# Reality of Essay Questions in Final Exams at the College of Education in Shaqra: An Evaluative Study

Thuraya Abdullah Hussein Al-Jar
Faculty of Education in Shaqra
Shaqra University

thaljaar@su.edu.sa

Received : 26/11/2025 Accepted : 02/02/2026

**Abstract:**

In light of global shifts toward enhancing higher-order thinking skills, the importance of evaluating university assessment tools, particularly essay questions, is highlighted, as they are key to measuring these skills. This study aims to evaluate the reality of essay questions in final exams at the College of Education in Shaqra, in light of the modified Bloom's Taxonomy, in the context of trends in developing university assessments. The study presents an applied model that can be generalized to colleges of education in Saudi universities, taking into account local characteristics. The study seeks to analyze the reality of essay questions in a typical Saudi educational environment, link them to educational assessment quality standards, and offer applicable solutions. It also aims to determine the diversity of questions and the percentages of representation of cognitive levels, and to analyze the reasons for neglecting higher-order questions (analysis-creativity).

**Keywords**: Bloom’s taxonomy, essay questions, higher-order thinking, assessment reform, Saudi education.

# واقع تطبيق الأسئلة المقالية في الاختبارات النهائية بكلية التربية بشقراء : دراسة تقويمية

ثريا عبدالله حسين الجار
كلية التربية بشقراء
جامعة شقراء
thaljaar@su.edu.sa

الاستلام :2025/11/26 القبول :2026/02/02

**الملخص:**

في ظل التحولات العالمية نحو تعزيز مهارات التفكير العليا، تبرز أهمية تقييم أدوات التقويم الجامعي، وخاصة الأسئلة المقالية التي تُعد مفتاحًا لقياس هذه المهارات. وتهدف هذه الدراسة إلى تقييم واقع الأسئلة المقالية في الاختبارات النهائية بكلية التربية بشقراء وفق تصنيف بلوم المعدل، في ظل توجهات تطوير التقويم الجامعي. وتُقدم الدراسة نموذجًا تطبيقيًا يمكن تعميمه على الكليات التربوية في الجامعات السعودية، مع مراعاة الخصائص المحلية. سعت الدراسة إلى التعرف على واقع استخدام الأسئلة المقالية بالاختبارات النهائية في قسم تعليم الطفولة المبكرة بكلية التربية بشقراء، وربطها بمعايير جودة التقويم التربوي، وتقديم حلول قابلة للتطبيق، وتحديد مدى تنوع الأسئلة، ونسب تمثيل المستويات المعرفية، وتحليل أسباب إهمال الأسئلة العليا (التحليل–الإبداع). استخدمت الدراسة المنهج الوصفي التحليلي، حيث اعتمدت على تحليل محتوى كمي باستخدام استمارة مُحكمة، بالإضافة إلى تحليل نوعي لآراء أعضاء هيئة التدريس عبر مقابلات شبه مقننة. وتم تحليل 217 سؤالاً من 47 اختباراً نهائياً، من اختبارات قسم تعليم الطفولة المبكرة، مع إجراء مقابلات مع 13 عضو هيئة تدريس من العاملين بكلية التربية بشقراء. أظهرت النتائج هيمنة أسئلة التذكر (81%)، وضعفاً شديداً في أسئلة التحليل (0.5%) والإبداع (0%)، مع عزوف أعضاء هيئة التدريس عن استخدامها بسبب صعوبة التصحيح. وأوصت الدراسة بتطوير دليل إرشادي لتصميم الأسئلة وفق بلوم المعدل، وتضمين ورش عمل دورية في خطط تطوير أعضاء هيئة التدريس، وتدريب أعضاء هيئة التدريس على صياغة الأسئلة العليا، وإنشاء بنك أسئلة معتمد.

**الكلمات المفتاحية:** تصنيف بلوم، أسئلة المقال، مهارات التفكير العليا، إصلاح التقييم، التعليم السعودي.

**المقدمة:**

يشهد ميدان التعليم تطورات متسارعة تطال جميع مكونات المنظومة التعليمية، ومنها التقويم الذي يُعدّ عنصراً مهماً منها، فهو الموجه الأساسي لتطويرها والعامل المؤثر بجميع مكوناتها، فالتقويم لم يعد مجرد إصدار حكم على شيء بالضعف أو القوة، بل أصبح اتخاذ قرارات ملموسة وعملية بالتعديل والتحسين وإعادة الصياغة. ورغم كل التطورات التي شهدها التقويم التربوي، في تنوع أساليبه وأدواته، إلا أن الاختبارات العامة، والاختبارات التحريرية خاصة، ما تزال الأداة الأكثر شيوعاً واستخداماً في مؤسسات التعليم العام والتعليم الجامعي. وتؤكد التوجهات الحديثة في التقويم أهمية الاختبارات، حيث تأتي أهميتها من كونها وسيلة موثوقة للحكم على فعالية العملية التعليمية، من خلال تشخيص العقبات والمشكلات التعليمية، وتقديم حلول مناسبة، تُمكّن صانعي القرارات من رسم سياسات تربوية هامة تتعلق بمراجعة أهداف الخطط التعليمية، وملاءمتها للمنهج، وتُعد الاختبارات أداة مهمة من أدوات التقويم خاصة في المرحلة الجامعية، التي تهدف إلى مدى نجاح الخطة الدراسية بمكوناتها، في تحقيق أهدافها، ومخرجاتها المتوقعة [10].

ولهذا فإن الاختبارات النهائية التي تجرى في نهاية الفصل الدراسي، مهمة جدًا للطالب الجامعي، حيث تهدف الاختبارات إلى نقل الطالب من مستوى دراسي إلى آخر أعلى، أو من مرحلة دراسية إلى أخرى أو توجيههم إلى شعب وتخصصات علمية، أو استحقاقه شهادات علميّة، أو حتى انسحابه بعد الرسوب أو التعثر؛ ومع هذه الأهمية للاختبارات النهائية بالمرحلة الجامعية، وفي كونها أدق المقاييس لكثير من مخرجات التعلم، إلا أن صدق المعلومات المحصلة من نتائج هذه الاختبارات يتوقف على مدى العناية والاهتمام التي تُولَى لإعدادها، فمع غياب معايير إعداد الاختبارات النهائية في الشكل والمضمون، وترك الأمر لاجتهادات أعضاء هيئة التدريس بالجامعات تبرز حاجة ملحة إلى إجراء دراسات تحليلية لأسئلة الاختبارات للتأكد من أنها تتصف بالضبط العلمي[4]،

وأنها على درجة من التنوع والشمول، وأنها تتناسب مع تصنيف بنجامين بلوم للأهداف المعرفية بكل مستوياته [5] وتُعد الاختبارات المقالية من أهم وأقدم الاختبارات التي تستخدم في تقويم أهداف لا يمكن تقويمها بالاختبارات الموضوعية مثل القدرة على تحليل الأفكار والربط بينها وتقييمها والقدرة على إنتاج أفكار جديدة، كما أنها لا تتأثر بعامل التخمين العشوائي، وهي وسيلة في قدرتها على قياس مستويات عقلية عليا وغيرها من المزايا[9].

ولذا ترى الباحثة أن الاختبارات بدلاً من أن تكون وسيلة للتحسين أصبحت هدفاً أكثر منها وسيلة للتعليم مما أثر على العملية التعليمية فأصبح اهتمام عضو التدريس ودوره هو تعليم الطالب المواضيع الهامة والتركيز على كيفية الإجابة على بعض الأسئلة، مما دفع بالطالب إلى الحفظ والاستظهار وإهمال بقية مكونات النمو الأخرى التي تسهم في بناء شخصية متوازنة قادرة على التعامل مع المتغيرات المستحدثة.

لهذا ينبغي مراجعة وتقويم الاختبارات المقالية؛ لإنه إذا لم تكن عملية التقويم على درجة من الدقة والإتقان والموضوعية فإنها ستعطي نتائج غير صحيحة أو مضللة يترتب عليها اتخاذ قرارات وإجراءات خاطئة تفسد المنظومة التعليمية، وتضرها أكثر مما تفيدها [37].

## مشكلة الدراسة وأسئلتها:

يفرض التطور المعرفي السريع وتنوع مصادر المعرفة على منظومة التعليم ضرورة تطوير الاختبارات التحصيلية، والانتقال بالتعليم من ثقافة الحفظ والاسترجاع إلى ثقافة الإبداع والابتكار، من أجل إعداد الفرد الكفء القادر على التعامل مع متطلبات الحياة المعقدة في مواجهة المشكلات.

وبالتالي ومن خلال عمل الباحثة كعضو هيئة تدريس لاحظت أن معظم الجامعات العربية تتبنى الاختبارات التحصيلية كوسيلة وحيدة للتقويم، فقد أصبحت هذه الاختبارات من أكثر الأسباب التي تسهم في تدني مستوى المنتج التعليمي من عدة جوانب كضعف التأهيل والإعداد، وعدم القدرة على مواجهة متطلبات التنمية، وضعف القدرة التنافسية الضرورية لاحتياجات سوق العمل، وتؤدي إلى نجاح وهمي دون فهم وتطبيق لما تعلموه في مواقف جديدة. ونظراً لأهمية الاختبارات التحصيلية النهائية في الجامعات والاعتماد عليها في الكشف عن مدى تحقق الأهداف التعليمية، وفي إصدار الأحكام واتخاذ القرارات الملائمة، فقد لقيت اهتماماً كبيراً من الباحثين على مستوى العالم، وأجريت العديد من الدراسات التي كشفت نتائج معظمها عن قصور في بناء الاختبارات من نواحي كثيرة، كدراسة نايف وعبدالكاظم [38]، ودراسة عبده[1]، ودراسة لايبي والمياحي[34]، ودراسة سعيد وجبار[42] ودراسة الشريفي[14]، وغيرها من الدراسات السابقة التي توصلت نتائجها إلى عدم الموازنة بين الأسئلة الموضوعية والمقالية، وعدم الالتزام بالأبعاد المختلفة لتصنيف بلوم للمجال المعرفي، والتركيز المستويات المعرفية الدنيا وإهمال المستويات العليا، من هذا المنطلق تأتي هذه الدراسة لمعرفة واقع الأسئلة المقالية في الاختبارات النهائية بقسم تعليم الطفولة المبكرة بكلية التربية بشقراء، بجامعة شقراء.

## أسئلة الدراسة:

تتحدد مشكلة الدراسة الحالية في السؤال الرئيس الآتي: ما واقع الأسئلة المقالية في الاختبارات النهائية ببعض كليات جامعة شقراء؟ وينبثق عن هذا السؤال الأسئلة الآتية:

1. ما مدى تضمين أسئلة الاختبارات النهائية للأسئلة المقالية، وتنوعها بكلية التربية بجامعة شقراء؟
2. ما مدى التزام أعضاء هيئة التدريس بالمعايير الشكلية لصياغة الأسئلة المقالية في الاختبارات النهائية بكلية التربية بشقراء؟
3. ما نسب تمثيل الأسئلة المقالية في الاختبارات النهائية لمستويات الأهداف المعرفية وفق تصنيف بلوم المعدل؟
4. ما هي أسباب ضعف نسبة تضمين الأسئلة المقالية ومستوياتها المعرفية العليا في الاختبارات النهائية بكلية التربية بجامعة شقراء من وجهة نظر أعضاء هيئة التدريس بكلية التربية بشقراء؟
5. ما هي مقترحات أعضاء هيئة التدريس بالجامعة لتضمين الأسئلة المقالية في الاختبارات النهائية لطلبة جامعة شقراء؟

## أهداف الدراسة:

تهدف الدراسة إلى تقويم الأسئلة المقالية في الاختبارات النهائية بكلية التربية بشقراء بجامعة شقراء، وذلك من خلال:

1. التعرّف على مدى تنوع أسئلة الاختبارات النهائية.
2. الكشف عن نسب تضمين الأسئلة المقالية في الاختبارات النهائية.
3. توضيح نسب تمثيل الأسئلة المقالية في الاختبارات النهائية للمستويات المعرفية.
4. تفسير أسباب ضعف تمثيل الأسئلة المقالية في الاختبارات النهائية للمستويات المعرفية العليا.

**أهمية الدراسة:**

تتمثل أهمية الدراسة في الآتي:

1. تناول الدراسة للاختبارات التحريرية النهائية باعتبارها المكون الرئيس للاختبارات التحصيلية التي تمثل أهم وسائل التقويم في الجامعات والمعاهد.
2. تعنى الدراسة بدراسة واقع الأسئلة المقاليّة في الاختبارات النهائية، وهو من المواضيع التي لم تلق الدراسة والبحث الكافيين –حسب علم الباحثة– بالجامعات السعودية.
3. قد تفيد نتائج هذه الدراسة المسؤولين والقائمين على إعداد الاختبارات النهائية في معرفة مستوى جودة الاختبارات التي يُعدها أعضاء هيئة التدريس، واتخاذ الإجراءات اللازمة لتجويد عمليات إعدادها.
4. يؤمل أن تكون هذه الدراسة مهادا لدراسات وأبحاث تحليلية تقويمية أخرى للاختبارات المختلفة.

**حدود الدراسة:**

تقتصر هذه الدراسة وتتحدد بالحدود الآتية:

**الحد الموضوعي:** تمثيل أسئلة الاختبارات النهائية بقسم تعليم الطفولة المبكرة بكلية التربية بجامعة شقراء للمستويات المعرفية حسب تصنيف بلوم المعدّل.

