسئلة المقالية، و مادة الرياضيات التي تناسبها الأسئلة الموضوعية و هكذا.

3- **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 خليطًا من الأسئلة المقالية و الموضوعية، أما من يفتقد إلى هذه المهارة و يتقاعس في بذل الجهد و الوقت لإعداد الأسئلة الموضوعية الجيدة فإنه يضطر إلى وضع أسئلة مقالية.

4- عمر التلميذ و مستواه التعليمي : يقل عدد الأسئلة لدى طلاب المراحل الدنيا، و ذلك لأنهم يحتاجون إلى وقت أطول في الاستجابة للفقرة الواحدة مما يحتاجه طلاب المراحل العليا.

5- الوقت المخصص لتحضيرها : يفضل وضع أسئلة الاختبار من متعدد عندما يكون الوقت كافيًا، و عندما يكون الوقت غير متاح تستخدم الأسئلة المقالية وأسئلة التكميل و الإجابة القصيرة.

6- **توفر الإمكانات المادية :** ارتفاع تكاليف طباعة و تصوير الأسئلة الموضوعية بعكس الاسئلة المقالية، و بالتالي يتوقف استخدام الأسئلة الموضوعية على توفر الإمكانات المادية. (الزيود، و عليان، 2002 م)، و(الحريري، 2007 م).

سادسًا : مراجعة أسئلة الاختبار

1- **نمط المفردة :** إذا احتوى الاختبار على أكثر من نمط من أنماط الفقرات فمن المفضل ترتيبها حسب النمط و الفقرات ذات النمط الواحد مع بعضها البعض.

2- **صعوبة المفردة :** و يتم 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.

3- **المحتوى :** و يقصد بهذا الترتيب تسلسل الفقرات في السؤال الواحد تسلسلاً منطقيًا لمحتوى المادة الدراسية.

4- **مستوى الأهداف :** و ذلك بوضع الفقرات التي تقيس مستوى معينًا من الأهداف بشكل متسلسل، مع إمكانية ملاحظة أن مستوى الأهداف قد يستخدم لتدريج الفقرات بناء على صعوبتها. (الشيخ و آخرون، 2004م)، (الردادي، 2006م).

ثامنًا : صياغة تعليمات الاختبار

تمثل التعليمات إرشادات هامة توجه التلميذ و ترشده إلى كيفية الاستجابة لفقرات الاختبار، و تلعب التعليمات الواضحة و المفهومة دورًا كبيرًا في تقليل أسئلة التلاميذ واستفساراتهم داخل قاعة الاختبار و تنقسم التعليمات إلى نوعين :

1- **التعليمات الخارجية :** و هي ورقة منفصلة توضع فوق أوراق أسئلة الاختبار، وعدد الأسئلة و أنواعها، و زمن الإجابة المخصص ، و طريقة و مكان الإجابة، و بعض الأمثلة التوضيحية و إمكانية الإجابة عن بعض الأسئلة و ترك البعض الآخر، و بيان عقوبة الغش.

2- **التعليمات الداخلية :** و توضع قبل كل سؤال، لتوضيح المطلوب من السؤال مع ذكر مثال لكيفية الإجابة. (الصمادي، و الزرابيع، 2004 م)، و (سيد، و سالم، 2005م).

تاسعًا : إخراج الاختبار و طباعته

يتكون الاختبار عادة من مجموعة من الأوراق (كراسة الاختبار)، و التي يجب على المعلم عند إخراجها ملاحظة أمور كثيرة، أهمها :

1- أن تحتوي كراسة الاختبار على تعليمات الإجابة عليه،وأن تكون طباعة الأسئلة واضحة وأن يراعى الفصل بين التعليمات و الأسئلة.

2- أن يراعى الفصل بين كل سؤال و الذي يليه بمسافة معقولة، وأن لا يجزئ السؤال على صفحتين متتاليتين.

3- أن يفصل بين كل نوع أو شكل من أشكال الأسئلة و الشكل الآخر بخط، وأن يراعى الفصل بين مقدمة السؤال و بدائله بمسافة معقولة،ويراعى الترتيب المنطقي لأوراق الاختبار، و تثبيتها بشكل يسهل تناول الاختبار و تداوله. (الردادي، 2006 م).

عاشرًا :تجريب الاختبار

و يشير الدوسري (2001 م : 207) إلى أن الاعتماد على نتائج اختبارات لم يتم تجريب أسئلتها يؤدي إلى ضعف مستواها من حيث الصدق و الثبات، و بالتالي عدم مقدرة تلك الاختبارات على تحقيق أغراضها، و يذكر أن هناك ثلاثة طرق لتجريب الاختبار،وهي:

1- تجريب الاختبار على أساس الاختبار الفعلي.
2- تجريب الاختبار على أساس اختبار تجريبي.
3- دمج الأسئلة ضمن اختبار فعلي.

أحد عشر : تطبيق الاختبار الفعلي

الأفضل أن يعقد الاختبار في نفس المكان الذي كان يتلقى التلاميذ فيه دروسهم، مع ضرورة بعده عن الضوضاء و الضجيج، و أن تتوفر فيها الإضاءة والتهوية الكافيتين، مع ضرورة تواجد معلم المادة الدراسية يوم الاختبار لتوفير نوع من الراحة النفسية و الطمأنينة لدى التلاميذ، و يجب عدم تشتيت انتباه التلاميذ وتركيزهم و يستحسن تذكيرهم بالوقت المتبقى للاختبار بين الحين و الآخر، و ينبغي عدم التهاون بالغش

بجميع أشكاله، لأنه يؤثر على مصداقية النتائج، و لذلك يجب وضع مساقات كبيرة بين التلاميذ عند تطبيق الاختبار و مع ضرورة المراقبة المشددة. (كاظم، 2001 م)، و (الصمادي، و الزرابيع، 2004 م)

الثاني عشر : تصحيح الاختبار

تعتمد طريقة التصحيح على نمط الفقرات، فعندما تكون الفقرات ذات استجابة مختارة أمكن تصحيحها يدويًا عن طريق عمل ورقة إجابة ؛ و قد يكون التصحيح آليًا، حيث يستخدم في هذا النوع من التصحيح آلات إلكترونية مبرمجة لهذا الغرض و في هذه الحالة يمكن استخدام نماذج إجابة تكون مصممة مسبقًا و ذات مقاسات معينة تتناسب مع أجهزة التصحيح. (الحريري، 2007 م).

أما إذا كانت الفقرات ذات استجابة منشأة فهناك طرقتان للتصحيح، هما :

الطريقة التحليلية، و الطريقة الكلية، و استخدام أي منهما يعتمد على عوامل كثيرة، كالهدف من الاختبار، و الوقت المتوفر لدى المعلم للقراءة، و نوع السؤال فيما غذا كان محددًا أو مفتوحًا.

و في الطريقة التحليلية تجزأ الإجابة على عناصر، و تكون 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 المصحح، و تكون درجة الثبات متدنية لأن الدرجة تمنح حسب ذاتية المصحح و العوامل النفسية أو المزاجية الخاصة به. (كاظم، 2001 م)، و (الدوسري، 2001 م).

الثالث عشر : تحليل فقرات الاختبار

عرف النبهان (2004 م : 191) تحليل الفقرات بأنه "عملية تتعلق باستقصاء الخصائص الإحصائية لاستجابات الطلبة على كل فقرة من فقرات الاختبار، بينما نجد الكبيسي (2007 م : 168) يقصد بتحليل فقرات الاختبار بأنه "استخراج معاملي الصعوبة و التمييز و تحديد فاعلية البدائل (المموهات)".

و بذلك تكون أهم الخصائص الإحصائية لمفردات الاختبار :

1- صعوبة المفردة Item Difficulty، و يرمز لها بالرمز (P).
2- تمييز المفردة Item Discriminatio، و يرمز لها بالرمز (D).
3- جاذبية البدائل (المموهات) DistractorAttractivness، و يرمز لها بالرمز (DA).

أوّلاً : معامل السهولة

أورد علام (2007 م : 251) بأن "سهولة المفردة يقصد بها النسبة المئوية لعدد الطلاب الذين أجابوا على المفردة إجابة صحيحة، و تكون المفردة صعبة جدًا إذا أجاب عليها جميع الطلاب بإجابة خاطئة و تكون هذه النسبة صفرًا، و عند إجابة جميع الطلاب بإجابة صحيحة تكون هذه المفردة سهلة جدًا عنده تكون النسبة المئوية 100%، و في الحالتين لا تفيد في الكشف عن الفروق الفردية بين الطلبة في امتلاكهم للصفة أو السمة التي يتوقع من الاختبار قياسها".

كنما ذكر علام (2007 م : 253) " بأن مدى معاملات السهولة المحصورة بين (0,30) و (0,70) تكون مقبول، و يفضل أن تكون قيمته (0,50) حيث أن هذه القيمة تجعل التمييز بين المجموعتين العليا و الدنيا من الطلاب أكبر ما يمكن".

طرق حساب معامل السهولة :

أ- في حالة النسب المئوية نستخدم المعادلة رقم (5) :

$$\text{معامل السهولة} = \frac{\text{عدد الإجابات الصحيحة}}{\text{عدد الإجابات الصحيحة + عدد الإجابات الخاطئة}} = \frac{\text{ص}}{\text{ص + خ}} \quad \ldots\ldots(5)$$

(إسماعيل، 2007 م).

ب- في حالة الترتيب

نستخدم المجموعتين الطرفيتين و فيها يتم استخراج معامل سهولة المفردة بالخطوات التالية :

1- ترتيب درجات الطلبة الكلية ترتيبًا تنازليًا.

2- تحديد الفئة العليا (27%) ؛ و هم من حصلوا على أعلى الدرجات، والفئة الدنيا (27%) ؛ و هم من حصلوا على أدنى الدرجات، تطبيق المعادلة التالية معادلة رقم (6) :

$$\text{معامل السهولة} = \frac{\text{ن1} + \text{ن2}}{2\ \text{ن}} \quad \ldots\ldots(6)$$

حيث أن :

ن1 : عدد الطلبة الذين أجابوا عن المفردة إجابة صحيحة من الفئة العليا.

ن2 : عدد الطلبة الذين أجابوا عن المفردة إجابة صحيحة من الفئة الدنيا.

ن : عدد الطلبة في إحدى المجموعتين (إسماعيل، 2007 م).

و حساب معامل السهولة بهذه الطريقة يتجاهل درجات المجموعة الوسطى و التي تشكل حوالي (46%) من مجمل الطلبة.

