

تطوير الاختبارات التحصيلية بكلية التربية جامعة دمنهور في ضوء معايير الجودة

د. حنان أحمد يوسف كرسون

دكتوراه علم النفس التربوي

أخصائي بوحدة القياس والتقويم بكلية التربية ،جامعة دمنهور

hanankarson748@gmail.com

مستخلص البحث

هدف البحث الحالي إلى إعداد قائمة بمعايير الجودة لإختبارات المقررات التربوية بكلية التربية، جامعة دمنهور، وتقديم معايير تساعد المعلمين في تقييم الإختبارات وتطويرها في ضوء معايير الجودة ، مع تقديم تقرير عن تطوير الإختبارات للمقررات التربوية بكلية التربية ، جامعة دمنهور، واشتملت أدوات البحث على قائمة لمعايير جودة الإختبارات للمقررات التربوية ، واستخدمت الباحثة المنهج الوصفي ، وتكون مجتمع البحث الحالي من جميع أسئلة الامتحانات النهائية النظرية التي وضعها المعلمون بكلية التربية ، وتم جمع (٨٠) اختبار من اختبارات العلوم التربوية والنفسية بكلية التربية ، جامعة دمنهور من عام (٢٠٢٠:٢٠٢٤) بواقع (٢٠) اختبار لكل عام جامعي ، واستخدمت الباحثة استمارة خاصة لإستخدامها معايير لجودة الإختبارات وذلك لتحليل الإختبارات، حيث تكونت هذه الاستمارات من جميع مستويات المجال المعرفي تبعًا لهم بلوم مقابل كل مجموعة من الأفعال والمحتويات التي تصفها ، وتوصلت الباحثة أنه يوجد تحسن في نسب استيفاء الاختبار لمعايير جودة الاختبار بداية من عام (٢٠٢١) ، وكذلك تجاوز نسب استيفاء المعايير الأساسية وتعليمات الاختبار ولغة الاختبار (٩٠%) ، وأيضًا يُعد معيار لغة الاختبار أعلى نسبة لمعايير جودة الاختبار على مدار أربعة أعوام ، بينما انخفضت نسبة استيفاء الاختبار لمعيار ارتباط الاختبار بالتوصيف ومضمون الاختبار عن (٩٠%).

الكلمات المفتاحية : التقويم ، الأسئلة ، الإمتحانات ، معايير الجودة

Abstract

The current research aimed to prepare a list of quality standards for educational course tests at the Faculty of Education, Damanhour University, and to provide standards that help teachers in evaluating and developing the tests in light of quality standards, along with presenting a report on the development of tests for educational courses at the Faculty of Education, Damanhour University. The research tools included a list of quality standards for educational tests, and the researcher used the descriptive approach. The study population consisted of all the final exam questions prepared by the teachers at the Faculty of Education. A total of (80) exams were collected from educational and psychological sciences tests at the Faculty of Education, Damanhour University, for the years (2020:2024), with (20) exams for each academic year. The researcher used a special form to apply quality standards for the tests in order to analyze the examinations. These forms were composed of all levels of the cognitive domain according to Bloom's taxonomy for each group of actions and the content it describes. The researcher found that there has been an improvement in the test's compliance rates with quality standards starting from 2021, as well as exceeding the basic compliance rates and test instructions and test language (90%). Additionally, the standard for test language has the highest percentage of quality standards for tests over four years, while the compliance rate of the test with regard to the description and content of the test decreased to below 90%.

Keywords: Evaluation, Questions, Exams, Quality Standards

مقدمة

مما لا شك فيه أن منظومة التعليم كنظام تعليمي اجتماعي يتأثر بما يدور حوله من تغيرات في جوانب الحياة المختلفة ، ولذا فنجاح المؤسسة التعليمية كمؤسسة اجتماعية وتربوية وأكاديمية يكمن في قدرتها على الأخذ بأساليب التطوير للعملية التعليمية مما يحقق التوازن بين ما تقوم به من أدوار ، وما يسود في المجتمع من أفكار ومعتقدات واتجاهات لرفع القدرة التنافسية وفتح آفاق جديدة للمعرفة في كافة المجالات المختلفة.

فعملية تقويم الإمتحانات عملية عامة وتقويم الطلاب عملية خاصة لتحقيق الجودة الشاملة لنظام العملية التعليمية وبخاصة في مرحلة التعليم الجامعي ، فإصلاح التقويم ونظم الإمتحانات يعتبر مدخلاً مهماً من مداخل اصلاح العملية التعليمية وجودة التعليم لا يتحقق إلا من خلال طرق تقويم نسبهم في قياس نواتج العملية التعليمية وعندما يتحقق ذلك يتحقق الهدف المنشود ومقارنة الأداء بمعايير داخلية أو خارجية ، ونحن بصدد قضية جديدة تكمن في كيفية بناء منظومة جديدة للقياس والتقويم التربوي تواكب الثورة المعلوماتية وتقويم المتعلمين والمشاريع التعليمية وتقويم النظام التعليمي بأكمله (صلاح الدين علام، ٢٠٠٩) .

فجودة العملية التعليمية وتحسين مستوياته لا يمكن تحقيقه إلا من خلال إصلاح عملية التقويم ، فإصلاح التقويم يُعد مدخلاً للإرتقاء بجودة العملية التعليمية ، ويعد عملية التقويم عملية مهمة تأثرت بالمؤسسة التعليمية إبان الثورة الصناعية ، ولذا انتشرت البحوث التجريبية والشبه تجريبية وكذلك البحوث السيكمترية ، وظهرت مفاهيم أساسية مثل الدقة ، والضبط، والقياس ، والاختبار ، والصدق ، والثبات ، والحيادية ، والموضوعية ، ومع انتشار التكنولوجيا فقد انتشرت مفاهيم جديدة مثل صناعة المعرفة ، والوسائط المتعددة، والتعلم النشط،، والإتصال ، والعمل الجماعي، والتعلم الذاتي وغيره من هذه المفاهيم.

فالاختبارات من أحد وسائل التقويم ويُعتمد عليه في صنع القرارات التي تهدف إلى تحسين عمليتي التعليم والتعلم وذلك لأن الإمتحانات جزء أساسي من العملية التعليمية تُستخدم في

صياغة سياسة تعليمية وإجراء تعديلات على المناهج الدراسية إذا استدعى ذلك) (Phillip,2014:23.

ولابد من رسم منظومة للقياس والتقييم التربوي للتعليم كمدخل لإصلاح وتحسين جودة العملية التعليمية (حسيب محمد حسيب ٢٠٢٠)

ويتطلب التقييم تقييم كافة مدخلات ومخرجات العملية التعليمية في منظومة متكاملة بكل مكوناتها من منهج ومعلم ومتعلم ونظام امتحان ، ويتطلب التقييم للإمتحانات أدوات قياس صالحة لتقييم تعلم الطلاب مثل " الإمتحانات التحصيلية " ولذلك لابد من التأكد من جودة هذه الإمتحانات وقدرتها على تحقيق الهدف منها والتعرف على مستوى الطلاب لتحقيق الهدف التعليمي والكشف عن مواطن القوة والضعف حتي يتم تحسين وتطوير العملية التربوية " فؤاد أبوحطب وآخرون ، ٢٠٠٨ ، (BacolodmTobias,2006, Brwon,2002) .

ولكي يتم تطوير الإمتحانات التحصيلية ينبغي الأخذ بالنظام الحديث بما ينعكس أثره على تطوير أداء المعلم الذي تتوفر فيه الخصائص التي تتناسب مع متطلبات المجتمع من حيث المرونة ، والإبداع ، والإبتكار بما يحقق معايير الجودة في العملية التعليمية لدى الطالب الجامعي (محمد محمد عبدالمعظم ، ٢٠٠٩:٢٠٨).

ومن خلال ماسبق اتضح أن مصطلح الجودة في التعليم ، والتميز ، والإبداع الأكاديمي من المصطلحات الحديثة لعملية التعلم ، والتي جاءت على التغير في طبيعة الأهداف التعليمية التي انتقلت من التركيز على الكفاءات المعرفية للطلاب وعلى اختبارات التحصيل كوسيلة للحكم على عملية التقييم لتصبح ملازمة للعملية التعليمية ومستمرة تحت طائلة التعلم مدى الحياة.

وأشارت دراسة (أمانى الغامدي ، سماح زغاري ، ٢٠١٨) لتقويم الإختبارات أكدت إلى وجود قصور في بناء معظم الإمتحانات التحصيلية وأنها تحتاج إلى كثير من الموضوعية كما أنها غير مقننة ويصعب الإعتماد على نتائجها في الحكم على الطلاب ولذا من الضروري أن تبني الإمتحانات على أسس ومعايير علمية حديثة ومعاصرة بحيث تقدم بيانات صادقة يمكن الإعتماد عليها .

كما أشارت دراسة (منال حسن ٢٠١٣) إلى خطورة إنحصار أسئلة الإمتحانات في المستويات المعرفية الدنيا حيث أن ذلك يعكس بعداً تربوياً خطيراً لأنهم يعتمدون على قياس المستويات المعرفية الدنيا من الممكن أن ينتقل لطلابهم وينهجون نفس النهج والإستراتيجية في إعداد أسئلتهم لأن ما يتم إستخدامه في إعداد المعلم من طرائق وأساليب تقييمية مختلفة يتبعه المعلم بعد تخرجه مع تلاميذه أثناء التدريس .

وتأسيساً على ذلك فترى الباحثة أنه تبين محدودية الأساليب المتبعة في تقييم الطلاب وقلة جودها في التشخيص الدقيق لمستويات الطلبة بالرغم من الجهود المبذولة في تطوير وتقييم الاختبارات التحصيلية النهائية مازالت هذه الجهود شكلية أكثر منها موضوعية فمن خلال الملاحظات الميدانية لعملية الاختبارات من خلال عملي بوحدة القياس والتقييم بكلية التربية تبين وجود تفاوت واضح في مستوى الاختبارات التحصيلية مابين شديدة الصعوبة وشديدة السهولة ، على الرغم من إعتمادهم على الاختبارات المختلطة مابين الأسئلة الموضوعية والمقالية معاً وأن معظمها غير مطابقة للمواصفات ولا يقيس إلا المستويات الدنيا فلا تقيس الإبتكار والجودة وسلامة الحكم وتعتمد على قياس الحفظ والإستدكار ، وبعد الإطلاع على التقارير الفنية وجد قصور في جودة المفردات الإختبارية من حيث معاملات السهولة والصعوبة وإنخفاض معاملات الثبات مما يؤدي إلى تفاوت كبير في نتائج الطلاب ويترتب على ذلك إختلافات في ردود فعل الطلاب بعد الإمتحانات وتفاوت آرائهم ومستوى رضاهم عنها .

مشكلة الدراسة

التقييم عملية تعد لإصلاح التعليم وهو عنصر مهم في إدارة البرامج الأكاديمية والوسيلة الفعالة في ضبط نوعية المخرجات وضمان جودتها وعلى الرغم من تعدد أدوات ووسائل القياس والتقييم إلا أن الإمتحانات التقييمية تعد من أهم وسائل التقييم وأكثرها إنتشاراً (فؤاد أبوحطب وآخرون ٢٠٠٨) .

وقد أجريت العديد من الدراسات لتقييم الاختبارات التحصيلية في مراحل التعليم المختلفة ، وأكدت معظم الدراسات أن الاختبارات ما زالت تركز على المستويات المعرفية الدنيا وتعمل المستويات المعرفية العليا ولا تهتم بالجوانب الوجدانية والمهارية ، كدراسة سالم أحمد

المجاهد (٢٠١٣)، ودراسة منال علي حسن (٢٠١٣)، ودراسة خديجة ضيف الله القرشي (٢٠١٠)، ومحمد محمد عبدالمنعم (٢٠٠٩)، ودراسة ثناء محمد حسن (٢٠٠٥).
وقد وضعت هيئات التقويم والقياس التربوي عملية التقويم متطلب أساسي للحصول على الإعتماد من خلال تقنين وتنويع أساليب التقويم للتحقق من مدى توافر معايير الجودة. وأسئلة الإمتحانات من الوسائل التقويمية للطلاب التي تساعدهم على توظيف المعلومات وتثبيتها وتكاملها مما يسهل عمليتي الإحتفاظ بهذه المعلومات وانتقالها وهما الهدفان الأساسيان للتعلم ، وهذه الأسئلة بمستوياتها المتعددة المختلفة المعرفية ، الإنفعالية ، المهارية ، ولذا الأسئلة الجيدة تولد أفكارًا أو عقلًا مدرّكًا لدى الطلاب ، بينما الأسئلة الرديئة التي تؤكد على الحفظ دون الفهم فإنها توصل الطلاب للنجاح دون فهم وتطبيق ما تعلمه في مواقف جديدة ، ولذلك عندما تكون الأسئلة هادفة تحقق عملية التقويم السليم لدى الطلاب وتشخيص نقاط الضعف لديهم من أجل الإرتقاء وتطوير المستوى من ناحية ومن ناحية أخرى يتم تحقيق الأهداف السلوكية أو الجودة وما يقدمه المعلم من نشاطات مختلفة تساعد على وضع الكفايات التحصيلية لدى الطالب الجامعي ، وقد أشارت بعض الدراسات إلى عدم اهتمام المعلم وضعف مهاراتهم في صياغة الأسئلة وتصميم الورقة الإمتحانية والتي تعد المعيار الأساسي لتصنيف قدرات الطلاب وربما يكون أحد الأسباب الرئيسية لتدني مستوى التحصيل الدراسي بشكل عام وتدني التحصيل في التعليم العالي بشكل خاص وكثرة رسوب الطلاب ترجع إلى عدم التعامل بشكل كاف مع فن وضع الأسئلة وكيفية صياغتها وشروطها ومستوياتها وتطبيقاتها ومن هنا تولدت الفكرة لدى الباحثة بإجراء دراسة تقويمية وتحليلية للأسئلة الإمتحانية في كلية التربية للعام الدراسي (٢٠٢٣ - ٢٠٢٤) على وفق تصنيف بلوم المعرفي ولذا تتمثل مشكلة البحث في محاولة الإجابة على السؤال التالي :

ما مدى ملائمة المؤشرات الكمية والكيفية لمعايير جودة الإمتحانات لتصنيف بلوم لدى عينة البحث الحالي؟

أهداف البحث

١- إعداد قائمة بمعايير الجودة لإختبارات المقررات التربوية بكلية التربية ،جامعة دمنهور .

- ٢- تقديم معايير تساعد المعلمين في تقييم الإختبارات وتطويرها في ضوء معايير الجودة.
- ٣- تقرير عن تطوير الإختبارات للمقررات التربوية بكلية التربية ،جامعة دمنهور
- ٤- تطوير أساليب حديثة للتقويم بالتعليم الجامعي في ضوء معايير الجودة

أهمية البحث

- ١- تقديم صورة واضحة عن الإختبارات التحصيلية للمقررات التربوية بكلية التربية ،جامعة دمنهور من خلال الكشف عن نقاط القوة والضعف .
- ٢- الكشف عن مدى توافر معايير الجودة للإختبارات التحصيلية للمقررات التربوية .
- ٣- بناء قائمة معايير لجودة الإختبارات تفيد المعلمين بكلية التربية في تطوير اختباراتهم والإبتكار والإبداع فيها.

مفاهيم البحث الاجرائية

التقويم Evaluation: هو عملية منظمة لجمع وتحليل المعلومات حول البرامج الأكاديمية المتعلقة بالمعلم والمتعلم والإدارة والوسائل والنشاطات التي تشكل مجموعها وحدة عملية التعليم والتعلم للتأكد من مدى تحقيق نواتج التعلم واتخاذ القرار بالبرامج الأكاديمية في تحقق أهدافها (أحمد عودة ، ٢٠٠٢).

والتعريف الإجرائي : هو عملية تقويم أسئلة الامتحانات التحصيلية لجميع أقسام الكلية وبيان نقاط القوة والضعف في الأسئلة.

الأسئلة : هي أداة يستخدمه معلمو البرامج في نهاية الفصلين الدراسين الأول والثاني لقياس مستوى المام الطالب للمعارف والمهارات المتعلقة بالمادة التدريسية ويتم التعرف أيضاً على مدى التغيرات التي تحدث عند الطلبة نتيجة العملية التعليمية.

الاختبارات التحصيلية : هي أداة إجراء منظم يتكون من عدد من الأسئلة المعدة مسبقاً لقياس تحصيل الطلبة في جمع المواد الإمتحانية النهائية والمهام التعليمية أو المخرجات التعليمية.

معايير الجودة: المعيار هو المستوى المرجعي الذي يستدل عليه من خلال المؤشرات التي تستخدم كمواصفات للأداء النموذجي (Arter& Mctighe,2001:12).

بينما تُعرف الجودة بأنها : التزام المؤسسة التعليمية بأنجازات وابتكار معايير حقيقية متعارف عليها وتتطابق عناصر المنظومة التعليمية مع المواصفات القياسية وفي إطار المعايير القومية للعملية التعليمية (إبراهيم سعد شحاتة، ٢٠٠٥: ١٠).