**الحد المكاني:** كلية التربية بجامعة شقراء

**الحد الزماني:** العام الجامعي 1446هـ.

**الحدود المنهجية:** ارتباط نتائج الدراسة بأداتي الدراسة (بطاقة تحليل الأسئلة وبطاقة المقابلة).

**الدراسات السابقة:**

اطلعت الباحثة على عدد من الدراسات السابقة المتعلقة بموضوع الدراسة الحالية، حيث تم ترتيبها من الأقدم للأحدث كالتالي:

سعت **دراسة نايف وعبد الكاظم**[38] إلى تقويم أسئلة الاختبارات النهائية لأربعة صفوف دراسية في كلية التربية للعلوم الإنسانية بجامعة كربلاء في ضوء مستويات بلوم للمجال المعرفي، بهدف الوقوف على مدى شمولية أسئلة التاريخ في الصفوف الأربعة. اعتمدت الدراسة المنهج الوصفي التحليلي واستخدمت التكرارات والنسب المئوية، واستخدام بطاقة التحليل أداة له، كشفت نتائج الدراسة عن اقتصار الاختبارات على الأسئلة المقالية. في حين حصل مستوى الفهم تكرر في الاختبارات بأعلى نسبة وصلت (52%)، تلاه مستوى المعرفة بنسبة (27%). بينما حصل كل من مستوى التطبيق والتقويم على نسبة (1%). أكدت الدراسة على ضرورة التركيز على المستويات العليا لتصنيف بلوم، مع تنويع الأسئلة بين الموضوعية والمقالية.

اهتمت **دراسة عبده**[1] بتقييم أنواع الأسئلة المتضمنة في أوراق الاختبارات التي يجريها طلاب قسم اللغة الإنجليزية بكلية التربية بجامعة الحديدة، لمعرفة مدى تمثيل هذه الأسئلة لمستويات بلوم للمجال المعرفي. اعتمدت الدراسة المنهج الوصفي التحليلي، واستخدمت التكرارات والنسب المئوية، واستخدمت بطاقة التحليل أداة لها. تم تحليل (80) ورقة اختبار، وكشفت نتائج الدراسة أن معظم الأسئلة في أوراق الاختبارات تركز على تقييم مهارات التفكير الدنيا، مثل: المعرفة والفهم والتطبيق، إلا أنها تتجاهل المستويات العليا كالتركيب والتقييم. وأوصت الدراسة بالاهتمام بالمستويات العليا لتصنيف بلوم للمجال المعرفي.

هدفت **دراسة لعيبي والمياحي**[34] إلى تحليل أسئلة الاختبار للدراسات العليا في ضوء معايير جودة أسئلة الاختبارات في قسم العلوم التربوية بكلية التربية بجامعة البصرة. وظفت الدراسة المنهج الوصفي التحليلي واستخدمت التكرارات والنسب المئوية. واستخدمت بطاقة تحليل تكونت من ستة معايير، كان من بينها معيار مدى تمثيل الأسئلة لمستويات تصنيف بلوم المعدّل، تم تطبيقها على (35) سؤالاً. أظهرت النتائج عدم تحقيق الأسئلة لجميع معايير جودة أسئلة الاختبارات، وعدم تنوعها، وأن معظمها اقتصر على المستويات الدنيا من تصنيف بلوم المعدّل. فقد حصل مستوى التذكر على نسبة (57%)، تلاه مستوى الفهم بنسبة (23%)، بينما حصل مستوى التطبيق على نسبة (8%)، وحصل مستوى التحليل على نسبة (4%). وأوصت الدراسة بالتأكد من تمثل الأسئلة لكافة مستويات المجال المعرفي لتصنيف بلوم المعدّل والتركيز على الأسئلة التي تحفّز الطلبة على التفكير.

هدفت **دراسة سعيد وجبار**[42] إلى التحقق من مدى تغطية أسئلة الاختبارات النهائية للأهداف المعرفية بمستوياتها المختلفة تبعاً لتصنيف بلوم، بقسمي التاريخ والجغرافيا بجامعة كرميان. اعتمدت الدراسة المنهج الوصفي التحليلي واستخدمت التكرارات والنسب المئوية. استخدمت الدراسة بطاقة التحليل أداة لها، وتم تطبيقها على (565) سؤالاً. توصلت نتائج الدراسة إلى أن مستوى المعرفة حصل على أعلى نسبة مئوية بلغت (47%)، وحصل مستوى الفهم على نسبة (29%)، بينما حصل مستوى التطبيق على نسبة (11%)، ونسبة (8%) لمستوى التحليل، و(1.9%) لمستوى التقويم. كما كشفت النتائج عن تفوق نسبة الأسئلة المقالية مقارنة بالأسئلة الموضوعية. وأوصت الدراسة

بضرورة العناية بالمستويات المعرفية لتصنيف بلوم مع تنويع الأسئلة ما بين موضوعية ومقالية.

هدفت **دراسة الجابري وحسين**[8] إلى تقويم أسئلة الاختبارات الإلكترونية لقسم اللغة العربية في كلية التربية الأساسية بجامعة ميسان وفقاً لتصنيف بلوم. اعتمدت الدراسة المنهج الوصفي التحليلي، واستمارة التحليل أداة لها. تم تطبيق الأداة على (207) سؤال إلكتروني، من مواد التخصص للغة العربية، واستخدمت التكرارات والنسب لاستخراج النتائج. توصلت نتائج الدراسة إلى أن الأسئلة تركز بدرجة كبيرة على المستويات الثلاث الأولى من مستويات بلوم المعرفية. حيث حصل مستوى التذكر على نسبة (40%) ومستوى الفهم على نسبة (33%)، (16%) لمستوى التطبيق. في حين تركز الأسئلة بنسب ضئيلة على مستويات التحليل والتركيب والتقويم. حيث حصل مستوى التحليل على نسبة (7%)، ونسبة (1%) لمستوى التركيب، ونسبة (0.9%) لمستوى التقويم. وأوصت الدراسة بضرورة جعل أسئلة الاختبارات الإلكترونية ممثلة لمستويات بلوم المعرفية الدنيا والعليا.

سعت **دراسة الشريفي**[14] إلى تقويم الاختبارات النهائية لقسم اللغة العربية في ضوء تصنيف بلوم المعرفي، ومعايير وضع الأسئلة الجامعية بكلية التربية بجامعة القادسية. اعتمدت الدارسة منهج تحليل المحتوى لتحليل (119) سؤالاً، باستخدام التكرارات والنسب المئوية، واستخدمت بطاقة تحليل كأداة لها. توصلت النتائج إلى أن الأسئلة قد ركزت على المستويات الدنيا للمجال المعرفي لبلوم، وهي التذكر والفهم والتطبيق، حيث حصلت على نسب (25%)، و (38%)، و(15%) على التوالي. في حين أهملت غالب الأسئلة المستويات العليا من التصنيف، وهي التحليل والتركيب والتقويم، حيث حصلت على نسب (10%)، و(5%)، و(7%) على التوالي. وأوصت الدراسة بعدم الركون إلى وضع أسئلة سطحية تركز على مستويات الفهم أو الاستيعاب، والتركيز على المستويات العليا التي تنمي روح التفكير.

سعت **دراسة غايتس وبلو**[29] إلى تقويم الاختبارات النهائية التحريرية للدارسات العليا بكلية الفيزياء وعلم الفلك في جامعة ليدز بالمملكة المتحدة. تم تحليل أوراق الاختبارات لمقررين من مقررات تخصص الفيزياء لعامي 2016م و2017م، باستخدام منهجية تعتمد على تصنيف بلوم. ناقشت الدراسة أن بعض الأفعال المستخدمة في الأسئلة تتناسب بشكل لا جدال فيه مع مستويات تصنيف بلوم، في حين يصعب تصنيف أغلب أفعال الأسئلة بشكل مؤكد. توصلت النتائج إلى أن الأسئلة ركزت على المستويات الدنيا للمجال المعرفي لبلوم، وهي التذكر والتطبيق والتحليل، وغفلت عن المستويات العليا. وأوصت الدراسة بضرورة تصميم أسئلة الاختبار باستخدام الأفعال القابلة للقياس مع الأخذ بالاعتبار للمستويات العليا من التفكير.

سعت دراسة قادر وزملائه[28] إلى تحديد مدى التزام أعضاء هيئة التدريس في كليات التمريض والعلوم بالجامعة الجمهورية بكردستان بالأبعاد المختلفة لتصنيف بلوم للمجال المعرفي. عالجت الدراسة الموضوع باستخدام المنهج الوصفي الكمي لتحليل (524) سؤالاً من (75) ورقة اختبار نهائي من إعداد أعضاء هيئة التدريس في كلا الكليتين. كشفت النتائج عن انخفاض تمثل المستويات العليا بالأسئلة، مقابل ارتفاع تمثل المستويات الدنيا. حيث حصل مستوى التذكر على نسبة تمثل بلغت (48.28%) تلاه مستوى الفهم بنسبة (18.7%)، والتطبيق بنسبة (17.9%). في حين حصل مستوى التحليل على نسبة (9.16%)، ونسبة (2.1%) لمستوي التقييم والابتكار. كما كشفت الدراسة إلى وجود علاقة طردية بين انخفاض مستوى السؤال وارتفاع النجاح فيه. وأوصت الدراسة بضرورة أخذ المستويات العليا لتصنيف بلوم بعين الاعتبار عند وضع الأسئلة، وإنشاء لجنة مختصة بفحص الأسئلة قبل طباعتها وتسليمها للطلبة، وإجراء مزيد من البحوث التي تشجع الجامعات على التركيز على خطط التقييم عند تصميم مناهج المقررات الدراسية، خاصة فيما يتعلق بشمولية الأهداف التدريسية لجميع مستويات تصنيف بلوم للمجال المعرفي.

هدفت **دراسة أمين**[19] إلى تحليل وتقويم أسئلة الامتحانات النهائية لأقسام كلية التربية الأساسية بجامعة صلاح الدين بأربيل. وظفت الدراسة المنهج الوصفي بشقيه المسحي والتحليلي لتحليل (90) ورقة اختبار من الاختبارات النهائية النظرية التي يعدها أعضاء هيئة التدريس بستة من الأقسام العلمية والإنسانية. لتحقيق أهداف الدراسة تم تصميم استمارة تحليل. كشفت نتائج الدراسة عن تركيز معظم الأسئلة على مستويات التذكر والفهم والتطبيق، وضعف تمثيلها لمستويات التحليل والتركيب والتقويم. حيث حصل مستوى التذكر على نسبة (39.24%)، ومستوى الفهم على نسبة (22.5%)، ومستوى التطبيق (16.44%). في حين حصلت المستويات الثلاثة العليا على نسبة (8.81%)، (2.04%)، و(0.58%) على التوالي. وأوصت الدراسة بضرورة التركيز على نوع الأسئلة التي تحفز الطلبة على المستويات العليا، وتدريب أعضاء هيئة التدريس على وضع الأسئلة بصياغة صحيحة وتنويعها.

استهدفت **دراسة زهور وزملائه**[51] تحليل المستوى المعرفي لأوراق الاختبارات النهائية بقسم العلاج الطبيعي بجامعة كراتشي الطبية الخاصة بالباكستان، باستخدام تصنيف بلوم. تم تقييم (1080) سؤالاً متعدد الاختيارات وفقاً لتصنيف بلوم المعدل من اختبارات المستويات الدراسية الخمسة بالقسم. استخدمت الدراسة المنهج الوصفي المسحي، وتم تصميم استمارة تحليل. وقارنت النتائج مدى تمثيل الأسئلة في أسئلة المستويات الخمسة. وتوصلت إلى أن معظم الأسئلة ذات المستويات الدنيا من التصنيف كانت من أسئلة المستوى الدراسي الأول والثاني بنسبة تتراوح بين (91%) إلى (94%). في حين جاءت المستويات الأعلى للمستويات الدراسية السنة الثالثة إلى الخامسة، بنسبة تتراوح بين (27.5%) إلى (38%). ولم تجد الدراسة سؤالاً واحداً من المستويين الخامس والسادس من مستويات تصنيف بلوم المعدل في أسئلة أي من المستويات الدراسية الخمسة. أوصت الدراسة بضرورة تركيز أعضاء هيئة التدريس على الأسئلة ذات المستويات المعرفية الأعلى، والاهتمام بتطوير التفكير النقدي لدى الطلبة.

هدفت **دراسة علي وزملائه**[7] إلى تحليل أسئلة الاختبارات النهائية على مدى أربع سنوات في ثلاث كليات، المقالية منها، لمعرفة مستويات التعلم وفق تصنيف بلوم للمجال المعرفي، في مقررات البرمجة في كلية علوم الحاسب بجامعة السودان. اعتمدت الدراسة المنهج الكمي والنوعي وتحليل المحتوى. ولجمع البيانات استخدمت الدراسة استمارة تحليل تم تطبيقها على (2318) سؤالاً مأخوذة من (110) ورقة اختبار نهائي، ومن مستويات مختلفة، بشرط أن يكون اختبار نفس المقرر في ثلاث سنوات على الأقل من 2015–2018م. كشفت النتائج أن مستوى المعرفة حصل على المركز الأول بنسبة (62.5%)، ومستوى التطبيق في المرتبة الثانية بنسبة (18.1%)، تلاه مستوى الفهم بنسبة (12.6%)، فيما بلغت نسبة مستوى التحليل (3.9%)، ومستوى التقييم بنسبة (1.6%)، وجاء مستوى التركيب في المرتبة الأخيرة بنسبة (1.3%). وأوصت الدراسة بتحسين جودة التقييم من خلال تعديل طرق التدريس، وتوجيه الأسئلة نحو المهارات المعرفية عالية المستوى وفقاً لتصنيف بلوم.

