تقييم امتحانات طلاب كلية التربية ومدى تحقيقها لنواتج التعلم المستهدفة

أ.د/محمود فتحي عكاشة* د/إيمان صلاح ضحا**

المستخلص:

هدفت الدراسة الحالية إلى دراسة مدي تحقيق الامتحانات التحصيلية النهائية في الجامعة (كلية التربية جامعة دمنهور كمثال) لمتطلبات جودة تقييمها في ضوء المواصفات القياسية للورقة الامتحانية ،ودراسة مدى صلاحيتها السيكومترية وتحقيقها لنواتج التعلم المستهدفة ،مع الاعتبار بآراء الطلاب وتحديد مستوي رضاهم عن تلك الامتحانات.

تم تطبيق استبيان رضا الطلاب عن الامتحانات النهائية بعد التحقق من صلاحيتة للتطبيق على عينة من (٦٤٩طالباوطالبة) ،كما تم استخدام قائمة المواصفات القياسية للورقة الامتحانية لفحص عينة من (٤٦) ورقة امتحانية اختيرت لمقررات تربوية لعموميتها علي جميع الشعب موزعة على أربع فصول دراسية متتالية ،كما تم فحص التقارير الفنية التي توفرها برامج التصحيح الآلي لعينة من (١٤) ورقة امتحانية للتحقق من صلاحيتها السيكومترية في قياس نواتج التعلم.وكشفت النتائج أنّ:

١. مستوى رضا الطلاب عن توفر المواصفات الشكلية للأوراق الامتحانية النهائية جاء مُرضيًا،في حين تباينت آرائهم نحو توافر المواصفات الكيفية بها ،كمابلغت النسبة العامة لتحقق المؤشرات الشكلية في عينة الامتحانات التي خضعت للدراسة ٧٣.٣١% بينما كانت نسبة تحقق المؤشرات الكيفية ٦٤.٤١% .

٢. كان متوسط معامل الثبات عبر عينة الامتحانات المصححة آليًا(٠.٧٤ قيمة متوسطة) وزادت فيها نسبة الأسئلة غير المُمَيِّزة التي لم يصل معامل تمييزها إلى (٠.٢)

٣. تراوحت النسبة المئوية لمتوسط تحقق نواتج التعلم (المحسوب من خلال نتائج الطلاب على مفردات الامتحانات ومطابقتها مع نواتج التعلم الموضوعة من أجلها) ما بين ٦٩.٧٦% إلى٩٠.٧٨% بمتوسط ٨٠.٤٨ % على المستوى الاجمالي للأوراق الامتحانية التي خضعت للدراسة وهي قيم مقبولة فقط في حالة ثبات الامتحانات وجودة مفرداتها.

تم تقديم مجموعة من التوصيات لمتخذي القرار في كل من الكلية والجامعة لادراجها في الخطط التصحيحية بالاضافة إلى بعض الأفكار البحثية المقترحة.

الكلمات المفتاحية : تقييم نواتج التعلم / الامتحانات التحصيلية/المواصفات القياسية للورقة الامتحانية.

* أستاذ علم النفس التربوي- العميد الأسبق لكلية التربية - جامعة دمنهور
البريد الالكتروني : m.okasha@edu.dmu.edu.eg
**مدرس علم النفس التربوي-كلية التربية-جامعة دمنهور
البريد الالكتروني : eman_daha@edu.dmu.edu.eg

مقدمة:

تُعد عملية التقويم عامةً و تقويم الطلاب خاصةً من العمليات الأساسية في تحقق الجودة الشاملة للنظام التعليمي وبالتحديد التعليم الجامعي فاصلاح التقويم التربوي ونظم الامتحانات يعتبر مدخلا هامًا من مداخل اصلاح النظم التعليمية ، فجودة التعليم لن تتحقق إلا من خلال طرق تقويم تُسهم في قياس نواتج التعلم المستهدفة وتتحقق في اختباراتها معايير الجودة مما يساعد على متابعة التقدم نحو الأهداف ومقارنة الأداء بمعايير داخلية أو خارجية.

فالتقويم لُب العملية التعليمية وجوهرها فإذا ما صلح التقويم صلُحت هي الأخرى بكافة عناصرها حيث أنه يُوَجّه مسار عمليتي التعليم والتعلم (صلاح الدين علام ،٢٠٠٩)؛ ويُرشِّد القرارات التربوية المتعلقة بالبرامج التعليمية وبالطلاب كانتقالهم إلى مستوى أعلى أو تخرجهم وحصولهم على مؤهل يمكّنهم من مزاولة المهنة أو رسوبهم لاستكمال ما يلزمهم من معارف ومهارات حيث يفترض أن لدى الخريج من المهارات والكفايات ما يُمَكِّنه من خدمة المجتمع بأقصى ما تؤهله قدراته "وفي حالة عدم دقة الامتحانات في قياس مستوى المتعلمين سيتخرج أفراد يواجهون المجتمع بمتطلباته ومشكلاته وليس لديهم من القدرات والكفايات ما يلبي احتياجاته ويتغلب على مشكلاته فلا يتناسب تأهيلهم مع متطلبات التنمية والقدرة التنافسية وفق احتياجات سوق العمل. (حمد الكلثم،٢٠١٥)

ويتطلب التقويم تقييم كافة مدخلات وعمليات ومخرجات العملية التعليمية في منظومة متكاملة بكل مكوناتها من منهج،معلم،متعلم، ادارة ...الخ ؛ويتطلب التقييم أدوات قياس صالحة ومن أهم أدوات القياس التي تُستخدم لتقييم تعلم الطلاب "الامتحانات التحصيلية " لذلك لابد من التأكد من جودتها وقدرتها على تحقيق الهدف منها وهو التعرف على مستوى الطلاب والكشف عن مدى تحقق الأهداف التعليمية التي وضعتها مسبقًا السياسات التعليمية في الخطط والبرامج والمقررات الدراسية والكشف عن مواطن القوة والضعف فيها حتى يمكن تحسين وتطوير العملية التربوية (فؤاد أبو حطب،وآخرون،٢٠٠٨ ; Bacolod ,Tobias ,2006 Brwon;2002;).

لذلك يؤكد التربويون وخبراء القياس ضرورة أن تبنى الامتحانات على أسس ومعايير علمية حديثة ومعاصرة بحيث تقدم بيانات صادقة يمكن الاعتماد عليها .

ولكن كشفت نتائج العديد من الدراسات عن وجود قصورًا في بناء معظم الامتحانات التحصيلية وافتقارها إلى الموضوعية والشمولية ،كما أنها غير مقننة ويصعب الاعتماد على نتائجها في الحكم على الطلاب من أمثلتها:(دراسة أماني الغامدي،سماح زغاري،٢٠١٨؛ ضياء الدين عبد الواحد ،٢٠١٥؛منال حسن،٢٠١٣؛ Entwistle & Entwistle, 1992; Frederickson & Collins, 1989; Ramsden, 1988;Wilson,et al ,2016).

كما استهدفت الكثير من الدراسات تحليل محتوى العديد من الاختبارات سواء كانت دورية أو نهائية في عدد من الجامعات وانتهت نتائجها لوجود كثير من أوجه القصور والأخطاء الشائعة في مفردات هذه الاختبارات التي من أهمها؛ انحصارها في قياس المستويات المعرفية الدنيا وندرة تطرقها لقياس المستويات العليا كالقدرات التحليلية والابداعية الضرورية جدا للتعلم المستمر مدى الحياة . (أماني الغامدي،سماح زغاري،٢٠١٨؛ إبراهيم الحسينان،محمود أحمد،٢٠١٤؛بندر الأمير،٢٠١٦ ؛حمد الكلثم ،٢٠١٥؛ عبد الكريم السوداني،٢٠٠٦؛ على صلاح حسن،٢٠١٨؛محمد عبد المنعم ،٢٠٠٩؛ ؛منار

الشامي،سهام الشافعي،٢٠١١؛منال حسن ،٢٠١٣؛يحيى الشريفي،٢٠١٣؛ Entwistle & Entwistle, 1992; Wilson,et al ,2016))

وقد أشارت (منال حسن ،٢٠١٣) إلى خطورة انحصار أسئلة أساتذة الجامعات في المستويات المعرفية الدنيا خصوصًا لأعضاء هيئة التدريس بكليات التربية حيث أن ذلك يعكس بعدًا تربويًا خطيرًا لأن تركيز أسئلتهم على قياس المستويات المعرفية الدنيا من الممكن أن ينتقل لطلابهم "معلمي المستقبل" ويتبعون نفس الاستراتيجية في اعداد أسئلتهم بعد التخرج لأن نتائج الدراسات أثبتت أن ما يتم استخدامه في اعداد المعلمين أثناء الخدمة من طرائق وأساليب تقويمية مختلفة يتبعه المعلم بعد تخرجه مع تلاميذه أثناء التدريس.

وبذلك فقد تبيّن محدودية الأساليب المتبعة في تقويم الطلاب وقلة جدواها في التشخيص الدقيق لمستوياتهم كما يبدو أنه على الرغم من الجهود المبذولة في تطوير منظومة الامتحانات التحصيلية النهائية فلا تزال هذه الجهود شكلية أكثر منها موضوعية فمن خلال الملاحظات الميدانية لعملية الامتحانات وفحص الأوراق الامتحانية [١] تبين وجود تفاوت واضح في مستوى الامتحانات التحصيلية ما بين شديدة الصعوبة وشديدة السهولة ،وما بين الاعتماد على نوع واحد فقط من الأسئلة إما مقالية أو موضوعية كما أن معظمها غير مطابق للمواصفات ولا يقيس إلا المستويات الدنيا فلا تقيس الابتكار وسلامة الحكم والنقد بل إنها تعتمد على النصوص الحرفية للكتب الدراسية وقياس الحفظ والاستظهار؛هذا بالاضافة إلى أن التقارير الفنية التي يتم استخراجها للاختبارات التي يتم تصحيحها آليا أفادت وجود قصور في جودة المفردات الاختبارية من حيث معاملات سهولتها وصعوبتها وتمييزها وانخفاض معاملات ثباتها فبعض المفردات بلغ معامل سهولتها ١٠٠% على العكس بلغ معامل صعوبة البعض الآخر ١٠٠% مما يؤدي في كثير من الأحيان لتفاوتات غريبة في نتائج الطلاب ونسب نجاح غير مقبولة، ويترتب على ذلك اختلافات في ردود فعل الطلاب بعد الامتحانات وتفاوت في أرائهم ومستوى رضاهم عنها.

وانطلاقًا من منطق أن التعليم المتميز يجب أن تتطابق مواصفات خريجيه مع المواصفات القياسية للإنتاج المتميز وتتواءم مع حاجات المستفيدين منه فلتحقيق الجودة الشاملة في التعليم لابد من التحقق من مدى التطابق بين مواصفات الخريج والمواصفات القياسية المنشودة من خلال عملية القياس والتقويم التي ينبغي أن تخضع هي الأخرى لعمليات التقويم للتحقق من جودتها وكفاءتها وصدق نتائجها في تقرير مدى التطابق ؛فالتقويم بكافة أدواته كأحد مكونات أي منهج أو برنامج تعليمي يجب أن يخضع للتقويم والتطوير المستمرين.

من هنا تتضح الحاجة إلى تقويم جودة الامتحانات التحصيلية باعتبارها الأداة الرئيسة لتقويم الطالب في كثير من الكليات لذلك جاءت الدراسة الحالية لتتعرف على واقع الامتحانات التحصيلية لطلاب كلية التربية جامعة دمنهور من وجهة نظر الطلاب ومدى رضاهم عنها فهم أهم مدخلات المنظومة التعليمية وأهم مخرجاتها في نفس الوقت ، بالاضافة إلى التعرف على واقع تلك الامتحانات في ضوء المعايير والمواصفات القياسية للورقة الامتحانية(الشكلية والفنية) ومدى صلاحيتها السيكومترية ،ودراسة مدى تحقق نواتج التعلم المستهدفة بعد تطبيق وإدارة تلك الامتحانات.

[١] **هذه الملاحظات بُنيت من خلال خبرة الباحثة وعملها كمدير تنفيذي وحدة القياس والتقويم بكلية التربية جامعة دمنهور لمدة ٣ أعوام بدءًا من تاريخ ١-٧-٢٠١٧والتي من مهامها الاشراف على لجان الاختبارات وتصحيحها آليًا واعداد التقارير الفنية الخاصة بها وجمع نماذج من الاختبارات النهائية لفحصها وارسالها إلى لجنة فحص أسئلة الامتحانات لاعداد التقارير حولها.**

مشكلة الدراسة:

التقويم مدخل لاصلاح التعليم فهو أهم مداخل تطوير التعليم وهو العنصر الأهم في تخطيط وإدارة البرامج الأكاديمية في الجامعات وهو الوسيلة الفعالة في ضبط نوعية المخرجات وضمان جودتها ،ولكن وعلى الرغم من تعدد أدوات وأساليب التقويم إلا أن الامتحانات التحصيلية تعد من أهم وسائل التقويم وأكثرها انتشارًا (فؤاد أبو حطب،وآخرون،٢٠٠٨).

وبرغم ذلك فهناك الكثير من الانتقادات توجه إلى الامتحانات التحصيلية التي تحولت من وسيلة تستهدف قياس تحصيل الطلاب وتحديد جوانب القوة وجوانب الضعف لديهم إلى غاية وهدف أساسي لدى جميع المهتمين بالعملية التعليمية مما يُفقد المنتج التعليمي قيمته وقدرته على المنافسة في المجتمع المعاصر بكل ما يحمل من تحديات.(أمطانيوس مخائيل،٢٠١٦؛صلاح علام ،٢٠٠٩؛ Entwistle & Entwistle, 1992; Frederickson & Collins, 1989; Ramsden, 1988; Marnat,2003;Wilson,et al ,2016).

وقد أشارت بعض الدراسات إلى ضعف مهارات أعضاء هيئة التدريس في الجامعات في صياغة الأسئلة وتصميم الأوراق الامتحانية ووجود بعض الأخطاء العلمية الشائعة في صياغة الأسئلة في الاختبارات النهائية . (أماني الغامدي&سماح زغاري،٢٠١٨؛ حمد الكلثم ،٢٠١٥؛ عبد الكريم السوداني،٢٠٠٦؛ محمد عبد المنعم ،٢٠٠٩؛ إبراهيم الحسينان&محمود أحمد ،٢٠١٤ ؛منار الشامي ،سهام الشافعي،٢٠١١؛منال حسن ،٢٠١٣؛يحيى الشريفي،٢٠١٣)

وانطلاقًا من مبدأ ضرورة اخضاع كافة حلقات العملية التعليمية للتقويم لتحقيق الجودة الشاملة في العملية التعليمية بما فيها عمليات التقويم نفسها بما تضمنه من أدوات وأساليب للتحقق من مدى صدقها وشمولها في قياس نواتج التعلم المستهدفة ؛ فهناك حاجة ملحّة لتسليط الضوء على الامتحانات التحصيلية كأهم أدوات التقييم وأكثرها شيوعًا والاهتمام بتطوير هذه الامتحانات في ضوء الشروط الفنية والمواصفات القياسية الواجب توافرها عند بنائها.

وفي ضوء ذلك كانت الحاجة لاعداد دراسة للتعرف على واقع الامتحانات التحصيلية النهائية بكلية التربية جامعة دمنهور من وجهة نظر الطلاب ،والتعرف على واقع هذه الامتحانات في ضوء المعايير والمواصفات القياسية للورقة الامتحانية من الناحية الشكلية والفنية بما تتضمنه من جودة مفرداتها وقدرتها على قياس نواتج التعلم المستهدفة ومدى تحقيقها لها.

١. ما مستوى رضا طلاب كلية التربية جامعة دمنهور عن عمليات التقييم والامتحانات التي يتعرضون لها أثناء دراستهم بالكلية؟
٢. ما نسبة تحقق المؤشرات الشكلية والكيفية لعينة من الأوراق الامتحانية النهائية بكلية التربية ؟
٣. ما مؤشرات الصلاحية السيكومترية للامتحانات النهائية المصحَّحَة آليًّا ؟
٤. إلى أي مدى تتحقق (نواتج التعلم المستهدفة من المقررات) من خلال نتائج الامتحانات النهائية المصحَّحَة آليًّا بالكلية ؟

أهداف الدراسة:

١. التعرف على واقع الامتحانات التحصيلية من وجهة نظر الطلاب ومدى رضاهم عنها.

٢. الكشف عن مدى توافر المواصفات القياسية للاختبار في عينة من الامتحانات النهائية لكلية التربية من الناحية الشكلية والفنية.
٣. الكشف عن الصلاحية السيكومترية للامتحانات النهائية المصححة آليًا لكلية التربية ومدى كفاءتها في قياس نواتج التعلم المستهدفة من المقررات المختلفة.

أهمية الدراسة:

١. يمكن الاستفادة من نتائجها في بناء خطط التحسين وتطوير نظم التقويم والامتحانات لتحقيق النواتج التعليمية المستهدفة بما يتوافق مع المعايير الأكاديمية والارتقاء بمستوى الطلاب كأحدمتطلبات الجودة والاعتماد.
٢. تُشعر الطلاب بأهمية أرائهم في الامتحانات وقيمتهم في المنظومة التعليمية.
٣. تفيد أعضاء هيئة التدريس في تحسين جودة أسئلتهم واختباراتهم وهي مهارات ينبغي اتقانها.
٤. تقديم تقرير للجهات المستفيدة بأهم نقاط القوة والضعف في الامتحانات النهائية بكلية التربية جامعة دمنهور للاستفادة منها في تحديد ورش العمل و الدروات التدريبية اللازمة لتحسين مهارات أعضاء هيئة التدريس وادراجها في الخطط التصحيحية للكلية والجامعة.
٥. تخدم معايير الجودة والتطوير وتفيد في التقويم الذاتي والدراسة الذاتية للكلية.
٦. تُفيد نتائجها في نشر ثقافة التقويم الجيد لدى الطلاب وأعضاء هيئة التدريس من خلال اعداد ورش العمل والمطبوعات المختصة بذلك.
٧. تُفيد نتائجها عند بناء البرامج التدريبية لتحسين مهارات أعضاء هيئة التدريس فيما يتعلق بعمليات تقويم الطلاب وذلك ضمن منظومة تنمية قدرات اعضاء هيئة التدريس ومعاونيهم بحيث يتدرب المستفيدين من نتائج واقعية يتعاملون معها.
٨. تُفيد نتائجها متخذي القرار في التحقق من عدالة الاختيار بين الطلاب عند تعيين المعيدين من دفعات مختلفة بالرجوع إلى الأدلة السيكومترية التي تفيد صلاحية وجودة الامتحانات واستقرار وثبات نتائجها.
٩. تُفيد المجتمع فعندما تعكس درجات الطلاب مستواهم الحقيقي يتخرج أفرادًا قادرين على خدمة المجتمع وحل مشكلاته.

مصطحات الدراسة:

نواتج التعلم المستهدفة The Intended Learning Outcomes(ILOs): هي مجموعة من المعارف والمهارات التي تضعها مؤسسة تعليمية لبرنامج/مقرر والتي تصف الأداء المتوقع من المتعلم كنتيجة لانتهائه من دراسة الأنشطة التعليمية . وحددتها الهيئة القومية لضمان جودة التعليم والاعتماد لأي برنامج /مقرر دراسي في أربعة مكونات رئيسية هي

١. **المعرفة والفهم:** وهي المعلومات الأساسية والمفاهيم التي يتوقع أن يكتسبها الطالب عند دراسته لمقرر معين والتي تتضمن المفاهيم الأساسية والمصطلحات والنظريات والحقائق.
٢. **المهارات الذهنية:** وهي القدرات التعليمية والادراكية التي تتضمن تطبيق وتحليل واعادة بناء وتقييم المعلومات كما تتضمن التفكير النقدي والابداعي .

٣. **المهارات المهنية :**وتتضمن تطبيق المعلومات والتدريبات العملية المتخصصة في مجال ما بهدف التطوير الناجح في المهنة أو التطوير الذاتي.
٤. **المهارات العامة أو الانتقالية :**تلك المهارات التي لا ترتبط بموضوع معين وغالبًا ما يحتاجها الطالب في التعلم المستمر وتنمية القدرات الذاتية وتتضمن القدرة على الاتصال والعمل في فريق والتعامل **مع** الأرقام والتعلم الذاتي والتواصل مع الآخرين والقدرة على حل المشكلات. (دليل الهيئة القومية لضمان جودة التعليم والاعتماد بجمهورية مصرالعربية ،٢٠١٣)

الأوراق الامتحانية : ويحددها الباحثان في الدراسة الحالية بأنها " الامتحانات التحصيلية النهائية التي يمكن من خلالها الكشف عن أثر ما تعلمه الطالب أو ما تدرب عليه من معرفة وفهم أومهارة مكتسبة أثناء دراسته لمقرر بأكمله".

الإطار النظري والدراسات السابقة:

١- نواتج التعلم المستهدفة :The Intended Learning Outcomes(ILOs)

تُعد نواتج التعلم واحدة من اللّبِنات الأساسية الضرورية لشفافية نظم التعليم بصفة عامة والتعليم العالي بصفة خاصة (Adam,2004)،فهي تصف ما هو متوقع من المتعلم أن يعرفه أو يستطيع القيام به بعد الانتهاء من المقرر.

ويتم التحقق من نواتج التعلم من خلال قياس قدرة الطلاب على أداء أنشطة التعلم المستهدف تحقيقها مع انتهاء وحدات المقرر.(Dodridge,1999; Harden et al,1999; Kennedy,et al,2007)

وقد تعددت تعريفات نواتج التعلم حيث تعرف بأنها "وصف صريح لما يجب على المتعلم معرفته وفهمه والقيام به نتيجة للتعلم " (Bingham,1999) .أوهي "عبارات تقرر ما يتوقع أن يقوم به المتعلم نتيجة لنشاط التعلم " (Jenkins& Unwin,2007). فهي "مجموعة من الجمل تحدد ما يرغب المتعلم تحقيقه بعد الانتهاء بنجاح من دراسته لمحتوى المقرر، وتعرفها (اقبال درندري،٢٠١٠) بأنها أداء الطلاب الملاحظ الذي يوضح أن الأهداف المقصودة للتعلم قد تحققت.

وترى كل من (أماني الغامدي،سماح زغاري ،٢٠١٨) أن معظم تعريفات نواتج التعلم تتمحور حول نواتج التعلم المستهدفة قبل التعلم أو نواتج التعلم المتحققة بعد التعلم أي أن الفرق يكمن في البعد الزمني أكثر من أي بعد آخر .

غالبا ما يتم وضع نواتج التعلم المستهدفة قبل تقديم مهام التعلم للمتعلمين (Anderson,et al,2005) حيث أن ذلك يساعد على مطابقة نواتج التعلم المستهدفة للمعلم والمتعلم فيتشاركان المحتوى التربوي وفقًا لأهدافيهما ومعارفيهما وكلًا من أنشطة التعليم و التعلم.

فتحدث الرؤية التطابقية عادة خلال فترة الدراسة، وتتكون النّية المشتركة بينهما التي تتمثل في اكتساب وفهم محتوى المادة التعليمية من خلال الأنشطة التربوية المناسبة ومن ثم فتلك الأهداف المشتركة تعتبر بمثابة المؤشر الذي يؤدي إلى تحسين القدرات المتعلمة.

من الجدير بالذكر أنه ينبغي أن تضع مؤسسات التعليم العالي نواتج التعلم المستهدفة بشكل هرمي تتمثل قاعدته في نواتج التعلم المستهدفة للكلية بشكل عام التي تنعكس في نواتج التعلم المستهدفة لبرامج

الكلية والتي تنعكس بدورها في مقررات البرامج ؛ وتتوزع نواتج التعلم لكل مقرر على الوحدات والدروس (Ahmedani,2017).