منهج البحث

يستخدم البحث الحالي المنهج الوصفي التحليلي في تقديم وصف الإختبارات للمقررات التربوية لتقويم الإمتحانات وجودتها، وكذلك يحلل الإختبارات للمقررات التربوية في ضوء معايير الجودة للعملية التعليمية.

أدوات البحث

قائمة لمعايير جودة الإختبارات للمقررات التربوية .

فروض البحث

- ١- هل توجد فروق ذات دلالة إحصائية بين متوسط درجات تحليل محتوى الإختبار للأعوام (٢٠٢٠:٢٠٢٤) وذلك في ضوء المعيار الأول (البيانات الأساسية) ؟
- ٢- هل توجد فروق ذات دلالة إحصائية بين متوسط درجات تحليل الإختبار للأعوام (٢٠٢٠:٢٠٢٤) في ضوء المعيار الثاني (تعليمات الإختبار) ؟
- ٣- هل توجد فروق ذات دلالة إحصائية بين متوسط درجات تحليل محتوى الإختبارات للأعوام (٢٠٢٠:٢٠٢٤) في ضوء المعيار الثالث (لغة الإختبار) ؟
- ٤- هل توجد فروق ذات دلالة إحصائية بين متوسط درجات تحليل الإختبار للأعوام (٢٠٢٠:٢٠٢٤) في ضوء المعيار الرابع (معيار ارتباط الإختبار بالتوصيف) ؟
- ٥- هل توجد فروق ذات دلالة إحصائية بين متوسط درجات تحليل الإختبار للأعوام (٢٠٢٠-٢٠٢٤) في ضوء المعيار الخامس من معايير جودة الإختبار (مضمون الإختبار) ؟

خطوات البحث واجراءته

١- الإطلاع على الأدبيات العربية والأجنبية في مجال القياس والتقويم ونواتج التعلم وجودة العملية التعليمية .

٢- عرض قائمة معايير جودة الاختبارات على مجموعة من الخبراء في مجال القياس والتقويم .

٣- حصر الإختبارات التربوية في كلية التربية جامعة دمنهور .

٤- تحليل الإختبارات في ضوء قائمة معايير جودة ونواتج التعلم المعدل من قبل الباحثة .

٥- المعالجة الإحصائية لعملية التحليل وتفسير النتائج ومعالجتها .

٦- تقديم التوصيات والمقترحات بشأن جودة نواتج التعلم الواجب توافرها بالإختبارات .

الاستعراض النظري والبحثي لمفاهيم البحث

توجد صعوبات في حياة الإنسان بدون التقويم ، والقياس، والتقدير، والتقييم ، ولذلك اختلف المعلمون حول تقويم الاختبارات وفي نظرتهم له وربما يرجع ذلك إلى التطورات العلمية من خلال الدراسات والبحوث التربوية الكثيرة فالتقويم هي العملية التي من خلالها تتحدد قيمة ما يحدث أو وصف شئ ما تم الحكم على قبول أو ملائمة ماوصف فأساس كلمة التقويم معرفة قيمة الشئ أو الفكرة أو المعني أو العادة أو الإستعداد أو الخدمة أو الحاجة إلى الوصول إلى الهدف أو النشاط التعليمي والتربوي الذي يتطلب الحكم عليه من خلال نجاحه أو فشله بالنسبة للأهداف وظهور مواضع الضعف والقوة معني هذا أن عملية الحكم أو التقويم هي جزء هام من العملية التعليمية والتربوية وأسلوباً هادفاً وأساساً لتطويرها .

وأصبح للتقويم أهداف متنوعة كتقويم كفاءة الطلاب ومراقبة تعلمهم وتطوير استراتيجيات تعلمهم وتقويم نموهم على مدار فترة زمنية معينة ودمج الطلبة في عملية التقويم من خلال عملية التقويم الذاتي ومعرفة حاجاتهم ونقاط القوة لديهم ، ومن هنا نري أنه لايمكن لنوع واحد من التقويم أن تحقق الأهداف المرجوة ولذلك تم الإهتمام بالتقويم الصفي والتعلم الواقعي فهو يتكامل مع عملية التدريس كما يشرك الطلاب في تقويم أعمالهم ويوفر لهم فرصة لإثبات قدراتهم وكفائتهم ومواطن القوة لديهم لدعمها (أحمد الثوابية ، ٢٠٠٤).

أهمية تقويم نظام الإمتحانات

١- ضمان العدالة والتأكد من أن جميع الطلاب لديهم الفرصة المتساوية لإظهار معرفتهم ومهاراتهم .

٢- تحسين جودة التعليم والكشف عن نقاط القوة والضعف في المناهج الدراسية .

٣- رفع مستوى الثقة في النتائج والموضوعية فيها .

٤- تقويم الإمتحانات وتطوير أدواته أكثر فاعليه ودقة (Brown,2010)

مؤشرات أداء نظام الإمتحانات

٥- معدل النجاح معقول ومتوقع .

٦- التوزيع الطبيعي للدرجات .

٧- الارتباط بين نتائج الإمتحان وأداء الطلاب في الحياه العلمية والعملية .

٨- جمع آراء الطلاب والمعلمين والمراقبين حول النظام .

٩- تحليل البيانات الإحصائية والمتعلقة بنتائج الإمتحانات (Crocker,Algina,1986)

.(

ولذلك أصبحت عمليتي القياس والتقويم من أهم التطورات التي حدثت في المواقف التعليمية من خلال إتخاذ العديد من القرارات التربوية ، وتعد عملية التقويم أشمل من عملية القياس بالرغم من أنهما مرتبطان إذا أن عملية التقويم تصف السلوك وصفاً كمياً ونوعياً ، بينما عملية القياس تصف السلوك وصفاً كمياً فقط ولذا يتحقق من خلالهما قياس مدى تحقيق الأهداف التعليمية وجودة التعليم وتقويم ما يحصله التلاميذ من معلومات من خلال الإختبارات التحصيلية .

تطوير نظم التقويم بمراحل التعليم المختلفة :

نظام التقويم في المراحل التعليمية المختلفة لم يلحقه تطور مستمر لأنه مازال تقويماً تقليدياً يعتمد على نظام الإمتحان الواحد في نهاية فترة التعليم ، ومفهوم التقويم مرادف لمفهوم الإمتحان ولا بد من تقويم الطلاب للمرحلة الجامعية لأنه يشهد أمور غير عادية منها :

١- القلق المستمر والتوتر الشديد للمتعلمين وأسرهم .

٢- تزايد الدروس الخصوصية و إنتشارها وزيادة التكلفة المادية على أسر المتعلمين .

٣- اللجوء إلي الغش بأنواعه المختلفة .

- ٤- غياب المتعلمين وإنقطاعهم عن المدرسة .
 - ٥- إزدياد مستويات العنف وقلة إنضباط المتعلمين .
 - ٦- إقتصار الإمتحانات على المستويات العقلية فقط كالحفظ والتذكر .
 - ٧- توقف المؤسسات التعليمية عن أداء دورها والقيام بمهامها ووظائفها .
- ولذا لابد من تطوير نظم التقويم بالمراحل التعليمية المختلفة وذلك من خلال مايلي :
- ١- تعديل نظام الدراسة بالمراحل التعليمية من حيث المسارات والزمن المخصص لدراسة كل مقرر .
 - ٢- تحديد الأوزان النسبية لدرجات المقررات الدراسية في ضوء الأهمية المجتمعية وعدد سنوات الدراسة.
 - ٣- عدم الإعتماد على الإمتحان ذات النظام الواحد والموحد في نهاية العام الدراسي ويجب الإعتماد على أساليب متعددة لقياس مستويات المتعلمين .
 - ٤- رفع كفاءة المعلمين بما يكفل لهم تطبيق عمليات القياس والتقويم بالصورة الأمثل .
 - ٥- نشر نظم التقويم الجديد بشكل موسع وإشراك المعلمين وأولياء الأمور والطلاب في مناقشة النظام وتعريفهم بمزاياه وعيوبه (حسيب محمد حسيب ، ٢٠٢٠)
- تقويم وتطوير عمليات المتابعة :**
- ١- بناء نظام فعال للمتابعة والتقويم يقيس مدي تطبيق السياسات وفعالية النظام التعليمي على كافة المستويات وفقاً لمعايير واضحة ومعلنة .
 - ٢- تسخير كافة جهود منظومة المتابعة والتنظيم نحو تشخيص حالة المؤسسات التعليمية ، وحصر مشكلاتها وتقديم الحلول المناسبة في صورة خطط إجرائية منها ما هو قابل للتنفيذ بشكل عاجل ومنها ما يدرج في خطط طويلة الأجل وهو ما يعرف بالتقويم من أجل التمكين .
 - ٣- تقويم السياق العام للنظام التعليمي من حيث المناخ العام ، والتمكين وتكافل الفرص والرضا عن العمل ، والمشاركة المجتمعية .

- ٤- تقويم أداء المؤسسات التعليمية باعتبارها وحدة واحدة من خلال توافر جودة المدخلات .
- ٥- إجراءات تنفيذية لعمليتي التعليم والتعلم وما تتضمنه من ممارسات للأداء المدرسي الفعال ويشمل تقويم الأداء المهني للإدارة المدرسية والمعلمين .
- ٦- تقويم مخرجات العملية التعليمية
- ٧- نشر ثقافة المتابعة والتقويم القائم على المعايير داخل الأجهزة والمؤسسات التعليمية .
- ٨- تشكيل هيكل تنظيمي مقترح للمتابعة والتقويم ويتضمن هذا الهيكل اللجنة العليا للمتابعة والتقويم برئاسة وزير التربية والتعليم .
- ٩- الإدارة المركزية للمتابعة والتقويم بديوان الوزارة كإدارة العامة لبحوث التقويم ، والإدارة العامة لمعايير مؤشرات الأداء ، والإدارة العامة لإعداد التقارير الفنية .

تقييم الإمتحانات التحصيلية :

استراتيجية تقييم الإمتحانات التحصيلية هي: قياس القدرات والمهارات العقلية الموضوعة لقياسها وقد أشار (Biggs,19973) حصول الطلاب الذين يتعلمون في شكل صم على درجات عالية في الإمتحانات وهو منتشر في عديد من النظم التعليمية ويرجع ذلك بتحليل الإختبار التحصيلي إذ أن كثير من المفردات لا تتطلب أكثر من استدعاء الحقائق والتفاصيل الموجودة في المحاضرات والكتب ولا تغطي كافة نواتج التعلم ، وقد أشار (Entwistle&Entwistle,1992) لأن هناك علاقة مرتبطة بين نوعية الأسئلة المستخدمة من قبل المعلمين ونوع التفكير السائد لدى الطلاب فالمعلمين الذين يستخدمون أسئلة تنير التفكير يصبح طلابهم قادرين على ممارسة العمليات العقلية العليا .

ولكي تؤدي الإختبارات التحصيلية وظائفها بكفاءة عالية لابد أن يتم التخطيط لها وبناءها وفق أسس علمية حيث تؤكد المنظومة التعليمية أن المعلم هو الذي يجب أن يقوم بإعداد الإختبارات وتطبيقه ، وتصحيحه باعتباره أن لديه القدرة الفعلية على ممارسة الإختبارات بأسلوب منظم وهادف ، والمحك في ذلك إمتلاك المعلم لكفايات بناء الإختبارات التحصيلية التي يكون قادراً على إتقانها وأصبحت المعلومات التي تقدمها والإختبارات أقرب إلي الصحة

والصواب والموضوعية ولذلك تزداد الثقة بصحة المعلومات ويكون إتخاذ القرار قريباً من الواقع العملي (عبدالرحمن عدس، ١٩٩٩: ٩).

وتعد أفضل وسيلة لتقويم الطلاب من خلال الإختبارات التحصيلية الجديدة إما عن طريق التخطيط الجيد للإختبار الذي يتطلب كتابة جدول المواصفات للإختبار قبل البدء به وتحديد الأهمية النسبية لكل وحده أو موضوع دراسي ، إما عن طريق المهارات الجيدة في كتابة الأسئلة التي تتضمن إمتلاك المعلم الجامعي مهارة وكتابة الأنواع المختلفة من الأسئلة سواء كان إختبار مختلط مابين الموضوعي أو المقالي أو المقالي فقط وقدرته على وضع عدد كاف وتحديد صعوبة فقرات الإختبار وكذلك قدرته على وضع أسئلة تحاكي الفروق الفردية للطلاب (Hopkins &Antas,1985:178) .

ولكى تؤدي الإمتحانات التحصيلية على أكمل وجه لابد أن تتصف بالصفات التالية:

أ- الموضوعية **Objectivity** وتعني بعدم تأثر نتائج التقويم بالعوامل الذاتية أو الشخصية للمصحح ولذلك فإن درجة الطالب لا تتوقف على من يصحح ورقته فلا تختلف درجته بإختلاف المصححين .

ب- الصدق **Validiity** : ويقصد به قدرة الإمتحان على قياس الشئ الذي وضع لقياسه فعلاً ولا يقيس غيره ، وطريقة التقويم الصحيحة تظهر في صدق الأداء الذي يقاس ولذلك لابد أن يشتمل الإمتحان على الدروس والوحدات التي درست في الفصل الدراسي .

ج- الثبات **Reliability** : ويسمى بالدقة ويقصد به أن يعطي الإمتحان نفس النتائج إذا ما كررت تطبيقها في قياس الشئ نفسه مرات عديدة ولذلك سؤال واحد لا يمثل عينه كافية لقدرة الطالب ولذا يستخدم عدد كبير من الأسئلة للحصول على مقياس موثوق به لمعرفة المهارة والقدرة ، وكلما زاد عدد الأسئلة كلما كان ضماناً أكبر .

د- القيمة التشخيصية : وهو يبين ما نجح فيه الطالب وما لم ينجح فيه ويحدد الخطأ الذي وقع فيه ويتطلب هذا التقويم إنتباهاً لعملية الأداء وتوضيح كيفية استجابة الطالب للسؤال ، ويعد من أفضل أساليب التقويم التي تبين بالتحديد ما يخطأ الطالب

في عمله ، فالمعلم قد يقرأ ورقة الإمتحان ويقدر درجات الطالب وتنتهي عملية التقويم بالنسبة له ، فإذا وجد للطالب علامة ضعيفة فقد يرغب في تحسينها مرة أخرى ولكن الإمتحان لم يبين له أساس التحسن . (ريان حسن فكري ١٩٩٥ : ٣٤٢) .

وتوجد وجهات نظر عديدة لتصنيف بلوم الذي استخدم المعيار المستخدم في هذا البحث حيث وضع تصنيف هرم بلوم علي أساس تصنيف الأهداف التعليمية في المجال المعرفي في ستة مستويات (التذكر المعرفة ، والفهم ، والإستيعاب ، والتطبيق ، والتحليل ، والتركيب ، و التقويم (زياد بركات وصباح عبدالهائي، ١٢: ٢٠٠٧)



شكل (١) يبين تصنيف الأهداف التعليمية تبعاً لتصنيف بلوم المعرفي

ومن خلال الشكل السابق يجب تطوير الأهداف التعليمية في المجال التعليمي وإصلاحها لمسايرة العملية التربوية في تقدمها مما يشغل التربويين ولكي يحدث عملية التعلم ينبغي الإهتمام بأسئلة الإمتحان التحصيلية إهتماماً بالغاً في العملية التعليمية وعملية إعداد الإختبارات أصبحت في بعض الأحيان غاية تسلطية بحد ذاتها تثير الرعب والقلق في نفوس كثير من الطلبة (عبد الحميد السورتي ، ١٩٩٨ : ٣٩) .

الإختبارات التحصيلية في الوضع الحالي لاتشجع على الإبتكار وإنما تشجع على الحفظ والصم الألي وتركز على ما يحفظه الطلبة لا على ما يستطيعون فعله (سناء إبراهيم أبودقه ، ١٤٧: ٢٠٠٧) .

ويحدد المتخصصون عدة أنواع للإختبارات التحصيلية تبعاً لطريقة الإستجابة إلي ثلاثة أنواع رئيسية وهي الإختبارات الشفوية ، والكتابية ، والأدائية وتسمى هذه الإختبارات بالإختبارات الصفية لإستخدام المعلم لها في بناء إختباراتهم لتقييم الطلاب

أولاً: الإختبارات الشفوية Oral axamination

استخدمت هذه الإختبارات منذ قديم الزمان من خلالها يوجه المعلم للطلاب عدة أسئلة ويطلب منه الإجابة عليها ومازالت هذه الإختبارات موجودة حتى الآن .

ثانياً: الإختبارات الكتابية Essay test Items أو التحريرية

ويهدف منها الترابط بين الأسئلة والمحتوي لتحقيق الأهداف وهي نوعان من الإختبارات هما إختبار الأداء الأقصى ، وإختبار تحديد الأداء المميز وهي إختبارات تقيس ما يتحمله أن يفعل المختبر في موقف معين أو نوع معين من المواقف وتتنوع الإختبارات ما بين إختبارات مقالية ، وإختبارات موضوعية

أ- الإختبارات المقالية Essay test

يستخدم هذا النوع من الإختبارات لتحقيق الأهداف المتمثلة في التركيب والتحليل حيث أنها تقيس الأهداف المعرفية كما صنفها بلوم بالرغم من أنها سهلة الإعداد إلا أنها تحتاج إلي جهد كبير في الإجابة والتصحيح وتنقسم الأسئلة المقالية إلي نوعين - الأسئلة المقالية المفتوحة وهي تبدأ بكلمات مثل تذكر تكلم أو تحدث ، وتتطلب الإجابة هنا بالإبتكار، والتنظيم في إيجاد موضوع متكامل ، وتعطي للطلاب حرية تحديد طبيعة إجابته ومداها مما يتيح له الفرصة لإظهار مهاراته في الشرح والتحليل والتقييم.

