مجلة جامعة عمران 05 (2025) ص 060-033

مجلة جامعة عمران

journal homepage: www.auy.edu.ye

تقويم جودة الاختبارات التحصيلية لمقررات الطب البشري بكلية الطب-جامعة عمران

سناء أحمد علي التويتي[^]

مركز التطوير الأكاديمي وضمان الجودة -جامعة عمران.

الملخص

هدفت الدراسة إلى تقويم اختبارات التحصيل لمقررات الطب البشري بكلية الطب- جامعة عمران في ضوء معايير جودة الاختبارات، ولتحقيق هذا الهدف أُستخدم المنهج الوصفي التحليلي، وتكونت عينة الدراسة من (26) اختباراً من اختبارات مقررات الطب البشري للعام الجامعي 2022-2023م، وتكونت أداتها من استمارة تحليل مكونة من (61) معياراً لجودة الاختبارات التحصيلية، منها (32) معياراً خاصاً بمضمون الورقة الاختبارية، و(29) معياراً خاصاً بجانبها الشكلي والفني، وأظهرت النتائج أن المعايير الضمنية للورقة الاختبارية توافرت بدرجة عالية وبنسبة (70.9%)، في حين حصل معيارا شمول الاختبار لمحتوى المقرر- وهو اختبار يغطي مخرجات التعلم في مجال المهارات الذهنية-على نسب ضعيفة (22%) و (27.8%) على التوالي، وتوافرت المعايير الشكلية بدرجة متوسطة وبنسبة (50.5%)، في حين توافرت المعايير الفنية بدرجة عالية جداً وبنسبة (90.3%)، كما أظهرت النتائج أن الأساتذة أكثر استخداماً لأسئلة الاختيار من متعدد؛ حيث توافرت في (22) اختباراً، بينما توافرت أسئلة الصواب والخطأ في ستة اختبارات، وأسئلة الإكمال والمطابقة في أربعة اختبارات فقط، وعليه أوصت الدراسة بضرورة شمول أسئلة الاختبار لمحتوى المقرر الدراسي، ومخرجات التعلم في المجال الذهني، وتنوع الأسئلة الموضوعية والمقالية، بالإضافة إلى تنفيذ ورش تدريبية لأعضاء هيئة التدريس في ضوء معايير جودة الاختبارات التحصيلية.

الكلمات المفتاحية: تقويم -الاختبار التحصيلي -معايير جودة الاختبارات.

Abstract: The study aimed to evaluate the achievement tests of the human medicine courses at the Faculty of Medicine-Amran University in the light of the quality standards of the tests. To achieve this goal, the descriptive analytical approach was used. The study sample consisted of (26) tests of the human medicine courses for the academic year 2022-2023, and its tool consisted of an analysis form consisting of (61) standards for the quality of the achievement tests, including (32) standards for the content of the test paper, and(29) standards for its formal and technical side. The results showed that the implicit standards of the test paper was available to a high degree and by (70.9%), while The standard of test coverage of the course content was also achieved. The test covers learning outcomes in the field of mental skills at low rates (22%) and (27.8%) respectively and the formal standards were available to a medium degree and by (50.5%), while the technical standards were met with a very high score of (90.3%). The results also showed that professors are more likely to use multiple choice questions, as they were available in (22) tests, while the right and wrong questions were provided in six tests, and the completion and matching questions were provided in only four tests. Accordingly, the study recommended that the test questions should include the course content, and learning outcomes in the mental field, and the diversity of objective and essay questions, in addition to the implementation of training workshops for members Faculty in light of the quality standards of the achievement tests.

المقدمة

يفرض التقدم العلمي والتقني المتسارع على الجامعات السعي لتطوير النظام التعليمي بشكل عام ونظام التقويم بشكل خاص، مما يعد مدخلاً لإصلاح النظام التعليمي؛ فجودة التعليم تتحقق من خلال جودة وسائل القياس

[*]**Email:** altwaty155@gmail.com

©Amran University Publishing

والتقويم التي تسهم في قياس تعلم الطلبة وتقييم فاعلية التدريس، وتوجيه القرارات التعليمية، بناءً على أسس علمية، فضلاً عن جودة وسائل القياس والتقويم ومنها الاختبارات التحصيلية التي أصبحت أحد متطلبات الاعتماد الأكاديمي؛ لذلك يؤكد خبراء القياس والتقويم ضرورة بناء هذه الاختبارات على أسس ومعايير علمية حديثة، بحيث تقدم بيانات صادقة يمكن الاعتماد عليها في تشخيص دقيق للعملية التعليمية التعلمية، واتخاذ قرارات موضوعية مناسبة لتطويرها، فالتعليم المتميز ينطلق من مبدأ تماثل مواصفات متخرجيه مع المواصفات القياسية للإنتاج المتميز، وتتواءم هذه المواصفات مع حاجات المجتمع وسوق العمل.

كما تسعى التوجهات المعاصرة إلى المواءمة بين مخرجات التعلم واستراتيجيات التدريس والتقويم، وذلك من خلال عملية القياس والتقويم التي ينبغي أن تخضع هي الأخرى لعمليات التقويم؛ للتحقق من جودتها وصدق نتائجها في تقدير مدى هذا التماثل، فالتقويم بأدواته كافة وبوصفه أحد مكونات أي منهج أو برنامج تعليمي يجب أن يخضع للتقويم والتطوير المستمر (عكاشة وضحا، 2020 ؛ Uluman, 2023)، وهذا ما جعل الاختبارات التحصيلية تحضي باهتمام الكثير من الباحثين، ودراسة واقع تطبيقها بين تقييم جودة تصميم الاختبارات، وتقييم كفاية معديها، وتقييم مدى شموليتها وصلاحيتها وتوافرها على معايير وشروط الاختبار الجيد (سراوي علي وفرحات فوزي، 2022، صـ299).

كما أكد علماء القياس أن تأثير عملية التقويم في جودة مخرجات التعلم تكمن في أساليب التقويم التي يستخدمها الأستاذ، الذي يتوجب عليه أن يقلل من التركيز على التقويم من أجل وضع الدرجات وتحديد الناجحين والراسبين، ويتجه إلى استخدام التقويم بوصفه وسيلة لتطوير القدرات والمعارف التي تبقى مع الطالب بعد التخرج (العلي، 2013، صـ43). لا سيما أن التعلم لم يعد قائماً على حفظ المعلومات واسترجاعها؛ إذ تطورت أدوات التقويم وتعددت أساليبه، وعلى وجه الخصوص تقويم تحصيل الطلبة، الذي شهد تطورات سريعة وتحولات جوهرية في منهجيات القياس والتقويم، وتحولاً نوعياً في أساليبها وأدواتها وتقنياتها وممارستها الميدانية (Al-Khawuldeh et al., 2022)، وعلى الرغم من ذلك فإن الاختبارات التحصيلية لا تزال أهم تلك الأدوات وأكثرها استخداماً في المؤسسات التعليمية بشكل عام والتعليم الجامعي بشكل خاص، فهي الوسيلة الأساسية في تقييم العملية التعليمية، وتسهم في تحقيق الأهداف التعليمية ومخرجات التعلم، وتلعب دوراً أساسياً في حياة الطلبة، فبناءً على نتائجهم يُحدد مستقبلهم، وتُتخذ العديد من القرارات نحوهم ويُختار الأفراد للمهن المناسبة لهم (AbuFoudeh, 2020).

كما أن نتائج الاختبارات التحصيلية تعكس بعض جوانب الأداء التدريسي للأستاذ الجامعي؛ باعتبار أن عملية إعداد الاختبارات أحد مؤشرات الكفاية التي ينبغي أن يمتلكها هذا الأستاذ؛ إذ إن هناك ترابطاً بين مخرجات التعلم وطرق التدريس والأنشطة والتقويم، فطرق التدريس التي تصلح لتحقيق مخرجات التعلم في مستوى المعرفة والفهم كالمحاضرة مثلاً قد لا تصلح لتحقيق مخرجات التعلم في مستويات التفكير العليا، كالقدرة على تحليل المعرفة وإعادة تركيبها وتقويمها وإصدار الأحكام بما يسهم في تطوير المهارات الذهنية كالتفكير الإبداعي

والابتكاري والقدرة على حل المشكلات؛ لذلك تكمن أهمية نتائج عملية التقويم للأستاذ الجامعي في تعديل طرق التدريس وأساليب التقويم التي يستخدمها وتحسينها، بشكل يجعل تدريسه أكثر فعالية، وأساليب تقويمه أكثر دقة وموضوعية. (جامعة المجمعة، 2024 ؛ القدسي وآخرون، 2020)

يتضح مما سبق أهمية بناء الاختبارات التحصيلية على وفق معاييرها العلمية، وضرورة إتقان الأستاذ الجامعي لمهارات بنائها؛ وهو ما أشارت إليه العديد من الدراسات التي أوصت بأن مهارة صياغة الأسئلة وإعداد الاختبارات تمثل أهم الكفايات التي ينبغي أن يكتسبها الأستاذ الجامعي ويسعى إلى إتقانها، وأن على الجامعات تدريبه على مهارات إعداد الاختبارات؛ إذ إن كفاية الأستاذ الجامعي وتنمية خبرته اللازمة لأداء مهامه وواجباته في تعليم الطلاب وتقويمهم ما تزال دون المستوى المطلوب.

على أن هناك جهوداً بُذلت لتحسين جودة التعليم العالي خلال العقدين الماضيين، ابتداءً من تأسيس مجلس الاعتماد الأكاديمي وضمان جودة التعليم العالي في العام (2004) وانتهاءً بتقديم وثيقة معايير الاعتماد الأكاديمي (البرامجي) في ثمانية معايير رئيسة، نص المعيار الثامن منها (تقييم أعضاء هيئة التدريس) على ضرورة توافر دليل لتقييم أداء أعضاء هيئة التدريس يهدف إلى تحسين الأداء التدريسي في البرامج على وفق أفضل الممارسات المحلية والإقليمية (مجلس الاعتماد الأكاديمي وضمان جودة التعليم العالي، 2021، ص19)، بالإضافة إلى إنشاء مراكز للتطوير وضمان الجودة في العديد من الجامعات اليمنية، وعلى الرغم من تلك الجهود فإن الدراسات ذات الصلة مثل دراسة (غليون، 2009)، و(بافطوم والسفياني،2019)، و (السفياني، 2020) و(الذري، 2021) أشارت إلى أنه لا تزال هذه الجهود شكلية أكثر منها موضوعية، وأن معظم الاختبارات التحصيلية غير مطابقة لمعايير الاختبارات الجيدة وأنها ليست في المستوى المطلوب.

ولذلك فإن عدم مراعاة هذه المواصفات والمعايير لا سيما في مؤسسات التعليم العالي يؤدي إلى ضعف موثوقية الدرجات التي يحصل عليها الطلبة في الاختبارات التحصيلية، ويقلل من أهميتها في تزويد القسم العلمي بمعلومات مفيدة عن مستوى تحقق مخرجات تعلم البرنامج الأكاديمي في المقرر (عكاشة وضحا، 2020، صـ12)؛ لذلك جاءت هذه الدراسة لتسهم في تشخيص عملية تقويم الاختبارات التحصيلية على وفق مواصفات ومعايير جودة الاختبارات، وتقديم توصيات لمتخذي القرار قد تساعد على تحسين عملية التقويم وتطويرها في الجامعة.

مشكلة الدراسة وأسئلتها

تمثل الاختبارات التحصيلية أهمية كبيرة في العملية التعليمية، بوصفها الوسيلة الأكثر استخداماً للحكم على مدى تحقق أهدافها واتخاذ قرارات بشأنها بناءً على أسس علمية موضوعية، فضلاً عن أن لها تأثيرات نفسية سلبية في الطلبة كالخوف والقلق والتوتر الذي يشكّل تهديداً لأدمغتهم مما ينعكس على أدائهم في الاختبارات، وفي الوقت نفسه تأثيرات إيجابية حيث يمثل التفوق في الاختبار حافزاً داخلياً يتولد لدى الطلبة مما يدفعهم إلى

النجاح المستمر(القدسي وآخرون، 2020)؛ لذلك ينبغي إعداد الاختبارات بمواصفات ومعايير الاختبار الجيد؛ لكي تسهم في الحد من تأثيراتها السلبية على الطلبة وتزيد من التأثيرات الإيجابية لديهم.