وأوصت الهيئة القومية لضمان جودة التعليم والاعتماد بأن يتم وضع نواتج التعلم المستهدفة الخاصة بالمقررات الدراسية في ضوء المعايير الأكاديمية المرجعية القومية National Academic Referenced Standers (NARS) بحيث تتحدد نواتج التعلم المستهدفة لأي برنامج أو مقرر دراسي في أربعة مكونات رئيسية هي(المعرفة والفهم) وهي المعلومات الأساسية والمفاهيم التي يتوقع أن يكتسبها الطالب عند دراسته لمقرر معين، و(المهارات الذهنية) وهي القدرات الذهنية التي تتضمن تطبيق وتحليل واعادة بناء وتقييم المعلومات كما تتضمن التفكير النقدي والابداعي،و(المهارات المهنية) وتتضمن تطبيق المعلومات والتدريبات العملية المتخصصة في مجال مابهدف التطوير الناجح في المهنة أو التطوير الذاتي،وأخيرًا (المهارات العامة أو الانتقالية) وهي تلك المهارات التي لا ترتبط بموضوع معين وغالبًا ما يحتاجها الطالب في التعلم المستمر وتنمية القدرات الذاتية وتتضمن القدرة على الاتصال والعمل في فريق والتعامل مع الأرقام والتعلم الذاتي والتواصل مع الآخرين والقدرة على حل المشكلات. (دليل الهيئة القومية لضمان جودة التعليم والاعتماد بجمهورية مصر العربية ،٢٠١٣)

تصاغ نواتج التعلم في صورة تقريرية نصية واضحة المعنى (Gilbert&Gale,2008; Kennedy,et al ,2007)، بحيث تكون مرتكزة حول المتعلم ،وقابلة للملاحظة والقياس فيما يتعلق بكل من المعارف والمهارات والسمات المحققة من قبل المتعلم .(Pilli,& Admiraal,2017) ولذلك ينبغي أن يتم استخدام أفعال Action Verbs محددة وواضحة تعبر عن مستوى الأداء أو السلوك المتوقع من الطالب ويراعى فيها تجنب الإطناب والتكرار.(Kennedy,et al,2007).وفي الغالب يأخذ مخرج التعلم النمط التالي :"*النهاية المقرر ،سوف يكون المتعلم قادرًا على .....س ،ص* "حيث (س) تعبر عن القدرة ،و(ص) هي محتوى الموضوع (Gilbert&Gale,2008) .

وقد أشار (Tangworakitthaworn,et al,2013) إلى أنّ نواتج التعلم المستهدفة تنقسم إلى مكونين رئيسيين **هما:القدرة ، ومحتوى الموضوع كالتالي:**

١. **يُعبّر مكون القدرة** عن قدرة المتعلم على أداء أنشطة التعلم ويُعبّر عنه بفعل يُعيّن القدرة المتعلمة المكتسبة learned capability verb (LCV) وهو كلمة اجرائية تعتبر العنصر اللُّغوي لنواتج التعلم المستهدفة.

وضع بنيامين بلوم منذ عام (Bloom,,et al,1956) تصنيفًا لغايات وأهداف التعلم التربوية في بنية هرمية تمثل أشكالًا ومستويات مختلفة للتعلم في مستويات متدرجة ومعقدة بحيث يُعد كل مستوى شرطًا لتحقيق المستوى الأعلى؛ ويرى (محمود إبراهيم ،محمد السيد ،٢٠١٥) أن تحديد مستويات التعلم المتدرجة في المجال المعرفي والمهاري والوجداني كان أساسًا بُنيت عليه فكرة تصنيف نواتج تعلم الطلاب فيما بعد

وبذلك يمثل تصنيف بلوم للمجال المعرفي (Bloom,,et al,1956) عنصر القدرة في نواتج التعلم المستهدفة ، فمُخرج مثل: "في نهاية المقرر سيكون الطالب قادرًا على تصميم مخطط للعلاقات للكيان موضع الدراسة " يكون فعل القدرة المكتسبة في هذا المثال هو"يُصمم".

وقد بدأت مستويات بلوم للمجال المعرفي بالمعرفة ثم الفهم ثم التطبيق ثم التحليل ثم التركيب وأخيرًا التقويم ؛ولكن طورت أندرسون (Anderson) مستويات التعلم المعرفية لتتلاءم مع متطلبات القرن الحادي والعشرون بحيث تم تعديل المستوى قبل الأخير ليُعبّر عن "التقويم" والمستوى الأخير ليُعبّر عن الابداع[٢] .وقد ركزت أندرسون على الأفعال بدلًا من الأسماء بمعنى ما يمتلكه الطالب من معارف أوما يستطيع أدائه من مهارات في نهاية المقرر ،بحيث يكون النموذج أكثر ديناميكية لتصنيف العمليات الفكرية التي يستخدمها الطلاب في اكتساب وتوظيف المعرفة (Anderson & Krathwohl,2001).

٢. أما فيما يتعلق **بمحتوى موضوع المادة** Subject Matter Content (SMC) فيتكون من بنية العلم المتمثلة في هرم قاعدته الحقائق وقمته النظريات .
ويرى (Tangworakitthaworn,et al,2013) أنه بناءً على نظرية عرض المكونات التي اقترحها (Merrill,,1973; Merrill,1994) ، هناك أربع فئات تمثل المحتوى ، وهي "الحقيقة والمفهوم والإجراء والمبدأ"؛ حيث تتكون الحقيقة من جزئين مرتبطين من المعلومات مثلًا "اسم وتاريخ محددين" أو "حدث واسم محدد " فسؤال مثل :من هو ألفريد بينيه؟" تُعبّر اجابتة عن حقيقة هي أن "ألفريد بينيه عالم نفس فرنسي ومُعد أول اختبار ذكاء"؛أما المفهوم فهو عنصر ملموس أو مجرد ذو خصائص معينة فسؤال مثل"حدد خصائص الاختبار التحصيلي الجيد" تُعبّر اجابته عن عدة مفاهيم تتضمن الموضوعية والشمول والصدق والثبات ، بينما يعتبر الإجراء (أو العملية) مجموعة من الخطوات لإنجاز بعض الأهداف كاجابة سؤال مثل :"حدد الخطوات المطلوبة لإعداد الاختبار التحصيلي"؛ أما المبدأ فهو علاقة السبب والنتيجة وابداء الأسباب كإجابة سؤال مثل " من خلال الأرقام المتاحة في جدول مواصفات الاختبار ، حدد عدد الأسئلة في كل مستوى مع ذكر السبب.

وقام (Tangworakitthaworn,et al,2013) بتحديد مصفوفة الأبعاد الثنائية لمحتوى الأداء **The Two-Dimensional Performance Content Matrix (2D-PCM)** باستخدام مخطط تصنيف (Merrill,1994) لتمثيل العلاقة بين أداء المتعلم بمستوياته المختلفة من ناحية والمواد التعليمية (المحددة في ضوء محتوى موضوع المادة بمستوياتها المتمثلة في الحقيقة ، المفهوم ، الإجراء ،والمبدأ) من ناحية أخرى ونتج عن ذلك تحديد العلاقات في مصفوفة الأبعاد الثنائية لمحتوى الأداء (2D-PCM) التي تُعبّر عن تشغيل كل مستوى من مستويات الأداء على مستويات محتوى الموضوع المختلفة ؛وقد تمكن من تمثيل و نمذجة كافة نواتج التعلم المستهدفة لمقرر أو درس أو برنامج محدد من خلال المخططات من خلال تحديد العنصرين الرئيسين من نواتج التعلم المستهدفة_**الأول** هو مكون القدرة الذي يشير إلى نشاط المتعلمين في أداء مهام التعلم المتمثل في الفئات الستة من تصنيف بلوم المعرفي التي تشكل التسلسل الهرمي المعرفي باستخدام الفعل الاجرائي للقدرة المتعلمة. **والثاني** هو مكون محتوى مادة التعلم التي يتم تمثيلها باستخدام مصفوفة الأبعاد الثنائية لمحتوى الأداء 2D-PCM .

٢ تتبنى هيئة ضمان جودة التعليم والاعتماد حاليا نموذج هرم بلوم الجديد الذي يتضمن المراحل الست التالية: (التذكر– الفهم– التطبيق – التحليل– التقويم– الإبداع) فى وثائق التدريب الحديثة (دورة نظم الامتحانات وتقويم الطلاب لكليات ومعاهد التعليم العالي)،وكذلك مراكز القياس والتقويم ومراكز ضمان الجودة بالجامعات المصرية

ويذكر (Tangworakitthaworn,et al,2013) أنّ هذا الهيكل المنطقي لنواتج التعلم المستهدفة "باستخدام المخططات البيانية التي تتضمن معلومات حول السياقات التعليمية وقدرات المتعلمين" يفيد كل من واضعي المناهج ومصممي التعليم في تصميم وتطوير المناهج الدراسية

وتجدر الاشارة إلى أنه من الضروري وضع سيناريو لتطبيق مخطط نواتج التعلم المستهدفة في التعليم مع توضيح كافة علاقات وقيود نواتج التعلم المستهدفة من أجل دعم تطوير المناهج وتسهيل تحقق نواتج التعلم المستهدفة؛ فوضوح نواتج التعلم المستهدفة ضروري لكل من المعلم والمتعلم (Biggs & Tang, 2011; Kenny&Desmarais,2012) حيث تساعد المعلم في تنظيم أعماله والتركيز على الأولويات المهمة بما يتناسب مع احتياجات الطلاب وتحديد الاستراتيجيات المناسبة وأساليب التقويم الملائمة للتحقق من مدى اكتساب الطالب لنواتج التعلم فضلًا على أنها تساعده على التنمية المهنية المستدامة في ضوء نتائج تقويم نواتج التعلم لدى طلابه ،كما تساعد المتعلم على تحقيق تعلم أفضل فيمكنه أن يتخير الأنشطة والمهام التي تساعد على تحقيق نواتج التعلم وفقًا لاستعداداته وميوله كما أنّ القواعد الواضحة تساعده على تحديد مستواه وتطوير أداؤه أولاً بأول ومن هنا تأتي ضرورة التقويم في ضوء نواتج التعلم.

٢. أساليب تقييم نواتج التعلم المستهدفة وأهمية التقييم فى ضوئها.

أشار كثير من الباحثون كما يذكر(Ewaiset al, 2020) أن هناك ثلاثة مكونات رئيسة للمقرر الدراسي هي: **المعلم** ثم **نموذج التعلم** بما يتضمنه من مواد التعلم ، نواتج التعلم ،.. إلخ ،ثم **التقويم** .ويعتبر التقويم عنصرًا حيويًا وحرجًا في المقرر الدراسي فمن خلاله يتم قياس عملية تعلم المتعلمين والحكم على مدى تحقق النواتج المستهدفة من التعلم وبالتالي توجيه عملية التعلم والتحكم فيها (Agnantis,2016).

وتتعدد أنواع التقييمات حسب توقيتها والغرض منها حيث أن التقييم القبلي يكون قبل التعلم والتقييم التكويني يتم من خلال الأسئلة والاختبارات التي تطبق أثناء عملية التعلم وبطريقة مستمرة ، بينما التقييم الختامي يتعلق بقياس تقدم التعلم في نهاية تعلم وحدة أو موضوع.

وقد أشار (Race & Brown, 2001) إلى أنه كلما زاد التنوع في طرق التقييم كلما كان التقييم أكثر عدالة بالنسبة للطلاب فالاستخدام الزائد لأحد الإستراتيجيات التقييمية يمكن أن يميز طلاب بعينهم مما يؤدي إلى تقييم محدود غير حقيقي لأداء الطلاب .

ومن الأساليب التى يستخدمها المعلمون في تقييم أداء الطلاب الاختبارات والمقابلات والملاحظات والتكليفات المنزلية إلا أن أكثرها استخدامًا هي الاختبارات التحصيلية؛ حيث تعتبر الاختبارات التحصيلية من أكثر أدوات التقويم شيوعًا في المؤسسات التعليمية لما لها من أهمية في عملية التقويم حيث تعطي فكرة واضحة عن قدرة وامكانيات الطلاب ومدى تقدمهم وتحديد مواطن القوة والضعف لديهم وإثارة دافعيتهم للتعلم .

ويعرف (محمود عمر ،وآخرون ،٢٠١٠) الاختبارات التحصيلية بأنها محاولة للكشف عن أثر ما تعلمه الطالب أو ما تدرب عليه أثناء دراسته لموضوع معين أو وحدة تعليمية معينة "
وغالبًا ما تُستخدم كلمة الاختبار كمرادف للامتحان التحصيلي باعتبارهما أدوات لقياس المعلومات والمهارات والخبرات السابق دراستها من خلال مادة دراسية محددة ،ويذكر (صلاح مراد،أمين سليمان

،٢٠٠٥)ان الفرق بين الاختبار والامتحان التحصيلي هو الفرق بين الاختبارات المقننة Standardized tests وغير المقننة Non standardized tests على الترتيب من حيث أن:

- الاختبارات المقننة (الاختبارات)Standardized tests :تقيس أهداف مشتركة عامة لمجموعة من المدارس وتتناول بالقياس أجزاء كبيرة من المعلومات والمهارات السابق دراستها وأسئلتها تم تجربتها ومراجعتها ومعالجتها احصائيًا بغرض حساب معاملات السهولة والصعوبة وفعالية التمييز وفعالية المشتتات ،والتخلص من أثر التخمين وتزود بمعايير الأداء لمجموعات مختلفة من الأفراد على نطاق واسع.
- الاختبارات غير المقننة (الامتحانات) Examination حيث تقيس أهدافًا خاصة بالفصل أو المدرسة وتقيس أجزاء محددة من المحتوى بمعنى أنها تتناول معلومات ومهارات معينة ومحددة و نادرًا ما يتم تجربة أسئلتها أومراجعتها أو تحليلها احصائيًا ،وقد تزود بمعايير الأداء على نطاق الفصل أو المدرسة فقط (عينة محدودة).

وتتنوع الاختبارات في طريقتها ما بين :شفوية حيث يمكن من خلال الأسئلة التي توجه للطالب استكشاف كثير من الجوانب الخاصة بتحصيله بالاضافة لجوانب أخرى منها طريقة وسرعة رد فعله ،**وكتابية** والتي تحذى بكثير من الاهتمام ويعوّل عليها في تشخيص قدرات الطالب **وأدائية** والتي تستهدف الكشف عن قدرته على القيام بأنشطة ومهارات في مهام واقعية من خلال بطاقات التقدير.(صلاح علام ،٢٠٠٠)

وتعتبر الاختبارات الكتابية من أكثر الاختبارات شيوعًا واستخدامًا وأهمية خصوصًا في الكليات النظرية وتتنوع مابين الاختبارات المقالية التي يترك للطالب حرية التعبير فيها وفق معايير قد تكون مفتوحة أو محددة باجابات قصيرة ومحددة، مقابل الاختبارات الموضوعية محددة الإجابة والتي تتضمن أسئلة الصواب والخطأ والاختيار من متعدد والمزاوجة.

ولكل نوع من الاختبارات سواء المقالية أو الموضوعية جوانب قوة وجوانب قصور فيشير (Boulton-Lewis, 1995) إلى أنه من خلال الامتحانات المقالية يمكن قياس مهارات التفكير العليا كالتفكير التحليلي والتفكير الناقد، كما تتطلب مهارات تواصل مما يسمح بالمعالجة العميقة من قبل الطلاب، و تسمح للطلاب بالتحكم في اختياراتهم، وتنظيم وعرض معارفهم وفهمهم وتتحدى قدراتهم لتقديم المعارف الشرطية وما وراء النظرية إذا كانوا يمتلكونها،لذلك أكّد (Oermann, 1999) إلى أنها أكثر صدقًا في قياس نواتج التعلم، كما أنها الأسهل في البناء بالنسبة للوقت والجهد، في الوقت نفسه أشار (Holsgrove, 1992) إلى أنها محدودة في ثباتها وموضوعيتها حيث تتطلب وضع عدد محدود من الأسئلة يجاب عنها في وقت محدد قد لا تغطي المحتوى المراد قياسه، وذلك يؤدي إلى تحيز وعدم عدالة في وضع الأسئلة، كما أنه في حالة عدم وجود معايير للأداء المتدرجة Rubrics أثناء تصحيح مفرداتها تكون النتائج غير ثابتة وغير موضوعية؛وبذلك فلها حدود في التطبيق خصوصًا مع الأعداد الكبيرة للطلاب (Gibbs,et al 1997) .

على الجانب الآخر تمتاز الاختبارات الموضوعية بالثبات والصدق وموضوعية التصحيح (Pepple,et al,2010) كما أنها تغطي أجزاء مختلفة من المحتوى المراد قياسه ومناسبة مع الأعداد الكبيرة

للطلاب(Hammond et al, 1998) ، فضلاً عن أنها تساعد الطلاب في مراقبة وتأكيد تعلمهم إذا صُحبت بالتغذية المرتدة التي تقدّم في وقتها المناسب وبذلك يمكنها تحفيز الطلاب إيجابيًا .

في الوقت نفسه ذكر (Masters,et.al, 2001) أن الأسئلة الموضوعية تقيس فقط قدرة الطلاب على التعرُّف والتعلم السطحي واستدعاء الحقائق والتي تعتبر غير مناسبة لكثير من المقررات التي تستخدم حل المشكلات ، كما أنها صعبة في إعدادها وصياغتها فقد قدّر (Farley, 1989) كتابة سؤال اختيار من متعدد بصورة جيدة يمكن أن يستغرق ساعة من الزمن.

ومن أهم مشكلات الاختبارات الموضوعية أيضًا استخدام الأسئلة ضعيفة البناء التي قد تتضمن مصطلحات غير دقيقة أو بدائل غير مناسبة، وبذلك تكون النتائج متحيزة وتكون التقييمات الناتجة عنها مشوهة؛ ومن ذلك يبدو أن الأسئلة الموضوعية وبخاصة أسئلة الاختيار من متعدد كغيرها من إستراتيجيات التقييم لها قيود وحدود وتطبق على بعض المحتويات أكثر من غيرها بشرط أن تفي بمعايير التقييم الجيد ؛وقد حددت (Quinn, 2000) عدة معايير أساسية لتقييم فعالية أسئلة الاختيار من متعدد كإستراتيجية تقييم هي :"القابلية للتطبيق والثبات والصدق والتمييز"،فإذا تحقّقت تلك المعايير يمكن لها أن تكون جزءًا لا يتجزأ من عملية التعليم والتعلم بحيث تقيّم الأداء في علاقته بنواتج التعلم المستهدفة ILOs. .In(Brady, 2005)

وتجدر الإشارة إلى أنه أيًا كانت الأساليب المستخدمة في التقويم فينبغي أن تتم في ضوء نواتج التعلم المستهدفة ويُعد ذلك تطبيقًا لمدخل تصميم المنهج المتمركز حول المتعلم ويعزز فلسفة المتابعة والتقييم والتحسُّن المستمر ويرفع من مستوى الدافعية للتعلم ويخلق الفرصة لتطبيق التخطيط الجماعي للمنهج أو البرنامج التعليمي وللتعاون بين أعضاء هيئة التدريس لتحديد الثغرات واقتراح طرق العلاج واختيار استراتيجيات التعليم وطرق التقويم وأدواته المناسبة مما يوفر أساسًا لتطوير البرامج الأكاديمية ويساعد على توفير الوضوح والتكامل بين المقررات التي يتضمنها (Biggs,1973).

وأكد (Biggs & Tang, 2011) على ضرورة التناظر الدقيق بين ما يتم تعليمه وما يتم قياسه،وما يتم تعلّمه واقترح مفهوم المحاذاة البنائية constructive alignment الذي يؤكد على ضرورة أن تصمم التقييمات المختلفة بعناية لتكشف عما إذا كان الطلاب قد حققوا نواتج التعلم المستهدفة أم لا؛فبذلك لا يعتبر التقييم فقط مقياسًا مفيدًا لتعلم الطلاب، ولكن أيضًا يعتبر بناءً طبيعيًا للتغذية المرتدة الفعالة والمثمرة ؛فيزيد من وضوح ما هو متوقع من الطلاب من ناحية ومن شفافية وضع الدرجات والأحكام من ناحية أخرى.

وبذلك فارتباط التقويم بنواتج التعلم يمثل اتجاهًا من الاتجاهات الحديثة لضمان جودة مؤسسات التعليم العالي حيث يحدد بدقة نواتج التعلم المستهدفة وطرائق اكتسابها وطرائق تقييمها بحيث تستوعب جميع المعارف والمهارات المطلوبة للمتعلم وتجعل التقويم عنصرًا فاعلًا لتطوير العملية التعليمية. (هدى صالح،حنان محمد،٢٠١٧)

٣. تقييم الامتحانات التحصيلية في ضوء نواتج التعلم المستهدفة:

تتحدد فاعلية إستراتيجيات التقييم في ضوء قدرتها على قياس القدرات والمهارات العقلية الموضوعة لقياسها واتساقها مع نواتج التعلم المستهدفة ILOs.

وقد أشار (Biggs, 1973) إلى أنه من المزعج أن يحصل الطلاب الذين يتعلمون في شكل صم Rout Learning عديم المعنى على درجات عالية في الامتحانات وهو السائد في عديد من النظم التعليمية ،وقد يرجع ذلك في ضوء ما ذكره كل من (Frederickson & Collins, 1989;Ramsden, 1988) فيما يتعلق بتحليلات الاختبارات التحصيلية القياسية أن كثير من المفردات لا تتطلب أكثر من استدعاء الحقائق والتفاصيل الموجودة في المحاضرات والكتب ولاتغطي كافة نواتج التعلم المستهدفة من المقرر، وأيد ذلك تحليلات (Entwistle & Entwistle, 1992) للإختبارات النهائية وأشار إلى *أن هناك علاقة وثيقة بين نوعية الأسئلة المستخدمة من قبل المدرسين ونوع التفكير السائد لدي طلابهم ،فالمدرسين الذين يستخدمون أسئلة تثير التفكير يصبح طلابهم قادرين علي ممارسة العمليات العقلية العليا.*

كما أن عدم مراعاة معايير الاختبار الجيد "كالصدق والثبات والموضوعية والشمولية لكافة نواتج التعلم المستهدفة "خصوصًا في مؤسسات التعليم العالي يؤدي إلى ضعف موثوقية الدرجات التي يحصل عليها الطلاب في الاختبارات التحصيلية ويقلل من أهميتها في تزويد القسم العلمي بمعلومات مفيدة عن مستوى تحقق نواتج تعلم البرنامج الأكاديمي في المقرر.

وبذلك فعدم التطابق والتوافق بين نواتج التعلم المستهدفة واستخدام طرق التقييم المناسبة لتقييم تحققها قد يؤدي لعدم العدالة في تقويم الطلاب فيُهمَل ذوي القدرات والمبدعين وترتفع الدرجات مقارنةً بالمهارات، مما يعني تفريغ العملية التعليمية من محتواها واعتبار الامتحانات وعملية تقويم الطلاب مجرد تحصيل حاصل وقد يكون هذا هو السبب في وجود شعورًا عامًا بأن التعليم يعاني من مشكلات عديدة وأوجه قصور شتى أهمها تدني مستوى نواتج وأكثرها وضوحًا ضعف مستوى الخريجين (فؤاد أبو حطب،وآخرون،٢٠٠٨)

وأكدت العديد من الدراسات كما يذكر (Wilson,et al ,2016) وجود فجوة وخلل جوهري بين التعلم المستهدف وأشكال التقييم الشائعة فكانت ممارسات التقييم لا تتماشى مع المعلومات الأساسية التي يكتسبها الطلاب الجامعيين في المقررات الدراسية ، وقد كشفت دراستهم التحليلية بعنوان "المحاذاة غير الصحيحة الأساسية: التعلم المستهدف والممارسات التقييمية في مشاريع البحث العلمي الجامعية" أن المداخل التقليدية للتقييم غير متطابقة بشكل أساسي مع معظم نواتج التعلم وقد اقترح أنه ربما يرجع ذلك إلى تركيز هذه الأساليب على منتجات المشاريع البحثية وليس على العمليات والخبرات اللازمة لعمل الأبحاث والمشروعات.

كما أشار (ضياء الدين عبد الواحد،٢٠١٥) أن الاختبارات التحصيلية النهائية لقسم المكتبات والمعلومات بجامعة عين شمس لا تعطي صورة حقيقية عن قياس الأهداف التعليمية كما تعكسها أهداف المحتوى الدراسي ،حيث تركز على المجال المعرفي في بعض مستوياته خاصة التذكر مع اغفال المجال المهاري والوجداني ،مما يدل على عدم صدق تلك الاختبارات ،وتوصلت دراسة (أماني الغامدي،سماح زغاري،٢٠١٨) إلى نتيجة مماثلة حيث كشفت من خلال مراجعة توصيف المقررات عن وجود فجوة بين الأهداف التعليمية المقصودة وأساليب التقويم المنصوص عليها من جانب وبين نتائج التحليل للاختبارات من جانب آخر، ففي الوقت الذي تنص فيه التوصيفات على تنمية الابداع ومناقشة القضايا واستخدام أساليب إبداعية في التقويم فإن واقع النتائج تكشف أن معظم عمليات قياس نواتج التعلم انحصرت على الأغلب في اختبارات موضوعية تقيس قدرات عقلية محدودة تنحصر في التذكر والاستيعاب.