**التعليق على الدراسات السابقة:**

- تتفق الدراسة الحالية مع أغلب الدراسات السابقة في تناولها لموضوع أسئلة الاختبارات النهائية، في المقررات الجامعية، وتمثيلها للأهداف المعرفية وفق تصنيف بلوم المعدّل، واستخدامها للمنهج الوصفي التحليلي، واستخدام استمارة تحليل المحتوى.
- وتختلف الدراسة الحالية عن الدراسات السابقة في مكان الدراسة وإجراءاتها، وتتباين في المقررات الدراسية المستخدمة في التحليل. كما تختلف في تركيز هذه الدراسة على الأسئلة المقاليّة، ودراسة أسباب تدني تمثيل الأسئلة المقالية للمستويات العليا، واقتراح إجراءات للتحسين.
- أفادت الباحثة من تلك الدراسات في اعتماد المنهج المستخدم، وبناء أداة الدراسة، وإجراءات الصدق والثبات، وصياغة أسئلة الدراسة، وتفسير النتائج.

**التحليل النقدي للدراسات السابقة:**

**نقاط القوة في الدراسات السابقة:**

اتفقت معظم الدراسات (مثل دراسة الجابري، وحسين، 2022؛ وقادر وزملائه، 2023) على:

- هيمنة الأسئلة الموضوعية في الاختبارات الجامعية.
- اقتصار الأسئلة المقالية على قياس التذكر والفهم (بنسبة 80–90%).

**أوجه القصور والفجوات البحثية:**

الدراسات السابقة مثل دراسة سعيد وجبار[42] ركزت على تحليل توزيع الأسئلة دون تفسير أسباب هذا التوزيع. ولذا تم إضافة مقابلات مع أعضاء هيئة التدريس لفهم الأسباب الكامنة (مثل: ضعف التدريب، وصعوبة التصحيح).

**محدودية العينات:**

العديد من الدراسات، مثل دراسة الشريفي[14]، حللت أسئلة من تخصص واحد (اللغة العربية). ولذلك تم توسيع العينة لتشمل تخصصات متعددة في كلية التربية (الطفولة المبكرة، والتربية الخاصة).

**إهمال الجانب التطبيقي:**

الدراسات السابقة مثل **دراسة نايف وعبد الكاظم** [38] قدمت توصيات عامة دون اقتراح أدوات عملية (مثل: نماذج لأسئلة إبداعية). ولذلك قدمت الباحثة دليلاً إجرائيًا لصياغة الأسئلة وفق مستويات بلوم المعدل (مرفق في الملاحق).

**الربط بين الإطار النظري ومشكلة الدراسة:**

"في ضوء نقص الدراسات التي تربط بين تصنيف بلوم المعدل وممارسات أعضاء هيئة التدريس، تهدف هذه الدراسة إلى تحليل الأسئلة المقالية ليس فقط كميًا، ولكن من خلال فهم السياق المؤسسي والتربوي الذي يحدد شكلها".

**التعريفات المفاهيمية والإجرائية:**

- **السؤال:** مشكلة تتطلب تفكيراً وجهداً وتنظيماً وإدراكاً للعلاقات وصولاً إلى الحل[46].
- ويُعرّف إجرائياً: بأنه مهمة تعليمية مصوغة على نحو يمكن معه قياس مدى تحقق الأهداف التدريسية لمقرر معين من خلال استجابة الطلبة لها.
- **الاختبار:** أداة لقياس عينة من السلوك، تتكون في مجموعها من عدد من البنود أو الأسئلة يجيب عنها الطالب، وعادة ما يتم تقدير الإجابات بمقاييس عددية أو درجات تقديرية48.
- ويُعرّف **إجرائياً بأنه:** أداة تتضمن مجموعة من الأسئلة التي أعدت لقياس التحصيل الدراسي لطلبة الكليات في الاختبارات النهائية بكليات جامعة شقراء للفصل الدراسي الأول من العام الجامعي 1446هـ– 2024م.

**الاختبارات النهائية:**

هي الاختبارات التي تُعطى عند نهاية البرنامج أو المادة الدراسية، بغرض إعطاء تقييم شامل للطلاب فيما درسوه، وذلك بإعطائهم درجات أو شهادة تفيد بانجازهم دراسة البرنامج أو المقرر[17].

وتُعرّف **إجرائياً** بأنها: الأسئلة النهائية التي يعدها أعضاء هيئة التدريس بكليات جامعة شقراء في نهاية الفصل الدراسي بعد الإنتهاء من تدريس مقرر دراسي، بهدف تقييم مدى تحقق الأهداف التدريسية.

**الأسئلة المقالية:** هي الأسئلة التي يُعطى فيها الطالب حرية الإجابة عن الأسئلة بالطريقة التي يريدها، واختيار المادة وتنظيمها بالطريقة التي يراها، مع مراعاة الصحة في التعبير والدقة في استخدام المفاهيم والأفكار والمصطلحات والقواعد العلمية، والقدرة على العرض والشرح والتحليل والاستنباط وربط المعلومات بعضها ببعض[49].

**التحليل:** القدرة على تفكيك المادة إلى مكوناتها وأجزائها من أجل فهم بنيتها التنظيمية [24].

ويُعرّف إجرائيًا بأنه: عملية تجزئة محتويات أسئلة الاختبارات النهائية من خلال الوصف الموضوعي المنظم للمضمون الظاهر في ضوء تصنيف بلوم المعدل للأهداف التعليمية المعتمدة في هذه الدراسة، بغرض تحديد مدى تمثل الأسئلة لهذه المستويات وعرض تكرار التمثل والنسب المئوية والتعليق عليها.

**تصنيف بلوم (Bloom) المعدّل:**

هو تصنيف الأهداف التعليمية لبعد العمليات المعرفية المتمثلة بمستويات ست في صيغة الأفعال الآتية: (يتذكر، ويفهم، ويطبق، ويحلل، ويقوّم، يبدع).

**جامعة شقراء:** جامعة حكومية سعودية، تأسست عام 2009م، ويقع مقرها الرئيس في مدينة شقراء، وتضم (21) كلية موزعة في ثلاث محافظات هي: شقراء، والدوادمي، والقويعية، وعفيف.

**كلية التربية بشقراء:** هي إحدى كليات الجامعة التي تقع في محافظة شقراء. تقدم الكلية برامج أكاديمية متنوعة في مرحلتي البكالوريوس والدراسات العليا. ووتقتصر برامج البكالوريوس المقدمة حالياً على برنامج تعليم الطفولة المبكرة .

**منهج الدراسة:**

استخدمت الباحثة في الدراسة الحالية المنهج الوصفي بشقيه المسحي والتحليلي لمناسبته لموضوع الدراسة، ومرجعية الدراسة المستخدمة في عملية التحليل، فهي تصنيف بلوم المعدّل للأهداف التعليمية الخاص بالمستويات المعرفيّة.

**مجتمع الدراسة وعينتها:**

تكون مجتمع الدراسة وعينتها من جميع أسئلة الاختبارات النهائية التحريرية التي وضعها أعضاء هيئة التدريس بقسم تعليم الطفولة المبكرة بكلية التربية بشقراء بجامعة شقراء، للعام الجامعي1446هـ. وعددها (47) اختباراً. بواقع (21) للفصل الدراسي الأول، و(26) للفصل الدراسي الثاني.

**ثامنا: المواد والتحليلات:**

**الاختبارات التحصيلية:**

رغم تعدد أساليب وأدوات التقويم إلا أن الاختبارات التحصيلية تبقى من أكثر الأدوات التقويمية استخداماً في المجال التعليمي[43]. فهي وسيلة للكشف عن قدرات وكفاية المتعلمين، والتقدم الذي يحققه المتعلم في العملية التعليمية[3]، وأسلوباً لقياس عملية التحصيل الدراسي [17]، ونواتج التعلم في الجوانب العقلية والمعرفية[2]، ومقياساً يكاد يكون وحيداً يجاز به الطلاب من صف إلى آخر، ومن مرحلة تعليمية إلى أخرى[49]، ووسيلة تساعد في تصنيف الطلاب حسب قدراتهم وميولهم على التخصصات العلمية، واكتشاف قدراتهم ومواهبهم، وتنمية مهارات التفكير العليا لديهم[37].

ولكي يؤدي الاختبار التحصيلي دوره في تقويم نتائج التعلم المعرفية لا بد أن تتوفر فيه عدد من الخصائص والمواصفات، هي كالآتي: **الصدق**، أي قدرة الاختبار على قياس السمة التي وضع لقياسها.

**الثبات:** وهو الاتساق في نتائج الاختبار عند تطبيقه مرة أخرى، على العينة نفسها، وتحت الظروف نفسها

**الموضوعية:** وهو الاتفاق بين عدد المصححين في الدرجة الممنوحة لإجابة الطالب الواحد على بنود الاختبار.

**التمييز:** وهي قدرة الاختبار على إبراز الفروق بين الأفراد ذوي المستوى العالي والمتوسط والضعيف.

**الشمول:** وهو تمثيل الاختبار جميع العناصر المكونة لموضوع الاختبار.

**السهولة:** في تطبيق الاختبار وتقدير الدرجات، وتفسير النتائج.

**قلة التكلفة الاقتصادية** [31،23، 32].

### الاختبارات المقالية:

**الاختبارات التحصيلية** قد تكون اختبارات شفهية، أو اختبارات تحريرية، وتتنوع أسئلة الاختبارات التحصيلية التحريرية بتنوع الأهداف ومعايير تصنيفها؛ فهناك:

**أسئلة استدعاء**، يستدعي فيها الطالب الإجابة من الذاكرة، وهي الأسئلة المقالية، والإجابات القصيرة والإكمال، وهي أسئلة تفكير، وتتطلب مستويات عالية من التفكير كالتحليل والتركيب والتقويم والتنظيم والمقارنة والتنبؤ، وهي بحاجة إلى قدرة للإجابة عليها.

**أسئلة تعرُّف أو تذكُّر**، وتحتاج قليل من التفكير ومقدار عالٍ من التخمين، يُطلب فيها من الطالب التعرف على الإجابة فقط، ومنها أسئلة الاختيار من متعدد، وأسئلة الصواب والخطأ، وأسئلة المطابقة، وهذه عادة تسمى الأسئلة الموضوعية[44،48].

وتنقسم الأسئلة المقالية إلى نوعين، هما:

**الأسئلة المقالية المفتوحة**، وفيها يُترك للطالب الحرية الكاملة لتحديد مدى إجابته وشمولها، وتمتاز هذه الأسئلة بإظهار قدرة المتعلم على التحليل والترتيب والتنظيم والربط بين الأفكار، وإظهار القدرات الابتكارية في التعبير[49].

**الأسئلة المقالية المقيدة**. وفي هذا النوع من الأسئلة توضع حدود وقيود على الإجابات المطلوبة من حيث: المادة المطلوبة في السؤال، أو تحديد المساحة المخصصة له، أو تحديد نقاط معينة للإجابة[22]. وعلى الرغم من مزايا الأسئلة المقالية المقيدة كسهولة الإعداد وقياسها لمستويات الفهم والتطبيق والتحليل، إلا أنها لا تسمح للطالب بإظهار قدراته على التنسيق والترتيب والربط والابتكار والتركيب والتقويم[31،32].

وتتفوق الأسئلة المقالية على الأسئلة الموضوعية في قياس نتائج التعلم المهمة ومهارات التفكير العليا، كالقدرة على استرجاع المعلومات، والتحليل والمقارنة والتلخيص، والنقد، والاستنتاج، وربط النتائج بالأسباب، وربط المعاني والمفاهيم والألفاظ [31]. كما أنها تعمل على تقييم المهارات غير المعرفية كالمهارات العاطفية، والتواصل المكتوب، والقدرة على التفكير التباعدي والابتكار[23]. وبهذا فهي تُبيّن الطالب المستوعب للمادة والطالب الذي يعتمد على حفظ المادة صماً[48].

إلا أنه يُعاب على الأسئلة المقالية افتقارها إلى صدق المحتوى[40]، وتكلفتها العالية[48]، وتأثرها بذاتية المصحح[17،23]. وحاجتها لوقت طويل وجهد كبير لتصحيحها[12]، وعدم مراعاتها للفروق الفردية بين الطلاب، وعدم عنايتها بالجوانب الوجدانية والمهارية. وفي اختلاف فهم الطلاب للسؤال المقالي، وصعوبة تحديد معايير واضحة لأداء الطلاب عليها[2]، وتأثر درجة المتعلم سلباً بقدرته على الكتابة والتعبير، وعدد الأخطاء الإملائية واللغوية التي يرتكبها أثناء تحرير الإجابة [48،49].

### الاختبارات الموضوعية:

وُضِعت؛ لتلافي عيوب الاختبارات المقالية. ويرجع تسميتها بالموضوعية لموضوعية تصحيح إجاباتها، فلا تتدخل ذاتية المعلم في تصحيحها؛ لأن إجاباتها محددة، ومعروفة، ويتفق فيها المصححون على الدرجة التي تُعطى للمتعلم[23، 49].