ومن خلال عمل الباحثة في مركز التطوير الأكاديمي وضمان الجودة في الجامعة، والاطلاع على عينات من نماذج الاختبارات في الكليات، لوحظ أن هذه النماذج لا تتوافق مع مواصفات الاختبارات ومعاييرها الجيدة، مما يؤثر سلباً في نتائج الطلبة وفي تحقيق المخرجات التعليمية بالشكل المطلوب، وهذا ما أشارت إليه العديد من الدراسات كدراسة (الرازحي، 2005) ودراسة (بافطوم والسفياني، 2019) ودراسة (السفياني، 2020) ودراسة (الذري، 2021)، لذلك تأتي هذه الدراسة لتقويم الاختبارات التحصيلية، وتقديم الأسس العلمية والمواصفات الجيدة للورقة الاختبارية بشكل عام، من خلال مواصفات ومعايير يمكن الاعتماد عليها، بناءً على الأسس العلمية لجودة الاختبارات وتتلخص مشكلة الدراسة في الأسئلة الآتية:

1. ما مواصفات جودة الاختبارات التحصيلية و معاييرها؟
2. ما نسبة توافر المعايير الضمنية للاختبارات التحصيلية لمقررات الطب البشري بكلية الطب – جامعة عمران للعام الدراسي 2022-2023م؟ ويتفرع هذا السؤال إلى أربعة فروع هي:

2.1. ما مدى جودة أسئلة الاختبارات المقالية لمقررات الطب البشري بكلية الطب – جامعة عمران للعام الدراسي 2022-2023م؟

2.2. ما مدى جودة أسئلة الاختبارات الموضوعية لمقررات الطب البشري بكلية الطب – جامعة عمران للعام الدراسي 2022-2023م؟

2.3. ما نسبة توافر المعايير العامة للاختبارات التحصيلية لمقررات الطب البشري بكلية الطب – جامعة عمران للعام الدراسي 2022-2023م؟

2.4. ما مدى تحقيق أسئلة الاختبارات التحصيلية لمخرجات التعلم المستهدفة لمقررات الطب البشري بكلية الطب – جامعة عمران للعام الدراسي 2022-2023م؟

3. ما نسبة توافر المواصفات الشكلية للورقة الاختبارية لمقررات الطب البشري بكلية الطب – جامعة عمران للعام الدراسي 2022-2023؟

4. ما نسبة توافر المواصفات الفنية للورقة الاختبارية لمقررات الطب البشري بكلية الطب – جامعة عمران للعام الدراسي 2022-2023؟

أهداف الدراسة

1. تحديد المواصفات والمعايير الفنية والشكلية للورقة الاختبارية الجيدة.
2. الكشف عن مدى توافر المعايير الضمنية لأسئلة الاختبارات الجيدة المعدة من قبل أعضاء هيئة التدريس في قسم الطب البشري بكلية الطب – جامعة عمران.

3. الكشف عن مدى توافر المواصفات الفنية والشكلية للورقة الاختبارية الجيدة في أوراق اختبارات مقررات الطب البشري بكلية الطب جامعة عمران للعام الدراسي 2022-2023م.

أهمية الدراسة

يمكن أن تفيد الدراسة في الآتي:

1. تقديم مجموعة من المواصفات والمعايير التي يمكن اتباعها عند إعداد الاختبارات التحصيلية وتقويمها على وفق معايير الاختبارات الجيدة.

2. قد تفيد نتائج الدراسة متخذي القرار في كلية الطب البشري ومركز التطوير الأكاديمي وضمان الجودة في الجامعة في معرفة جودة الاختبارات في هذه الكلية ومدى قياسها لمخرجات التعلم المستهدفة، والاستفادة من ذلك في تجويد الاختبارات التحصيلية مستقبلاً.

3. يمكن الاستفادة من نتائج هذه الدراسة في بناء خطط التحسين والتطوير لنظم التقويم والامتحانات لتحقيق النواتج التعليمية المستهدفة بوصفه أحد متطلبات الاعتماد الأكاديمي.

5. تساعد الدراسة أعضاء هيئة التدريس ومعاونيهم على تقييم أدائهم لإعداد الاختبارات التحصيلية على وفق مواصفات ومعايير الاختبارات الجيدة.

محددات الدراسة

تتحدد الدراسة في أوراق الاختبارات التحصيلية للعام الدراسي 2022-2023م لمقررات الطب البشري بجامعة ع مران، ومواصفات جودة هذه الاختبارات ومعاييرها التي تشمل (61) مواصفة ومعياراً أعدتها الباحثة.

مصطلحات الدراسة

1. التقويم

يعرّف التقويم بأنه: عملية جمع معلومات عن ظاهرة ما، وتصنيفها وتحليلها وتفسيرها، سواء كانت كمية أو كيفية؛ بهدف إصدار حكم أو قرار لغرض تحسين العمل. (شحاته وآخرون، 2003)

ويعرّف التقويم في هذه الدراسة بأنه: عملية وصفية تجمع فيها المعلومات حول مواصفات الاختبارات التحصيلية لمقررات الطب البشري (عينة الدراسة) ومعايير جودتها في ضوء قائمة مواصفات الورقة الاختبارية ومعايير الجودة للاختبارات المعدة في هذه الدراسة؛ لتحديد نقاط القوة والضعف فيها؛ لمساعدة متخذي القرار على عمل الإجراءات المناسبة لتحسينها.

2. الاختبار التحصيلي

هو أداة تقيس المعلومات والمهارات التي سبق دراستها من خلال مادة دراسية محددة. (سليمان وأبو علام، 2010، صـ191). ويعرف إجرائياً بأنه الأداة التي تشمل مجموعة من الأسئلة أُعدت من قبل أعضاء هيئة التدريس في قسم الطب البشري بكلية الطب -جامعة عمران، ونُفذت على الطلبة خلال العام الدراسي 2022-2023م.

3. معايير جودة الاختبارات

تُعرّف معايير جودة الاختبارات في هذه الدراسة بأنها مجموعة المعايير والأسس العلمية التي تبنى عليها الاختبارات التحصيلية، وتتحدد بقائمة المعايير المستخدمة للتقويم في هذه الدراسة.

الإطار النظري والدراسات السابقة:

1. الإطار النظري

1.1. أهمية تقويم الاختبارات:

يمثل التقويم بشكل عام عملية منهجية تتطلب جمع بيانات موضوعية ومعلومات صادقة من مصادر متعددة، باستخدام أدوات قياس متنوعة في ضوء أهداف محددة؛ بغرض التوصل إلى تقديرات كمية، وأدلة كيفية يُستند إليها في إصدار أحكام، ويمتد التقويم ليشمل كل العناصر المرتبطة بالعملية التعليمية (علام، 2000)؛ لذا فالتقويم يمثل أحد الأركان الأساسية للعملية التعليمية، ونتائجه تشكل إطاراً مرجعياً لكل عمليات التحسين والتطوير التي يجب أن تصاحب العملية التعليمية بعناصرها كافة، كما تمثل نتائج التقويم تحصيل الطلبة، وتعد بالنسبة للأستاذ الجامعي عاملاً مساعداً على التخطيط السليم للخبرات التعليمية المناسبة، وتحديد مدى تحقيق طلبته لمخرجات التعلم، وتمكنه من تقويم أسلوبه وطريقته في التدريس، ولذا فإن القدرات التي يتخذها الأستاذ الجامعي تكون مبنية على أسس وأساليب صحيحة (مجيد، 2014، صـ35).

ولما تمثله الاختبارات التحصيلية ونتائجها من أهمية بالغة في عملية التقويم، ولأنها تعد مؤشراً جيداً عن مستوى التحصيل لطلبة وقدراتهم ونشاطهم، فمن الأهمية بمكان أن تبنى هذه الاختبارات على وفق الأسس العلمية التي تساعد على تحقيق الجودة المطلوبة فيها؛ لكي تؤدي وظائفها بكفاية عالية، ولذا فإن عملية تقويم الاختبارات تسهم في تشخيص نقاط القوة والضعف فيها والعمل على تحسينها وتطويرها.

2.1. جودة الاختبارات التحصيلية:

تتمثل جودة الاختبارات التحصيلية في مطابقة هذه الاختبارات لشروط بناء أسئلتها في أنواعها المتعددة بحسب ما هو موصف في الأدب النظري للاختبارات والمقاييس، بالإضافة إلى المواصفات الفنية والشكلية لتصميم الورقة الاختبارية، وتوصف هذه الشروط بمعايير الجودة للاختبارات التحصيلية، ويسرد الأدب النظري العديد من المعايير لكل نوع من أسئلة الاختبارات كلاً على حدة كما يأتي:

2.1.1. الأسئلة المقالية

تكمن أهمية الأسئلة المقالية في أنها تقيس قدرة الطالب على التعبير بأسلوبه الخاص، وطرح الأفكار، والتعرف على وجهات النظر، ومن ثم فهي تقيس التعلم ذا المعنى، كما تحقق مبدأ من مبادئ الدماغ: وهو كل دماغ فريد بذاته أي إن عملية اكتساب المعلومات وتنظيمها ومعالجتها في الدماغ تختلف من شخص إلى آخر؛ لذلك يمكن من خلالها التعرف على مدى اكتساب الطلبة لمهارات التفكير العليا مثل التقويم والإبداع. كما تتميز هذه الأسئلة بأنها سهلة الإعداد، وتلغي احتمالية التخمين، وتعطي حرية الإجابة التي توفرها للطالب وتتمثل في القدرة على

انتقاء المعلومات وتنظيمها وتركيبها، والقدرة على الإبداع والنقد والابتكار. (سليمان وآخرون، مشروع تطوير نظم تقويم الطلاب والامتحانات، بدون)، وفي المقابل يؤخذ على الأسئلة المقالية سلبيات منها أنها تحتاج وقتاً طويلاً لتصحيحها، كما أنها منخفضة الصدق والثبات؛ لأنها تتأثر بذاتية المصحح، وغير شاملة للمحتوى. (عثمان، 2010)

وتكمن معايير جودة الأسئلة المقالية في تنويع الأسئلة لقياس المستويات الدنيا والعليا من التفكير، على أن تكون الأسئلة غير مركبة أي لا تحتوي على أكثر من فكرة في السؤال نفسه، وأن تكون الإجابة واضحة ولا تحتمل أكثر من تفسير، كما يجب أن تكون الأسئلة شاملة لعمليات عقلية وأهداف سلوكية (مخرجات تعلم) مختلفة، ويفضل أن تكون موزعة باستعمال أفعال مثل عدَد، انقد، اقترح، صمم، حلل... إلخ، حيث يمكن للأسئلة المقالية أن تقيس مخرجات التعلم في مجالاتها المختلفة، ففي مجال المعرفة والفهم تستخدم أفعالاً دالة مثل عرِّف ، وضح، بين، اذكر، عدد، صف، اشرح، فسر، وفي مجال المهارات الذهنية تستخدم أفعالاً مثل حلل، ركب، اربط، خطط، قارن، اقترح، عدّل، و في مجال المهارات المهنية تستخدم أفعالاً مثل: صمَم، طور، استخدم، اتقن، كما يجب أن تصاغ الأسئلة بما يتناسب مع الوقت المحدد للإجابة. (القدسي وآخرون، 2020)

2.1.2. الأسئلة الموضوعية:

يقصد بالأسئلة الموضوعية التي تكون إجاباتها محددة، ولها إجابة صحيحة واحدة لكل سؤال، وتعتمد على الإجابة النموذجية بوصفه معياراً للتصحيح يعتمد عليه جميع المصححين. وتتميز هذه الأسئلة بالشمول، وبدرجة عالية من الصدق والثبات، ولا تستغرق وقتاً وجهداً عند تصحيحها، كما أنها تلغي احتمالية التورية، ولا تتأثر بذاتية المصحح، وفي المقابل يؤخذ على هذا النوع من الأسئلة: أنها لا تقيس قدرة الطالب على التعبير الكتابي، وتعجز عن قياس الجانب الانفعالي من السلوك، والعمليات العقلية العليا، وتستغرق وقتاً وجهداً في إعدادها، وتحتاج مهارة عالية لصياغتها، كما أنها تشجع على تجزئة المعرفة. (سليمان وأبو علام، 2010) ويوجد أربعة أنواع من الاختبارات الموضوعية هي:

1. أسئلة الصواب والخطأ

هي عبارات إما أن تكون صحيحة أو غير صحيحة، ويطلب من الطالب تحديد ما إذا كانت العبارة صحيحة أو خاطئة. وتتمثل معايير جودة هذه الأسئلة في تجنب العبارات التي تحتوي على النفي بقدر الإمكان، وإذا لزم الأمر ذلك فإنه لا بد من إبراز كلمة النفي، وعدم استخدام كلمات مثل غالباً ، نادراً ، دائماً ، أبداً ، في كثير من الأحيان، إلى حد ما ...إلخ، ويفضل أن تكون عدد العبارات الصحيحة مساوياً لعدد العبارات الخاطئة تقريبا، ولكن دون مطابقة تامة، وأن ترتب بشكل عشوائي، كما ينبغي أن تتضمن العبارة أو السؤال فكرة واحدة فقط، وأن تصاغ العبارات بإحكام بحيث تكون صحيحة تماماً أو خاطئة تماماً، وألا تتطلب الإجابة عليها قيام الطالب بعمليات تحريرية كثيرة وطويلة، بالإضافة إلى عدم نقل عبارات من المقرر الدراسي واستخدامها جزئياً في فقرات الصواب والخطأ؛ لأن ذلك يشجع الطلاب على الحفظ دون الفهم. (القدسي وآخرون، 2020)

2. أسئلة الاختيار من متعدد:

هي عبارة عن سؤال مكون من جزأين الأول: أصل السؤال والثاني: بدائل متعددة تقدم إجابات ممكنة للمشكلة التي عرضت في الأصل، وعلى الطالب أن يختار البديل الصحيح من بينها.