ثانياً : معامل التمييز

أورد علام (2007 م : 254) تعريفًا لمعامل تمييز المفردة بأنه " يشير إلى درجة تمييز المفردة بين مرتفعي التحصيل و منخفضي التحصيل من الطلاب في الاختبار".

و بمعنى آخر تكون الفقرة مميزة جدًا إذا كان أداء الطلبة عليها يرتبط ارتباطًا موجبًا بأدائهم على الاختبار.

و يذكر (إسماعيل، 2007 م : 11) أن حدود معامل التمييز يتراوح بين -1 و +1، و الحد الأدنى لمؤشر التمييز هو 0,30 على الأقل، و عندما تصل إلى 0,40 فأكثر فإنها تدل على قوة تمييز بين المجموعتين، و من 0,20 إلى 0,30 تعتبر قوة تمييز منخفضة و لكن لابأس بها، و يعتبر التمييز ضعيفًا و تحتاج الفقرة إلى تعديل إذا تراوح مؤشر التمييز بين 0,10 إلى 0,20، بينما تحذف الفقرة أو تعدل جذريًا إذا كان مؤشر التمييز أقل من 0,10.

طرق حساب معامل التمييز :

أ- طريقة المجموعتين الطرفيتين :

عند استخدام هذه الطريقة يتطلب أولاً ترتيب الطلبة ترتيبًا تنازليًا حسب درجاتهم الكلية، و من ثم تحديد أعلى 50% و أقل 50% إذا كان عدد الطلبة قليلاً، أو أعلى 27% و أقل 27%إذا كان العدد كبيرًا.

و يتم حساب معامل التمييز بهذه الطريقة باستخدام المعادلة التالية، معادلة رقم(7) :

$$\text{معامل التمييز} = \frac{\text{ص ع} - \text{ص د}}{0,27\text{x العدد الكلي}} \quad \ldots\ldots(7)$$

حيث : ص ع : هي عدد الإجابات الصحيحة في المجموعة العليا.
ص د : هي عدد الإجابات الصحيحة في المجموعة الدنيا.
(اسماعيل، 2007 م)

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

و تتخلص هذه الطريقة في حساب معامل الارتباط بين الدرجة على الفقرة من جهة و الدرجة الكلية على الاختبار من جهة أخرى ؛ حيث إن معامل التمييز الناتج بهذه الطريق يأخذ بعين الاعتبار استجابات جميع الطلبة، و ذلك باستخدام معامل الارتباط الثنائي الأصيل، أو باستخدام معامل الارتباط الثنائي المتسلسل الحقيقي (Point BiserialCorrelatio)، و ذلك عندما تشكل الدرجات على المفردات متغيرًا ثنائيًا (0 و 1) و تشكل الدرجات ككل متغيرًا متصلاً، و عند معامل الارتباط الثنائي الأصيل نقوم بتطبيق المعادلة رقم (8) :

$$\text{ر ث ص} = \frac{\text{م1} - \text{م2}) \times \text{بزأ؟ ثثأ "}}{\text{ع}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (8)$$

حيث :
م1 : هي متوسط درجات الأفراد الذين أجابوا إجابة صحيحة على المفردة.
م2 : هي متوسط درجات الأفراد الذين أجابوا إجابة خاطئة على المفردة.
أبز : معامل السهولة.
أ ثث : معامل الصعوبة.
ع : الانحراف المعياري. (اسماعيل، 2007 م).

ثالثاً : معامل جاذبية (فاعلية) المموهات :

المموهات أو المشتتات : هي البدائل أو الخيارات غير الصحيحة في سؤال الاختيار من متعدد.
و يفترض في المموه أن يكون جذابًا للمفحوصين و خاصة ممن ينتمون إلى الفئة الدنيا، و أن لا تقل نسبة الطلبة الذين يختارونه عن 5% منهم، أكثرهم من فئة مجموعة الأداء المنخفض.
و المموه الجيد و الفعال هو ذلك البديل الذي يتمتع بمعامل جاذبية سالب و كبير؛ إذ يجب مراجعة أو تطوير أو استبدال أي بديل لا يختاره أحد من الطلبة، أو أن تكون جاذبية موجبة أو صفرًا.

طريقة حساب معامل جاذبية المموه :

يمكن حساب معامل جاذبية المموه بنفس طريقة حساب معامل تمييز الفقرة حسب المعادلة رقم (9) :

$$\text{معامل فاعلية المموه} = \frac{\text{ن ع} - \text{ن د}}{\text{ن}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (9)$$

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. (اسماعيل، 2007 م).

شروط الاختبار الجيد :

يرى كل من ثورانديك وهيجين (1986 م : 54)، و عودة (1993 م : 334) أن الاختبار الجيد لابد أن يتصف بعدة صفات و هي :

الموضوعية و الصدق و الثبات (و نعني الخصائص السيكومترية للاختبار) والمعايير و القابلية للاستعمال.

أما القابلية للاستعمال فتشمل مجموعة من العوامل هي :

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

بينما يذكر أبو علام (1987 م : 274) "بأن خاصيتي الثبات و الصدق من أهم الخصائص التي يجب أن تتوفرا في الاختبار كي يتم الوثوق و الاعتماد عليه و على نتائجه التي نحصل عليها من تطبيقنا له".

أوّلاً :الثبات

هناك أكثر من معنى للثبات فقد يعني الاستقرار أي تماثل نتائج أداة القياس عند تطبيقه أكثر من مرة تحت نفس الظروف، و قد يعني الاتساق أي اتساق نتائج الاختبار مع بعضها البعض أو الانسجام بين أجزاء الاختبار (عودة، 1993 م) (ثورانديك، هيجن، 1986 م).

و يتفق الباحث مع التعريف الذي ذكره إسماعيل (2007 م : 16) عن الثبات وهو "الحصول على نفس النتائج تقريبًا عندما يتم إعادة نفس الاختبار على نفس المجموعة و في نفس الظروف تقريبًا".

طرق حساب معامل الثبات :

1- طريقة إعادة الاختبار (Test-Retest) :

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

2- طريقة الصور المتكافئة (Parallel) :

نقوم بإعداد صورتين متكافئتين من الاختبار و يتم تطبيقها على عينة من الأفراد في الوقت نفسه ثم نقوم بإيجاد معامل الارتباط بينهما و يسمى معامل التكافؤ.

3- التجزئة النصفية (Split-Half) :

نقوم بتقسيم الاختبار إلى نصفين متكافئين إحصائيًا و بدون علم الأفراد المختبرين و بعد التطبيق نرصد نتائج النصف الأول على حدة و الثاني على حده ونوجد معامل الارتباط بينهما و نحصل بذلك على ثبات نصف الاختبار فأجد سبيرمان براون معادلة لتصحيح الثبات من النصف إلى الكل و هي معادلة رقم(11).

$$\text{معامل ثبات الاختبار} = \frac{2\,ر}{1 + ر} \quad \text{.....(11)}$$

(علام، 2006 م).

4- معادلة رولون(Rulon Formula) :

افترض رولون أن التباين الكلي للاختبار ناتج من جزأين :

- أحدهما ناتج عن الخطأ التجريبي.
- و الآخر ناتج عن الفروق الفردية.

و تأخذ معادلة رولون الصورة التالية و هي معادلة رقم (12) :

$ع^2ف$

معامل الثبات = 1- $\frac{ع^2 ف}{ر + 1}$ .. $ع^2ك$ .....(12)

(إسماعيل،2007 م).

حيث : $ع^2ف$ = تباين الفرق بين درجات نصفي الاختبار.

و $ع^2ك$ = تباين الاختبار ككل.

5- معادلة جتمان(Gutman Formula) :

يرى جتمان إمكانية إهمال شرط تساوي التباين بين نصفي الاختبار عند سبيرمان براون و خرج بالمعادلة التالية رقم (13) :

و تأخذ معادلة رولون الصورة التالية و هي معادلة رقم (13) :

معامل الثبات = 2 [1 - $\frac{ع^2 ف}{+ 1}$ $\frac{ع^2_1 + ع^2_2}{ع^2 ك}$ .....(13)

حيث : $ع^2_1$= تباين النصف الأول من الاختبار.

$ع^2_2$=تباين النصف الثاني من الاختبار.

$ع^2ك$ =تباين الاختبار ككل.(إسماعيل، 2007 م).

6- معادلة كيودر– ريتشاردسون :

تستخدم هذه المعادلة لحساب التجانس الكلي للاختبار و تقوم على اعتبار كل سؤال من الاختبار بندًا و جزءًا منه و نستخدمها لشمول الاتساق بين هذه الأجزاء (الاسئلة) و نسميه معامل الاتساق الداخلي (معامل التجانس)، و لها صيغتان :

الصيغة الأولى (صيغة 20) :

و هي الأكثر استخدامًا و يشترط لاستخدامها أن يقيس الاختبار بعد واحد و لا يكون الاختبار اختبار سرعة و درجة كل سؤال إما صفر او واحد و هي المعادلة رقم (14) :

ن ($ع^2$ ك – مج أ بز xأثث)

معامل الثبات = ------- [----------] .....(14)

ن – 1 $ع^2ك$

حيث : ن = عدد مفردات الاختبار.

$ع^2ك$ =التباين الكلي لدرجات الاختبار.

أبز : معامل السهولة.

أثث : معامل الصعوبة. (إسماعيل، 2007 م).

الصيغة الثانية (صيغة 21) :

و يشترط لاستخدامها بالإضافة إلى الشروط في صيغة 20 تساوي جميع مفردات المقياس في درجة الصعوبة و يكون متوسط درجة الصعوبة لجميع المفردات 0,5، و هذا الشرط أدى إلى شيوع استخدام صيغة 20 أكثر من صيغة 21، و صيغة 21 هي المعادلة رقم (15) :

ن م (ن - م)

معامل الثبات = ------- [1 - ---------] .....(15)

ن – 1 ن ع²ك

حيث :م = متوسط الدرجات الكلية،

و ن : عدد مفردات الاختبار (إسماعيل، 2007 م).

7- معادلة كرونباخ (α) (Cronbach) :

و تستخدم لحساب ثبات الاختبار من جزأين أو أجزاء متعددة و هي المعادلة رقم (16) :

ن مج ع² ن

معامل الثبات = ------- [1- ---------] .....(16)

ن – 1 ع²ك

حيث : ن = عدد الأجزاء التي ينقسم إليها الاختبار.

ع²ك =تباين درجات الاختبار ككل.

ع²ن =تباين المفردات (أي مفردة ن) من الاختبار.

مج ع²ن =مجموع تباينات الأجزاء (إسماعيل، 2007 م).

