كلية التربية
كلية معتمدة من الهيئة القومية لضمان جودة التعليم
إدارة: البحوث والنشر العلمي (المجلة العلمية)
=======

بناء أداة لتقويم الاختبارات التحصيلية
في الجامعات السعودية
من وجهة نظر الطلاب

إعـداد

د/ محمد بن حسن يحيى الزبيدي

أستاذ القياس والتقويم المساعد

قسم علم النفس – جامعة الطائف

﴿ المجلد الرابع والثلاثون– العدد الثاني – جزء ثانى – فبراير ٢٠١٨م ﴾

http://www.aun.edu.eg/faculty_education/arabic

الملخص:

يهدف البحث الحالي لبناء أداة لتقويم الاختبارات التحصيلية التي يقوم بإعدادها اساتذة الجامعات وذلك اسهاما من الباحث في تطوير وجودة اساليب التقويم التي يتبعها الاستاذ الجامعي بما يتفق مع المعايير المعدة من قبل الجامعات السعودية في بناء ادوات التقويم.

وقد تم الاطلاع على الاساليب التي يتبعها الاساتذة الجامعيين في تقويم تحصيل طلابهم ووجد أن الاختبار التحصيلي اكثر الادوات استخداما، لذا قام الباحث بتصميم اداة تشمل المعايير الفنية والعلمية في بناء الاختبارات التحصيلية وتم تحكيم الاداة من قبل مجموعة من الخبراء ثم طبقت على عينة استطلاعية لاستخراج بعض معايير الاختبار التحصيلي الجيد، حيث حصل الباحث على اجماع المحكمين على صلاحية هذه الاداة لقياس مهارات الاساتذة الجامعيين في بناء الاختبارات التحصيلية ،وحسب معامل الاتساق الداخلي بين عبارات الأداة فكان ٥٤% وهو مؤشر مناسب لاستخدام هذه الاداة . كذلك وجد ان ثبات هذه الاداة باستخدام معامل الفا كان ٧٢% . وقد اتبع الباحث منهج التحليلي الوصفي.

الكلمات المفتاحية : الاختبارات التحصيلية

Abstract

The current research aims to build a tool for evaluating the achievement tests are prepared by university professors. This is as a contribution of the researcher in developing and making sure of quality of the evaluation methods. These evaluation methods are followed by the university professor in accordance with the standards that are prepared by the Saudi universities in constructing evaluation tools.

Through reviewing the literature, the researcher has found that the achievement test is the common used tool by lecturers in their evaluation to their students. So, the researcher designed a tool including technical and scientific standards in constructing achievement tests. The tool was arbitrated by a group of experts. Then, it was applied on a pilot sample to extract some criteria of the good achievement test. The researcher obtained the consensus of the arbitrators on the validity of this tool to measure the skills of university professors in the constructing of achievement tests, and according to the coefficient of internal consistency between items of the tool was 54%, which is regarded as a suitable indication for using the tool. Also, It was found that the reliability of this tool using *Cronbach's alpha* which was 72%. The descriptive analytical approach was employed in the current study..

Keywords: Examinations

مقدمة :

يعتبر التعليم الجامعي أحد أهم مرتكزات التنمية البشرية؛ ذلك أنه يتعلق بإعداد الكفاءات المتخصصة في مختلف مجالات الحياة، وبقدر جودة التعليم الجامعي بقدر ما نضمن جودة هذه الكفاءات. والتدريس الجامعي باعتباره أحد الأهداف الأساسية للجامعات والمعاهد العليا، يتصل به مجموعة من العوامل التي تتعلق بالأستاذ الجامعي، والطلبة، والمناهج الجامعية، وإدارة الجامعات. وهذه العوامل تتداخل معاً لتؤثر على نوعية وجودة التدريس الجامعي سلباً أو إيجاباً، وبقدر توفر متطلبات الجودة في كل هذه العوامل بقدر ما تكون جودة الاستاذ الجامعي.

ولقد جاء الاهتمام بموضوع الجودة في التعليم نتيجة لما أشارت إليه معظم التقاريرالعالمية بوجود شكوى من مستوى جودة التعليم في الدول النامية والمتقدمة على حد سواء، ففي منتصف القرن العشرين ركزت الولايات المتحدة الأمريكية وفرنسا ودول أخرى على دراسة جوانب الجودة في التعليم، بسبب انخفاض مستوى التعليم فيها ، وعلى المستوى العربي فقد كانت الشكوى عامة حول التعليم العالي، ففي اللقاء الذي نظمه منتدى الفكر العربي بالتعاون مع برنامج الأمم المتحدة الإنمائي Undp حول التعليم العالي في البلدان العربية عام ١٩٩٧م، ذكر صبحي القاسم أنه لم يتم التعامل مع الجامعات العربية كمشروع استثماري وطني، وإن حدث، فهو كما يذكر إبراهيم بدران، استثمار في الكم وفي الحجم وفي العدد، وهو غالباً ما يمثل حالة من الاستثمار غير الرشيد، بسبب التوسع الأفقي في كل اتجاه على حساب العمق، لذلك أوصى في ورقته البحثية اعتبار المشروع الجامعي مشروعاً اقتصادياً بالمعنى المقصود، وتقديم التمويل الكافي والحوافز التي تجعل الاستثمار في التعليم الجامعي أكثر جدوى وقابلية للتطوير (منتدى الفكر العربي، ١٩٩٧م) .

ومن أجل جعل التعليم العالي مشروعاً استثمارياً وطنياً، لابد من التأكيد على مفاهيم الجودة وبرامجها في التعليم العالي، وهذا لا يعني التركيز على نوعية مخرجات منظومة التعليم العالي فحسب، بل لابد من التأكيد على كل عنصر من عناصر هذه المنظومة، بما في ذلك جودة هيئة التدريس، والخدمات التعليمية، والإدارية.

وتسعى الجامعات بشكل عام على تحقيق خريجيها للأهداف التعليمية المرسومة من قبل السياسة التعليمية العليا في الدولة او البلد، ولان الطالب الجامعي المخرج الأساسي للجامعات ، وبقدر تأهيله وتزويده بالمعارف والمهارات الأساسية التي تهيئه لتولي دوره في بناء المجتمع وتطوير النشء ،مما بعكس جودة وقدرة المؤسسة التي تخرج منها.لذا توجب على اعضاء هيئة التدريس في هذه الجامعات الحرص على قياس وتقويم مخرجات التعلم لدى طلابهم بالاساليب العلمية المناسبة وفق التوجهات الحديثة في القياس والتقويم، والسعي جادين للتزود والتدرب على اساليب التقويم المتعددة.

مشكلة البحث:

من خلال تدريس الباحث لبعض مقررات علم النفس بجامعة أم القرى وجامعة الطائف واحتكاكه بالطلاب وما وصل إليه من شكاوى من الطلاب حول طرق التقويم التي يتبعها أساتذة الجامعة واعتمادهم على الاختبارات التحصيلية بشكل كبير،وأن انخفاض معدلاتهم يرجع إلى طريقة التقويم التي يتبعها الأستاذ الجامعي، تولد لدى الباحث الاحساس بالمشكلة وضرورة بحثها لذا فقد تمثلت مشكلة البحث في بناء اداة تسعى الى تقويم الاختبارات التي يعدها الاستاذ الجامعي وذلك من وجهة نظر الطالب وفق الاعتبارات التالية:

- ما نوع الأسئلة التي يستخدمها الأستاذ الجامعي في تقويم تحصيل طلابه ؟
- هل تتوفر المهارات الأساسية في بناء الاختبارات التحصيلية التي يعدها الأستاذ الجامعي؟
- هل يتم تدريب الطلاب على وسائل التقويم التي يعدها الأستاذ الجامعي؟

كما تتناول مشكلة البحث دراسة واقع الاسئلة التي يعدها الاستاذ الجامعي من خلال الاعتبارات التالية:

- هل تتمتع الأسئلة التي يعدها الأستاذ الجامعي بدرجة مناسبة من الشمولية ؟
- هل تتمتع الأسئلة التي يعدها الأستاذ الجامعي بدرجة مناسبة من الموضوعية ؟
- هل توجد تعليمات واضحة لأداة التقويم يفهمها الطلاب؟

وعليه جاءت هذه الدراسة لتشرك المحور المهم (الطلبة) في العملية التعليمية في بيان رأيه وذلك لكونه على اتصال مباشر بالاستاذ الجامعي . وتقييم الطلبة له مساهمة ايجابية في تطوير أداء الأستاذ الجامعي وفي المساعدة على اتخاذ القرار بالنسبة للجهات المسؤولة خاصة.

واعتمد البحث في إطاره النظري على أحدث التوجهات والنماذج العلمية والرسمية لتأهيل الاستاذ الجامعي وتطويره، ومن أبرزها قدرته على استخدام ادوات القياس المناسبة، ودرجة تمكنه منها. وعلى الاستاذ الجامعي أن يبلغ المستوى والدرجة الدالة على تمكنه من خلال تطبيقه وممارسته. ثم قدم البحث مجموعة من التوصيات والمقترحات لتطوير أداء أساتذة الجامعت في مجال تقويم تحصيل طلابهم.

اهمية البحث:

يعد قياس وتقويم تحصيل الطلاب من المشكلات الرئيسة التي شغلت ولا تزال تشغل اهتمام المربين وخبراء القياس التربوي حتى وقتنا الحالي، والسبب في ذلك أن أساليب القياس والتقويم تلعب دوراً مهماً في تطوير العملية التعليمية ويمكن أبراز أهمية الدراسة من خلال النقاط التالية :

١- زيادة الاهتمام بتطوير الاختبارات التي يعدها المدرسون لقياس التحصيل الدراسي منذ السبعينات وحتى وقتنا الحاضر، وذلك التطور قد ظهر على يد العديد من علماء القياس التربوي أمثال . بابام (Popham,١٩٧٥) وبيرك (Berk , ١٩٧٨) وبلوم (Bloom١٩٧١) في (علام، ٢٠٠٦)

وبناءاً على ذلك فإنه من الطبيعي أن يواكب ذلك التطور تغيراً في أساليب القياس والتقويم، وبالتالي في الاختبارات المدرسية.