وقد اهتم العديد من الباحثين بدراسة واقع الاختبارات سواء الدورية منها أو النهائية وذلك بفحص مفرداتها وتحديد مستويات التفكير التي تستهدف قياسها في ضوء مواصفات ومعايير جودة الاختبارات وذلك للعديد من المراحل التعليمية منها:

دراسة (عبد الكريم السوداني،٢٠٠٦) وهدفت الي معرفة مستويات التفكير (تصنيف بلوم للمجال المعرفي) التي تستثيرها أسئلة الاختبارات النهائية التي يتعرض لها طلاب كلية التربية قسم علوم الحياة،وقد اعتمد على أسلوب تحليل المحتوي في معرفة هذه المستويات بعد التحقق من صدق التحليل بحساب معامل الاتفاق بين المحللين حيث بلغ (٠.٨١) وبلغت عدد الأسئلة المحللة (١٦٦) سؤال؛ وقد كشفت الدراسة عن خلو أسئلة الصف الأول والثاني من مستويات التطبيق والتركيب والتقويم وبالنسبة للصف الثالث والرابع كان مستوي التذكر أعلي المستويات تحققًا.

دراسة (محمد عبد المنعم ،٢٠٠٩) التي هدفت إلى الكشف عن واقع أسئلة الإختبارات الفصلية للمواد الاكاديمية والتربوية للفصل الدراسي الثاني للطلاب الخريجين من التعليم الابتدائي في كلية المعلمين جامعة الملك فيصل " التخصصات الأدبية والعلمية " في ضوء معايير جودة التقويم حيث تم تحديد قائمة لمعايير الاختبارات التحصيلية الكتابية ثم تقييم أسئلة تلك الاختبارات في ضوء تلك المعايير، وقد كشفت النتائج عن وجود العديد من الأخطاء اللُّغوية والنَّحوية والطباعية في أسئلة كثير من الاختبارات ،ووجود كثير من الأخطاء العلمية والفنية في جودة الأسئلة بالاضافة إلى تكرار بعض الأسئلة بصورة كاملة أو جزئية من اختبارات سابقة ،كما أن كثير من أسئلة تلك الاختبارات قد ترِد إجاباتها في نفس الاختبار في نصوص أسئلة أخرى؛ وقد استدلّ من هذه النتائج على أن معظم الاختبارات لم تبنى في ضوء مواصفات الاختبار الجيد وفي ضوء جدول المواصفات .

دراسة(منار الشامي ،سهام الشافعي،٢٠١١) هدفت لدراسة مدى جودة الاختبارات التحصيلية بقسم الاقتصاد المنزلي جامعة تبوك وتحديد الأهمية النسبية لأنواع الأسئلة المختلفة في الاختبارات التحصيلية . وكشفت النتائج عن انخفاض جودة الاختبارات التحصيلية حيث لم تصل أي منها لدرجة الجودة المطلوبة والتي تم تحديدها بنسبة ٨٥%وقد تم تقييم جودة الاختبارات في ضوء أربعة معايير رئيسية هي *معيار التحليل العام* (الذي يتضمن شكل الورقة الإختبارية وخلوها من الأخطاء اللغوية وكتابة تعليمات الإجابة والدرجة الكلية للإختبار) *ومعيار مضمون الأسئلة* (الذي يتضمن مدى وضوح عبارات الأسئلة وعلاقتها بأهداف المادة ومدى شموليتها لموضوعات المقرر ومدى مناسبتها للزمن) ثم معيار *الأسئلة المقالية وأخيرًا معيار الأسئلة الموضوعية* وكشفت النتائج أنه : كان لمعيار التحليل العام الأثر الأكبر في درجات جودة الاختبارات يليه معيار مضمون الأسئلة ثم معيار جودة الأسئلة المقالية وأخيرًا معيار جودة الأسئلة الموضوعية ، كما كان هناك تفاوت في نسب أنواع الأسئلة في الاختبارات وركزت بشكل كبير على الأسئلة الموضوعية حيث بلغت نسبتها (٨٣.٢%) في مقابل الأسئلة المقالية التي بلغت نسبتها (١٦.٨%)، كما وجدت علاقة ارتباطية عكسية بين ضغوط العمل التي تواجه عضوات هيئة التدريس وبين جودة الاختبارات اللاتي يقمن باعدادها.

دراسة (منال حسن ،٢٠١٣) هدفت إلى تقويم أسئلة الاختبارات النهائية للأقسام الأدبية والعلمية لجميع المستويات بكلية التربية بحفر الباطن جامعة الدمام في ضوء معايير الجودة وقد كشفت النتائج أنّ: معظم أسئلة اختبارات المواد التخصصية العلمية والأدبية والمواد التربوية تقيس المستويات المعرفية الدنيا فأغلبها تقيس أدنى المستويات المعرفية وهو مستوى التذكر وفي أحسن الأحوال قد تصل إلى مستوى

الفهم ،كما كشفت النتائج أيضًا أن معظم الأسئلة أهملت معيار الشمول ولم تلتزم بالمعايير التي خصصت لتنظيم وإخراج ورقة الاختبار.

دراسة (يحيى الشريفي،٢٠١٣) هدفت إلى دراسة مدى شمول أسئلة الاختبارات النهائية لمادة الأدب في كليات التربية لمستويات بلوم المعرفية ، وقد شملت عينته (٩٢) سؤالاً من أسئلة مادة الأدب موزعة بين المراحل الدراسية الأربعة في أقسام اللغة العربية وقد توصل إلى أن مستوى الفهم حصل على أعلى التكرارات بنسبة (٤٣.٤٨%) بينما حصل مستوى التقويم على أدنى التكرارات بنسبة (١.٠٩%)، أما تنوع الأسئلة فلم يجد الباحث أي أثر للاختبارات الموضوعية في نماذج الأسئلة لمادة الأدب واقتصرت جميعها على الأسئلة المقالية، أما معيار الشمول فكان تمثيل محتوى المادة الدراسية ضعيفاً في نماذج الأسئلة النهائية ولم يصل إلى العتبة التي افترضها الباحث وهي (٨٠%).

دراسة (علي الحديبي ، هناء إبراهيم،٢٠١٤) هدفت إلى وضع قائمة بمعايير الجودة في تصميم أوراق الأسئلة بالجامعة الاسلامية بالمدينة المنورة تضمنت ثلاثة معايير أساسية ينبثق عنها خمسة وعشرون مؤشرًا ،بحيث يتضمن *المعيار الأول* البيانات الأساسية عن الجامعة والمقرر الدراسي *والمعيار الثاني* يركز على الجوانب الفنية التي تتعلق باخراج ورقة الاختبار وتنظيمها وترتيبها ،*والمعيار الثالث* يركز على الصياغة العلمية الصحيحة للأسئلة سواء كانت مقالية أم موضوعية ، وبفحص و تقويم أوراق الاختبارات على ضوء تلك المعايير اتضح أن المعيار الثاني (الاهتمام باخراج وتنظيم الورقة الامتحانية) جاء في الترتيب الأول حيث حصل على نسبة توافر ٨٤.١٤% يليه المعيار الأول (البيانات الأساسية للجامعة وللمقرر) حيث حصل على نسبة توافر ٨٢.١٢% وأخيرًا المعيار الثالث (الصياغة العلمية الصحيحة للأسئلة) وحصل على نسبة توافر ٨٠.٧٧% وهذا يدل على أن أعضاء هيئة التدريس بتلك الجامعة يهتمون باخراج الورقة وتنظيمها ثم بكتابة البيانات الأساسية ويأتي في آخر اهتمامهم صياغة الأسئلة صياغة علمية.

دراسة (إبراهيم الحسينان، محمود أحمد ،٢٠١٤) وهدفت إلى التحقق من استيفاء معايير جودة الاختبارات الموزعة على خمسة أبعاد هي(البيانات الأساسية للاختبار،تعليمات الاختبار،لغة الاختبار ،ارتباط الاختبار بالتوصيف،مضمون الاختبار) في اختبارات قسم العلوم التربوية في كلية التربية جامعة المجمعة ؛وقد شملت الدراسة(٨٠) اختبارًا موزعة على أربعة أعوام بواقع(٢٠) اختبار لكل عام منذ عام ١٤٣١ه حتى ١٤٣٥ه؛ وقد توصلت الدراسة إلى أنه: تحسنت نسب استيفاء الاختبارات التربوية لمعايير الجودة بدءًا من عام ١٤٣٢ ه،كما وجدت فروق ذات دلالة احصائية بين الاختبارات التربوية لعام ١٤٣١ ه وبين الأعوام التالية لصالح الأعوام التالية ،وتجاوزت نسب استيفاء معايير (البيانات الأساسية وتعليمات الاختبار ولغة الاختبار) نسبة ٩٠% ،وقد حصل معيار (لغة الاختبار) على أعلى نسب استيفاء على مستوى السنوات الأربع حيث تجاوزت نسبته ٩٦% ،بينما كانت أدنى نسبة مئوية للاستيفاء لمعيار (مضمون الاختبار) حيث حصل على نسبة ٦٨% في اختبارات عام ١٤٣١ه.

دراسة (حمد الكلثم ،٢٠١٥) وهدفت إلى تحديد معايير جودة الورقة الامتحانية من حيث الشكل والمضمون وجودة الأسئلة المقالية والموضوعية، ثم التعرف على مدى توافر هذه المعايير في الأوراق الامتحانية في اختبارات تخصص التربية الإسلامية في جامعتي أم القرى والدمام بواقع (١٥) ورقة اختبار لكل جامعة،ودراسة الفروق في نسب تحقق هذه المعايير في الأوراق الامتحانية للجامعتين وقد كشفت النتائج أن الأسئلة التي تستهدف مستوى التذكر كانت بالمرتبة الأولى مقارنة بالأسئلة التي

تستهدف باقي المستويات المعرفية؛ وقد وجدت فروق ذات دلالة احصائية فيما يتعلق بتحقق معايير جودة الاختبار وبتنوع الأسئلة لتشمل العديد من المستويات المعرفية وبمعايير جودة الأسئلة المقالية والموضوعية لصالح جامعة أم القرى.

دراسة (أماني الغامدي،سماح زغاري،٢٠١٨) وهدفت إلى تحليل عينة من اختبارات التقويم المستمر (اختبارات المنتصف) في كلية التربية بجامعة الامام عبد الرحمن بن فيصل وذلك لدراسة مدى استيفائها للمعايير الكيفية والشكلية التي تجعلها قادرة على قياس نواتج و نواتج التعلم لبرامج الكلية وقد توصلت الدراسة أن النسبة العامة لتحقق المؤشرات الشكلية ٨٣% بينما كانت نسبة المؤشرات الكيفية ٤٤% وهي نسبة غير مرضية في التعبير عن جودة مضمون وبناء اختبارات المنتصف التي تم تحليلها وكان عددها عشرين اختبارًا،حيث لم يكن هناك نوع من التوازن بين الأسئلة المقالية والموضوعية و لم يتم التوظيف الأمثل للأنواع المختلفة من الأسئلة الموضوعية ،كما كانت جميع الأسئلة المقالية ذات اجابات محددة ليست مفتوحة النهاية فلم تسمح بقياس نواتج التعلم في صورتها المتقدمة المبنية على عمليات عقلية نوعية توضح التمايز بين المستويات المتباينة للمتعلمين.

بالإضافة إلى تلك الدراسات التي تناولت واقع الاختبارات سواء الدورية منها أو النهائية وفحص مدى جودتها فقد اهتم بعض الباحثون بدراسة الأخطاء الشائعة في بناء الاختبارات منهم من ركز على أسئلة محددة مثل أسئلة الاختيار من متعدد كدراسة **(بندر بن زيد الأمير،٢٠١٦)** التي هدفت إلى دراسة الأخطاء الشائعة في بناء مفردات الاختبارات الموضوعية من نمط "الاختيار من متعدد" لدى معلمي الرياضيات بمدينة مكة المكرمة وتأثيرها على الخصائص السيكومترية للإختبارات التحصيلية ،حيث تم فحص عينة مكونة من (٥١) نموذج من الأسئلة لمعلمي الصف الثاني المتوسط وهو ما يمثل ٣٩.٢% من اجمالي نماذج أسئلة نهاية الفصل الدراسي الأول ،وقد تم فحصها وفقا لقائمة تدقيق الأخطاء التي تكونت من ٢٢ خطأ مقسمة على ثلاثة أبعاد رئيسية هي "الأخطاء في مدخل السؤال،الأخطاء في الإجابة الصحيحة ،الأخطاء في البدائل.

و كشفت الدراسة أن الأخطاء الشائعة في بناء مفردات الاختبارات من نمط الاختيار من متعدد خمسة أخطاء توزعت على الأقسام الثلاثة حيث وُجد خطآن شائعان في المشتتات وهما أن عدد البدائل غير مناسب لعمر الطالب والبدائل غير مرتبة من الأعلى إلى الأسفل أو العكس وكانت نسبة تكرارها على التوالي ٥١.٢% ، ٢٩.٩% ،كما وُجد خطآن شائعان في مدخل السؤال وهما الصياغة غير الواضحة للمفردة والنقل النصي للمفردة من الكتاب وكانت نسبة تكرارهما على التوالي ٢٧.٢% ، ٢٢.٧% ،بالإضافة إلى خطأ شائع واحد في الإجابة الصحيحة وهو عدم توزيع الإجابة الصحيحة توزيعًا عشوائيًا وكانت نسبة تكراره ١٠٠%؛ بالاضافة إلى ذلك فقد وجدت فروق ذات دلالة إحصائية في متوسطات معاملات الصعوبة والتمييز و فعالية المشتتات بين مفردات الاختبار الجيد (جميع فقراته سليمة البنية) ومفردات الاختبار غير الجيد (جميع فقراته تحتوي على الأخطاء الشائعة في بناء المفردات من نمط الاختيار من متعدد) لصالح النموذج الأول(الاختبار الجيد) ،كماوجدت فروق ذات دلالة إحصائية بين معاملات الثبات والصدق التلازمي بين الاختبارين لصالح الاختبار الجيد.

أما دراسة (هدى صالح ،حنان محمد ،٢٠١٧) فقد استهدفت تحديد الأخطاء الشائعة لمعلمات علم النفس في بناء مفردات أسئلة الصواب والخطأ ، وذلك بفحص عينة مكونة من (٤٠) نموذجًا من أسئلة معلمات علم النفس للفصل الدراسي الثاني لتحديد الأخطاء الشائعة فيها والتي تكررت بنسبة ٣٠% فأكثر، وقد

كشفت النتائج عن وجود أربعة أخطاء شائعة أولها " نقل العبارة نصًا من الكتاب المدرسي، والثلاثة الأخرى مرتبطة باخراج أسئلة الصواب والخطأ وهم "عبارات الصواب والخطأ غير متقاربين في الطول" ، "التوزيع غير العشوائي للعبارات" " نسبة عبارات الصواب أقل من ٤٠ % وأكثر من ٦٠%"؛ وقد كشفت النتائج أيضًا عن عدم وجود فروق ذات دلالة إحصائية في متوسطات معاملات السهولة بين مفردات الاختبار الجيد ومفردات الاختبار غير الجيد(جميع مفرداته تحتوي على الأخطاء الشائعة في بناء مفردات أسئلة الصواب والخطأ) بينما وجدت الفروق في معاملات التمييز بين مفردات الاختبارين تعود إلى الأخطاء الشائعة في بناء مفردات أسئلة النموذج غير الجيد ؛كما وجدت فروق ذات دلالة إحصائية بين معاملات الثبات للاختبارين لصالح الاختبار الجيد بينما لم توجد الفروق بينهما في معاملات صدق المحك.

ومن الدراسات التي هدفت لدراسة تلك الأخطاء على الأسئلة الموضوعية بصفة عامة دراسة (علي صلاح حسن ،٢٠١٨) حيث قام باعداد قائمة لتدقيق الأخطاء في بناء الاختبارات الموضوعية ثم فحص في ضوئها مجموعة عشوائية من الاختبارات الموضوعية لأعضاء هيئة التدريس من كليات وأقسام مختلفة بجامعة أسيوط ؛وقد كشفت النتائج أنه بالنسبة لأسئلة الصواب والخطأ كانت نسبة الأخطاء الإملائية واستخدام صيغة السؤال في الفقرات أقل من ٢٠% بينما كانت نسب باقي الأخطاء تتراوح بين ٢٥% : ٩٤% وهي نسب كبيرة أما بالنسبة لأسئلة الاختيار من متعدد فتراوحت نسب الأخطاء ما بين ٢٥% : ١٠٠% موزعة ما بين مدخل الفقرة والبدائل والاجابات الصحيحة وهي نسب كبيرة أيضًا مما يدل على عدم اهتمام كثير من أعضاء هيئة التدريس بالخطوات العلمية لبناء الاختبارات ومراعاة شروط وقواعد بناء الاختبارت الموضوعية؛كما كشفت النتائج أيضًا عن وجود فروق دالة احصائيًا بين (الاختبار الموضوعي الجيد والاختبار المتضمن على أخطاء) في الخصائص السيكومترية للاختباركل من حيث(متوسطات معاملات الثبات والصدق للاختبار) ولفقراته من حيث (متوسطات معاملات السهولة والصعوبة والتمييز) لصالح الصورة الجيدة وذلك في مادة الاختبارات النفسية لطلاب الفرقة الرابعة شعبة علم النفس.

ومن جانب آخر فهناك دراسات اهتمت بالأخطاء اللُّغوية في الامتحانات منها دراسة (سليمان خاطر،٢٠٠٨) التي هدفت إلى فحص الأخطاء اللُّغوية في أوراق أسئلة الاختبارات النهائية بكلية المعلمين في حائل وذلك على عينة عشوائية من خمسة فصول دراسية متتالية تشمل جميع التخصصات التي تدرس باللغة العربية في الأقسام العلمية بالكلية على أن يكون معد الاختبار في كل ورقة مختلفًا عنه في الورقة الأخرى، وقد شمل بالفحص الأخطاء اللغوية في جميع مستوياتها إملائية وتعبيرية وصرفية ونحوية ودلالية وأسلوبية، وقد خلص من ذلك إلى أن الأخطاء اللغوية في أوراق الإختبارات كثيرة ، ولم تعد حالات نادرة ، وأضاف أن معظم تلك الأخطاء من بين القواعد الأولية للعربية التي تدرس في المراحل الأولى من التعليم العام ، ويتكرر تدريسها إلى بدايات التعليم الجامعي وهذا يؤكد أن المشكلة ليست معرفية علمية من الجهل بتلك القواعد السهلة ، وإنما هي من قلة العناية والاهتمام والحرص على سلامة اللغة عند عامة أهلها ، وقد اقترح بعض المقترحات التي تساعد على حل تلك المشكلة حتى لا تستفحل ويُعتاد عليها بمرور الزمن فيصعب علاجها منها: اقامة الدورات التذكيرية لأعضاء هيئة التدريس ، وتبادل النظر في أوراق الأسئلة بعد انتهاء الاختبار لمراجعة الأخطاء وعدم تكرارها ، ودراسة الأسئلة علميًا وأسلوبيًّا وتربويًّا في شكلها ومضمونها من قبل القسم أو لجنة على مستوى الكلية لفحص الأوراق والتنبيه على الأخطاء لعدم تكرارها.

ومن الجدير بالذكر أنه ينبغي أن يصاحب تقويم الامتحانات وفحص جودة مفرداتها والتحقق من مطابقتها للمواصفات القياسية قياس رأي المستهدفين وتحديد مدى رضاهم عنها والتعرف على مستوى تقييمهم لتحقق النواتج المستهدفة من المقررات الدراسية ومدى جودة الاختبارات في قياس تلك النواتج وقد ***كشفت*** دراسة **(محمود إبراهيم ،محمد السيد ،٢٠١٥)** التي استهدفت التعرف على مستوى تقييم الطلاب لدراسة مقرر الأصول الفلسفية للتربية في ضوء نواتج التعلم المستهدفة (المعارف والمهارات الذهنية،والمهارات المهنية،والمهارات العامة والمنقولة) ***عن*** تحقق نواتج التعلم المعرفية والذهنية والمنقولة بدرجة كبيرة والمهنية بدرجة متوسطة ،كما أسفرت النتائج عن وجود فروق دالة احصائيًا حسب متغير الشعبة في إجمالي المحاور،ولم تظهر الفروق وفق متغير التقدير، وقد وجدت علاقة موجبة بين تحقق نواتج تعلم المقرر والعوامل المؤثرة في تحقيقها.

بالاضافة إلى رأي الطلاب أضاف **(محمد عبد القادر الشرقاوي،٢٠١٦)** رأى والمعلمين والموجهين وأولياء الأمور في دراسته التي هدفت إلى تقويم مخرجات التعلم المستهدفة لمادة التربية الرياضية لتلاميذ المرحلة الاعدادية في ضوء المعايير القومية لضمان الجودة والاعتماد ،وكانت أهم النتائج أن مخرجات التعلم المستهدفة تتحقق من وجهة نظر التلاميذ بنسبة ٦٨% ومن وجهة نظر المعلمين والموجهين بنسبة ٦٠.٩% أما من وجهة نظر أولياء الأمور فكانت تتحقق بنسبة ٦٥.٩% .

وأخيرًا ومع الاتجاه نحو التصحيح الآلي للاختبارات الموضوعية باستخدام أجهزة التصحيح الحديثة هدفت دراسة **(فاتن موسى،احسان حجازي،٢٠٢٠)**إلى تقويم الاختبارات المصححة آليًا والاختبارات المصححة يدويًا واتجاه طلبة الفرقة الرابعة شعبة علم النفس التربوي نحوهما حيث أوضحت النتائج أن اتجاهات الطلبة كانت ايجابية نحو الاختبارات المصححة آليًا ووجدت فروقًا بين نتائج الطلاب لصالح الاختبارات المصححة آليًا.

ومما تقدم يمكن القول أنه:

للوصول إلى التّميز في التعليم يجب أن تتطابق الأهداف المنشودة مع نواتج العملية التعليمية ،كما يلزم التحقق من هذا التطابق بعمليات التقويم الصحيحة والتي ينبغي أن تخضع هي الأخرى لعمليات التقويم للتحقق من جودتها وكفاءتها وصدق نتائجها في تقرير هذا التطابق ؛ومن ذلك تظهر أهمية تقويم واقع الامتحانات النهائية وفحص مدى جودة مفرداتها ومدى مطابقتها للمواصفات القياسية ،والتحقق من مدى صلاحيتها في قياس نواتج التعلم لاسيما أن يصحب ذلك برأي الطلاب فيها ومدى رضاهم عنها باعتبارهم المُدخل والمُخرج الأهم في المنظومة التعليمية وبذلك تحددت أسئلة الدراسة الحالية كالآتي:

١. ما مستوى رضا طلاب كلية التربية جامعة دمنهور عن عمليات التقييم والامتحانات التي يتعرضون لها أثناء دراستهم بالكلية؟
٢. ما نسبة تحقق المؤشرات الشكلية والكيفية لعينة من الأوراق الامتحانية النهائية بكلية التربية ؟
٣. ما مؤشرات الصلاحية السيكومترية للامتحانات النهائية المصحَّحَة آليًّا ؟
٤. إلى أي مدى تتحقق (نواتج التعلم المستهدفة من المقررات) من خلال نتائج الامتحانات النهائية المصحَّحَة آليًّا بالكلية ؟

اجراءات الدراسة: يتناول الجزء التالي من الدراسة بناء بطاقة المواصفات القياسية للورقة الامتحانية وكيفية استخدامها لفحص عينة الأوراق الامتحانية النهائية وبناء استبيان رضا الطلاب عن الامتحانات

التحصيلية والتحقق من صلاحيته للتطبيق انتهاء باجراءات تطبيق تلك الأدوات على عينة الأوراق الامتحانية والطلاب.