وتتمثل معايير جودة أسئلة الاختيار من متعدد في شروط خاصة بأصل السؤال والبدائل، وذلك بأن يحتوي السؤال مشكلة محددة تماماً بحيث يفهم المتعلم المقصود من المشكلة دون حاجة إلى الاستعانة بالبدائل لتوضيحها، ولا يتضمن تلميحات تساعد الطالب على التوصل إلى الإجابة الصحيحة، وتستعمل صيغة النفي فقط عندما تكون هي الأنسب، وفي حالة وجودها يجب أن تبرز بشكل واضح (السنداني، 2009). وفيما يخص شروط صياغة البدائل فتتمثل في أن يكون لكل سؤال بديل واحد فقط يمثل الإجابة الصحيحة، وألا يكون موقع البديل الصحيح ثابتاً في جميع الفقرات، وأن يكون عدد البدائل ما بين 3-5، وأن تتجانس البدائل الخاطئة مع البديل الصحيح، وتتفق نحويّاً مع أصل السؤال، وألا تحتوي على كلمات مكررة، حيث توضع الكلمات المشتركة في مقدمة السؤال، كما يجب أن تكون جميع البدائل محتملة ومعقولة ولها ارتباط بالسؤال ومتجانسة من حيث الشكل والصياغة، وأن ترتب الأعداد والتواريخ فيها (إن وجدت) تصاعديًّا / تنازليًّا. (سليمان وأبو علام، 2010)

3. أسئلة إكمال الفراغ:

هي عبارة ناقصة ولا يتم المعنى إلا بوضع الكلمة أو الجملة المكملة لها. وتتمثل معايير جودة هذا النوع من الأسئلة في أن يتناول كل سؤال أو عبارة هدفاً واحداً فقط، وأن تكون العبارات متقاربة في الطول والصياغة العامة، وأن تحتوي العبارة الواحدة على فراغ واحد أو اثنين على الأكثر؛ لأن كثرة الفراغات تفقد العبارة وضوحها ومعناها، وينبغي أن يكون موقع الفراغ عند قرب العبارة أو نهايتها، وأن يكون ملء الفراغ بجملة أو كلمة معينة لا يمكن وضع غيرها وألا يحتمل الفراغ أكثر من إجابة، وألا تنقل العبارة بالنص من الكتاب حتى لا تقيس مستوى التذكر فقط، كما ينبغي أن تصاغ العبارة الناقصة أو السؤال بإحكام، بحيث لا يحتمل الفراغ سوى إجابة صحيحة واحدة، وأن تحذف الكلمات أو المصطلحات الرئيسة وليس الكلمات غير المهمة أو الإنشائية التي تساعد على تكوين الجملة. (مجيد، 2014، صـ251)

4. أسئلة التوصيل (المطابقة)

هي مجموعة من العبارات (مقدمات) وإجابات تنظم في عمودين يحتوي الأول: على عدد من العبارات، ويحتوي الثاني: على الإجابات، ويطلب من الطالب أن يختار من العمود الثاني الإجابة المناسبة للعبارة المعطاة في العمود الأول. وتتمثل معايير جودة أسئلة التوصيل في أن يكون عدد فقرات المقدمات غير متساوٍ مع عدد فقرات الإجابات، وأن تكون التعليمات واضحة ببيان مكان لوضع رقم أو حرف الإجابة، وألا يكون ترتيب المقدمات بترتيب الإجابات نفسه، وألا تزيد العبارات في كل عمود عن ست أو سبع، وأن يوجد عنوان لعمود المقدمات، وآخر لعمود الإجابات، وأن تكون جميع العبارات متجانسة، وأن توضع كل من القائمتين في صفحة واحدة. (السنداني، 2009)

3.1. المعايير الضمنية للورقة الاختبارية:

تشمل المعايير الضمنية للورقة الاختبارية في هذه الدراسة معايير صياغة الأسئلة المقالية والموضوعية وقد أشير لتلك المعايير سابقاً، بالإضافة إلى معايير عامة للاختبار ومخرجات التعلم، حيث تتمثل المعايير العامة في التنوع في فقرات الاختبار لمحتوى المقرر الدراسي، وتناسب عدد الأسئلة مع الزمن الكلي للاختبار، وتجنب الأسئلة الاختيارية، ومراعاة الفروق الفردية بين الطلبة، واستقلال الأسئلة فلا يعتمد أحدهما على الآخر، وقياس الأسئلة جوانب مختلفة من مخرجات التعلم (القدسي وآخرون، 2020، صـ6). ويمكن أن نضيف معيار عدم تكرار نموذج الاختبار من أعوام سابقة أو كليات متعددة؛ لكيلا يعتمد الطلبة على الأسئلة نفسها في النموذج.

4.1. مخرجات التعلم.

ترّكز مخرجات التعلم على ما هو متوقع من الطالب تحقيقه من اكتساب المعرفة وفهمها والقدرة على تطبيقها بعد عملية التعلم، وكيف يمكن إظهار هذا التعلم، بدلاً من التركيز على ما درسه الطالب، فمخرجات التعلم تمثل مؤشرات عما يتوقع من الطالب تحقيقه بعد الانتهاء من عملية التعلم (Kennedy, 2007, p21)؛ لذلك فإن الصيغة السليمة لمخرجات التعلم لمقرر دراسي تعتمد على استخدام أفعال دقيقة مستقاة إلى حد كبير من تصنيف بلوم لمستويات التعلم المعرفي والوجداني والمهارى؛ وبناءً على ذلك فإن أسئلة الاختبار ينبغي أن تعكس مخرجات التعلم التي يراد قياسها، وهذا بدوره يفرض على أستاذ المقرر أن يصيغ مخرجات التعلم بصورة دقيقة قابلة للقياس.

وتتكون مخرجات التعلم من مكونين رئيسين: الأول المعرفة والفهم الذي يتمثل فيما ينبغي أن يعرفه الطالب ويفهمه بعد دراسة المقرر الدراسي، ويشمل اكتساب الحقائق، والمفاهيم، والمبادئ، والقواعد، والقوانين، والنظريات التي يتطلبها دراسة المقرر الدراسي، والثاني المهارات وتنقسم إلى ثلاثة أقسام الأولى: المهارات الذهنية التي تتمثل في المستويات العليا من التفكير كالتحليل والتركيب والتقويم والإبداع، والثانية: المهارات التخصصية أو المهنية: وتتمثل في القدرة على استخدام المعرفة المكتسبة بعد دراسة المقرر الدراسي بنجاح في التطبيقات المهنية، والثانية المهارات العامة والانتقالية: وهي مهارات عامة غير مرتبطة بمجال التخصص فحسب، بل إنها قابلة للتطبيق في مجالات الحياة المختلفة، بالإضافة إلى أهمية استخدامها عند ممارسة المهنة، على سبيل المثال لا الحصر: القدرة على التواصل والعمل في فريق، ومهارات التفاوض، واستخدام تكنولوجيا المعلومات. (القدسي وآخرون، 2021؛ كنيدي، 2013)

والجدير ذكره أن المهارات المهنية تقاس من خلال الأنشطة، والاختبارات العملية، وتقديم العروض، وعمل المشاريع؛ لذلك اقتصرت هذه الدراسة على مخرجات التعلم في مجالي: المعرفة والفهم، والمهارات الذهنية؛ لأن عينة الدراسة الاختبارات النظرية فقط.

5.1.المعايير الفنية للورقة الاختبارية: تتمثل هذه المعايير في المواصفات الآتية: (العلي، 2013؛ القدسي وآخرون،2020؛ النور وآخرون 2021)

1. خلو ورقة الاختبار من الأخطاء اللغوية والمطبعية وطباعتها على الكمبيوتر.
2. سلامة ترقيم الأسئلة وفروعها وصفحات الاختبار، و تحديد درجة كل سؤال رئيس وفرعي.
3. ترتيب مفردات كل نوع من أنواع الأسئلة معاً (صواب وخطأ، اختيار من متعدد، إكمال...).
4. تنظيم ورقة الاختبار ووجود هوامش مناسبة، ومكان مناسب للإجابة في حال كانت الإجابة في الورقة نفسها.
5. الفصل بين التعليمات العامة والأسئلة بفاصل(خط)، ووضوح تعليمات الإجابة، وكتابتها بخط بارز.

6.1. المعايير الشكلية للورقة الاختبارية:

تنقسم الورقة الاختبارية من حيث الشكل إلى أربعة أجزاء هي: رأس، متن، حواشي جانبية، ذيل، ولكل جزء معايير خاصة به تتمثل كما في الآتي: (جيوسي، 2016؛ السفياني، 2020؛ دليل جودة الاختبارات، 2022)

6.1.1 الرأس (الإكليشة) ويتضمن البيانات التعريفية الآتية:

1. الجهة اليمنى من الإكليشة التي تشمل اسم : البلد والوزارة والجامعة والكلية والقسم.
2. الجهة اليسرى من الإكليشة وتشمل: المستوى الدراسي، اسم المقرر، تاريخ الاختبار، زمن الاختبار.
3. الوسط ويشمل شعار الجامعة أو الكلية.
4. تحت الشعار يكتب بخط بارز نوع الامتحان(نصفي / نهائي)، والعام الدراسي.

ونرى أن يضاف إلى الإكليشة اسم أستاذ المقرر بدلاً عن الهامش السفلي؛ ليسهل التوضيح للكنترول.

6.1.2 متن الورقة ويشمل بُعدين هما:

1 .التعليمات العامة: تكتب بعد الإكليشة مباشرة وتشمل: مكان وطريقة الإجابة عن الأسئلة، والأدوات المسموح باستخدامها في الاختبار، بالإضافة إلى بعض القيم الثابتة التي يحتاجها الطالب، وعدد الأسئلة، والأوراق.

2 . الأسئلة المطلوب الإجابة عنها من قبل الطالب.

6.1.3 الهوامش الجانبية وتشمل على بُعدين: الأول الهامش الأيمن يكتب فيه أرقام الأسئلة الرئيسة، والثاني الهامش الأيسر يكتب فيه الدرجات التفصيلية للأسئلة.

6.1.4 الذيل (الهامش السفلي): ويتضمن بيانات تعريفية تتمثل في أرقام الصفحات، واسم أستاذ المقرر، وكتابة (تابع) أو للأسئلة بقية، في حال استخدام أكثر من ورقة، وعبارة التمنيات بالتوفيق، إلى جانب عبارة انتهت الأسئلة.

2. الدراسات السابقة:

دراسة الصلاحي (2008) هدفت إلى تقويم الاختبارات النهائية في التخصصات الإنسانية بجامعة صنعاء، لمعرفة مدى مطابقتها للمعايير العلمية للاختبارات التحصيلية وتكونت عينة الدراسة من (86) ورقة اختبارية

للفصل الأول من العام الجامعي 2004-2005م، وتكونت أداة الدراسة من استمارة تقييم الاختبارات، واستُخدم المنهج الوصفي التحليلي، وأظهرت النتائج أن معظم الأسئلة تركزت في النوع المقالي، وأن الاختبارات لم تكن شاملة للمقررات الدراسية وقواعد بناء الاختبارات، وأن بعض الأوراق أغفلت الجوانب الفنية للورقة الاختبارية، وأوصت الدراسة بتدريب أعضاء هيئة التدريس لتطوير قدراتهم في إعداد الاختبارات وبنائها.