- الأسئلة المقالية المحددة : ويكون السؤال فيها طويلاً والإجابة محددة وتستلزم من الطالب الفهم والإستيعاب وعلى الرغم من شيوع وإنتشار الإختبارات المقالية إلا أنها لاتعطي جميع المادة وتحتاج إلي قدرة كتابية وتدخل ذاتية المصحح وتأثير ضعف لغة الطالب في حالة الكتابة مما يؤثر على تقييم الطالب وكذلك تحتاج وقت وجهد كبير في الإجابة

ب- الإختبارات الموضوعية Objection test :

وهي أسئلة لا تتأثر بذاتية المصحح ومن أشهرها إختبارات الصواب والخطأ ، الإختيار من متعدد ، المزوجة أو المطابقة ، التكميل.

ثالثاً: إختبارات الأداء Performance Tisting

وهذا الاختبار يهدف إلى الكشف عن قدرة الطالب على القيام بأنشطة معينة وأدائه لمهارات في موقف مقارب ومشابه للحياة بصورتها الواقعية ومناسبة للمادة الدراسية وتحقق هذه الإختبارات العديد من الأهداف منها:

- ١- قياس مدى فهم الدراسة النظرية المرتبطة بالمادة العلمية كما هو في العلوم الطبيعية.
 - ٢- قياس بعض المهارات .
 - ٣- التنبؤ بنجاح الطلاب في المستقبل.
 - ٤- تشخيص جوانب القوة والضعف في المجالات المهارية .
- ويمكن تصحيح الإختبارات الأدائية وتفسير نتائجها من خلال استخدام بطاقة الاختبار ، تحدد عليه جوانب المهارت واعطاء أرقام لتقدير مستوى أداء الطالب المفحوص (صلاح الدين علام، ٢٠٠١).

ثانياً: جودة الإختبارات

عملية التقويم للإمتحانات من العناصر الأساسية للجودة ، وإصلاح نقاط الضعف حول الجودة والإعتماد في التعليم العالي تركز على نواتج العملية التعليمية ، وتشتمل عملية الجودة على إعادة تشكيل العلاقة بين المعلم والطالب من عملية تعليم وتعلم إلى عملية العمل بروح الفريق الواحد (على محمد زهيد الغامدي، ٢٠٠٨: ٧).

فالمؤسسة التعليمية تضع معايير للأداء وللتقويم له دور في استيفاء لهذه المعايير ولذلك ينبغي جعل مهام عملية التقويم في نفس الخط مع نواتج التعلم في المقررات التربوية والبرامج الخاصة بالعملية التعليمية ، ويتم اصدار الحكم على جودة التقويم في المؤسسة التعليمية من خلال : نظم التقييم وذلك بهدف اجراء المقارنات ، ونظم الاختبارات التي تقيس معارف الطلاب للمناهج الدراسية (Somerset,2011:23).

وتعتبر الجودة في المجال التربوي مجموعة من المعايير والإجراءات والمواصفات والخصائص المتوقعة في الخدمة التعليمية وفي العمليات والأنشطة التي تتحقق من خلالها المواصفات التي تحسن فب المنتج التعليمي فالجودة الشاملة توفر أدوات وأساليب تساعد على تحقيق نواتج التعلم (أحمد علي كنعان، ٢٠٠٥: ٦٦).

فنتقويم الإمتحانات من منظور الجودة يتعدى مجرد إجراء الإختبارات التي تشمل الأسئلة الكتابية والتي تقيس المعارف والتي تجري في فترات محددة إلى مفهوم أكثر شمولية والانتقال من الإختبارات التحريرية البسيطة إلى التقويم المتعدد ومن قياس معارف ومهارات بسيطة إلى قياس المستويات العليا من التفكير ومن قياس مستوى المعارف إلى كيفية توظيفها في الحياة المهنية والعامة فضلاً عن تقويم الجوانب الوجدانية كالقيم والاتجاه (J, et al , Kulieke , 1990). وتقويم نتائج الطلاب تعطي مؤشرات حول مدى الحاجة إلى تفعيل العملية التعليمية والجوانب التي ينبغي تطويرها من تحديد الأهداف كنواتج التعلم إلى طرق التعليم والتعلم والوسائل الموظفة فيها إلى اختبار أدوات التقويم التي تتناسب لتحقيق الأهداف المنشودة ، ومع ظهور الجودة والإعتماد أصبحت الحاجة ملحة إلى لوجود أساليب حديثة لتقويم الإمتحانات تتواءم مع ظروف العصر وتتمثل فيما يلي:

أ- تقييم الأداء : Performance Assessment

تقييم الأداء أداة للانتقالات الحادة التي وجهت إلى الإختبارات المقالية والموضوعية بطريقتها التقليدية المعتادة والتي تقيس العمليات المعرفية في أدنى مستوياتها ، وذلك لأن تنمية القدرات والعمليات المعرفية تمثل هدفاً من أهداف أى نظام تعليمي ، وعدم وجود اختبارات مناسبة لقياس تلك القدرات ولا يمكن اصدار أحكام صادقة موضوعية ودقيقة على مدى امتلاك المتعلمين لهذه القدرات ومدى نموها لديهم (إسماعيل صبري ، محب كامل الرفاعي ، ٢٠٠٣: ٣٦٢).

ويشتمل أداء الطالب مجموعة من المعايير لأداء مهمة وتسمح هذه المعايير بقياس مستوى أداء المتعلم ونوعية استجابته بصورة شاملة ، ويفضل استخدام تقويم الأداء من أربعة إلى ستة مرات أثناء مدة الدراسة ، وتتطلب مهام الأداء الإبتكار لتحقيق مستويات جودة معينة ، ويتم

تقييم الابتكارات معتمداً على محكات تتضمن خصائص الجودة التي ينبغي أن تتوفر فيها ، ولذا لا يقتصر الإهتمام بالابتكار وإنما يعتمد على خصائص الابتكار والنتائج ونوعيتها (صلاح الدين محمود علام ، ٢٠٠٤: ١١).

ب- سجلات العمل Portfolios

تتمثل أعمال الطالب في الملف الذي ينجزه لإستخدامه في عرض أفضل انجازاته أو رصده نموه التعليمي خلال فترة زمنية معينة بحيث تقدم صورة واضحة عن مستوى تقدم التعلم والتفكير الذي يحقق الأداء الجيد للطالب (خديجة ضيف الله القرشي ، ٢٠١٢: ٣٠) ويُعد السجلات عملية من عمليات التقييم المستمر للطالب والمشاركة الفعالة من قبل الطالب ، والذي يتم فيه تحديد الأهداف وتقويم مدى تحققها وتحديد ما يجب تعلمه مستقبلاً وتحدد درجات لكل سجل من سجلات تقييم الطالب بتنظيم الطالب لسجل عمله ووضوح تفكيره وشرح المفاهيم وتحليل المشكلات ، ومن خلال ذلك يتم تقييم الطالب من قبل المعلم وتوضيح نقاط القوة والضعف لديه (Coates, 1995: 23).

ج- مقاييس التقدير Rating Systems

وهي اختبار يقوم فيه الطالب بتبويب الفقرات إلى فئات وفقاً لبعد أنفق أو لا أنفق ، ويتضمن عبارات وصفية في عدد ثابت من المجموعات (Colman, 2006: 13). وتستخدم مقاييس التقدير في عملية التقويم الذاتي للطالب أو من خلال عمليات الملاحظة في سمة من السمات أو مجموعة من السمات ووصفها كمياً والحكم على وجودها لدى الطالب المتعلم (أكرم عادل الشير ، أريج عادل ، ٢٠١٢: ١٤). وتقدر مقاييس التقدير تقديرًا كمياً تعبر عن إنطباع أداء الطالب ، أو تقدر بمقاييس التحليل والتي تعطي تفصيلات لمكونات القدرات العقلية كالطلاقة والمرونة للسلوك القابل للقياس والملاحظة والحكم الموضوعي الذي يعكس مدى كفاءة العمليات المعرفية المحددة التي يعتقد أنها أساسية لكفاءة الأداء المعرفي النمائي والأكاديمي والمهاري (فتحى مصطفى الزيات ، ٢٠٠٤: ١٥).

واهتمت العديد من الدراسات بتقويم الإختبارات التحصيلية بهدف التحقق من استيفائها معايير جودة الإختبار وتمثل فيما يلي:

وقد أشارت دراسة (عبدالكريم السوداني ، ٢٠٠٦) أن مستويات التفكير (تصنيف بلوم للمجال المعرفي التي تستثيرها أسئلة الإختبارات النهائية التي يتعرض لها الطلاب كلية التربية قسم علوم الحياه ، وقد إعتد على أسلوب تحليل المحتوي في معرفة هذه المستويات وبعد التحقق من صحة التحليل وحساب معامل الإتفاق بين المحللين وبلغت عدد الأسئلة المحللة (١٦٦) سؤال فقد كشفت الدراسة عن خلو الأسئلة من مستويات التطبيق والتركيب والتقويم بينما الصفوف الأعلى كان مستوي التذكر أعلى المستويات تحققاً

وأشار (فؤاد أبوحطب و آخرون ، ٢٠٠٨) عدم مراعاة معايير الإختبار الجيد كالصدق والثبات والموضوعية والشمولية لكافة نواتج العملية التعليمية خاصة في مؤسسات التعليم العالي يؤدي إلي ضعف موثوقية الدرجات التي يحصل عليها الطلاب في الدرجات التحصيلية ويقلل من أهميتها في تزويد القسم العلمي بمعلومات مفيدة عن مستوى تحقق نواتج تعلم البرنامج الأكاديمي ، وتعتبر الإمتحانات وعملية تقويم الطلاب مجرد تحصيل حاصل وقد يكون هذا هو السبب في وجود شعوراً عاماً بأن التعليم يعاني من مشكلات وأوجه قصور أهمها تدني مستوى نواتج التعلم وأكثرها وضوحاً ضعف مستوي الخريجين كما أشارت دراسة (ضياء الدين عبدالواحد ، ٢٠١٥) أن الإختبارات التحصيلية النهائية لقسم المكتبات والمعلومات بجامعة عين شمس لاتعطي صور حقيقية عن قياس الأهداف التعليمية كما تعكسها أهداف المستوي الدراسي ، حيث تركز على المجال المعرفي خاصة التذكر مع إغفال المجال المهاري والوجداني مما يدل على عدم صدق تلك الإختبارات .

وأشارت دراسة (محمد عبدالمنعم ، ٢٠٠٩) والتي هدفت للكشف عن واقع اسئلة الإختبارات التحصيلية للمواد الأكاديمية والتربوية للفصل الدراسي الثاني للطلاب الخريجين من التعليم الأساسي في كلية المعلمين جامعة الملك فيصل التخصصات الأدبية والعلمية في ضوء معايير جودة التقويم حيث تم تحديد قائمة لمعايير الإختبارات التحصيلية الكتابية ثم تقييم أسئلة تلك الإختبارات في ضوء تلك المعايير، وتوصلت نتائج الدراسة عن وجود العديد من الأخطاء

اللغوية والنحوية والطباعية في أسئلة كثير من الإختبارات ووجود كثير من الأخطاء العلمية والفنية في جودة الأسئلة بالإضافة إلي تكرار بعض الأسئلة بصورة كاملة أو بصورة جزئية من إختبارات سابقة كما أن كثير من أسئلة تلك الإختبارات قدر ترد إجاباتها في نفس الإختبار في نصوص أسئلة أخرى وقد إستدل من هذه النتائج على أن معظم الإختبارات لم تبني في ضوء مواصفات الإختبار الجيد وفي ضوء جدول المواصفات

ودراسة (يحي الشريفي ، ٢٠١٣) هدفت إلى مدى شمولية أسئلة الإختبارات التحصيلية لمادة الأدب في كليات التربية لمستويات بلوم المعرفية وتوصل مستوي الفهم على أعلى تكرارات بينما حصل مستوي التقويم على أدنى التكرارات أما تنوع الأسئلة لم يجد الباحث أي أثر للإختبارات الموضوعية في نماذج الأسئلة و إقتصرت الأسئلة كلها على الأسئلة المقالية ، ومعيار الشمولية لمحتوي المادة الدراسية ضعيفاً في نماذج الأسئلة النهائية ولم يصل إلي النسبة التي إفتراضها الباحث وهي (٨٠ %)

أما دراسة (على الحديبي ، وهناء إبراهيم ، ٢٠١٤) هدفت إلى وضع قائمة لمعايير الجودة التعليمية في تصميم أوراق الأسئلة الإختبارية بالجامعة الإسلامية وتضمنت ثلاثة معايير أساسية ينبثق عنها (٢٥) مؤشراً حيث يتضمن المعيار الأول البيانات الأساسية عن الجامعة والمقرر الدراسي ، والمعيار الثاني يركز على الجوانب الفنية التي تتعلق بإخراج ورقة الإختبار وتنظيمها وترتيبها ، والمعيار الثالث يركز على الصياغة العلمية والصحيحة للأسئلة سواء كانت مقالية أو موضوعية وبفحص وتقويم أوراق الإختبارات في ضوء تلك المعايير إتضح أن المعيار الثاني (الإهتمام بإخراج وتنظيم الورقة الإمتحانية) جاء في الترتيب الأول يليه المعيار الأول (البيانات الأساسية للجامعة والمقرر) ويليه المعيار الثالث (الصياغة العلمية الصحيحة للأسئلة) مما يدل على أن معلم المادة يهتم بإخراج الورقة وتنظيمها ثم بكتابة البيانات الأساسية ويأتي في آخر إهتمامهم صياغة الأسئلة صياغة علمية .

دراسة (منال حسن ، ٢٠١٣) هدفت إلى تقويم أسئلة الإختبارات النهائية للأقسام الأدبية والعلمية لجميع المستويات بكلية التربية بحفر الباطن جامعة الدمام في ضوء معايير الجودة ، وتوصلت نتائجه أن

أغلب أسئلة الإختبارات التحصيلية في المواد العلمية والأدبية و التربوية تقيس المستويات المعرفية الدنيا فأغلبها تقيس أدنى المستويات المعرفية وهو مستوى التذكر وفي أحسن الأحوال قد تصل إلي مستوى الفهم ، كما توصلت إلى أن معظم الأسئلة أهملت معيار الشمول ولم تلتزم بالمعايير التي خصصت لتنظيم وإخراج ورقة الإختبار .

كما أشارت دراسة (حمد الكلثوم ، ٢٠١٥) التي تحدد معايير درجة الورقة الإمتحانية من حيث الشكل والمضمون وجودة الأسئلة المقالية والموضوعية ، ثم التعرف علي مدى توافر هذه المعايير في الورقة الإمتحانية ودراسة الفروق في نسب تحقق هذه المعايير في الأوراق الإمتحانية وقد كشفت النتائج أن الأسئلة التي تقيس مستوى التذكر كانت بالمرتبة الأولى مقارنة بالأسئلة التي تقيس باقي المستويات المعرفية ، وكذلك توجد فروق ذات دلالة إحصائية فيما يتعلق بتحقيق معايير جودة الإختبار، وتنوع الأسئلة لتشتمل على العديد من المستويات المعرفية ومعايير جودة الأسئلة المقالية والموضوعية لصالح جامعة أم القرى .

دراسة (أماني الغامدي ، سماح زغاري ، ٢٠١٨) وهدفت هذه الدراسة إلى تحليل عينة من إختبارات التقويم المستمر في كلية التربية بجامعة الإمام عبدالرحمن بن فيصل لدراسة مدى استيفائها للمعايير الكيفية والشكلية التي تقيس نواتج التعلم وتوصلت الدراسة أن النسبة العامة لتحقيق المؤشرات الشكلية أعلى من نسبة المؤشرات الكيفية وهي نسبة غير مرضية في التعبير عن جودة مضمون وبناء الإختبارات التي تم تحليلها وكان عددها (٢٠) إختبارًا حيث لم يكن هناك نوعاً من التوازن بين الأسئلة المقالية والموضوعية ولم يتم التوظيف الأمثل للأنواع المختلفة للأسئلة الموضوعية كما كانت جميع الأسئلة المقالية محددة بإجابات ليست مفتوحة فلم تسمح بقياس نواتج التعلم في صورتها المتقدمة المبينة على عمليات عقلية نوعية توضح التمايز بين المستويات المتباينة للمتعلمين .

