Scientific Journal
University of Saba Region
ISSN :2709- 2747 (Online)

المجلة العلمية
لجامعة إقليم سبأ
ISSN :2709- 2739 (Print)

تقويم اختبارات القبول لمادة الأحياء للمتقدمين لكلية الطب، في جامعة إقليم سبأ، في ضوء معايير الاختبار الجيد

Evaluation of Admission Tests in Biology for the Applicants to the Faculty of Medicine at University of Saba Region in Light of Good Test Standards

ماجدة علي سعيد الحميدي [1]
Majdah Ali Saeed Al-Humaidi
عبدالله حسن عبد الرب [2]
Abdullah Hassan Abdulrab

المجلد(8) العدد(2) ديسمبر 2025م
https://doi.org/10.54582/TSJ.2.2.126

(1) باحثة
عنوان المراسلة : maj1993al@gmail.com

(2) أستاذ مناهج وطرق تدريس العلوم – كلية التربية جامعة إب
عنوان المراسلة : dr.abdullahyem2013@gmail.com

الملخص:

هدفت الدراسة إلى تقويم اختبارات قبول مادة الأحياء للمتقدمين لكلية الطب في جامعة إقليم سبأ في ضوء معايير الاختبار الجيد، وتم استخدام أسلوب تحليل المحتوى كأحد أساليب المنهج الوصفي. تكونت عينة الدراسة من (125) فقرة، تندرج ضمن خمسة نماذج من اختبارات القبول للأعوام (2021–2024)، وصُممت أداة التحليل الخاصة بمعايير الاختبار الجيد، حيث تكونت من (33) مؤشرًا فرعيًا، موزعة على أربعة معايير رئيسة، هي: مواصفات ورقة الاختبار الجيد، ووضوح التعليمات، الجوانب الفنية للاختبار الجيد، أسئلة الاختيار من متعدد، والمستوى المعرفي للأسئلة. واستخدمت التكرارات والنسب المئوية لتحليل بيانات الدراسة. أظهرت نتائج الدراسة أن اختبارات القبول التزمت ببعض المعايير الأساسية، مثل: الترقيم، وتضمين البيانات الأساسية بدقة، وسلامة اللغة، وجودة الطباعة، بالإضافة إلى خلوها من الأخطاء العلمية. ومع ذلك، كشفت الدراسة عن وجود قصور في جوانب أخرى، أهمها: تحديد الدرجات، والتعليمات الإرشادية، والتدقيق الإملائي، وتوحيد الخط. فيما يخص أسئلة الاختيار من متعدد، تبين التزامها بالمعايير بشكلٍ جيد، مع وجود ملاحظات تتعلق بإبراز النفي، وتوزيع البدائل الصحيحة. أما فيما يتعلق بالمستوى المعرفي للأسئلة، فقد ركزت الأسئلة بشكلٍ أساسي على مستويات التفكير الدنيا (التذكر والفهم)، وأغفلت المستويات العليا للتفكير.

الكلمات المفتاحية: التقويم، اختبارات القبول، معايير الاختبار الجيد.

Abstract:

The study is aimed to evaluate the biology admission tests for the applicants to the Faculty of Medicine at University of Saba Region in light of the standards of a good test. The content analysis method, which is one of the descriptive approach techniques, was used. The study sample consisted of (125) items distributed on five admission test models from the years 2021 to 2024. A content analysis tool of (33) sub-indicators was designed based on the standards of a good test, and distributed on four domains: first, specifications of the test paper and clarity of instructions; second, technical aspects of a good test; third, multiple-choice questions; fourth, cognitive level of the questions. Frequencies and percentages were used to analyze the data. The study's findings showed that the admission tests adhered to certain basic standards, such as numbering, accurate inclusion of essential data, linguistic accuracy, and print quality, in addition to being free of academic errors. However, the study revealed shortcomings in other areas, most notably: grading, guidance instructions, spelling, and font consistency. Regarding multiple-choice questions, they generally met the required standards, though some issues were noted concerning the highlighting of negation and the distribution of correct alternatives. As for the cognitive level of the items, the questions primarily focused on lower-order thinking skills (recall and understanding) while neglecting higher-order thinking levels.

Keywords: Evaluation, Admission Tests, Standards of a Good Test

Copyright: © 2025 Majdah Ali Saeed Al-Humaidi, and Abdullah Hassan Abdulrab. This article is an open-access article distributed under the terms and conditions of Creative Commons Attribution (CC BY 4.0) license.

المقدمة:

أسهم التقدم العلمي في تطوير العملية التعليمية؛ حيث أدى إلى ابتكار أساليب تعليمية حديثة تُحسِّن من جودة التعلم، وتعد العملية التعليمية التعلمية ضرورية لضمان تكامل المعرفة بين المعلم والمتعلم؛ حيث لم يعد التعليم يقتصر على نقل المعلومات، وإنما يعتمد على بناء التفكير النقدي والابتكار.

ونتيجة للتغييرات العلمية والتكنولوجية التي طرأت على جميع مجالات الحياة، دعت جميع دول العالم لتغيير منهاج التقويم لديها؛ وذلك لإصلاح نظام التعليم والاستفادة من التطورات لخلق منظومة تعليمية حديثة (آغا وآخرون، 2023).

والتقويم أحد العناصر المهمة في العملية التعليمية والتعلمية؛ حيث يؤدي دورًا محوريًا في تطويرها، فالعملية التربوية بحاجة إلى تقويم مستمر في جميع مكوناتها، بدءًا من الأهداف وانتهاءً بالتقويم؛ إذ تتصف عملية التقويم بالشمول والتنظيم، وتشمل كل العناصر المرتبطة بالعملية التعليمية؛ حيث تسهم في جمع معلوماتها وبياناتها؛ ولذلك تتسع مجالات التقويم لتشمل كل جوانب العملية التربوية (الجلبي، 2005).