ملاحظة : عندما ن = 2 نحصل على معادلة جتمان.

ثانيًا :الصدق Validity

يعتبر أبو لبدة (1985 م : 233) "الصدق أهم خاصية من خواص الاختبار ويعني ذلك أن يقيس الاختبار بالفعل ما وضع لقياسه".

و الصدق يكون نسبيًا أي أنه لا يوجد اختبار صادق بشكل كامل و لا يوجد اختبار غير صادق، أيضًا الصدق يكون نوعيًا أي صادقًا للغرض الذي أعد من أجله ضمن العينة و الظروف المحيطة. (عودة، 1993 م).

الطرق المختلفة للتأكد من الصدق :

أوّلاً : الصدق الظاهري (صدق المحكمين) Face Validity :

و يقصد به إمكانية التعرف على ما يقيسه الاختبار من خلال العنوان و بالتالي فهو فحص سطحي و لا يعتد به في الأبحاث و يتم تحديد معامل الصدق الظاهري عن طريق صدق المحكمين. (إسماعيل، 2007 م).

ثانياً : صدق المحك Criterion-relatedValidity :

يرتبط هذا الصدق بمدى الارتباط بين درجات الاختبار المراد معرفة صدقه ومحك له صدق و ثبات موثوق بهما و مطبق على نفس العينة، و هو ينقسم إلى صدق تلازمي و صدق تنبؤي.

أ- الصدق التلازميCongcuentValidity :

و نقصدبه مدى تطابق و ارتباط نتائج تطبيق الاختبار و نتائج محك في نفس الوقت أي في الوقت الحاضر و ذلك بإيجاد معامل الارتباط بين الاختبار وبين المحك.

و كلما كانت قيمة معامل الارتباط (+ 0,5) أو أكثر دل على الصدق والأفضل أن يكون (+ 0,8) و كلما زاد عدد العينة كان أفضل. (إسماعيل، 2007 م).

ب- الصدق التنبؤي (المستقبلي) PredictiveValidity :

يشير الصدق التنبؤي إلى مدى نجاح الاختبار و قدرته على التنبؤ بأداء الفرد في المستقبل و يتم حسابه بإيجاد معامل الارتباط بين درجات الاختبار الحالي و درجات محك آخر بعد فترة زمنية مناسبة مثلاً كاختبار القدرات أو اختبارات القبول في الجامعات و المعدل التراكمي للطالب.

و يعتبر معامل الصدق التلازمي قيمة قصوى لمعامل الصدق التنبؤي. (إسماعيل، 2007 م).

رابعًا : صدق التكوين الافتراضي (صدق البناء) :

و ينقسم إلى ثلاثة أقسام :

أ- **التحليل المنطقي :**و ينقسم إلى ثلاثة أقسام :

1. **النظرية و البحوث التجريبية :** يتم تصميم الاختبار في ضوء نظرية معينة يتم التأكد من صحة فروض هذه النظرية بالبحوث التجريبية.
2. **التحليل المتعدد :**مقارنة نتائج الاختبار الحالي بنتائج البحوث و الدراسات السابقة التي تناولت نفس الموضوع و إذا وجد اتساق فيمل بينها كان ذلك دليلاً على الصدق.
3. **طريقة تطور الأعمار** تناسب هذه الطريقة بعض السمات و القدرات المرتبطة في نموها مع العمر و بخاصة الذكاء، و قد لا تتلائم مع باقي الصفات الإنسانية.

ب- **طريقة الارتباطات :**و ينقسم إلى ثلاثة أقسام :

1. **الصدق العاملي :FactorialValidity:**ذكر أبو لبدة (1985 م : 252) "أنه عبارة عن البحث عم 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 و أي اختبار آخر أو مجموعة اختبارات أخرى تقيس الشيء نفسه".
2. **المقارنة الطرفية (الاتساق الداخلي) :**يتم تطبيق الاختبار المراد حساب صدقه ثم نحسب الدرجة الكلية للاختبار، ثم نحصل على مجموعتين 27% أعلى و 27% أدنى، ثم نحسب الفروق بين المجموعتين المرتفعة و المنخفضة في درجات مفردات الاختبار عن طريق اختبار(ت) و يوجد احتمالين الأول وجود فروق ذات دلالة إحصائية بين المجموعتين و يكون الاختبار بذلك صادقًا لأنه استطاع التمييز بين المجموعتين، و الثاني عدم وجود فروق ذات دلالة إحصائية بين المجموعتين يكون الاختبار غير صادق لأنه لم يستطيع التمييز بين المجموعتين. (إسماعيل، 2007 م).
3. **الصدق التقاربي و التمييزي (التمايزي) :**إذا ارتبط الاختبار ارتباطًا عاليًا أو مرتفعًا بالمتغيرات التي يجب أن يرتبط بها نظريًا أو تتسق معه يسمى هذا صدق تقاربي.

أما إذا ارتبط الاختبار ارتباطًا منخفضًا بالمتغيرات التي لا يجب أن يرتبط بها نظريًا أو لا تتسق معه فهذا صدق تمييزي (تمايزي). (إسماعيل، 2007).

4-ا **لصدق 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 و نجاحهم في الاختبار البعدي و بفحص المفردات يتم الحكم بصعوبة المفردة عند فشل معظم الطلاب في التطبيق على حلها و بسهولة المفردة عند نجاح معظم الطلاب في التطبيقين على حلها، و يكون نجاح معظم الطلاب في الاختبار القبلي و فشلهم في الاختبار البعدي حالة شاذة و تشير إلى وجود خطأ ما في الصياغة أو خطأ التعليم أو كليهما معًا. (إسماعيل، 2007).

ثانيا :الخصائص السيكومترية.

❖ **معنى مصطلح سيكومتري(Psychometric) :**يتكون مصطلح سيكومتري من مقطعين هما :

- سيكو**(Psycho)**: وهي صفة مأخوذة من الكلمة (Psychic)وتعني نفسي، وهي كلمة تأخذ إحدى الاستعمالات الآتية :
  - كل ما يتعلق بظواهر روحية.
  - اسم عام لكل الظواهر التي يتكون منها موضوع أو مادة علم النفس؛ المتصل بالعقل أو الشخص، أو الذات.
  - مرادف مبهم لما هو نفسي المنشأ أولما هو وظيفي المنشأ.
- **متري أوقياسي(Metric)** : وهي لاحقة تتصل عمومًا وأكثر تخصيصًا، القياس المعتمد على وحدات المتر والغرام (دسوقس، 1990).

وبتتبع المعاني التي يدل عليها مصطلح الخصائص السيكومتريةPsychometric عند كثير ممن استخدموا المصطلح مثل : (فرج، 1980)، و(العلي، 1991)، و(إمريزق، 1992)، و(ظاظا، 2000)، و(سليمان، 2001)، و(النبهان، 2004)، يمكن رصد الدلالات التالية لهذا المصطلح :

- يقترن استخدام مصطلح الخصائص السيكومترية للاختبار بالموضوعات التي تتناول مواصفات أوخصائص الاختبار الجيد.
- تركز كثير من الدراسات والكتب على الصدق والثبات بشكل رئيس، للدلالة على الخصائص السيكومترية.
- تصنيف دراسات أخرى بعض الخصائص السيكومترية للفقرة مثل معاملات الصعوبة، ومعاملات التمييز.
- أيضًا خاصية أخرى للخصائص السيكومترية للفقرة؛ مثل فعالية المشتتات أوفعالية المموهات.
- تهمل بعض الدراسات مقاييس النزعة المركزية، ومقاييس التشتت عند تعرضها لدراسة الخصائص السيكومترية، بينما تثبته دراسات أخرى ولوبشكل غير مباشر أوبشكل مقصود.

❖ **تصنيفات الخصائص السيكومترية :**

يمكن تصنيف الخصائص السيكومترية للاختبار إلى ثلاثة أقسام رئيسية هي :

- **الخصائص السيكومترية التي تقدر من خلال الدرجة الكلية للاختبار :**

ويمكن التعرف عليها بشكل رئيسي من خلال تحليل درجات الطلاب الكلية عن الاختبار. ومن هذه الخصائص : مقاييس النزعة المركزية، ومقاييس التشتت، ومقاييس التماثل والاعتدالية.

- **الخصائص السيكومترية التي تقدر من خلال درجة الفقرة الاختبارية :**

ويمكن التعرف عليها من خلال تحليل درجات الطلاب عن فقرة محددة من فقرات الاختبار. ومن هذه الخصائص : صعوبة الفقرة، وتمييزها، وتباينها، وفعالية المشتتات لكل فقرة.

- **الخصائص السيكومترية المشتركة :**

ويقصد بها الخصائص التي يتم تقديرها من خلال درجة كل فقرةاختبارية، أومن خلال الدرجة الكلية للاختبار، أومن خلالهما معًا. ومن هذه الخصائص الصدق والثبات للاختبار وهما من أهم الخصائص السيكومترية التي يجب التركيز عليها عند البحث عن الخصائص السيكومترية. (غيث، 2007 : 36-37).

❖ **الخصائص السيكومترية لفقرات الاختبار في إطار القياس الكلاسيكي :**

عند بناء وتحليل الاختبارات لابد من تحديد بعض الشروط السيكومترية للاختبار، ويقصد بالشروط السيكومترية للاختبار تلك الخصائص الضرورية والمتعلقة بصعوبة الفقرة، وتمييز الفقرة، ومن أهم الخصائص السيكومترية لفقرات الاختبار في إطار القياس الكلاسيكي ما يلي

- **صعوبة الفقرة (Item Difficulty) :**

يعرف القياس الكلاسيكي مفهوم صعوبة الفقرة أنه "متوسط احتمالات الإجابة الصحيحة عن الفقرة لجميع أفراد المجموعة و التي تظهر ما إذا كانت الفقرة سهلة أو صعبة بالنسبة لأفراد المجموعة الذين يجيبون عن فقرات الاختبار"،.(عبد الحافظ، 1999 :81).

وقد أشار "دوران" المشار إليه في (عودة، 1993) بأن أية فقرة ضمن توزيع معاملات الصعوبة من (0,30) إلى (0,80) يمكن أن تكون مقبولة وينصح بالاحتفاظ بها، ويرى "هنرسن" المشار إليه في (1989، Allen & Yen) أنّه إذا كانت القدرة التمييزية للفقرات تتراوح بين (0,30) إلى (0,40) فإن مدى الصعوبة المقبول للفقرات يكون من (0,40) إلى (0,60).