أولًا:بناء قائمة المواصفات القياسية للورقة الإمتحانية للاختبارات النهائية لكلية التربية جامعة دمنهور وفق الخطوات التالية:ملحق (١)

١. الرجوع إلى نواتج التعلم المستهدفة ومواصفات خريج كليات التربية المعلنة من قبل الهيئة القومية لضمان الجودة والاعتماد.
٢. المحاولات السابقة في مجال القياس والتقويم والدراسات السابقة التي عنيت باعداد بطاقات وقوائم لمعايير الاختبار الجيد .
٣. اعداد قائمة مبدئية مكونة من (٣٠) مؤشر تتعلق بالبيانات الرئيسية للاختبار واخراج الاختبار ومحتواه ومدى ملاءمتة لوقت الاختبار والجودة الفنية للأسئلة.
٤. التحقق من صدق القائمة من خلال عرضها على مجموعة من السادة المحكمين المتخصصين في (المناهج والقياس والتقويم) بكلية التربية جامعة دمنهور للتحقق من صدقها من حيث دقة صياغة كل مؤشر، ووضوح المعني المقصود بكل مؤشر، بالإضافة إلى ارتباط كل مؤشر بالمحور الخاص به وإبداء آرائهم حول مدى كفاية تلك المؤشرات ومقترحاتهم لتحسين القائمة وسهولة تطبيقها.
٥. اجراء التعديلات والمقترحات التي قام بها السادة المحكمون ،وقد تكونت القائمة في صورتها النهائية من محورين رئيسين وخمسة محاور فرعية (٢٧) مؤشر بحيث يشتمل المحور الأول (المؤشرات الشكلية) على (٣ محاور فرعية) هي **البيانات الأساسية للامتحان،وتعليمات الامتحان،** ثم **الاخراج العام للورقة الامتحانية** ،أما المحور الثاني (المؤشرات الفنيّة) يشتمل على (محورين فرعيين) هما **مضمون الأسئلة** الذي يستهدف مدى وضوح الأسئلة وشمولها وعلاقتها بأهداف المادة ثم **جودة صياغة الأنواع المختلفة للأسئلة** سواء الأسئلة المقالية أو الموضوعية (الاختيار من متعدد والصواب والخطأ) .
٦. تم التحقق من ثبات القائمة من خلال تحليل الأوراق الامتحانية المتضمنة في الدراسة الحالية والبالغ عددها ٤٦ ورقة امتحانية حيث تم التحليل من قبل الباحثة ومحللين آخرين ثم حساب نسبة الاتفاق بين المحللين باستخدام معادلة هلوستي

نسبة الاتفاق(P^o) = (عدد وحدات التحليل المُتفَق عليها / عدد وحدات التحليل الكلية) × 100%

وقد بلغت نسبة الاتفاق ٩٧.٦ مما يعد مؤشرًا على ثبات التحليل.

ثانيًا: استبيان رضا الطلاب عن الامتحانات التحصيلية النهائية:ملحق(٢)

تم إعداد استبيان رضا الطلاب عن الامتحانات التحصيلية النهائية وفقاً لعدد من الخطوات يمكن إيجازها فيما يلي :

١- تحديد الهدف من الاستبيان " حيث يهدف هذا الاستبيان إلى قياس رضا طلاب كلية التربية جامعة دمنهور عن الامتحانات التحصيلية النهائية.

٢- تحديد الجوانب المستهدف قياسها في ضوء الهدف من الدراسة من خلال توزيع استبيان مفتوح على عينة من الطلاب لتوضيح آرائهم حول الأوراق الامتحانية وأهم المشكلات أو الأخطاء فيها من وجهة نظرهم من حيث وضوح الأسئلة وعددها وتمثيلها للمقرر وتنوع المهارات المقاسة من

خلالها ومدى ملاءمتها لزمن الامتحان والدرجة المخصصة عليها بالاضافة إلى ملاحظاتهم حولها من حيث الصياغة اللغوية ومقترحاتهم لتحسينها.

٣- في ضوء نتائج الاستبيان المفتوح واستجابات الطلاب عليه وقائمة مواصفات الورقة الامتحانية تم صياغة مفردات المقياس وقد تضمن المقياس على ٢٦ مفردة موزعة بواقع (١١) مفردة للبعد الأول "المؤشرات الشكلية" (بواقع ٤مفردات لبعد البيانات الأساسية،٤ مفردات لتعليمات الامتحان،٣ مفردات للاخراج العام للامتحان)،(١٥) مفردة للبعد الثاني "المؤشرات الفنية"(بواقع ٩ مفردات لمضمون الأسئلة ،٦ مفردات لمدى جودة صياغة الأسئلة)يستجيب عليها المفحوص بطريقة التقرير الذاتي وفقًا لمقياس ليكرت خماسي التدريج (موافق بشدة،موافق ،غير متأكد ،غير موافق،غير موافق بشدة) .

٤- تم وضع سؤالًا مباشرًا مفتوحًا في نهاية الاستبانة يحدد فيه الطالب مدى رضاه عن الامتحانات التحصيلية النهائية ويعقب بتعليق لتوضيح سبب ذلك.

٥- للتحقق من صلاحية الاستبانة للإستخدام: تم القيام بالآتي:

أ- عرض الاستبانة على مجموعة من المحكمين المتخصصين فى المجال لاستطلاع آرائهم حول مدى كفاية المفردات وانتماء كل فقرة للبعد الذي تقيسه ومدى سلامة صياغتها اللغوية؛ وقد تم التعديل في ضوء الهدف من الدراسة.

ب- تم تطبيق الاستبانة على عينة استطلاعية من ٥٠ فردًا من طلاب كلية التربية (بواقع ٢٥ طالبًا و٢٥ طالبة) موزعين على الفرق الأربعة من الشعب المختلفة العلمية والأدبية في الفصل الدراسي الثاني لعام ٢٠٢٠/٢٠١٩ امتدت أعمارهم من ثمانية عشر سنة وستة شهور إلى اثنين وعشرين سنة وشهرين بمتوسط (٢٠سنة وشهرين) وإنحراف معيارى (٠.٢٩) .

ت- حساب معامل ألفا كرونباخ لمفردات كل بعد على حدة بعدد مفردات كل بعد وفى كل مرة يتم حذف درجات احدى المفردات من الدرجة الكلية للمهارة التى تنتمي إليها ومقارنة القيم الناتجة بمعامل ألفا الكلي للمهارة بدون حذف أى مفردة؛ وقد كانت قيم معاملات الثبات لجميع المفردات أقل من معاملات ثبات المحاور التى تنتمي إليها والتي كانت على الترتيب (٠.٨٨٥ ،٠.٨٨٠ ،٨٧٧، ٨٨٩ ،٧٩٩) أى أن جميع المفردات ثابتة حيث إن تدخل المفردة لا يؤدى إلى خفض معامل الثبات الكلي للمحور الذى تنتمي إليه المفردة ، كما أن معامل الثبات الكلي للاستبانة بطريقة ألفا يساوى (٠.٨٩٥) وهى قيمة مرتفعة .

ث- حساب القدرة التمييزية لكل مفرده من مفردات الاستبانة ،وامتدت بين (٠.٥٤١ ،٠.٦٩٧) وجميعها قيم موجبة ومرتفعة مما يعطي ثقة في صدق المفردات وقدرتها على التمييز.

ج- التحقق من الاتساق الداخلي للاستبانة : بحساب معامل الارتباط بين درجة كل مفرده ودرجة المحور الذى تنتمي إليه ، ثم حساب معاملات الارتباط بين درجة المحور والدرجة الكلية للاستبانة ،وقد كانت قيم معاملات الارتباط مرتفعة حيث امتدت بين

(٠.٧٣٩إلى٠.٩٢٧) مما يدل على وجود ارتباط قوى بين درجة المفردة ودرجة المحور الذى تنتمي إليه ، كما كانت معاملات الارتباط بين درجة كل محور والدرجة الكلية أكبر من (٠.٨٥٥) مما يدل على وجود اتساق داخلي للمقياس ككل وهذا مؤشر معقول على صدقة .

٦- تم وضع تعليمات مفصلة للاستبيان لطمأنة الطلاب من أن آرائهم موضع تقدير ويقصد منها التحسين والتطوير وهي سرية و لن تستخدم لأية أغراض أخرى سوى البحث العلمي بالإضافة إلى توضيح الهدف من الاستبانة وطريقة الاستجابة على المفردات:

- **فإذا كان البند متوفرًا في كافة الامتحانات التي تعرض لها فعليه اختيار" موافق بشدة "**
- **فإذا كان البند متوفرًا في معظم الامتحانات التي تعرض لها فعليه اختيار"موافق"**
- **فإذا كان لا يعرف بالضبط ولم يهتم بمعرفة مدى توفر البند فعليه اختيار "غير متأكد"**
- **فإذا كان البند متوفرًا في بعض الامتحانات فقط فعليه اختيار"غيرموافق"**
- **فإذا كان البند غيرمتوفرًا مطلقًا في أي الامتحانات التي تعرض لها فعليه اختيار "غير موافق بشدة "**

ثالثًا: تطبيق أدوات الدراسة على عينة الطلاب والأوراق الامتحانية

١. تم تطبيق استبيان رضا الطلاب عن الامتحانات النهائية بشكل ورقي والكتروني عبر شبكة الانترنت بالنسبة لطلاب كلية التربية من خلال الصفحة الرسمية لكلية التربية بدمنهور على شبكة الانترنت وكافة المنصات التعليمية التي أتيحت للباحثة، في بداية الفصل الدراسي الثاني من العام الجامعي ٢٠١٩/ ٢٠٢٠، وبلغ عدد الطلاب الذين استجابوا طالبا منهم (٦٤٩طالبا و طالبة). وتضمنت العينة طلابًا من الفرق المختلفة من جميع التخصصات عام وأساسي وطفولة وكانت الأغلبية لطلاب الشعب العامة بنسبة ٦٣.٧٩% وللفرق النهائية بنسبة ٥٢.٢٣% من اجمالي العينة على النحو المبين بجدول (١):

جدول(١) توزيع لأعداد الطلاب المستجيبن على استبيان الرضا عن الامتحانات النهائية من الشعب والفرق المختلفة

الشعبة/الفرقة	الفرقة الاولى		الفرقة الثانية		الفرقة الثالثة		الفرقة الرابعة		الاجمالي	النسبة المئوية بالنسبة للعينة
	ذكور	اناث	ذكور	اناث	ذكور	اناث	ذكور	اناث		
عام	١٩	٤١	١٣	٦٥	٧	٢٧	٣٢	٢١٠	٤١٤	%٦٣.٧٩
أساسي	٢	٣٦	٢	٣٦	١١	٣١	٢٣	١١	١٥٢	%٢٣.٤٢
طفولة	-	٨	-	١٠	-	٢	-	٦٣	٨٣	%١٢.٧٩
الاجمالي	١٠٦		١٢٦		٧٨		٣٣٩		٦٤٩	
النسبة المئوية بالنسبة للعينة	%١٦.٣٣		%١٩.٤١		%١٢.٠٢		%٥٢.٢٣			

٢. استخدام قائمة المواصفات القياسية للورقة الامتحانية لفحص عينة الأوراق الامتحانية النهائية موضع الدراسة وعددها(٤٦) ورقة امتحانية اختيرت من المقررات التربوية العامة التي يدرسها معظم الطلاب في الكلية وتم مراعاة أن تكون جميعها لأساتذة مختلفين وكان من بينهم خمسة أوراق

امتحانية تكررت على دفعات مختلفة حيث كان يدرّسهم نفس الأستاذ ولكنها اعتبرت ورقة امتحانية واحدة وتأخذ الأرقام (٢ ،٣٠،٣٢، ٣٤، ٣٩) في قائمة الفحص وتوزعت الأوراق الامتحانية على الفرق الأربعة بنسبة ٧٧.٤٢% من الشعب العامة ، ٤٢.٣١ % من الأساسي، ٤٥.٧١ % من الطفولة وذلك من إجمالي المقررات التربوية المقررة عليهم خلال الأربع سنوات وكانت الأوراق موزعة على أربع فصول دراسية تبدأ من الفصل الدراسي الثاني ٢٠١٨إلى الفصل الدراسي الأول ٢٠٢٠ (ملحق ٣)

٣. تحديد عينة من الامتحانات التي صححت أوراق الاجابة الخاصة بها آليًا (حيث تتوفر فيها درجات الطلاب على كل سؤال " وربط كل مفردة فيها بنواتج التعلم المستهدفة من المقرر) ثم استخراج التقارير الفنية للتحقق من مدى صلاحيتها السيكومترية في قياس تلك النواتج وكان عددها (١٤) اختبارًا .

٤. حساب مدى تحقق نواتج التعلم المستهدفة فعليًا لعينة الامتحانات التي صححت آليًا (المحددة في الخطوة السابقة) من خلال التقارير الفنية الخاصة بذلك والتي يوفرها برنامج التصحيح الآلي Remark classic (OMR)

رابعاً : نتائج الدراسة وتفسيرها ومناقشتها :

١. **تحليل نتائج السؤال الأول ومناقشتها :** للإجابة عن السؤال الأول الذي ينص على " ما مستوى رضا طلاب كلية التربية جامعة دمنهور عن عمليات التقييم والامتحانات التي يتعرضون لها أثناء دراستهم بالكلية؟" تم تطبيق استبيان رضا الطلاب عن الأوراق الامتحانية التي تُعرض عليهم في نهاية كل فصل دراسي على عينة الدراسة ،وحصر التكرارات ومعالجة بياناتها باستخدام مجموعة من الاحصاءات التي تستهدف وصف استجابات الطلاب على عبارات الاستبيان وهي:

- النسب المئوية لحساب التكرارات وكذلك المتوسط والانحراف المعياري لدرجة المقابلة لاستجابات الطلاب
- تحديد مستوى الموافقة على كل عبارة من عبارات المقياس وترتيبها على حسب وزنها النسبي ،ويتم حساب الوزن النسبي لكل عبارة عن طريق اعطاء درجة لكل استجابة من الاستجابات الخمسة وفقًا لطريقة ليكرت من (٥ :١) للعبارات الموجبة ومن (١: ٥) للعبارات السالبة وهم المفردات (١٦ ،٢٣، ٢٤ ، ٢٥) ثم تضرب هذه الدرجات في التكرار المقابل لكل استجابة وجمعها ثم قسمتها على اجمالي أفراد العينة ويعطي ما يسمى بالوسط المرجح الذي يعبر عن الوزن النسبي لكل عبارة على حدة كما يلي:

$$\text{الوزن النسبي للعبارة} = \frac{(٥\times\text{تكرار موافق بشدة})+(٤\times\text{تكرار موافق})+(٣\times\text{تكرار غير متأكد})+(٢\times\text{تكرار غير موافق})+(١\times\text{تكرار غير موافق بشدة})}{\text{عدد أفراد العينة}}$$

وتحدد مستوى ومدى الموافقة على المفردة من خلال المعادلة(مستوى الموافقة=(ن-١)/ن)حيث تشير (ن)إلى عدد الاستجابات وتساوي (٥) وبذلك يتحدد مستوى ومدى الموافقة لكل استجابة للعبارات الموجبة كالتالي:(غير موافق بشدة من ١ إلى ١.٨٠)،(غير موافق من ١.٨١ إلى ٢.٦٠)،(غير متأكد من

٢.٦١إلى ٣.٤٠)، (موافق من ٣.٤١ إلى ٤.٢٠)، (موافق بشدة من ٤.٢١ إلى ٥) وينعكس هذا الترتيب للعبارت السالبة

وكانت نتائج آراء الطلاب حول المواصفات الشكلية للاختبارات (البيانات الأساسية ،تعليمات الامتحان ،الاخراج العام للورقة الامتحانية) موضحة بجدول(٢)

جدول (٢) استجابات الطلاب على أبعاد الاستبيان الخاصة بالمواصفات الشكلية للاختبارات ن=٦٤٩

رقم المفردة	العبارة		درجة الموافقة					مجموع الأوزان	الانحراف المعياري	المتوسط والوزن النسبي	مستوى الموافقة	الترتيب
			موافق بشدة	موافق	غير متأكد	غير موافق	غير موافق بشدة					
١	البيانات الأساسية للجامعة والكلية والمقرر	ك	٢٤٥	٣١٦	٦٦	١٧	٥	٢٧٢٦	٠.٧٨٤	٤.٢٠	موافق	٢.٥
		%	٣٧.٨	٤٨.٧	١٠.٢	٢.٦	٠.٨					
٢	البيانات الأساسية للقسم والمقرر /كود المقرر	ك	٢٠٨	١٩٠	١٧١	٦٣	١٧	٢٤٥٦	١.٠٧٩	٣.٧٨	موافق	٩
		%	٣٢.٠	٢٩.٣	٢٦.٣	٩.٧	٢.٦					
٣	البيانات المتعلقة بزمن الامتحان والدرجة الكلية له	ك	٢٢٦	٣٠٦	٥٩	٤٦	١١	٢٦٣٤	٠.٩٣٥	٤.٠٦	موافق	٥
		%	٣٤.٨	٤٧.١	٩.١	٧.١	١.٧					
٤	الدرجة الخاصة بكل سؤال، و كل مفردة	ك	١٤٢	٢٢٩	١٠١	١٤٥	٣٢	٢٢٥١	١.١٩٦	٣.٤٧	موافق	١١
		%	٢١.٩	٣٥.٣	١٥.٦	٢٢.٣	٤.٩					
٥	طريقة الاجابة عن أسئلة الامتحان	ك	١٨٢	٢٦٧	٦٩	٩٤	٣٧	٢٤١٠	١.١٨٣	٣.٧١	موافق	١٠
		%	٢٨.٠	٤١.١	١٠.٦	١٤.٥	٥.٧					
٦	يكتب رأس السؤال بخط مميز وواضح عن باقي المفردات	ك	٢٦٣	٣٠٤	٤٢	٣٢	٨	٢٧٢٩	٠.٨٥٩	٤.٢٠	موافق	٢.٥
		%	٤٠.٥	٤٦.٨	٦.٥	٤.٩	١.٢					
٧	تمييز أية تعليمات مراد تنبيه الطلاب لها	ك	٢٠٤	٢٥٤	٧٩	٩٣	١٩	٢٤٧٨	١.١١٤	٣.٨٢	موافق	٧
		%	٣١.٤	٣٩.١	١٢.٢	١٤.٣	٢.٩					
٨	كافة التعليمات المفصلة في حالة التصحيح الالي	ك	٢٦٥	٢٣٠	٩٠	٤٩	١٥	٢٦٢٨	١.٠٢٨	٤.٠٥	موافق	٦
		%	٤٠.٨	٣٥.٤	١٣.٩	٧.٦	٢.٣					
٩	يكتب الامتحان على الكمبيوتر بخط واضح وتراعى المسافة البينية	ك	٢٨٠	٢٧٦	٣٨	٤٨	٧	٢٧٢١	٠.٩٢٠	٤.١٩	موافق	٤
		%	٤٣.١	٤٢.٥	٥.٩	٧.٤	١.١					
١٠	الورقة الامتحانية واضحة الطباعة و سهلة القراءة	ك	٣١٤	٢٦٧	٣٥	٢٩	٤	٢٨٠٥	٠.٨١٨	٤.٣٢	موافق بشدة	١
		%	٤٨.٤	٤١.١	٥.٤	٤.٥	٠.٦					
١١	الورقة الامتحانية خالية من الأخطاء اللغوية في الصياغة	ك	١٨٩	٢٥٤	١١٤	٧٦	١٦	٢٤٧١	١.٠٥٩	٣.٨١	موافق	٨
		%	٢٩.١	٣٩.١	١٧.٦	١١.٧	٢.٥					
الاجمالي		ك	٢٥١٨	٢٨٩٣	٨٦٤	٦٩٢	١٧١	٢٨٣٠٩		٣.٩٧	موافق	
		%	٣٥.٣	٤٠.٥	١٢.١	٩.٧	٢.٤					

يتضح من جدول (٢)أن:

١. جميع العبارات كانت بدرجة موافق ماعدا العبارة العاشرة " الأوراق الامتحانية واضحة الطباعة وسهلة القراءة كانت بدرجة موافق بشدة وتراوحت الأوزان النسبية بين (٣.٤٧) و(٤.٣٢) حيث

أكدت عينة الدراسة توافر المواصفات الشكلية للأوراق الامتحانية التي يتعرضون لها في نهاية كل فصل دراسي ؛حيث بلغ متوسط استجابة أفراد العينة على المحور ككل (٣.٩٧).

٢. كانت نسبة الموافقة بدرجة موافق تساوي(٤٠.٥%) وهي أعلى نسبة ،وبدرجة موافق بشدة (٣٥.٣%) وهي نسبة مرتفعة،وبدرجة غير موافق بشدة (٢.٤%) وهي نسبة منخفضة جدا وهذا مؤشر على أن آراء الطلاب نحو المواصفات الشكلية للأوراق الامتحانية كانت في الاتجاه الإيجابي .

٣. كانت العبارة العاشرة(الورقة الامتحانية واضحة الطباعة وسهلة القراءة) الأعلى بين عبارات المحور بنسبة (٤.٣٢) وهذا يوضح حرص أعضاء هيئة التدريس (أساتذة المواد)على الطباعة الواضحة للامتحانات لتقليل العبء على الطلاب أثناء الامتحانات وتقليل الاستفسارات والمشتتات وحرص وادارة الكلية على توفير الأدوات لذلك .

٤. كان الترتيب الثاني للعبارات (١ ، ٦) حيث حصلتا على نفس الوزن النسبي (٤.٢٠) وبالتالي كان الترتيب بمتوسط الرتبتين يساوي(٢.٥) ومن ذلك يرى الطلاب أنه تتوفر بالأوراق الامتحانية البيانات الأساسية للجامعة والكلية والمقرر كما تحدد فيها الأسئلة الرئيسية بوضوح وهذا يؤكد وضوح الورقة الامتحانية وسهولة قراءتها.

٥. كانت العبارة الخامسة(تتضمن الورقة الامتحانية طريقة الاجابة عن أسئلة الامتحان) في الترتيب العاشر بنسبة (٣.٧١) والعبارة الرابعة (تتضمن الورقة الامتحانية الدرجة الخاصة بكل سؤال وكل مفردة) في الترتيب الحادي عشر والأخير بنسبة (٣.٤٧)؛ وقد اتضح ذلك أيضًا من تعليقات الطلاب على السؤال المفتوح حيث تمحورت بعض التعليقات حول الرغبة الشديدة لدى بعض الطلاب في توضيح طريقة الاجابة المرغوبة عند أستاذ الماة هل كتابة كل التفاصيل الموجودة نصًا في الكتاب أم توضيح العناصر وإضافة التعليقات الشخصية ،هل كتابة كل العناصر بالترتيب أم وضع المخططات والرسوم التوضيحية والجداول،البدء بالمقدمات ثم التحديد أم البدء بالعناصر ثم الاجمال وكلها تعليقات توحي بأنهم يفتقدون هذا البند في اختباراتهم؛ كما تمحور البعض الآخر من التعليقات حول رغبتهم في تحديد الدرجات الخاصة بكل سؤال رئيسي وكل مفردة فرعية لأن ذلك يفيدهم في تحديد أولوياتهم وتنظيم وقت الاجابة وقد علّق بعض الطلاب أن كثير من الامتحانات تتضمن الدرجة الكلية للاختبار دون التفاصيل وقد يفسّر ذلك كون المفردة الثالثة الخاصة بتضمن الأوراق الامتحانية على زمن الامتحان والدرجة الكلية في ترتيب متقدّم (الترتيب الخامس).