وهدفت دراسة غليون (2009) إلى تقويم الاختبارات النهائية للدراسات العليا بكلية التربية جامعة صنعاء في ضوء معايير الاختبار الجيد، وتكونت العينة من (77) ورقة اختبارية لبرنامج الماجستير في مناهج وطرق تدريس، وأصول التربية، وعلم النفس، للعام الجامعي 2007-2008م، وتكونت أداة الدراسة من استمارة تحليل شملت معايير التقييم المعدة في الدراسة، واستُخدم المنهج الوصفي التحليلي، وأظهرت النتائج عدم شمول الاختبارات لموضوعات المقرر، ووجود قصور في الأسئلة المقالية النقدية التي تتمثل في وجهات النظر من قبل الطالب، وأوصت الدراسة بتدريب أعضاء هيئة التدريس المشاركين في برنامج الدراسات العليا على مبادئ التقويم الحديث، والاهتمام بتحليل نتائج الاختبارات سنوياً، وتضمينها في خطة تطوير عملية التعليم والتعلم.

كما هدفت دراسة العلي (2013) إلى تقويم جودة الاختبارات التحصيلية لأعضاء هيئة التدريس بكلية التربية – جامعة سلمان بن عبد العزيز، وتكونت أداة الدراسة من استمارة تقييم الاختبارات، وطبقت على (74) اختباراً، واستُخدم المنهج الوصفي التحليلي، وأظهرت النتائج أن أغلب الأسئلة الاختبارية غير شاملة؛ إذ إنها تقيس مدى تحقق الأهداف المعرفية للمقررات المختلفة، ولم تتطرق إلى الأهداف التعليمية الأخرى، وأوصت الدراسة بإجراء تقويم دوري لواقع جودة الاختبارات التحصيلية لأعضاء هيئة التدريس وتدريبهم على كتابة الأسئلة التي تقيس المستويات العليا. وفي السياق نفسه هدفت دراسة ساعد وعامر (2017) إلى التعرف على مدى توافر معايير الاختبار التحصيلي الجيد في الاختبارات التي يعدها أساتذة التعليم الجامعي، وتكونت عينة الدراسة من (36) ورقة اختبارية للعام الجامعي 2012-2013م في جامعة محمد بسكرة، واستُخدم المنهج الوصفي، وأظهرت النتائج أن نسبة الأسئلة المقالية أكثر من الموضوعية، وفيما يخص معايير طباعة الاختبارات فقد كانت بنسبة متوسطة، أما كتابة تعليماته فقد كانت ضعيفة، بالإضافة إلى إن إعداد فقراته جاءت بنسبة ضعيفة، وأوصت الدراسة بتدريب الأساتذة على بناء الاختبارات التحصيلية الموضوعية.

ومن حيث شمول الاختبارات ومستويات الصعوبة والتمييز فقد هدفت دراسة بافطوم والسفياني (2019) إلى التعرف على مدى تمثيل الاختبارات الفصلية لأوزان الوحدات الدراسية، وأهداف بلوم المعرفية، ومستويات الصعوبة والتمييز فيها، وتكونت عينة الدراسة من (13) اختباراً للفصل الدراسي الأول للعام الجامعي 2018-2019م بكلية التربية – جامعة حضرموت، وتكونت أداة الدراسة من استمارة تحليل، واستخدمت المنهج الوصفي التحليلي، وتوصلت الدراسة إلى أن نسبة الأسئلة الموضوعية أكثر من المقالية، وأن هناك تبايناً في تمثيل الاختبارات الفصلية لأوزان الوحدات الدراسية، بالإضافة إلى ارتفاع نسبة تمثيل أهداف مستويي: (الفهم والتذكر)

حيث بلغت (67.3%) في حين انخفضت نسبة تمثيل المستويات العليا حيث بلغت (32.3%)، وأوصت الدراسة بضرورة بناء الاختبارات الفصلية في ضوء جدول المواصفات ومعايير الاختبار الجيد.

وفيما يخص الورقة الاختبارية بشكل عام فقد هدفت دراسة السفياني (2020) إلى التعرف على مواصفات إعداد الورقة الاختبارية ومعايير صياغة فقراتها، وتكونت عينة الدراسة من (8) امتحانات نهائية لمقررات تربوية (متطلبات الكلية) للفصل الدراسي الثاني للعام الجامعي 2018-2019م بكلية التربية- محافظة المهرة، واستخدمت استمارة تحليل لتحقيق هدف الدراسة، كما استخدام المنهج الوصفي التحليلي، وأظهرت النتائج أن توافر المواصفات الجيدة للورقة الاختبارية التي يعدها أعضاء هيئة التدريس في الكلية جاءت بدرجة عالية، كما توافرت معايير صياغة الفقرات الموضوعية والمقالية والموصفات الشكلية للورقة الاختبارية بنسبة عالية حيث حصلت على الترتيب الأول والثاني والثالث على التوالي، في حين جاءت الموصفات الضمنية للورقة الاختبارية في الترتيب الأخير وبدرجة متوسطة، وأوصت الدراسة بضرورة مراعاة أعضاء هيئة التدريس لمواصفات الورقة الاختبارية الشكلية والضمنية، ومراعاة صياغة الفقرات الاختبارية بشكل سليم.

وهدفت دراسة الذري (2021) إلى تقويم اختبارات التحصيل التربوي بكلية التربية- جامعة ذمار في ضوء معايير جودة الاختبارات، وتكونت عينة الدراسة من (15) اختباراً من اختبارات المقررات التربوية للفصل الأول 2020-2021م، وتوصلت الدراسة إلى أن معايير جودة الورقة الاختبارية في اختبارات التحصيل للمقررات التربوية بالكلية توافرت بدرجة متوسطة بلغ نسبة توافرها (58.9%)، وأن معايير جودة الأسئلة ككل توافرت بدرجة متوسطة، إذ بلغ نسبة توافرها بشكل إجمالي (59.1%)، في حين تفاوتت نسب توافر المحاور المتعلقة بها، فقد جاءت معايير المحور الخاص بالأسئلة الموضوعية بدرجة عالية جداً وبنسبة توافر (82.2%) ، وجاءت معايير المحور الخاص بالمعايير العامة بدرجة متوسطة وبنسبة توافر (50.8%)، وجاءت معايير المحور الخاص بالأسئلة المقالية بدرجة ضعيفة وبنسبة توافر (44.4%)، وأوصت الدراسة بضرورة عقد دورات تدريبية لتنمية مهارات أعضاء هيئة التدريس في إعداد اختبارات التحصيل للمقررات التربوية في ضوء معايير جودة الأسئلة الاختبارية.

التعليق على الدراسات السابقة:

تتفق الدراسة الحالية مع جميع الدراسات السابقة من حيث استخدامها للمنهج الوصفي التحليلي، واستمارة تحليل المحتوى، وفي تناولها لموضوع تقويم الاختبارات التحصيلية الجامعية على وفق معايير الاختبار الجيد ومواصفاته، الذي يمثل هدف الدراسة، فيما تتميز هذه الدراسة عن الدراسات السابقة في مجتمعها وعينتها، حيث نجد أن جميع الدراسات السابقة تناولت مقررات تربوية واختصت بكلية التربية، فيما تناولت هذه الدراسة مقررات الطب البشري، وهذا ما اختلفت به الدراسة الحالية عن الدراسات السابقة، لا سيما أن أعضاء هيئة التدريس ومعاونيهم في هذه الكلية قد يفتقرون إلى معايير الاختبارات الجيدة بحكم تخصصاتهم العلمية، وقد استفادت الباحثة من جميع الدراسات في بناء قائمة المواصفات والمعايير والاستشهاد بها في مناقشة النتائج.

منهجية الدراسة وإجراءاتها:

1. منهج الدراسة: استُخدم المنهج الوصفي التحليلي في استقراء الأدب النظري والدراسات السابقة المرتبطة بجودة الاختبارات؛ بهدف تحديد معايير جودة الاختبارات التحصيلية، وبناء أداة الدراسة ومتطلباتها، كما استُخدم أسلوب تحليل المضمون في تحليل الأوراق الاختبارية.

2. مجتمع الدراسة: تكون مجتمع الدراسة من جميع اختبارات مقررات الطب البشري للعام الجامعي 2022-2023م بكلية الطب جامعة عمران، وعددها (44) مقرراً كما ورد في توصيف برنامج الطب والجراحة.

3.عينة الدراسة: تكونت عينة الدراسة من (26) اختباراً، اختيرت بطريقة العينة العشوائية الطبقية من أصل (44) اختباراً من اختبارات العام الدراسي 2022-2023م لمقررات الطب البشري أي بنسبة 59% كما ورد في توصيف برنامج الطب والجراحة حيث احتوى التوصيف على (44) مقرراً، منها ثمانية مقررات متطلبات جامعة، وستة مقررات متطلبات كلية، والباقي متطلبات برنامج، حيث اعتمد توزيع الطبقات في العينة على نوع المتطلب: (كلية، جامعة، برنامج)، و الجدول الآتي يوضح خصائص عينة الدراسة:

جدول (1): يوضح خصائص عينة الدراسة من حيث نوع المتطلب والمستوى الدراسي.

م	المقرر	المتطلب	المستوى	م	المقرر	المتطلب	المستوى
1	لغة إنجليزية	جامعة	الأول	14	طب مجتمع (3)	برنامج	الثالث
2	لغة عربية	جامعة	الأول	15	طب مجتمع (5)	برنامج	الرابع
3	الثقافة الإسلامية	جامعة	الأول	16	علم الطفيليات (1)	برنامج	الثالث
4	مهارات حاسوب	جامعة	الأول	17	علم الطفيليات (2)	برنامج	الثالث
5	مصطلحات طبية	كلية	الأول	18	باطنة (1)	برنامج	الرابع
6	بيولوجيا الإنسان	كلية	الأول	19	باطنة (2)	برنامج	الرابع
7	كيمياء عامة	كلية	الأول	20	طب أطفال (1)	برنامج	الرابع
8	علم المناعة والجينات الوراثية	برنامج	الثاني	21	طب أطفال (4)	برنامج	الخامس
9	علم النفس	برنامج	الخامس	22	جراحة (2)	برنامج	الرابع
10	علم الأحياء الدقيقة (1)	برنامج	الثالث	23	علم الأمراض العام	برنامج	الثاني
12	علم الأحياء الدقيقة (2).	برنامج	الثالث	24	علم أمراض (1)	برنامج	الثالث
13	علم وظائف الأعضاء (2).	برنامج	الثاني	25	طب النساء والتوليد (1)	برنامج	الرابع
13	طب مجتمع (2)	برنامج	الثاني	26	أنسجة (3).	برنامج	الثاني

4. أداة الدراسة:

بُنيت أداة الدراسة وهي استمارة التحليل من خلال الاسترشاد بالأدب النظري والدراسات السابقة والأدلة، مثل دليل عضو هيئة التدريس في بناء الاختبارات التحصيلية بجامعة صنعاء للعام 2020م ذات الصلة بمواصفات الاختبار الجيد ومعاييره، وتمثلت الأداة في قائمة معايير تكونت من (61) مواصفة ومعياراً توزعت على ثلاثة محاور رئيسة هي:

1. المحور الشكلي للورقة الاختبارية: وشمل على (23) مواصفة كما هو موضح في جدول (11)، حيث توزعت هذه المواصفات على أربعة محاور فرعية الأول: المواصفات المطلوبة في الهامش العلوي للاختبار (الإكليشة) وتكون من اثنتي عشرة مواصفة للتقييم، والثاني المواصفات المطلوبة في كتابة تعليمات الاختبار وتكون من خمس مواصفات للتقييم، الثالث: المواصفات المطلوبة في الهوامش الجانبية: (الأيمن والأيسر) وتكون من ثلاث مواصفات، الرابع: المواصفات المطلوبة في الهامش السفلي للاختبار وشمل على ثلاث مواصفات للتقييم.

2.المحور الضمني للورقة الاختبارية وشمل على (32) معياراً توزعت على ثلاثة محاور فرعية الأول: معايير صياغة الأسئلة وتكون من خمسة محاور فرعية أخرى هي معايير الأسئلة: المقالية، والصواب والخطأ، والاختيار من متعدد، وأسئلة التكميل، وأسئلة المطابقة، وشملت جميعها (27) معياراً كما هو موضح في جداول: (3، 4، 5، 6، 7) على التوالي، أما المحور الثاني فقد تكون من ثلاثة معايير ضمنية عامة، في حين تضمن المحور الثالث مدى قياس الاختبار لمخرجات التعلم وشمل معيارين، كما هو موضح في جدول (9).