القوة التمييزية للفقرة (Item Discrimination) :

يعرف القياس الكلاسيكي مفهوم القوة التمييزية للفقرة بأنه مدى فاعلية فقرة الاختبار في التمييز بين المستويات المختلفة للسمة المراد قياسها لدى الممتحنين (علام، 2001 : 216).

و تفصل مؤشرات تمييز الفقرة بين الممتحنين ذوي الدرجات المرتفعة والممتحنين ذوي الدرجات المنخفضة. و باستخدام هذه المؤشرات يستطيع الفرد أن يتعرف على الممتحنين الأقوياء و يمتلكون احتمالية أعلى للإجابة بطريقة صحيحة، و الضعفاء المنخفضة درجاتهم و يمتلكون احتمالية أقل للإجابة بطريقة صحيحة (2001، Crislip& Chin-Chance).

- **صدق القياس (MeasurementValidity) :**يعرف القياس الكلاسيكي مفهوم صدق القياس بأنه يتعلق بما يقيسه الاختبار و إلى أي مدى ينجح في قياسه. فالصدق يحدد قيمة الاختبار وصلاحيته في قياس ما وضع لقياسه (خطاب، 2001 : 159).
- **ثبات القياس (MeasurementReliability) :**يعرف القياس الكلاسيكي مفهوم ثبات القياس ؛ بأنه مدى اتساق درجات الاختبار من قياس إلى آخر، أو القدرة على قياس الظواهر لعدة مرات، والحصول على نفس النتائج أو قريبًا جدًا من بعضها. (761 : 1995، El-Korashy).

❖ **الخصائص السيكومترية لفقرات الاختبار في إطار نموذج راش :**

و من أهم الخصائص السيكومترية لفقرات الاختبار في إطار نموذج "راش" ما يلي :

- **صعوبة الفقرة (Item Difficulty) :**

يدل مفهوم صعوبة الفقرة في ضوء نموذج "راش" بمستوى القدرة (Θ) التي يكون عندها المنحنى المميز للفقرة أكثر انحذارًا، و بعبارة أخرى كلما زادت صعوبة الفقرة زادت فرصة قدرة الفرد على اجتياز الفقرة بنسبة (50%). (237، 1999 : Yen &Edwardson).

و يعرف مفهوم صعوبة الفقرة أيضًا بأنه "النقطة على متصل السمة الكامنة التي يُتوقع أن يكون احتمال الإجابة الصحيحة عن الفقرة دون تخمين تساوي (0,50). (علام، 2001 : 217).

و قد أشار هامبلتون و سومیناثان (1975، Hambelton&Swaminathan) إلى أن قيمة الصعوبة (b) تتراوح من (-∞) إلى (+∞) نظريًا ؛ إلا أنه عمليًا فقيمتها تتراوح بين (-3) إلى (+3).

- **قدرة الفرد Ability:**

و يرى هامبلتونو سومیناثان (Hambelton&Swaminathan ،1975) أن القدرة تتدرج من (-∞) إلى (+∞) ؛ إلا أنه نادرًا ما لوحظ تجاوزها للمدى من (-3) إلى (+3).

- **القوة التمييزية للفقرة (Item Discrimination) :**

يعرف مفهوم القوة التمييزية للفقرة في ضوء نموذج "راش" هو معدل التغير في احتمال الاستجابة الصحيحة للأفراد على الفقرة بالنسبة لمستوى القدرة، و تقدر هذه القوة بمعامل التمييز الذي يقصد به الميل النسبي للمنحنى المميز للفقرة على محور القدرة (كاظم، 1988 أ : 74).

- **صدق القياس (MeasurementValidity) :**

يعرف مفهوم صدق القياس في ضوء نموذج "راش"بأن تعرف الفقرات فيما بينها متغيرًا واحدًا ؛ أي أن فقرات الاختبار تتدرج من حيث صعوبتها بحيث تعرف متغيرًا واحدًا، كما يعني تدرج قدرات الأفراد على المتغير محددة تقديرات أدائهم على هذا الاختبار، و هذا يوضح أن كلاً من صعوبات الفقرات، و قدرات الأفراد تتدرج على متصل واحد يمثل متغيرًا واحدًا (المرجع السابق، 1988 أ : 98).

- **ثبات القياس (MeasurementReliability) :**

يعرف مفهوم ثبات القياس في ضوء نموذج "راش" بتحقق ما بقى من مطالب الموضوعية في القياس ؛ عندما تستخدم أداة القياس التي أنشئت باستخدام هذا النموذج، حيث يتحقق :

- استقلال القياس عن الاختبار المستخدم.
- استقلال القياس عن مجموعة الأفراد المؤدية للاختبار.

أي أن استقلالية القياس و تحرره الذي يوفره نموذج "راش" تتيح الفرضة لثبات القياس بحيث لا يختلف القياس (سواء لقدرة الفرد أو صعوبة الفقرة) باختلاف عينة التدريج و باختلاف الاختبار المستخدم لقياس قدرة ما (كاظم، 1988 أ : 99) ؛ (اسماعيل و آخرون، 1994 : 131).

كما أن نموذج "راش" لا يعطي فقط معامل ثبات الاختبار الكلي، و لكنه يعطي أيضًا معامل ثبات لكل فرد و فقرة (El-Korashy ،1995 : 760).

خلاصة الفصل :

إن أهمية التعرف على الاختبارات التحصيلية وطريقة بنائها ومعرفة الخصائص السيكومترية للاختبارات التحصيلية يؤدي الى تقويم النتاجات التعليمية بشكل فعال وعلى إثر انعدام الإلمام بهذه الخصائص تواجه المدرسة الجزائرية مشاكل عصيبة تزداد خطورة و انتشارا مع مرور الزّمن، فضعف المستوى أصبح السّمة الغالبة عند أغلبية المتمدرسين ،ويتضح ذلك من نتائج الامتحانات المحدّدة لمستقبل التلاميذ والتي بنبغي أن يراعى في بنائها على الخصائص السيكومترية للاختبار الجيد.

الفصل الخامس :عرض وتفسير النتائج

أولا :منهج وإجراءات الدراسة .

تمهيد :

يشكل الفصل الخامس مدخلا للدراسة الميدانية بحيث يتضمن هذا الفصل وصفا للطريقة والإجراءات التي اتبعها الطالب الباحث في تحديد العينة واستخدام أداة الدراسة وخطوات التحقق من صدق الأداة وثباتها، إضافة إلى وصف تصميم الدراسة والأساليب لإحصائية المتبعة في تحليل البيانات .

- **منهج الدراسة :**

استخدم الطالب المنهج الوصفي التحليلي منهجاً للدراسة وذلك لملائمته لطبيعة الدراسة، وهو أسلوب من أساليب البحث العلمي حيث يعتمد على دراسة الواقع أو الظاهرة كما توجد في الواقع ، ويهتم بوصفها وصفًا دقيقًا ويعبر عنها تعبيرًا كيفيًا ؛ من حيث وصف الظاهرة وتوضيح خصائصها .

- **الأساليب والبرامج الإحصائية المستخدمة :**

تم تحديد الأساليب والبرامج الإحصائية المستخدمة بالدراسة في ضوء مشكلة الدراسة وأهدافها وذلك للإجابة على تساؤلات الدراسة ، وتتمثل الأساليب والبرامج الإحصائية المستخدمة في هذه الدراسة ،البرنامج الإحصائي للرزمة الإحصائية للعلوم الاجتماعية (SPSS)وذلك لحسابةمعامل الفا كرونباخ ،لقياس الثبات ، معاملات الارتباط بين الفقرات، اختبار ت ، اختبار (ت) لدراسة الفروق ، إضافة إلى استخراج معامل السهولة والصعوبة ومعامل التمييز للفقرات ،والمتوسط الحسابي النسب المئوية .....الخ.

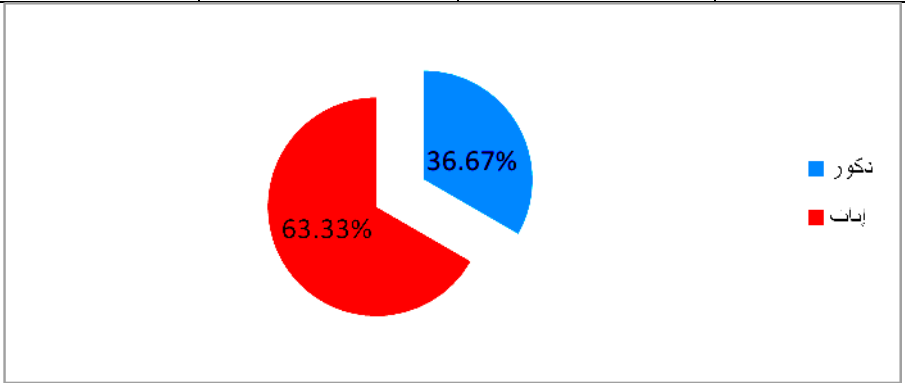
- **عينة الدراسة :**

تم اختيار عينة الدراسة بطريقة قصدية ، وتكونت عينة البحث الحالي من مستوى السنة الثالثة ثانوي شعبة أداب وفلسفة بثانوية بوعمامة للعام الدراسي 2015/2014 وقدر عددهم 30 تلميذ.

3-1- مواصفات عينة الدراسة:

الجدول رقم (01)يوضح مواصفات عينة الدراسة تبعا لمتغير الجنس :

المتغير	التصنيف	العدد	النسبة المئوية %
الجنس	ذكور	11	36.67%
	إناث	19	63.33%



يتضح من الجدول رقم (01)والشكل أن نسبة الذكور قدرت ب (36.67 %) من حجم العينة مقارنة بنسبة الإناث والتي بلغت (63.33 %) وهذا ما يوضح أن نسبة الإناث هي أعلى من نسبة الذكور .