وكانت نتائج آراء الطلاب حول المواصفات الكيفية للأوراق الامتحانية (مضمون الأسئلة،مدى جودة صياغتها) موضحة بجدول (٣):

جدول (٣) استجابات الطلاب على أبعاد الاستبيان الخاصة بالمواصفات الفنِّية للاختبارات ن=٦٤٩

رقم المفردة	العبارة		درجة الموافقة					مجموع الأوزان	الانحراف المعياري	المتوسط والوزن النسبي	مستوى الموافقة	الترتيب
			موافق بشدة	موافق	غير متأكد	غير موافق	غير موافق بشدة					
١٢	صياغة واضحة المعنى ومحددة ودقيقة المعني للمطلوب	ك	١٦٣	٢٥١	٩٨	١٠٧	٣٠	٢٣٥٧	١.١٥٩	٣.٦٣	موافق	٦
		%	٢٥.١	٣٨.٧	١٥.١	١٦.٥	٤.٦					
١٣	الأسئلة خالية من المصطلحات الغامضة	ك	١٣٩	٢١١	١١٣	١٤٩	٣٧	٢٢١٣	١.٢١٤	٣.٤١	موافق	١٠
		%	٢١.٤	٣٢.٥	١٧.٤	٢٣.٠	٥.٧					
١٤	تغطى أسئلة الامتحان كل محتوى المقرر	ك	١٢٨	٢٦٨	١١٠	١٠٩	٣٤	٢٢٩٤	١.١٣٨	٣.٥٣	موافق	٨
		%	١٩.٧	٤١.٣	١٦.٩	١٦.٨	٥.٢					
١٥	عدد الأسئلة ملائما	ك	١٤٦	١٩٩	٧١	١٥٥	٧٨	٢١٢٧	١.٣٦١	٣.٢٨	غير متأكد	١١

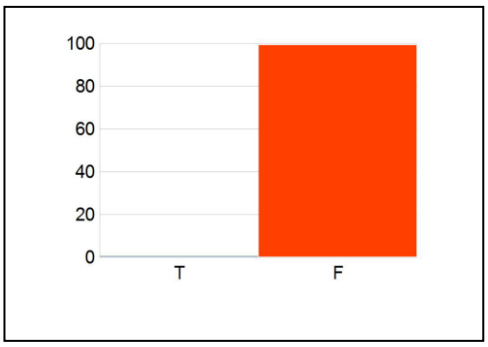
رقم المفردة	العبارة		درجة الموافقة					مجموع الأوزان	الانحراف المعياري	المتوسط والوزن النسبي	مستوى الموافقة	الترتيب
			موافق بشدة	موافق	غير متأكد	غير موافق	غير موافق بشدة					
	لزمن الاجابة على الامتحان	%	٢٢.٥	٣٠.٧	١٠.٩	٢٣.٩	١٢.٠					
١٦	تكرر الأسئلة بصورة كاملة من امتحانات سابقة	ك	١٣١	١٩٦	١٦١	١٢٥	٣٦	١٦٨٦	١.١٦٨	٢.٦٠	موافق	١٣
		%	٢٠.٢	٣٠.٢	٢٤.٨	١٩.٣	٥.٥					
١٧	ترتبط جميع أسئلة الامتحان بالمقرر الذي تم تدريسة	ك	٢٠٤	٢٩٥	٦١	٦٤	٢٥	٢٥٣٦	١.٠٦٩	٣.٩١	موافق	١
		%	٣١.٤	٤٥.٥	٩.٤	٩.٩	٣.٩					
١٨	تقيس أسئلة الامتحان مستويات التفكير المختلفة	ك	١٣٠	٣٢٦	٧٩	٦٥	٤٩	٢٣٧٠	١.١٣٣	٣.٦٥	موافق	٣.٥
		%	٢٠.٠	٥٠.٢	١٢.٢	١٠.٠	٧.٦					
١٩	تتنوع أسئلة الامتحان مابين (مقالية انشائية ، موضوعية بأنواعها)	ك	١٣٥	٣٢٢	٦١	٩٣	٣٨	٢٣٧٠	١.١٣٣	٣.٦٥	موافق	٣.٥
		%	٢٠.٨	٤٩.٦	٩.٤	١٤.٣	٥.٩					
٢٠	يوجد توازن بين عدد الأسئلة الموضوعية والمقالية وتوزيع الدرجات بينهم	ك	٩٨	٢٦٢	١٣٨	١٢٣	٢٨	٢٢٢٦	١.٠٨٩	٣.٤٣	موافق	٩
		%	١٥.١	٤٠.٤	٢١.٣	١٩.٠	٤.٣					
٢١	الأسئلة المقالية مصاغة بدقة ومحدد فيها كافة العناصر المطلوبة	ك	١٤٧	٣٢٥	٨٢	٧٣	٢٢	٢٤٤٩	١.٠٢٩	٣.٧٧	موافق	٢
		%	٢٢.٧	٥٠.١	١٢.٦	١١.٢	٣.٤					
٢٢	تتضمن الورقة الامتحانية بعض الأسئلة المقالية ذات النهايات المفتوحة تسمح بالابداع والابتكار	ك	١٠٩	٣١٠	٩٨	٨٧	٤٥	٢٢٩٨	١.١٢٧	٣.٥٤	موافق	٧
		%	١٦.٨	٤٧.٨	١٥.١	١٣.٤	٦.٩					
٢٣	الأسئلة الموضوعية تتضمن بعض الأسئلة والبدائل البديهية.	ك	٩٧	٣١٥	١٢٥	٨٥	٢٧	١٥٧٧	١.٠٢٨	٢.٤٣	موافق	١٤
		%	١٤.٩	٤٨.٥	١٩.٣	١٣.١	٤.٢					
٢٤	يوجد تداخل وتكرار بين اجابات الأسئلة فقد يجيب سؤال عن سؤال آخر	ك	٧٩	٢٤٣	١٣٠	١٤٦	٥١	١٧٩٤	١.١٦١	٢.٧٦	غير متأكد	١٢
		%	١٢.٢	٣٧.٤	٢٠.٠	٢٢.٥	٧.٩					
٢٥	بعض الفقرات المنقولة نصًا من الكتاب	ك	١٢٩	٣١٥	٩١	٨٩	٢٥	١٥١٣	١.٠٦١	٢.٣٣	موافق	١٥
		%	١٩.٩	٤٨.٥	١٤.٠	١٣.٧	٣.٩					
٢٦	التوازن والتوزيع العشوائي للإجابات بين البدائل	ك	٨٣	٣٤٦	١٣٩	٦٦	١٥	٢٣٦٣	٠.٩١١	٣.٦٤	موافق	٥
		%	١٢.٨	٥٣.٣	٢١.٤	١٠.٢	٢.٣					
الاجمالي		ك	١٩١٨	٤١٨٤	١٥٥٧	١٥٣٦	٥٤٠	٣٤٦٠٩		٣.٥٦	موافق	
		%	١٩.٧	٤٣.٠	١٦.٠	١٥.٨	٥.٥					
٢٧	هل أنت راضٍ عن الامتحانات التحصيلية النهائية؟		راض		غير متأكد		غير راض		٠.٩٢٣	٢.٢٩		
		ك	٣٩٧		٤٥		٢٠٧	١٤٨٨			غيرمتأكد	
		%	٦١.٢		٦.٩		٣١.٩					

يتضح من جدول (٣)أن:

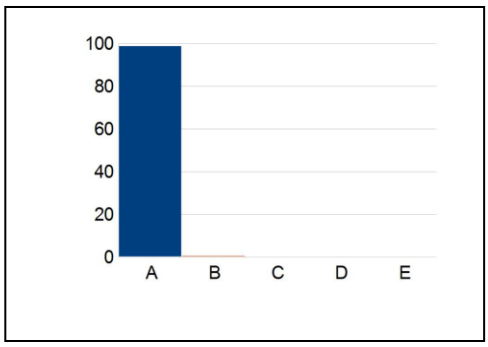
١. تراوحت درجة الموافقة على جميع العبارات بين درجة (موافق) ودرجة (غير متأكد) وقد تراوحت الأوزان النسبية بين (٢.٣٣) و(٣.٩١) و بلغ متوسط استجابة أفراد العينة على المحور ككل (٣.٥٦) بدرجة موافق.
٢. كانت نسبة الموافقة بدرجة موافق تساوي(٤٣%) وهي أعلى نسبة ،وتقاربت نسب الموافقة بدرجة موافق بشدة وغير متأكد وغير موافق حيث كانت على الترتيب (١٩.٧ %، ١٦%، ١٥.٨%) وكانت درجة غير موافق بشدة (٥.٥%) وهي نسبة منخفضة جدا وهذا مؤشر على تباين آراء الطلاب نحو توافر المواصفات الكيفية للأوراق الامتحانية وعدم حسم ١٦% منهم لموقفهم نحوها.
٣. جاءت العبارة رقم (١٧)"ترتبط جميع أسئلة الامتحان بالمقرر الذي تم تدريسه" الأعلى بين عبارات المحور بنسبة (٣.٩١) وهذا يعني موافقتهم ورضاهم عن محتوى الامتحانات؛ في الوقت نفسه أقر الكثير من الطلاب في تعليقاتهم حول هذا البند على وجه التحديد أن ذلك لا يعني أنهم لا يجدون مفاجئات في أسئلة الامتحانات ففي الغالب تكون الأسئلة من المقرر الذي تم تدريسه ولكن منها ما يكون غير متوقع ؛ولم يتم التأكيد على ادراجه ضمن أسئلة الامتحان من قبل أساتذه المواد.
٤. كانت العبارات(١٦، ٢٣، ٢٥) وهي عبارات سالبة في الترتيب الأدني حيث احتلت الترتيب (١٣ ، ١٤ ،١٥) على التوالي و كانت على درجة (موافق) وهذا يدل على عدم رضا الطلاب عن الامتحانات فيما يتعلق بهذه البنود فقد وافق أكثر من ٥٠% من الطلاب بدرجات (موافق بشدة/موافق) على أن الأسئلة تتكرر بصورة كاملة من اختبارات سابقة بينما لم يحسم ما يقارب ٢٥% منهم موقفهم من ذلك، وفيما يتعلق بجودة الاسئلة خصوصًا الموضوعية منها فقد وافق ٦٣.٤% من الطلاب بدرجات (موافق بشدة/موافق) على عدم جودة الأسئلة ووجود بعض الأسئلة والبدائل البديهية بينما لم يحسم ١٩.٣% منهم موقفهم من ذلك ،كماوافق ٦٨.٤% من الطلاب بدرجات (موافق بشدة/موافق) على وجود بعض الفقرات منقولة نصًا من الكتاب بينما لم يحسم ١٤% منهم موقفهم من ذلك؛ وقد تأكدت تلك الآراء من بعض تعليقات الطلاب التي دارت حول أن بعض الأسئلة تكون صعبة جدا وتأتي من بين السطور "وقد يجدوها مكتوبة نصًا في الكتاب " وبعضها الآخر غامض جدًا ويمكن أن "يحمل أكثر من معنى" .
٥. حصلت العبارة (٢٧)المحددة في هدفها نحو مدى الرضا عن الامتحانات النهائية والمحددة في بدائلها على نسبة (٢.٢٩) بدرجة (غير متأكد) مما يدل على عدم حسم الطلاب لرضاهم عن الامتحانات النهائية وحاجتها إلى التطوير واتضح ذلك من تعليقات بعض الطلاب التي دارت حول "البحث عن طرق بديلة عادلة لقياس قدرات ومهارات الطلاب بعدالة وموضوعية ومحاولة اثبات الشفافية في التصحيح والاعلان الفوري للنتائج"
٢. **تحليل نتائج السؤال الثاني ومناقشتها :** للإجابة عن السؤال الثاني الذي ينص على:"ما نسبة تحقق المؤشرات الشكلية والكيفية لعينة من الأوراق الامتحانية النهائية بكلية التربية ؟" تم استخدام قائمة المواصفات القياسية للورقة الامتحانية لفحص عينة الأوراق الامتحانية النهائية موضع الدراسة كالتالي :-
   أ- تم تحليل كل اختبار وفق المؤشرات باعطاء ٣ درجات عند توفر المعيار كاملًا ، درجتان إذا توفر المؤشر في معظم بنوده ، ودرجة واحدة إذا توفر المؤشر إلى حدما في أحد بنوده فقط ، صفر إذا لم يتوفرمطلقًا.

ب- تم فحص الأوراق الامتحانية ومطابقتها بتوصيف المقررات الموجودة بوحدة ضمان الجودة بالكلية للتحقق من استيفاء مؤشر شمول الورقة الامتحانية لكل محتوى مقرر وارتباطة بنواتج التعلم المستهدفة.

ت- بالنسبة للمؤشرات الخاصة بكل من (الصياغة الواضحة للأسئلة ،جودة الأسئلة الموضوعيه والابتعادعن البدائل البديهية) وبالتحديد في المقررات التي صححت آليًا تم الاستعانة بالمعاملات الاحصائية للبدائل وتكرارات اختيارات الطلاب في التقارير الفنية للاختبارات فإذا كانت نسب اختيار الطلاب للاجابة الصحيحة ١٠٠% دل ذلك على وضوحها وتأخذ الأشكال (١، ٢) في التقارير الفنية مما يدل على وضوحها وعدم كفاءة البدائل والمشتتات

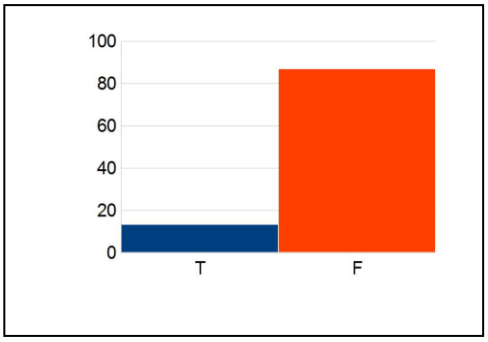


شكل(١)
مفردة غيرجيدة من نوع الصواب والخطأ

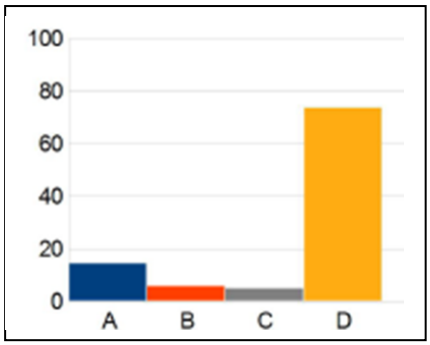


شكل (٢)
مفردة غيرجيدة من نوع الاختيار من متعدد

أما إذا كانت صياغة الأسئلة والبدائل جيدة فمن الطبيعي أن تلفت انتباه الطلاب ويختارها البعض بنسب معينة مع بقاء الاجابة الصحيحة النسبة الأعلى وتأخذ الأشكال(٣،٤)



شكل(٣)
مفردة جيدة من نوع الصواب والخطأ



شكل (٤)
مفردة جيدة من نوع الاختيار من متعدد

فالمشتت الفعال يجذب أكبر عدد من أفراد المجموعة الدنيا ،والمشتت غير الفعال هو الذي لا يختاره أحد ؛حيث تكون نسبة اختياره من أفراد المجموعة الدنيا والعليا قليلة ،أونسبة اختياره من أفراد المجموعة العليا أكبر منه في المجموعة الدنيا (صلاح مراد،أمين سليمان،٢٠٠٥)

ث- تم حذف درجة أي محور مرتبط بنوع معين من الأسئلة إذا لم يتوفر في ورقة الامتحان ، فإذا اعتمد على الأسئلة الموضوعية دون المقالية تحذف درجة الأسئلة المقالية،وهكذا كما تم

حذف درجة مؤشر "وجود كافة التعليمات في حالة التصحيح الآلي" إذا كان التصحيح يدويًا وبذلك تكون الدرجة النهائية للاختبار إذا توفر فيه كافة المؤشرات (٨١ درجة) موزعة تبعًا للمحاور الفرعية كالتالي:

جدول(٤)الدرجات الفرعية والكلية للمؤشرات الشكلية والكيفية للورقة الامتحانية

المواصفات	المؤشرات الشكلية			المؤشرات الكيفية			
المحاور الفرعية	البيانات الأساسية	تعليمات الامتحان	الاخراج العام	مضمون الأسئلة	جودة صياغة الأسئلة		
					المقالية	الاختيار من متعدد	الصواب والخطأ
عدد المؤشرات	٤	٤	٣	٧	٢	٤	٣
درجة المحور	١٢	١٢	٩	٢١	٦	١٢	٩
الدرجة الكلية	٣٣			٤٨			

ويوضح ملحق (٤، ٥) كل من مصفوفة تحليل محتوى عينة الأوراق الامتحانية التربوية في الدراسة الحالية من حيث المؤشرات الشكلية والكيفية

ج- استخراج درجة كل ورقة امتحانية على قائمة المواصفات واستخراج النسب المئوية لمدى توافرها من خلال الدرجة الكلية للاختبار حيث تفاوتت الدرجة الكلية للاختبار على قائمة المواصفات وفقا لنوعية الأسئلة المتضمنة وطبيعة التصحيح آلي أم يدوي وفقا لجدول (٥)

جدول (٥)الدرجات الكلية للمؤشرات الشكلية والكيفية لأشكال الامتحانات المختلفة

التصحيح	نوع الأسئلة المتضمنة في الامتحان	المؤشرات الشكلية	المؤشرات الفنيّة	الدرجة الكلية
يدوي	مقالية فقط	٣٠	٢٧	٥٧
يدوي	مقالية+صواب وخطأ	٣٠	٣٦	٦٦
يدوي	مقالية +صواب وخطأ+اختيار من متعدد	٣٠	٤٨	٧٨
يدوي	صواب وخطأ فقط	٣٠	٣٠	٦٠
آلي	مقالية+صواب وخطأ+اختيار من متعدد	٣٣	٤٨	٨١
آلي	صواب وخطأ+اختيار من متعدد	٣٣	٤٢	٧٥
آلي	صواب وخطأ	٣٣	٣٠	٦٣
آلي	اختيار من متعدد	٣٣	٣٣	٦٦

ويوضح جدول (٦)النسب المئوية لتوفر المواصفات القياسية لكل اختبار من الناحية الشكلية والكيفية على مستوى الأبعاد الفرعية والدرجة الكلية.

جدول (٦)النسب المئوية لتوفر المواصفات القياسية لكل اختبار من الناحية الشكلية والكيفية

رقم الورقة في الفحص	طريقة التصحيح	البيانات الأساسية	التعليمات	الاخراج العام	النسبة المئوية للمؤشرات الشكلية	مضمون الأسئلة	الأسئلة المقالية	أسئلة الاختيار من متعدد	أسئلة الصواب والخطأ	النسبة المئوية للمؤشرات الكيفية	النسبة المئوية لاجمالي المؤشرات
١	يدوي	%٣٣.٣٣	%٥٠	%١٠٠	%٦٣.٣٣	%٣٣.٣٣	%٥٠	-	-	%٣٧.٠٤	%٥٠.٨٨
٢	يدوي	%٤١.٦٧	%٥٠	%١٠٠	%٦٦.٦٧	%٣٨.١٠	%٥٠	-	-	%٤٠.٧٤	%٥٤.٣٩
٣	يدوي	%٤١.٦٧	٥٨.٣٣ %	%١٠٠	%٧٠.٠٠	%٤٢.٨٦	%٦٦.٦٧	-	-	%٤٨.١٥	%٥٩.٦٥
٤	يدوي	%٥٨.٣٣	%٢٥	%١٠٠	%٦٣.٣٣	%٤٢.٨٦	%٣٣.٣٣	-	-	%٤٠.٧٤	%٥٢.٦٣
٥	يدوي	%٥٨.٣٣	%٥٠	%١٠٠	%٧٣.٣٣	%٤٧.٦٢	%٣٣.٣٣	-	-	%٤٤.٤٤	%٥٩.٦٥
٦	يدوي	%٦٦.٦٧	%٧٥	%١٠٠	%٨٦.٦٧	%٩٥.٢٤	%١٠٠	-	%٨٨.٨٩	%٩٤.٤٤	%٩٠.٩١
٧	يدوي	%٧٥	٤١.٦٧ %	%١٠٠	%٧٦.٦٧	%٣٣.٣٣	%١٦.٦٧	-	-	%٢٩.٦٣	%٥٤.٣٩
٨	يدوي	%٥٠	٦٦.٦٧ %	%١٠٠	%٧٦.٦٧	%٥٢.٣٨	%٥٠	-	-	%٥١.٨٥	%٦٤.٩١

رقم الورقة في الفحص	طريقة التصحيح	البيانات الأساسية	التعليمات	الاخراج العام	النسبة المئوية للمؤشرات الشكلية	مضمون الأسئلة	الأسئلة المقالية	أسئلة الاختيار من متعدد	أسئلة الصواب والخطأ	النسبة المئوية للمؤشرات الكيفية	النسبة المئوية لاجمالي المؤشرات
٩	يدوي	%٣٣.٣٣	٥٨.٣٣ %	%٨٨.٨٩	%٦٣.٣٣	%٦١.٩٠	%٣٣.٣٣	-	-	%٥٥.٥٦	%٥٩.٦٥
١٠	يدوي	%٥٠	٣٣.٣٣ %	%١٠٠	%٦٣.٣٣	%٤٧.٦٢	%٣٣.٣٣	-	-	%٤٤.٤٤	%٥٤.٣٩
١١	يدوي	%٢٥	٤١.٦٧ %	%٤٤.٤٤	%٤٠.٠٠	%٤٧.٦٢	%٥٠	%٧٧.٧٨	-	%٥٥.٥٦	%٤٨.٤٨
١٢	يدوي	%٧٥	%٧٥	%١٠٠	%٩٠.٠٠	%١٠٠	%٨٣.٣٣	%٧٧.٧٨	%٧٥	%٨٧.٥٠	%٨٨.٤٦
١٣	يدوي	%٦٦.٦٧	%٥٠	%١٠٠	%٧٦.٦٧	%٥٢.٣٨	%٦٦.٦٧	-	-	%٥٥.٥٦	%٦٦.٦٧
١٤	يدوي	%٢٥	٥٨.٣٣ %	%١٠٠	%٦٣.٣٣	%٥٢.٣٨	-	%٧٧.٧٨	-	%٦٠.٠٠	%٦١.٦٧
١٥	يدوي	%٤١.٦٧	٥٨.٣٣ %	%١٠٠	%٧٠.٠٠	%٥٧.١٤	%٨٣.٣٣	-	-	%٦٢.٩٦	%٦٦.٦٧
١٦	آلي	%١٠٠	٤١.٦٧ %	%٦٦.٦٧	%٦٩.٧٠	%٩٠.٤٨	%٦٦.٦٧	%٨٣.٣٣	%٨٨.٨٩	%٨٥.٤٢	%٧٩.٠١
١٧	آلي	%١٠٠	٩١.٦٧ %	%٨٨.٨٩	%٩٣.٩٤	%٧١.٤٣	-	%٩١.٦٧	%١٠٠	%٨٣.٣٣	%٨٨.٠٠
١٨	آلي	%١٠٠	٦٦.٦٧ %	%٦٦.٦٧	%٧٨.٧٩	%٤٧.٦٢	-	%٧٥	%٧٧.٧٨	%٦١.٩٠	%٦٩.٣٣
١٩	آلي	%١٠٠	%١٠٠	%٨٨.٨٩	%٩٦.٩٧	%٧١.٤٣	-	%٩١.٦٧	%٨٨.٨٩	%٨٠.٩٥	%٨٨.٠٠
٢٠	آلي	%١٠٠	%٧٥	%٨٨.٨٩	%٨٧.٨٨	%٦٦.٦٧	-	%٩١.٦٧	%٨٨.٨٩	%٧٨.٥٧	%٨٢.٦٧
٢١	آلي	%٨٣.٣٣	٥٨.٣٣ %	%٨٨.٨٩	%٧٥.٧٦	%٦١.٩٠	-	%٧٥	%٧٧.٧٨	%٦٩.٠٥	%٧٢.٠٠
٢٢	آلي	%١٠٠	٣٣.٣٣ %	%٥٥.٥٦	%٦٣.٦٤	%٥٢.٣٨	-	%٧٥	%٧٧.٧٨	%٦٤.٢٩	%٦٤.٠٠
٢٣	آلي	%١٠٠	٦٦.٦٧ %	%٧٧.٧٨	%٨١.٨٢	%٦٦.٦٧	-	%١٠٠	%٨٨.٨٩	%٨٠.٩٥	%٨١.٣٣
٢٤	آلي	%١٠٠	٥٨.٣٣ %	%٧٧.٧٨	%٧٨.٧٩	%٦١.٩٠	-	%٦٦.٦٧	%٧٧.٧٨	%٦٦.٦٧	%٧٢.٠٠
٢٥	آلي	%١٠٠	%١٠٠	%١٠٠	%١٠٠.٠٠	%٦١.٩٠	-	%٩١.٦٧	%١٠٠	%٧٨.٥٧	%٨٨.٠٠
٢٦	آلي	%٥٨.٣٣	%٥٠	%١٠٠	%٦٦.٦٧	%٤٢.٨٦	-	-	%٨٨.٨٩	%٥٦.٦٧	%٦١.٩٠
٢٧	آلي	%٣٣.٣٣	%٥٠	%١٠٠	%٥٧.٥٨	%٤٧.٦٢	-	%٩١.٦٧	-	%٦٣.٦٤	%٦٠.٦١
٢٨	آلي	%٥٠	٥٨.٣٣ %	%١٠٠	%٦٦.٦٧	%٥٢.٣٨	-	%٨٣.٣٣	%١٠٠	%٧١.٤٣	%٦٩.٣٣
٢٩	آلي	%٦٦.٦٧	٥٨.٣٣ %	%١٠٠	%٧٢.٧٣	%٥٢.٣٨	-	%٩١.٦٧	%٨٨.٨٩	%٧١.٤٣	%٧٢.٠٠
٣٠	آلي	%٦٦.٦٧	٨٣.٣٣ %	%١٠٠	%٨١.٨٢	%٥٢.٣٨	-	%٩١.٦٧	%٨٨.٨٩	%٧١.٤٣	%٧٦.٠٠
٣١	آلي	%٨٣.٣٣	٨٣.٣٣ %	%١٠٠	%٨٧.٨٨	%٥٢.٣٨	-	%٨٣.٣٣	%٦٦.٦٧	%٦٤.٢٩	%٧٤.٦٧
٣٢	آلي	%٩١.٦٧	٩١.٦٧ %	%١٠٠	%٩٣.٩٤	%٥٢.٣٨	-	%٧٥	%٧٧.٧٨	%٦٤.٢٩	%٧٧.٣٣
٣٣	يدوي	%٥٠	٣٣.٣٣ %	%١٠٠	% ٦٣.٣٣	%٤٧.٦٢	%٥٠	-	-	%٤٨.١٥	%٥٦.١٤
٣٤	يدوي	%٤١.٦٧	%٢٥	%١٠٠	%٥٦.٦٧	%٥٧.١٤	%٨٣.٣٣	-	-	%٦٢.٩٦	%٥٩.٦٥
٣٥	يدوي	%٤١.٦٧	%٥٠	%١٠٠	%٦٦.٦٧	%٥٢.٣٨	%٨٣.٣٣	-	-	%٥٩.٢٦	%٦٣.١٦
٣٦	يدوي	%٥٠	٦٦.٦٧ %	%١٠٠	%٧٦.٦٧	%٥٢.٣٨	%٨٣.٣٣	-	-	% ٥٩.٢٦	%٦٨.٤٢
٣٧	يدوي	%٥٠	٥٨.٣٣ %	%١٠٠	%٧٣.٣٣	%٥٧.١٤	%٨٣.٣٣	-	-	%٦٢.٩٦	%٦٨.٤٢
٣٨	آلي	%٥٠	٣٣.٣٣ %	%٧٧.٧٨	%٥١.٥٢	%٤٧.٦٢	-	-	%٧٧.٧٨	% ٥٦.٦٧	% ٥٣.٩٧
٣٩	آلي	%٧٥	٦٦.٦٧ %	%٨٨.٨٩	%٧٥.٧٦	%٦١.٩٠	-	%٦٦.٦٧	%٧٧.٧٨	% ٦٦.٦٧	% ٧٠.٦٧
٤٠	آلي	%٥٠	٦٦.٦٧ %	%٨٨.٨٩	%٦٦.٦٧	%٥٧.١٤	-	%٨٣.٣٣	%٨٨.٨٩	% ٧١.٤٣	% ٦٩.٣٣
٤١	آلي	%٧٥	٥٨.٣٣ %	%٧٧.٧٨	%٦٩.٧	%٥٧.١٤	-	%٨٣.٣٣	%٨٨.٨٩	% ٧١.٤٣	%٧٠.٦٧
٤٢	آلي	%٨٣.٣٣	٨٣.٣٣ %	%١٠٠	%٧٨.٧٩	%٥٧.١٤	-	%٩١.٦٧	%٨٨.٨٩	% ٧٣.٨١	% ٧٦
٤٣	آلي	%٧٥	٩١.٦٧ %	%٨٨.٨٩	%٨٤.٨٥	%٧١.٤٣	-	%٩١.٦٧	%٨٨.٨٩	% ٨٠.٩٥	% ٨٢.٦٧