كما أشارت (منار الشامي ، سهام الشافعي ، ٢٠١١) أن مدي جودة الإختبارات التحصيلية وتحديد الأهمية النسبية لأنواع الأسئلة المختلفة في الإختبارات التحصيلية فقد تم إنخفاض جودة الإختبارات التحصيلية حيث لم تصل أي منها لدرجة الجودة المطلوبة وقد تم تقييم جودة الإختبارات في ضوء أربعة معايير إلي معيار التحليل العام والذي يتضمن شكل الورقة

الإختبارية و أهم الأخطاء اللغوية وكتابة تعليمات الكتابة والدرجة الكلية للإختبارات ، معيار مضمون الأسئلة والذي يتضمن وضوح عبارات الأسئلة وعلاقتها بأهداف المادة ومدي شموليتها لموضوعات المقرر ، معيار الأسئلة المقالية و كذلك معيار الأسئلة الموضوعية ، وتوصلت نتائج الدراسة إلى أن لمعيار التحليل العام الأثر الأكبر في درجات جودة الإختبار يليه معيار مضمون الأسئلة ثم معيار جودة الأسئلة المقالية ثم معيار جودة الأسئلة الموضوعية كما أن هناك تفاوت في نسب أنواع الأسئلة في الإختبارات وركزت بشكل كبير على الأسئلة الموضوعية حيث بلغت نسبتها (٨٣.٠٢ %) في مقابل الأسئلة المقالية التي بلغت نسبتها (١٦.٠٨ %) كما وجدت علاقة إرتباطية عكسية بين ضغوط العمل وبين جودة الإختبارات .

دراسة (بندر بن زيد الأمير ، ٢٠١٦) والتي هدفت إلى دراسة الأخطاء الشائعة في بناء مفردات الإختبار من نمط الإختبار من متعدد لدى ملعمي الرياضيات وتأثيرها على الخصائص السيكمترية للإختبارات التحصيلية ، حيث تم فحص عينة مكونة من (٥١) نموذج من الأسئلة لمعلمي الصف الثاني المتوسط وبعد فحص الأسئلة وفقاً لقائمة تدقيق الأخطاء التي تكونت من ثلاثة أبعاد رئيسية وهي الأخطاء في مدخل السؤال ، الأخطاء في الإجابة الصحيحة ، الأخطاء في البدائل ، ولذا أن الأخطاء الشائعة في بناء مفردات الإختبارات من نمط الإختبار من متعدد خمسة أخطاء فيهما خطئان شائعان وهما أن عدد البدائل غير مناسب لعمر الطالب والبدائل غير مرتبة من أعلى إلى أسفل أو العكس وكانت نسبة تكرارها بنسبة (٥١ %) ، (٢٩.٠٩ %) كما وجد خطئان شائعان في مدخل السؤال وهما الصياغة غير الواضحة لمفردة بالإضافة إلى خطأ واحد في الإجابات الصحيحة وهو عدم توزيع الإجابة الصحيحة توزيعاً عشوائياً وكانت نسبة تكرارها (١٠٠ %) بينما توجد فروق ذات دلالة إحصائية في متوسطات معاملات الصعوبة والتمييز وفاعلية المشتتات بين مفردات الإختبار الجيد ومفردات الإختبار غير الجيد كما وجدت فروق ذات دلالة إحصائية بين معاملات الثبات والصدق التلازمي بين الإختبارين لصالح الإختبار الجيد .

وأشارت دراسة (هدي صالح ، حنان محمد ، ٢٠١٧) والتي هدفت إلى تحديد الأخطاء الشائعة في بناء مفردات أسئلة الصواب والخطأ وذلك بفحص عينة مكونة من (٤٠) نموذج

من أسئلة معلمات علم النفس للفصل الدراسي الثاني لتحديد الأخطاء الشائعة فيها وقد كشفت النتائج عن وجود أربعة أخطاء أولها نقل العبارة نصاً من الكتاب المدرسي ، والثلاثة الأخرى مرتبطة بإخراج أسئلة الصواب والخطأ وهم عبارات غير متقاربين في الطول والتوزيع غير العشوائي للعبارات فنسبة عبارات الصواب أقل من عبارات الخطأ ، وكذلك توصلت إلي عدم وجود فروق ذات دلالة إحصائية في متوسطات معاملات السهولة بين مفردات الإختبار الجيد والإختبار غير الجيد بينما وجدت الفروق في معاملات التمييز بين مفردات الإختبارين تعود إلى الأخطاء الشائعة في بناء مفردات أسئلة النموذج غير الجيد ، بينما وجدت فروق ذات دلالة إحصائية بين معاملات الثبات للإختبارين لصالح الإختبار الجيد بينما لا توجد الفروق بينهما في معاملات صدق المحك .

ومن جانب آخر هناك دراسات إهتمت بالأخطاء بصفة عامة مثل دراسة (على صلاح حسن ، ٢٠١٨) ، حيث قام بإعداد قائمة بتدقيق الأخطاء في بناء الإختبارات الموضوعية وفحص مجموعة عشوائية من الإختبارات الموضوعية لأعضاء هيئة التدريس من كليات وأقسام مختلفة بجامعة أسيوط ، وقد كشفت النتائج نسبة أسئلة الصواب والخطأ كانت نسبة الأخطاء الإملائية أقل من (٢٠%) بينما كانت نسب بعض الأخطاء تتراوح ما بين (٢٥%) و (٩٤%) وهي نسب كبيرة ، أما بالنسبة لأسئلة الإختبار من متعدد فتراوحت نسبة الأخطاء فيها ما بين (٢٥%) و (١٠٠%) موزعة ما بين مدخل الفقرة والبدايل ، والإجابات الصحيحة وهي نسب كبيرة أيضاً مما يدل على إهتمام كثير من أعضاء هيئة التدريس بالخطوات العلمية لبناء الإختبارات ومراعاة شروط وقواعد بناء الإختبارات الموضوعية ، كما كشفت نتائج الدراسة عن وجود فروق ذات دلالة إحصائية بين إختبار الموضوع الجيد والإختبار المتضمن على الأخطاء في الخصائص السيكمترية للإختبار ككل من حيث متوسطات معاملات الثبات والصدق للإختبار ولفقراته من حيث معاملات متوسطات السهولة والصعوبة والتمييز لصالح الصورة الجيدة .

ويتبين من العرض السابق للدراسات السابقة ما يلي:

قامت بعض الدراسات بتحليل محتوى الإختبارات في ضوء معايير جودة أداء التعلم ، وكذلك تنوعت المراحل التعليمية منها ما اهتم بتقويم الإختبارات في مرحلة ما قبل التعليم الجامعي ، ومنها ما اهتم بتقويم الإختبارات في مرحلة التعليم الجامعي، وأكدت نتائج معظم الدراسات أن الإختبارات تركز على المجال المعرفي على حساب المجالين المهاري والوجداني، وأشارت بعض النتائج كدراسة (محمد عبدالمنعم، ٢٠٠٩) التي توصلت إلى وجود العديد من الأخطاء اللغوية والنحوية والطباعية في أسئلة كثير من الإختبارات ووجود كثير من الأخطاء العلمية والفنية في جودة الأسئلة بالإضافة إلى تكرار بعض الأسئلة بصورة كاملة أو بصورة جزئية من إختبارات سابقة كما أن كثير من أسئلة تلك الإختبارات قدر ترد إجاباتها في نفس الإختبار في نصوص أسئلة أخرى وقد إستدل من هذه النتائج على أن معظم الإختبارات لم تبني في ضوء مواصفات الإختبار الجيد وفي ضوء جدول المواصفات ، كما اتضح من دراسة (على الحديبي ، وهناء إبراهيم ، ٢٠١٤) والتي تضمنت ثلاثة معايير أساسية ينبثق عنها (٢٥) مؤشراً حيث يتضمن المعيار الأول البيانات الأساسية عن الجامعة والمقرر الدراسي ، والمعيار الثاني يركز على الجوانب الفنية التي تتعلق بإخراج ورقة الإختبار وتنظيمها وترتيبها ، والمعيار الثالث يركز على الصياغة العلمية والصحيحة للأسئلة سواء كانت مقالية أو موضوعية وبفحص وتقويم أوراق الإختبارات في ضوء تلك المعايير اتضح أن المعيار الثاني (الإهتمام بإخراج وتنظيم الورقة الإمتحانية) جاء في ترتيب الأول يليه المعيار الأول (البيانات الأساسية للجامعة والمقرر) ويليه المعيار الثالث (الصياغة العلمية الصحيحة للأسئلة) مما يدل على أن معلم المادة يهتم بإخراج الورقة وتنظيمها ثم بكتابة البيانات الأساسية ويأتي في آخر إهتمامهم صياغة الأسئلة صياغة علمية.

وتأسيساً على ما سبق اتضح أن الوصول إلى التميز في العملية التعليمية يجب أن تتطابق الأهداف مع نواتج العملية التعليمية ، كما يتم التحقق من هذا التطابق بعملية التقويم الصحيحة للتحقق من جودتها وكفائتها وصدق نتائجها في تقرير هذا التطابق و من ثم تظهر أهمية تقويم واقع الإمتحانات التحصيلية وفحص مدى جودة مفرداتها ومدى مطابقتها للمواصفات القياسية

والتحقق من مدى صلاحيتها في قياس نواتج التعلم ولاسيما أن يظهر رأي الطلاب فيها ومدي رضاهم عنها بإعتبارهم المدخل والمخرج في العملية التعليمية، ولأهمية التقويم في انجاح العملية التعليمية لابد من توظيفه التوظيف السليم من أجل إنجاح النظام الجامعي الجديد التي تبحث من خلاله الدولة على الجودة وترقية الكفاءات ، وعلى القائمين من مسئولين وأساتذة العمل والتطبيق الصارم الذي يعد أحد ركائز التكوين الهادف وتحقيق الهدف والمعتمد في الأصل على التدريس بالكفاءة لا بالمهارة .

إجراءات البحث

منهج البحث

من أجل تحقيق هدف البحث الحالي استخدمت الباحثة المنهج الوصفي التحليلي الذي من خلاله وصفت الباحثة موضوع البحث وتحليل بياناته والآراء التي تطرح حول العمليات التي تتضمنها فالمنهج الوصفي يتناسب مع طبيعة البحث الحالي لأنه اعتمد هنا على جمع بيانات وتبويبها والوصول إلى الهدف التي تسهم في فهم الواقع وتصوره .

مجتمع البحث

تكون مجتمع البحث الحالي من الأوراق الإختبارية التي أعدها أعضاء هيئة التدريس في التخصصات العلمية والإنسانية ، وتم اختيار (كلية التربية) بالأسلوب العشوائي بشعبها الثلاث (عام وأساسي وطفولة) كعينة للبحث الحالي.

أداة البحث

قامت الباحثة باعداد معايير للمواصفات القياسية للتحليل و للتقويم وعدة جوانب من الإمتحانات من وجهة نظر طلاب كلية التربية ، وكيفية استخدامها لفحص معايير للمواصفات القياسية ، وبناء إستبيان رضا الطلاب عن الإمتحانات النهائية والتحقق من صلاحيتها للتطبيق ، وتطبيق تلك الأدوات على عينة من الأوراق الإمتحانية النهائية ، ولذلك قامت الباحثة ببناء أداة معايير الجودة لتحليل الإختبارات ومكون من (٥٠) معياراً من خلال تصميمه بناء على الأسس التالية: نواتج التعلم ، المراجع والدراسات السابقة المتعلقة بمعايير الجودة، المراجع والدراسات في مجال القياس والتقويم ، ومن خلال المعايير الأساسية في

صورته الأولى يندرج تحت كل معيار مجموعة من المعايير الفرعية ، وحدد لإرتباط المحور بخيارين هما (مرتبط - غير مرتبط) ، بينما حدد لصحة الشرط علمياً خيارين أيضاً (صحيح - غير صحيح) ، وحدد الصياغة اللغوية خيارين هما : (واضحة ، وغير واضحة) ، بينما حدد لأهمية العبارة ثلاث خيارات (دائماً ، أحياناً ، مطلقاً) وذلك لحساب الوزن المئوي لفقرات المعيار .

وتم اشتقاق معايير الإختبارات التحصيلية من خلال مراجعة الأدبيات المتخصصة في القياس والتقويم التربوي والدراسات التي تناولت تحليل وتقويم اختبارات التحصيل ، ولذا تم إعداد قائمة بمعايير الإختبارات التحصيلية وتكون قابلة للتحقيق في البحث الحالي وتشتمل على التنظيم والإخراج في الورقة الإمتحانية ، وكذلك تنوع الإختبارات وعدم اقتصارها على نوع واحد ، وكذلك شمول الإختبار لموضوعات المقرر ، وكذلك المستوي المعرفي الذي يقيسه كل سؤال ، وذلك بحسب تصنيف بلوم فيما يلي التذكر : يتضمن تذكر المعلومات المتعلقة بالمقرر ، الفهم : ويتضمن قدرة الطالب على توضيح المعلومات الموجودة في المقرر ، التطبيق : ويقصد به قدرة الطالب على تطبيق القواعد والمبادئ في مواقف معينة للوصول إلى حل مشكلة في موقف ما ، وأيضاً تحليل المعلومات ، وتركيب وإعادة تنظيم المعلومات وجعلها مترابطة ذات فكرة جديدة ، التقويم وإصدار الأحكام القيمية على المواقف الجديدة مع مراعاة قواعد صياغة أنواع الإختبارات مثل الإختبارات المقالية ، إكمال الفراغات ، الصواب والخطأ ، الإختيار من متعدد ، المطابقة ومراعاة دقتها من الناحية العلمية التي ترد في الأوراق الإختبارية .

صدق أداة البحث

وتم حساب الصدق الظاهري لمعيار (مدى ملائمة العبارات مع ما وضعت له) إلى جانب صدق المحتوى (مدى تغطية الفقرات للمجال المراد قياسه) وذلك وفق الإجراءات التالية:

الصدق الظاهري : قامت الباحثة بعرض المعيار على مجموعة متخصصة من خبراء علم النفس والقياس والتقويم التربوي وبلغ عدد المحكمين (١٠) ، وفي ضوء نتائج تحكيم القائمة تم إجراء التعديلات التي اقترحتها السادة المحكمين مثل دمج معيار البيانات الأساسية مع معيار بيانات الطالب ، ودمج معيار لغة الإختبار مع معيار الشكل العام للإختبار ، وكذلك إعادة

تسمية معيار ارتباط الإختبار بالهدف إلى ارتباط الإختبار بالتوصيف، وبناء على هذه التعديلات أصبحت القائمة في صورتها النهائية خمسة معايير أساسية وإثنان وأربعون معيار فرعي .

خطوات تحليل الأوراق الإختبارية والأسئلة المتضمنة فيها

تحليل كل ورقة إختبارية وتحليل كل سؤال على حسب بنود المعايير الآتية :

١- تحليل كل ورقة إختبارية وتحليل كل سؤال في هذه الورقة من خلال التنظيم والإخراج ورصد البيانات الأساسية في الأماكن المخصصة لها .

٢- تحديد الهدف من الإستبيان ويهدف إلى قياس رضا الطلاب عن الإمتحانات التحصيلية النهائية .

٣- تحديد الجوانب المراد قياسها في ضوء الهدف من البحث الحالي وأهم المشكلات التي تواجههم والأخطاء الواردة من وجهة نظرهم من حيث وضوح الأسئلة وعددها وتمثيلها للمقرر وتنوع المهارات من خلالها ومدى ملائمتها لزمان الإمتحان والدرجة المخصصة على كل سؤال وكذلك ملاحظتهم من حيث الصياغة اللغوية ومقترحاتهم لتحسين كل هذه الملاحظات .

٤- في ضوء نتائج الإستبيان واستجابات الطلاب وردودهم على الإستبيان وقائمة معايير للمواصفات الورقة الإمتحانية تم صياغة المفردات المقياس ، وقد تضمن المقياس على (٥٠) عبارة موزعة (البيانات الأساسية للإمتحان (٥) عبارات ، وتعليمات الإختبار (٦) عبارات ، لغة الإختبار (٦) عبارات ، ارتباط الإختبار بالتوصيف (٨) ، مضمون الإختبار (٩) بما يتناسب (٣٤) معيار فرعي .

الوسائل الإحصائية

استخدمت الباحثة الوسائل الإحصائية لتحليل بيانات البحث الحالي من خلال تطبيق البرنامج

spss

١- استخدام قائمة معايير لجودة الإختبار لتحليل محتوى الإختبار .

٢- حساب تكرار استيفاء معيار جودة الإختبار ، وكذلك حساب النسبة المئوية لتكرار المعيار .

٣- إعادة عملية التحليل مرة ثانية بعد مرور شهر على التحليل الأول وتطبيق معادلة هولستي (Holesty) وذلك لحساب معامل الإتفاق بين التحليل الأول والثاني .

٤- معادلة بيرسون لاستخراج الثبات لأداة البحث بطريقة إعادة الإختبار .

٥- معادلة ألفا كرونباخ لاستخراج معامل الثبات لأداة البحث الحالي .

٦- تحليل التباين لتحديد دلالة الفروق بين المتوسطات .