- **وصف الأداة :**

تم اعتماد اختبار تحصيلي في مادة اللغة العربية لتلاميذ الثالثة ثانوي شعبة آداب وفلسفة،الذي تم إجراؤه في الفصل الأول خلال السنة الدراسية 2015/2014 والذي يتكون من ثلاثة أسئلة مقسمة كالأتي:

السؤال الأول:ويحتوي على ستة أسئلة '(08 نقاط)

السؤال الثاني: ويحتوي على ستة أسئلة'(09 نقاط)

السؤال الثالث: ويحتوي على سؤالين'(03 نقاط)

الجدول رقم (02) يوضح وصف الاختبار مع التنقيط

الاسئلة	عدد الفقرات	عدد الدرجات/نقاط	مجموع الدرجات/نقاط
السؤال الأول	الفقرة 01	2	08نقاط
	الفقرة02	1	
	الفقرة03	1	
	الفقرة04	1	
	الفقرة05	1	
	الفقرة06	2	
السؤال الثاني	الفقرة 01	3	09 نقاط
	الفقرة02	1	
	الفقرة03	1	
	الفقرة04	1	

	الفقرة05	1	
	الفقرة06	2	
السؤال الثالث	الفقرة 01	1	03 نقاط
	الفقرة02	2	
المجموع	14 فقرة	20 درجة	20 نقطة

يتضح من خلال الجدول أن أسئلة الاختبار ألتحصيلي في مادة اللغة العربية ذات طابع مقالي، وبالاستعانة من أساتذة من نفس التخصص، تم ضبط جدول مواصفات هذا الاختبار بأنها مناسبة للمرحلة العمرية وطبيعة المادة ،كما هو موضح في الجدول رقم (04) :

الجدول رقم (03)يوضح جدول مواصفات الاختبار.

الأسئلة	الفهم	التحليل	التركيب	التقويم
السؤال الأول	% 30	% 40	%20	%20
السؤال الثاني	%30	%40	%20	%10
السؤال الثالث	%10	%20	%30	%40

الجدول أعلام يوضح مواصفات الاختبار بناءا على المستويات المهارية التي يقيسها كل سؤال ، وقد تم توزيع النسب المئوية بالتنسيق مع أستاذ المادة والموافقة عليها قبل بدء عملية التحليل.

5-إجراءات الدراسة :

تم إعداد الدراسة وفق الإجراءات الآتية :

1- تحديد المستوى الدراسي والتخصص والمادة الدراسية(سنة ثالثة ثانوي. شعبة آداب وفلسفة، واختبار مادة اللغة العربية).

2- الحصول على أسئلة الاختبار ألتحصيلي وعلى أوراق التصحيح المنجزة خلال الفصل الأول.

3- إدخال البيانات في جهاز الكمبيوتر ومعالجتها إحصائيا باستخدام الحزمة الإحصائية للعلوم الاجتماعية(SPSS)

4- استخراج النتائج وتحليلها ومناقشتها ومقارنتها مع الدراسات السابقة واقتراح التوصيات المناسبة.

5-كما قام الطالب العمليات الإجرائية للتحليل :

دراسة جميع أوراق الإجابة للمتعلمين الذين طبق عليهم الاختبار.

ترتيب أوراق الإجابات طبقاً لمجموع درجات كل منهم ترتيباً تنازلياً أو تصاعدياً .

تحدد المجموعتان الطرفيتان (العليا – الدنيا) و تترك أوراق الإجابات الباقية جانباً .

إعداد جدول يسجل فيه رقم السؤال وتكرار الإجابات الصحيحة للمجموعتين وعمود لحساب معامل السهولة والصعوبة ومعامل التمييز.
تحليل أسئلة الاختبار من حيث معامل السهولة ومعامل الصعوبة ومعامل التمييز.

ثانيا :عرض ومناقشة نتائج الدراسة :

يتناول هذا الجزء من الجانب الميداني عرضا لنتائج الدراسة وتفسيرها ، وقد قام الطالب فيه بعرض نتائج الخصائص السيكومترية لاختبار مادة اللغة العربية وفق القياس الكلاسيكي ،ثم مناقشة نتائج الدراسة الحالية وتفسيرها، وفقًا للتساؤلات الآتية:

السؤال الأول :

ما قيم مؤشرات صدق وثبات اختبار مادة اللغة العربية وفق القياس الكلاسيكي ؟

أولا : صدق الأداة :

يمكن التحقق من صدق الاختبار بعدة طرق ومنها:
الرجوع إلى جدول المواصفات ،وتحليل المستويات المهارية وهذا بهدف التأكد من مطابقة نسبة الأهداف السلوكية مع أسئلة الاختبار ألتحصيلي،وبناءا على الجدول رقم (04) ،يمكن القول أن:
باعتبار أن المادة الدراسية –لغة عربية- مادة أساسية ومحددة لمستوى التلاميذ في شعبة أدأب وفلسفة،يفترض أن يكون الاختبار ألتحصيلي يركز على الفهم أكثر في السؤال الاول نسبة 50%"باعتباره يتعلق بالبناء الفكري للنص ، ثم التحليل 30 %ثم التركيب 20% ثم وبالنسبة للسؤال الثاني المتعلق بالبناء اللغوي ،فيجب التركيز أيضا الفهم بنسبة 50%ثم على التحليل 30%ثم التركيب 20% ،أما السؤال الثالث المتعلق بالبناء اللغوي فمن المفروض التركيز على الفهم 30% ثم التحليل 30% ثم التقويم 40%
فما يمكن قوله أن الخلل الحاصل في تقسيم النسب المئوية للمستويات المهارية يؤثر على سلم التنقيط وعلى أداء التلاميذ وبالتالي على طول الاختبار ألتحصيلي و عدم الدقة في إجراءات تصحيح الاختبار .

ب - **صدق المحكمين:**

للتحقق من صدق الأداة ،تم تقديم أسئلة الاختبار ألتحصيلي لمادة اللغة العربية على مجموعة من المحكمين من أساتذة التعليم الثانوي ،للتأكد من صدق المضمون ،وبهدف إبداء أرائهم للتأكد من أن الفقرات التي شكلت الاختبار تمثل فعلا السلوك المراد قياسه ،وبهدف إطلاعهم على موضوع الدراسة وأهدافها وتقديم الإشكالية وصياغة الفرضيات ،وسنعرضهم في الجدول الآتي:

الجدول رقم (04)يوضح صدق المحكمين.

المحكم	الدرجة	الخبرة المهنية	التخصص

1	أستاذة تعليم ثانوي (ثانوية بوعمامة)	07سنوات	ماستر أدب عربي
2	أستاذ تعليم ثانوي (ثانوية محمد بوضياف)	12 سنة	ليسانس أدب عربي
3	أستاذة تعليم ثانوي (ثانويةالسايح محمد)	10سنوات	ليسانس أدب عربي
4	أستاذة تعليم ثانوي (ثانوية السايح محمد)	08سنوات	ماستر أدب عربي
5	أستاذ تعليم ثانوي (ثانوية بوعمامة)	10سنوات	ليسانس أدب عربي
6	أستاذ تعليم ثانوي (ثانوية بوعمامة)	06 سنوات	ماستر أدب عربي

يوضح الجدول أعلاه عينة من أساتذة التعليم الثانوي الذين لديهم خبرة مهنية لا بأس بها وبمعدل أكثر من 10 سنوات عمل ومؤهل علمي نسبة50 %ليسانس و50 شهادة ماستر في الأدب العربي ، ومن حيث وضوح وملائمة الفقرات وشموليتها وقصد مراجعته، و تحديد سلامة الصياغة اللغوية، وعدد الفقرات هل هي كافية أو غير كافية ،ووضوح العبارات من حيث المعنى في كل عبارة ،و الحكم على مدى ملائمة العبارات وقد كانت النتيجة كما يلي :

الجدول رقم (05)يوضح أراء المحكمين حول بناء الاختبار.

المعايير/ الأسئلة	السؤال 1			السؤال 2			السؤال 3		
معايير صياغة السؤال الجيد	نعم	لا	تعدل	نعم	لا	تعدل	نعم	لا	تعدل
-هل المحتوى المراد قياسه مغطى في المنهاج ؟	*			*			*		
-هل الأهداف المراد قياسها موجودة في المنهاج ؟	*			*			*		
-هل السؤال واضح ومستوى لغته ملائم للتلاميذ؟	*			*			*		
-هل إجابة السؤال محددة ؟			*	*			*		
-هل صياغة السؤال جديدة ومختلفة عن نص الكتاب ؟		*		*			*		
-هل طريقة عرض السؤال غير ملائمة ؟		*		*			*		
-هل المعلومات المتضمنة في السؤال ضرورية ؟	*			*			*		
-هل هناك ترتيب منطقي في الفقرات الفرعية للسؤال ؟	*			*			*		
- هل تسلسلها بحسب مستوى الأهداف ، أو التدرج في الصعوبة؟ .	*			*			*		
-هل ترتبط الفقرات الفرعية جميعها بالموضوع ؟	*					*	*		
-هل تعليمات السؤال واضحة ؟	*			*			*		
-هل تم تحديد العلامة الكلية والفرعية لكل سؤال وفقرة ؟	*					*	*		
-هل تم تجنب الكلمات والمصطلحات التي تحمل اكثر من معنى ؟	*			*			*		
-هل تم تسليط الضوء على المعلومات الهامة في السؤال بطريقة واضحة ؟	*			*	*				*
-هل المدة الزمنية المخصصة للإجابة كافية ؟		*		*					*

بعد فحص فقرات الاختبار في مادة اللغة العربية ،من طرف المحكمين من حيث الوضوح والدقة وسهولة أو صعوبة التعبير وطباعة الاختبار ،تم الأخذ برأي لأغلبية حيث

كانت نسبة اتفاق المحكمين80% و هي نسبة تعد معيارا مقبولا عند كثير من الباحثين ،مما يدل أن الاختبار مناسب للمستوى وللفئة العمرية.

ثانيا ثبات الأداة :

لقد تم استخراج معامل الثبات باستخدام معادلة ألفا كرونباج (Cronbacha Alph) ، كما هو موضح في الجدول رقم (06) :

الجدول رقم (06) يوضح ثبات الاختبار.

عدد الفقرات	معامل الثبات بطريقّة ألفا كرونباج
14	0.117

يتضح من الجدول رقم (04) أن ثبات الاختبار قد بلغ (0.117) فهو ثبات غير عالي مما يجعله غير قابلا ومناسبا لأغراض البحث .

1-الصدق الذاتي : تم حساب الصدق الذات لأداة الدراسة عن طريق حساب الجذر التربيعي لمعامل الثبات الخاص بفقرات الاختبار حسب ما هو مبين في الجدول أدناه:

جدول **رقم (07) يوضح القيم الخاصة بالصدق الذاتي للاختبار**.

معامل الثبات	الصدق الذاتي	عدد الفقرات
0.117	0.34	14

يتضح من الجدول رقم (5) أن معامل الصدق الذاتي للاختبار بلغ (0.34) في حين أنه إذا كان محصورا بين (0.50 – 1) يعتبر الاختبار صادقا وعليه يمكن القول أن الاختبار غير صادق ويدل بوضوح على عدم تمتع الاختبار بدرجة عالية من الصدق الذاتي.