وبهذا فهي تمثل مقياساً عادلاً يمكن الوثوق به لقياس معلومات الطالب في المادة الدراسية، وإلمامه بالمصطلحات، وتحليل الموضوعات، والربط بين العناصر، ومدى فهمه لها، وما وصل إليه من مهارات نتيجة تعلمها، ومدى انتفاعه بالمعلومات في حل المشكلات[31]. ومن أنواع الأسئلة الموضوعية:

**أسئلة الصح والخطأ** التي تتطلب من الطالب الحكم على العبارة إذا كانت صحيحة أو خاطئة، وتقيس هذه العبارات قدرة الطالب على الحفظ والاستظهار للحقائق والمفاهيم العلمية.

**أسئلة الاختيار من متعدد،** التي تتطلب من الطالب تحديد الإجابة الصحيحة من بين إجابات متعددة محتملة. وتقيس هذه الأسئلة اكتساب المعلومات وفهمها والاستفادة منها والقدرة على التمييز والمقارنة.

**أسئلة التكملة التي تتكون من جملة أو عبارة مفيدة محذوفاً منها المعلومة التي يُراد تذكر الطالب لها.**

**أسئلة المزاوجة،** وتتكون من عمودين متقابلين يضم كل منهما مجموعة من العناصر، ويُطلب من الطالب أن يربط العنصر في العمود الأول بما يناسب في العمود الثاني [2، 49].

وتتميز الأسئلة الموضوعية بقدرتها على تغطية مجالات الأهداف المعرفية والمهارية والوجدانية، وشمولها لأجزاء كبيرة من المقرر [2]، وانخفاض تكلفتها [48]، ومناسبتها للطلبة الذين يفتقرون لمهارات الكتابة والتعبير اللغوي[9]. كما يفضلها الأساتذة الجامعيين بحكم توفيرها للعدالة الكافية في الحكم على أداء

الطلاب، وتقليلها لطلبات إعادة التصحيح من قبل الطلبة الجامعيين [39].

إلا أنه يسجل على الأسئلة الموضوعية بعض المآخذ أو العيوب يتمثل في عجزها عن قياس العمليات العقلية العليا عند الطلاب وخاصة في مجال التحليل والتركيب والتقويم وحاجتها لوقت طويل للوضع أو الإعداد، وقياسها في كثير من الأحيان للحقائق والأسماء والمعارف السطحية [40]، وتشجيعها لهم على الحفظ والتخمين[21]، وسهولة الغش فيها[48]، واعتمادها على الجزئيات والتفاصيل دون الاهتمام بالنظرة الكلية المتكاملة لموضوعات المقرر، إضافة إلى صعوبة وضع أسئلتها، وحاجتها إلى خبرة ومهارة تربوية وعلمية ووقت كافٍ لإعدادها[2].

**هرم بلوم للأهداف المعرفية:**

يعدُّ تصنيف بلوم (Bloom, 1956) من أكثر التصنيفات شهرة في مجال الأهداف التعليمية، فقد أسهم في تطوير نظام الأهداف التعليمية مما ساعد المهتمين من التربويين والمعلمين في قياس العملية التعليمية من خلال صياغة الأهداف على شكل عبارات سلوكية واضحة [16].

ويتكون تصنيف بلوم من ثلاثة مجالات، هي: المجال المعرفي، والمجال الوجداني، والمجال النفسحركي، أما المجال المعرفي فهو المجال الأهم في مجال تطوير الاختبارات، وقد قسمه بلوم إلى ستة مجالات، هي: التذكر والفهم والتطبيق والتحليل والتقويم.

صنف "بلوم" المجال المعرفي إلى ستة مستويات عقلية بشكل هرمي تراكمي، بحيث تتدرج هذه المستويات أو العمليات العقلية من السهل إلى الصعب، ومن البسيط إلى المعقد، بحيث يتضمن المستوى الأصعب المستوى الأبسط منه بالضرورة، ويصبح المستوى الأسهل متطلباً سابقاً للمستوى الأعلى منه. مبيناً أن مساحة العمليات العقلية التي تقوم بها الذاكرة البشرية والمخصصة للتذكر تكون هي أكبر مساحة تحتلها قاعدة الهرم، ثم تتضاءل هذه المساحة تدريجياً صعوداً إلى أن تصل إلى قمة الهرم التي تمثل أعلى القدرات العقلية التي يقوم بها الدماغ البشري كعملية التقويم من وجهة نظر بلوم[26]. ولكن بعد ما يقارب الخمسة عقود من طرح بلوم لتصنيفه ظهر ما يُعرف بهرم بلوم المعدَّل. حيث أعيد تسمية ثلاث فئات والتعبير عن جميع الفئات بالأفعال لتصبح ست فئات، تمثل العمليات المعرفية التي من خلالها يتفاعل المفكرون مع المعرفة، ويعملون معها على الترتيب كالآتي: تذكر، وتفهم، وتطبق، وتحلل، وتقيم، وتبتكر[45].

**الفرق بين تصنيف بلوم الأصلي والمُعدَّل:**

يمكن تلخيص ثلاثة اختلافات مهمة بين هرم بلوم الأصلي والهرم المُعدل والمطور من قبل اندرسون وزملائها[20]، وهي:

- استخدام الأفعال بدلاً من الأسماء، وهذا يمنح قوة ويؤثر إيجابياً في الطريقة التي تظهر بها هذه القدرات كأشياء تطبق.
- أزيلت مهارة التركيب، وأضيفت مهارات الإبداع (الابتكار) كأعلى مستوى في التصنيف، وهذا يعطي دلالة على أن الإبداع هو المهارة المعرفية الأكثر تعقيداً، وهو أعلى شكل من أشكال التعلم، وذروة المهام المعرفية[45].

وبشكل عام فإن تصنيف بلوم ساعد الباحثين والمعلمين على فهم الطرق الأساسية التي تُكتسب بها المعارف والمهارات، وتجنب الجهود المكررة، أو الزائدة عن الحاجة عند تطوير اختبارات متعددة لقياس نفي الأهداف التعلمية[45].

ويحتوي بعد العمليات المعرفية في تصنيف بلوم المعدَّل على ست مهارات تتدرج من الأبسط إلى الأكثر تعقيداً، ذكرها الهدور[6] **،كالآتي:**

- يتذكر: وتعني التعرف على معلومات معينة واسترجاعها من الذاكرة طويلة المدى.
- يفهم: وتمثل القدرة على إدراك المعاني والتعبير عنها بلغة الفرد الخاصة.
- يطبق: وتعني قدرة المتعلم على تطبيق المعارف والمعلومات العلمية والطرق التي تمت دراستها وفهمها في مواقف جديدة.
- يحلل: وتعني القدرة على تجزئة الفكرة إلى عناصرها الفرعية، وتحديد العلاقات بين هذه العناصر والتمييز بينها، والقيام بتزيمها وإعادة ترتيبها وصياغتها.
- يُقوِّم: وتتعلق بإصدار حكم على موضوع، أو ظاهرة أو موقف معين وفق معايير أو محكات محددة.
- يبدع: وتمثل القدرةعلى تركيب وتوليف العناصر المختلفة مع بعضها لتكوين شيء جديد.

**تاسعا: إجراءات الدراسة:**

**الأدوات والمنهجية:**

**أدوات جمع البيانات:**

**استمارة تحليل المحتوى (الأسئلة المقالية):**

**-عدد الأسئلة المحل** 217 سؤالاً مقالياً من 47 اختباراً.

**-معايير التصنيف:**

**المستوى المعرفي (حسب بلوم المعدل):**

يُصنف كل سؤال بناءً على الفعل المستخدم (مثال: "اذكر" = تذكر، "حلل" = تحليل).

**المعايير الشكلية:**

- الوضوح (السؤال خالٍ من الغموض– نعم/لا).
- التنسيق (يحدد مساحة الإجابة – نعم/لا).
- السلامةاللغوية(خال من الأخطاء الإملائية/النحوية).

**طريقة التطبيق:**

تحليل كل سؤال بشكل مستقل من قبل (الباحثة + باحث آخر (أستاذ دكتور) لمراجعة التحليل).

تسجيل النتائج في جدول إكسل لاحتساب التكرارات والنسب.

**المقابلات مع أعضاء هيئة التدريس:**

**عدد المشاركين** :12 عضواً (2 أستاذ، 2 أستاذ مشارك، أستاذ مساعد، 3 محاضر).

**معايير الاختيار**:

- خبرة لا تقل عن 3 سنوات في التدريس الجامعي.
- مشاركتهم في وضع أسئلة الاختبارات سابقاً.

**نوع المقابلة:** شبه مقننة (أسئلة مفتوحة + متابعة).

**طريقة اختيار العينة:**

**عينة الأسئلة:**

**طريقة الاختيار** :عشوائية طبقية (بعد تصنيف الاختبارات حسب المستوى الدراسي: السنة الأولى إلى الرابعة).

**التبرير:** ضمان تمثيل جميع المستويات الأكاديمية.

**عينة أعضاء هيئة التدريس:**

**طريقة الاختيار**: عمدية (بعد استبعاد من ليس لهم خبرة في وضع الاختبارات).

**سبب اختيار كلية التربية بشقراء:**

تعتمد على نظام الاختبارات التحريرية بنسبة 100.%

تشهد نقاشات مستمرة حول تطوير التقويم (مما يجعلها بيئة غنية للدراسة).

**تعزيز الثبات والصدق:**

**الثبات (الموثوقية):**

**اختبار كرونباخ ألفا:**

تم تطبيقه على استمارة التحليل لقياس اتساقها الداخلي.

و**النتيجة** ($\alpha = 0.87$) مما يدل على ثبات عالٍ.

**الاتفاق بين المحللين:**

نسبة الاتفاق = 91% (حسب معادلة كوبر).

**خطوات التحقق:**

1. تدريب المحلل الثاني على تصنيف بلوم المعدل.
2. تحليل عينة عشوائية (15% من الأسئلة) بشكل مستقل.

**الصدق (الصلاحية): صدق المحتوى:**

- عرضت الأداة على 5 خبراء في المناهج وطرق التدريس.
- **نسبة الاتفاق** :90 %على تصنيف الفئات.

**صدق البناء:**

ربط نتائج التحليل بنتائج المقابلات (مثل: إذا ذكر المعلمون صعوبة وضع أسئلة إبداعية، يجب أن تظهر النتائج انخفاضاً في أسئلة مستوى "الإبداع").

**تجنب التحيز:**

**التحيز في التحليل:**

- استخدام محللين مستقلين (الباحث + زميل متخصص (أستاذ دكتور) للدقة.
- إخفاء أسماء واضعي الأسئلة أثناء التحليل.

**التحيز في المقابلات:**

- تسجيل المقابلات صوتياً وكتابة النصوص حرفياً.
- استخدام أسئلة محايدة (مثال: بدلاً من "هل تعتقد أن الأسئلة المقالية صعبة؟"، اسأل: "كيف تصف تجربتك في وضع الأسئلة المقالية؟").

**منهج الدراسة:**

استخدمت الباحثة في الدراسة الحالي المنهج الوصفي بشقيه المسحي والتحليلي لمناسبته لموضوع الدراسة. ومرجعية الدراسة المستخدمة في عملية التحليل فهي تصنيف بلوم المعدّل للأهداف التعليمية الخاص بالمستويات المعرفيّة.

**مجتمع الدراسة وعينتها:**

تكون مجتمع الدراسة وعينتها من جميع أسئلة الاختبارات النهائية التحريرية التي وضعها أعضاء هيئة التدريس بكلية التربية بشقراء لمقررات قسم تعليم الطفولة المبكرة، خلال العام الجامعي 1446هـ، بفصليه الأول والثاني، وعددها (47) نموذجاً.

كما اختارت الباحثة عينة مقصودة تكونت من عدد (13) عضواً من أعضاء هيئة التدريس القائمين على تدريس مقررات قسم

تعليم الطفولة المبكرة، للاستجابة عن أسئلة المقابلة كالآتي: (2) أستاذ، (3) أستاذ مشارك، (5) أستاذ مساعد، و(3) محاضر.

**أدوات الدراسة:**

**استمارة التحليل:**

تم إعداد استمارة تحليل خاصة لاستخدمها معياراً لتحليل أسئلة الاختبارات النهائية، حيث استندت إلى عدد من الدراسات والمراجع التربوية ذات الصلة بأهداف الدراسة، منها دراسة الشيخ[15]، ودراسة الهدور[6]، ودراسة شاكونة[45]، ودراسة أمين[18].