نتائج البحث وتفسيرها

سوف يتم عرض النتائج التي توصلت إليها الباحثة ومن ثم تفسيرها وذلك وفق أهداف البحث وفق التسلسل الآتي :

تم جمع (٨٠) اختبار من اختبارات العلوم التربوية والنفسية بكلية التربية ،جامعة دمنهور من عام (٢٠٢٠:٢٠٢٤) بواقع (٢٠) اختبار لكل عام جامعي ، استخدام معايير لجودة الإختبارات وذلك لتحليل الإختبارات، ومرت عملية التحليل بالمراحل التالية:

١- تحديد هدف التحليل: ويهدف إلى الحكم على مدى استيفاء معايير جودة الإختبارات .
٢- تحديد فئة التحليل: وهي مجموعة من التصنيفات تقوم بها الباحثة طبقاً لنوعية المضمون والمحتوى والهدف منه ومن خلال ذلك يتم وصف المضمون وبما يتيح إمكانية التحليل واستخراج النتائج ، وتم استخدام قائمة معايير جودة الإختبار التي أعدت كأداة للتحليل.

٣- تحديد مساحة التحليل : وهي تقوم بها الباحثة بتحليلها وتشمل الورقة الإختبارية وصفحة التعليمات.

٤- ثبات أداة التحليل : باستخدام قائمة معايير جودة الإختبارات لتحليل محتوى الإختبار ، وحساب تكرار استيفاء معيار جودة الإختبار ، وكذلك حساب النسبة المئوية لتكرار المعيار، ولذلك يتم إعادة عملية التحليل مرة ثانية بعد مرور شهر على التحليل الأول وتطبيق معادلة هولستي (Holesty) وذلك لحساب معامل الإتفاق بين التحليل الأول

$\Delta V V$

الأساسية								
كتابة بيانات المقرروال شعبة	٢,٧٠٠ ٠	٠ ٤٧٠١٦	٢,٩٥٠ ٠	٣٩٥٠٠٠ ٠	٣٩٠٠٠ ٠	٢٢٣٦ ١٠	٣,٠٠٠ ٠	٣٠٧٧٩
تحديد زمن الإختبار	٢,٣٠٠ ٠	٧٣٢٧٠ ٠	٢,٩٥٠ ٠	٢٢٣٦١ ٠	٣,٠٠٠ ٠	٢٢٣٦ ١	٣,٠٠٠ ٠	٣٠٧٧٩
تحديد الدرجة الكليةللإخ تبار	٢,٢٠٠ ٠	٥٠٢٦٢ ٠	٢,٨٠٠ ٠	٤١٠٣٩ ٠	٢,٨٥٠ ٠	٣٦٦٣ ٥	٢,٨٥٠ ٠	٣٦٦٣٥
استيفاء بيانات أستاذ المادة ورئيس القسم	١,٣٥٠ ٠	٤٨٩٣٦ ٠	٢,٣٠٠ ٠	٤٧٠١٦ ٠	٢,٣٠٠ ٠	٥١٢٩ ٩	٢,٦٥٠ ٠	٤٨٩٣٦
المعيار الأول (البيانات الأساسية)	١٢,٦٥ ٠٠	١,٣٠٨ ٨٨	١٣,٨٠ ٠٠	٩٥١٤٥٠ ٠٠	١٤,١٥ ٠٠	٨٧٥٠ ٩٠	١٤,٢٥ ٠٠	١,٠١٩ ٥٥

ومن خلال الجدول السابق اتضح أن المعيار الأول لعام ٢٠٢١-٢٠٢٢ بلغ (١٢,٦٥٠٠) بانحراف معياري (١,٣٠) بنسبة تحقق (٨٣%) وقد حازت الفقرة الأولى والثانية نسبة عالية ، في حين جاءت العبارة الخامسة نسبة قليلة بما يعني ضعف توفر معيار استيفاء توقيع أستاذ المادة ورئيس القسم على ورقة الإختبار .

كما تشير بيانات المعيار الأول لعام (٢٠٢١-٢٠٢٢) أن متوسط تحقق المعيار بلغ (١٣,٨٠٠٠) بانحراف معياري (٩٥١٤٥٠) بنسبة تحقق (٩٢%) ، بينما حازت الفقرة الثانية والثالثة على نسبة تحقق عالية ، في حين حازت العبارة الخامسة بنسبة تحقق قليلة.، وكذلك تشير بيانات المعيار الأول لعام (٢٠٢٢-٢٠٢٣) أن متوسط تحقق المعيار بلغ (١٤,١٥٠٠) بانحراف معياري (٨٧٥٠٩٠) بنسبة تحقق (٩٤%) ، وحازت العبارة الثالثة نسبة تحقق عالية ، بينما العبارة الخامسة حازت على نسبة تحقق قليلة.

في حين تشير بيانات المعيار الأول لعام (٢٠٢٣-٢٠٢٤) إلى أن متوسط تحقق المعيار بلغ (١٤,٢٥٠٠) بانحراف معياري (٨٧٥٠٩٠) ، وحازت العبارة (١-٢-٣) على نسبة تحقق عالية ، بينما العبارة الخامسة في المرتبة الأخيرة ونسبتها قليلة ، وللتحقق من وجود فروق بين متوسطات درجات تحليل محتوى الاختبار للأعوام (٢٠٢٠-٢٠٢٤) للمعيار الأول ، وتم استخدام تحليل التباين ذي الأربع اتجاهات) .

كما تم استخدام تحليل التباين لدرجات مجموع البحث الحالي وذلك لتحديد دلالة الفروق بين المتوسطات في كل سنة من سنوات الاختبار كما هو موضح بالجدول التالي.

جدول (٣) دلالة الفروق بين المتوسطات لدرجات تحليل محتوى الإختبارات لأعوام (٢٠٢٠:٢٠٢٤) في ضوء المعيار الأول (معيار جودة الإختبارات)

الأعوام الدراسية الجامعية	الفرق بين المتوسطات عام ٢٠٢٠ ٢٠٢١	الفرق بين المتوسطات عام ٢٠٢١ ٢٠٢٢	الفرق بين المتوسطات عام ٢٠٢٢ ٢٠٢٣	الفرق بين المتوسطات عام ٢٠٢٣ ٢٠٢٤	مستوى الدلالة
-٢٠٢٠ ٢٠٢١	-	- ٢.١٥٠٠٠	٢.١٥٠٠٠	-٢.٦٠٠٠٠	دال إحصائياً
-٢٠٢١ ٢٠٢٢	٢.١٥٠٠٠	-	-٣٥٠٠٠٠	-٤٥٠٠٠٠	غير دال إحصائياً
-٢٠٢٢	٢.٥٠٠٠٠	٣٥٠٠٠٠	-	-٤٥٠٠٠٠	غير دال

إحصائياً					٢٠٢٣
غير دال	-	-٤٥٠٠٠٠	-٢.١٥٠٠٠	٢.٦٠٠٠٠	-٢٠٢٣
إحصائياً					٢٠٢٤

ومن خلال الجدول السابق تبين أن قيمة فرق متوسط درجات تحليل الإختبار لعام ٢٠٢٠ في ضوء المعيار الأول وبين اختبارات (٢٠٢١) يساوي (٢,١٥) وبذلك يكون دال إحصائياً ، كما أن الفرق بين اختبارات عام (٢٠٢٠) واختبارات عام (٢٠٢٢) يساوي (٢.٥) وهو دال إحصائياً ، كما أن الفرق بين اختبارات عام (٢٠٢٠) واختبارات عام (٢٠٢٣) يساوي (٢.٦) وهو دال إحصائياً .

كما تبين من جدول (٣) أن قيمة فرق متوسط درجات تحليل الإختبار لعام (٢٠٢١) في ضوء المعيار الأول وبين اختبارات (٢٠٢٢) واختبارات عام (٢٠٢٣) يساوي (٤.٥٠) وهو غير دال إحصائياً.

كما تبين أيضاً من الجدول السابق أن قيمة فرق متوسط درجات تحليل الإختبار لعام ٢٠٢٢ وبين اختبارات (٢٠٢٣) في ضوء المعيار الأول يساوي (٤.٥٠) وهو غير دال إحصائياً، وبذلك تم قبول الفرض الأول وبذلك توجد فروق ذات دلالة إحصائية بين متوسط درجات تحليل محتوى الإختبار لأعوام (٢٠٢٠:٢٠٢٤) في ضوء المعيار الأول من معايير جودة الإختبارات .

٢- للتحقق من صحة الفرض الثاني : بأنه توجد فروق ذات دلالة إحصائية بين متوسط درجات تحليل الإختبار للأعوام (٢٠٢٠:٢٠٢٤) في ضوء المعيار الثاني (تعليمات الإختبار) ، وتم استخدام المتوسط الحسابي والانحراف المعياري كما هو موضح بجدول (٤)

جدول (٤) المتوسطات والانحراف المعياري لدرجات تحليل محتوى الإختبار للأعوام (٢٠٢٠:٢٠٢٤) في ضوء المعيار الثاني (تعليمات الإختبار)

العبارة	المتوسط	الانحراف	المتوسط	الانحراف	المتوسط	الانحراف	المتوسط	الانحراف
٢٠٢٠	٢٠٢١	٢٠٢٠	٢٠٢١	٢٠٢١	٢٠٢٢	٢٠٢١	٢٠٢٢	٢٠٢١
٢٠٢١	٢٠٢٢	٢٠٢٢	٢٠٢٣	٢٠٢٣	٢٠٢٣	٢٠٢٣	٢٠٢٣	٢٠٢٣
٢٠٢٢	٢٠٢٣	٢٠٢٣	٢٠٢٣	٢٠٢٣	٢٠٢٣	٢٠٢٣	٢٠٢٣	٢٠٢٣
٢٠٢٣	٢٠٢٣	٢٠٢٣	٢٠٢٣	٢٠٢٣	٢٠٢٣	٢٠٢٣	٢٠٢٣	٢٠٢٣
٢٠٢٤	٢٠٢٣	٢٠٢٣	٢٠٢٣	٢٠٢٣	٢٠٢٣	٢٠٢٣	٢٠٢٣	٢٠٢٣

.....	٣.٠٠٠		٣.٠٠٠	٣.٧٧.	٢.٩٥٠	٣.٧٧٩.	٢.٣٠٠	واضحة وخالية من الغموض
.٤٤٤٢٦	٢.٧٥٠	٤٤٤٢٦.	٢.٧٥٠	٤٧٠١٦.	٢.٧٠٠	٦٨٦٣٣.	٢.٠٥٠	ترك مسافة بين تعليمات الاختبار والأسئلة
.٣٦٦٣٥	٢.٨٥٠٠	٤٤٤٢٦.	٢.٧٥٠٠	٥٠٢٦٢.	٢.٦٠٠٠	٦٤٠٧٢.	١.٩٠٠	خالية من التكرار
.٢٢٣٦١	٣.٠٠٠	٢٢٣٦١.	٢.٩٥٠٠	٤١٠٣٩.	٢.٨٠٠٠	٤٧٠١٦.	٢.٧٠٠	تحديد عدد الأسئلة
.٣٦٦٣٥	٢.٨٥٠٠	٤٤٤٢٦.	٢.٢٥٠٠	٤٨٩٣٦.	٢.٦٥٠٠	٥٠٢٦٢.	٢.٦٠٠	تحديد كيفية الإجابة
.٧٨٦٤٠	٢.٢٥٠٠	٧٨٦٤٠.	٢.٢٥٠٠	٦١٥٥٩.	٢.٢٠٠٠	٦٨٠٥٦.	١.٤٠٠	تحديد درجات الإجابة على كل سؤال
١.٣٥٣٣٦	٢٢.٤٠٠٠	١.٢٦٨٠٣	٢٢.١٥٠٠	١.٦٩٨٣٠	٢١.٤٠٠٠	١.٥٧٢٨٠	١٨.٥٠	المعيار الثاني

تبين من الجدول السابق أن متوسط تحقق المعيار الثاني (تعليمات الإختبار) لعام (٢٠٢١-٢٠٢٠) بلغ (١٨.٥٠٠) بانحراف معياري (١.٥٧) بنسبة تحقق (٧٧%) ،بينما حازت الفقرة رقم (٣) على المرتبة الأولى بنسبة تحقق عالية ، في حين حازت العبارة رقم (٦) في المرتبة الأخيرة بنسبة تحقق (٤٦%) مما يُعني ضعف توفر معيار تحديد درجة الإجابة لكل سؤال ، في حين تشير بيانات تحقق المعيار الثاني لعام (٢٠٢١-٢٠٢٢) إلى أن متوسط تحقق المعيار بلغ (٢١.٤٠٠) بانحراف معياري (١,٦٩) بنسبة تحقق

(٨٩%) ، بينما حازت الفقرة رقم (١) المرتبة الأولى وبنسبة عالية ، في حين جاءت العبارة رقم (٦) في المرتبة الأخيرة بنسبة تحقق (٧٣%) .
وتشير بيانات تحقق المعيار الثاني لعام (٢٠٢٢-٢٠٢٣) إلى أن متوسط تحقق المعيار بلغ (٢٢.١٥٠٠) بانحراف معياري (١.٢٦) بنسبة تحقق (٩٢.٢%) ، وقد حازت العبارة رقم (١) المرتبة الأولى بنسبة تحقق عالية ، في حين حازت العبارة (٥-٦) في المرتبة الأخيرة بنسبة تحقق (٧٥%) بما يعني ضعف توفر معياري تحديد كيفية الإجابة على السؤال وتحديد درجة الإجابة لكل سؤال .

كما تشير بيانات تحقق المعيار الثاني لعام (٢٠٢٣-٢٠٢٤) إلى أن متوسط تحقق المعيار بلغ (٢٢.٤٠٠) بانحراف معياري (١.٣٥) بنسبة تحقق (٩٣.٣%) ، وقد حازت العبارة رقم (١-٣-٤) على المرتبة الأولى بنسبة تحقق عالية ، في حين حازت العبارة رقم (٦) المرتبة الأخيرة بنسبة تحقق (٧٥%) ، وللتحقق من وجود فروق بين متوسطات درجات تحليل محتوى الإختبارات للأعوام (٢٠٢٠-٢٠٢٤) في ضوء المعيار الثاني من معايير جودة الإختبار تم استخدام (تحليل التباين ذي الأربع اتجاهات) .

كما تم استخدام تحليل التباين لدرجات مجموعة الدراسة لتحديد دلالة الفروق بين المتوسطات في كل عام من أعوام الإختبار كما هو موضح بجدول (٥) .

جدول (٥) دلالة الفروق لمتوسطات درجات تحليل محتوى الإختبارات للأعوام (٢٠٢٠-٢٠٢٤) في ضوء المعيار الثاني من معايير جودة الإختبار

مستوى الدلالة	الفروق بين المتوسطات				الأعوام
	٢٠٢٤-٢٠٢٣	٢٠٢٣-٢٠٢٢	٢٠٢٢-٢٠٢١	٢٠٢١-٢٠٢٠	
دال إحصائياً	-٣.٩٠٠٠٠	-٣.٦٥٠٠٠	-٢.٩٠٠٠٠	-	٢٠٢١-٢٠٢٠
غير دالة إحصائياً	-١.٠٠٠٠٠	-٠.٧٥٠٠٠	-	٢.٩٠٠٠٠	٢٠٢٢-٢٠٢١
غير دالة إحصائياً	-٠.٢٥٠٠٠	-	٠.٧٥٠٠٠	٣.٦٥٠٠٠	٢٠٢٣-٢٠٢٢
غير دالة إحصائياً	-	٠.٢٥٠٠٠	١.٠٠٠٠٠	٣.٩٠٠٠٠	٢٠٢٤-٢٠٢٣

تبين من جدول (٥) أن قيمة متوسط درجات تحليل الإختبار لعام (٢٠٢٠) في ضوء المعيار الثاني وبين اختبارعام (٢٠٢١) يساوي(٢.٩٠) وهو دال احصائياً ، كما أن الفروق بين اختبارات عام (٢٠٢٠) واختبارات عام (٢٠٢٢) يساوي (٣.٦٥) وهو غير دال إحصائياً ، كما ان الفرق بين اختبارات عام (٢٠٢٠) واختبارات عام (٢٠٢٣) يساوي (٣.٩٠) وهو غير دالة إحصائياً، كما تبين من الجدول السابق أن قيمة فرق متوسط درجات تحليل الإختبار لعام (٢٠٢١) في ضوء المعيار الثاني وبين اختبارات عام (٢٠٢٢) يساوي (٠.٧٥٠٠) وهو غير دالة إحصائياً ، كما أن الفرق بين اختبارات عام (٢٠٢١) واختبارات عام (٢٠٢٣) يساوي (١.٠٠) وهو غير دالة إحصائياً.

وبلغت قيمة فرق متوسط درجات تحليل الإختبارعام (٢٠٢٢) واختبارات (٢٠٢٣) في ضوء المعيار الثاني (٠,٢٥٠٠) وهو غير دال إحصائياً.

وبذلك تم قبول الفرض الثاني وتوجد فروق ذات دلالة إحصائية بين متوسط درجات تحليل الإختبارللأعوام (٢٠٢٠:٢٠٢٤) في ضوء المعيارالثاني (تعليمات الإختبار) من معايير جودة الإختبار .