3.المحور الفني للورقة وشمل ستة معايير للتقييم، تحددت كما هو موضح في جدول (12).

5. **صدق الأداة:** اعتمدت الباحثة على الصدق الظاهري، من خلال عرض الأداة على مجموعة من الخبراء في القياس والتقويم وعلم النفس التربوي والمناهج وطرق التدريس؛ لإبداء ملحوظاتهم وقد أخذت الباحثة بهذه الملحوظات.

6. ثبات الأداة: لقياس الثبات حُللت أوراق اختبار مقرري مصطلحات طبية نموذج (B) وعلم النفس الطبي لدور التكميلي من قبل الباحثة بوصفها عينة استطلاعية، وحسب معامل الاتفاق على وفق المعادلة الآتية: (حنان ونعيم، 2020، صـ120)

$$\text{معامل الاتفاق} = \frac{\text{عدد الحالات المتفق عليها}}{\text{عدد الحالات الكلية}}$$

جدول (2): يوضح معامل الاتفاق

الاتفاق في التحليل	عدد حالات الاتفاق	نسبة معامل الاتفاق
الاتفاق عبر الزمن	55	90%
الباحثة مع باحث آخر	54	89%
متوسط الثبات	90%	

7. إجراءات التطبيق: للقيام بعملية التحليل اتبعت الباحثة الإجراءات الآتية:

1. خوطب عميد كلية الطب للتنسيق والتعاون مع الباحثة للحصول على البيانات والاختبارات اللازمة للدراسة.

2. حللت الباحثة كل اختبار على حدة باستمارة تحليل خاصة بكل محور رئيس حيث حللت الباحثة محتوى أوراق الاختبارات في ضوء كل معيار بشكل مستقل، ورصدت التوافر أو عدم التوافر في استمارة التحليل الخاص بكل محور.

3. تُرجمت الاختبارات المطبوعة باللغة الإنجليزية والاستعانة بمختص بهذه اللغة، وأنجزت قائمة تضم الأخطاء اللغوية والمطبعية والإملائية، ثم عرضت هذه القائمة على ذوي الاختصاص في اللغة الإنجليزية وصوبت، كما أنجزت قائمة الأخطاء اللغوية والمطبعية والإملائية التي اشتملت عليها الاختبارات المطبوعة باللغة العربية، ثم صوبها متخصصون في اللغة العربية.

4. لتحليل محتوى أوراق الاختبارات في ضوء المواصفات الشكلية للورقة الاختبارية قرأت الباحثة المعلومات الواردة في الهوامش العلوية والسفلية والجانبية، بالإضافة إلى المعلومات الواردة عن تعليمات الاختبار، ورصدت توافر المواصفات أو عدم توافرها في استمارة التحليل.

5. لتحليل محتوى أوراق الاختبارات في ضوء المواصفات الفنية للورقة الاختبارية قرأت الباحثة كل ورقة اختبار وتأكدت من سلامة ترقيم الأسئلة وفروعها دون تداخل، ومن وجود مسافات مناسبة بين الأسئلة وبين فقرات السؤال الواحد، ووجود مكان مناسب وكافٍ للإجابة في حال كانت الإجابة في الورقة نفسه، كما تأكدت الباحثة من خلو الاختبار من الأخطاء اللغوية والمطبعية، من خلال عمل قائمة تضم جميع الكلمات والجمل التي لوحظ أنها تحتوي على نوع من الأخطاء السابقة، ثم عرضتها على ذوي الاختصاص في هذه المجال؛ من أجل تصويبها ومن ثم رصد كل ذلك في استمارة التحليل.

6. لتحليل محتوى الأسئلة في ضوء المعايير الضمنية فيما يخص جودة الأسئلة المقالية والموضوعية فقد قرأت الباحثة كل فقرة بدقة وتأكدت من تحقق المعايير أو عدم تحققها، ورصدت ذلك في استمارة التحليل. وفيما يخص المعايير العامة من حيث تناسب عدد الأسئلة مع الزمن الكلي للاختبار فقد تم قرأت الباحثة زمن الاختبار في الأوراق الاختبارية وقارنته مع الزمن المحدَد للمحاضرات في توصيف المقرر، وبالنسبة للأوراق التي لم تحتوِ على زمن الاختبار، فقد تواصلت الباحثة تم التواصل مع المختصين في الكنترول أو مع أستاذ المقرر. وفيما يخص شمول الاختبار لمحتوى المقرر، فقد قرأت الباحثة أسئلة كل اختبار على حدة، ومن ثم قرأت ولاحظت الموضوعات المتعلقة بالاختبار في توصيف المقرر، حيث اعتمدت الباحثة التوصيف للعام 2021-2022م الذي توفر بوصفه محكاً لمعرفة مدى شمول أسئلة الاختبار لموضوعات المقرر، من خلال حساب عدد الموضوعات التي تغطيها أسئلة الاختبار وقسمتها على العدد الكلي لموضوعات المقرر وضرب الناتج × (100)، فإذا كان الناتج (80%) فما فوق عُد المعيار متوافراً، وإذا كان أقل من هذه النسبة عُد غير متوافر.

7.اعتمدت الباحثة على مخرجات التعلم المحددة في توصيف كل مقرر لتحليل معياري، بحيث يغطي الاختبار مخرجات التعلم في مجال المعرفة والفهم، والمهارات الذهنية، حيث قارنت أسئلة الاختبار وفقراته في ضوء

مخرجات التعلم المحددة في التوصيف والدراسة، ومن ثم رصدت المخرجات المتوافرة وغير المتوافرة في الاختبارات في استمارة التحليل.

8. للحكم على درجة توافر المعيار اعتمدت الباحثة على المحك الآتي:

المعيار المتوافر من	100-80	80-65	65-50	أقل من 50
درجة التوافر	عالية جداً	عالية	متوسطة	ضعيفة

الأساليب الإحصائية: استخدمت الباحثة التكرارات والنسب المئوية، ومعادلة (Scott) لحساب معامل الاتفاق.

نتائج الدراسة ومناقشتها:

للإجابة عن السؤال الأول الذي ينص على: ما مواصفات جودة الاختبارات التحصيلية ومعاييرها؟

للإجابة عن هذا السؤال صممت الباحثة قائمة بمواصفات جودة الاختبارات التحصيلية و معاييرها تضمنت (61) مواصفة و معياراً فُصِّل في أداة الدراسة سابقاً.

إجابة السؤال الثاني الذي ينص على: ما نسبة توافر المعايير الضمنية للاختبارات التحصيلية لمقررات الطب البشري بكلية الطب – جامعة عمران للعام الدراسي 2022-2023م؟

للإجابة عن هذا السؤال قُسِّم إلى أربعة أسئلة فرعية، وكانت الإجابة عن كل سؤال على حدة من خلال إجراء عملية التحليل الإحصائي ببرنامج الـ spss كما يأتي:

1. ما مدى جودة أسئلة الاختبارات المقالية لمقررات الطب البشري بكلية الطب – جامعة عمران للعام الدراسي 2022-2023م؟ وتظهر إجابة هذا السؤال كما في جدول (3).

جدول (3) التكرارات والنسب المئوية والترتيب لنسبة توافر معايير جودة الأسئلة المقالية.

المعايير	متوافر		غير متوافر		الترتيب	درجة التوافر.	عدد الاختبار التي فيها أسئلة مقالية	عدد الاختبارات الكلية
	تكرار	%	تكرار	%				
1. تجنب الأسئلة المركبة	12	80	3	20	4	عالية جداً	15	26
2. تجزئة محتوى السؤال الواحد إلى العديد من الأسئلة القصيرة التي تحمل كل منها فكرة واحدة	12	80	3	20	2	عالية جداً	15	
3. تحديد كم الإجابة أو عدد الأسطر في حال الأسئلة الطويلة.	9	60	6	40	3	متوسط	15	
4. صياغة الفقرة بحيث لا يكون للفقرة أكثر من إجابة.	15	100	-	-	1	عالية جداً	15	
الإجمالي	48	80	12	20		عالية جداً		

يتضح من جدول (3) أن الأسئلة المقالية توافرت في خمسة عشر اختباراً بنسبة (57.7%) من العدد الكلي للاختبارات، كما يتضح أن المعايير الضمنية للورقة الاختبارية فيما يخص الأسئلة المقالية توافرت بنسبة إجمالية (80%) وهي نسبة عالية جداً، حيث احتل المعيار رقم (4) الخاص بصياغة الفقرة بحيث لا يكون للفقرة أكثر من إجابة المرتبة الأولى بنسبة توافر (100%) يليه معياري رقم (1، 2) اللذان ينصان على تجنب الأسئلة المركبة، وتجزئة محتوى السؤال الواحد إلى العديد من الأسئلة القصيرة يدور كل سؤال منها حول فكرة واحدة حيث توافر بنسبة (80%) وهي نسبة عالية جداً، فيما حصل على المرتبة الثالثة معيار رقم (3) الخاص بتحديد كم الإجابة أو الأسطر في حال الأسئلة الطويلة بنسبة توافر (60%) وهي نسبة متوسطة. ويمكن أن تعزى هذه النتيجة إلى أن أغلب أسئلة الورقة الاختبارية في الكلية تشمل أسئلة مقالية قليلة جداً؛ إذ تتكون من سؤالين إلى ثلاثة أسئلة، وباقي الأسئلة تكون موضوعية يغلب عليها أسئلة الاختيار من متعدد والصواب والخطأ مما قد يسهم في رفع نسبة توافر معايير الأسئلة المقالية، بالإضافة إلى أن عدد هذه المعايير قليلة. وتتفق هذه النتيجة مع دراسة بافطوم والسفياني (2019)، بينما تختلف عن دراسة الذري (2021).

2. ما مدى جودة أسئلة الاختبارات الموضوعية لمقررات الطب البشري بكلية الطب – جامعة عمران للعام الدراسي 2022-2023م؟ للإجابة عن هذا السؤال قسمت الإجابة إلى أربعة فروع هي:

2.1. معايير جودة أسئلة الصواب والخطأ.

جدول (4) التكرارات والنسب المئوية والترتيب لنسبة توافر معايير جودة أسئلة الصواب والخطأ.

م	المعايير	متوافر		غير متوافر		الترتيب	درجة التوافر	عدد الاختبار
		تكرار	%	تكرار	%			
.1	صياغة العبارات بدقة، فإما أن تكون صحيحة تماماً أو خاطئة تماماً.	6	100	-	-	1	عالية جداً	6
.2	أن تتضمن العبارة فكرة واحدة فقط.	4	33.3	2	66.7	2	ضعيف	6
.3	تجنب النمطية في ترتيب العبارات الصحيحة والخاطئة.	6	100	-	-	1	عالية جداً	6
.4	تجنب تضمين الفقرة كلمات توحي بصواب العبارة أو خطئها مثل (فقط، دائماً، غالباً، جميع)	6	100	-	-	1	عالية جداً	6
.5	ألا تبدأ العبارات بالنفي	6	100	-	-	1	عالية جداً	6
الإجمالي		28	86.6	2	13.3			

يلاحظ من جدول (4) أن المعايير الضمنية لأسئلة الصواب والخطأ توافرت بدرجة عالية جداً وبنسبة إجمالية (86.6%)، حيث اشتركت المعايير رقم (1، 3، 4، 5) والخاصة بصياغة العبارات في الدقة، وأن تكون العبارات صحيحة تماماً أو خاطئة تماماً، وتجنب النمطية في ترتيب العبارات الصحيحة والخاطئة، وتجنب تضمين الفقرة كلمات توحي بصواب العبارة أو خطئها مثل: فقط ، دائماً... إلخ بدرجة عالية جداً من التوافر وبنسبة مئوية (100%) وبذلك جاءت في الترتيب الأول، وقد تعزى هذه النتيجة إلى سهولة إعداد فقرات

الصواب والخطأ، بالإضافة إلى عدد الاختبارات القليل جداً التي توافر فيها هذا النوع من الأسئلة، مما قد يسهم في رفع النسبة العامة، في حين جاء معيار رقم (2) الخاص بأن تتضمن العبارة فكرة واحدة فقط في الترتيب (2) وبنسبة مئوية بلغ توافرها (33.3%) بدرجة ضعيفة. وتتفق هذه النتيجة مع دراسة السفياني (2020) ودراسة الذري (2021).

2.2. معايير جودة أسئلة الاختيار من متعدد.

جدول (5) التكرارات والنسب المئوية والترتيب لنسبة توافر معايير جودة أسئلة الاختيار من متعدد.