رقم الورقة في الفحص	طريقة التصحيح	البيانات الأساسية	التعليمات	الاخراج العام	النسبة المئوية للمؤشرات الشكلية	مضمون الأسئلة	الأسئلة المقالية	أسئلة الاختيار من متعدد	أسئلة الصواب والخطأ	النسبة المئوية للمؤشرات الكيفية	النسبة المئوية لاجمالي المؤشرات
٤٤	آلي	٢٥%	٥٠%	١٠٠%	٥٤.٥٥%	٥٧.١٤%	-	٩١.٦٧%	٧٧.٧٨%	٧١.٤٣ %	٦٤ %
٤٥	آلي	٥٠%	٨٣.٣٣ %	١٠٠%	٧٥.٧٦%	٩٠.٤٨%	٨٣.٣٣%	٩١.٦٧%	٧٧.٧٨%	٨٧.٥ %	٨٢.٧٢ %
٤٦	آلي	٦٦.٦٧%	٤١.٦٧ %	٦٦.٦٧%	٥٧.٥٨%	٥٧.١٤%	-	٨٣.٣٣%	٧٧.٧٨%	٦٩.٠٥ %	٦٤ %
الاجمالي		٦٣.٧٧%	٦٠.١٤ %	٩١.٧٩%	٧٣.٣١%	٥٧.٢٥%	٦١.١١%	٨٤.٦٧%	٨٤.٢٩%	٦٤.٤١ %	٦٨.٦٨ %

يتضح من جدول (٦) ما يلي:

١. كانت النسبة العامة لتحقق المؤشرات القياسية للورقة الامتحانية (الشكلية والكيفية) معًا أقل من ٧٥% حيث كانت ٦٨.٦٨% وهي نتيجة غير مرضية على وجه العموم في التعبير عن جودة الامتحانات النهائية التي تم تحليلها .
٢. بلغت النسبة العامة لتحقق المؤشرات الشكلية٧٣.٣١% في حين كانت نسبة تحقق المؤشرات الكيفية أقل منها حيث كانت ٦٤.٤١% وهي نتيجة غير مرضية أيضًا في التعبير عن جودة مضمون الامتحانات النهائية التي تم تحليلها. وتتفق هذه النتيجة مع نتائج دراسة كل من (منار الشامي ،سهام الشافعي،٢٠١١؛أماني الغامدي،سماح زغاري،٢٠١٨).
٣. بالنسبة للمؤشرات الشكلية يتضح:

أ- بالنسبة لمؤشر البيانات الأساسية كانت النسبة العامة لتحققه٦٣.٧٧% وقد كان تحقق هذا المؤشر أقل من ٥٠% في ٢٣.٩% من الأوراق الامتحانية وتحقق بنسبة تتراوح ما بين ٥٠ :٧٥ % في ٥٠% من الأوراق وتجاوز الـ٧٥ % في ٦.٥%من الأوراق وبلغ ١٠٠% في ١٩.٦ %من الأوراق ولكن في المجمل تحقق هذا البند من ٥٠% فأكثر في أكثر من ٧٦% من الأوراق الامتحانية ؛ ويمكن تفسير تفاوت النسب في ذلك إلى أولويات أساتذة المواد في ورقة الامتحانات فكثير منهم لا يهتم بوضع شعار الجامعة والكلية ويكتفي بأحدهم وهكذا بالنسبة للقسم والبرنامج فمعظم الأوراق الامتحانية التي تم فحصها لم تتضمن البرنامج ،كما يكتفي البعض بوضع اسم المقرر دون الكود الخاص به ،كما أن بعض الأوراق لم يكن متضمن فيها الفصل الدراسي على الاطلاق وكان التاريخ من العناصر المهملة في معظم الأوراق بالاضافة إلى أن معظم الأوراق تضمنت الزمن والدرجة الكلية للاختبار دون تفصيل الدرجات الخاصة بكل سؤال وهذا يتفق مع آراء الطلاب حيث كانت العبارة الرابعة في مقياس آراء الطلاب (تتضمن الورقة الامتحانية الدرجة الخاصة بكل سؤال وكل مفردة) في الترتيب الحادي عاشر والأخير بنسبة (٣.٤٧) كما وُجدت بعض التعليقات حول رغبتهم في ذلك.

ب- بالنسبة لمؤشر تعليمات الامتحان كانت النسبة العامة لتحققه٦٠.١٤% وقد كان تحقق هذا المؤشر أقل من ٥٠% في ٢١.٧% من الأوراق الامتحانية وتحقق بنسبة تتراوح ما بين ٥٠ :٧٥ % في ٥٨.٧% من الأوراق وتجاوز الـ٧٥ % في ١٥.٢%من الأوراق وبلغ ١٠٠% في ٤.٣% من الأوراق ولكن في المجمل تحقق هذا البند من ٥٠% فأكثر في ٧٨.٢% من الأوراق الامتحانية حيث تضمنت طريقة الاجابة على الامتحان خصوصًا إذا كانت الاجابة في (قالب مصمم للاجابة) سواء تصحيح يدوي أم آلي ،كما قام معظم أساتذة المواد بالتنبيه على طلابهم بقراءة تعليمات الامتحان قبل الشروع في

الحل من حيث طريقة التعبئة الصحيحة والأخطاء التي يجب تجنبها واهتم الكثير منهم بكتابة معلومات تفصيلية عن طريقة الحل في اختباراتهم.

ت- بالنسبة لمؤشر الإخراج العام للاختبار كانت النسبة العامة لتحققه٩١.٧٩% وكان المؤشر الأعلى على الإطلاق في نسبة تحققه وقد بلغ ١٠٠% في أكثر من٦٣% من الامتحانات وهذا يدل حرص أعضاء هيئة التدريس على وضوح الورقة وسهولة قراءتها ومراجعتها للتأكد من خلوها من الأخطاء اللغوية ،كما أن كافة الامتحانات باستثناء اختبار واحد فقط كانت مكتوبة على الكمبيوتر وجيدة الطباعة، وهذا يؤكد ويدعم النتيجة التي تم الحصول عليها في السؤال الأول حيث كانت المفردة(١٠)الخاصة بوضوح الطباعة وسهولة قراءتها ذات الترتيب الأول بين عبارات المحور بنسبة (٤.٣٢).

٤. بالنسبة للمؤشرات الكيفية يتضح:

أ- بالنسبة مضمون الأسئلة كانت النسبة العامة لتحققه٥٧.٢٥% وقد كانت النسبة الأقل على مستوى جميع المؤشرات الفرعية والرئيسية وقد كان تحقق هذا المؤشر أقل من ٥٠% في ٢٨.٣% من الأوراق وتحقق بنسبة تتراوح ما بين ٥٠ :٧٥ % في ٦٣.٠٤% من الأوراق وتجاوز الـ٧٥ % فقط في ٨.٧%من الأوراق حيث كانت أكثر البنود المهملة في هذا البعد هي تنوع الأسئلة بين المقالية الانشائية وبين الموضوعية بأنواعها ، والتوازن بينها وبين الدرجات المخصصة لها كما يتضح من جدول (٧)

جدول (٧)نسب الأنواع المختلفة للأسئلة في الأوراق الامتحانية موضع الدراسة

أنواع الأسئلة المتضمنة في الورقة الامتحانية	عدد الامتحانات التي تتضمن هذه الانواع	نسبة الامتحانات على مستوى الامتحانات المفحوصة
مقالية فقط	١٦	٣٤.٧٨%
اختيار من متعدد+صواب وخطأ	٢١	٤٥.٦٥%
اختيار من متعدد فقط	١	٢.١٧%
صواب وخطأ فقط	٣	٦.٥٢%
مقالية + أحد الأنماط فقط (اما اختيار من متعدد وإما صواب وخطأ)	٢	٤.٣٤%
مقالية + اختيار من متعدد +صواب وخطأ	٣	٦.٥٢%

فكانت نسبة الأوراق الامتحانية التي تضمنت أسئلة مقالية وأنواع مختلفة من الأسئلة فقط ٦.٥٢% ونسبة الأوراق التي تضمنت أسئلة مقالية وأحد الأنواع فقط من الأسئلة الموضوعية حوالي ٤.٣٤% في حين تضمنت معظم الأوراق الامتحانية على أحد الأنواع من الأسئلة فقط إما المقالية ٣٤.٧٨% أو الموضوعية (صورة٨.٦٩% أو صورتين فقط منها وهما الاختيار من متعدد والصواب والخطأ٤٥.٦٥%) وكان السائد أن الامتحانات التي صححت آليًا اكتفت بهذين النوعين فقط من الأسئلة مما يؤدي إلى خلو هذه الامتحانات من الأسئلة التي تقيس مستويات التفكير المختلفة مثل التقويم والابداع حيث أنه توجد محددات للأسئلة الموضوعية في قياس كل مستويات التفكير ،فالابداع يصعب قياسة باختيار أحد البدائل وانما بكتابة الطالب واسترساله وابداء رأيه في الأسئلة التي تسمح له بذلك .

بالاضافة إلى ذلك فمعظم الامتحانات التي صححت يدويا تضمنت أسئلتها المقالية عبارات عامة نمطية تقيس المستويات الدنيا من التفكير فقد تمحورت حول استدعاء الحقائق والنقاط المذكورة نصًا في الكتاب دون اعطاء الحرية في ابداء الرأي واستكشاف المهارات العليا ؛كما أن بعض الامتحانات المقالية لم تغطي

كل محتوى المقرر وذلك لأنها تتطلب الكثير من الوقت وتتفق هذه النتيجة مع نتائج دراسة (عبد الكريم السوداني،٢٠٠٦).

أما الامتحانات الموضوعية والتي صححت آليًا فكان معظمها يتضمن على عدد أسئلة غير ملائمًا لزمن الاجابة على الامتحان حيث تضمنت كثير من الامتحانات على (٨٠) سؤال وربما أقل في بعض الامتحانات في حين كان الزمن الكلي ساعتان وهذا غير مناسب تمامًا فاذا كان متوسط الاجابة على السؤال دقيقة واحدة فيجب أن يحتوي الامتحان على ١٢٠ سؤالًاعلى الأقل.

ومن ناحية أخرى استوفت جميع الامتحانات تقريبًا مؤشر "ترتبط جميع أسئلة الامتحان بنواتج التعلم المستهدفة (ILOs) فلم تخرج أسئلة الامتحانات عن الموضوعات المقررة بتوصيفات المقررات وهو ما اتفق مع رأي الطلاب حيث جاءت العبارة رقم (١٧) في مقياس آراء الطلاب"ترتبط جميع أسئلةالامتحان بالمقرر الذي تم تدريسه" الأعلى بين عبارات المحور بنسبة (٣.٩١) وهذا يعني عدم اعتراضهم على ذلك ولكن في الوقت نفسه أقر بعضهم أن المفاجئات في أسئلة الامتحانات تكون أسئلة من المقرر الذي تم تدريسه ولكن غير متوقعه ؛ولم يتم التأكيد على ادراجها ضمن أسئلة الامتحان من قبل أساتذه المواد.

ب-بالنسبة للمؤشرات الثلاثة " جودة الأسئلة المقالية ،وأسئلة الاختيار من متعدد وأسئلة الصواب والخطأ "فكانت النسبة العامة لتحققهم (٦١.١١% ، ٨٤.٦٧% ، ٨٤.٢٩%) على الترتيب وبذلك فكان مؤشر جودة الأسئلة المقالية هو الأقل بالنسبة للأنواع الأخرى من الأسئلة الموضوعية و شملت هذه النسبة كافة الأوراق الامتحانية التي تضمنت أسئلة مقالية سواء كانت منفردة أم معها صورة أو أكثر من الأسئلة الموضوعية وقد شكّلت ٤٥.٦٥% من الأوراق و كانت صياغة كثير من الأسئلة دقيقة وواضحة المعنى ومحددة للعناصر المطلوبة في الاجابة ولكن كان البعض منها أكثر عمومية وتنقصه الدقة في الصياغة ولكن كان من الشائع ندرة الأسئلة ذات النهايات المفتوحة التي تسمح بالابداع والابتكار.ويتفق ذلك مع نتائج دراسة كل من(أماني الغامدي،سماح زغاري،٢٠١٨؛ حمد الكلثم ،٢٠١٥؛منال حسن ،٢٠١٣).

ت-أما بالنسبة لمؤشر جودة أسئلة الاختيار من متعدد وأسئلة الصواب والخطأ فقد تقاربت النسب حيث التزمت أغلب الامتحانات "بأن لا يقل عدد البدائل عن أربعة" كما اتضح "التوزيع العشوائي للاجابة الصحيحة سواء بالنسبة لأسئلة الاختيار من متعدد أو أسئلة الصواب والخطأ"،كما تقاربت العبارات الصحيحة والخاطئة في الطول ،أما بالنسبة لمؤشر جودة الأسئلة والبدائل والابتعاد عن العبارات البديهية المنقولة نصًا من الكتاب " فقد لوحظ في الامتحانات التي صُحِّحت آليًا (٢٦ اختبار) وتمثل ٥٦.٥٢% من الامتحانات وجود بعض البدائل الواضحة التي وصلت الاستجابة الصحيحة عليها ١٠٠% مما يدل على وضوحها كما وجدت بعض العبارات والتعريفات المنقولة نصًا من الكتاب ويتفق هذا مع نتائج رأي الطلاب حيث وافق الطلاب على كل من العبارة (٢٥)"تتضمن الامتحانات بعض العبارات المنقولة نصًا من الكتاب وعلى العبارة (٢٣)"تتضمن الامتحانات بعض الأسئلة والبدائل البديهية"ويؤكد ذلك ما أشار إليه (علي صلاح حسن ،٢٠١٨) عن نتائج دراسته حول حاجة أعضاء هيئة التدريس للاهتمام بالخطوات العلمية لبناء الامتحانات ومراعاة شروط وقواعد بناء الاختبارت الموضوعية.

بالاضافة إلى ذلك فقد وجدت في بعض الامتحانات قليل من الأسئلة تتضمن في نصوصها اجابة لأسئلة أخرى ويتفق هذا مع نتائج دراسة (محمد عبد المنعم ،٢٠٠٩) ؛ وهذا يعتبر من أخطاء إعداد الأسئلة حيث يساعد الطلاب الذين يتمتعون بمستوى عال من الحكمة الاختبارية في الاجابة عن تلك الأسئلة بسهولة

مما يفقدها القدرة على التمييز بين الطلاب ويؤدي إلى نسب غير موضوعية في اجابات الطلاب. (Dodeen,2003; Dolly & Willams, 1986;Milleman, et al, 1965; Sarnacki, 1979)

٣- **تحليل نتائج السؤال الثالث ومناقشتها :** للإجابة عن السؤال الثالث الذي ينص على

"ما مؤشرات الصلاحية السيكومترية للامتحانات النهائية المصححة آليًّا ؟" تم القيام بالتالي:

أ- تحديد عينة من الأوراق الامتحانية التي صححت آليًّا (حيث تتوفر كافة البيانات الخاصة بالطلاب وبالمفردات والاحصائيات الخاصة بكل مفردة وبالامتحان بأكمله) وبلغ عددها (١٤)اختبار منهم (٨) من تخصص الباحثة .

ب- استخراج التقارير الفنية الخاصة بالخصائص السيكومترية للاختبار ولمفرداته من خلال برنامج OMR للتصحيح الآلي وهو تقرير (100) Test Analysis; (203) Item Analysis Graph; ((310) Test Statistics

ت- استخراج الخصائص السيكومترية لكل اختبار ومعاملات الثبات ونسب المفردات غير المُمَيِّزة بين الطلاب ويوضح جدول (٨) النتائج التي تم التوصل إليها:

جدول (٨)الخصائص السيكومترية لعينة الامتحانات المصحّحة آليًا

مسلسل	رقم الورقة في قائمة الفحص	النسبة العليا	النسبة الدنيا	النسبة المئوية للمتوسط	مدى الدرجات	نسبة الأسئلة ذات التمييز الضعيف (point biserial < 0.2)	الانحراف المعياري	التباين	الثبات بمعادلة كيودر ٢٠	الثبات بمعادلة كيودر ٢١	ألفا كرونباخ	الخطأ المعياري للقياس
١	٢٥	%٩٢.٨٦	%٤٢.٨٦	%٧٣.١٩	٣٥	%٤٢.٨٦	٥.٩٧	٣٥.٦٤	٠.٧٤	٠.٦٢	٠.٧٤متوسط	٣.٠٧
٢	٢٤	%٩١.٠٠	%٥٠.٠٠	%٧٨.٩٨	٤١	%٦٦	٥.٢٥	٢٧.٥٦	٠.٦٤	٠.٤٠	٠.٦٤ متوسط	٣.١٤
٣	١٧	%٩٤.١٧	%٤٧.٥٠	%٧١.٦٢	٥٦	%٤٢.٥	١٠.١٢	١٠٢.٤٧	٠.٨٥	٠.٧٧	٠.٨٥مرتفع	٣.٩٥
٤	٢٠	%٩٠.٠٠	%٥٨.٧٥	%٧٢.٦٩	٢٥	%٦٧.٥	٤٠٤٤	١٩.٧١	٠.٥٧	٠.٢٠	٠.٥٧منخفض	٢.٩١
٥	٢١	%٩٥.٠٠	%٦٢.٥٠	%٨٣.٠٢	٢٦	%٦٧.٥	٤.٣٥	١٨.٩٦	٠.٥٧	٠.٤١	٠.٥٧منخفض	٢.٨٥
٦	٣٢	%١٠٠	%٦١.٢٥	%٨٩.٧٧	٣١	%٤٥	٥.٤٢	٢٩.٣٣	٠.٨٠	٠.٧٦	٠.٨٠ مرتفع	٢.٤٢
٧	٢٩	%١٠٠	%٤٥.٠٠	%٩١.٣٤	٤٤	%٣٢.٥	٤.٦٢	٢١.٣١	٠.٧٦	٠.٧١	٠.٧٦متوسط	٢.٢٥
٨	٢٢	%١٠٠	%٧.٠٦	%٨٠.٣٧	٧٩	%٢٧.٠٦	٨.٧٥	٧٦.٥٨	٠.٨٧	٠.٨٣	٠.٨٧مرتفع	٣.١٦
٩	٢٣	%٨٤.٦٢	%٤٣.٠٨	%٦٩.٧٦	٢٧	%٤١.٥٤	٥.٦٧	٣٢.١٧	٠.٦٧	٠.٥٨	٠.٦٧متوسط	٣.٢٧
١٠	٢٦	%٩١.٦٧	%٠.٠٠	%٧٨.١٤	٥٥	%١٨.٣٣	٤.٧٤	٢٢.٤٥	٠.٧٥	٠.٥٥	٠.٧٥متوسط	٢.٣٥
١١	٢٧	%٩٨.٣٣	%٠.٠٠	%٨٣.٥٩	٥٩	%٣.٣٣	٨.٨٠	٧٧.٤٢	٠.٩٣	٠.٩١	٠.٩٣ مرتفع ج	٢.٤١
١٢	٣١	%٩٤.٢٩	%٥١.٤٣	%٧٥.٣٢	٣٠	%٥٧.١٤	٤.٨٣	٢٣.٢٩	٠.٦٢	٠.٤٥	٠.٦٢ متوسط	٢.٩٧
١٣	٣٢	%١٠٠	%٦٧.٥٠	%٩٠.٧٨	٢٦	%٥٢.٥	٥.٣٣	٢٨.٤٥	٠.٨١	٠.٧٧	٠.٨١مرتفع	٢.٣١
١٤	٣٨	%١٠٠	%٥١.٤٣	%٨٨.١١	٣٤	%٤٢.٨٦	٤.٩٩	٢٤.٩٣	٠.٧٧	٠.٧٢	٠.٧٧متوسط	٢.٤١

يتضح من جدول (٨) أن:

١. تجاوزت النسبة العليا في عينة أوراق الاجابة المصححة آليًّا ٩٥% في ٥٠% من الامتحانات ووصلت إلى ١٠٠% في ٣٥٠٧% منها ،بينما وصلت النسبة الدنيا إلى صفر% في بعض الامتحانات بينما كانت ٥٠% فأعلى في ٥٠% منها وهذا قد يكون مؤشر على سهولة الامتحانات والنسب العالية للنجاح ولكن لا يمكن الاعتماد على النسبة العليا والدنيا في التفسير؛ فالدرجة التي يحصل عليها الفرد في الاختبار غالبًا لا تكون هي الدرجة الحقيقية إنما تتكون من الدرجة الحقيقية

مضافًا إليها درجة الأخطاء العشوائية ، فالدرجة الملاحظة للفرد في الاختبار لاتساوى درجته الحقيقية إلا إذا كانت درجة الخطأ صفرا، ولكنها تزيد عن درجاته الحقيقية إذا كانت درجة الخطأ موجبة (مثلا في حالات التخمين) ، وتقل عنها إذا كانت درجة الخطأ سالبة (مثلا في حالة المفردات أو التعليمات الغامضة للامتحان) ،ولكن هذه الأخطاء الموجبة والسالبة يلاشى بعضها البعض الآخر على المدي البعيد استنادا إلى قوانين الصدفة فيصبح متوسط درجات الخطأ صفرًا، ولذلك فإن متوسط الدرجات الملاحظة لمجموعة الأفراد الذين يطبق عليهم الاختبار يُعد أفضل تقدير لمتوسط درجاتهم الحقيقية في الاختبار(صلاح علام ،٢٠٠٠) .