٣- **للتحقق من صحة الفرض الثالث** بانه توجد فروق ذات دلالة إحصائية بين متوسط درجات تحليل محتوى الإختبارات لأعوام (٢٠٢٠:٢٠٢٤) في ضوء المعيار الثالث (لغة الإختبار) ، واستخدم المتوسط الحسابي والانحراف المعياري كما هو موضح بجدول (٦)

جدول (٦) المتوسطات والانحرافات المعيارية لدرجات تحليل محتوى الإختبارات للأعوام (٢٠٢٠-٢٠٢٤) في ضوء المعيار الثالث من معايير جودة الإختبار

العبارة	المتوسط	الانحراف	المتوسط	الانحراف	المتوسط	الانحراف	المتوسط	الانحراف
	عام	المعيار	عام	المعيار	عام	المعيار	عام	المعيار
	٢٠٢٠	٢٠٢٠	٢٠٢١	٢٠٢١	٢٠٢٢	٢٠٢٢	٢٠٢٣	٢٠٢٣
	٢٠٢١	٢٠٢١	٢٠٢٢	٢٠٢٢	٢٠٢٣	٢٠٢٣	٢٠٢٤	٢٠٢٤
يكون حجم الخط	٢.٧٠٠٠	٤٤٢٦.	٢.٧٠٠٠	٤٧٠١٦.	٢.٧٥٠٠	٤٨٩٣٦	٢.٧٥٠٠	٤٨٩٣٦

مناصب								
خالية من الأخطاء الطباعة	٣.٠٠٠		٢.٨٥٠٠	٤٨٩٣٦.	٢.٨٥٠٠	٢.٨٥٠٠	٤٨٩٣٦.	٢.٨٥٠٠
خالية من الأخطاء الإملائية	٢.٨٠٠٠	٤١٠٣٩.	٢.٨٠٠٠	٤١٠٣٩.	٢.٨٠٠٠	٢.٩٠٠٠	٣٠.٧٧٩	٢.٩٠٠٠
خالية من الكلمات الغامضة	٢,٨٥٠٠	٣٦٦٣٥.	٢.٨٥٠٠	٣٦٦٣٥.	٢.٨٥٠٠	٢.٩٥٠٠	٢٢٣٦١	٢.٩٥٠٠
خالية من المصطلحات التي لها أكثر من معنى	٢.٥٠٠٠	٤١٠٠٠٠.	٢.٢٠٠٠	٦٥٠٠٠.	٢.٢٠٠٠	٢.٠٠٠٠	١٣٢٠٠	٢.٠٠٠٠
مناسبة لمستوى الطلاب	١.٩٥٠٠	٤٨٩٣٦.	٢.١٠٠٠		٢.٣٥٠٠	٢.٥٠٠٠	٣٢٤٠٠	٢.٥٠٠٠
المعيار الثالث	١٦.٩٥	٠.٦٠٩١٦	١٧.٩	٧١٣٥١.	١٧.٢	٥٦٢٩٤.	١٧.٢٥	٠.٦٢٥٥١

وتشير بيانات الجدول السابق أن متوسط تحقق المعيار الثالث (لغة الاختبار) لإختبار للعام (٢٠٢٠-٢٠٢١) بلغ (١٩.٧٥٠) بانحراف معياري (١.٠١٥) بنسبة تحقق عالية ، وقد حازت العبارة رقم (٢) على المرتبة الأولى بنسبة تحقق عالية ، في حين حازت العبارة (٦) المرتبة الأخيرة بنسبة تحقق (٦٥%) بما يُعني ضعف توفر معيار مناسبة مستوى الاختبار ، بينما تشير بيانات تحقق المعيار الثالث لعام (٢٠٢١-٢٠٢٢) إلى أن تحقق المعيار بلغ (٢٠.٠٠٠) بانحراف معياري (١.٢١) بنسبة تحقق عالية كما في العبارة رقم (٣) ، وحازت العبارة رقم (٦) المرتبة الأخيرة بنسبة تحقق (٧٠%) بما يُعني ضعف توفر معياري.

وتشير بيانات تحقق المعيار الثالث لعام (٢٠٢٢-٢٠٢٣) أن متوسط تحقق المعيار بلغ (٢٠.٠٠٠) بانحراف معياري (٠.٩٧١) بنسبة تحقق (٩٥%) ، وحازت العبارة رقم (٣) على

المرتبة الأولى بنسبة تحقق عالية، وحازت العبارة رقم (٥-٦) المرتبة الأخيرة بما يُعني ضعف توفر معياري خالية من المصطلحات التي لها أكثر من معنى ومعياري مناسبة لمستوى الاختبار.

في حين تشير بيانات تحقق المعيار الثالث لعام (٢٠٢٣-٢٠٢٤) أن متوسط تحقق المعيار بلغ (٢٠١٥٠) (بأنحراف معياري ٠.٩٥) بنسبة تحقق (٦٠%) ، وحازت العبارة رقم (٣) المرتبة الأولى بنسبة تحقق عالية ، بينما حازت العبارة (٥) المرتبة الأخيرة ونسبة تحقق منخفضة بما يُعني ضعف توفر معيار خلو المصطلحات التي لها أكثر من معنى ، وللتحقق من وجود فروق ذات دلالة بين متوسطات درجات تحليل محتوى الاختبار بأعوام (٢٠٢٠:٢٠٢٤) في ضوء المعيار الثالث من معايير جودة الاختبار ، وتم استخدام (تحليل التباين ذو الأربع اتجاهات).

وتم استخدام تحليل التباين لدرجات مجموعة الدراسة لتحديد دلالة الفروق بين المتوسطات في كل عام من أعوام الاختبار كما هو موضح بجدول (٧).

جدول (٧) دلالة الفروق لمتوسطات درجات تحليل الاختبار لأعوام (٢٠٢٠:٢٠٢٤) في ضوء المعيار الثالث (لغة الاختبار)

الأعوام	الفروق بين المتوسطات لعام ٢٠٢١-٢٠٢٠	الفروق بين المتوسطات لعام ٢٠٢٢-٢٠٢١	الفروق بين المتوسطات لعام ٢٠٢٣-٢٠٢٢	الفروق بين المتوسطات لعام ٢٠٢٤-٢٠٢٣	مستوي الدلالة
٢٠٢١-٢٠٢٠	-	-٢٥٠٠٠	-٢٥٠٠٠	-٤٠٠٠٠	٠.٩٧٤
٢٠٢٢-٢٠٢١	٢٥٠٠٠	-	-١٢٠٠٠	-١٥٠٠٠	٠.٩٧٤
٢٠٢٣-٢٠٢٢	٢٥٠٠٠	٢٥٠٠٠	-	-١٥٠٠٠	١.٠٠٠
٢٠٢٤-٢٠٢٣	٤٠٠٠٠	١٥٠٠٠	١٥٠٠٠	-	٠.٩٩٣

تبين من الجدول السابق أن قيمة فرق متوسط درجات تحليل الاختبار لعام (٢٠٢٠) في ضوء المعيار الثالث وبين إختبارات (٢٠٢١) يساوي (٠.٢٥٠٠-) وبذلك يكون دال إحصائياً ، كما أن الفرق بين إختبارات عام (٢٠٢٠) و إختبارات عام (٢٠٢٢) يساوي (٠.٢٥٠٠-) وهو غير دال إحصائياً ، كما أن الفرق بين إختبارات عام (٢٠٢٠) وإختبارات عام (٢٠٢٣) يساوي (٠.٤٠٠٠-) وبذلك يكون غير دال إحصائياً ، وأن قيمة فرق متوسط درجات

الإختبار (٢٠٢١) في ضوء المعيار الثالث وبين اختبارات (٢٠٢٢) يساوي (-١٢٠٠٠٠) وهو غير دال إحصائياً ، كما أن الفرق بين اختبارات عام (٢٠٢١) واختبارات عام (٢٠٢٣) يساوي (-١٥٠٠٠٠) وهو غير دال إحصائياً .

وبذلك يتضح من الجدول السابق أن قيمة فرق متوسط درجات تحليل الإختبارات لعام ٢٠٢٢ وبين اختبارات عام (٢٠٢٣) في ضوء المعيار الثالث يساوي (١٥٠٠) و هو غير دال إحصائياً .

وبذلك تم قبول الفرض الثالث وتوجد فروق ذات دلالة إحصائية بين متوسط درجات تحليل الإختبار للأعوام (٢٠٢٠:٢٠٢٤) في ضوء المعيار الثالث (لغة الإختبار) .

٤- للتحقق من صحة الفرض الرابع للبحث وهو توجد فروق ذات دلالة إحصائية بين

متوسط درجات تحليل الإختبار للأعوام (٢٠٢٠:٢٠٢٤) في ضوء المعيار الرابع

(معيار ارتباط الإختبار بالتوصيف) ، وتم استخدام المتوسط الحسابي والانحراف

المعياري كما هو موضح بجدول (٨)

جدول (٨) المتوسطات والانحراف المعياري لدرجات تحليل الإختبار للأعوام (٢٠٢٠:

٢٠٢٤) في ضوء المعيار الرابع من معايير جودة الإختبار

العبارة	المتوسط لعام ٢٠٢٠ - ٢٠٢١	الانحراف المعياري لعام ٢٠٢٠ - ٢٠٢١	المتوسط لعام ٢٠٢٢ - ٢٠٢٣	الانحراف المعياري لعام ٢٠٢٢ - ٢٠٢٣	المتوسط لعام ٢٠٢٣ - ٢٠٢٤	الانحراف المعياري لعام ٢٠٢٣ - ٢٠٢٤
المعلومات الهامة في المقرر	٢.٤٠٠٠	٧٥٣٩٤.	٢.٨٠٠٠	٤١٠٣٩.	٢.٩٥٠٠	٢٢٣٦١.
تراعي الأمثلة الوزن النسبي للمقرر	١.٦٠٠	٦١٥٥٩.	٢.٠٠٠٠	٤٧٠١٦.	٢.٨٠٠٠	٤١٠٣٩.
الأهداف مناسبة حسب توصيف	٢.٧٠٠٠	٤٧٠١٦.	٢.٨٠٠٠	٤١٠٣٩.	٣.٠٠٠٠	٠.٠٠٠٠

المقرر							
شمولية الاختبار للمقرر	٢.٨١٠٠		٣.٠٠٠٠		٣.٠٠٠٠		٣.٠٠٠٠
الأهداف المعرفية المحددة بالمقرر (التذكر - الفهم - التطبيق - التحليل - التقويم)							
الاختبار يقيس الأهداف الوجدانية	١.٣٠٠٠	٤٧٠١٦.	١.٤٥٠٠	٦٠٤٨١.	١.٥٥٠٠	٦٠٤٨١.	١.٦٠٠٠
الاختبار يقيس الأهداف المهارية	١.٤٠٠٠	٥٠٢٦٢.	١.٥٥٠٠	٣٦٦٣٥.	١.١٥٠٠	٣٦٦٣٥.	١.٦٠٠٠
السؤال يستدعي استجابة محددة	٢.٨٥٠٠	٣٦٦٣٥.	٢.٩٥٠٠	٢٢٣٦١.	٣.٠٠٠٠		٣.٠٠٠٠
الأسئلة تتناسب مع أهداف المادة	٢.٨٠٠٠	٤١٠٣٩.	٣.٠٠٠٠		٣.٠٠٠٠		٣.٠٠٠٠
المعيار الرابع	٢٠.٣	١.٦٨١٧٧	٢٢.٧٥	٠.٧٣٥٥٤	٢٣.٠٥	٨٤٣٢٦.	٢٣.٤٥

وأشارت بيانات جدول (٨) أن متوسط تحقق المعيار الرابع (ارتباط الاختبار بالتوصيف) لعام (٢٠٢٠ - ٢٠٢١) وحازت العبارة رقم (٩) المرتبة الأولى ونسبة تحقق عالية ، وحازت العبارة رقم (٦) المرتبة الأخيرة بنسبة تحقق (٤٣%) بما يُعني ضعف توفر معيار يقيس الاختبار الأهداف الوجدانية المحددة للمقرر ، كما حازت العبارة رقم (٧) نسبة تحقق منخفضة بما يُعني

ضعف توفر معيار يقيس الاختبار الاهداف المهارية المحددة بالمقرر في مستوياتها (الذكر - الفهم - التطبيق - التحليل - التقويم).

بينما تشير بيانات تحقق المعيار الرابع لعام (٢٠٢١ - ٢٠٢٢) نسبة تحقق (٨٥ %) ، وحازت العبارة رقم (٤ - ٩) المرتبة الاولى بنسبة تحقق عالية ، في حين حازت العبارة (٦) المرتبة الأخيرة بما يُعني ضعف توفر معيار يقيس الاختبار الأهداف الوجدانية المحددة بالمقرر ، كما حازت العبارة (٧) على نسبة تحقق منخفضة بما يُعني ضعف توفر معيار يقيس الاختبار الأهداف المهارية المحددة بالمقرر ، كما حازت العبارة (٥) على نسبة تحقق منخفضة بما يُعني ضعف توفر معيار يقيس الاختبار الأهداف المعرفية المحددة بالمقرر في مستوياتها المختلفة (التذكر - الفهم - التطبيق - التحليل - التقويم) .

تشير بيانات تحقق المعيار الرابع لعام (٢٠٢٢ - ٢٠٢٣) نسبة تحقق المعيار بلغ (٩٥.٢٥) وبانحراف معياري (١.٢) ، وحازت العبارة رقم (٣ - ٤ - ٨ - ٩) المرتبة الأولى بنسبة تحقق عالية ، في حين حازت العبارة (٧) المرتبة الأخيرة بما يُعني ضعف توفر معيار يقيس الاختبار الأهداف المهارية المحددة بالمقرر ، كما حازت العبارة (٦) على نسبة تحقق منخفضة بما يُعني ضعف توفر معيار يقيس الاختبار الأهداف الوجدانية المحددة بالمقرر ، كما حازت العبارة (٥) على نسبة تحقق منخفضة بما يُعني ضعف توفر معيار يقيس الاختبار الاهداف المعرفية المحددة بالمقرر في مستوياتها المختلفة (التذكر - الفهم - التطبيق - التحليل - التقويم) .

كما تشير بيانات تحقق المعيار الرابع لعام (٢٠٢٣ - ٢٠٢٤) أن متوسط تحقق المعيار بلغ (٢٦.٤٥) وانحراف معياري (١.١٤) نسبة تحقق (٨٥ %) ، وحازت العبارة رقم (٣ - ٤ - ٨ - ٩) المرتبة الاولى بنسبة تحقق عالية ، في حين حازت العبارة (٦) المرتبة الأخيرة بما يُعني ضعف توفر معيار يقيس الاختبار الأهداف الوجدانية المحددة بالمقرر ، كما حازت العبارة (٧) على نسبة تحقق منخفضة بما يُعني ضعف توفر معيار يقيس الاختبار الأهداف المهارية المحددة بالمقرر ، كما حازت العبارة (٥) على نسبة تحقق منخفضة بما يُعني ضعف توفر معيار

يقيس الاختبار الاهداف المعرفية المحددة بالمقرر في مستوياتها المختلفة (التذكر - الفهم - التطبيق - التحليل - التقويم) .

وللتحقق من وجود فروق بين متوسطات درجات تحليل الاختبار للأعوام (٢٠٢٠ : ٢٠٢٤) في ضوء المعيار الرابع من معايير جودة الاختبار وتم استخدام (تحليل التباين ذي الأربع اتجاهات) كما هو موضح بجدول (٩) .

جدول (٩) دلالة الفروق لمتوسطات درجات تحليل الاختبار للأعوام (٢٠٢٠ - ٢٠٢٤) في ضوء المعيار الرابع من معايير جودة الاختبارات

الأعوام	الفروق بين المتوسطات لعام ٢٠٢٠-٢٠٢١	الفروق بين المتوسطات لعام ٢٠٢١-٢٠٢٢	الفروق بين المتوسطات لعام ٢٠٢٢-٢٠٢٣	الفروق بين المتوسطات لعام ٢٠٢٣-٢٠٢٤	مستوي الدلالة
٢٠٢٠-٢٠٢١	-	-٢.٧٠٠٠	-٣.٠٥٠٠	-٣.٦٠٠٠	دالة إحصائية
٢٠٢١-٢٠٢٢	-٢.٧٠٠٠	-	-٣.٥٠٠٠	-٩.٠٠٠	غير دالة احصائيا
٢٠٢٢-٢٠٢٣	-٣.٠٥٠٠	-٣.٥٠٠٠	-	-٥.٥٠٠٠	غير دالة احصائيا
٢٠٢٣-٢٠٢٤	-٣.٦٠٠٠	-٩.٠٠٠	-٥.٥٠٠٠	-	غير دالة احصائيا

تبين من جدول (٩) أن قيمة الجدول السابق أن قيمة فرق متوسط درجات تحليل الاختبار لعام ٢٠٢٠ في ضوء المعيار الرابع، وبين إختبارات (٢٠٢١) يساوي (٢.٧٠) وبذلك يكون دال إحصائياً ، كما أن الفرق بين إختبارات عام (٢٠٢٠) و إختبارات عام (٢٠٢٢) يساوي (٣.٠٥) وهو غير دال إحصائياً ، كما أن الفرق بين إختبارات عام (٢٠٢٠) وإختبارات عام (٢٠٢٣) يساوي (٣.٦٠) وبذلك يكون غير دال إحصائياً ، وأن قيمة فرق متوسط درجات الإختبار (٢٠٢١) في ضوء المعيار الرابع وبين إختبارات ٢٠٢٢ يساوي (٠.٣٥) وهو غير دال إحصائياً ، كما أن الفرق بين إختبارات عام (٢٠٢١) وإختبارات عام (٢٠٢٣) يساوي (٠.٩٠) وهو غير دال إحصائياً .