**صدق الاستمارة:**

للتأكد من صدق الاستمارة عُرضت على مجموعة من الخبراء والمختصين في المناهج وطرق التدريس. بلغت نسبة اتفاقهم على صدق الاستمارة (91%). وبعد تعديل بنود البطاقة أصبحت البطاقة في صورتها النهائية تحتوي على (23) بنداً موزعة على المحاور الآتية:

1. محور التنوع: 13 بنداً
2. محور المعايير الشكلية: 4 بنود
3. محور مهارات التفكير: 6 بنود

**ثبات الاستمارة:**

تم استخدام طريقة التطبيق وإعادة التطبيق للتأكد من قدرة الاستمارة على إخراج نفس النتائج لو أُعيد تطبيقها عل العينة نفسها، وفي الظروف نفسها، وفق الخطوات الآتية:

1. اختيار عينة عشوائية من أسئلة الاختبارات النهائية، بلغت نسبتها (15%)، من عينة أوراق الاختبارات.
2. تحليل أسئلة الاختبارات من الباحثة للمرة الأولى، ثم تحليلها مرة ثانية بعد مرور أسبوعين باستعمال نفس الأداة.
3. تم عرض أسئلة الاختبارات على محلل آخر يمتلك الكفاية والقدرة على التحليل بعد توضيح أهداف الدراسة وطريقة التحليل.
4. تجميع نتائج التحليل التي أجرتها الباحثة والمحلل الآخر على شكل تكرارات ونسب مئوية وحساب نسبة الاتفاق بينهما.
5. استخراج نسبة الاتفاق، وعدد مرات التحليل باستخدام معادلة كوبر:
6. معامل الاتفاق = عدد مرات الاتفاق /عدد مرات الاتفاق + عدد مرات الاختلاف * 100
7. التأكد من أن معامل الاتفاق في كل تحليل على درج عالية من الثبات، كما هو مبين في جدول (1)

**الجدول (1)** ثبات تحليل أسئلة الاختبارات النهائية باستخدام معادلة كوبر:

| نوع الاتفاق | الإجراءات | معامل الاتفاق |
|---|---|---|
| الاتفاق عبر الزمن | بين الباحث ونفسه بفاصل زمني | %93 |
| الاتفاق بين المحللين | بين الباحث ومحلل آخر | %89 |
| المتوسط | | %91 |

يتبين من الجدول (1) أن قيمة معامل الثبات انحصرت بين (89% – 93 %). إذ كانت قيمية الثبات بين الباحث ونفسه عبر الزمن (93%)، في حين كانت قيمة معامل الثبات بين الباحث وزميله هي (89%) وبمتوسط مقداره (91%). وتدل هذه النسب على درجة عالية من الثبات تمكن من الوثوق بدقة بنتائج التحليل.

**فئات التحليل:** تم اعتماد مستويات تصنيف بلوم المعدل كفئات رئيسية للتحليل في هذه الدراسة، في حين تم اعتماد العمليات الدالة على هذه المستويات كفئات تحليل فرعية.

**وحدة التحليل:** اعتبار كل فقرة من فقرات أسئلة الاختبارات كوحدة تحليل، وذلك من أجل تحليل تلك الأسئلة ليظهر من خلالها تكرار المستوى المراد تحليل المحتوى في ضوئه.

**قواعد التحليل:**

1. اعتماد جملة السؤال، ووحدة السياق كوحدات للتحليل.
2. إذا احتوى السؤال الرئيس على أكثر من فقرة تعامل كل فقرة على أنها سؤال مستقل.
3. في حال احتواء السؤال على عدة متطلبات معطوفة بعضها على البعض الآخر، فإنه يُعامل كل مطلوب على أنه سؤالٌ مستقلٌ.
4. إذا كان هناك أسئلة اختيارية يتم التعامل معها بوصفها أسئلة إجبارية، ويتم احتساب الدرجة الكلية للاختبار من خلال جمع درجات كل الأسئلة الاختيارية والإجبارية.

**إجراءات الدراسة:**

1. جمع أسئلة الاختبارات النهائية من القسم المحدد.
2. إعداد بطاقة التحليل وفق تصنيف بلوم المعد للأهداف المعرفية.

3. الاتفاق مع المحلل الآخر على قواعد التحليل.
4. تطبيق بطاقة التحليل في صورتها النهائية بقراءة الأسئلة قراءة متأنية، لتحديد نوع السؤال، وللتعرف على المستويات المعرفية التي تمثلت بالأسئلة.
5. رصد نتائج التحليل وتصنيفها في جداول إلكترونية.
6. استخراج تكرارات الأسئلة حسب فئات التحليل، ثم استخراج النسب المئوية لكل فئة.
7. إجراء المعالجات الإحصائية.
8. إعداد بطاقة المقابلة.
9. إجراء المقابلات مع أعضاء هيئة التدريس.
10. مناقشة النتائج وتقديم التوصيات والمقترحات.

**بطاقة مقابلة:**

استخدمت الباحثة المقابلة شبه المقننة، التي اعتمدت على سؤال جميع أفراد العينة نفس الأسئلة المعّدة مسبقاً وفقاً لدليل محدد، بنفس الطريقة والترتيب، مع طرح أسئلة للاستيضاح دون الخروج عن جوهر السؤال.

ولتحديد الأسئلة والمحاور التي سيتم تنفيذ مقابلة أفراد العينة وسؤالهم عنها، قامت الباحثة بطرح سؤال مفتوح على (5) من أعضاء هيئة التدريس، تم فيه سؤالهم عن أسباب ارتفاع نسبة تضمين الأسئلة الموضوعية في أسئلة الاختبارات، مقابل ضعف تضمين الأسئلة المقالية، وكذلك أسباب ضعف تضمين الأسئلة المقالية ذات مستويات التفكير العليا في الاختبارات النهائية، وعن الإجراءات المقترحة لتحسين الوضع.

**صدق الأداة:**

للتحقق من صدق الأداة قامت الباحثة بعرض دليل المقابلة على (5) من المحكمين من تخصص المناهج وطرق التدريس، لمعرفة رأيهم حول ما ورد فيها من أسئلة رئيسة ومحاور فرعية. تم تعديل الدليل وفقاً لآراء المحكمين، وأصبح دليل المقابلة متضمناً ثلاثة أسئلة رئيسة. كما تم تحري صدق الأداة من خلال تسجيل الإجابات، وكتابة الملاحظات آلياً ويدوياً، والتعامل مع جميع الإجابات التي وردت بشفافية، وموضوعية.

**ثبات الأداة:**

تم التحقق من ثبات الأداة من خلال طرح الأسئلة بصيغ مختلفة، للتأكد من تكرار نفس الإجابة، والتحقق من مدى اتساق إجابات أفراد العينة، ومنطقيتها، وتمثيلها للواقع تمثيلاً حقيقياً.

**إجراءات المقابلة:**

1. إعداد دليل المقابلة، ونموذج تسجيل الإجابات اليدوي.
2. التواصل مع أفراد العينة التي تم اختيارها بشكل عشوائي، وتحديد الزمان والمكان المناسب لهم لإجراء المقابلة.
3. تسجيل المقابلات يدوياً في نموذج الإجابة، وبشكل آلي بالهاتف الجوال بعد أخذ موافقتهم.
4. طرح الأسئلة على أفراد العينة بنفس الترتيب، وبحيادية تامة.
5. تفريغ النماذج اليدوية والتسجيلات الآلية.
6. حساب نسب الاتفاق والاختلاف في استجابات أفراد العينة من خلال حساب التكرارات والنسب المئوية.
7. تحليل النتائج وتفسيرها ومناقشتها.

**توقيت المقابلة:**

بلغت الساعات الكلية لتنفيذ المقابلة (6) ساعات، بمعدل 30 دقيقة لكل فرد، تخللها بعض التوقف والمعاودة حسب الظروف.

**التحليل والنتائج ومناقشتها: التحليل الإحصائي**

**البرامج المستخدمة:**

- SPSS لتحليل التكرارات والنسب المئوية.
- اختبار **كاي–سكوير** (Chi-square)لمقارنة توزيع الأسئلة بين المستويات المعرفية.

**أولاً: إجابة السؤال الأول الذي نصه: ما مدى تضمين أسئلة الاختبارات النهائية للأسئلة المقالية وتنوعها بكلية التربية بجامعة شقراء؟** وللإجابة عن السؤال تم حساب التكرارات والنسب المئوية للأسئلة الموضوعية والمقاليه، فكانت كما هي موضحة في الجدول (2):

**الجدول (2)** توزيع الأسئلة حسب تنوعها (مقالية وموضوعية):

| م | نوع السؤال | التكرار | النسبة |
|---|---|---|---|
| 1 | أسئلة مقالية فقط | صفر | صفر |
| 2 | أسئلة موضوعية فقط | 10 | %21 |
| 3 | أسئلة مقالية موضوعية | 37 | %79 |
| 4 | الأسئلة الموضوعية | 47 | %100 |

| | | | |
|---|---|---|---|
| **5** | **الصواب والخطأ** | 40 | %85 |
| **6** | **الاختيار من متعدد** | 39 | %53 |
| **7** | **المطابقة** | 22 | %47 |
| **8** | **إكمال الفراغ** | 14 | %30 |

يتضح من جدول (2) الأسئلة التي تم تحليلها، وعددها (217) سؤالاً، قد تنوعت بين مقالية وموضوعية. حظيت الأسئلة الموضوعية بالنصيب الأكبر من أسئلة الامتحان، حيث توفرت الأسئلة الموضوعية بجميع أوراق الاختبارات عينة الدراسة بنسبة (100%)، في حين توفرت الأسئلة المقالية بنسبة (79%) من أوراق الاختبار. فلم تنفرد أي من أوراق الاختبارات بأسئلة مقالية فقط، في حين اقتصرت ما نسبته (21%) من الاختبارات على أسئلة موضوعية فقط.

تتفق نتيجة هذا السؤال مع نتيجة دراسة الصلاحي [13]، ودراسة غليون [30]. وتتوافق مع نتائج دراسة المزوغي[11]، ودراسة الشريفي[14]، ودراسة سعيد وجبار[41].

كما يوضح جدول (2) أيضاً تنوع الأسئلة الموضوعية في هذه الاختبارات، تفوّقت فيها أسئلة الصواب والخطأ على بقية الأنواع، وجاءت بنسبة (85%)، تلتها أسئلة الاختيار من متعدد بنسبة (53%)، وجاءت أسئلة المطابقة بنسبة (47%)، في حين جاءت أسئلة إكمال الفراغ بأقل نسبة، بلغت (30%). وربما يُعزى هذا إلى امتلاك أعضاء هيئة التدريس لمستوى جيد من الكفايات المتعلقة بإعداد الأسئلة الموضوعية. وتشير هذه النتيجة إلى أن مؤشر تنوع الأسئلة بين موضوعية ومقالية جاء بدرجة عالية. وتُعزى هذه النتيجة إلى انحياز عدد من أعضاء هيئة التدريس القائمين على تدريس المقررات بقسم تعليم الطفولة المبكرة بكلية التربية للأسئلة الموضوعية، واستغنائهم عن الأسئلة المقالية بشكل كامل. وربما يعزى هذا الانحياز لسهولة إعداد الأسئلة الموضوعية، وسرعة تصحيحها وموضوعيتها المرتفعة مقارنة بالأسئلة المقالية، دون إدراك إلى أن اقتصار الاختبار على الأسئلة الموضوعية لا تعكس قدرات الطالب التعبيرية، وتنظيم الأفكار والربط بينها، والتحليل وتطبيق القوانين والقواعد على موقف جديدة. أو ربما لعدم اهتمام عدد من أعضاء هيئة التدريس بالتمحيص القرائي المتعمق للمحتوى الدراسي ونقده مثل اهتمامهم بالمفاهيم الأساسية للمحتوى الذي يدفعهم إلى التركيز على الأسئلة الموضوعية دون المقالية

**أما تنوع أسئلة الاختبارات النهائية بكلية التربية بجامعة شقراء، فكانت كما في الجدول (3):**

**الجدول (3)** توزيع الأسئلة حسب عدد أنواع الأسئلة:

| م | محتوى ورقة الاختبار | التكرار | النسبة |
|---|---|---|---|
| **1** | نوع واحد من الأسئلة | 1 | %2 |
| **2** | نوعين من الأسئلة | 9 | %19 |
| **3** | ثلاثة أنواع من الأسئلة | 19 | %40 |
| **4** | أربعة أنواع من الأسئلة | 14 | %30 |
| **5** | خمسة أنواع من الأسئلة | 4 | %9 |
| **المجموع** | | **47** | **%100** |

يتضح من الجدول (3) أن أوراق الاختبار التي تم تحليلها جاءت أسئلتها متنوعة بدرجة ضعيفة. فقد احتوى عدد (4) أوراق اختبار على خمسة أنواع من الأسئلة بنسبة (9%). في حين جاءت الأوراق التي احتوت على أربعة أنواع بنسبة (30%)، ونسبة (40%) على ثلاثة أنواع. وقد تُشير هذه النتيجة إلى ضعف إدراك أهمية تنوع أسئلة الاختبارات لقياس قدرات متنوعة لدى الطلاب، وما قد يشكله الاعتماد إلى نوعين أو ثلاثة أنواع من الأسئلة من خلل في موضوعية الأسئلة، ومصداقيتها اللذين يعتبران من أهم معايير الاختبارات الجيدة. تتفق هذه النتيجة مع نتيجة دراسة لعيبي والمياحي[34]، التي أظهرت نتائجها عدم تحقيق الأسئلة لجميع معايير جودة أسئلة الاختبارات، وعدم تنوعها.

كما تتفق مع نتائج دراسة أمين [19] التي توصلت إلى عدم تنوع أسئلة الدراسة وعدم شموليتها.

وتختلف هذه النتيجة مع ما توصلت إليه دراسة الصلاحي [13]، ودراسة غليون[30].

**ثانياً: إجابة السؤال الثاني الذي نصه: ما مدى التزام أعضاء هيئة التدريس بالمعايير الشكلية لصياغة الأسئلة المقالية في الاختبارات النهائية بكلية التربية بشقراء؟**

للإجابة عن السؤال تم حساب التكرارات والنسب المئوية لمدى تحقق المعايير الشكلية للصياغة، على النحو الآتي:

- **الدقة والوضوح:**

**الجدول (4)** معيار الدقة والوضوح:

| م | المعيار | التكرار | النسبة |
|---|---|---|---|
| 1 | صياغة جيدة | 79 | %36 |
| 2 | صياغة غير جيدة | 138 | %64 |
| 3 | صياغة غير واضحة | 53 | %24 |
| 4 | صياغة غير محددة | 41 | %19 |
| 5 | صياغة مركبة | 22 | %10 |
| 6 | صياغة غير مباشرة | 15 | %7 |

يوضح الجدول (4) أن معيار الدقة والوضوح تحقق بنسبة ضعيفة في صياغة الأسئلة المقالية في الاختبارات النهائية، بنسبة بلغت (36%). فجاءت الأسئلة ذات الصياغة غير الجيدة بنسبة (64)%. وجاءت صياغة عدد (53) سؤالا بنسبة (24%) غير واضحة، و(41) سؤالا بنسبة (19%) غير محددة، و(22) سؤالاً بنسبة (10%) مركبة، و(15) سؤالاً بنسبة (7%) غير مباشرة.