المعايير	متوافر		غير متوافر		الترتيب	درجة التوافر	عدد الاختبارات
	تكرار	%	التكرار	%			
1. تجانس البدائل في الشكل والمضمون والطول	13	59.1	9	40.9	4	متوسطة	22
2. ترتيب البدائل الصحيحة بطريقة عشوائية.	22	100	-	-	1	عالية جداً	22
3. لا يتضمن أصل الفقرة ما يشير إلى الإجابة الصحيحة.	22	100	-	-	1	عالية جداً	22
4. عدد البدائل من (4-3).	20	95.9	2	9.1	3	عالية جداً	22
5. تقيس الفقرة نتاجاً تعليمياً واحداً.	22	100	-	-		عالية جداً	22
6. خلو البدائل من عبارات مثل: (كل ما ذكر، لا شيء مما ذكر)	8	36.4	14	63.6	5	ضعيف	22
7. وجود إجابة صحيحة واحدة فقط للفقرة.	22	100	-	-	1	عالية جداً	22
8. يصاغ أصل الفقرة على شكل سؤال، أو جملة ناقصة.	22	100	-	-	1	عالية جداً	22
9. يخلو أصل السؤال من أي تلميح أو إشارة للإجابة.	22	100	-	-	1	عالية جداً	22
10. وضوح معنى أصل الفقرة بحيث تدل على إجابة واحدة	21	95.5	1	4.5	2	عالية جداً	22
الإجمالي	194	88.6	26	11.8		عالية جداً	

يتضح من جدول (5) أن المعايير الضمنية فيما يخص أسئلة الاختيار من متعدد توافرت بدرجة عالية جداً وبنسبة توافر (% 88.6)، حيث اشتركت المعايير رقم (2، 3، 5، 7، 8، 9) والخاصة بترتيب البدائل الصحيحة بطريقة عشوائية، ولا يتضمن أصل الفقرة ما يشير إلى الإجابة الصحيحة، وعدد البدائل من (4-3)، وتقيس الفقرة نتاجاً تعليمياً واحداً، ووجود إجابة صحيحة واحدة فقط للفقرة، ويصاغ أصل الفقرة على شكل سؤال أو جملة ناقصة، ويخلو أصل السؤال من أي تلميح أو إشارة للإجابة على الترتيب الأول وبنسبة مئوية بلغ توافرها (100%)، في حين حصل معياري: وضوح معنى أصل الفقرة بحيث تدل على إجابة واحدة، وعدد البدائل من (4-3) على درجة عالية، وبالترتيب الثاني والثالث وبنسبة توافر (95.5%) و (59.1%) على التوالي، وحصل معيار رقم (1) والخاص بتجانس البدائل في الشكل والطول والمضمون على درجة متوسطة

وبنسبة توافر (59.1%)، بينما حصل معيار رقم (6) والخاص بخلو البدائل من عبارات كل ما ذكر أو لا شيء مما سبق على درجة ضعيفة وبنسبة توافر (36.4%).

ويمكن أن تعزى هذه النتيجة إلى أن أسئلة الاختيار من متعدد أصبحت شائعة لاسيما مع التوجه نحو أتمتة الاختبارات وعمل بنك الأسئلة في عدد من الجامعات، وهو التوجه الذي يبدأ من اختبارات الشهادة الأساسية والثانوية، بالإضافة إلى اختبارات القبول لا سيما في كليتي الطب والهندسة؛ مما يسهم في معرفة معايير هذا النوع من الأسئلة وانتشارها لدى أعضاء هيئة التدريس، بالإضافة إلى ميلهم نحو إعدادها لتميزها بالعديد من المميزات منها: أنها سهلة التصحيح، وشاملة لمحتوى المقرر أكثر من باقي أنواع الأسئلة، ونسبة التخمين فيها قليلة، ويمكن أن تشمل مستويات متعددة من مستويات المعرفة لهرم بلوم، مما يدفع أعضاء هيئة التدريس لاستخدامها بشكل متكرر، وذلك يسهم بدوره في تنمية مهاراتهم في تطبيق معايير هذا النوع من الأسئلة. وتتفق هذه النتيجة مع دراسة السفياني (2020) ودراسة الذري (2021).

2.3. معايير جودة أسئلة الإكمال.

جدول (6) التكرارات والنسب المئوية والترتيب لنسبة توافر معايير جودة أسئلة الإكمال.

المعايير	متوافر		غير متوافر		الترتيب	درجة التوافر	عدد الاختبارات
	تكرار	%	تكرار	%			
1. وضع الفراغ في نهاية العبارة أو قرب نهايتها.	2	50	2	50	2	متوسطة	4
2.عدم كتابة جزء من الجملة يوحي بالجزء الناقص.	4	100	-	-	1	عالية جداً	4
3. أن يكون في كل عبارة فراغ أو فراغان كحد أقصى.	1	25	3	75	3	ضعيف	4
4.أن تكون الفراغات متساوية في الطول.	4	100	--	-	1	عالية جداً	4
الإجمالي	11	68.7	5	31.2		متوسطة	

نجد من جدول (6) أن المعايير الضمنية الخاصة بأسئلة الإكمال توافرت بدرجة متوسطة وبنسبة توافر (68.7%) في الأربعة الاختبارات التي حللت، حيث اشترك معياران هما رقم (4،2) في الترتيب الأول وبنسبة توافر (100%) وهما المعياران اللذان ينصان على: عدم كتابة جزء من الجملة يوحي بالجزء الناقص، وأن تكون الفراغات متساوية في الطول، في حين حصل المعيار رقم (1) الخاص بوضع الفراغ في نهاية العبارة على الترتيب الثاني وبدرجة متوسطة وبنسبة توافر (50%)، بينما حصل معيار رقم (3) الخاص بأن يكون في كل عبارة فراغ أو فراغان كحد أقصى على الترتيب الثالث وبدرجة ضعيفة وبنسبة توافر (25%). وتختلف هذه النتيجة مع دراسة السفياني (2020) ودراسة الذري (2021).

2.4. معايير جودة أسئلة المطابقة.

جدول (7) التكرارات والنسب المئوية والترتيب لنسبة توافر معايير جودة أسئلة المطابقة.

م	المعايير	متوافر		غير متوافر		الترتيب	درجة التوافر	عدد الاختبارات
		تكرار	%	تكرار	%			
.1	عبارات إحدى المجموعتين أكثر من الأخرى.	3	75	1	25	2	عالية	4
.2	تجانس عناصر المجموعتين وتوافقهما.	3	75	1	25	2	عالية	
.3	أن تظهر جميع بنود السؤال في الورقة نفسها.	3	75	1	25	2	عالية	
.4	وضوح كيفية الإجابة بدقة، وضع الأرقام، التوصيل بخط.	4	100	-	-	1	عالية جداً	
الإجمالي		13	81.2	3	18.7			

يلاحظ من جدول (7) أن المعايير الضمنية الخاصة بأسئلة المطابقة توافرت في أربعة اختبارات فقط من عينة الدراسة وبدرجة عالية جداً وبنسبة توافر (81.2%)، حيث حصل المعيار رقم (2) الخاص بتجانس عناصر المجموعتين وتوافقهما على درجة عالية جداً وبنسبة توافر (100%)، بينما حصلت المعايير رقم (1، 3، 4) الخاصة بأن تكون عبارات إحدى المجموعتين أكثر من الأخرى، وأن تظهر جميع بنود السؤال في الورقة نفسها، ووضوح كيفية الإجابة بدقة مثل: وضع الأرقام أو التوصيل بخط على الترتيب الثاني وبدرجة متوسطة وبنسبة توافر (75%). ويمكن أن تشير هذه النتيجة إلى معرفة واضعي هذا النوع من الأسئلة بمعايير كتابتها، وقد توافرت في أربعة اختبارات فقط؛ لأن أسئلة المطابقة تقتصر على قياس العلاقة بين المفاهيم. وتتفق هذه النتيجة مع دراسة الذري (2021). ويمكن تلخيص الجداول السابقة في الجدول الآتي:

جدول (8): نسبة ودرجة التوافر لمعايير الأسئلة الموضوعية بشكل إجمالي.

	الفروع	عدد المعايير الفرعية	النسبة المئوية		الترتيب	درجة التوافر
			التوافر	عدم التوافر		
محور الأسئلة الموضوعية	أسئلة الصواب والخطأ	5	86.6%	13.3%	2	عالية جداً
	أسئلة الاختيار من متعدد	10	88.6 %	11.8%	1	عالية جداً
	أسئلة الإكمال	4	68.7%	31.2%	4	متوسطة
	أسئلة المطابقة	4	81.2%	18.7%	3	عالية جداً
الإجمالي		23	81.3%	18.7 %		عالية جداً

نجد من جدول (8) أن معايير الأسئلة الموضوعية ككل توافرت بدرجة عالية جداً وبنسبة مئوية (81.3%) حيث توافرت كلٌ من أسئلة الصواب والخطأ والاختيار من متعدد وأسئلة المطابقة بدرجة عالية جداً وبالترتيب الثاني والأول والثالث على التوالي، في حين توافرت أسئلة الإكمال بدرجة متوسطة وبالترتيب الرابع. وبذلك تمت الإجابة عن الفرع (ب) من السؤال الثاني الذي ينص على: ما مدى جودة أسئلة الاختبارات الموضوعية لمقررات الطب البشري بكلية الطب- جامعة عمران للعام الدراسي 2022-2023م؟

3. **ما نسبة توافر المعايير العامة ومخرجات التعلم للاختبارات التحصيلية لمقررات الطب البشري بكلية الطب- جامعة عمران للعام الدراسي 2022-2023م؟** تظهر نتائج هذا السؤال في جدول (9):

جدول (9): التكرارات والنسب المئوية والترتيب لنسبة توافر المعايير العامة ومخرجات التعلم للاختبارات التحصيلية لمقررات الطب البشري.

م	المعايير	متوافر		غير متوافر		الترتيب	درجة التوافر	عدد الاختبارات
		تكرار	%	تكرار	%			
المعايير العامة	1. التنوع في الفقرات المقالية والموضوعية.	14	53.8	12	46.2	2	متوسطة	26
	2. تناسب عدد الأسئلة مع الزمن الكلي للاختبار.	26	100	-	-	1	عالية جداً	26
	3. شمول الاختبار لمحتوى المقرر الدراسي.	4	22.2	14	77.8	3	ضعيفة	18
الإجمالي		44	58.6	26	41.3		متوسطة	
مخرجات التعلم	1. يغطي الاختبار مخرجات التعلم في مجال المعرفة والفهم.	18	100	-	-	1	عالية جداً	18
	2. يغطي الاختبار مخرجات التعلم للمهارات الذهنية.	5	27.8	13	72.2	2	ضعيفة	
الإجمالي		23	63.8	13	36.1		متوسطة	

يتضح من جدول (9) أن المعايير الضمنية الخاصة بالمعايير العامة توافرت بنسبة إجمالية (58.6%) وبدرجة متوسطة، حيث حصل معيار (2) الذي ينص على: تناسب العديد من الأسئلة مع الزمن الكلي للاختبار على المرتبة الأولى وبنسبة توافر (100%)، في حين حصل معيار (1) الخاص بتنوع الفقرات المقالية والموضوعية على الترتيب الثاني وبنسبة توافر (53.8%) وهي درجة متوسطة، وقد يعزى ذلك إلى طبيعة الأسئلة المقالية والتي تحتاج وقتاً طويلاً في التصحيح، وعدم شمولها جميع مفردات المقرر، بالإضافة إلى الأعداد الكثيرة للطلبة وضعف مهارات الكتابة لديهم، وكل ذلك قد يدفع أعضاء هيئة التدريس إلى تفضيل الأسئلة الموضوعية، حيث نجد أن الأسئلة المقالية غابت تماماً عن أحد عشر اختباراً، في حين توافرت في باقي الاختبارات على استحياء، حيث كانت من سؤال إلى ثلاثة أسئلة في أغلب عينة الدراسة.

بينما جاء معيار شمول الاختبار لمحتوى المقرر في المرتبة الثانية وبنسبة توافر (22.2%) وهي درجة ضعيفة، وقد يعزى ذلك إلى عدم معرفة بعض أعضاء هيئة التدريس لتوصيف المقررات التي يدرسونها، لاسيما الأعضاء الجدد، بالإضافة إلى صعوبة الوضع المالي والاقتصادي لأعضاء هيئة التدريس مما يدفعهم إلى ضغط المحاضرات، وعدم القدرة على إكمال مفردات المقررات. وتتفق هذه النتيجة مع دراسة الصلاحي (2008) وغليون (2009) والسفياني (2020) والذري (2021).