٢- تحتل عملية قياس وتقويم تحصيل الطلاب جانبا كبيراً من اهتمام المفكرين التربويين في مختلف دول العالم، فقد أجريت دراسات متعددة تناولت الاختبارات المدرسية التي بعدها المعلم للتعرف على جوانب الضعف والقوة لدي المعلمين في مجال الاختبارات،ومن هذه الدراسات، دراسة كيربر(Kerber 1976)، ودراسة (علام، ١٩٨٠).

٣- تظهر أهمية هذا البحث بشكل واضح في الجهود التي بذلتها المنظمة العربية للتربية والثقافة والعلوم وما عقدته من مؤتمرات وحلقات في هذا المجال، وقد اكدت هذه الحلقات والمؤتمرات في توصياتها على:

إعطاء أهمية خاصة للاختبارات التي تقيس الميول والاهتمامات والقدرات الاستعداديه.

عقد دورات تدريبية في إعداد الاختبارات وطرق تقويم الطلاب ومتابعة نموهم .

كما عقد مؤتمر خبراء تطوير الاختبارات بالكويت في الفترة من ٧-١٢ ديسمبر (١٩٧٤م) وقد شاركت فيه وفود عديدة من الدول العربية، وهدف المؤتمر مناقشة موضوع الامتحانات بالدول العربية، من حيث عيوبها وإمكانات تقويمها وتطويرها، ومن أهم ما نوقش بالمؤتمر العلاقة بين الأهداف والامتحانات، والعلاقة بين المحتوى الدراسي والامتحان، ومستويات الامتحانات المعرفية، وأختتم المؤتمر أعماله بعدة توصيات تركزت على :

- ضرورة إجراء بحوث تتناول الاختبارات الجامعية.

- ضرورة الاهتمام بتحسين الأساليب الفنية والإدارية للامتحانات.
- الاهتمام باستخدام الأساليب الموضوعية في الامتحانات.
- ضرورة شمول الامتحانات للمستويات المعرفية العليا وعدم الاقتصار على قياس المستويات الدنيا.

أهداف البحث:

يهدف البحث الحالي إلى بناء اداة لتقويم اسئلة الاستاذ الجامعي من وجهة نظر الطلاب .

وذلك من خلال التعرف على:

- نوع الأسئلة التي يستخدمها الأستاذ الجامعي في تقويم تحصيل طلابه.
- توفر المهارات الأساسية في بناء الاختبارات التحصيلية التي يعدها الأستاذ الجامعي.
- تدريب الطلاب على وسائل التقويم التي يعدها الأستاذ الجامعي.
- تمتع الأسئلة التي يعدها الأستاذ الجامعي بدرجة مناسبة من الشمولية.
- تمتع الأسئلة التي يعدها الأستاذ الجامعي بدرجة مناسبة من الموضوعية .
- وجود تعليمات واضحة لأداة التقويم يفهمها الطلاب.

مصطلحات البحث :

التقويم :

التقويم لغة :من قوم أي صحح وأزال الاعوجاج وقوم السلعة بمعنى سعرّها. وعرف بلوم التقويم بأنه إصدار حكم لغرض ما على قيمة الأفكار أو الأعمال أو الحلول أو الطرق أو المواد، وأنه يتضمن استخدام المحكات والمستويات والمعايير، لتقدير مدى كفاية الأشياء ودقتها وفعاليتها، ويكون التقويم كمياً أو كيفياً.(علام،٢٠٠٣)

وعرف جرونلند التقويم بأنه عملية منهجية، تحدد مدى ما تحقق من الأهداف التربوية من قبل الطلبة، وأنه يتضمن وصفاً كميا وكيفياً، بالإضافة إلى إصدار حكم على القيمة. (علام،٢٠٠٣)

الاختبار التحصيلي:

التحصيل عند حمادين(٢٠٠٥م) هو" مدى ما تحقق من أهداف التعلم في موضوع أو مساق أو عدة مساقات سبق للطالب دراستها أو التدرب عليها من خلال المشاركة في الأعمال المبرمجة.

ذكر علام (٢٠٠٦م) " أن بروان يرى الاختبار بأنه إجراء منظم لقياس عينة من سلوك الفرد، بينما يرى شيس الاختبار بأنه أسلوب منظم لمقارنة أداء الفرد بمعيار أو مستوى أداء محدد" ويعين الاختبار عند الكبيسي (٢٠٠٧م) بأنه" قياس لأداء الفرد في مجالات التعليم وعلم النفس".

ويعرف القرني (١٤٢١هـ) الاختبار التحصيلي بأنه" عبارة عن أداة من أدوات القياس التي يستخدمها المعلم بطريقة منظمة، لتحديد مستوى تحصيل الطلاب في مادة دراسية معينة"

والاختبار التحصيلي يعرفه بامشموس وآخرون (١٤٠٥هـ) بأنه" الأداة التي تستخدم في قياس المعرفة والفهم والمهارات في تعلم مادة دراسية معينة أو مجموعة من المواد الدراسية بوضح أسئلة تراعي في صياغتها شروط معينة، بحيث تمر في خطوات تجريبية تصل بها إلى تحقيق أهداف الاختبار"

التعريف الاجرائي لتقويم التحصيل:

ومن خلال العرض السابق لمساهمة بعض علماء وباحثي علم النفس التربوي والقياس والتقويم ، يمكن للباحث تعريف تقويم التحصيل الدراسي بما يلي :

يدل تقويم التحصيل الدراسي على قياس مقدار ما يكتسبه الطالب من معلومات ومعارف وخبرت ومهارات نظرية أو تطبيقية ، نتيجة لأدائه للاختبارات التحصيلية خلال مرحلة دراسية معينة بطريقة منظمة، ويمكن قياس هذا التحصيل من خلال التعرف على متوسط الدرجات التي حصل عليها الطالب.

الإطار النظري :

يعرض الباحث في هذا الفصل المفاهيم النظرية المرتبطة بمتغيرات البحث مثل التقويم والتحصيل الدراسي والعوامل المؤثرة فيه وإلقاء الضوء على أهميته وأسسه ومكوناته وأبعاده .

الاتجاهات الحديثة في التقويم :

نتيجة للتطور في نظريات وأساليب القياس والتقويم النفسي والتربوي واستجابة للتغير الاجتماعي والتقدم التقني ومن خلال الجهود المبذولة لتطوير وتعديل نظم الامتحانات ظهرت مجموعة من الاتجاهات في تقويم الطالب يمكن أن نوجز أهمها فيما يلي:

- التقويم الشامل الذي تنادي به التربية الحديثة ،والذي يمتد ليشمل جميع جوانب شخصية الطالب في إطار برنامج شامل ومتوازن للتقويم ، لا يقتصر على نوع واحد من أساليب التقويم وجمع البيانات وإنما تتنوع فيه الأساليب حسب تنوع الأهداف وأنماط السلوك المراد قياسها •
- ينظر الاتجاه الحديث إلى الامتحانات وعملية التقويم على أنها جزء متكامل من العملية التربوية.
- يعتمد التقويم في الاتجاه الحديث على نماذج ونظريات للقياس النفسي، وقد ظل القياس النفسي ولفترة طويلة يعتمد على المقاييس معيارية المرجع (Norm referenced) والتي تعتمد على مقارنة أداء الفرد بأداء غيره من الأفراد وفي السبعينات بدأ التوسع في استخدام المقاييس محكية المرجع (Criterion referenced) وهي تفسر درجة الفرد في ضوء مدى تحقيقه أهدافاً تعليمية محددة ، وهي بذلك تعطي دلالة على مدى كفاءة الفرد في القيام بأعمال معينة.
- يميل التقويم الحديث إلى الابتعاد عن المركزية والشكلية ويدعو إلى التقليل من الاعتماد على الاختبارات المقننة والاتجاه إلى الاختبارات التي يعدها المعلم ، وهي بذلك تنقل مسؤولية التقويم من الإدارة المركزية بالوزارة إلى إدارة التعليم والمدرسة والمعلمين وفي ذلك تخفيف الأعباء وتحقيق المزيد من المرونة والتنوع في تحقيق الأهداف ويقتصر دور الإدارة المركزية في التخطيط والتطوير وتبادل المعلومات.
- استنباط أنواع جديدة من أساليب التقويم التي تتسم بالمرونة وتعتمد على الثقة المتبادلة بين المعلم والطالب ، وتزيد من ثقة الطالب بنفسه كاختبار الكتاب المفتوح والاختبار المنزلي والأبحاث والمشروعات.
- الارتقاء بمستوى الإعداد الفني لأساليب التقويم ، بحيث يتم بناؤها وفقا للأسس العلمية المتطورة للقياس النفسي.
- الاستفادة من التقدم الذي أحرزته بعض الأجهزة الحديثة لإكساب عملية التقويم المزيد من المرونة والسهولة والسرعة.
- سيهتم نظام التقويم في المستقبل بمدى قدرة الطالب على تقويم المعلومات وتقدير علاقتها بالموضوع الذي يدرسه وأن يعرف نوع المعلومات التي يحتاجها ومصادرها وكيفية استخدامها •

- تركز أساليب التقويم الحالي على التفكير التقاربي Convergent thinkingوالذي يفترض وجود إجابات صحيحة محددة للأسئلة ، هذه المهارات ستصبح في المستقبل قليلة الفائدة ،وسيتجه التقويم إلى قياس الصفات والمهارات الأكثر أهمية لعالم متغير كالمبادرة والأصالة والتكيف والمثابرة والقدرة على القيادة والتعاون.
- السعي لابتكار أساليب جديدة لتقدير أداء الطلاب بدلا من نظام الدرجات والعلامات ، والذي يعطي درجة للأداء في مقرر معين ولا يعطي تفصيل لإمكانات الطالب ومواهبه ومهارته.

ومن أمثلة البدائل المقترحة التقارير التفصيلية التي ترسل إلى أولياء الأمور، ومن المقترح أيضا المخطط التحصيلي للطالب وهو ذلك الذي يتضمن بيان الجوانب المختلفة الهامة التي اكتسبها الطالب من المهارات والصفات الشخصية والأنشطة خارج المنهج وتحديد مستويات أداء الطالب في كل منها (القرشي، ١٤٠٦هـ)

أهمية الاختبارات التحصيلية :

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

تعد الاختبارات من مقومات العملية التعليمية ، وهي جزء من القياس ، وترتبط به ارتباطا وثيقا ، تستخدم للحكم على مدى النجاح في تحقيق الأهداف المنشودة، وتحظى الامتحانات بأهمية كبيرة في جميع المؤسسات التربوية التقليدية والمتطورة ، وكثيرا ما ينعكس هذا على عملية التدريس.