وقد تُعزى هذه النتيجة إلى ضعف قدرات عدد من أعضاء هيئة التدريس على صياغة أسئلة مقالية دقيقة وواضحة ومحددة بعيدة عن الغموض والتأويلات والتفسيرات المشتتة لتفادي الفهم الخاطئ للأسئلة، أو إلى عدم إدراكهم أن الصياغة المعقدة أو المربكة تؤدي إلى إرباك الطلاب، ودفعهم لتقديم إجابات بعيدة عن المطلوب من السؤال، علاوة على زيادة عامل القلق والتوتر أثناء تأدية الاختبار، وأن رسوب الطلاب قد يعود في كثير من الأحيان إلى ضعف كفاية ورقة الاختبار أكثر من ضعف مستوى الطالب نفسه.

**السلامة النحوية واللغوية والإملائية:**

**الجدول (5)** معيار السلامة النحوية واللغوية والإملائية:

| م | المعيار | التكرار | النسبة |
|---|---|---|---|
| 1 | صياغة سليمة | 80 | %37 |
| 2 | صياغة غير سليمة | 137 | %63 |
| 3 | أخطاء بعلامات الترقيم | 82 | %38 |
| 4 | خطأ باستخدام (؟) | 54 | %25 |
| 5 | أخطاء إملائية | 40 | %18 |
| 6 | خطأ باستخدام الهمزة | 27 | %12 |
| 7 | خطأ بنظام الأرقام | 24 | %11 |
| 8 | أخطاء طباعية | 7 | %3 |
| 9 | أخطاء نحوية | 7 | %3 |

يوضح الجدول (5) أن معيار السلامة اللغوية والنحوية والإملاية تحقق بنسبة ضعيفة في صياغة الأسئلة المقالية في الاختبارات النهائية، بنسبة بلغت (38%). فجاءت الأسئلة ذات الصياغة غير السليمة بنسبة (63)%. وجاء عدد (82) سؤالاً بنسبة (38%) بخطأ بعلامات الترقيم، سجل منها عدد (54) سؤالاً بنسبة (25%) خطأ باستخدام علامة الاستفهام (؟) مكان النقطة (0) في الأسئلة المقالية التي تبدأ بفعل أمر. وسجل عدد (40) سؤالاً بنسبة (18%) خطأ إملائياً، من (27) بنسبة (12%) بكتابة همزة الوصل وهمزة القطع. كما جاء عدد (24) سؤالاً بنظام ترقيم غير صحيح.

وتُعزى هذه النتيجة إلى ضعف عدد من أعضاء هيئة التدريس في المهارات اللغوية، وإلى عدم إدراكهم لتأثير السلامة اللغوية والنحوية والإملائية على فهم الطلاب للأسئلة، وعلى قدراتهم على تقديم إجابات صحيحة، وعلى تأثيرها في مهارات الطلاب اللغوية واحتمالية نقلها لطلابهم بعد أن ينضموا لمهنة التعليم.

**ج. بدء الأسئلة بأفعال سلوكية:**

**الجدول (6)** معيار بدء الأسئلة بأفعال سلوكية:

| م | المعيار | التكرار | النسبة |
|---|---|---|---|
| 1 | أسئلة تبدأ بأفعال سلوكية | 198 | %91 |
| 2 | أسئلة لا تبدأ بأفعال سلوكية | 19 | %9 |
| 3 | أكمل | 54 | %25 |

| | | | |
|---|---|---|---|
| 4 | عدد | 51 | %23.5 |
| 5 | اذكر | 31 | %14 |
| 6 | علل | 14 | %6.5 |
| 7 | اكتب | 10 | %5 |
| 8 | قارن | 8 | %4 |
| 9 | أجب | 8 | %4 |
| 10 | عرف | 7 | %3 |
| 11 | وضح | 5 | %2 |
| 12 | احسب | 4 | %2 |
| 13 | أوجد | 2 | %1 |
| 14 | تحدث | 2 | %1 |
| 15 | اشرح | 1 | %0.4 |
| 16 | ناقش | 1 | %0.4 |
| 17 | ما | 11 | %5 |
| 18 | كيف | 5 | %2 |
| 19 | كم | 1 | %0.4 |
| 20 | هل | 1 | %0.4 |
| 21 | متى | 1 | %0.4 |

يبين الجدول (6) أن معيار بدء الأسئلة المقالية بأفعال سلوكية قابلة للملاحظة والقياس قد تحقق بنسبة عالية بلغت (91%). جاء الفعل (أكمل) بنسبة (25%)، يليه الفعل (عدد) بنسبة جاء الفعل (أكمل) بنسبة (25%)، يليه الفعل (عدد) بنسبة (23.5%)، ثم الفعل (اذكر) بنسبة (14%). في حين جاء الفعل (اشرح) و(ناقش) بنسبة ضعيفة بلغت (0.4%). كما جاءت الأسئلة التي تبدأ بأدوات استفهام جاء أعلاها نسبة الأداة (ما) بنسبة (5%)، ولعل هذه النتيجة تُعزى إلى وقوع معظم الأسئلة المقالية التي تم تحليلها في مستوى التذكر الذي يتطلب الحفظ والاسترجاع من الذاكرة، أو الفهم دون التفكير أو التحليل.

**تحديد مسافة مناسبة للإجابة:**

**الجدول (7)** معيار تحديد مسافة مناسبة للإجابة:

| م | المعيار | التكرار | النسبة |
|---|---|---|---|
| 1 | مسافة محددة للإجابة | 194 | %89 |
| 2 | بدون تحديد | 23 | %11 |
| 3 | مسافة بعدد أسطر | 77 | %35 |
| 4 | مسافة بعدد نقاط | 71 | %33 |
| 5 | مسافة بفراغات بين الجمل | 46 | %21 |

**يبين الجدول (7) أن معيار تحديد مسافة مناسبة للإجابة قد تحقق بنسبة متوسطة بلغت (89%). في حين** جاء عدد (33) سؤالا بنسبة (11%) بدون تحديد، إما بترك مساحة بيضاء في الورقة، أو بطلب الكتابة خلف الورقة. جاءت المسافة بعدد أسطر متوافق مع المطلوب من السؤال بأعلى نسبة تكرار بلغت (35%)، يليها مسافة بعدد نقاط مرقمة بأرقام أو بشرط بنسبة (33%)، في حين جاءت المسافة بفراغات بين الجمل في الأسئلة التي تبدأ بفعل (أكملي) بنسبة (21%). ويشير هذا إلى اهتمام معظم أعضاء هيئة التدريس بتوجيه الطلاب للإجابة المطلوبة بتحديد المسافة المناسبة للإجابة. وربما يُعزى ذلك إلى كون معظم الأسئلة من الأسئلة المقالية المقيدة وليست المفتوحة. على عملية التقويم وعملية التعليم. والجدير بالذكر هنا أن الفعل (أكمل) جاء بصيغة لا تناسب الأسئلة المقالية، مثل :(عدد)، أو (اذكر)، جاء بصيغة مثل: (أكملي الفراغات فيما يأتي)، أو (أكملي العبارات التالية). مما يشير إلى ضعف في مهارات صياغة أسئلة الاختبار بشكل عام، رغم كونها مسؤولية كبيرة وإن عدم اتصافها بالمعايير التي تناسب أهميتها قد تصبح أداة مضللة تقدم بيانات غير دقيقة، قد تقود إلى أحكام غير صادقة، وتتوافق هذه النتيجة مع نتائج دراسة غايتس وبلو[29].

**ثالثاً: إجابة السؤال الثالث الذي نصه: ما نسب تمثيل الأسئلة المقالية في الاختبارات النهائية لمستويات الأهداف المعرفية وفق تصنيف بلوم المعدّل؟** للإجابة عن هذا السؤال تم تحليل أسئلة الاختبارات عينة الدراسة، للتعرف على مدى توافر مستويات أهداف المجال المعرفي حسب تصنيف بلوم المعدل في الأسئلة عينة الدراسة. كما تم توزيع مستويات أهداف المجال المعرفي على أسئلة الاختبارات، واحتساب التكرارات والنسب المئوية لكل تصنيف.

**الجدول (8)** التكرار والنسب المئوية للمستويات المعرفية حسب تصنيف بلوم المعدّل:

| المستوى المعرفي | عدد الأسئلة | النسبة المئوية | أمثلة على الأفعال المستخدمة |
|---|---|---|---|
| **التذكر** | **175** | **% 81** | **عرّف، اذكر، أكمِل** |
| **الفهم** | **23** | **%10,5** | **فسّر، قارن، لخّص** |
| **التطبيق** | **18** | **% 8** | **طبّق، استخدم، حلّل** |
| **التحليل** | **1** | **%0,5** | **حلّل، ناقش، فرّق** |
| **التقويم** | **0** | **%0** | **قوّم، انقد، ارجع** |
| **الإبداع** | **0** | **%0** | **صمّم، ابتكر، طوّر** |

ويبين الجدول (8) أن نسبة (81%) من الأسئلة المقالية تقيس نسبة التذكر، ونسبة (10.5%) تقيس مستوى الفهم، ونسبة (8%) تقيس مستوى التطبيق، ونسبة (0.5%) تقيس نسبة التحليل. في حين جاءت مستويات التقويم والإبداع بنسبة (صفر). وتدل هذه النتيجة على أن الأسئلة المقالية تُهمل قياس المستويات المعرفية العليا لدى الطلاب، وتعطي مؤشرات سلبية عن نوعية مخرجات التعلم وطبيعتها. كما قد يشير إلى غلبة الطابع التقليدي في التدريس الذي يهدف إلى التركيز على الحفظ والاسترجاع فقط، وتهمل إثارة التفكير وتنمية قدرات الطلاب على التخيل والتفسير واتخاذ القرار وحل المشكلات.

**الجدول (9)** اختبار (T-test) لمقارنة متوسطات توزيع الأسئلة:

| المتغير | المجموعة 1 (أسئلة التذكر/الفهم) | المجموعة 2 (أسئلة التحليل/الإبداع) | (p-value) قيمة | النتيجة |
|---|---|---|---|---|
| متوسط عدد الأسئلة لكل اختبار | 12،3 | 1،2 | 0،001 | فرق ذو دلالة إحصائية |

ويتبين من جدول (9) أن هناك فرقا ذا دلالة إحصائية ($p < 0.05$) بين عدد الأسئلة التي تقيس المستويات الدنيا (التذكر/الفهم) والمستويات العليا (التحليل/الإبداع). مما يدل على أن التركيز على المستويات الدنيا في الاختبارات ليس عشوائيًا، بل يعكس تحيزًا منهجيًا.

وربما تُعزى هذه النتيجة إلى عدم تقيد أعضاء هيئة التدريس بتحديد الأهداف السلوكية عند التدريس وإعداد جدول مواصفات عند إعداد الاختبارات لتغطية جميع المستويات التي يغطيها محتوى المقررات، أو إلى أن تنمية المستويات العليا من التفكير لم تكن هدفاً أساسياً لمقرراتهم، أو إلى تعوّد معظم أعضاء هيئة التدريس على إعداد أسئلة تقيس قدرة الطالب على حفظ المعلومات، أو إلى تقيد الاختبار النهائي بوقت محدد، وهو ساعتان، وحاجة المستويات العليا في التصنيف لوقت إضافي طويل، مما يضطرهم لصياغة أسئلة تناسب الوقت المتاح.

وقد تُعزى هذه النتيجة إلى أسباب أخرى كالثقافة التعليمية السائدة في مراحل التعليم العام، والمرحلة الجامعية التي تركز في مجملها على الحفظ والاسترجاع، حتى أن معظم الاختبارات الوطنية، تعتمد بشكل كلي على الأسئلة الموضوعية. ولا شك أن نظام "التلقين" في المدارس والجامعات يقلل من فرص ممارسة مهارات التفكير العليا أمام الطلاب. كما أن المناهج المعتمدة لا تشجع معظمها على التفكير الإبداعي، مع غياب الحوافز للمعلمين ولأعضاء هيئة التدريس لتطوير أسئلة إبداعية في الاختبارات التي يعدونها.

تتوافق هذه النتيجة مع نتائج دراسة كل من الصلاحي[13] ، ودراسة نايف وعبد الكاظم[38] التي توصلت نتائجها إلى أن الاختبارات تركز على المستويات الدنيا من تصنيف بلوم وتهمل المستويات العليا، كما أن الأسئلة لا تتمتع بالتنوع، وتتفق مع دراسة عبده[1] ودراسة لعيبي والمياحي[34] التي أظهرت نتائجهما عدم تحقيق الأسئلة لجميع معايير جودة أسئلة الاختبارات، وعدم تنوعها، وأن معظمها اقتصر على المستويات الدنيا من تصنيف بلوم المعدّل، ودراسة سعيد وجبار[41] ، ودراسة الجابري وحسين[8] ، ودراسة قادر وزملائه[28] ، ودراسة أمين[19] ، ودراسة زهور وزملائه [50]، ودراسة علي وزملائه[7] التي توصلت نتائجها جميعا إلى أن الاختبارات تركز على المستويات الدنيا من مستويات بلوم المعدل.