- **السؤال الثاني :ما قيم مؤشرات السهولة و الصعوبة لفقرات اختبار مادة اللغة العربية وفق القياس الكلاسيكي ؟.**

<u>أولا :حساب معامل الصعوبة والسهولة للفقرات.</u>

- بعد الحصول على جميع أوراق الإجابة للمتعلمين الذين طبق عليهم الاختبار. تم ترتيب أوراق الإجابات طبقاً لمجموع درجات كل منهم ترتيباً تنازلياً أو تصاعدياً يتم تحديد المجموعتان الطرفيتان (العليا – الدنيا) كما يلي:

الجدول رقم (08)يوضح تحديد المجموعتان(العليا – الدنيا).

المجموعة العليا	المجموعة الدنيا
من التلميذ رقم 01 الى التلميذ رقم 15	من التلميذ رقم 16 الى التلميذ رقم 30

بعد تقسيم العينة إلى مجموعتين-العليا،الدنيا- يتم تحديد عدد الإجابات الصحيحة للمجموعتين لحساب معامل السهولة والصعوبة .

الجدول رقم (09) يوضح تكرار الاجابات الصحيحة للمجموعتين

رقم السؤال	تكرار الإجابات	مجوع الإجابتين
		الصحيحة والخاطئة

	الصحيحة	%	الخاطئة	%	عدد	%
1	104	49.52%	76	49.52%	180	18.10%
2	84	40.00%	96	40.00%	180	22.86%
3	22	10.48%	38	10.48%	60	9.05%
مج	210	100%	210	100%	420	50%

- يعتمد معامل الصعوبة المطلوب على الغرض من الاختبار ، وفي الاختبارات التحصيلية العادية فإن أفضل معامل صعوبة للسؤال أو الفقرة هو 50% وما حولها، ويمكن حساب معامل الصعوبة للأسئلة المقالية باستخدام المعادلة التالية :

$$\text{معامل الصعوبة} = \frac{\text{مجموع الدرجات المحصلة على السؤال}}{\text{عدد الطلاب} \times \text{درجة السؤال}}$$

الجدول رقم (10) يوضح معامل الصعوبة لكل فقرات الأسئلة.

الفقرات	م الصعوبة	م السهولة
فقرة1	0.27	0.73
فقرة2	0.63	0.37
فقرة3	0.67	0.33
فقرة4	0.60	0.40
فقرة5	0.43	0.57
فقرة6	0.30	0.70
فقرة7	0.19	0.81
فقرة8	0.60	0.40
فقرة9	0.43	0.57
فقرة1	0.47	0.53
فقرة11	0.37	0.63
فقرة12	0.18	0.82
فقرة13	0.43	0.57
فقرة14	0.15	0.85

يوضح الجدول أعلاه معاملات الصعوبة لمفردات الاختبار ألتحصيلي التي تراوحت ما بين (0.15 – 0.67) وهي مقبولة باعتبار أنه :

يعتبر السؤال (المفردة) مقبولا إذا تراوحت قيمة معامل الصعوبة له بين (0.15 – 0.85) (أبو جلالة ، 1999: 221) ، كون المفردة التي يقل معامل الصعوبة لها عن 0.15 تكون شديدة الصعوبة ، والمفردة التي يزيد معامل الصعوبة لها عن 0.85 تكون شديدة السهولة .

الجدول رقم (11) يوضح معامل الصعوبة لكل الأسئلة.

الاسئلة	السؤال الاول	السؤال الثاني	السؤال الثالث
مجموع الإجابات الصحيحة	104	84	22
مجموع الإجابات الخاطئة	106	126	188
معامل صعوبة السؤال	0.43	0.31	0.24
معامل سهولة السؤال	0.57	0.69	0.76

يوضح الجدول أعلاه معاملات الصعوبة لأسئلة الاختبار ألتحصيلي التي تراوحت ما بين (0.24 – 0.43) ويعتبر السؤال (المفردة) مقبولا إذا تراوحت قيمة معامل الصعوبة له بين (0.15 – 0.85) (أبو جلالة ، 1999: 221) ، وعليه فأسئلة الاختبار ألتحصيلي مقبولة وهي في متناول التلاميذ ولا يجدون صعوبة في الإجابة عليها

ثانيا:حساب معامل التمييز.

تم حساب مؤشرات التمييز لفقرات الاختبار، باستخدام طريقة المقارنة الطرفية، حيث تم ترتيب علامات التلاميذ الكلية ترتيبا تنازليا وتحديد 27% من المجموعة العليا في مقابل 27% من المجموعة الدنيا ،من التلاميذ الذين أجابوا إجابة صحيحة في كل مجموعة عن كل مفردة على حدة، وذلك بطرح متوسط كل فقرة من فقرات الاختبار للمجموعة الدنيا من متوسط كل فقرة من فقرات الاختبار للمجموعة العليا ، ثم تطبيق المعادلة التالية :

معامل التمييز = مجموع الدرجات في الفئة العليا – مجموع الدرجات في الفئة الدنيا

الدرجة الكلية للسؤال × عدد أفراد إحدى المجموعتين .

والجداول الآتية توضح مؤشرات تمييز فقرات الاختبار .

جدول رقم (12)يوضح عدد الاجابات الصحيحة في الفئة العليا.

عدد الاجابات الصحيحة في الفئة العليا														
س1						س2				س3		مج
فقرة 1	فقرة 2	فقرة 3	فقرة 4	فقرة 5	فقرة 6	فقرة 7	فقرة 8	فقرة 9	فقرة 10	فقرة 11	فقرة 12	فقرة 13	فقرة 14	15 فقرة
8	7	7	6	4	8	6	5	9	7	6	6	5	6	90
2	1	1	1	1	2	3	1	1	1	1	2	1	2	20

من خلال الجدول الإحصائي يمكن القول إن مجموع الإجابات الصحيحة للفئة العليا بلغ 90 إجابة صحيحة على العينة التي بلغ عددها 10 تلاميذ من الفئة الدنيا

جدول رقم(13) يوضح عدد الإجابات الصحيحة في الفئة العليا.

عدد الاجابات الصحيحة في الفئة الدنيا														
س1						س2				س3		مج
فقرة 1	فقرة 2	فقرة 3	فقرة 4	فقرة 5	فقرة 6	فقرة 7	فقرة 8	فقرة 9	فقرة 10	فقرة 11	فقرة 12	فقرة 13	فقرة 14	15 فقرة
4	3	6	4	2	5	3	4	5	4	5	4	3	4	56

2	1	1	1	2	1	1	3	1	1	1	2	1	2	1	20

من خلال الجدول الإحصائي يمكن القول إن مجموع الإجابات الصحيحة للفئة الدنيا بلغ 56 اجابة صحيحة على العينة التي بلغ عددها 10 تلاميذ من الفئة الدنيا

جدول رقم (14)يوضح معامل التمييز لكل الفقرات.

معامل التمييز للفقرات														
س1							س2						س3	
فقرة 1	فقرة 2	فقرة 3	فقرة 4	فقرة 5	فقرة 6	فقرة 7	فقرة 8	فقرة 9	فقرة 10	فقرة 11	فقرة 12	فقرة 13	فقرة 14
0.59	0.29	0.29	0.29	0.29	0.59	0.88	0.29	0.29	0.29	0.29	0.59	0.29	0.59

يتضح من خلال الجدول أن:

الفقرات (13.11.10.9.8.5.4.3.2) بلغ معامل التمييز أقل من (0.30)يعني أنه يمكن تعديلها (أبو جلالة ، 1999: 221) ،والملاحظ أن معامل التمييز أكثر من (صفر) باتجاه موجب ، فإن هذا **يعني أن عدد من أجابوا على الفقرة إجابة صحيحة من الفئة العليا أكبر من عدد الذين أجابوا عليها إجابة صحيحة من الفئة الدنيا ، أي أن تمييز الفقرة تميز موجب وهذا ما دل عليه الاختبار ألتحصيلي** ،كما يلاحظ أنه **لا توجد فقرة معامل تمييزها أقل من(0.19**).

وبالنسبة لمعامل التمييز من (0.29- 0.30) فتعدل الفقرة ، وهذا ما دل عليه الاختبار ألتحصيلي والفقرات المعنية هي : (13.11.10.9.8.5.4.3.2) وبالتالي يجب تعديل الفقرات وإذا كان معامل التمييز أكثر من (0.30) ، فتقبل الفقرة (أبو جلالة ، 1999: 221) ، وهذا ما دل عليه الاختبار ألتحصيلي و**الفقرات المعنية هي : (14.12.7.6.1.) وبالتالي تقبل الفقرات** .

السؤال الثالث:ما درجة تحقيق بيانات اختبار مادة اللغة العربية لافتراضات نموذج راش ؟

- لمعرفة ما مدى تحقيق بيانات اختبار مادة اللغة العربية لافتراضات نموذج راش نقوم بتطبيق الأسلوب الإحصائي كا2 وعن طريق ادخال البيانات الى SPSS للاجابة عن فرضية البحث:
- H0 : إن بيانات اختبار مادة اللغة العربية غير مناسبة لافتراضات نموذج راش
- H1 : إن بيانات اختبار مادة اللغة العربية مناسبة لافتراضات نموذج راش
- نحصل على البيانات الاتية في الجدول أدناه:

جدول رقم (15)يوضح نتائج الأسلوب الإحصائي كا2 .

الفقرات	Khi-deux	درجة الحرية	قيمةSig
الفقرة 1	,133[a]	1	.715
الفقرة 2	2,133[a]	1	.144

الفقرة 3	3,333[a]	1	.068
الفقرة 4	1,200[a]	1	.273
الفقرة 5	,533[a]	1	.465
الفقرة 6	1,200[a]	1	.273
الفقرة 7	,533[a]	1	.465
الفقرة 8	1,200[a]	1	.273
الفقرة 9	,533[a]	1	.465
الفقرة 10	,133[a]	1	.715
الفقرة 11	2,133[a]	1	.144
الفقرة 12	2,133[a]	1	.144
الفقرة 13	,533[a]	1	.465
الفقرة 14	4,800[a]	1	.028

يتضح من خلال الجدول الإحصائي أن :

أن قيمة sig المعنوية لكل الفقرات أكبر من (0.05) ،إذن نقبل الفرض الصفري ونرفض فرض البحث ونقول **إن بيانات اختبار مادة اللغة العربية غير مناسبة لافتراضات نموذج راش**

السؤال الرابع :هل توجد فروق دالة إحصائيا بين مرتفعي ومنخفضي التحصيل في التحصيل الدراسي ؟

لدراسة الفروق بين مرتفعي ومنخفضي التحصيل نقوم بتطبيق اختبار(ت).