تأتي أهمية الاختبارات من كونها وسيلة للحكم على فاعلية العملية التعليمية والتربوية، من خلال ما تقوم به من دور فعال فيها ؛ حيث تساعد على تشخيص العقبات والمشكلات التعليمية ، وتقدم الحلول المناسبة وأوجه حلها . وهي مهمة للمؤسسة التعليمية بحيث تدفعها لمراجعة أهدافها، ومدى ملاءمة المنهج لتحقيق هذه الأهداف ؛ كما تأتي أهميتها في أنها تساعد على معرفة مدى تحقيق الخطة التعليمية للأهداف الخاصة بها .

لقد ارتبط تطور الاختبارات في كافة مجالاتها وأنواعها بالتطور الذي أصاب جوانب حياة الإنسان في كافة المجالات ، خاصة في الربع الأخير من القرن الماضي ، بفضل جهود علماء التربية الحديثة ، مما زاد من أهمية الامتحانات بشكلها العام ؛ إذ أصبحت هذه الامتحانات تتطور باستمرار مع كل تطور يصيب حياة البشر ، مما يزيد من أهميتها في العملية التعليمية (أبو لبدة ، ١٩٨٥).

أغراض الاختبارات التحصيلية:

ذكر (القرني، ١٤٢٠هـ) عدداً من أغراض الاختبارات التحصيلية ومنها:

- **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 المناهج أو المرافق التعليمية المختلفة ومصادر التعلم، مع الإشارة إلى أن هناك اختبارات خاصة بالتشخيص، ولكن هذا لا ينفي عن الاختبارات التحصيلية الصفية العادية وظيفة التشخيص.
- **التصنيف:** أي تصنيف الطلاب إلى تخصصات مختلفة: أكاديمي – تجاري – صناعي – زراعي – وما إلى ذلك، أو تصنيفهم إلى مجموعات اعتماداً على قدراتهم العقلية أو ميولهم. وبطبيعة الحال لا يكون هذا التصنيف ممكناً إلا بالاعتماد على نتائج الطلاب في اختبارات تحصيلية، أو اختبارات خاصة، أو وسائل قياس أخرى من مقابلات شخصية أو استبانات وقوائم وما إلى ذلك.
- **قياس مستوى التحصيل:** مدى تحقق الأهداف التعليمية لدى المتعلم في مادة دراسية بعينها، وفي المواد الدراسية جميعها، والاختبارات التحصيلية في معظمها إنما تنصب على تحقيق هذا الهدف بقصد الأخذ بنتائجه في تحسين مستوى التعليم وترفيع الطلاب إلى صفوف أعلى وما إلى ذلك.
- **تنشيط دافعية التعلم،** والنقل من صف إلى آخر ومنح الدراجات والشهادات.
- **التعرف على مجالات التطوير** للمناهج والبرامج والمقررات الدراسية.

تصنيف الاختبارات التحصيلية:

هناك أكثر من أساس لتصنيف الاختبارات التحصيلية، ذكرها كل من: الدوسري(٢٠٠١م) (مراد، وسليمان، ٢٠٠٢م) (علام ، ٢٠٠٦م) ولعل أهمها ما يلي:

التصنيف وفق طريقة أداء الطلبة (طريقة الاستجابة):

أ- اختبار شفوي:

وهو اختبار غير مكتوب يقدم للمفحوصين في صورة أسئلة شفوية ويجيب عليها بنفس الطريقة، وهي من أقدم أنواع الاختبارات وتستخدم في عدد من المجالات كالقراءة الجهرية، وإلقاء الشعر وتلاوة القرآن الكريم.

ب- اختبار تحريري:

وهو اختبار مكتوب يقدم للمفحوص في صورة أسئلة تحريرية ويجيب عليها كتابياً، والغرض منها معرفة قدرة الطالب على الربط والتنظيم وإدراك العلاقات.

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

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

التصنيف وفق تفسير درجات الاختبار:

الاختبارات مرجعية المعيارcriterion –referenced tests :

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

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

الاختبارات مرجعية المحك norm – referenced tests :

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

اولا: الاختبارات معيارية المرجع norm – referenced tests :

خصائص الاختبارات معيارية المرجع :

ذكر (علام،٢٠٠٦م) مجموعة من الخصائص للاختبارات المعيارية منها:

١ . تحدد هذه الاختبارات بواسطة لجان متخصصة ، بحيث تكون عامة على مستوى الدولة.

٢ . يمكن من خلال هذه الاختبارات مقارنة أداء الطلاب ببعضهم بعضاً ، وهنا يمكن القول: الطالب الأول والطالب الثاني ، وهكذا . فمثلاً الطالب الذي يحصل على معدل ٨٠% يمكن القول بأنه يتفوق على ٧٩% من الطلبة .

٣ . تعتبر الاختبارات المعيارية المرجع من وسائل التقويم الختامي ، وبالتالي تجري مرة واحدة سنوياً أو كل فصل دراسي ، وتشمل المادة كلها في الغالب .

٤. في الاختبارات معيارية المرجع ، ينظر للمادة الدراسية ككل لا كأجزاء فامتحان اللغة العربية مثلاً يتضمن القواعد ، والقراءة ، والمحفوظات ... وعلامة الطالب عليها كلها يعني علامة الطالب في اللغة العربية . ويمكن في هذه الاختبارات أن يغطي نجاح الطالب في بعض الفروع على رسوبه في فروع أخرى ، فنجاح الطالب في القواعد مثلاً قد يغطي رسوبه في القراءة ، وبالتالي ينظر لنجاحه في مادة اللغة العربية ككل .

٥. في هذه الاختبارات تحدد في العادة علامة النجاح والعلامة العليا ، فمثلاً لو كانت العلامة الكلية (١٠٠) تكون علامة النجاح (٥٠) وهي متوسط العلامة والحد الأعلى (١٠٠) .

ثانيا: الاختبارات محكية المرجع

خصائص الاختبارات محكية المرجع :

تتصف الاختبارات محكية المرجع بعدد من الخصائص أهمها :

١- يحدد المحك لهذه الاختبارات بناء على خبرة المعلم ومعرفته بتلاميذه ، ومثال ذلك أن معلم الرياضيات قد يحدد إمكانية حل الطالب لعشرة أسئلة على الجذر التربيعي ، حلاً صحيحاً ، دليلاً على استيعابه بموضوع الجذر التربيعي . وهنا إذا أعطي الطالب خمسة عشر سؤالاً واستطاع أن يحل عشرة منها بشكل صحيح أو أكثر من عشرة فإنه يعني أنه استوعب الموضوع (الجذر التربيعي) . أما إذا لم يستطع حل ذلك العدد من الأسئلة كأن يحل تسعة أسئلة . أو أقل فهذا يعني أنه لم يستوعب الدرس .

٢- في الاختبارات المحكية المرجع يقارن أداء الطالب بالمحك ، وليس بغيره ، كما هو متعارف عليه في الاختبارات المبنية على المعيار المرجعي . ففي المثال السابق يقارن أداء الطالب بعدد الأسئلة الصحيحة اللازمة لاعتبار تحقيق هدف الاستيعاب .

٣- الاختبارات المحكية المرجع من وسائل التقويم التكويني :

وهذا يعني أنه قد تجري عدة مرات للمادة الواحدة ، والتي قد يحدد المعلم محكات لكل اختبار في المادة نفسها . وبالتالي فإن نجاح الطالب في أحدها لا يعني نجاحه بالمادة جميعها.

ومثال ذلك ، أن نجاح الطالب في استيعاب الجذر التربيعي ، كما في المثال السابق ، لا يعني نجاحه في مادة الرياضيات جميعها .(علام،٢٠٠٦)

وحتى يحكم على استيعاب الطالب لمادة الرياضيات ، لابد أن يجتاز المحكات المرجعية جميعها والمحددة لكل دروس الرياضيات . وهذا أيضاً يختلف عن الاختبارات التي تبنى على معيارية المرجع التي تحقق الأهداف العامة للمادة .

ومثال آخر على ذلك من مادة اللغة العربية ، فإذا حدد المعلم بأن حفظ قصيدتين شعريتين من بين ثلاث قصائد موجودة في مادة اللغة العربية ، هو شرط نجاح الطالب في مادة المحفوظات ، وإذا حفظ طالب ما قصيدتين من بين القصائد الثلاث المذكورة فإنه يكون قد حقق هدف الحفظ للمحفوظات فقط ، ولا يعتبر قد استوعب مادة اللغة العربية ككل .. وهكذا.

٣ : التصنيف وفق طريقة تصميم وبناء الاختبار:

يمكن تصنيف الاختبارات التحصيلية إلى:

الاختبارات المقننة:

وهي الاختبارات التي تعد بواسطة خبراء القياس ومؤسسات نشر الاختبارات وتستخدم في كافة المدارس وتتفق مع الأهداف العريضة للمادة الدراسية، وبذلك فهي تقيس معلومات اتفق خبراء التربية على أنها ضرورة لمتوسط الأداء في مادة دراسية معينة، ويطبق على عينة كبيرة نسبياً من الطلاب لها خصائص معلومة وتنتقي بطريقة عشوائية لإيجاد معايير استناداً إلى متوسط أداء هذه العينة وتصبح هذه المعايير الوسيلة التي تقارن في ضوءها كل طالب بمتوسط أداء هذه العينة.

الاختبارات التي يكتبها المعلم لطلابه:

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

التصنيف حسب نمط الإجابة:

- الإجابة المختارة (انتقائية).
- الإجابة المصاغة (منشأة).

ومن أكثر هذه الاختبارات شيوعاً:

أولاً: اختبارات الاختيار من متعدد:

ذكر عوده (١٩٩٣م)" بان الفقرة في اختبارات من متعدد تتكون من جزأين رئيسيين: الأول يسمى المتن، وقد يكون على شكل سؤال أو عبارة تتضمن قضية معينة تحتاج إلى إجابة، ويسمي الجزء الثاني: بالبدائل والتي هي بمثابة حلول أو إجابات محتملة للسؤال أو القضية الواردة في المتن وتتضمن البدائل إجابة صحيحة وعدداً من الإجابات الخاطئة والتي يطلق عليها بالمشتتات أو المموهات".