**رابعاً: إجابة السؤال الرابع الذي نصه: ما هي أسباب ضعف نسبة تضمين الأسئلة المقالية ومستوياتها المعرفية العليا في الاختبارات النهائية بكلية التربية بجامعة شقراء من وجهة نظر أعضاء هيئة التدريس بكلية التربية بشقراء؟**

للإجابة عن هذا السؤال تم حساب التكرارات والنسب المئوية ومناقشة النتائج وتحليلها، على النحو الآتي:

**صعوبة إعداد وتصحيح الأسئلة المقالية:** أجمع جميع أفراد العينة وعددهم (13) بنسبة (100%) على أن أحد أهم أسباب ضعف نسبة تضمين الأسئلة المقالية أو الاستغناء عنها تماماً في أسئلة الاختبارات النهائية هو صعوبة تصحيحها، وعدم ضمان الدقة والعدالة في التصحيح كالأسئلة الموضوعية، يقول أحدهم: (الأسئلة المقالية صعبة في التصحيح، وتحتاج قراءة وإعادة قراءة لإجابات الطالبات)، ويقول آخر: (لا مقارنة بين سهولة تصحيح الأسئلة الموضوعية مقابل صعوبة الأسئلة المقالية، في دقائق أضع الأسئلة الموضوعية، وفي دقائق انتهي من تصحيحها). ويقول آخر: (لا مجال للأسئلة المقالية مع الأعداد الكبيرة من الطالبات، الأسئلة الموضوعية سهلة وسريعة التصحيح بالنسبة للمعلم، وسهلة وسريعة في الإجابة بالنسبة للطالب، وينهي معظم الطلبة الإجابة عنها في أقل من نصف الوقت المحدد للاختبار). وقد تكون صعوبة التصحيح نتيجة ثانوية لنقص التدريب، وليس سبباً جذرياً بحد ذاته. فنقص التدريب على كتابة الأسئلة التي تتطلب مهارات عليا، وعلى إعداد معايير واضحة للتصحيح، وعدم اشتراط اجتياز دورات تدريبية معتمدة في تصميم الاختبارات كجزء من سياسات التوظيف في وزارة التعليم بالمملكة قد يخلق خوفا لدى الكثير من أعضاء هيئة التدريس بالجامعات من إعداد هذا النوع من الأسئلة وتصحيحها بموضوعية ودقة.

**ضعف المستوى العام للطالبات:** أجمع جميع أفراد العينة وعددهم (13) بنسبة (100%) على أن ضعف مستوى تحصيل الطالبات العام في المقررات الدراسية، وعدم امتلاك معظمهن لأساسات قوية في اللغة العربية، وتحديداً في مهارات التعبير الكتابي والإنشاء، والإملاء، إضافة إلى رداءة الخط. فقد أجمعوا على أن معظم الطالبات لا يمتلكن المهارات اللازمة لتنظيم الأفكار وصياغة جمل وعبارات صحيحة ومترابطة. يقول أحدهم: لو ضمنا الاختبار أسئلة مقالية طويلة أو كثيرة معناه أننا حكمنا على معظم الطالبات بالرسوب المؤكد. فالأسئلة المقالية تتطلب تحليل للمعلومات، وفهما عميقا للمقرر. ويقول آخر: (الأسئلة المقالية تحدي كبير للطالبات. قد تعرف طالبة إجابة السؤال، لكنها لا تستطيع التعبير عنه كتابة، ربما تخاف من ارتكاب الأخطاء، أو تحرج من أسلوبها في الكتابة). ويقول آخر:( لا بد أن أراعي مستوى الطالبات العام. عادة الأسئلة المقالية قليلة العدد والدرجات المخصصة لها كبيرة، وعدم التمكن من الإجابة على سؤال مقالي يعني خسارة درجات كثيرة، على عكس الأسئلة الموضوعية التي لا يكلف الخطأ فيها أكثر من درجة). وأكدوا جميعاً على كثرة الأخطاء الإملائية، ورداءة خطوط معظم الطالبات يؤثر على جودة إجاباتهن، ويقلل من قدرة المصحح على فهم ما تكتبه الطالبة، مما قد يؤثر على تقييم الإجابات. يقول أحدهم فيما يخص رداءة الخط: (يصعب عليّ قراءة حتى اسم الطالبة). ويقول آخر: (من الظلم أن نحكم على إجابة طالب بسبب سوء خطه، أو بسبب أخطائه الإملائية). ويقول آخر: (لهذه الأسباب نختار الأسئلة الموضوعية التي لا تتطلب من الطالب أكثر من وضع إشارة على الإجابة الصحيحة). وقد تُعزى هذه النتيجة إلى قلة ممارسة الطلاب في المرحلة الجامعية، وقبلها بمراحل التعليم العام لهذا النوع من الأسئلة، ولاعتيادهم على أساليب التقييم الكمية على حساب الجودة والمهارات اللغوية، إضافة إلى ضعف التأسيس في المراحل الابتدائية، والمتوسطة، والثانوية، مع غياب التدريب العلمي، والأساليب الفاعلة لتنمية المهارات اللغوية الأساسية كالإملاء، والنحو، والتركيب، والإنشاء، مما يترك أثراً سلبياً على مستوى الطلاب في المرحلة الجامعية، ويدفع إلى استسلام معظم أعضاء هيئة التدريس أمام هذا المستوى، وعدم اتخاذ أي إجراء لعلاجه.

**نظام الاختبار المعمول به:** أجمع عدد (12) فرداً من أفراد العينة وبنسبة (92%) على أن مطالبة الأستاذ الجامعي باستكمال جميع إجراءات تصحيح أوراق الاختبار ومراجعتها ورصد الدرجات خلال ثلاثة أيام من تاريخ عقد الاختبار، هو من الأسباب التي تجعل معظم أعضاء هيئة التدريس مضطرا للاعتماد على الأسئلة الموضوعية، وتجنب الأسئلة المقالية في إعداد أسئلة الاختبارات النهائية، يقول أحدهم: (أعداد الطالبات كثيرة جداً خاصة إذا كنت أدرس أكثر من شعبة لنفس المقرر، ثلاثة أيام لا تكاد تكفي لتصحيح الأسئلة الموضوعية). ويقول آخر: (قد لا نجد من يقبل بمراجعة التصحيح، أو التدقيق لاختباراتنا من الزملاء في حال وجود سؤال مقالي طويل).

وقد يُعزى قلق معظم أعضاء هيئة التدريس من قلة الوقت المتاح للتصحيح؛ لضعف قدرة بعض الأعضاء في إعداد نماذج إجابة دقيقة، وقلة تدريبهم على كيفية تحديد معايير التصحيح والأفكار الأساسية، واقتصار معظم الدورات التدريبية الخاصة بإعداد أسئلة الاختبارات على توجيهات لإعداد الأسئلة الموضوعية بشكل شبه دائم، دون التطرق لطرق تصحيح الأسئلة المقالية فضلاً عن طرق

إعدادها. مما يخلق صورة ذهنية خاطئة لديهم عن تصحيح الأسئلة المقالية المفتوحة، تدفعهم إلى تجنب تضمينها ضمن أسئلة الاختبارات.

**تأصل ثقافة الاختبارات التقليدية التي تركز على الحفظ والاسترجاع:** أجمع (9) من أفراد العينة بنسبة (70%) على أن أسباب ضعف تضمين الأسئلة المقالية في الاختبارات النهائية يعود إلى ثقافة الاختبارات الجامعية، وتعود أعضاء هيئة التدريس والطالبات على التعامل مع لأسئلة الموضوعية أكثر من المقالية. يقول أحدهم: (تعودنا نحن الأعضاء على الأسئلة الموضوعية أثناء دراستنا، ونقلنا هذا التعود لطلابنا). ويقول آخر: (اطلعت على نماذج أسئلة كثيرة عندما بدأت بالعمل بالجامعة، ووجدتها أسئلة موضوعية كثيرة، وسؤال مقالي بسيط واحد). واتفق جميعهم على صعوبة تقبل الطالبات للتعامل مع الأسئلة المقالية حتى وإن تم تدريبهن، بسبب تعودهن عليها، وهن في المرحلة الثانوية. فيقول أحدهم: (حاولت أن أضع أسئلة مقالية كثيرة في الاختبار، وتفاجأت بأن عدد كبير من الطالبات أجبن على الأسئلة الموضوعية فقط وتركن الأسئلة المقالية دون إجابة). ويقول آخر:(تعودت الطالبات على التقدم بشكاوى كثيرة إذا كانت الأسئلة المقالية كثيرة. تعودن على ألا يزيد السؤال المقالي عن تعريف أو تعليل من سطر إلى سطرين. علماً بأن أكثر طلبات إعادة التصحيح تكون للاختبارات المقالية). وقد تُعزى هذه النتيجة إلى ضعف قدرات بعض أعضاء هيئة التدريس على صياغة الأسئلة المقالية، وانتقائها بدقة تلائم مستوى الطلاب، وعلى إعداد نماذج إجابة تتضمن النقاط المطلوبة بالتفصيل، وتوزيع درجة السؤال عليها، مما يُصعب عليهم الخروج من دائرة التعود إلى دائرة التحدي وتجربة إدراج الأسئلة المقالية في الاختبارات التي يعدونها، وتشجيع الطالبات على تقبلها، وتعويدهن على التعامل معها.

**ومن جانب آخر فإن من أسباب ضعف تمثيل الأسئلة المقالية في الاختبارات النهائية للمستويات المعرفية العليا من وجهة نظر أعضاء هيئة التدريس بكلية التربية بشقراء:**

**ضعف مستوى الطالبات على الفهم والتحليل والتنظيم والربط بين المفاهيم:** اتفق (11) فرداً من أفراد العينة بنسبة (85%) بأن سبب ضعف تمثيل الأسئلة المقالية في الاختبارات النهائية للمستويات المعرفية العليا إلى ضعف الطالبات العام، وعدم ألفتهن بها في المراحل الدراسية المبكرة مما أثر على أدائهن في المرحلة الجامعية اللاحقة. الأمر الذي يجعل أعضاء هيئة التدريس مضطرين للاعتماد على الأسئلة التي تتطلب استرجاع المعلومات فقط وتجنب الأسئلة التي تتطلب مهارات تفكير عليا: كالتحليل والتقويم والإبداع. فيقول أحد أفراد العينة: (ليس لدى الطالبات خبرة في تحليل الأسئلة المقالية التي تتطلب مهارات تفكير عليا. قد لا تستطيع الطالبة فهم السؤال وتحليله وتحديد المطلوب ومن ثم صياغة إجابة منظمة مدعومة بالبراهين والأدلة والأمثلة). ويقول آخر: (المستوى الذي نراه لا يحفزنا على كتابة أسئلة تتطلب أكثر من استرجاع معلومة).

**استخدام استراتيجيات تدريس تقليدية:** تفرد (2) من أفراد العينة بنسبة (15%) وأكدا على أن اعتماد معظم أعضاء هيئة التدريس على الأسئلة التي لا تتطلب أكثر من الحفظ والفهم إلى افتقار معظمهم إلى استخدام استراتيجيات التدريس التي تشجع على التفكير النقدي، وحل المشكلات التي تتطلب مهارات تفكير عليا التي تعدُّ أساسية في الأسئلة المقالية. فيقول الأول: (اعتماد الممارسات التدريسية التقليدية التي تركز على الحفظ والتلقين بدلاً من الفهم العميق قد يؤدي ضعف الطلاب في هذه المهارات، وعدم قدرتهم على حل الأسئلة المقالية، ومن ثم التركيز على أسئلة الحفظ والفهم). ويقول الآخر: (لم يتدرب بعض أعضاء هيئة التدريس على استراتيجيات التدريس التي تعزز مهارات التفكير العليا. إضافة إلى قلة فرص التجريب والابتكار في التدريس تؤدي إلى الاعتماد على طرق التدريس التقليدية التي تشجع على الحفظ والاسترجاع، وبالتالي ظهور الاختبارات بهذه الصورة).

وتُعزى هذه النتيجة إلى قلة اهتمام أعضاء هيئة التدريس بصياغة الأهداف السلوكية بجميع المستويات المعرفية أثناء التخطيط للتدريس والإفادة منها عند وضع الأسئلة. إضافة إلى أن إعداد الأسئلة ذات المستويات المعرفية العليا وتصحيحها قد تحتاج إلى خبرات وقدرات قد لا تتوفر لدى الكثير منهم.

تتفق هذه النتيجة مع نتيجة دراسة نايف وعبد الكاظم [38]، ودراسة سعيد وجبار [41]، ودراسة عبده[1]، ودراسة الجابري وحسين[8] ، التي توصلت نتائجها إلى قلة الاختبارات المقالية ذات المستويات العليا لصعوبة تصحيحها ولصعوبة كتابتها.

وتختلف هذه النتيجة مع نتيجة دراسة نايف وعبد الكاظم[38] التي كشفت نتائجها عن اقتصار الاختبارات على الأسئلة المقالية فقط. كما اختلفت مع دراسة سعيد وجبار[41] التي كشفت نتائجها عن تفوق نسبة الأسئلة المقالية.