- الفرضية الصفرية (**H0**): **لا توجد فروق** دالة إحصائيا عند مستوى الدلالة (0.05) في التحصيل الدراسي بين مرتفعي ومنخفضي التحصيل في التحصيل الدراسي.
- الفرضية البديلة (**H1**): **توجد فروق** دالة إحصائيا عند مستوى الدلالة (0.05) في التحصيل الدراسي بين مرتفعي ومنخفضي التحصيل في التحصيل الدراسي.

وعن طريق إدخال البيانات إلى برنامج spss وحساب اختبار (ت) فيspss نحصل على النتائج الاتية في الجدول أدناه:

الجدول رقم (18)يوضح الفروق بين مرتفعي ومنخفضي التحصيل .

فقرات/اسليب	مستوى	العينة	المتوسط الحسابي	الانحراف المعياري	اختبار ت	درجة الحرية	قيمةSig
ITEMS1	0-9	12	.3333	.49237	.14213	28	.077
	10 +	18	.6667	.48507	.11433		
ITEM2	0-9	12	.5000	.52223	.15076	23.480	.080
	10 +	18	.7222	.46089	.10863		
ITEM3	0-9	12	.6667	.49237	.14213	28	.230
	10 +	18	.6667	.48507	.11433	21.618	.245
ITEM4	0-9	12	.6667	.49237	.14213	28	1.000
	10 +	18	.5556	.51131	.12052	23.480	1.000
ITEM5	0-9	12	.2500	.45227	.13056	28	.559
	10 +	18	.5556	.51131	.12052	24.357	.557

ITEM6	0-9	12	.5000	.52223	.15076	28	.105
	10 +	18	.6667	.48507	.11433	25.672	.098
ITEM7	0-9	12	.4167	.51493	.14865	28	.379
	10 +	18	.6667	.48507	.11433	22.480	.388
ITEM8	0-9	12	.4167	.51493	.14865	28	.188
	10 +	18	.7222	.46089	.10863	22.720	.196
ITEM9	0-9	12	.3333	.49237	.14213	28	.101
	10 +	18	.5000	.51450	.12127	21.854	.111
ITEM10	0-9	12	.4167	.51493	.14865	28	.384
	10 +	18	.5000	.51450	.12127	24.458	.381
ITEM11	0-9	12	.3333	.49237	.14213	28	.667
	10 +	18	.3889	.50163	.11824	23.717	.668
ITEM12	0-9	12	.2500	.45227	.13056	28	.767
	10 +	18	.4444	.51131	.12052	24.042	.766
ITEM13	0-9	12	.4167	.51493	.14865	28	.295
	10 +	18	.4444	.51131	.12052	25.672	.284
ITEM14	0-9	12	.1667	.38925	.11237	28	.885
	10 +	18	.3889	.50163	.11824	23.613	.886
Sex	0-9	12	1.50	.522	.151	28	.206
	10 +	18	1.67	.485	.114	27.238	.184

يتضح من خلال الجدول الإحصائي أن :

أن قيمة sig المعنوية لكل الفقرات أكبر من (0.05) ،إذن نقبل القرض الصفري ونرفض فرض البحث ونقول **لا توجد فروق بين مرتفعي ومنخفضي التحصيل في التحصيل الدراسي .**

السؤال الخامس :هل توجد فروق دالة إحصائيا بين الذكور والإناث في التحصيل الدراسي ؟

لدراسة الفروق بين الذكور والإناث في التحصيل نقوم بتطبيق اختبار(ت).

- ❖ الفرضية الصفرية (**H0**): **لا توجد فروق** دالة إحصائيا عند مستوى الدلالة (0.05) في التحصيل الدراسي بين الذكور والإناث.
- ❖ الفرضية البديلة (**H1**): **توجد فروق** دالة إحصائيا عند مستوى الدلالة (0.05) في التحصيل الدراسي بين الذكور والإناث.

- وعن طريق إدخال البيانات إلى برنامج spss نحصل على الجدول الأتي:

- جدول رقم (17) يوضح الفروق بين الجنسين في التحصيل .

فقرات/أسليب	الجنس	العينة	المتوسط الحسابي	الانحراف المعياري	اختبار ت	درجة الحرية	قيمةsig
الفقرة1	ذكر	12	.5833	.51493	.434	28	.667
	أنثى	18	.5000	.51450	.434	23.717	.668
الفقرة2	ذكر	12	.6667	.49237	.299	28	.767
	أنثى	18	.6111	.50163	.300	24.042	.766
الفقرة3	ذكر	12	.7500	.45227	.772	28	.447
	أنثى	18	.6111	.50163	.789	25.391	.438
الفقرة4	ذكر	12	.6667	.49237	.592	28	.559

	أنثى	18	.5556	.51131	.596	24.357	.557
الفقرة5	**ذكر**	**12**	**.1667**	**.38925**	**-2.588**	**28**	**.015**
	أنثى	**18**	**.6111**	**.50163**	**-2.725**	**27.238**	**.011**
الفقرة6	ذكر	12	.5833	.51493	-.147	28	.884
	أنثى	18	.6111	.50163	-.146	23.290	.885
الفقرة7	ذكر	12	.6667	.49237	.884	28	.384
	أنثى	18	.5000	.51450	.892	24.458	.381
الفقرة8	ذكر	12	.4167	.51493	-1.698	28	.101
	أنثى	18	.7222	.46089	-1.660	21.854	.111
الفقرة9	**ذكر**	**12**	**.1667**	**.38925**	**-2.588**	**28**	**.015**
	أنثى	**18**	**.6111**	**.50163**	**-2.725**	**27.238**	**.011**
الفقرة10	ذكر	12	.3333	.49237	-1.183	28	.247
	أنثى	18	.5556	.51131	-1.192	24.357	.245
الفقرة11	ذكر	12	.3333	.49237	-.299	28	.767
	أنثى	18	.3889	.50163	-.300	24.042	.766
الفقرة12	ذكر	12	.2500	.45227	-1.067	28	.295
	أنثى	18	.4444	.51131	-1.094	25.672	.284
الفقرة13	ذكر	12	.5000	.52223	.585	28	.563
	أنثى	18	.3889	.50163	.580	23.051	.568
الفقرة14	ذكر	12	.1667	.38925	-1.294	28	.206
	أنثى	18	.3889	.50163	-1.362	27.238	.184

يتضح من خلال الجدول الإحصائي أن :

1 - قيمة sig المعنوية للفقرتين (9) و(.5) أصغر من (0.05) ،إذن نرفض القرض الصفري ونقبل فرض البحث ونقول توجد فروق بين الذكور والإناث في التحصيل الدراسي

2 - قيمة sig المعنوية للفقرات (13 .11..10.8.7.6.4.3.2.1) أكبر من (0.05) ،إذن نقبل القرض الصفري ونرفض فرض البحث ونقول لا توجد فروق بين الذكور والإناث.في التحصيل الدراسي .

<u>وعليه يمكن القول أنه لا توجد فروق بين الذكور والإناث.في التحصيل الدراسي .</u>

❖ عرض ومناقشة نتائج الفرضيات :

❖ كمناقشة عامة للفرضيات،يمكن القول أن الدراسة الميدانية هدفت إلى معرفة الخصائص السيكومترية لاختبار مادة اللغة العربية وفق مؤشرات القياس الكلاسيكي ونموذج راش لمستوى السنة الثالثة ثانوي شعبة أدأب وفلسفة , والكشف عن مدى مطابقة الاختبار التحصيلي للفصل الأول لنموذج راش . وقد تم التحقق من الخصائص السيكومترية حسب

النظرية التقليدية - الصدق , الثبات ,التمييز -. كما تم تحليل الاختبار حسب النظرية الحديثة في القياس (نموذج راش) عن طريق البرنامج الإحصائي (spss) .

❖ حيث بينت نتائج الدراسة أنه:

1- لايتمتع الاختبار التحصيلي لمادة اللغة العربية بدرجة عالية من الصدق والثبات باستخدام النظرية التقليدية ,مما يعني أن فقرات الاختبار التحصيلي غير صالحة لقياس مستوى التلاميذ وبالتالي عدم تحقق الفرضية الاولى.

2- إن معاملات الصعوبة لأغلب فقرات أسئلة الاختبار ألتحصيلي التي لم تتعدى (0.50) وهي مقبولة،تدل على أن التلاميذ لايجدون صعوبة في الاجابة عن أسئلة الاختبار،فكما ذكر(أبو جلالة ، 1999ص 221) ، أن قيمة معامل الصعوبة إذا تراوحت ما بين (0.24 – 0.43) يعتبر السؤال مقبولا ، وعليه فأسئلة الاختبار ألتحصيلي مقبولة وهي في متناول التلاميذ ولا يجدون صعوبة في الإجابة عليها وبالتلي تتحقق الفرضية الثانية.

3- بيانات اختبار مادة اللغة العربية غير مناسبة لافتراضات نموذج راش ،.

4- لا توجد فروق بين مرتفعي ومنخفضي التحصيل في التحصيل الدراسي .

5- لا توجد فروق بين الذكور والإناث.في التحصيل الدراسي

خلاصة (تفسير نتائج الدراسة)

من خلال التطرق إلى موضوع الدراسة الميدانية ، نستكشف أن الاختبار ألتحصيلي لمادة اللغة العربية في شعبة آداب وفلسفة لمستوى السنة الثالثة ثانوي لا يتمتع بمؤشرات القياس الكلاسيكي، الصدق والثبات، وأن فقرات أسئلة الاختبارالتحصيلي ، لا تكشف عن المستوى الحقيقي المعرفي –التحصيل الدراسي – لدى التلاميذ، في مستوى السنة الثالثة ثانوي،ومن هنا أقول أن ما هو جدير بالذكر، أن أربعة أساتذة في التعليم الثانوي من بين الاساتذة الذين تم

التنسيق معهم في الدراسة الميدانية ، صرحوا أنه يتم بناء الاختبار ألتحصيلي في المادة بناءا على أضعف تلميذ في القسم والشعبة ، وبالتالي فأسئلة الامتحان تكون سهلة ،والجدير بالذكر أن المادة الدراسية –لغة عربية –مادة أساسية في الشعبة وعليه يجب بناء اختبار تحصيلي وفق أسس موضوعية للكشف عن المستوى الحقيقي للتلاميذ ، خصوصا أن المستوى الدراسي سيتوج بامتحان شهادة البكالوريا .