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

أنماط اختبار الاختيار من متعدد:

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

نمط الإجابة الصحيحة:

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

نمط الإجابة الخاطئة:

في هذا النمط يحتوي السؤال على إجابة واحدة خاطئة وبقية البدائل صحيحة ويرى جفري " أن النمط يكون مناسباً حينما تكون معرفة الاستثناء لها نفس أهمية معرفة القاعة" (العريفي، ١٤٢٧هـ).

ذكر أبو لبدة (١٩٨٥م) أن هذا النوع من الأسئلة قد تكن الأجوبة جميعها صحيحة بصورة جزئية ولكمن أحدهما أكثر صحة من البقية أو أفضل من غيره من الأجوبة، وتؤكد ميري" على أنه حينما يستخدم هذا النمط لابد وأن يتضح أن الخيار الصحيح هو الخيار الأفضل" (العريفي، ١٤٢٧هـ:١٩) ويكن أن تستخدم هذه الصورة عند أبو علام (١٩٨٧م) لقياس الفهم والتطبيق" وذكر (علام ٢٠٠٧م) أن هذا النوع من المفردات يكون أكثر صعوبة في إجابته من النوع السابق وذلك لأنه يتطلب إجراء تمييزات دقيقة بين البدائل المعطاة".

نمط أسوأ الأجوبة:

ويشبه أبو لبدة (١٩٨٥م)" هذا النمط بالنمط السابق إلا انه في هذه الحالة يتطلب من المفحوصين اختيار أسوأ الأجوبة".

نمط يتطلب إكمال فراغات باستخدام بدائل معطاة:

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

قواعد صياغة اختبارات الاختيار من متعدد:

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

قواعد متعلقة بأصل السؤال:

- يجب أن يحتوي أصل السؤال على مشكلة محددة وواضحة ومفهومة لدى المفحوص.
- يفضل تجنب استخدام صيغ النفي، لأن ذلك يربك المفحوص وإلا وجب وضع خط تحت أداة النفي أو إبرازها حتى ينتبه المفحوص.
- استقلالية الأسئلة وعدم اعتماد سؤال على آخر.
- يراعى عدم وجود أي تلميح أو إشارة في المتن يمكن أن تشير إلى الجواب أو تدل عليه.
- إذا كان السؤال متعلق بتعريف مصطلح معين فمن الأفضل وضع المصطلح في أصل السؤال.

ويذكر عودة (١٩٩٣م) " أنه كلما زاد عدد البدائل قلت فرصة التخمين، ففي حالة وجود خمسة بدائل للفقرة تكون نسبة التخمين ٢٠%، بينما تكون نسبة التخمين ٢٥% في حالة أربعة بدائل، و٣٣% في حالة وجود ثلاثة بدائل"

ثانياً: اختبارات الصواب والخطأ:

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

ثالثاً: اختبارات المزواجة (المطابقة):

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

رابعاً : الإجابة المصاغة (المنشأة):

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

أولاً: الأسئلة الإنشائية المحددة (المقيدة):

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

ثانياً: الأسئلة الإنشائية المفتوحة:

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

خطوات بناء الاختبار التحصيلي الجيد:

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

أولاً: تحديد الغرض من الاختبار:

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

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

ثانياً: تحديد الأهداف التعليمية:

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

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

المجال المعرفي:

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

- **التذكر**: وتتمثل في قدرة الطالب على تذكر المعارف أو المعلومات السابقة واسترجاعها، وهي أدنى مستويات المجال المعرفي، ومن الأفعال السلوكية التي تستخدم في صياغة هذا المستوى، يذكر، يسمي، يصف، يعدد.
- **الفهم**: وهو القدرة على إدراك واستيعاب معنى المادة أو النص الذي يدرسه الطالب، ومن أفعاله السلوكية التي تستخدم في صيغة هذا المستوى: يفسر، يترجم، يوضح، يعطي، يستخرج.

- **التطبيق:** ويعني استخدام ما تم تعلمه من قوانين ومفاهيم أو قواعد أو مبادئ أو تعليمات أو إجراءات وغيرها في مواقف جديدة أو مواقف واقعية، ومن الأفعال السلوكية التي تستخدم في صياغة هذا المستوى: يطبق، يحل، يربط، يشكل، يصحح.
- **التحليل:** ويعني القدرة على تجزئة أو تحليل المادة إلى عناصرها أو مكوناتها، ومن الأفعال السلوكية التي تستخدم في صياغة هذا المستوى: يحلل، يصنف، يستنبط، يميز.
- **التركيب:** ويعني القدرة على جمع عناصر أو أجزاء لتكوين كل متكامل أو نمط أو تركيب غير موجود أصلاً، ومن الأفعال السلوكية التي تستخدم في صياغة هذا المستوى: يركب، يؤلف، يصمم، يكون، يبني.
- **التقويم:** هو أعلى مستويات المجال المعرفي، ويعني القدرة على إصدار أحكام حول قيمة الأفكار أو اتخاذ قرارات مناسبة، ومن الأفعال السلوكية التي تستخدم في صياغة هذا المستوى: يقوم، يحكم، يقارن، يعطي الطلب حكماً، يصدر الطالب حكماً، ينتقد (الحريري، ٢٠٠٧م).

المجال الوجداني:

ويهتم هذا المجال بكل ما يرتبط بالعاطفة الإنسانية و قيم واتجاهات وميول وتقبل واستجابة.

المجال النفس حركي (المهاري):

يتضمن هذا المجال الأهداف التي تؤكد على نتاجات التعلم ذات العلاقة بالمهارات الحركية، ويشتمل هذا المجال على عدة مستويات: الملاحظة والتقليد، والتجريب، والممارسة، والإتقان والإبداع.(الحريري، ٢٠٠٧م).

ثالثاً: تحليل المحتوى:

يقصد بتحليل المحتوى الوصول إلى مفردات المقرر الدراسي، أو إحصاء المعلومات الأساسية في المقرر الدراسي، أو هو إحصاء المعارف والمهارات الأساسية المتضمنة في الدروس، وذلك بهدف ضمان تمثيل فقرات الاختبار لمفردات المحتوى تمثيلاً متكافئاً. (الكبيسي، ٢٠٠٧م)، (الحريري، ٢٠٠٧م).

رابعاً: إعداد جدول المواصفات:

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

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

خامساً: اختيار نمط أسئلة الاختبار:

اختيار النمط المناسب للاختبار يتوقف على مجموعة من العوامل ومنها:

- الهدف المراد قياسه: إذا كان الهدف الكشف عن الحقائق والنظريات فالأفضل أن يختار المعلم أسئلة الخطأ والصواب، ولقياس الأهداف التربوية المتقدمة كالفهم والاستيعاب فيفضل في هذه الحالة الأسئلة ذات الاختيار من متعدد ، بينما تستخدم الأسئلة المقالة لقياس الحقائق والأفكار المترابطة.
- طبيعة المحتوى المراد قياسه: من الممكن استخدام جميع أنماط الفقرات في أي مادة دراسية إلا أن بعض تلك الأنماط تكون أكثر ملائمة لمحتوى مادة معينة، كمادة النصوص الأدبية التي تناسبها الأسئلة المقالية، ومادة الرياضيات التي تناسبها الأسئلة الموضوعية وهكذا.
- مهارة المعلم في كتابة الأنواع المختلفة من الأسئلة: يتباين المعلمون في نوعية المهارات التي يمتلكها كل منهم في وضع الأسئلة، فمن يمتلك مهارة عالية يضع خليطاً من الأسئلة المقالية والموضوعية، أما من يفتقد إلى هذه المهارة ويتقاعس في بذل الجهد والوقت لإعداد الأسئلة الموضوعية الجيدة فإنه يضطر إلى وضع أسئلة مقالية.

- عمر التلميذ ومستواه التعليمي: يقل عدد الأسئلة لدى طلاب المراحل الدنيا، وذلك لأنهم يحتاجون إلى وقت أطول في الاستجابة للفقرة الواحدة مما يحتاجه طلاب المراحل العليا.
- الوقت المخصص لتحضيرها: يفضل وضع أسئلة الاختيار من متعدد عندما يكون الوقت كافياً، وعندما يكون الوقت غير متاح تستخدم الأسئلة المقالية وأسئلة التكميل والإجابة القصيرة.
- توفر الإمكانات المادية: ارتفاع تكاليف طباعة وتصوير الأسئلة الموضوعية بعكس الأسئلة المقالية، وبالتالي يتوقف استخدام الأسئلة الموضوعية، على توفر الإمكانات المادية. (الحريري، ٢٠٠٧م).

جميع العوامل السابقة ينظر إليها بصورة كلية، ثم يتخذ بعد ذلك القرار باختيار النمط المناسب للاختبار، أما في حالة تساوي تأثير هذه العوامل على أنماط مختلفة من الاختبارات فإن اختيار أنسبها يعتمد على الخصائص السيكومترية التي تتعلق بكل نمط من الاختبارات، وعلى معد الاختبار أن يختار النمط الأكثر صدقاً وثباتاً وتمييزاً مما يزيد من تحقيق الاختبار لأهداف القياس التربوي.

سادساً: مراجعة أسئلة الاختبار:

اقترح تنكلمان أن تتم مراجعة الأسئلة في ضوء ثلاثة معايير وهي(علام،٢٠٠٣)

- المحتوى المتعلق بالمادة الدراسية ومدى صحته وصحة نموذج الإجابة.
- الجانب التحريري وسلامة الصياغة، ومدى التناسق في الأسلوب والشكل مع بقية الأسئلة.
- الناحية الفنية المتعلقة بالجوانب القياسية في السؤال وخصائص نوعه.
- ومراجعة الأسئلة في ضوء هذه المعايير تتطلب توفر الخبرة في مجال القياس والتقويم .

سابعاً: ترتيب أسئلة الاختبار:

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

- نمط المفردة: إذا احتوى الاختبار على أكثر من نمط من أنماط الفقرات فيمن المفضل ترتيبها حسب النمط والفقرات ذات النمط الواحد مع بعضها البعض.
- صعوبة المفردة: ويتم ذلك بأن تتدرج الأسئلة في المجموعة الواحدة من السهل إلى الصعب، حيث يوفر هذا الترتيب الدافعية للطالب للاستمرار في محاولة الإجابة عن الأسئلة الأولى السهلة.