كما يتضح من جدول (9) أن المعايير الضمنية الخاصة بمخرجات التعلم توافرت بنسبة إجمالية (63.8%) وبدرجة متوسطة، حيث توافر معيار رقم (1) الخاص بمخرجات التعلم في مجال المعرفة والفهم بنسبة (100%) وبدرجة عالية جداً، في حين توافر معيار رقم (2) الخاص بمخرجات التعلم في مجال المهارات الذهنية بنسبة (27.8%) وبدرجة ضعيفة، ويمكن أن يعزى ذلك إلى عدم امتلاك المعرفة الكافية لأعضاء هيئة التدريس لمخرجات التعلم في مجال المهارات الذهنية التي تمثل مستويات التفكير العليا في هرم بلوم، حيث تتطلب هذه المهارات قدرات عالية في صياغة فقرات الأسئلة على وفق المهارات الذهنية، كما تتطلب المهارات الذهنية طرق تدريس حديثة وأنشطة تفاعلية ووقتاً طويلاً لتنميتها لدى الطلبة، وكل ذلك يصعب تنفيذه من قبل بعض أعضاء

هيئة التدريس؛ لانشغالهم بأعمال أخرى إلى جانب مهامهم التدريسية؛ بسبب الأوضاع الاقتصادية التي يعانون منها، مما قد يدفع هؤلاء الأعضاء إلى التدريس التقليدي الذي يتمحور حول مستويات التفكير الدنيا المتمثلة في مجال المعرفة و الفهم، بالإضافة إلى سهولة صياغة فقرات الأسئلة على وفق هذا المجال.

وتتفق هذه النتيجة مع دراسة العلي (2013) ودراسة بافطوم والسفياني (2019) ودراسة السفياني (2020) والذري (2021) فيما يخص المعيار الأول في حين تختلف عن الدراستين الأخيرتين فيما يخص المعيار الثاني.

ويمكن تلخيص إجابة السؤال الثاني في الجدول الآتي:

جدول (10) التكرارات والنسب المئوية والترتيب لنسبة توافر المعايير الضمنية للاختبارات التحصيلية لمقررات الطب البشري.

	المعايير	عدد المعايير الفرعية	النسبة المئوية %		الترتيب	درجة التوافر
			للتوافر	عدم التوافر		
محور المعايير الضمنية	الأسئلة المقالية.	4	80%	20%	1	عالية جداً
	الأسئلة الموضوعية.	23	81.3%	18.7 %	1	عالية جداً
	المعايير العامة.	3	58.6%	41.3%	2	عالية
	مخرجات التعلم.	2	63.8%	36.1%	3	متوسطة.
الإجمالي		32	70.9%	29.0%		عالية

يتضح من جدول (10) أن المعايير الضمنية ككل للاختبارات التحصيلية لمقررات الطب البشري بكلية الطب – جامعة عمران توافرت بنسبة مئوية بلغت (70.9%) وبدرجة عالية، في حين تفاوتت نسب توافر المحاور المتعلقة بها، حيث حصل معيارا الأسئلة المقالية والموضوعية على الترتيب الأول وبدرجة توافر عالية جداً وبنسبة مئوية (80%) و (81.3%) على التوالي، كما حصل معيار المعايير العامة بالترتيب الثاني وبدرجة توافر عالية وبنسبة توافر بلغت (58.6%)، بينما جاء معيار مخرجات التعلم في المرتبة الثالثة وبدرجة متوسطة وبنسبة مئوية (63.8%). وتختلف هذه النتيجة عن دراسة السفياني (2020). وبذلك أجابت الباحثة عن السؤال الثاني الذي ينص على: ما نسبة توافر المعايير الضمنية للاختبارات التحصيلية لمقررات الطب البشري كلية الطب – بجامعة عمران للعام الدراسي 2022-2023م؟

إجابة السؤال الثالث الذي ينص على: ما نسبة توافر المعايير الشكلية للورقة الاختبارية لمقررات الطب البشري بكلية الطب – جامعة عمران للعام الدراسي 2022-2023م؟ تظهر إجابة هذا السؤال من خلال الجدول الآتي:

جدول (11): التكرارات والنسب المئوية والترتيب لنسبة توافر محور المواصفات الشكلية في الورقة الاختبارية.

	المعايير	متوافر		غير متوافر		الترتيب	درجة التوافر
		تكرار	%	تكرار	%		
الأكليشة	1. كتابة اسم الجامعة.	26	100	–	–	1	عالية جداً
	2. كتابة اسم الكلية.	24	92.2	2	7.7	3	عالية جداً
	3. توفر شعار الجامعة أو الكلية.	21	80.8	5	19.2	5	عالية جداً
	4. كتابة اسم القسم أو التخصص المقدم له الاختبار.	22	84.6	4	15.4	4	عالية جداً
	5. كتابة اسم المقرر.	24	92.3	2	7.7	3	عالية جداً
	6. بيان المستوى الدراسي.	21	80.8	5	19.2	5	عالية جداً
	7. بيان العام الجامعي والفصل الدراسي.	20	76.9	6	23.1	6	عالية
	8. كتابة يوم الاختبار.	–	–	26	100	16	منعدمة
	9. كتابة تاريخ الاختبار.	4	15.4	22	84.6	14	ضعيفة
	10. كتابة زمن الاختبار وفق اللائحة.	15	57.7	11	42.3	8	متوسطة
	11. بيان نوع الاختبار (نصفي/ نهائي).	19	73.1	7	26.9	7	عالية
	12. كتابة اسم مدرس المقرر	19	73.1	7	26.9	7	عالية
	مجموع النسب والتكرارات	215	68.9	97	31.0		عالية
تعليمات الاختبار	1. تعليمات الاختبار مكتوبة بخط مناسب وواضح.	9	34.6	17	65.4	11	ضعيفة
	2. بيان عدد الأسئلة.	4	15.4	22	84.6	14	ضعيفة
	3.بيان عدد أوراق الاختبار إذا كانت أكثر من ورقة واحدة.	5	19.2	21	80.8	13	ضعيفة
	4. تحديد مكان الإجابة إذا كانت الإجابة في الورقة نفسها.	21	80.8	5	19.2	5	عالية جداً
	5.يوجد خط فاصل بين التعليمات العامة والأسئلة.	4	15.4	22	84.6	14	ضعيفة
	مجموع التكرارات والنسب	43	33.1	87	66.9		ضعيفة
الهوامش الجانبية	1. كتابة أرقام الأسئلة في الهامش الأيسر والعكس للاختبارات باللغة العربية.	25	96.2	1	3.8	2	عالية جداً
	2.كتابة الدرجة الكلية للاختبار.	1	3.8	25	96.2	15	ضعيفة
	3. كتابة درجة كل سؤال في الهامش الأيمن والعكس للاختبارات باللغة العربية.	10	38.5	16	61.5	10	ضعيفة
	مجموع التكرارات والنسب	36	46.2	42	53.8		ضعيفة
الهامش السفلي	1. ترقيم الصفحات إذا كان الاختبار أكثر من صفحة.	13	50	13	50	9	متوسطة
	2. كتابة انتهت الأسئلة أو تمنياتي لكم التوفيق في نهاية الاختبار.	21	80.8	5	19.2	5	عالية جداً
	3. كتابة للأسئلة بقية وفي الصفحة التالية تابع بقية الأسئلة في حال كان الامتحان أكثر من ورقة.	8	30.3	18	69.2	12	ضعيفة
مجموع التكرارات والنسب		42	53.8	36	46.1		متوسطة
الإجمالي الكلي لمحور المعايير الشكلية		المعايير الفرعية	متوافر	غير متوفر	درجة التوافر		
		23	50.5%	49.5 %	متوسطة		

يتضح من جدول (11) أن محور المعايير الشكلية توافر بنسبة إجمالية بلغت (% 50.5) بدرجة متوسطة، في حين تفاوتت المحاور الفرعية لهذا المحور حيث حصل محور الإكليشة على درجة عالية وبنسبة مئوية (68.9%) وحصل على الترتيب الأول، بينما اشتركا محوري تعليمات الاختبار والهوامش الجانبية بدرجة توافر

ضعيفة وبنسبة توافر (% 33.1) و (% 46.2) في الترتيب الرابع والثالث على التوالي. وقد يعزى ذلك الضعف إلى عدم معرفة أعضاء هيئة التدريس بالمعايير الشكلية؛ لعدم حصولهم على توجيهات وإرشادات من قبل الكنترول في الكلية، وعدم تنفيذ دورات تدريبية أو ورش عمل خاصة بمعايير جودة الاختبارات بشكل عام، بالإضافة إلى اهتمام أعضاء هيئة التدريس بنوع أسئلة الاختبار أكثر من المعايير الشكلية؛ اعتقاداً منهم أن طلبة الجامعة يعدون في مستوى متقدم ولا يحتاجون تعليمات خاصة بالاختبار. وتختلف هذه الدراسة عن دراسة الصلاحي (2008) ودراسة السفياني (2020) ودراسة الذري (2021).

إجابة السؤال الرابع الذي ينص على: ما نسبة توافر المعايير الفنية للورقة الاختبارية لمقررات الطب البشري بكلية الطب – جامعة عمران للعام الدراسي 2022-2023م؟ تظهر إجابة هذا السؤال من خلال الجدول الآتي:

جدول (12): التكرارات والنسب المئوية والترتيب لنسبة توافر المعايير الفنية في الورقة الاختبارية.

المعايير الفنية	متوافر		غير متوافر		الترتيب	درجة التوافر	عدد الاختبار
	تكرار	%	تكرار	%			
1. خلو الاختبار من الأخطاء اللغوية والمطبعية.	14	53.8	12	46.2	4	متوسط	26
2. توجد مسافات مناسبة بين الأسئلة.	25	96.2	1	3.8	1	عالية جداً	26
3. توجد مسافات مناسبة بين فقرات السؤال الواحد.	26	100	–	–	1	عالية جداً	26
4. وجود مكان مناسب وكافٍ للإجابة في حال الإجابة في الورقة نفسها.	26	100	–	–	1	عالية جداً	26
5. توضع الأسئلة المتجانسة معاً.	26	100	–	–	1	عالية جداً	26
6. سلامة ترقيم الأسئلة وفروعها دون تداخل.	24	92.3	2	7.7	3	عالية جداً	26
الإجمالي	141	90.3	15	9.61		عالية جداً	

يتضح من جدول (12) أن المعايير الفنية ككل توافرت بنسبة (90.3%) وبدرجة عالية جداً حيث حصلت المعايير رقم (2، 3، 4، 5،6) على الترتيب الأول وبدرجة توافر عالية جداً، وتشير هذه النتائج إلى معرفة أعضاء هيئة التدريس بالمعايير الفنية للورقة الاختبارية، بالإضافة إلى استعانة بعضهم بمختصين في الطباعة والتنسيق الفني للأوراق الاختبارية مما يسهم في تحسين المعايير الفنية للورقة الاختبارية.

وتتفق هذه النتيجة مع دراستي: السفياني (2020) والذري (2021)، وباستثناء معيار رقم (1) الخاص بخلو الاختبار من الأخطاء اللغوية والمطبعية، حيث حصل هذا المعيار على درجة متوسطة وبنسبة (53.8%) وحاز الترتيب الرابع، وقد يعزى ذلك إلى عدم المراجعة اللغوية والمطبعية للاختبارات قبل تطبيقها، بالإضافة إلى التخصصات العلمية لأعضاء هيئة التدريس وتدريس المقررات الطبية باللغة الإنجليزية الذي يتطلب التعمق في قواعد هذه اللغة، وقد يسهم ذلك في بُعد أعضاء هيئة التدريس عن قواعد اللغة العربية. حيث تنوعت الأخطاء ما بين أخطاء لغوية وطباعية، مثل كتابة الهمزة، وتكرار بعض الكلمات دون فائدة، واستعمال حروف الجر، ويمكن استعراض هذه الأخطاء كما يأتي:

1. اختر الإجابة الصحيحة من دخال المربع والأصح من داخل المربع في اختبار رقم (4).