وبذلك يتضح من الجدول السابق أن قيمة فرق متوسط درجات تحليل الإختبارات لعام (٢٠٢٢) وبين اختبارات عام (٢٠٢٣) في ضوء المعيار الرابع يساوي (٠.٥٥٠) و هو غير دال إحصائياً .

وبذلك تم قبول الفرض الرابع وتوجد فروق ذات دلالة إحصائية بين متوسط درجات تحليل الإختبار للأعوام (٢٠٢٠: ٢٠٢٤) في ضوء المعيار الرابع من معايير جودة الاختبار (ارتباط الاختبار بالتوصيف) .

٥- للتحقق من صحة الفرض الخامس للبحث وهو توجد قفروق ذات دلالة احصائية بين متوسط درجات تحليل الاختبار للأعوام (٢٠٢٠ - ٢٠٢٤) في ضوء المعيار الخامس من معايير جودة الاختبار (مضمون الاختبار ، وتم استخدام المتوسط الحسابي والانحراف المعياري كما هو موضح بجدول (١٠))

جدول (١٠) المتوسطات والانحراف المعياري لدرجات تحليل الإختبار للأعوام (٢٠٢٠ :

٢٠٢٤) في ضوء المعيار الخامس من معايير جودة الإختبار

العبارة	المتوسط لعام	الانحراف المعياري لعام	المتوسط لعام	الانحراف المعياري لعام	المتوسط لعام	الانحراف المعياري لعام	المتوسط لعام	الانحراف المعياري لعام
	٢٠٢٠ -	٢٠٢١	٢٠٢٢ -	٢٠٢٣	٢٠٢٢ -	٢٠٢٣	٢٠٢٣ -	٢٠٢٤
تضمن الاختبار أسئلة مختلطة ما بين المقالي والموضوعي	٢.١٠٠٠	٨٥٢٢٤.	٢.٨٠٠٠	٤١٠٣٩.	٢.٨٥٠٠	٣٦٦٣٥.	٢.٨٥٠٠	٣٦٦٣٥.
تحديد الإجابة على السؤال	٢.٦٠٠٠	٥٠٢٦٢.	٢.٨٥٠٠	٣٦٦٣٥.	2.7000	٤٧٠١٦.	٢.٨٥٠٠	٣٦٦٣٥.
ترتيب منطقي لل فقرات الفرعية للسؤال	٢.٨٠٠٠	٤١٠٣٩.	١.٦٠٠٠	٠.٠٠٠٠	٢.٧٠٠٠	١.٠٠٠٠	٢.٧٠٠٠	٧.٠٠٠٠
يستخدم واضع الإمتحان لألفاظ (قارن - انقد -	١.٦٥٠٠	٥٨٧١٤.	٢.٠٠٠	٥٨٧١٤.	٢.٢٥٠٠	٥٧١٢٤	٢.٦٥٠٠	٥٨٧١٤.

صنف - لخص (حدد)							
تحديد السؤال تحديد واضح	٢.٧٥٠٠	٤٤٤٢٦.	٢.٨٥٠٠	٢٢٣٦١.	٢.٩٥٠٠	٢٢٣٦١.	٢.٩٥٠٠
تتفاوت الأسئلة من حيث الصعوبة	٢.١٠٠	٥١٠٤٢.	٢.٤٥٠٠	٤٤٤.٢٦	٢.٦٠٠٠	٤١٠٣٩.	٢.٧٠٠٠
قدرة الطالب على توظيف المعلومات	١.٣٥٠٠	٥٥٠١٢.	١.٦٥٠٠	٤٥٠١٢.	١.٧٥٠٠	٤٥٠١٢.	١.٧٥٠٠
خلو الأسئلة من العبارات الموحية بالإجابة	٢.٤٠٠٠	٣٧٠١٦.	٢.٦٥٠٠	٤٤٤٢٦.	٢.٧٥٠٠	٤٤٤٢٦.	٢.٨٥٠٠
يتناسب زمن الاختبار مع عدد الأسئلة	١.٨٠٠٠	٣٠٧٧٩.	١.٩٥٠٠	٣٠٧٧٩.	٢.٣٠٠٠	٤٧٠١٦.	٢.٤٠٠٠
المعيار الخامس	١٨.٥٥	٤٧.٠١٥					

تشير البيانات الواردة في الجدول السابق أن متوسط تحقق المعيار الخامس (مضمون الاختبار بالتوصيف) لعام (٢٠٢٠-٢٠٢١) بلغ (٢٤.٨٥) بانحراف معياري (١,٤٦) بنسبة تحقق (٦٨%) ،وحازت العبارة (٣) المرتبة الأولى ونسبة تحقق عالية ، وحازت العبارة (٧) على المرتبة الأخيرة بما يُعني ضعف توفر معيار بقدرة الطالب على توظيف المعلومات ، كما حازت العبارة (٤) على نسبة تحقق منخفضة بما يُعني توفر معيار وتجنب واضح الاختبار لأسئلة معينة مثل ناقش ، أو اكتب ، أو اشرح ويستخدم بدلاً منها لخص - حدد- انقد وهكذا ، كما تشير لبيانات الواردة في الجدول السابق أن متوسط تحقق المعيار الخامس (مضمون الاختبار بالتوصيف) لعام (٢٠٢١-٢٠٢٢) بلغ (٢٧.٨٥) بانحراف معياري (١,٠٤) بنسبة تحقق (٧٧%) ،وحازت العبارة (٢) المرتبة الأولى ونسبة تحقق عالية ، وحازت العبارة (٣) على المرتبة الأخيرة بما يُعني ضعف توفر معيار بقدرة الطالب على توظيف المعلومات ، كما حازت العبارة (٤) على نسبة تحقق منخفضة بما يُعني ضعف توفر معيار مضمون الاختبار بالتوصيف.

وتشير البيانات الواردة في الجدول السابق أن متوسط تحقق المعيار الخامس (مضمون الاختبار بالتوصيف) لعام (٢٠٢٢-٢٠٢٣) بلغ (٢٧.٨٠) بانحراف معياري (١,٧) بنسبة تحقق (٧٧%) ،وحازت العبارة (٣) المرتبة الأولى ونسبة تحقق عالية ، وحازت العبارة (٧) على المرتبة الأخيرة بما يُعني ضعف توفر معيار بقدرة الطالب على توظيف المعلومات ، كما حازت العبارة (٤) على نسبة تحقق منخفضة بما يُعني ضعف توفر معيار تقيس الأسئلة بقدرة الطالب على توظيف المعلومات.

وتشير البيانات الواردة في الجدول السابق أن متوسط تحقق المعيار الخامس (مضمون الاختبار بالتوصيف) لعام (٢٠٢٣-٢٠٢٤) بلغ (٢٨.٣٠) بانحراف معياري (٤٤٦) بنسبة تحقق (٧٨%) ،وحازت العبارة (١) المرتبة الأولى ونسبة تحقق عالية ، وحازت العبارة (٧) على المرتبة الأخيرة بنسبة تحقق (٥٨%) بما يُعني ضعف توفر معيار بقدرة الطالب على توظيف المعلومات .

وللتحقق من وجود فروق بين متوسطات درجات تحليل محتوى الاختبار للأعوام (٢٠٢٠-٢٠٢٤) في ضوء المعيار الخامس من معايير الجودة ، وتم استخدام (تحليل التباين ذي الأربع اتجاهات) كما هو موضح بجدول (١١).

جدول (١١) دلالة الفروق لمتوسطات درجات تحليل المحتوى للأعوام (٢٠٢٠-٢٠٢٤) في ضوء المعيار الخامس

الأعوام	الفروق بين المتوسطات لعام ٢٠٢١-٢٠٢٢	الفروق بين المتوسطات لعام ٢٠٢٢-٢٠٢٣	الفروق بين المتوسطات لعام ٢٠٢٣-٢٠٢٤	مستوي الدلالة
٢٠٢١-٢٠٢٢	-	-٣.٩٥٠٠٠	-٣.٤٥٠٠٠	٠.٠٠٠
٢٠٢٢-٢٠٢٣	-٣.٠٠٠٠	٠.٥٠٠٠٠	-٤.٥٠٠٠٠	٠.٩٩٩
٢٠٢٣-٢٠٢٤	-٢.٩٥٠٠٠	-	-٠.٥٠٠٠٠	٠.٦٢٨
٢٠٢٤-٢٠٢٣	٣.٤٥٠٠٠	-٤.٥٠٠٠٠	-	٠.٧٠٢

وتبين من جدول (١١) أن قيمة فروق متوسط درجات تحليل الاختبار لعام (٢٠٢٠) في ضوء المعيار الخامس وبين اختبارات عام (٢٠٢١) يساوي (٣.٠٠) وهو دال احصائيًا ، كما أن الفرق بين الاختبار لعام (٢٠٢٠) في ضوء المعيار الخامس وبين اختبارات عام (٢٠٢٢)

يساوي (٢٠٩٥) وهو غير دال احصائياً ، واختبارات عام (٢٠٢٣) يساوي (٣.٤٥) وهو غير دال احصائياً، ومن خلال البيانات في الجدول السابق تبين أن قيمة فرق متوسط درجات تحليل الاختبار لعام (٢٠٢١) في ضوء المعيار الخامس وبين اختبارات عام (٢٠٢٢) يساوي (٠.٥٠٠) وهو دال احصائياً ، كما أن الفرق بين الاختبار لعام (٢٠٢١) في ضوء المعيار الخامس وبين اختبارات عام (٢٠٢٣) يساوي (٠.٤٥٠٠) وهو دال احصائياً .

كما قيمة فرق متوسط درجات تحليل الاختبار لعام (٢٠٢٢) في ضوء المعيار الخامس وبين اختبارات عام (٢٠٢٣) يساوي (٠.٥٠٠) وهو غير دال احصائياً . وبذلك تم قبول الفرض الخامس وتوجد فروق ذات دلالة احصائية بين متوسط تحليل محتوى الاختبار للأعوام (٢٠٢٠ : ٢٠٢٤) في ضوء المعيار الخامس (مضمون الاختبار بالتوصيف) من معايير جودة الاختبار .

وتوصلت نتائج البحث إلى أن هناك تحسن ملحوظ في نسب استيفاء الاختبار لمعايير جودة الاختبار بداية من عام (٢٠٢١) ، وكذلك تجاوز نسب استيفاء المعايير الأساسية وتعليمات الاختبار ولغة الاختبار (٩٠%) ، وأيضاً يُعد معيار لغة الاختبار أعلى نسبة لمعايير جودة الاختبار على مدار أربعة أعوام ، بينما انخفضت نسبة استيفاء الاختبار لمعيار ارتباط الاختبار بالتوصيف ومضمون الاختبار عن (٩٠%).

وكذلك تشير نتائج التحليل الإحصائي بصفة عامة إلى استيفاء الاختبارات لمعايير جودة الاختبار الفرعية بنسبة تزيد (٨٠%) إلا أن هناك معايير فرعية تنخفض نسبة استيفائها عن ذلك وفيما يلي عرض لهذه المعايير :

المعيار الأول (البيانات الأساسية) : معاييرها الفرعية : توقيع أستاذ المادة ورئيس القسم على ورقة الامتحان.

المعيار الثاني (تعليمات الاختبار) : معاييرها الفرعية ترك مسافة بين كل سؤال وما يليه، وكذلك تحديد كيفية الإجابة عن السؤال، وتحديد درجة الإجابة لكل سؤال .

المعيار الثالث (لغة الاختبار) : معيارها الفرعي خلو الاختبار من المصطلحات التي لها أكثر من معنى ، ومناسبة مستوى الاختبار ، ووضع علامات الترقيم بالسؤال.

المعيار الرابع (ارتباط الاختبار بالتوصيف) معايير الفرعية يقيس الاختبار الأهداف المعرفية والوجدانية والمهارية المحددة بالمقرر، وكذلك مراعاة التوازن النسبي لموضوعات المقرر .
المعيار الخامس (مضمون الاختبار) معايير الفرعية هنا الملاحظة والمعرفة السابقة للمقرر، وتجنب الأسئلة المفتوحة كناقش - اكتب - اشرح واستخدام بدلاً منها قارن - انقد - صنف وهكذا ، وأيضاً تقيس الأسئلة قدرة الطالب على توظيف المعلومات ، وعدد الأسئلة يتناسب مع زمن الاختبار .

تفسير النتائج

تتفق نتائج البحث الحالي مع نتائج دراسة كل من دراسة (محمد محمد عبد المنعم ، ٢٠٠٩) (ومحمد الزهراني ، ٢٠٠٩) بوجود أخطاء عديدة لغوية ونحوية ووجود العديد من الأخطاء اللغوية والنحوية والطباعية في أسئلة كثير من الإختبارات ووجود كثير من الأخطاء العلمية والفنية في جودة الأسئلة بالإضافة إلي تكرار بعض الأسئلة بصورة كاملة أو بصورة جزئية من إختبارات سابقة كما أن كثير من أسئلة تلك الإختبارات قدر ترد إجاباتها في نفس الإختبار في نصوص أسئلة أخرى وقد إستدل من هذه النتائج على أن معظم الإختبارات لم تبني في ضوء مواصفات الإختبار الجيد وفي ضوء جدول المواصفات ، وكذلك تتفق مع دراسة (عبدالكريم السوداني ، ٢٠٠٦) أن مستويات التفكير (تصنيف بلوم للمجال المعرفي التي تستثيرها أسئلة الإختبارات النهائية التي يتعرض لها الطلاب كلية التربية قسم علوم الحياه ، وقد إعتد على أسلوب تحليل المحتوي في معرفة هذه المستويات وبعد التحقق من صحة التحليل وحساب معامل الإتفاق بين المحللين وبلغت عدد الأسئلة المحللة (١٦٦) سؤال فقد كشفت الدراسة عن خلو الأسئلة من مستويات التطبيق والتركيب والتقويم بينما الصفوف الأعلى كان مستوي التذكر أعلى المستويات تحققاً ، وأيضاً دراسة (علي الحديبي ، وهناء إبراهيم ، ٢٠١٤) هدفت إلي وضع قائمة لمعايير الجودة التعليمية في تصميم أوراق الأسئلة الإختبارية بالجامعة الإسلامية وتضمنت ثلاثة معايير أساسية ينبثق عنها (٢٥) مؤشراً حيث يتضمن المعيار الأول البيانات الأساسية عن الجامعة والمقرر الدراسي ، والمعيار الثاني يركز على الجوانب الفنية التي تتعلق بإخراج ورقة الإختبار وتنظيمها وترتيبها ، والمعيار الثالث يركز على

الصياغة العلمية والصحيحة للأسئلة سواء كانت مقالية أو موضوعية وبفحص وتقويم أوراق الإختبارات في ضوء تلك المعايير إتضح أن المعيار الثاني (الإهتمام بإخراج وتنظيم الورقة الإمتحانية) جاء في ترتيب الأول يليه المعيار الأول (البيانات الأساسية للجامعة والمقرر) ويليه المعيار الثالث (الصياغة العلمية الصحيحة للأسئلة) مما يدل على أن معلم المادة يهتم بإخراج الورقة وتنظيمها ثم بكتابة البيانات الأساسية ويأتي في آخر إهتمامهم صياغة الأسئلة صياغة علمية ، وتختلف نتائج البحث الحالي مع دراسة (منال حسن ، ٢٠١٣) والتي هدفت إلي تقويم أسئلة الإختبارات النهائية للأقسام الأدبية والعلمية لجميع المستويات بكلية التربية بحفر الباطن جامعة الدمام في ضوء معايير الجودة وتوصلت نتائج الدراسة أن أغلب أسئلة الإختبارات التحصيلية في المواد العلمية والأدبية والتربوية تقيس المستويات المعرفية الدنيا فأغلبها تقيس أدنى المستويات المعرفية وهو مستوى التذكر وفي أحسن الأحوال قد تصل إلي مستوى الفهم ، كما توصلت إلي أن معظم الأسئلة أهملت معيار الشمول ولم تلتزم بالمعايير التي خصصت لتنظيم وإخراج ورقة الإختبار ، وكذلك دراسة (حمد الكلثوم ، ٢٠١٥) التي تحدد معايير درجة الورقة الإمتحانية من حيث الشكل والمضمون وجودة الأسئلة المقالية والموضوعية ، ثم التعرف علي مدى توافر هذه المعايير في الورقة الإمتحانية ودراسة الفروق في نسب تحقق هذه المعايير في الأوراق الإمتحانية وقد كشفت النتائج أن الأسئلة التي تقيس مستوى التذكر كانت بالمرتبة الأولى مقارنة بالأسئلة التي تقيس باقي المستويات المعرفية ، وكذلك توجد فروق ذات دلالة إحصائية فيما يتعلق بتحقيق معايير جودة الإختبار وتنوع الأسئلة لتشتمل على العديد من المستويات المعرفية ومعايير جودة الأسئلة المقالية والموضوعية لصالح جامعة أم القرى ، وأيضًا دراسة (منار الشامي ، سهام الشافعي ، ٢٠١١) أن مدى جودة الإختبارات التحصيلية وتحديد الأهمية النسبية لأنواع الأسئلة المختلفة في الإختبارات التحصيلية فقد تم إنخفاض جودة الإختبارات التحصيلية حيث لم تصل أي منها لدرجة الجودة المطلوبة وقد تم تقييم جودة الإختبارات في ضوء أربعة معايير إلي معيار التحليل العام والذي يتضمن شكل الورقة الإختبارية و أهم الأخطاء اللغوية وكتابة تعليمات الكتابة والدرجة الكلية للإختبارات ، معيار مضمون الأسئلة والذي يتضمن وضوح عبارات الأسئلة وعلاقتها بأهداف