- المحتوى: ويقصد بهذا الترتيب تسلسل الفقرات في السؤال الواحد تسلسلاً منطقياً لمحتوى المادة الدراسية.
- مستوى الأهداف: وذلك بوضع الفقرات التي تقيس مستوى معيناً من الأهداف بشكل متسلسل، مع إمكانية ملاحظة أن مستوى الأهداف قد يستخدم لتدريج الفقرات بناء على صعوبتها .(الردادي، ٢٠٠٦م).

ثامناً: صياغة تعليمات الاختبار:

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

- التعليمات الخارجية: وهي ورقة منفصلة توضع فوق أوراق أسئلة الاختبار، وعدد الأسئلة وانواعها، وزمن الإجابة المخصص للاختبار، وطريقة ومكان الإجابة، وبعض الأمثلة التوضيحية وإمكانية الإجابة على بعض الأسئلة وترك البعض الآخر، وبيان عقوبة الغش.
- التعليمات الداخلية: وتوضع قبل كل سؤال، لتوضيح المطلوب من السؤال مع ذكر مثال لكيفية الإجابة.

تاسعاً: إخراج الاختبار وطباعته:

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

- أن تحتوي كراسة الاختبار على تعليمات الإجابة عليه.
- أن تكون طباعة الأسئلة واضحة خالية من الأخطاء المطبعية والإملائية.
- أن يراعي الفصل بين التعليمات والأسئلة.
- أن يراعي الفصل بين كل سؤال والذي يليه بمسافة معقولة.
- أن لا يجزئ السؤال على صفحتين متتاليتين.
- أن يفصل بين كل نوع أو شكل من أشكال الأسئلة والشكل الآخر بخط.
- أن يراعي الفصل بين مقدمة السؤال وبدلائله بمسافة معقولة.
- يراعي الترتيب المنطقي لأوراق الاختبار، وتثبيتها بشكل يسهل تناول الاختبار وتداوله (الردادي، ٢٠٠٦م).

عاشراً: تجريب الاختبار:

يطلق على التجربة الأولية للاختبار التجريب الاستطلاعي ، ويتم بتطبيق الاختبار على عينة صغيرة من التلاميذ بهدف التعرف على مدى وضوح التعليمات والأسئلة، ومدى مناسبتها لمستوى التلاميذ بشكل عام، والزمن اللازم للإجابة عليه، وبعد ذلك تجري التعديلات اللازمة.

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

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

أحد عشر: تطبيق الاختبار الفعلي:

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

والأفضل أن يعقد الاختبار في نفس المكان الذي كان يتلقى التلاميذ فيه دروسهم، مع ضرورة بعده عن الضوضاء والضجيج، وأن تتوفر فيها الإضاءة والتهوية الكافيتين، مع ضرورة تواجد معلم المادة الدراسية يوم الاختبار لتوفير نوع من الراحة النفسية والطمأنينة لدى الطلاب، ويجب عدم تشبب انتباه التلاميذ وتركيزهم ويستحسن تذكيرهم بالوقت المتبقي للاختبار بين الحين والآخر، وينبغي عدم التهاون بالغش بجميع أشكاله، لأنه يؤثر على مصداقية النتائج، ولذلك يجب وضع مسافات كبيرة بين التلاميذ عند تطبيق الاختبار ومع ضرورة المراقبة المشددة

الثاني عشر: تصحيح الاختبار:

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

أما إذا كانت الفقرات ذات استجابة منشأة فهناك طريقتان للتصحيح هما:

- الطريقة التحليلية، والطريقة الكلية، واستخدام أي منهما يعتم على عوامل كثيرة، كالهدف من الاختبار، والوقت المتوفر لدى المعلم للقراءة، ونوع السؤال فيما إذا كان محدداً أو مفتوحاً.
- وفي الطريقة التحليلية تجزأ الإجابة على عناصر، وتكون درجة السؤال هي مجموع درجات تلك العناصر، إذ أن المصحح أثناء قراءته يحدد الدرجة التي يستحقها كل عنصر من عناصر الإجابة.
- أما الطريقة الكلية فيتم إعطاء الدرجة للتلميذ بناء على الانطباع العام الذي تشكله إجابة التلميذ لدى المصحح، وتكون درجة الثبات متدنية لأن الدرجة تمنح حسب ذاتية المصحح والعوامل النفسية أو المزاجية الخاصة به.(الدوسري، ٢٠٠١م).

الثالث عشر: تحليل فقرات الاختبار:

ذكر الكبيسي (٢٠٠٧م) بان تحليل فقرات الاختبار يقصد بها استخراج معاملات الصعوبة والتمييز، وتحديد فاعلية البدائل (المموهات)".

ويذلك تكون أهم الخصائص الإحصائية لمفردات الاختبار:

- صعوبة المفردة ويرمز لها بالرمز (P).
- تمييز المفردة ، ويرمز بها بالمرمز (D).
- جاذبية البدائل (المموهات) ويرمز لها بالرمز (DA).

فوائد تحليل المفردات:

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

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

- تنمية مهارات المعلمين والقائمين بإعداد الاختبارات نتيجة التدرب والفحص المستمر للمفردات ، والتخلص من الأخطاء والعيوب التي تصيب الفقرة من حيث الصياغة أو الأخطاء الفنية الأخرى .
- الحصول على مفردات جيدة من حيث الصعوبة والتمييز والاحتفاظ بها وإعادة استخدامها في الاختبارات القادمة .
- التعرف على درجات الفقرات يحدد مواقع الفقرات في الاختبارات ووضع كل فقرة في المكان المناسب لها .
- تحديد فعالية البدائل وإعادة النظر في المموهات التي يختارها أفراد المجموعة العليا أكثر من الدنيا أو لم يختارها أحد من الطلبة .
- يمكن أن تكون خبرة تعليمية جيدة للطلبة إذا أسهم الطلبة في بحث وتحليل الفقرات ؛ لأنها تخلق لديهم نوعا من الدقة والانتباه في فحص البدائل قبل اختيار الجواب .

شروط الاختبار الجيد :

يرى كل من ثوراندیك وهيجن (١٩٨٦ م) وعودة (١٩٩٣م) أن الاختبار الجديد لابد أن يتصف بعدة صفات وهي :

الموضوعية والصدق والثبات (ونعني الخصائص السيكومترية للاختبار) والمعايير والقابلية للاستعمال .

أما القابلية للاستعمال فتشمل مجموعة من العوامل هي :

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

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

أولا : الثبات

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

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

ونلاحظ أن درجات الاختبار التي تتمنع بقدر مرتفع من الثبات وكانت تعكس الجوانب الحقيقية للسمة أو القدرة المراد قياسها ولا تعكس عوامل الخطأ العشوائية يزداد ثبات هذه الدرجات إذا قلت هذه العوامل وتوجد مصادر متعددة تسهم في أخطاء القياس وقد تؤدي إلى انخفاض ثبات درجات الاختبار بعضها يتعلق بأداة القياس ذاتها وبعضها يتعلق بإجراءات وظروف تطبيقها وبعضها الآخر يتعلق بالأفراد المختبرين ،.

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

أي أن الدرجة الملاحظة = الدرجة الحقيقية + درجة الخطأ .

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

$$\text{أي أن معامل الثبات} = \frac{\text{تباين الدرجات الحقيقية}}{\text{تباين الدرجات الملاحظة}}$$

ثانيا : الصدق : Validity

يعتبر أبو لبدة (١٩٨٥م) " الصدق أهم خاصية من خواص الاختبار ويعنى ذلك أن يقيس الاختبار بالفعل ما وضع لقياسه " .

والصدق يكون نسبيا أي أنه لا يوجد اختبار صادق بشكل كامل ولا يوجد اختبار غير صادق أيضا يكون نوعيا أي للغرض الذي أعد من أجله ضمن العينة والظروف المحيطة (عودة ، ١٩٩٣ م) .

الطرق المختلفة للتأكد من الصدق :

أولا : الصدق الظاهري (صدق المحكمين) facre validity

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

ثانيا : صدق المحتوى content validity

يدل صدق المحتوى على مدى تمثيل محتوى الاختبار للنطاق السلوكي الشامل للسمة المراد قياسها ويكون واضحا في الاختبارات التحصيلية بأنواعها (علام،٢٠٠٣)، نستطيع التأكد من صدق المحتوى في الاختبارات النفسية عن طريق صدق المحكمين بينما في الاختبارات التحصيلية فإننا نستخدم جدول المواصفات بعد تحليل المقرر الدراسي والتعرف على أهدافه .

ثالثا : صدق المحك crierion – relateqity valipity

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

أ- الصدق التلازمي

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

وكلما كانت قيمة معامل الارتباط (+ ٠.٥) أو أكثر دل على الصدق والأفضل أن يكون (+ ٠.٧) وكلما زاد عدد العينة كان أفضل (علام ، ٢٠٠٦م) .

ب – الصدق التنبؤي (المستقبلي)

يشير الصدق التنبؤي إلى مدى نجاح الاختبار وقدرته على التنبؤ بأداء الفرد في المستقبل ويتم حسابه بإيجاد معامل الارتباط بين درجات الاختبار الحالي ودرجات محك آخر بعد فترة زمنية مناسبة مثلا كاختبار القدرات أو اختبارات القبول في الجامعات والمعدل التراكمي للطالب .

ويعتبر معامل الصدق التلازمي قيمة قصوى لمعامل الصدق التنبؤي (علام، ٢٠٠٣م) .

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

١ – النظرية والبحوث التجريبية .

يتم تصميم الاختبار في ضوء نظرية معينة يتم التأكد من صحة فروض هذه النظرية بالبحوث التجريبية .

التحليل المتعدد :

مقارنة نتائج الاختبار الحالي بنتائج البحوث والدراسات السابقة التي تناولت نفس الموضوع وإذا وجد اتساق فيما بينها كان ذلك دليلا على الصدق.

طريقة تطور الأعمار .

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

طريقة الارتباطات وتنقسم إلى ثلاث أقسام :

١- الصدق العاملي :

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

٢- المقارنة الطرفية (الاتساق الداخلي) :

يتم تطبيق الاختبار المراد حساب صدقة ثم نحسب الدرجة الكلية للاختبار ، ثم نحصل على مجموعتين ٢٧% أدنى ، ثم نحسب الفروق بين المجموعتين المرتفعة والمنخفضة في درجات مفردات الاختبار عن طريق اختبار (ت) ويوجد احتمالين الأول وجود فروق ذات دلالة إحصائية بين المجموعتين ويكون الاختبار بذلك صادقا لأنه استطاع التمييز بين المجموعتين ، والثاني عدم وجود فروق ذات دلالة إحصائية بين المجموعتين يكون الاختبار غير صادق لأنه لم يستطيع التمييز بين المجموعتين . (إسماعيل ، ٢٠٠٧ م) .