ومن صعوبات الدراسة الميدانية ما يلي :

- ❖ قلة المراجع الخاصة بموضوع الدراسة خاصة العربية منها افتقار المكتبات الى هذا النوع من الدراسات وحتى الكتب الخاصة بموضوع الدراسة.
- ❖ عدم توفر البرامج الخاصة بنموذج راش في حدود اطلاعي ،كما لاحظت عدم تجاوب الأساتذة أثناء إشراكهم في موضوع الدراسة .
- ❖ أن المدة الزمنية المخصصة لموضوع الدراسة – سنة كاملة مع الدراسة – غير كافية بحكم إزدواجية العمل والمهنة- مستشار التوجيه والارشاد –وكحديث عهد بالتوظيف .

التوصيات والاقتراحات :

- ❖ اجراء دورات تدريبية للعامليين في مجال التقويم حول نموذج راش والتركيز على التدريب في استخدام البرامج الاحصائية الخاصة بنموذج راش
- ❖ توفير البرامج الإحصائية الحديثة التي تستخدم في تحليل بيانات الاختبارات النفسية والتربوية الخاصة بنموذج راش مثل ASCAL و RESCAL , وبرنامجي ITEMAN و هي ITEMEN32 وLOGIST وRUMM2020 و Quest و Microscal و miltilog و BILOG وغيرها .
- ❖ استخدام نموذج راش في بناء الاختبارات التحصيلية .
- ❖ بناء الاختبارات وأعدادها وفق الطرق الحديثة في المؤسسات التعليمية.
- ❖ اجراء دراسة مماثلة على مستوى التعليم المتوسط والابتدائي

قائمة المراجع

أ- أولا :المراجع باللغة العربية .

1-الكتب :

1- أحمد عبدا لرحمن (2011م): تصميم الاختبارات , الاردن, دار اسامة.

2- عبد الرحمن الطريرى (1997)، *القياس النفسى والتربوى 0نظريتـه ، أسسه ، تطبيقاته* ، الرياض ، مكتبة الرشد .

3- صلاح مراد وأمين سليمان (2002)، *الاختبارات والمقاييس فى العلوم النفسية والتربوية – خطوات إعدادها وخصائصها* ، القاهرة ، دار الكتاب الحديث.

4- سعد عبد الرحمن (1998)0 *القياس النفسى (النظرية والتطبيق)* ، ط3 ، القاهرة ، دار الفكر العربى

5- صلاح الدين علام (2000)،*القياس والتقويم التربوى والنفسى – أساسياته وتطبيقاته وتوجهاته المعاصرة* ، القاهرة ، دار الفكر العربى .

6- السيد أبو هاشم (2004) 0*الدليل الإحصائى فى تحليل البيانات باستخدام SPSS* الرياض ، مكتبة الرشد .

7- محمود محمد البستنجي(2010م) : كتاب القياس والتقويم للمعلم بين النظرية والتطبيق, المملكة العربية السعودية جدة, خوارزم العلمية للنشر والتوزيع.

8- صلاح الدين محمود علام(2006م) : الاختبارات والمقاييس التربوية والنفسية, عمَّان, دار الفكر للنشر والتوزيع

9-موسى النبهان (2004م): أساسيات القياس في العلوم السلوكية . عمان : دار الشروق للنشر والتوزيع.

10فريد كامل أبو زينة(1998م) : أساسيات القياس و التقويم في التربية , مكتبة الفلاح للنشر و التوزيع .

11-عبدا لواحد الكبيسي(2007م): القياس والتقويم تجديدات ومناقشات.عمان, دار الجرير للنشر والتوزيع

2-المجلات:

1- أحمد عودة (1992) ،مدى التوافق بين نموذج راش والمؤشرات التقليدية فى اختيار فقرات مقياس اتجاه سباعى التدريج ، *مجلة كلية التربية بجامعة الإمارات ، العدد (8) يونيو ، 153-179 .*

2- - محمد غنيم ووليد القفاص (2000)، إدراك الطلاب للمناخ الأكاديمى وعلاقته بمداخلهم للدراسة ، *المجلة المصرية للدراسات النفسية* ، *العدد (25) يناير ،120 -162 .*

3- عبد الرحمن الطريرى (1996) ، الخصائص السيكومترية لاختبار الذكاء الإعدادى باستخدام نموذج راش، *مجلة دراسات نفسية ، العدد (4) أكتوبر ،457-473*

4- صلاح الدين علام (1987)،دراسة موازنة ناقدة لنماذج السمات الكامنة ، والنماذج الكلاسيكية فى القياس النفسى والتربوى جامعة الكويت ، *المجلة العربية للعلوم الإنسانية ، العدد(27) ، 18-44.*

5- أحلام الشربينى ورضا حجازى (2002)، تقويم تحصيل تلاميذ مرحلة التعليم الأساسى فى العلوم باستخدام نظرية السمات الكامنة ، مجلة *بحوث المؤتمر العلمى السادس للتربية العلمية ، المجلد الثانى ، 611-648.*

6- صلاح الدين علام (1985)، تحليل بيانات الاختبارات العقلية باستخدام نموذج راش اللوغاريتمى الاحتمالى (دراسة تجريبية) ، *جامعة الكويت ، المجلة العربية للعلوم الإنسانية ، العدد(17) ، 100-124*

3-الرسائل والمذكرات :

1- إسماعيل الوليلى (2001)، دراسة سيكومترية مقارنة لبعض نماذج الاستجابة للمفردة فى انتقاء مفردات الاختبارات مرجعية المحك ،*رسالة دكتوراه ، جامعة الأزهر ، كلية التربية .*

2- أمينة كاظم (1996) ، *نماذج السمات الكامنة* فى (أنور الشرقاوى وآخرون (1996)0 *اتجاهات معاصرة فى القياس والتقويم النفسى والتربوى*،القاهرة ، الانجلو المصرية.

3- احمد محمد عبد الرحمن (2011م) : تصميم الاختبارات أسس نظرية وتطبيقات عملية , عمان , دار أسامة للنشر والتوزيع.

4- شحتة عبد المولى (1999)، تقويم بناء الاختبارات المرجعة إلى المحك / المعيار فى ضوء نظرية الاستجابة للمفردة والنظرية التقليدية ، *رسالة دكتوراه ، جامعة عين شمس ، كلية التربية*

5- عماد عبد المسيح (1991)،استخدام نموذج راش اللوغاريتمى أحادى البارامتر فى تحليل مفردات الاختبارات المعرفية مرجعية المعيار ثنائية القطب (دراسة تجريبية) ، *جامعة المنيا ، كلية التربية ، مجلة البحث فى التربية وعلم النفس ، العدد(4) ابريل ، 443-475 .*

ب-المراجع باللغة الأجنبية:

16- Abouserie ,R. (1993) . Approaches to studying and Levels of Processing : Relationship with self esteem and Achievement Motivation . *Egypt ion Journal Psychology Studies* ,4,1-20.

17- Clarke , R. (1986) . Students Approaches To Learning in AN Innovative Medical School : A Cross – Sectional Study . *British Journal of Educational Psychology*,56,3,309-321.

18- Duff ,A. (1997) . A note on the reliability and validity of a 30- item version of Entwistle & Tait's Revised Approaches to Studying Inventory . *British Journal of Educational Psychology*,67,4,529-539.

19- Entwistle, N. & Ramsden, P. (1983). Understanding Student Learning . London : Groom Helm.

20- Entwistle , N. & Kozeki , B. (1985). Relationships Between School Motivation , Approaches To Studying , and Attainment , among British and Hungarian Adolescents . *British Journal of Educational Psychology*,55,2,124-137.

21- Entwistle , N. (1981).*Styles Of Learning and Teaching* . New York , John Wiley &Sons.

22- Gregory , R. (2004) . *Psychological testing : History , Principles , applications* . 4th . Ed. , New York , Pearson Education Group Inc.

23- Hambleton , R. & Zaal , J. (1991). *Advances in Educational and Psychological Testing* . MA : Kluwer Academic Publishers , Boston .

24- Hambleton, R. & Swaminathan ,H.(1989).*Item response theory Principles and applications* . Boston , Kluwer Nijhoff Publishing.

25- Harper , G. & Kember , D. (1989) . Interpretation Of Factor Analyses From The Approaches To Studying Inventory . *British Journal of Educational Psychology*,59,1,66-74.

26- John , R. & Woodley , A. (1999). Approaches to studying in people with hearing loss. *British Journal of Educational Psychology*,69,4,533-546.

27- John , R. (1995). Mature Students in Higher Education : II.AN Investigation of Approaches to Studying and Academic Performance . *Studies in Higher Education* , 20,1,5-17.

28- John , R. (1994). Mature Students in Higher Education : I. A Literature Survey on Approaches to Studying . *Studies in Higher Education* , 19,3,309-325.

29- John , R. (1993). Gender Differences in Responses to the Approaches to Studying inventory . *Studies in Higher Education* , 18,1,3-13.

30- John ,R. (1990). Reliability and Replicability of the Approaches to studying Questionnaire . *Studies in Higher Education* , 15,2,155-168.

31- Kember , D. & Gow , L. (1990) . Cultural Specificity Of Approaches To Study . *British Journal of Educational Psychology*,60,3,356-363.

32- Ramsden , P. & Entwistle , N. (1981). Effects Of Academic Departments On Student Approaches To studying . *British Journal of Educational Psychology*,51,3,368-383.

33- Randall ,S.(1998). Comparing Measurement Theories . *Paper Present at the Annual Meeting of the American Educational Research Association (San Diego , CA , April 13-17).*

34- Smith , E. & Tasng , F. (1998). A comparative study of approaches to studying in Hong Kong and the United Kingdom . *British Journal of Educational Psychology*,68,1,81-93.

35- Smith , S. (2001). Approaches to Study of Three Chinese national groups . *British Journal of Educational Psychology*,71,4,429-441.

36- Smith, E. (1999). Intuition – Analysis Style and Approaches to Studying . *Educational Studies* , 25,2,159-173.

37- Watkins , D. , Hattie , J. & Astilla , E. (1986). Approaches to Studying by Filipino Students : A Longitudinal Investigation . ***British Journal of Educational Psychology*****,56,3,357-362.**

38- Waugh , R. & Addison , P. (1998). A Rasch measurement model analysis of the Revised Approaches to Studying Inventory. ***British Journal of Educational Psychology*****,68,1,95-112.**

39- Waugh , R. (2002). Measuring self – reported studying and learning for university students : Linking attitudes and behaviours on the same scale . ***British Journal of Educational Psychology*****,72,4,573-604.**

40- Waugh , R.(1999). Approaches to studying for students in higher education : A Rasch measurement model analysis . ***British Journal of Educational Psychology*****,69,1,63-79.**