في الوقت نفسه تراوحت النسبة المئوية للمتوسط ما بين ٦٩.٧٦% (تقدير مقبول) إلى ٩١.٣٤%(تقدير ممتاز) وكانت أعلى من ٨٠%(تقدير جيد جدًا) في ٥٠ % من الامتحانات وبذلك تتأكد سهولة الامتحانات والنسب العالية للنجاح وتؤكد هذه النتيجة ما توصلت إليه دراسة (فاتن موسى،احسان حجازي،٢٠٢٠) في تفضيل الطلاب للامتحانات المصححة آليًا وارتفاع درجاتهم فيها .

٢. أدت الزيادة في الدرجات المتطرفة إلى اتساع مدى الدرجات الذي تراوح من ٢٥ إلى ٧٩ في عينة الامتحانات المفحوصة ولكن من سلبيات المدى أنه إذا كانت هناك درجة واحدة عالية جدًا أو منخفضة جدًا فإن ذلك يؤثر تأثيرًا كبيرًا ويعطي انطباعًا خاطئًا عن حقيقة التشتت في هذه المجموعة ،ولذلك يعتبر الانحراف المعياري دليل أفضل على التشتت في أية مجموعة لا يتأثر كثيرًا بالدرجات المتطرفة الوحيدة فكل درجة في التوزيع تُسهم في قيمة الانحراف المعياري، ففي التوزيعات التي تتجمع فيها الدرجات حول المتوسط يكون الانحراف المعياري صغيرًا أما في التوزيعات التي تنتشر فيها الدرجات كثيرًا على جانبي المتوسط فإن الانحراف المعياري يكون كبيرًا (ليونا تايلر،١٩٨٨).ويتضح من جدول (٨)تفاوت قيم الانحراف المعياري بين الامتحانات حيث امتدت من (٤.٣٥) إلى (١٠.١٢) مما يؤكد انتشار درجات الامتحانات وتوزعها على المستويات المختلفة.

٣. **بالنسبة لجودة الأسئلة** يتضح من جدول (٨) أن نسبة الأسئلة غير المُميّزة التي لم تصل إلى معامل تمييز مناسب كانت عالية في معظم الامتحانات حيث وصلت في بعض الامتحانات إلى ٦٧.٥% مما يعطي مؤشر عن عدم الدقة في وضع الأسئلة وعدم فاعلية المشتتات حيث أن الغرض من تحديد معامل التمييز لكل سؤال هو معرفة قدرة السؤال على التمييز بين المجموعتين العليا والدنيا فالسؤال الذي تكون درجة تمييزه عاليه تعني أن نسبة من أجابوا عليه إجابة صحيحة من أفراد المجموعة العليا أكبر من نسبة من أجابوا عليه اجابة صحيحة من أفراد المجموعة الدنيا ويوضح جدول (٩) دلالة قيم معاملات التمييز المختلفة.

جدول (٩)دلالات قيم معامل التمييز لمفردات الاختبار
في:(صلاح مراد،أمين سليمان،٢٠٠٥)

قيم معامل التمييز	دلالة معامل التمييز
أكبر من +٠.٣	الحد الأدنى لمؤشر التمييز هو +٠.٣ على الأقل ٠.٤ فأكثر تدل على قوة تمييز عالية بين أفراد المجموعتين العليا والدنيا
٠.٢ -٠.٣	قوة تمييز منخفضة
٠.١ -٠.٢	معامل تمييز غير مقبول ويحتاج السؤال إلى التعديل
قيم سالبة	يحذف السؤال لأنه مضلل ،نسبة من أجابوا عليه بالصواب من المجموعة الدنيا أكبر من أقرانهم في المجموعة العليا.

وبزيادة نسب الأسئلة التي يقل معامل تمييزها عن ٠.٢ بالرجوع إلى دلالة القيم في جدول (٩) تزيد الشكوك حول نتائج تلك الامتحانات وقد يكون ذلك مؤشرًا على معاملات صعوبة غير مناسبة للمفردات حيث أن القيمة المثلى لمعامل الصعوبة تساوي٠.٥ كما تذكر (سلمى عبد الرحيم،٢٠٠١) في حين كثرت الأسئلة التي قل معامل صعوبتها عن ٠.٥ وبالتالي زاد معامل سهولتها حتى وصلت نسبة الاجابة على كثير منها ١٠٠% .

٤. **بالنسبة لمعاملات الثبات** يتضح من جدول (٨) معاملات الثبات المحسوبة عن طريق تحليل البنية الداخلية للاختبارInternal Structure للتعرف على مدى تجانس مفردات الاختبار Tests Items homogeneity من خلال معامل كيودر وريتشاردسون kuder-Richardson الصيغة (٢٠) KR20 والصيغة (٢١)KR21 ومعامل ألفا كرونباكCronbach وتعتبر طريقة تحليل البنية الداخلية للاختبار كما يرى (صلاح علّام،٢٠٠٠) مناسبة لاختبارات التحصيل الدراسي حيث يُستدل من درجات الطلاب في عينة من المفردات التي تشتمل عليها درجة تمكنهم من معارف ومهارات معينة لذلك فإنه من المناسب في هذه الحالة تقدير ثبات الدرجات باستخدام أي من طرق الاتساق الداخلي.

و تهدف طريقة كيودر إلى التوصل إلى قيمة تقديرية لمعامل ثبات الاختبارات التي تكون درجاتها ثنائية إما واحد صحيح أو صفر مثل مفردات الاختيار من متعدد أو مفردات الصواب والخطأ وتتغلب هذه الطريقة على مشكلة تعدد طرق التجزئة النصفية للاختبار بأن توجد قيمة تقديرية لمتوسط قيم معامل الارتباط بين كل من نصفي الاختبار لجميع طرق التجزئة الممكنة دون القيام بهذه التجزئة فعلًا ،وتؤكد هذه الطريقة العلاقات القائمة بين المفردات التي يشتمل عليها الاختبار أكثر من تأكيدها اتساق درجات الاختبار بمرور الزمن أو تكافؤها على الرغم من اختلاف صيغ الاختبار وقد اتضح أن:

- معاملات الثبات المحسوبة بالصيغة (٢١) كانت دائما أقل من معاملات الثبات المحسوبة بالصيغة (٢٠) ولكن يمكن الاعتماد على معامل الثبات المحسوب بالصيغة (٢٠) حيث يتطلب استخدام الصيغة (٢١) توافر شرط ضروري وهو تساوي جميع مفردات الاختبارات والمقاييس في درجة صعوبتها أو على الأقل يتراوح متوسط درجة صعوبة جميع المفردات ٠.٥ ولكن نظرًا لصعوبة تحقق هذا الشرط في كثير من الاختبارات او المقاييس فإن الصيغة (٢٠) ربما تكون أكثر ملاءمة لهذه الاختبارات بالاضافة إلى أن الصيغة (٢١)KR21 تعطي قيمًا تقديرية لمعامل التجانس أقل من حقيقتها وربما أقل من أي أسلوب آخر من أساليب التجزئة النصفية وبخاصة إذا استخدمت دون مراعاة للشروط التي تتطلبها (سوسن مجيد ،٢٠١٤)

- كما اتضح تساوي معاملات الثبات بمعادلة كيودر (٢٠) وبطريقة ألفا كرونباخ حيث اشتق كرونباخ معامل ألفا من الصيغة (٢٠) لكيودر حيث أن المعادلتين متماثلتين ولكن تم التعبير عن تباين الدرجات الكلية للاختبار في كلا المعادلتين بصيغتين مختلفتين.
- تراوح معامل الثبات بطريقة ألفا من (٠.٥٧)إلى(٠.٩٣) بمتوسط (٠.٧٤ وهي قيمة متوسطة) و**ربما تكون هذه القيم المتوسطة لأن** معامل ألفا يمثل متوسط قيم المعاملات الناتجة عن تجزئة الاختبار إلى أجزاء بطرق مختلفة وبذلك فهو يمثل معامل الارتباط بين أي جزئين من أجزاء الاختبار(امطانيوس مخائيل،٢٠١٦؛ صفوت فرج،٢٠٠٧)، كما أنه يتأثر بطول الاختبار ومدى تجانس مفرداته ويعطي الحد الأدنى للقيمة التقديرية لمعامل ثبات درجات الاختبار ؛أي أن قيمة معامل الثبات بعامة لا تقل عن قيمة معامل ألفا فإذا كانت قيمة معامل ألفا مرتفعة فإن هذا يدل بالفعل على ثبات درجات الاختبار أما إذا كانت منخفضة فربما يدل ذلك على أن الثبات يمكن أن تكون قيمته أكبر من ذلك باستخدام طرق أخرى (صلاح علام ،٢٠٠٠).

ويمكن القول أن هذه القيم المتوسطة لمعاملات الثبات متوقعة **لأن الاختبارات موضع الفحص غير مقننة** ؛ فالاختبارات المقننة Standardized tests :تقيس أهداف مشتركة عامة لمجموعة من المدارس وتتناول بالقياس أجزاء كبيرة من المعلومات والمهارات السابق دراستها وأسئلتها تم تجربتها ومراجعتها ومعالجتها احصائيًا بغرض حساب معاملات السهولة والصعوبة والتمييز وفعالية المشتتات ،والتخلص من أثر التخمين وتزود بمعايير الأداء لمجموعات مختلفة من الأفراد على نطاق واسع؛أماالاختبارات غير المقننة Examination فتقيس أهدافًا خاصة بالفصل أو المدرسة وتقيس معلومات ومهارات معينة ومحددة ،ونادرًا ما يتم تجربة أسئلتها أو مراجعتها أو تحليلها احصائيًا ،وقد تزود بمعايير الأداء على نطاق الفصل أو المدرسة فقط أي التطبيق على عينة محدودة(صلاح مراد،أمين سليمان ،٢٠٠٥).

وتشير (سوسن مجيد ،٢٠١٤) أن معاملات الثبات المقبولة تختلف من اختبار لآخر وذلك حسب الغرض من الاختبار ودقة القرار المترتب عليه فالاختبارات المقننة تتطلب معاملات ثبات لا تقل عن (٠.٨٠) وذلك عندما تكون القرارات على مستوى الأفراد أما على مستوى الجماعات فيقبل معامل ثبات (٠.٦٥) فقرارات الانتقاء تختلف عن قرارات التصنيف في تأثيرها على مستقبل الفرد ؛ومعاملات ثبات الاختبارات التحصيلية المقننة يجب ألا تقل معاملات ثباتها عن (٠.٨٥) بينما يمكن أن تقل عن ذلك في اختبارات الشخصية ؛كما أن الاختبارات التحصيلية التي لا يراعى في تطبيقها الخطوات التي يتم مراعاتها في الاختبارات المقننة معظمها تتراوح معاملات ثباتها بين (٠.٢٠ - ٠.٤٠) ونادرا ما تصل إلى (٠.٦٠) وهذا ما تحقق في الدراسة الحالية *ولكن ينبغي الحذر من خطورة عدم استقرار نتائج الامتحانات لأن هذا يجعل مسألة تعميمها على عينات مختلفة أمرًا غير عادًا خصوصًا عند الاختيار والمفاضلة بين الطلاب في دفعات مختلفة.*

٥. **بالنسبة للخطأ المعياري** الموضح بالعمود الأخير بجدول (٨) ويعتبر عاملا أساسيًا في تقرير نتائج الاختبارات والمقاييس وتفسيرها كدليل على مقدار عدم الدقة في تفسير الدرجات واتخاذ القرارات بالنسبة للأفراد (ليونا تايلر،١٩٨٨) .

ويشير (صلاح علام ،٢٠٠٠) إلى أن الخطأ المعياري للقياس مرتبطًا ارتباطًا وثيقًا بمفهوم الثبات كما أنه لا يتأثر كثيرًا بتشتت الدرجات على عكس الثبات الذي يعتمد على التشتت ؛لذلك تهتم أدلة الاختبارات

بتحديد القيم التقديرية لهذا الخطأ لكي يتمكن مستخدم الاختبار من تفسير الدرجات تفسيرًا احتماليًّا؛ فإذا كانت الدرجات الملاحظة خالية من الخطأ فان مقدار تباين الخطأ يساوى صفرًا، ويصبح مقدار تباين الدرجات الملاحظة مساويا مقدار تباين الدرجات الحقيقية، أي تكون الدرجات متسقة اتساقا تامًا ؛وبحساب الانحراف المعياري لدرجات الخطأ ينتج الخطأ المعياري للقياسي Standard error of measurement , فإذا كان مقداره صفرًا يكون ثبات الاختبار صحيحًا ١٠٠% ، ولكن عادة في القياس التربوي والنفسي يوجد قدرًا معينًا من الخطأ المعياري مما يؤدي إلى خفض ثبات درجات الاختبارات بقيم متفاوتة بحسب مقدار هذا الخطأ (سوسن مجيد ،٢٠١٤)

فإذا حصل فرد في اختبار ما على درجة معينة فإن هذه الدرجة ليست نقطة محددة وإنما يمكن أن تتذبذب في مدى يتراوح بين قيم معينة على متصل الدرجات تحددها القيمة التقديرية للخطأ المعياري للقياس ودرجة الثقة المطلوبة فإذا حصل الفرد في اختبار معين على ١٠٠ درجة وكان الخطأ المعياري لهذا الاختبار ٥ فتكون درجة الفرد ما بين ٩٥-١٠٥ وهذا نتيجة نقص ثبات الاختبار.(عباس عوض ١٩٩٨،)

ويلاحظ من جدول (٨) أن الخطأ المعياري لعينة الامتحانات تراوح ما بين (٢.٢٥) إلى (٣.٩٥) بمتوسط (٢.٨٢) وهذا يُعني في حالة الاختبار ذو الخطأ المعياري (٣.٩٥) أنّ هناك احتمال بمقدار ٣/٢ بأن الدجة الحقيقية للفرد تقع في مدى (٣.٩٥)نقطة من الدرجة التي حصل عليها فعلًا فإذا كان الفرد(أ) حصل على درجة ١٠٠ فإنه يمكن القول بدرجة من الثقة أن "درجته الحقيقية " خالصة من تأثير عوامل الصدفة تقع بين ١٠٣.٩٥ ، ٩٦.٠٥ وهذا يُبنى على حقيقة أنه في أي توزيع اعتدالي فإن ثلثي الحالات٦٨% تقريبًا تقع في مدى انحراف معياري واحد من المتوسط وتكون الثقة في هذا الاستدلال ٦٨% تقريبا ولكن مازال هناك احتمال بمقدار الثلث بأن"درجته الحقيقية " قد تكون أعلى أو أقل من ذلك ويمكن القول أنه يمكن الاستدلال بدرجة ثقة ٩٥%تقريبًا على أن درجته الحقيقية سوف تقع بين ١٠٧.٩-٩٢.١ (ذلك لأنه في التوزيع الاعتدالي يقع ٩٥% من الحالات في مدى انحرافين معياريين من المتوسط.

ويمكن اعزاء قيم الخطأ المعياري ومعاملات الثبات المتوسطة لعينة الامتحانات موضع الدراسة لعدم جودة مفرداتها وانخفاض معاملات تمييزها مما يُنبأ بحاجة أعضاء هيئة التدريس لكثير من التدريب على اعداد أسئلة امتحاناتهم في ضوء معايير الجودة والمواصفات القياسية لها وكذلك تقنينها والتأكد من ثباتها قبل التطبيق.

٤-تحليل نتائج السؤال الرابع ومناقشتها : للإجابة عن السؤال الرابع الذي ينص على" إلى أي مدى تتحقق (نواتج التعلم المستهدفة من المقررات) من خلال نتائج الامتحانات النهائية المصحَّحَة آليًّا بالكلية ؟

تم القيام بالتالي:

١. تحديد عينة أوراق الاجابة على الامتحانات النهائية الموضوعية بشرط أن:
   a. تكون صُحّحَت آليًّا من خلال وحدة القياس والتقويم (حيث تتوفر درجات الطلاب بالتفصيل على كل مفردة من مفردات الامتحان) .
   b. يتوفر فيها توصيف محدّث للمقرر بوحدة ضمان الجودة بالكلية.

٢. تحديد نواتج التعلم المستهدفة ILOs من كل مقرر على حدة والتي استُهدف قياسها من خلال أسئلة الامتحانات علمًا بأنه توجد نواتج تنتمي للمهارات المهنية والمنقولة غير مستهدف قياسها في الأوراق الامتحانية لأن تحتاج لأدوات تقييم تناسب طبيعتها .

٣. تحديد مفردات الأسئلة التي تقيس كل ناتج بمساعدة أستاذ المقرر؛وساعد في ذلك أن معظم تلك المقررات من تخصص الباحثة.

٤. استخراج التقارير الفنية الخاصة بنواتج تعلم الطلاب من خلال برنامج OMR للتصحيح الآلي وهو تقرير (101)Student Statistics; (103)Class Frequency Distribution; (111)Class Learning Objectives;(310) Test Statistics لاستخراج نسب تحقق نواتج التعلم بكل مقرر ويوضح جدول (١٠) النتائج التي تم التوصل إليها:

جدول (١٠) نسب تحقق نواتج التعلم المستهدفة بالأوراق الامتحانية

المسلسل	رقم الورقة في الفحص	عدد الطلاب	عدد المفردات	عدد الأهداف في التوصيف	عدد الأسئلة لكل هدف من-الى	تحقق الأهداف تراوح من - الى	متوسط اجمالي الأهداف	نسبة الحاصلون علىA	نسبة الحاصلون علىB	نسبة الحاصلون علىC	نسبة الحاصلون علىD	نسبة الراسبون E
١	٢٥	٦٨٨	٧٠	١٠	٥-٨	٨٧.٧٩-٤٩.٨٨	%٧٣.١٩	١.٠٢	٢٣.٦٩	٤٦.٥١	٢٢.٦٧	٦.١٠
٢	٢٤	٦٨٦	١٠٠	١٠	٨-١٠	٩٠.٤٨-٦٧.٨٤	%٧٨.٩٨	٠.٨٧	٤٩.٤٢	٤٤.٦١	٤.٨١	٠.٢٩
٣	١٧	٦٤٢	١٢٠	١٠	١٢	٨٧.١٦-٥٣.٦٢	%٧١.٦٢	١.٧١	١٥.٢٦	٤٤.٥٥	٢٩.٩١	٨.٥٧
٤	٢٠	٣٣٧	٨٠	١٩	٤-٥	٩٥.٩٦-٢٤.٤٨	%٧٢.٦٩	٠.٣٠	١٢.٤٦	٦٠.٥٣	٢٦.١١	٠.٥٩
٥	٢١	٥٢٠	٨٠	١٩	٤-٥	٩٥.٩١-٦٤.٦٦	%٨٣.٠٢	١١.٥٤	٦٤.٠٤	٢٣.٢٧	١.١٥	٠.٠٠
٦	٣٢	٩٤٣	٨٠	٨	٧-١١	٩٤.٩٠-٨٤.٢٤	%٨٩.٧٧	٦٠.٤٥	٣١.٣٩	٦.٧٩	١.٣٨	٠.٠٠
٧	٢٩	٦٦٢	٨٠	١٦	٤-٦	٩٧.٤٦-٨٠.٦٦	%٩١.٣٤	٧١.٣٠	٢٥.٣٨	٢.٨٧	٠.١٥	٠.٣٠
٨	٢٢	١١٤١	٨٥	١٥	٥-٧	٩٠.٢٩-٦٨.٠٥	%٨٠.٣٧	١٩.٢١	٣٥.٩٨	٢٨.٥٦	١٤.٠٦	٢.١٨
٩	٢٣	١٣٣	٦٥	١٥	٣-٥	٩٦.٩٩-٥٠.٨٨	%٦٩.٧٦	٠.٠٠	١٢.٠٣	٤٣.٦١	٣٠.٨٣	١٣.٥٣
١٠	٢٦	٧٠١	٦٠	٨	٥-٨	٩٢.٠٣-٦٨.٨٧	%٧٨.١٤	٠.٨٦	٤٩.٢٢	٤٣.٧٩	٥.٢٨	٠.٨٦
١١	٢٧	٧٠١	٦٠	٢٠	٢-٤	٩٢.٨٢-٥٣.٨٨	%٨٣.٥٩	٤٧.٠٨	٢٥.٦٨	١٣.٦٩	٧.٨٥	٥.٧١
١٢	٣١	٦٨٧	٧٠	١٠	٥-٨	٩١.٧٧-٥٠.٠٣	%٧٥.٣٢	١.٧٥	٢٨.٠٩	٥١.٨٢	١٦.٤٥	١.٨٩
١٣	٣٢	٧٩٢	٨٠	٨	٨-١٢	٩٥.٥٢-٨٥.٣٤	%٩٠.٧٨	٦٥.٤٠	٢٧.٧٨	٦.٥٧	٠.٢٥	٠.٠٠
١٤	٣٨	٥٨٥	٧٠	١٠	٥-٨	٩٥.٨٧-٧٧.٩٥	%٨٨.١١	٤٨.٣٨	٣٩.٨٣	١٠.٤٣	١.٢٠	٠.١٧
المتوسط العام		٦٥٨.٤٣	٧٨.٥٧	١٢.٧١	٥.٥-٧.٥	٩٣.٢١-٦٢.٨٨	%٨٠.٤٨	٢٣.٥٦	٣١.٤٥	٣٠.٥٤	١١.٥٨	٢.٨٧

يتضح من جدول (١٠) ما يلي:

١. تراوح عدد مفردات عينة الامتحانات ما بين ٦٥ إلى ١٢٠ مفردة بمتوسط ٧٨.٥٧ مفردة علمًا بأن هذه المفردات من النوع الموضوعي (اختيار من متعدد ،الصواب والخطأ،التوصيل) وهذا مؤشر على عدم ملائمة عدد الأسئلة لزمن الامتحانات فكافة الامتحانات المفحوصة لم يقل الزمن المخصص لها عن ساعتين وإذا كان متوسط اجابة المفردة دقيقة واحدة فينبغي أن لا يقل عدد الأسئلة عن ١٢٠ مفردة .

٢. تراوح عدد نواتج التعلم المستهدف قياسها من خلال تلك الامتحانات ما بين ٨ إلى ٢٠ هدف بمتوسط ١٢.٧١ هدف كما تراوح متوسط عدد المفردات التي تستهدف كل هدف على حدة ما بين ٥.٥ إلى ٧.٥ مفردة؛ مع الاشارة إلى أنّ جميع المفردات موضوعية مصححة آليًا وربما تكون غير كافية وحدها لقياس مدى تحقق نواتج التعلم ،أضف على ذلك فقد تضمنت توصيفات المقررات وجود كثير من نواتج التعلم حول تنمية مهارات عليا وخُصص للتحقق من بعضها مفردات من الأسئلة الموضوعية وهذا يتعارض مع فكرة وجود محددات للأسئلة الموضوعية في قياس تلك المهارات

وكان ينبغي أن يخصص لها مفردات مقالية مثلًا تسمح بذلك. ; (Boulton-Lewis, 1995 Masters,et.al, 2001;Oermann, 1999)

٣. يتضح من العمود السابع بجدول (١٠) تفاوت النسب المئوية لتحقق نواتج التعلم على مستوى الامتحانات المفحوصة وكان أقل مدي لهذه النسب يساوي ١٠.١٨ % في أحد الامتحانات بينما كان هذا المدى كبيرًا في بعض الامتحانات ففي أحد الامتحانات كانت النسبة المئوية لأقل الأهداف تحققًا تساوي ٢٤.٤٨% بينما كانت النسبة المئوية لأكثر الأهداف تحققًا تساوي ٩٥.٩٦ % بفارق ٧١.٤٨% .

٤. تراوحت النسبة المئوية "لمتوسط تحقق اجمالي الأهداف" في الامتحانات المفحوصة ما بين ٦٩.٧٦% إلى٩٠.٧٨% بنسبة ٨٠.٤٨ % على المستوى العام للاختبارات وهي نسبة مقبولة .