الصدق التقاربي والتمييزي (التمايزي) .

إذا ارتبط الاختبار ارتباطا عاليا أو مرتفعا بالمتغيرات التي يجب أن يربط بها نظريا أو تتسق معه يسمى هذا صدق تقاربي .

أما إذا ارتبط الاختبار ارتباطا منخفضا بالمتغيرات التي يجب أن تربط بها نظريا أو لا تتسق معه فهذا صدق تمييزي (تمايزي) (إسماعيل ، ٢٠٠٧ م) .

ج – الصدق التجريبي :

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

الدراسات السابقة :

لقد برز بشكل واضح أهمية هذا الموضع من خلال دراسات عديدة منها ما هو دراسات عربية ومنها دراسات أجنبية وفيما يلي عرض لأهم تلك الدراسات :

أولاً : الدراسات العربية :

وتم تقسيمها إلى قسمين دراسات ذات علاقة مباشرة بموضوع الدراسة الحالية ودراسات ذات علاقة غير مباشرة بموضوع الدراسة وفيما يلي عرض لها :

١- دراسات ذات علاقة مباشرة بالموضوع :

دراسة (علام، ١٩٨٠م) والتي هدفت إلى بناء وتقدير صدق محتوى اختبار تشخيصي مرجعي الميزان، لقياس مهارة المعلمين لبناء الاختبارات المدرسية، وقد تم بناء صورتين متكافئتين من الاختبار، كل صورة تتكون من ٩١مفردة قائمة على ٦٩هدفاً إجرائياً، ويتكون كل اختبار من خمسة اختبارات فرعية يقيس كل منها إحدى المهارات التالية:

- مهارة صياغة الأهداف التعليمية.
- مهارة تصميم خطة الاختبار.
- مهارة بناء وتصحيح مفردات الاختبار.
- مهارة تحليل البيانات الناتجة عن تطبيق الاختبار.
- مهارة تقرير نتائج الاختبار.

وقد بينت دراسة (علام، ١٩٨٠م) حاجة المعلم لتطوير الكفايات الأساسية في بناء الاختبارات المدرسية.

دراسة المركز العربي للبحوث التربوية لدول الخليج (١٩٨١م) تحت عنوان (دراسة مقارنة لأساليب التقويم والامتحانات في دول الخليج العربي) ، كان الهدف منها هو إجراء مقارنة لواقع أساليب التقويم في الدول الخليجية بالإضافة إلى وضع بعض النماذج لتطوير واقع أساليب تقويم

الطلبة، وقد تم جمع البيانات لهذه الدراسة بواسطة تحليل الوثائق الرسمية الخاصة بالتقويم والامتحانات في الدول الأعضاء، بالإضافة إلى ثلاثة استبيانات ثم توجيهها إلى المسئولين والمدرسين ومدراء المدارس، إذ تضمن الاستبيان الأول فقرات حول مهام إدارات التقويم وتنظيمها وأساليب التقويم المختلفة، أما الاستبيان الثاني فقد تضمن أنواع الاختبارات وكيفية بناءها، أما الاستبيان الثالث فقد تضمن الجوانب التي تم تقويمها في شخصية الطالب، بالإضافة إلى الفترات التي تتم فيها عملية التقويم ومن هذه الاختبارات هي الاختبارات الشفوية والتحريرية والعملية بنسب مختلفة حسب المراحل الدراسية ، وقد أوصت الدراسة. بتنوع أساليب التقويم بحيث تركز على الجوانب المختلفة في شخصية الطالب والاهتمام بإعداد الاختبار التحصيلي لكي يظهر الفروق الفردية بين الطلبة.

دراسة الخياط، و ذياب (١٩٩٦) والتي استهدفت الكشف عن الجوانب الإيجابية والسلبية في نظام تقويم كفاءة المعلم حول مدى تحقيق نظام التقويم للأهداف التي وضع من أجلها، وللتعرف على المعوقات التي تواجه تنفيذه.وطبقت استبانة لاستطلاع آراء العاملين في الحقل التربوي من معلمين وإداريين وموجهين حول هذا النظام على عينة بلغت (٥٢٣) فرداً، وبتحليل بيانات الدراسة توصلت إلى عدد من النتائج منها:

- المسؤولية التربوية الرئيسة لنظام تقويم كفاءة المعلم هي التثبت من مدى تحقيق المعلمين للأهداف المرجوة من المنهج الذي يدرسونه.
- أهم بنود استطلاع الرأي في قياس الكفاءة هو المحافظة على الدوام والالتزام بأخلاقيات المهنة.
- وجود معارضة بين أفراد العينة على إدراج جدول العقوبات والخصومات ضمن التقرير السنوى للمعلم.

وأوصت الدراسة بإعادة النظر في هذا النظام، وضرورة مراعاة خصائص المرحلة التدريسية في بناء بطاقة التقويم.

دراسة عفانة (١٩٩٨) استهدفت تحديد الكفاءات التدريسية التي يمارسها أساتذة الجامعة الإسلامية بغزة كما يراها طلبة الجامعة. وبلغت عينة الدراسة (٣٢١) طالباً وطالبة. وأظهرت النتائج قصوراً في الكفايات التي يمارسونها، حيث وصلت إلى (٣٦) كفاية من أصل (١٠٠) كفاية. كما بينت الدراسة وجود فروق ذات دلالة إحصائية في بعض الكفايات التدريسية ترجع إلى الجنس ونوع الكلية التي ينتمي إليها الطلاب.

دراسة الحكمي (١٤٢٣هـ) هدفت هذه الدراسة إلى إعداد معيار للكفاءات المهنية المتطلبة للأستاذ الجامعي، ومعرفة أكثر الكفاءات المهنية تفضيلاً لدى الأستاذ الجامعي من

وجهة نظر الطلاب بجامعة أم القرى فرع الطائف كما هدفت إلى الكشف عن المتغيرات التي يمكن أن يكون لها تأثير في الأحكام الصادرة من الطلاب على الكفاءة المهنية المتطلبة لمعلمهم ، ومن اهم النتائج أنه تتمحور الكفاءات المهنية المتطلبة للأستاذ الجامعي من وجهة نظر الطلاب حول ست كفاءات رئيسية هي (الشخصية، والإعداد للمحاضرة وتنفيذها، والعلاقات الإنسانية، والأنشطة واساليب التقويم، والتمكن العلمي والنمو المهني، وأساليب الحفز والتعزيز).

دراسة النور(٢٠١٤) هدفت الدراسة الى تقويم مستوى تطبيق اعضاء هيئة التدريس في اللغة العربية بجامعة السودان للعلوم والتكنلوجيا لمعايير تصميم الاختبار التحصيلي الجيد من وجهة نظر الطلبة ومتغيرات الجنس والمستوى والتخصص الدراسي والمعدل التراكمي ، واظهرت النتائج ان مستوى تطبيق اعضاء هيئة التدريس كان متوسطاً.

دراسة الغزيوات(٢٠٠٥) حيث هدفت الدراسة لمعرفة كفايات أعضاء هيئة التدريس في جامعة مؤتة الذين يدرسون طلاب الدراسات الجامعية من وجهة نظر الطلاب أنفسهم في كلية العلوم التربوية ، وكشفت الدراسة على أن الطلاب غير راضين عن استخدام أعضاء هيئة التدريس لطرق التقويم والتدريس التقليدية والتعامل غير الإنساني والتعصب في الرأي وأوصت الدراسة بضرورة اتباع طرق التقويم والتدريس الحديثة والتعامل بشكل إنساني مع الطلاب وتنمية العلاقات البين شخصية بين الطرفيين .

٢- دراسات ذات علاقة غير مباشرة بالموضوع :

دراسة العاني (١٩٨٠م) والتي هدفت إلى بناء معايير علمية يفترض أن تقوم على أساسها وفي ضوئها الاختبارات المدرسية، وإلى تحليل أسئلة الاختبارات النهائية للصف السادس الابتدائي في جميع مدارس بغداد ولجميع المواد الدراسية(اللغة العربية، التربية الدينية ، اللغة الإنجليزية، العلوم الرياضية، العلوم الاجتماعية، العلوم العامة) وقد حددت الدراسة بعض المعايير لاعتمادها في تحليل الاختبارات النهائية وذلك للوقوف على واقعها النوعي والعلمي ومدى أبعاد او اقتراب الأسئلة المقدمة للطلبة من هذه المواصفات اللازم توافرها في الاختبارات المدرسية وهذه المعايير:

١- الشمول. ٢- صياغة الأسئلة.

٣- الأسئلة المباشرة وغير المباشرة. ٤-استقلالية السؤال.

٥- التوافق بين توزيع الدرجات ونسب الشمول.

٦- التوافق بين التوزيع الزمني للإجابة ونسب الشمول.

٧- نسبة التطابق مع أسئلة الكتاب المدرسي ٨- النواحي الفنية.

وقد أظهرت نتائج الدراسة أن الأسئلة الاختيارية المقدمة للطلبة قد فشلت فشلا كبيراً في تقديمها أسئلة اختباريه جيدة يمكن أن يقاس تحصيل الطالب عن طريقها، كما أنها كانت بعيدة عن أقل النسب المعتمدة في معايير قياس التحصيل وتقويمه.

دراسة المركز العربي للبحوث التربوية لدول الخليج (١٩٨٥م) تحت عنوان ((تطوير أساليب تقويم الطلبة بمراحل التعليم العالم بالدول الأعضاء ، دراسة تحليلية)) والوقوف على بعض الاتجاهات العالمية المعاصرة في مجال التقويم والاستفادة منها في تطوير أساليب تقويم الطلبة ولتحقيق أهداف الدراسة، تم تصميم استمارتي معلومات حول الأساليب المستخدمة في تقويم الطلبة وجهت إحداها إلى المشرفين التربويين والأخرى إلى المسئولين عن إدارات التقويم والامتحانات بالدول الأعضاء، كما تضمنت الدراسة تحليل عينات من الاختبارات التي يعدها المعلمون لطلبتهم في الصفوف الدراسية بالدول الأعضاء في ضوء بعض المحكات، وقد توصلت الدراسة إلى العديد من النتائج منها.