**خامساً: إجابة السؤال الخامس الذي نصه: ما هي مقترحات أعضاء هيئة التدريس بالجامعة لتضمين الأسئلة المقالية في الاختبارات النهائية لطلبة جامعة شقراء؟**

للإجابة عن هذا السؤال تم حساب التكرارات والنسب المئوية ومناقشة النتائج وتحليلها، على النحو الآتي:

**1.اشتراط اجتياز اختبار تحريري للقبول بكلية التربية: اتفق (10)** أفراد من أفراد العينة بنسبة (77%) على ضرورة تعديل سياسة القبول بكلية التربية وتضمين الشروط اجتياز اختبار تحريري للمتقدمين لتقييم قدرات الطلاب المعرفية، ومهاراتهم الكتابية، ومهارات التفكير العليا.

**2.تدريب أعضاء هيئة التدريس على استراتيجيات التدريس التي تنمي مهارات التفكير العليا لدى الطلاب:** اجمع (8) من أفراد العينة على ضرورة تنظيم برامج تدريبية لأعضاء هيئة التدريس على كيفية تدريس مهارات التفكير العليا وأساليب تقييمها.

**إعداد دليل لجودة أسئلة الاختبار:** أكد (7) من أفراد العينة على ضرورة توفير دليل لجودة أسئلة الاختبار، يعمل كمرجع موحد للمعلمين يوفر إرشادات إجرائية حول كيفية إعداد الأسئلة بجميع أنواعها

**توصيات ومقترحات:**

- **بناء على نتائج الدراسة توصي الباحثة بالآتي تحت إشراف عمادة التطوير والجودة بالجامعة وبالتعاون مع خبراء في التقويم التربوي وخبراء في المناهج وطرق التدريس:**

1. إنشاء لجنة جامعية لمراجعة جودة الاختبارات الجامعية.
2. عقد ورش تدريبية سنوية إلزامية في صياغة الأسئلة تقيس المستويات العليا من تصنيف بلوم المعدل.
3. إلزام الأقسام بنسبة 30% لبناء أسئلة تفكير عليا.
4. تطوير دليل إرشادي لتصميم الأسئلة وفق بلوم المعدل لإعداد وتقييم جودة الأسئلة الاختبارات وفق بنوده.
5. إنشاء بنك أسئلة معتمد بمشاركة أعضاء هيئة التدريس من جميع كليات التربية التابعة لجامعة شقراء.
6. توظيف الذكاء الاصطناعي في تحليل الأسئلة لقياس مستوى بلوم تلقائيًا.

- **بناء على نتائج الدراسة تقترح الباحثة إجراء الدراسات المستقبلية الآتية:**

1. **دراسة طولية:** تأثير الأسئلة المقالية على التحصيل الأكاديمي الطويل المدى لقياس ارتباط استخدام الأسئلة العليا (التحليل/الإبداع) بتحسن مهارات التفكير لدى الخريجين.
2. **دراسة تجريبية:** فعالية بنك الأسئلة المعتمد في تحسين جودة الاختبارات لتقييم أثر الاعتماد على البنك في تنوع الأسئلة وموضوعية التصحيح.
3. **دراسة نوعية:** تحليل معوقات تطبيق تصنيف بلوم المعدل في البيئات التعليمية المحافظة لدراسة العوامل الثقافية المؤثرة على قبول الأسئلة الإبداعية.
4. **دراسة مقارنة:** دور الذكاء الاصطناعي في تصحيح الأسئلة المقالية، وتقليل التحيز لاستكشاف إمكانية استخدام أدوات مثل ChatGPT في التقييم

## References:


1. Abdu, M. Y. M. (2015). Evaluating Examination-Papers of EFL Students at Hodeidah University -Yemen. Abhath, 2(3), 1-14.
2. Abdulhamid, K. S. & Al-Bastami, D. A. (2012). Teaching strategies. Dammam: Al-Mutanabbi Library.
3. Albadi, B. H. , & Al-Hidabi, D. A. (2021). The Adequacy of the Islamic Institutes Teachers Efficiency in Constructing Achievement Tests According to Good Test Criteria from Their Perspective. Journal of Educational and Psychological Sciences, 5 (21), 76-96.
4. Al-Dalu, M. M. (2021). Analysis of the final exam questions for Islamic Education for the twelfth grade in Palestine. Journal of the Islamic University for Educational and Psychological Studies, 29(6), 334-352.
5. Al-Ghazal, M. O. (2020). Evaluating postgraduate examinations in the Department of Arabic Language and Literature, Faculty of Arts, Misurata University, in light of comprehensive quality standards. Journal of the Faculty of Arts, 15, 31-551.
6. Al-Hadour, Z. A. (2017). Analysis of the General Secondary Certificate Examinations in Mathematics in the Republic of Yemen in light of modern classifications of educational objectives. Al-Jami' Journal of Psychological Studies and Educational Sciences, 7(7), 232-255.
7. Ali, B. A., Hassan, A. A., & Ali, O. B. (2023). Electronic Evaluation of Quality of Exams' Questions Written in Arabic language Based on Bloom's Taxonomy. International Journal of Advanced Research in Computer and Communication Engineering, 12(8), 50-61.
8. Al-Jabri, R. H.& Hussein, A. A. (2022). Evaluating the questions of the electronic tests for the Arabic Language Department at the College of Basic Education according to Bloom's Taxonomy. Nasq Journal, 36(2), 1176-1189.
9. Al-Khalifa, H. J. & Mutawa, D. (2012). An Introduction to teaching. Riyadh: Al-Rushd Library.
10. Almasudi, M., Algaburi, A., & Algaburi,M.(2015). Teaching diversification protocols in teaching strategies and methods (my values charter). Al-dar al-manhgia for Publishing and Distribution. , Amman, Jordan.
11. Al-Mazoughi, I. S. (2017). Evaluating final exam questions in light of Bloom's Taxonomy of Cognitive Objectives: A case study of the Psychology Department, Faculty of Arts, Sabratha University. Sabratha University Scientific Journal, (3), 91-107.
12. Al-Saeed, S. M. (2009). Basic teaching skills for teachers. Riyadh: Al-Rushd Library.
13. Al-Salahi, A. (2008). Evaluating final exams at Sana'a University in light of the criteria for good exams. Journal of University Researcher, (17), 61-88.
14. Al-Sharifi, Y. K. (2021). Evaluating the final exams for the Arabic language department in light of Bloom's Taxonomy and university question-setting standards. Ahl al-Bayt Journal, 1 (28), 340-360.
15. Al-Sheikh, A. Y.(2015). An analytical study of the extent to which the questions in the ninth-grade basic science textbook in Yemen take into account the modified Bloom's Taxonomy. Arab Journal of Science and Technology Education, 3, 23-43
16. Al-Shurfat, M. S. M. (2020). The degree of Taking into Account the Questions of the Physics Book for the Twelfth Grade for the Scientific and Industrial Branches in Jordan, According to Bloom's Modified Levels of Knowledge. Journal of Educational and Psychological Sciences, 4(35), 59-76.
17. Al-Tariri, A. S. (2014). Psychological and Educational Measurement (2nd ed.). Al-Rushd Library, Riyadh, Kingdom of Saudi Arabia.
18. Amin, S. A. (2023a). Preparing final exam questions for scientific and humanities departments according to Bloom's Taxonomy of Creativity at the College of Basic Education, Salahaddin University - Erbil. Tikrit University Journal of Humanities 3(6), 324-345.
19. Amin, S. A. (2023b). The Evaluation of Final Exam Questions for Scientific and Humanities Departments according to Blooms Cognitive Classification in the College of Basic Education, Salahuddin University -Erbil. Journal of Tirkit University for Humanities, 30(6), 324-345.
20. Anderson,A., Lorin, W.,& David ,R. Krathwohl, eds.( 2001). A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives. New York: Addison Wesley Longman, Inc.

21. Ata, A. & Ghazy, M.(2023). Exam Questions from Essay to MCQ with Different Levels of Difficulty in Engineering Mechanics. Journal of Education, Teaching and Social Studies,5(2), 19-26.
22. Aziza, I. F. ( 2016). Essay Test as an Evaluation Tool for Teaching Arabic Language. AL-WIJDAN: Journal of Islamic Education Studies1(1), 14-26.
23. Badawi, R. M. (2011). Curriculum and teaching methods. Amman: Dar Al-Fikr.
24. Bahri, M, Y. (2012). The Educational Curriculum (Its Foundations and Analysis). Amman: Dar Safaa for Publishing and Distribution.
25. Brookhart, S. M. (2018). How to assess higher-order thinking skills in your classroom. ASCD.
26. Darwazeh, A. (2011). The degree to which teachers in Qalqilya Governorate schools consider Bloom's Taxonomy of Cognitive Objectives when planning their lessons. An-Najah University Journal for Research (Humanities), 25(10), 2559-2582.
27. Dar wazeh, A. N. (2020). Dar wazeh's classification of educational objectives: A modification of "Amdersen's" classification of "Bloom's" classification of educational objectives. International Journal of Educational and Psychological Studies, 8(1), 77-90.
28. Qadir, S. M., Omar, R. M. , Rasheed, M. H. ,& Mohammed, C. J. (2023). Assessing the End-of-Semester Examination Papers During the Implementation of the Bologna Process: Bloom's Taxonomy as a Framework. Koya University Journal of Humanities and Social Sciences,6(1), 77-89.
29. Gates, J., & Pugh, S. (2021). The Application of Bloom's Taxonomy to Higher Education Examination Questions in Physics. New Directions in the Teaching of Physical Sciences, 16(1), 1-11.
30. Ghalyoun, A. M. (2009). Evaluating the final postgraduate examinations at the Faculty of Education, Sana'a, in light of the criteria for a good examination. Scientific Journal of the Faculty of Education, 1(8), 174-198.
31. Hashim, K. & Al-Khalifa, H. J. (2015). Educational evaluation: Its Concept, Methods, Fields, and Modern Trends (5th ed.). Riyadh: Al-Rushd Library.
32. Hashim, K. (2006). Educational evaluation (Its Concept - Methods - Fields - Modern Directives). Riyadh: Al-Rushd Library.
33. Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. Theory Into Practice, 41(4), 212–218.
34. Laibi, W. A. & Al-Mayahi, M. A. (2023). Analysis of postgraduate examination questions in light of examination question quality standards in the Department of Educational Sciences. Basrah Journal of Human Sciences Research, 1(A), 222-239.
35. McMillan, J. H. (2018). Classroom assessment: Principles and practice (7th ed.). Pearson.
36. Mohammed, B. (2021). Achievement tests as an effective method for educational assessment. Al-Khaldounia Journal of Humanities and Social Sciences, 13(1), 53-61.
37. Muhammad, A. A. (2006). Curricula from a general and contemporary perspective. Riyadh: Al-Rushd Library.
38. Nayef, A. K. & Abdul Kadhim, M. H. (2014). Evaluating the final exam questions for the History Department at the College of Education for Humanities in light of Bloom's Taxonomy. Karbala University Journal of Science, 12(1), 91-110.
39. Opstad, L. (2021). Can Multiple-Choice Questions Replace Constructed Response Test as an Exam Form in Business Courses? Evidence from a Business School. Athens Journal of Education, 8(4), 349-360.
40. Sa'adah, J. A. & Ibrahim, A. M. (2023). The Contemporary school curriculum (11th ed.). Amman: Dar Al-Fikr.
41. Saeed, B. H. & Jabbar, H. I. (2022). Evaluating final exam questions in light of Bloom's Taxonomy of Cognitive Objectives: The History and Geography Departments as a model. Journal of Kalar University, 9(4), 504-516.
42. Saeed,A.H.(2009). Analysis of tests in the secondary school mathematics in the light of the general levels of knowledge Diltz (Dilts) ]Mastar[. Taez University Republic of yemen.
43. Sarawi, A. & Farhat, F. (2022). The extent to which semester achievement tests conform to the instructions of the test construction guide - Natural and life science tests for the fourth year of middle school as a model. Al-Shamel Journal of Educational and Social Sciences, 5(02), 397-411.
44. Saud, A. M. & Mahmoud, W. Q. (2022). Evaluation of the final exam questions for the Department of Educational and Psychological Sciences for the academic year 2018-2019. Psychological Research Center, 30(4), 45-78.

45. Shakouna, K. S. (2022). Bloom's Taxonomy of Cognitive Objectives (Revised). Al-Qalam Al-Mubeen Journal, 22(1), 1-18.
46. Shehata, H., Al-Najjar, Z. & Ammar, H. (2003). Dictionary of Educational and Psychological Terms. Egyptian-Lebanese House, Cairo: Arab Republic of Egypt.
47. Shibata, T., & Uto, M. (2022: October). Analytic Automated Essay Scoring Based on Deep Neural Networks Integrating Multidimensional Item Response Theory. Proceeding of the 29th International Conference on Computational Linguistic, pages 2917-2926. October 12-17-2022.
48. Tamam, S., & Salah, A. (2016). Modern Curricula and Methods of Teaching and Learning. Amman: De Bono Center for Thinking Skills.
49. Yusuf, M.I. & Al-Rifai, M. M. (2005). (Educational Evaluation: Its Foundations and Procedures (4th ed.). Riyadh: Al-Rushd Library.
50. Zahoor, A., Farooqui, S. I., Khan, A., Kazmi, S. A., Qamar, N., & Rizvi, J. (2023). Evacuation of Cognitive Domain in Objective Exam of Physiotherapy Teaching Program by Using Bloom's Taxonomy. Journal of Health and Allied Sciences NU-, 13(02), 289-293.