المادة ومدي شموليتها لموضوعات المقرر ، معيار الأسئلة المقالية وكذلك معيار الأسئلة الموضوعية، وتوصلت نتائج الدراسة إلى أن لمعيار التحليل العام الأثر الأكبر في درجات جودة الإختبار يليه معيار مضمون الأسئلة ثم معيار جودة الأسئلة المقالية ثم معيار جودة الأسئلة الموضوعية كما أن هناك تفاوت في نسب أنواع الأسئلة في الإختبارات وركزت بشكل كبير على الأسئلة الموضوعية حيث بلغت نسبتها (٨٣.٠٢ %) في مقابل الأسئلة المقالية التي بلغت نسبتها (١٦.٠٨ %) كما وجدت علاقة ارتباطية عكسية بين ضغوط العمل وبين جودة الإختبارات ، وأيضاً دراسة (هدي صالح ، حنان محمد ، ٢٠١٧) والتي هدفت إلى تحديد الأخطاء الشائعة في بناء مفردات أسئلة الصواب والخطأ وذلك بفحص عينة مكونة من (٤٠) نموذج من أسئلة معلمات علم النفس للفصل الدراسي الثاني لتحديد الأخطاء الشائعة فيها وقد كشفت النتائج عن وجود أربعة أخطاء أولها نقل العبارة نصاً من الكتاب المدرسي ، والثلاثة الأخرى مرتبطة بإخراج أسئلة الصواب والخطأ وهم عبارات غير متقاربين في الطول والتوزيع غير العشوائي للعبارات فنسبة عبارات الصواب أقل من عبارات الخطأ وكذلك توصلت إلى عدم وجود فروق ذات دلالة إحصائية في متوسطات معاملات السهولة بين مفردات الإختبار الجيد والإختبار غير الجيد بينما وجدت الفروق في معاملات التمييز بين مفردات الإختبارين تعود إلى الأخطاء الشائعة في بناء مفردات أسئلة النموذج غير الجيد ، بينما وجدت فروق ذات دلالة إحصائية بين معاملات الثبات للإختبارين لصالح الإختبار الجيد بينما لا توجد الفروق بينهما في معاملات صدق المحك .

ومن جانب آخر هناك دراسات إهتمت بالأخطاء بصفة عامة مثل دراسة (على صلاح حسن ، ٢٠١٨) حيث قام بإعداد قائمة بتدقيق الأخطاء في بناء الإختبارات الموضوعية وفحص مجموعة عشوائية من الإختبارات الموضوعية لأعضاء هيئة التدريس من كليات وأقسام مختلفة بجامعة أسيوط ، وقد كشفت النتائج نسبة أسئلة الصواب والخطأ كانت نسبة الأخطاء الإملائية أقل من ٢٠ % بينما كانت نسب بعض الأخطاء تتراوح ما بين ٢٥ % و ٩٤ % وهي نسب كبيرة ، أما بالنسبة لأسئلة الإختبار من متعدد فتراوحت نسبة الأخطاء فيها ما بين (٢٥ %) و (١٠٠ %) موزعة ما بين مدخل الفقرة والبدائل والإجابات الصحيحة وهي نسب كبيرة أيضاً مما

يدل على إهتمام كثير من أعضاء هيئة التدريس بالخطوات العلمية لبناء الإختبارات ومراعاة شروط وقواعد بناء الإختبارات الموضوعية كما كشفت نتائج الدراسة عن وجود فروق ذات دلالة إحصائية بين إختبار الموضوع الجيد والإختبار المتضمن على الأخطاء في الخصائص السيكمترية للإختبار ككل من حيث متوسطات معاملات الثبات والصدق للإختبار ولفقراته من حيث معاملات متوسطات السهولة والاصعوبة والتمييز لصالح الصورة الجيدة ، ولذلك تبدو نتائج البحث منطقية

توصيات البحث

- ١- تنوع الأسئلة ما بين المقالي والموضوعي بما يتناسب مع طبيعة كل مقرر وأداء الطلاب بكل مستوياتهم.
- ٢- تدريب وتقديم ورش عمل لأعضاء هيئة التدريس وتقديم الدعم الفني والبرامج الالكترونية لإعداد امتحانات جيدة خالية من أخطاء القياس لضمان موضوعية الامتحانات.
- ٣- تدريب أعضاء هيئة التدريس على كيفية إعداد بنوك أسئلة وتدرجها وفقاً لنظرية القياس الموضوعية واختبار مدى جدواها وفق معايير جودة الامتحانات، وكيفية إعداد أدوات القياس اللازمة.
- ٤- الإرشاد الأكاديمي للطلاب والتعريف بأهمية التقويم في ضوء معايير جودة الاختبار .
- ٥- الإهتمام بأراء الطلاب وتقييم اتجاهاتهم حول الطرق المستخدمة في الإمتحانات لتعزيز جوانب القوة وبناء خطط تحسينية لجوانب الضعف.
- ٦- تشكيل لجنة مراجعة داخلية بالكليات لتقييم الاختبارات في ضوء معايير جودة الاختبارات وتقديم الدعم لأعضاء هيئة التدريس ، وتقديم مكافآت تشجيعية لأفضل اختبار استوفى الشروط والمعايير.

البحوث المقترحة

- ١- بناء برامج تدريبية لتحسين مهارات أعضاء هيئة التدريس في تطبيق التكنولوجيا الحديثة لتقويم الإمتحانات .

- ٢- دراسات استكشافية لتقويم الإمتحانات التحصيلية لدى طلاب المرحلة الابتدائية.
- ٣- دراسة أثر نوع ومستوى مشاركة الطلاب في وضع اسئلة على محتوى المقرر في ضوء معايير جودة الاختبارات.
- ٤- تقويم الاختبارات بالكليات العملية في الجامعة في ضوء معايير جودة الاختبار.

المراجع

- ١- إبراهيم سعد شحاتة (٢٠٠٥). البحوث العلمية والتربوية بين النظرية والتطبيق . القاهرة: مكتبة الدار .
- ٢- إبراهيم عبدالله الحسينان ، محمود جابر حسن أحمد (٢٠١٤) . تقويم إختبارات المقررات التربوية في كلية التربية جامعة المجمعة في ضوء معايير الجودة ، مجلة كلية التربية جامعة الأسكندرية ، (٤) ، ٩٠-٢٣ .
- ٣- أحمد الثوبية (٢٠٠٤). استراتيجيات التقويم وأدواته ،الإطار النظري تحت إشراف الكنديين لمشروع الإقتصاد المبني على المعرفة .
- ٤- أحمد علي كنعان (٢٠٠٥). الإعتماد الأكاديمي لمؤسسات التعليم العالي في الجمهورية العربية السورية بين الواقع والمأمول ، دراسة ميدانية في كلية التربية في جامعة دمشق ، برنامج إعداد المعلم أنموذجًا ، المؤتمر التربوي الخامس - جودة التعليم الجامعية ، مملكة البحرين ، ١١-١٣ أبريل .
- ٥- أحمد عودة (٢٠٠٥) . القياس والتقويم في العملية التدريسية ، الأردن ، دار الأهل للنشر والتوزيع .
- ٦- إسماعيل صبري ، محب كامل الرافعي (٣٦٢) . التقويم التربوي: أسسه وإجراءاته ، الرياض ، مكتبة الرشيد .
- ٧- أكرم عادل الشير ، أريج عادل ، (٢٠١٢). أثر استخدام إستراتيجيات التقويم البديل وأدواته في تقويم تعلم الرياضيات واللغة العربية ، في الأردن ، مجلة العلوم النفسية والتربوية ، مارس ، ٢٤١-٢٧٠ .
- ٨- أماني الغامدي ، سماح زغاري (٢٠١٨) . نواتج التعلم واختبارات المنتصف في كلية التربية بجامعة الإمام عبدالرحمن فيصل :دراسة تحليلية ،مجلة الحكمة للدراسات التربوية والنفسية ، (١٦) ، ٣٦-٨ .
- ٩- بندر بن زيد الأمير (٢٠١٦) . الأخطاء الشائعة في بناء مفردات الاختبار في الاختيار من متعدد وتأثيرها على الخصائص السيكمترية للاختبارات التحصيلية ،المجلة العربية للعلوم الإجتماعية ، (٩) ، ١٣٧-١٧٥ .
- ١٠- حسيب محمد حسيب (٢٠٢٠) . تطوير نظام القياس والتقويم للتعليم قبل الجامعي " رؤى مقترحة " ، المجلة العربية للقياس والتقويم ، (٢) ، ص ص ٦٣-٨٣ .4

- ١١- حمد الكلثوم (٢٠١٥). تقويم بعض اختبارات تخصص تربية اسلامية في جامعتي أم القرى والدمام في ضوء معايير جودة الاختبار ، دراسات عربية في التربية وعلم النفس ، (٦٦) ، ٢١١-٢٣٦.
- ١٢- خديجة ضيف الله القرشي (٢٠١٢). تقويم الورقة الإختبارية لطالبات الكالوريوس بكلية التربية ، جامعة الطائف وفق معايير الجودة، المؤتمر الإقليمي الثاني لعلم النفس ، من ٢٩ / ١١:١٢/٢٠١٠ ، رابطة الإخصائية النفسية المصرية ، ١٩-٤٩.
- ١٣- ريان حسن فكري(١٩٩٥). التدريس :أهدافه ، أسسه ، أساليبه ، تقويم نتائجه وتطبيقاته .القاهرة: عالم الكتب.
- ١٤- زياد بركات وصباح عبدالهيا(٢٠٠٧). مدى تحقيق أسئلة الإمتحانات النهائية في جامعة القدس المعروضة للأهداف التعليمية تبعاً لهرم بلو ، جامعة القدس المفتوحة للإبحاث والدراسات ، (٩) .
- ١٥- سناء إبراهيم أبودقه (٢٠٠٧) . القياس والتقويم الصفّي ، المفاهيم والإجراءات للتعليم الفعال - غزة مكتبة أفاق .
- ١٦- صلاح الدين علام (٢٠٠٩) . القياس والتقويم التربوي في العملية التدريسية ، كلية التربية ، جامعة الأزهر ، دار المسيرة للنشر والتوزيع (الأردن - عمان) ١٠.
- ١٧- صلاح الدين محمود علام (٢٠٠٤) . التقويم التربوي المؤسس ، القاهرة، دار افكر العربي.
- ١٨- ضياء الدين عبدالواحد (٢٠١٥).تقويم الطلاب بأقسام المكتبات والعلومات في ضوء معايير الجودة والإعتماد : دراسة تحليلية للاختبارات التحصيلية ، المجلة الدولية لعلوم المكتبات والمعلومات ، (٢) ، ٤٠، ١١-٣٩.
- ١٩- عبد الحميد السورتي (١٩٩٨). المعجم الموسوعي للمصطلحات الإدارية والإجتماعية ، القاهرة : دار الفكر العربي.
- ٢٠- عبد الرحمن عدس (١٩٩٩).علم النفس التربوي : نظرة معاصرة .عمان : دار الفكر.
- ٢١- عبدالكريم عبدالصمد السوداني (٢٠٠٦). مستويات التفكير في الاختبارات التحصيلية النهائية لقسم علوم الحياة ، مجلة القادسية في الأدب والعلوم التربوية ، (٥) ، ١ ، ٧٠١٧.
- ٢٢- على الحديبي ، هناء إبراهيم (٢٠١٤).تقويم أوراق الاختبارات بالجامعات الإسلامية بالمدينة المنورة على ضوء معايير الجودة ، مجلة الجمعية التربوية بجامعة أسيوط ،مجلة كلية التربية ،جامعة أسيوط ، (٣٤) ، ٢٠ ، ٣٦٥-٤٠٦.

- ٢٣- على صلاح حسن (٢٠١٨). الأخطاء الشائعة في بناء الاختبارات الموضوعية وتأثيرها على الخصائص السيكمترية للاختبارات التحصيلية لدى طلاب كلية التربية بجامعة أسيوط ، مجلة كلية التربية ، جامعة أسيوط ، (٣٤) ، ٢ ، ٣٦٥-٤٠٦ .
- ٢٤- على محمد زهيد الغامدي (٢٠٠٨) . تصور مقترح انطباق نظام الجودة الشاملة في المؤسسات التربوية والتعليمية في ضوء المواصفة الدولية للجودة ، بحث مقدم في لقاء الجمعية السعودية للعلوم التربوية والنفسية ، فرع القصيم من ٢٨ - ٢٩ ربيع الآخر ١٤٢٩هـ .
- ٢٥- فتحي مصطفى الزيات (٢٠٠٤) . القيمة التنبؤية لتحديد وتشخيص صعوبات التعلم بين نماذج التحليل الكمي ونماذج التحليل الكيفي ، المؤتمر الدولي لصعوبات - ١٤/١١/٢٠٠٦ ، التعلم ، الرياض ، المملكة العربية السعودية .
- ٢٦- فؤاد أبوحطوب ، سيد عثمان ، أمال صادق (٢٠٠٨) . التقويم النفسي ، الطبعة الرابعة ، القاهرة ، مكتبة الانجلو المصرية ، ١٧-٧٥ .
- ٢٧- محمد محمد عبد المنعم (٢٠٠٩) . تقويم أسئلة الاختبارات التحصيلية بكلية المعلمين ، جامعة الملك فيصل في ضوء معايير جودة التقويم ، مجلة البحوث النفسية التربوية ، كلية التربية ، جامعة المنوفية ، (١) ، ٢٠٥-٢٣٦ .
- ٢٨- منار مرسي الدسوقي الشامي ، سهام أحمد رفعت الشافعي (٢٠٠١) . جودة التحقيق في التحليل وعلاقتها بضغوط العمل لدى أعضاء هيئة التدريس في جامعة تبوك ، مجلة بحوث التربية النوعية ، جامعة المنصورة ، (٢٣) ، ١٦٨ - ٢٠٧ .
- ٢٩- منال علي حسن (٢٠١٣) . تقويم أسئلة اختبارات طالبات كلية التربية بحفر الباطن ، القاهرة ، مجلة القراءة والمعرفة ، ١٣٦ ، ٢٦٥-٣٢٢ .
- ٣٠- هدي محمد صالح ، حنان فوزي محمد (٢٠١٧) . تقويم برنامج ماجستير المناهج وطرائق التدريس بمساري اللغة العربية والعلوم بجامعة القصيم في ضوء نواتج التعلم اللازم واكتسابها من وجهة نظر أعضاء هيئة التدريس والطلاب : دراسات في المناهج وطرق التدريس ، كلية التربية ، جامعة عين شمس ، (٢٢٥) ، ١٦-٦٤ .
- ٣١- يحي خليفة الشريفي (٢٠١٣) . تقويم أسئلة الإختبارات النهائية لمادة الأدب في كليات التربية في ضوء مستويات بلوم المعرفية ، مجلة القادسية للعلوم الإنسانية ، (١٦) ، ٢٧٧-٣-٤٢٠ .
- ٣٢- يزيد عيسى السورطي (١٩٩٨) . السلطوية في التربية العربية (المظاهر ، الأسباب ، والنتائج) ، المجلة التربوية ، جامعة الكويت ، (٤٦) .

- 33- Arter,J&Mctighe,j.(2001).Scoring rubrics in the classroom .Using performance criteria for assessing and improving student performance .Thousand Oaks.Corwing press,INC.
- 34- Bacolod,m&Tobias.(2006).School quality and achievement Growth evidence from the Philippines economics of education review dissertation abstract international vol 25 (6) pp63-619.
- 35- Brown,J.D.(2010).Assessment of learning Outcomes in higher education.Routledge.
- 36- Crock,L.&Algina,J.(1986). Introduction to classical and modern test theory.Holt ,Rinehart.and Winston.
- 37- Hopkins,C &Antens,R(1985).Classrom Measurement &Evaluation,2nd Ed,pesaok publisbers Incillonis .
- 38- Phillip.E.(2014).Measuring student Achievement :A study of standardized Testing and is Effect on student Learning Retrieved from, <http://my.jessup.edu/publicpolicy/wp-content/uploads/sites/39/2014/04/Gawthrop-Jeremiah-Final.bdf>.