١- التركيز على الجوانب المعرفية وخاصة في مستوياتها الدنيا في تقويم الطلبة وإهمال العمليات العقلية العليا من الجانب المعرفي.

٢- الاختبارات الأكثر استخداماً في عملية التقويم هي الاختبارات الشفوية والمقالية في جميع المواد وقلة استخدام الاختبارات الموضوعية.

٣- قلة استخدام الجوانب الوجدانية والمهارية في معظم المواد الدراسية، وقد أوصت هذه الدراسة على تدريب المعلمين على كيفية تحديد وصياغة الأهداف السلوكية، وكذلك تدريبهم على بناء واستخدام وتحليل نتائج أدوات التقويم المتعددة، والإعداد لإنشاء بنوك للأسئلة في مختلف المواد الدراسية لتطوير أساليب التقويم.

دراسة (جرادات، ١٩٨٨م) : والتي هدفت إلى التعرف على مدى معرفة معلمي ومعلمات العلوم للمرحلة الإعدادية بكفايات بناء الاختبارات المدرسية، وممارستهم لها في ضوء متغيرات الدرجة العلمية والخبرة والجنس، وقد أظهرت الدراسة نتائج هامة من بينها أن معرفة المعلمين وممارستهم لكفايات بناء الاختبارات المدرسية لم تصل إلى المستوى المقبول تربوياً، وأن معرفة المعلمين من حملة البكالوريوس فما فوق أعلى من معرفة المعلمين حملة الدبلوم ، وقد فسر الباحث تدني مستوى الكفايات لدى المعلمين وممارستهم لها، وعلى أن المعلومات المعطاة للمعلمين عن الاختبارات أثناء دراستهم للجامعة أو كليات المجتمع غير كافية ، مما يشير إلى تطوير برنامج تأهيلي يفي بحاجات المعلمين في هذا المجال قبل أو أثناء الخدمة.

دراسة القرشي(١٤١٦هـ) هدفت إلى التعرف إلى مدى توفر القواعد الأساسية اللازمة في بناء الاختبارات الموضوعية لدى معلمي الرياضيات بالمرحلة المتوسطة وقد خلصت هذه الدراسة إلى مجموعة من النتائج من أهمها :

- توفر ثلاث قواعد فقط لدى المعلمين عينة البحث من أصل أحد عشر قاعدة متعلقة ببناء اختبار الصواب والخطأ.
- توفر خمس قواعد فقط لدى المعلمين عينة البحث من أصل ثمان عشر قاعدة متعلقة ببناء اختبار الاختبار من متعدد.
- توفر قاعدة واحدة فقط لدى المعلمين عينة البحث من أصل خمس قواعد متعلقة ببناء اختبار المزاوجة.
- توفر قاعدة واحدة فقط لدى المعلمين عينة البحث من أصل ست قواعد متعلقة ببناء اختبار الإكمال والإجابة القصيرة.

دراسة الزهراني(١٤١٨هـ) هدفت إلى الكشف عن إمكانيات معلمي العلوم العلمية والعملية في المرحلة المتوسطة في مجال بناء الاختبارات ومدى إتقانهم لها وكذلك الكشف عن مدى ممارستهم لها في إعدادهم للاختبارات المدرسية، وكان من نتائجها:

- إن مستوى إتقان وممارسة معلمي العلوم للمرحلة المتوسطة لمهارات بناء الاختبارات المدرسية لم يصل إلى درجة القطع المحددة، أي أن هناك تدني في مستوى الإتقان والممارسة .
- مستوى إتقان حملة البكالوريوس أعلى من حملة الدبلوم.
- مستوى إتقان ذوي الخبرة القصيرة أعلى من ذوي الخبرة المتوسطة ومستوى إتقان ذوي الخبرة المتوسطة أعلى من ذوي الخبرة الطويلة.
- ضعف العلاقة الارتباطية بين كمية الملعومات المتوفرة لدى المعلم ودرجة ممارسته لها وأن هذه العلاقة تشير إلى أنه كلما كان مستوى المعرفة عاليا كلما كانت هناك ممارسة فعلية لهذه المعرفة في بناء الاختبارات.

ثانياً : الدراسات الأجنبية :

دراسة كيربر(Kerber1976) والتي هدفت لتحديد أثر كل من الخبرة والمؤهل الدراسي والتدريب على معرفة المعلمين بمفاهيم القياس والاختبارات، وإلى أي مدى يختلف المعلمون في الخدمة عن معلمي المستقبل، وقد توصلت الدراسة إلى أن معرفة المعلم لمفاهيم القياس تزداد بزيادة الخبرة والمؤهل والتدريب وعدد المقررات التي يتلقاها المعلم، كما أظهرت نتائج الدراسة أنه لا توجد فروق ذات دلالة إحصائية بين المعلمين في الخدمة ومعلمي المستقبل.

دراسة (Wesman,1971) وتوصلت إلى مجموعة من الإرشادات والقواعد تساعد معلمي الرياضيات في بناء اختبارات الاختبار من متعدد، وقد عرضت العديد من الأمثلة السيئة والأمثلة الجيدة من مفردات هذا النوع من الاختبارات الموضوعية ، وكذلك أورد في بعض الحالات التغيرات المطلوبة لإصلاح الخلل الوارد في الأمثلة السيئة.

دراسة (Tamir.1990) ووصفت هذه الدراسة كيفية بناء مفردات اختبار الاختبار من متعدد والأخذ بالمميزات الأساسية لهذا النوع من الاختبارات، حيث نوقشت فيها عدة نقاط منها:

الإجابات الصحيحة مقابل الأفضل – شروط تحليل مفردات اختبار الاختبار من متعدد– مشكلة التخمين– المفردات الموجبة مقابل المفردات السالبة.

التعليق على الدراسات السابقة:

من خلال استعراض الدراسات والبحوث السابقة لاحظ الباحث الآتي:

١- إن جميع الدراسات السابقة تشترك مع الدراسة الحالية في اهتمامها بطرائق التقويم و بقواعد بناء الاختبارات والكفاءات التعليمية في إعداد وتصحيح وتحليل الاختبارات ، أي أنها جميعها تهتم بأساليب التقويم. فدراسة جرادات(١٩٨٨م) ودراسة الحميضي (١٤١٤هـ) ودراسة الزهراني (١٤١٨هـ) وقد هدفت إلى التعرف على مدى معرفة المعلمين بكفايات بناء الاختبارات بصفة عامة.

٢- أما دراسة (القرشي ١٤١٦هـ) فهدفت إلى التعرف على مدى توفر قواعد بناء الاختبارات الموضوعية ، واهتمت دراسة (Wesman,1971) ودراسة (Tamir,1990) بكيفية بناء الفقرات الاختيارية من نوع الاختيار من متعدد.

٣- اتفقت نتائج بعض الدراسات السابقة على وجود تدني في مستوى المعرفة بالقواعد الأساسية اللازمة لبناء الاختبارات لدى المعلمين ، مثل دراسة جرادات (١٩٨٨م) ودراسة الحميضي (١٤١٤هـ) ودراسة القرشي(١٤١٦هـ) ودراسة الزهراني (١٤١٨هـ).

٤- بحثت بعض الدراسات السابقة مدى معرفة المعلمين بقواعد بناء الاختبارات في ضوء متغيرين هما الخبرة والمؤهل فقط مثل دراسة الزهراني (١٤١٨هـ) أو في ضوء ثلاثة متغيرات هي الخبرة والمؤهل والتخصص مثل دراسة الحميضي (١٤١٤هـ) أو في ضوء ثلاثة متغيرات مختلفة هي الخبرة والمؤهل والجنس مثل دراسة جرادات (١٩٨٨م).

٥- ركزت غالبية الدراسات التي بحثت في مدى معرفة المعلمين بقواعد بناء الاختبارات على مرحلة واحدة فقط من مراحل التعليم.

٦- أهملت الدراسات السابقة البرامج التدريبية والدورات التي حضرها أو تلقاها المعلم أثناء الخدمة ولم تدخلها ضمن متغيرات الدراسة.

٧- تختلف هذه الدراسة عن الدراسات السابقة في أنها بحثت تقويم الكفاءات التعليمية لدى الاستاذ الجامعي في بناء الاختبارات التحصيلية.

٨- كان أحد أهم الجوانب في كفايات عضو هيئة التدريس هو كفاية التقويم والتي تفرد البحث الحالي بالتعامل معها بشكل مستقل وأكثر عمقاً .

٩- كثير من الدراسات أجري في الوطن العربي مما يبين مزيد الاهتمام والعناية التي يلقاها هذا الجانب من الباحثين .

١٠- ركزت الدراسات السابقة على أهمية كفاية التقويم وهو ما استفاد منه الباحث في الدراسة الحالية

١١- كشفت غالبية الدراسات عن تدني مستوى الكفاءات التدريسية للمعلم الجامعي، حيث لم تصل في أحسن حالاتها إلى نسبة ٥٠% من أي قائمة للكفاءات قد تم استخدامها في تلك الدراسات، وأن الأستاذ الجامعي لم يصل بعد في مستوى أدائه إلى درجة الكفاية المتوقعة منه.

اجراءات البحث:

مجتمع البحث:

يسعى الباحث في هذه الدراسة الى استطلاع آراء طلاب جامعة الطائف المسجلين في جميع كليات الجامعة المختلفة – وهم طلاب البكالوريوس وطلاب الدبلوم التربوي وطلاب الدراسات العلياء- حول الوضع القائم في بناء الاساتذة للاختبارات التحصيلية.

العينة الاستطلاعية لأداة البحث:

حيث ان البحث الحالي يسعى لبناء أداة لتقويم الاستاذ الجامعي في بناء الاختبارات التحصيلية ، لذا وجب على الباحث تجريب الاداة على عينة استطلاعية من المجتمع الاصلي. وقد اختار الباحث شعبة من طلاب كلية التربية بجامعة الطائف في برنامج الدبلوم التربوي في العام الدراسي ١٤٣٨/١٤٣٧ الفصل الدراسي الثاني ، كان عددهم ٣٧ طالب يدرسون مقرر القياس والتقويم ، وتم توزيع الاداة بهدف التعرف على العبارات الغامضة وكذلك معرفة استيعاب الطلاب للألفاظ العبارات المكتوبة في الأداة وتعديل مايمكن تعديلة والخروج بالصيغة النهائية لعبارات الأداة.