٥. على المستوى الاجمالي لكافة الامتحانات كان متوسط نسب الطلاب في المستويات المختلفة (ممتاز،جيد جدًا،جيد،مقبول،راسب) على الترتيب تساوى (٢٣.٥٦ ،٣١.٤٥ ،٣٠.٥٤ ،١١.٥٨ ، ٢.٨٧) ونظرًا لأن هذه النسب على المستوى الاجمالي لاختبارات مختلفة وعينات مختلفة من الطلاب ليست على عينة محددة أو اختبار بعينه أو خاصة بأحد الأعضاء على وجه التحديد فيمكن القول أنها لا تحقق التوزيع الطبيعي الاعتدالى الذي يفترض أن تكون أغلبية الطلاب في المستوى المتوسط والأقلية عند الطرفين حيث زادت نسبة الطلاب الحاصلون على التقدير "ممتاز" وقلت نسبة الطلاب الراسبون كثيرًا وقد كان ذلك ملحوظًا على مستوى الامتحانات حيث وصلت نسبة النجاح في بعض الامتحانات ١٠٠% وخلت تمامًا من الطلاب الراسبون حتى بلغت نسبة الطلاب الحاصلون على التقدير "ممتاز" (٧١.٣٠%) في أحد الامتحانات ،وعلى العكس خلت أحد الامتحانات تماما من التقدير ممتاز وكانت نسبته (٠%) مما يدل على وجود خلل في هذه الامتحانات.

ومما سبق يتضح ارتفاع نسب نجاح الطلاب والتواء المنحني في الاتجاه الإيجابي وارتفاع نسب تحقق نواتج التعلم وهذه النتيجة الايجابية **إذا كانت حقيقية** فترجع الى وضوح نواتج التعلم في ذهن كل من المعلم والمتعلم واتباع كافة الأساليب والطرق والاستراتيجيات التدريسية المناسبة لتحققها ثم استخدام طرق التقييم المناسبة للتحقق منها وفقًا لنموذج المحاذاة البنائية لـ (Biggs & Tang, 2011) الذي يفترض أنه حينما تتوازى مكونات نظام التعليم ،خاصة مهام التقييم مع أنشطة التعلم والتعليم فمن السهل الوصول لنواتج التعلم المستهدفة، **ولكن إذا كانت غير حقيقة وزائفة** فقد تؤدي إلى عواقب وخيمة لاسيما أن هناك الكثير من أوجه القصور أهمها أنّ هذه العينة من الامتحانات من النمط الموضوعي ولم تشتمل على نماذج فربما مع الأعداد الكبيرة للطلاب والتي بلغ متوسطها (٦٥٨.٤٣) طالب لم يتم ضبط عملية الغش لاسيّما مع أن معظم الامتحانات لم تتضمن نماذج مختلفة للامتحان مما يُسهل عملية الغش خصوصًا مع عدد الأسئلة غير الملائم والذي كان في أغلب الأحوال أقل بكثير من زمن الامتحان أضف إلى ذلك أيضًا سهولة المفردات وزيادة نسبة المفردات غير المُمَيِّزة التي يقل معامل تمييزها عن (٠.٢)هذا بالاضافة إلى قياس نواتج التعلم الخاصة بالمهارات العليا بالأسئلة الموضوعية ذات الاجابة المحددة وعدم ادراج أنواع أخرى من الأسئلة المفتوحة والتي تسمح بذلك.

وهذه القضية تحتاج لمزيد من الدراسات الاحصائية لتحديد النسبة الحقيقة لتحقق نواتج التعلم المستهدفة بجميع مستوياتها بحيث تكون خالية من كافة أخطاء القياس مع ابتكار أنماط جديدة من الأسئلة لقياسها بموضوعية وشفافية.

توصيات الدراسة:

١. تقديم تقارير دورية لصُنَّاع القرار تشتمل على أهم نقاط القوة والضعف في الامتحانات النهائية بكلية التربية جامعة دمنهور لتحديد ورش العمل و الدروات التدريبية اللازمة لتحسين مهارات أعضاء هيئة التدريس وادراجها في الخطط التصحيحية للكلية والجامعة.
٢. تقديم التغذية الراجعة لأعضاء هيئة التدريس عن جودة أسئلتهم واختباراتهم وخصائصها السيكومترية وتدريبهم على كيفية اعداد الامتحانات في ضوء المواصفات والمعايير القياسية.
٣. الارشاد الأكاديمي للطلاب للتعريف بأهمية التقويم الجيد في ضوء نواتج التعلم المستهدفة من المقررات وعناصر جودة الامتحانات.
٤. تنويع أسئلة الامتحانات بين المقالية والموضوعية،المصححة آليًا ويدويًا بما يناسب طبيعة كل مقرر بحيث لا يُفرض على أستاذ المادة طريقة تقييم بعينها للحصول على تقويم دقيق وشامل لأداء الطلاب بكل مستوياته.
٥. التدريب المستمر لأعضاء هيئة التدريس وتقديم الدعم الفني والبرامج الالكترونية المتخصصة لاعداد الامتحانات الجيدة الخالية من أخطاء القياس وتطوير أنواع من الأسئلة التي تستهدف قياس المهارات العليا " والمعتمدة على Rubrics في التصحيح لضمان الموضوعية.
٦. المراجعة المستمرة لتقييم تحقق نواتج التعلم على مستوى البرامج والكلية والجامعة وبناء خطط تحسين لنقاط الضعف وتطوير أساليب تقويم البرامج في ضوء آراء الطلاب ونتائج الامتحانات.
٧. تفعيل عمل لجان فحص الامتحانات بعد تطبيقها ودراسة الأخطاء الشائعة والتدريب على كيفية تجنبها.
٨. ضرورة تدريب أعضاء هيئة التدريس على اعداد بنوك الأسئلة وتدريجها وفقًا لنظريات القياس الموضوعية واختبار مدى جدواها في قياس نواتج التعلم.

البحوث المقترحة:

١. أثر بناء الخطط التحسينية التي تستهدف تطوير نظم القياس والامتحانات في ضوء نتائج تقييم الامتحانات .
٢. بناء أدوات وتصميم برامج للقياس الكمي والكيفي لنواتج التعلم المستهدفة بكافة مستوياتها.
٣. دراسة أثر المراجعة غير الالزامية لمفردات الامتحان من الزملاء على جودتها وقابليتها للتطبيق.
٤. دراسة أثر إعلام الطلاب بنواتج التعلم المستهدفة وطرق ومعايير التقييم ومشاركتهم في تحديد مكونات الامتحان على مداخل تعلمهم وأدائهم الأكاديمي .
٥. دراسة أثر نوع ومستوى مشاركة الطلاب في وضع أسئلة على المحتوى في ضوء نواتج التعلم المستهدفة على أساليب تعلمهم وآدائهم الأكاديمي.

المراجع العربية:

١- إبراهيم الحسينان،محمود أحمد(٢٠١٤).تقويم اختبارات المقررات التربوية في كلية التربية جامعة المجمعة في ضوء معايير الجودة.مجلة كلية التربية جامعة الاسكندرية ،(٢٤)٤، ٢٣-٩٠.

٢- إقبال زين العابدين درندري (٢٠١٠).تقييم نواتج التعلم :نحو اطار مفاهيمي حديث في ضوء الاتجاهات المعاصرة للتقييم وجودة التعليم .مركز بحوث كلية التربية ،عمادة البحث العلمي ،جامعة الملك سعود ،المملكة العربية السعودية.
٣- أماني الغامدي،سماح زغاري(٢٠١٨).نواتج التعلم واختبارات المنتصف في كلية التربية بجامعة الإمام عبد الرحمن بن فيصل:دراسة تحليلية.مجلة الحكمة للدراسات التربوية والنفسية ،(١٦) ،٨-٣٦.
٤- امطانيوس مخائيل(٢٠١٦).بناء الاختبارات والمقاييس النفسية والتربوية وتقنينها.الطبعة الأولى،عمان ،دار الاعصار العلمي،٢١٧-٢٢٥.
٥- بندر بن زيد الأمير (٢٠١٦).الأخطاء الشائعة في بناء مفردات اختبار الاختيار من متعدد وتأثيرها على الخصائص السيكومترية للاختبارات التحصيلية .المجلة العربية للعلوم الاجتماعية ،(٩)٤، ١٣٧-١٧٥.
٦- حمد بن مرضي الكلثم (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.
٧- دليل الهيئة القومية لضمان الجودة والاعتماد الأكاديمي(٢٠١٣).معايير ضمان الجودة والاعتماد لبرامج التعليم العالي .جمهورية مصر العربية،٣٢-٣٣.
٨- سلمى زكي عبد الرحيم(٢٠٠١).دليلك في تصميم الاختبارات .الطبعة الأولى،عمان،الأردن للنشر والتوزيع،١٥٤-١٥٦.
٩- سليمان بن يوسف خاطر(٢٠٠٨).الأخطاء اللغوية في صياغة الاختبارات النهائية :أسبابها وأنماطها وعلاجها: دراسة وصفية تحليلية تقويمية.مجلة الجمعية العلمية السعودية للغة العربية ،(٢) ، ٢٥٧-٣٠٤.
١٠- سوسن شاكر مجيد(٢٠١٤).أسس بناء الاختبارات والمقاييس النفسية والتربوية. الطبعة الثالثة ،عمان ،مركز ديبونو لتعليم التفكير،١٢١-١٥٨.
١١- صفوت فرج (٢٠٠٧).القياس النفسي.الطبعة السادسة،القاهرة،الأنجلو المصرية،٣٠٧-٣٤١.
١٢- صلاح أحمد مراد،أمين على سليمان (٢٠٠٥).الاختبارات والمقاييس في العلوم النفسية والتربوية خطوات إعدادها وخصائصها.الطبعة الثانية ،القاهرة ،دار الكتاب الحديث، ٤٣، ٢١٥-٢١٩.
١٣- صلاح علام (٢٠٠٠).القياس والتقويم التربوي والنفسي أساسياته وتطبيقاته وتوجهاته المعاصرة .الطبعة الأولى، القاهرة ،دار الفكر العربي،١٣٢-١٧٩.
١٤- صلاح علام (٢٠٠٩).التقويم التربوي المؤسسي ؛أسسه ومنهجياته وتطبيقاته في تقويم المدارس.القاهرة ،دار الفكر العربي،٢١٢-٢٣٠.
١٥- ضياء الدين عبد الواحد (٢٠١٥).تقويم الطلاب بأقسام المكتبات والمعلومات في ضوء معايير الجودة والاعتماد :دراسة تحليلية للاختبارات التحصيلية .المجلة الدولية لعلوم المكتبات والمعلومات ،(٢)٤، ١١-٣٩.
١٦- عباس محمود عوض(١٩٩٨).القياس النفسي بين النظرية والتطبيق. الاسكندرية،دار المعرفة الجامعية،٥٨-٥٩.

١٧- عبد الكريم عبد الصمد السوداني(٢٠٠٦). مستويات التفكير في الاختبارات التحصيلية النهائية لقسم علوم الحياة. مجلة القادسية في الآداب والعلوم التربوية،(٥)١، ٧-١٧.
١٨- علي الحديبي،هناء إبراهيم(٢٠١٤).تقويم أوراق الاختبارات بالجامعات الاسلامية بالمدينة المنورة على ضوء معايير الجودة.مجلة الجمعية التربوية للدراسات الاجتماعية ،(٥٦) ،١٤٦- ٢٣٨.
١٩- علي صلاح حسن (٢٠١٨).الأخطاء الشائعة في بناء الاختبارات الموضوعية وتأثيرها على الخصائص السيكومترية للاختبارات التحصيلية لدى طلاب كلية التربية بجامعة أسيوط.مجلة كلية التربية جامعة أسيوط،(٣٤)٢، ،٣٦٥-٤٠٦.
٢٠- فاتن فاروق موسى،احسان شكري حجازي(٢٠٢٠).تقويم الاختبارات المصححة آليًا والاختبارات المصححة يدويًا واتجاه طلبة الفرقة الرابعة شعبة علم النفس التربوي نحوهما.مجلة كلية التربية جامعة بني سويف،عدد ابريل(١)،١-٢٦.
٢١- فؤاد أبو حطب ،سيد عثمان ،آمال صادق(٢٠٠٨).التقويم النفسي.الطبعة الرابعة، القاهرة ،مكتبة الأنجلو المصرية،١٧-٧٥.
٢٢- ليونا تايلر(١٩٨٨).الاختبارات والمقاييس .ترجمة: سعد عبد الرحمن ،محمد نجاتي ،الطبعة الثانية ،القاهرة ،دار الشروق،٣٥-٦٠.
٢٣- محمد عبد القادر الشرقاوي(٢٠١٦).دراسة تقويمية لمخرجات التعلم المستهدفة لمادة التربية الرياضية لتلاميذ المرحلة الاعدادية طبقا للمعايير القومية لضمان جودة التعليم والاعتماد.المجلة العلمية للبحوث والدراسات في التربية الرياضية،(٣٢)، ٢٢١-٢٤٥.
٢٤- محمد عبد المنعم (٢٠٠٩).تقويم أسئلة الاختبارات التحصيلية بكلية المعلمين جامعة الملك فيصل في ضوء معايير جودة التقويم .مجلة البحوث النفسية والتربوية ،كلية التربية جامعة المنوفية ،(٢٤)١ ،٢٠٦ – ٢٣٦.
٢٥- محمود إبراهيم،محمد السيد(٢٠١٥).تقييم الطلاب لدراسة مقرر الأصول الفلسفية للتربية في ضوء نواتج التعلم المستهدفة :دراسة ميدانية.دراسات عربية في التربية وعلم النفس ،(٦١)، ٢٣٩- ٢٩٦.
٢٦- محمود عمر،حصة فخرو،تركي السبيعي،آمنه تركي(٢٠١٠).القياس النفسي والتربوي. الطبعة الأولى ،عمان دار المسيرة للنشر والتوزيع،٩٦-٩٧.
٢٧- منار الشامي،سهام الشافعي(٢٠١١).جودة الاختبارات التحصيلية وعلاقتها بضغوط العمل لدى عضوات هيئة التدريس بجامعة تبوك.مجلة بحوث التربية النوعية،(٢٣) ،١٦٩-٢٠٧.
٢٨- منال حسن(٢٠١٣).تقويم اختبارات طالبات كلية التربية بحفر الباطن وفق معايير الجودة.مجلة القراءة والمعرفة ،كلية التربية جامعة عين شمس،(١٣٦)، ٢٦٦ -٣٢٢.
٢٩- هدى محمد صالح،حنان فوزي محمد(٢٠١٧).تقويم برنامج ماجستير المناهج وطرائق التدريس بمساري اللغة العربية والعلوم بجامعة القصيم في ضوء نواتج التعلم اللازم اكتسابها من وجهة نظر أعضاء هيئة التدريس والطلاب.دراسات في المناهج وطرق التدريس ،كلية التربية جامعة عين شمس ،(٢٢٥)، ١٦-٦٤.

٣٠- يحيى خليفة الشريفي(٢٠١٣).تقويم أسئلة الإختبارات النهائية لمادة الأدب في كليات التربية في ضوء مستويات بلوم المعرفية.مجلة القادسية للعلوم الإنسانية،(١٦)٣، ٣٧٧-٤٢٠.

المراجع الاجنبيه

31- Adam, S. (2004). "Using Learning Outcomes: A Consideration of the Nature, Role,Application and Implications for European Education of Employing Learning Outcomes at the Local, National and International levels", Report on United Kingdom Bologna Seminar, Herriot-Watt University,1-23.

32- Agnantis, K., Alexiadis, A., Refanidis, I. (2016)." COURSR2: an integrated time management system for lifelong learners", *international journal on artificial intelligence tools*,25(06),1650029.

33- Ahmedani,S.M. (2017)." Direct Methods Used to measure Program Learning Outcomes" ,available @ https://www.researchgate.net/publication/316190040 ,1-9.

34- Anderson, H.M., Moor, D.L., Anaya, G., Bird, E.(2005)." Student Learning Outcomes Assessment: A Component of Program Assessment", *American Journal of Pharmaceutical Education*, 69(2), 256–268 .

35- Anderson, L.W., Krathwohl, D.R.(2001)." A Taxonomy for Learning, Teaching and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives", Addison, Wesley, Longman ,New York .

36- Bacolod,P. M. & Tobias,L.J .(2006). "School quality and achievement Growth evidence from the Philippines ",*economics of education review*, 25 (6) , 619 – 632.

37- Biggs ,J.B. (1973). "Study behaviour and performance in objective and essay formats",*Australian Journal of Education*. 17, 157–167.

38- Biggs, J., and C. Tang. (2011). "Teaching for Quality Learning at University: Assessing for learning quality: II. Practice ", available @ https://www.researchgate.net/publication/215915395 ,1-36.

39- Bingham, J. (1999)." How to write learning outcomes" In:"AGuide to developing learning outcomes",The learning and Teaching Institute Sheffled Hallam University, Sheffled,1-18.

40- Bloom, B.S., Engelhart, M.D., Furst, E.J., Hill, W.H., Krathwohl, D.R.(1956)." Taxonomy of Educational Objectives", Handbook 1 Cognitive Domain. David McKay, New York ,201-207.

41- Boulton-Lewis, G.M. (1995). "The SOLO taxonomy as a means of shaping and assessing learning in higher education", *Higher Education Research and Development* .14, 143–154.

42- Brady,B.M.(2005). "Assessment of learning with multiple-choice questions", *Nurse Education in Practice* .5, 238–242.

43- Brown, C. A. (2002)." Portfolio Assessment: How Far Have We Come?"Available@ https://files.eric.ed.gov/fulltext/ED477941.pdf.

44- Dodeen,H.(2003)."Assessing test –taking strategies of university students:developing ascale and estimating its psychometric indices",*Assessment &Evaluation in Higher Education*.33(4)409-419.

45- Dodridge, M.(1999)." Learning Outcomes and Their Assessment in Higher Education" ,*EngineeringScience and Education Journal*, 161–168 .
46- Dolly, J.P., & Williams, K.S. (1986). "Using Test Taking Strategies ToMaximize ultiple-choice Test Scores", *Journal of Educational &Psychological Measurement*. (46) 3: 619 – 625.
47- Entwistle, A., & Entwistle, N. (1992). "Experiences of understanding in revising for degree examinations", *Learning and Instruction* .2, 1–22.
48- Ewais,A.&Awad,M.& Hadia,Kh.(2020)" Aligning Learning Materials and Assessment with Course Learning Outcomes in MOOCs Using Data Mining Techniques", in: " Hatzilygeroudis I., Perikos I., Grivokostopoulou F," Advances in Integrations of Intelligent Methods", Smart Innovation, Systems and Technologies,1-25.
49- Farley, J.(1989)." The multiple-choice test: writing the question",*Nurse Educator*. 14 (6), 3–5.
50- Frederickson, J.R. , & Collins, A. (1989). "A systems approach to educational testing", *Educational Researcher*. 18, 27–31.
51- Gibbs, G., Lucas, L., & Spouse, J., (1997). "The effects of class size and form of assessment on nursing student's performance,approaches to study and course perceptions", *Nurse EducationToday*. 17 (4), 311–318.
52- Gilbert, L., Gale, V.(2008)." Principles of E-Learning Systems Engineering", Chandos Publishing,Oxford .
53- Hammond, E.J., McIndoe, A.K., Sansome, A.J., & Spargo, P.M.(1998) ."Multiple-choice examination: adopting an evidencebased approach to exam technique", *Anaesthesia*. 53, 1105-1108.
54- Harden, R.M., Crosby, J.R., Davis, M.H.(1999)." AMEE Guide No.14: Outcome-Based Education":Part1 – An Introduction to Outcome-Based Education. *Medical Teacher* 21(1), 7–14.
55- Holsgrove, G.(1992). "Guide to post graduate exams: multiplechoicequestions British", *Journal of Hospital Medicine* .48(11), 757–761.
56- Jenkins, A., Unwin, D. (2007)." Writing and Using Learning Outcomes: A Practical Guide ", available@ https://www.researchgate.net/publication/238495834 ,1-30.
57- Kennedy, D., Hyland, A., Ryan, N.(2007)." Writing and Using Learning Outcomes", A PracticalGuide. University College Cork ,available@ : https://www.researchgate.net/publication/238495834
58- Kenny, N ., Desmarais, S., (2012). "A Guide to Developing and Assessing Learning Outcomes at the University of Guelph",office of the associate Vice-president(Academic).available@ https://docplayer.net/8753501-A-guide-to-developing-and-assessing-learning-outcomes-at-the-university-of-guelph.html.
59- Marnat,G.G.(2003)."Hand book of psychological assessment", (4)ed, wiley,New jersey,1-10.
60- Masters, J., Hulsmeyer, B., Pike, M., Leichty, K., Miller, M., &Verst, A., (2001)." Assessment of multiple-choice questions in selected test banks accompanying test books used in nursing education", *Journal of Nursing Education*. 40 (1), 25–32.

61- Merrill, M.D.(1973)." Content and Instructional Analysis for Cognitive Transfer Tasks". *Audio Visual Communications Review*, 21, 109–125 (1973).

62- Merrill, M.D.(1994)." The Descriptive Component Display Theory". In: Twitchell, D.G. (ed.) "Instructional Design Theory", *Educational Technology Publications*, New Jersey.111-157.

63- Millman, J., Bishop, C. H., & Ebel, R. (1965). "An Analysis of Test Wiseness", *Journal of Educational & psychological Measurement*.(25) 3: 707 -726.

64- Oermann, M.(1999). "Developing and scoring essay tests", *NurseEducator* .24 (2), 29–32.

65- Pepple,D.J., Young,L.E, & Carroll,R.G.(2010). "A comparison of student performance in multiple-choice and long essay questions in the MBBS stage I Physiology examination at the University of the West Indies (Mona Campus)", Adv Physiol Educ .34: 86–89.

66- Pilli, O.,Admiraal,W. (2017)." Students ' learning outcomes inmassive open online courses (MOOCs):some suggestions for course design", *J. Higher Educ*. 7(1), 46–71.

67- Race, P., & Brown, S. (2001). "The Lecturers Toolkit: A Practical Guide to Assessment, Learning and Teaching ", Kogan, London,1-28.

68- Ramsden, P. (1988b). "Studying learning: Improving teaching", in Ramsden, P. (ed.)," Improving Learning: New Perspectives", Kogan, London, pp. 13–31.

69- Sarnacki, R. E. (1979). "An Examination of Test Wiseness in the Cognitive Test Domain", *Journal of Review of Educational Research*. (49) 2: 252 – 279.

70- Tangworakitthaworn P., Gilbert L., Wills G.B. (2013)." A Conceptual Model of Intended Learning Outcomes Supporting Curriculum Development", In: Ng W., Storey V.C., Trujillo J.C. (eds) Conceptual Modeling, ER 2013. *Lecture Notes in Computer Science*, (8217). Springer, Berlin, Heidelberg,161-173.

71- Wilson ,A. , Howitt,S.,& Higgins,D. (2016)." A fundamental misalignment: intended learning and assessment practices in undergraduate science research projects", *Assessment & Evaluation in Higher Education*, 41:6, 869-884, DOI: 10.1080/02602938.2015.1048505.

" Assessing the exams of the Faculty of Education students, and the extent it achieve the intended learning outcomes"

Fathi Okasha, Mahmoud
Faculty of education/Damanhour university/educational psychology department

Eman Salah Daha
damanhour university-faculty of education-educational psychology department

Abstract

the current study aimed to study the extent to which the achievement tests in the university(Faculty of Education, Damanhur University as example) achieved the demands of assessment quality in terms of the standard specifications of the exam paper , studying its psychometric evidences , and the extent it achieved the intended learning outcomes; Considering viewpoint of the students and their satisfaction with it .

A student satisfaction questionnaire about the final tests was applied after checking its validity on a sample consisting of (649 male and female students), and the list of exam papers standard specifications was used to examine a sample of (46) exam papers chosen for educational courses as they general for all specializations , distributed on four consecutive semesters, and the technical reports provided by the automated graded programs for a sample of (14) tests were examined to verify their psychometric evidences in measuring the intended learning outcomes. The results revealed:

1. Students' satisfaction about the formal specifications for the final exam papers was sufficient, while their opinions regarding the qualitative specifications varied; The overall verification percentage for the formal indicators was 73.31%, while the verification percentage for the qualitative indicators was 64.41%.

2. The average reliability coefficient across the sample of the mechanically graded tests was (0.74 average value) and the percentage of non-discriminative questions that have a discrimination factor (less than 0.2) increased in it.

3. The percentage of the achievement average of learning outcomes (calculated by the results of students on each of the test items and their conformity with the learning outcome that the item set for measuring it) ranged from 69.76% to 90.78% with an average of 80.48% at the total level of exam papers and this result consider acceptable only in The state of stability of the tests and the quality of their items.

A set of recommendations was presented to decision makers in both the college and the university for inclusion in corrective plans in addition to some future research had been suggested.

Keywords: ILOs assessment /Achievement tests / Standard specifications for exam paper