2. تكرار كملة (يقوم) في فقرة: يقوم بتزويد الشبكة بالموارد والخدمات ويقوم بالتحكم فيها في اختبار رقم (4).
3. ماذا تعني المصطلحات التالية بالإنجليزي: (WWW,WAN) والصياغة السليمة ماذا تعني المصطلحات الإنجليزية الآتية؟ في اختبار رقم (5).
5. زيادة (الجزء العملي) في فقرة: أكمل الفراغات فيما يلي بدون خدش أو شطب الجزء العملي في اختبار رقم (4).
6. من التطبيقات برنامج .... والأصح من تطبيقات برنامج في اختبار رقم (4).
7. اجب على جميع الأسئلة والصحيح أجب عن جميع الأسئلة في اختبار رقم (3).
الأخطاء الإملائية في:
8. كلمة الأتية في فقرة: عرّف المفاهيم الأتية تعريفاً علمياً، والصواب الآتية في اختبار رقم (9).
9.عدم كتابة الهمزة في الكلمات: (اهداف، اهم، الامة) في فقرة: من أهداف الثقافة الإسلامية التعرف على اهم التحديات التي تواجه الثقافة الإسلامية، من اهم أدوار الامة في صناعة ثقافتنا، ابعاد الامة عن مصادر ثقافتنا، من أدوار الشباب في بناء الامة في اختبار رقم (3).

الاختبارات المطبوعة باللغة الإنجليزية:

1. صياغة سؤال الصواب والخطأ صيغت باللغة الإنجليزية
(Put (√) or (x) in front of the following statement) و تعني ضع علامة صح أو خطأ مقابل الجمل الآتية، والأصح ضع علامة (√) مقابل العبارة الصحيحة و علامة (×) مقابل العبارة الخاطئة، وتصاغ باللغة الإنجليزية (.Tick (√) for the true statement or (x) for the false statement)، وتكرر ذلك في اختبارين رقم (6) و (15).
2. في بعض البدائل كتبت باللغة الإنجليزية (ALL , NON) و تعني (كل و لاشيء)، والصواب: كل ما ذكر، أو لا شيء مما ذكر ويقابل ذلك في اللغة الإنجليزية (All the above answers , ALL is No thing correct)، وتكرر ذلك الخطأ في ثلاثة اختبارات.
3. صياغة بعض أسئلة الاختيار من متعدد بعبارة عامة، كما في اختبار رقم (15) باللغة الإنجليزية مثل:
(The following are true for measles reservoir) وتعني ما يلي صحيح بالنسبة لخزان الحصبة، والصياغة الأدق (which of the following are true for measles reservoir) ويقابلها باللغة العربية أي مما يلي صحيح بالنسبة لخزان الحصبة؟
4. تكرار جملة (symptomatic improvement) في سؤال (11) في اختبار رقم (22)، بالإضافة إلى وضع نقاط نهاية أكثر من جملة مما يشير إلى وجود العديد من الخيارات في الفقرة الواحدة، وفي سؤال (48،46) وضع النقطة في غير محلها، وزيادة (the) في أحد بدائل سؤال رقم (49).
5. زيادة الحرف (or) في سؤال رقم (35) في اختبار رقم (10).

6. في اختبار رقم (8) في سؤال رقم (2) وجود أقواس حول الفراغات، وفراغات دون أقواس، وزيادة علامات الشرطة، والفاصلة.

7. في اختبار رقم (23): في سؤال رقم (23) في الاختيار a يوجد نقطة بين المصطلحين: المكورات العنقودية الذهبية (staph.aureuse) ، والصواب عدم وجود نقطة بينهما.

توصيات الدراسة:

1. ضرورة مراعاة مواصفات ومعايير بناء الاختبارات الجيدة من قبل أعضاء هيئة التدريس ومعاونيهم.
2. إعطاء إرشادات وتوجيهات من قبل الأقسام والكنترول وبإشراف وحدة ضمان الجودة في الكلية بمراعاة المواصفات الفنية والشكلية للورقة الاختبارية.
3. تنفيذ ورش عمل ودورات تدريبية حول معايير الاختبارات الجيدة.
4. تنفيذ ورش عمل حول تقييم مخرجات التعلم لاسيما في مجال المهارات الذهنية من خلال الاختبارات التحصيلية.
5. ضرورة شمول أسئلة الاختبار محتوى المقرر الدراسي ومخرجات التعلم بمجالاتها المختلفة.

المقترحات:

1. إجراء دراسات مماثلة في اقسام وكليات أخرى.
2. بناء برامج تدريبية قائمة على جودة الاختبارات التحصيلية وتطبيقها.
3. إجراء دراسات تحليلية للاختبارات التحصيلية في ضوء متغيرات أخرى مثل مستويات بلوم المعرفية، ومخرجات التعلم ...إلخ
4. إعداد دليل متخصص بجودة الاختبارات التحصيلية خاص بجامعة عمران وتوزيعه على الكليات والأقسام.

المراجع:

[1] بافطوم، سالم أحمد؛ السفياني، هلال محمد (2019): مدى تمثيل الاختبارات الفصلية لأوزان الوحدات الدراسية وأهداف بلوم المعرفية ومستويات الصعوبة والتمييز بكلية التربية محافظة المهرة دراسة تحليلة، مجلة مركز جزيرة العرب للبحوث التربوية والإنسانية، المجلد (1)، العدد (2).

[2] حنان بثينة، ونعيم بوعوشه (2020): الصدق والثبات في البحوث الاجتماعية، مجلة دراسات في علوم الإنسان والمجتمع- جامعة جيجل، مجلد (3)، العدد (2).

[3] جيوسي، مجدي (2016): مدى تحقيق أسئلة الامتحانات النهائية في جامعة فلسطين التقنية للأهداف التعليمية تبعاً لهرم بلوم في ضوء معايير الورقة الامتحانية الجيدة، المجلة الدولية التربوية المتخصصة، المجلد (5)، العدد (8).

[4] الذري، حسين علي (2021): تقويم اختبارات التحصيل للمقررات التربوية بكلية التربية جامعة ذمار في ضوء معايير جودة الاختبارات، مجلة بحوث ودراسات تربوية، العدد (15).

[5] الرازحي، عبد الوارث عبده سيف (2005): معايير تقويم كفاءة نظام إعداد الاختبارات التحصيلية بجامعة صنعاء، المؤتمر العلمي السابع عشر – مناهج التعليم والمستويات المعيارية، المجلد (1)، الجمعية المصرية للمناهج وطرق التدريس، القاهرة.

[6] السنداني، عادل إبراهيم (2009): التقويم والقياس ومهارات الاتصال في العلوم الاجتماعية، دار الكتاب الحديث، القاهرة، جمهورية مصر العربية.

[7] السفياني، هلال محمد علي (2020): تقييم أوراق امتحانات المقررات التربوية بكلية التربية محافظة المهرة وفقاً لمواصفات إعدادها ومعايير صياغة فقراتها، مجلة جزيرة العرب للبحوث التربوية والإنسانية، المجلد (1)، العدد (5).

[8] ساعد، صباح وعامر، وسيلة (2017): تقييم كفاية بناء الاختبارات التحصيلية لدى أساتذة التعليم الجامعي وفق معايير الاختبار الجيد، مجلة العلوم الإنسانية والاجتماعية، العدد (28).

[9] سليمان، أمين محمد، أبو علام رجاء محمود (2010): القياس والتقويم في العلوم الإنسانية أسسه وأدواته وتطبيقاته، القاهرة، دار الكتب الحديث، ط1.

[10] شحاتة، حسن وآخرون (2003): معجم المصطلحات التربوية والنفسية، الدار المصرية اللبنانية، ط1.

[11] سراوي علي، فرحان فوزي (2022): مدى مطابقة الاختبارات التحصيلية الفصلية لتعليمات دليل بناء الاختبارات – اختبارات علوم الطبيعة والحياة للسنة الرابعة متوسط نموذجاً، مجلة الشامل للعلوم التربوية والاجتماعية، المجلد 5، العدد 2.

[12] سليمان، مريم الحاج وآخرون (بدون): دليل ممارسات نظم تقويم الطلاب والامتحانات، مشروع تطوير نظم تقويم الطلاب والامتحانات – مركز تطوير نظم تقويم الطلاب والامتحانات – جامعة دمنهور.

[13] الصلاحي، عبد السلام (2008): تقويم الاختبارات النهائية في جامعة صنعاء في ضوء معايير الاختبار الجيد، مجلة الباحث العلمي، جامعة إب، العدد (17).

[14] علام، صلاح الدين محمود (2000): القياس والتقويم التربوي النفسي أساسياته وتطبيقاته وتوجهاته المعاصرة، دار الفكر العربي، القاهرة، ط1.

[15] عبد الرحمن، أماني مكاوي (2019): مدى جودة الاختبارات التحصيلية في التعليم الجامعي في ضوء معايير الجودة الشاملة بكلية التربية بعفيف – جامعة شقراء، مجلة إدارة الجودة الشاملة، مجلد (20)، العدد (1).

[16] عثمان، محمد (2010): أساليب التقويم التربوي، دار أسامة للنشر والتوزيع، عمان، الأردن.

[17] عكاشة، محمود فتحي، ضحا، إيمان صلاح (2020): تقييم امتحانات طلاب كلية التربية ومدى تحقيقها لنواتج التعلم المستهدفة، مجلة البحث العلمي في التربية، العدد (21).

[18] العلي، ريم عبد العزيز محمد (2013): تقويم جودة الاختبارات التحصيلية لأعضاء هيئة التدريس بكلية التربية جامعة سلمان بن عبد العزيز، المجلة العلمية بكلية التربية بالود الجديد، الجزء الأول، العدد الثاني عشر.

[19] غليون، أزهار محمد (2009): تقويم الاختبارات النهائية للدراسات العليا بكلية التربية جامعة صنعاء في ضوء معايير الاختبار الجيد، المجلة العلمية لكلية التربية جامعة ذمار، المجلد (1)، العدد (8).

[20] القدسي، أحمد محمد وآخرون (2020): دليل عضو هيئة التدريس في إستراتيجيات التدريس الجامعي، سلسلة إصدارات مركز التطوير الأكاديمي وضمان الجودة للأدلة، مركز التطوير الأكاديمي وضمان الجودة ،جامعة صنعاء: الجمهورية اليمنية.

[21] القدسي، أحمد محمد وآخرون (2020): دليل عضو هيئة التدريس في بناء الاختبارات التحصيلية، سلسلة إصدارات مركز التطوير الأكاديمي وضمان الجودة للأدلة، مركز التطوير الأكاديمي وضمان الجودة ،جامعة صنعاء: الجمهورية اليمنية.

[22] القدسي، أحمد محمد وآخرون (2021): دليل مخرجات التعلم للبرامج والمقررات الدراسية، سلسلة إصدارات مركز التطوير الأكاديمي وضمان الجودة للأدلة، جامعة صنعاء، الجمهورية اليمنية.

[23] كنيدي، ديكلان (2013): صياغة مخرجات التعلم واستخدامها: دليل تطبيقي، ترجمة سعيد بن محمد الزهراني، عبد الحميد بن محمد، وزارة التعليم العالي، مركز البحوث والدراسات، المملكة العربية السعودية.

[24] مجيد، سوسن شاكر (2014): أسس بناء الاختبارات والمقاييس النفسية والتربوية، مركز ديبونو لتعليم التفكير، عمان، الأردن، ط3.

[25] مجلس الاعتماد الأكاديمي وضمان جودة التعليم العالي (2021): معايير الاعتماد الأكاديمي البرامجي، الجمهورية اليمنية، وزارة التعليم العالي والبحث العلمي.

[26] النور، عبد الغني مطهر وآخرون (2021): مقياس تقييم أسئلة الاختبارات، سلسلة إصدارات مركز التطوير الأكاديمي وضمان الجودة للمقاييس، جامعة صنعاء.

[27] وكالة الجامعة للشؤون التعليمية (2024): دليل وتقويم نواتج التعلم، الإصدار الثاني، جامعة المجمعة، المملكة العربية السعودية.

المراجع الأجنبية:

1-Abu Foudeh, Basel, K (2020): A Training Program Based on the Quality standards of Achievment Tests to Improve the competence of teaching staffing constricting Tests at Middle East University in Jordan Universal Journal of Educational Research 8(4): 1212-1217, http: // www.hrpub.org

2-Al Khawaldeh, S, & Alawneh, Y, Al, & Alzboun, M (2022): The Availability of quality standards for the construction of science achievement tests from of view of the examination committees Journal of Hunan university , Nutural Seinces , vol 49, No 9

3-Kennedy, D, (2007): Writing and using learning outcomes (A partial guide), university college cork.

4-Uluman, M (2023), A Review of Achievement test development in turkiye Regarding the Achievement test development process, journal of Human and social science (JOHASS), 6(2), p377-405.