أداة البحث وخطوات اعدادها:

أعد الباحث استبانة "تقويم الاختبارات التحصيلية لدى أساتذة الجامعات"، وقد استند في بنائها على مجموعة من الدراسات والمراجع التي بحثت في التدريس الجامعي، من أهمها: "معايير لتقويم جودة التعليم لدى المدرسين في الجامعات والمعاهد العليا" و"أساليب التقويم الجامعي".و اشتملت الاستبانة على بعدين هما : الشكل والإخراج العام والمحتوى ومستوى الاسئلة وقد بلغ عدد فقرات الاستبانة ٣٢ فقرة، موزعة على البعدين على التوالي: ١٦ فقرة للبعد الاول، ١٦فقرة للبعد الثاني، . وقد اختار الباحث سلَّماً تقديرياً مكوناً من خمسة خيارات هي: دائما ، غالباً، أحياناً ، نادراً، ابداً. وقد أعطيت هذه الخيارات القيم 5 ،4 ،3 ،2 ، 1على التوالي. وقد قام الباحث بإجراءات حساب صدق وثبات الأداة، حيث تم عرضها على مجموعة من أساتذة الجامعة من حملة الدكتوراه في التربية وعلم النفس، بلغ عددهم ٥ أساتذة ، وقد تم الاستفادة من توجيهات المحكمين وآراءهم في ابعاد وعبارات المقياس مما اخرج للباحث اداة تتسم بدرجة مقبولة من الصدق المتفق عليه من قبل المحكمين لقياس ما اعدت لقياسه.

الخصائص السيكومترية لأداة الدراسة:

الصدق:

تم عرض أداة الدراسة على خمسة من الاساتذة الجامعيين لتحكيم الأداة وقد اتفق جميع المحكمين على تمثيل الاداة للغرض الذي اعدت من أجله مع تعديل بعض الصياغات لبعض العبارات، وهذا يعطي الباحث الثقة في الاداة لاعتبرها صالحة لقياس قدرات الاساتذة الجامعيين في بناء الاختبارات التحصيلية.

بعد ذلك تم تطبيق الاداة على عينة استطلاعية قدرها ٣٧ طالبا من احدى الشعب التي تدرس مقرر القياس والتقويم بكلية التربية في جامعة الطائف ،في الفصل الثاني للعام الدراسي ١٤٣٨/١٤٣٩هـ وقد حصل الباحث على مايلي:

معامل الاتساق الداخلي للاستئناس به كمؤشر لصدق الأداة وكانت قيمته على النحو التالي

جدول(١)

معامل الاتساق للابعاد مع المجموع الكلي للعبارات المقياس

قيم الاتساق الداخلي	البعد الاول	البعد الثاني	المحموع الكلي
Pearson Correlation (Sig. (2-tailed البعد الاول N	1 37	.560** ,000 37	.902** .000 37
Pearson Correlation (Sig. (2-tailed البعد الثاني N	.560** ,000 37	1 37	.٨٦٣** .000 ٣٧
Pearson Correlation (Sig. (2-tailed المجموع N	.902** .000 37	.٨٦٣ ** .000 ٣٧	١ ٣٧

من الجدول (١) يتضح ان معامل الارتباط بين البعد الاول ومجموع عبارات الاداة كان 902. وهو معامل قوي عند مستوى معنوية 000. ، كذلك نجد ان معامل الارتباط بين البعد الثاني ومجموع عبارات الأداة ايضاً معامل قوي عند مستوى معنوية000. وبالمجمل نجد ان معامل ارتباط العبارات ككل هو ٥٦. ، وهو مؤشرمناسب لحد ما للحكم بصدق الأداة ولتحقيق الهدف المحدد من الدراسه.

كذلك قام الباحث بحساب القيمة الحرجة عن طريق حساب قيمة (t) بواسطة المقارنة الطرفية بين المجموعة الاعلى أداءً على المقياس والمجموعة الاقل أداءً وكانت على النحو التالي

جدول(٢)

المتوسط والانحراف المعياري للمجموعتين العليا والدنيا

المجموعتين الطرفية	العدد	المتوسط	الانحراف المعياري	خطأ الانحراف عن المتوسط
العليا	19	95.0000	4.96655	1.13941
الدنيا	18	78.2222	3.82800	.90227

جدول(٣)

قيمة(ف) ومستوى الدلالة وقيمة اختبار(ت) ومستوى المعنوية

f	sig	t	Df	Sig. (2-tailed)
١.٧٤٣	.195	11.462	35	.000

يتضح من الجدول (٣) وجود فروق معنوية بين درجات (المجموعة الأعلى) ودرجات (المجموعة الأدنى) وكانت الفروق بين المجموعتين ذات دلالة إحصائية عند مستوي معنوية ٠.٠٠٠ مما يدل على قدرة الاداة في ايجاد الفروق بين اصحاب القدرات العليا واصحاب القدرات الدنيا مما يعطينا مؤشر جيد بصدق الاداة في قياس ما اعدت لقياسه.

الثبات:

تم حساب الثبات عن طريق معامل الفا كرونباخ بعد تطبيق الأداة على العينة الاستطلاعية من الطلاب وكان معامل الثبات يساوي ٠.٧٢ بمعامل الفا كرنباخ ، كذلك تم حساب الثبات بمعامل التجزئة النصفية وكانت النتيجة على النحو التالي

جدول(٤)
معامل الثبات

Cronbach's Alpha	Part 1	Value	.655
		N of Items	16[a]
	Part 2	Value	.606
		N of Items	16[b]
	Total N of Items		32
Correlation Between Forms			.561
Spearman-Brown Coefficient	Equal Length		.719
	Unequal Length		.719
Guttman Split-Half Coefficient			.708

من الجدول (٤) يتضح أن معامل الثبات بالتجزئة النصفية يساوي 0.71.

وهم معامل جيد يوحي بثبات الاداة واستقرارها لتطبيقها بشكل اوسع على عينات اخرى، وبذلك أخذت الاداة صورتها النهائية.

التوصيات والمقترحات

- يوصي الباحث باستخدام هذه الاداة تزامننا مع تقديم الاسئلة التي يعدها الاستاذ الجامعي ثم عمل التحليل الاحصائي لفقراتها لمعرفة مدى تحقيق اسئلة الاستاذ الجامعي للاشتراطات الفنية للاختبارات التحصيلية.
- استخدام هذه الاداة مع غيرها من الادوات لتقويم الاستاذ الجامعي في الجوانب المختلفة.
- اقترح على الاساتذة ان يستخدموا هذه الاداة مع كل اختبار يقدمونه لطلابهم كأسلوب تقويم ذاتي.
- عمل دراسات تقوم على هذه الاداة من اجل اعادة احتساب الخصائص السيكومترية للأداة على عينات مختلفة من الطلاب.
- إقامة حلقات نقاشية تتناول قضايا تقويم الطلاب والصعوبات المتعلقة بذلك وكيفيه التعامل مع ألطلبه وفق الطرق الموضوعية في التقويم.
- إعداد برامج تدريبيه في بناء الاختبارات واساليب التقويم المختلفة لاعضاء هيئة التدريس.

قائمة المراجع

أبو باسل، محمد عبد الكريم (٢٠٠٢م): قياس وتقويم تعلم الطلبة ، دار الفرقان : عمان.

ابو زينة فريد كامل (١٩٩٨م) : أساسيات القياس والتقويم في التربية . ط٢ ، الكويت :مكتبة الفلاح .

أبو علام،رجاء محمود(١٩٨٧ م): قياس وتقويم التحصيل الدراسي، الكويت:دار القلم

أبو لبدة سبع محمد (١٩٨٥ م) : مبادئ القياس النفسي . ط٣ كلية التربية ، الجامعة الأردنية ، عمان .

علام ، صلاح الدين محمود (٢٠٠٠م) : القياس والتقويم التربوي (أساسياته وتطبيقاته وتوجيهاته المعاصرة) . ط١ ، القاهرة : دار الفكر العربي .

علام ، صلاح الدين محمود (٢٠٠٦م) : الاختبارات والمقايس التربوية والنفسية . الطبعة الأولى ، عمان: دار الفكر .

القرشي،عبد الفتاح(١٤٠٦هـ):"اتجاهات جديدة في أساليب تقويم الطلاب" كتب التربية العربي لدول الخليج العربي،رسالة الخليج العربي،السنة السادسة، العدد الثامن عشر.

القرني ، ناصر بن صالح (١٤٢١ هـ) : دليل المعلمين والمعلمات في تقويم الاختبارات التحصيليه. الطبعة الثانية ، الرياض:الإدارة العامة للاختبارات .

الكبيسي ، عبد الواحد حميد (٢٠٠٧ م) : القياس والتقويم . عمان : دار جرير .

الكبيسي ، عبد الواحد وربيع ، هادي مشعان (٢٠٠٨ م) : الاختبارات التحصيلية المدرسية . الأردن : مكتبة المجتمع العربي .

محمد، فتحية وسليمان، سعيد (١٩٩٣م): برنامج تدريبي مقترح لإعداد المعلم الجامعة. مجلة كلية التربية، جامعة الأسكندرية، المجلد السادس، العدد الأول.

مراد ، أحمد صلاح وسليمان ، أمين على (٢٠٠٢ م) : الاختبارات والمقاييس في العلوم النفسية والتربوية . دار الكتاب الحديث

النور، احمد يعقوب(٢٠١٤):تقويم مستوى تطبيق اعضاء هيئة التدريس بجامعة السودان للعلوم والتكنولوجيا لمعايير تصميم الاختبار التحصيلي الجيد كم وجهة نظر الطلبة وبعض المتغيرات الاخرى. المجلة التربوية المتخصصة، المجلد(٣)،العدد(٥)-ايار،٢٠١٤.

منتدى الفكر العربي، وبرنامج الأمم المتحدة الإنمائي، "التعليم العالي في البلاد العربية : السياسات والآفاق"، تحرير فاتن البستاني، سلسلة الحوارات العربية ، ١٩٩٧،ص ٤١ ،١٢٦.

Tamir, P (1990) 'Some issues related to the use of justifications to multiple choice items', J Biol Educ 24, 285-292

Thorndik,R.M.(2005).Measurement&Evaluation in Psychology saddle River.NJ :pearson Education pinc.

Wesman, A (1971) 'Writing the Test Item', in Thorndike R L (Editor) Educational Measurement, American Council of Education, Washington, DC, pp 81-129