وللتقويم أدوات عدة، أهمها الاختبارات؛ حيث تحتل موقع الصدارة بين وسائل التقويم المستخدمة –سواء في مراحل التعليم المختلفة أو في المرحلة الجامعية–التي قد تكون الوسيلة الوحيدة أحيانًا، وهذا يعني أن الأحكام التي تصدر في الجامعة تتم في ضوء تلك الاختبارات بالدرجة الأولى (أبو جامع والأشقر، 2021؛ الصلاحي، 2008).

ويعد التقويم لاختبارات القبول في الجامعات عملية أساسية في تطويرها؛ حيث يهدف إلى تحليل جودة وفعالية هذه الاختبارات في تحقيق أهدافها، كما أنها تقيس مدى موثوقية الاختبارات وعدالتها في تقييم الطلبة، وتساعد في تحديد مدى ارتباط نتائجها بأداء الطلبة الفعلي في الجامعة.

وأشار القواس (2017) إلى أنه لابد من تحسين جودة التعليم الجامعي، والاهتمام بأحد مدخلاته (الطالب) الذي يعد من مخرجات التعليم بعد فترة زمنية؛ وذلك لما للتعليم الجامعي من أهمية في إعداد الأفراد، لتلبية حاجات المجتمع ومواجهة مشكلاته، فاختيار نوعية الطالب الجامعي الجيد، يساعد في الحصول على مخرجاتٍ جيدة، ومن أهم الطرق التي تتبعها مؤسسات التعليم الجامعي لاختيار مدخلاتها الطلابية هي اختبارات القبول التي تقدم للطلبة المتقدمين، لاسيما المتقدمين لكلية الطب.

وتُعد اختبارات القبول في كليات الطب أداةً حاسمةً في قياس استعداد المتقدمين وقدراتهم الأكاديمية والمعرفية؛ حيث تُؤثر بشكلٍ مباشرٍ في جودة مخرجات العملية التعليمية. وتحتل اختبارات مادة الأحياء أهمية خاصة؛ نظرًا لدورها المحوري في إعداد الطلبة للبرامج الطبية، إذ تُقيِّم مدى إلمامهم بالمفاهيم البيولوجية الأساسية التي تُعد حجر الأساس للدراسات الطبية.

ويؤكد بركات (2020) والشيخ وآخرون (2020) أن الاختبار الجيد هو الذي يتمتع بالشمولية والصدق والثبات، وخلوه من الأخطاء العلمية والنحوية، ويقيس بدقة ما يُفترض قياسه، وفق معايير

واضحة؛ لذا من الضروري الالتزام بالشروط الأساسية في إعداد الأسئلة لضمان جودتها؛ نظرًا لتكرار الشكاوى حول مشكلات الاختبارات، سواء في الإعداد أو المحتوى، أو ضعف التمييز، وهذا يؤثر في دقتها، ويقلل من موثوقية الاختبارات وفاعليتها في التقييم الصحيح.

وقد أجريت العديد من الدراسات لتحليل الاختبارات في مراحل التعليم المختلفة، منها دراسة كل من: أبو جامع والأشقر (2021) الرحيمي (2023) بركات (2020) القواس (2017)، التي أوصت جميعها بإجراء تقويم دوري للاختبارات وتحليلها، وفق معايير الاختبار الجيد للتعرف على أوجه القصور فيها، ووضع التدابير العلاجية بشأنها؛ وذلك لاختيار أفضل الطلبة من المتقدمين.

وعلى المستوى المحلي سعت الجامعات اليمنية إلى سن العديد من الشروط لقبول الطلبة في كلية الطب، منها اجتياز اختبار مواد الأحياء والكيمياء واللغة الإنجليزية، مستخدمة في ذلك التقويم المعياري المرجع.

ونظرًا لأهمية إعداد الاختبارات، بوصفها أكثر وسائل التقويم شيوعًا؛ لذا فهناك حاجة إلى معايير جيدة لاتخاذ القرارات حول جودة أسئلة اختبارات المواد بشكلٍ عام ومادة الأحياء بشكلٍ خاص، وفق معايير الاختبار الجيد.

مشكلة الدراسة:

تحرص الجامعات اليمنية في ظل ارتفاع أعداد خريجي المرحلة الثانوية، وتزايد الإقبال على التعليم الجامعي إلى قبول الطلبة للدراسة فيها –لا سيما في الكليات الطبية–بعد اجتيازهم لاختبارات القبول؛ وذلك لضمان جودة التعليم الأكاديمي، آخذةً في الاعتبارٍ ضعف كفاءة خريجي التعليم الثانوي؛ لذا غدت عمليتا تقويم اختبارات القبول وتطويرها، ضرورةً ملحة، تفرضها التحديات التي تواجه التعليم الجامعي.

لذا لا بد أن تتصف هذه الاختبارات بالمعايير العلمية، وإلا أصبحت أداة مضللة، تقدم بيانات غير دقيقة، قد تقود إلى نتائج غير موثوقة، ولكي تؤدي الاختبارات دورها ينبغي تقويمها بشكل مستمر (الصلاحي، 2008).

وبالنظر إلى واقع اختبارات القبول في الجامعات اليمنية، لُوحظ أنها تعاني من مشكلات تتعلق بجودتها ومدى تحقيقها للأهداف المرجوة؛ ما قد يؤثر في دقة نتائجها وعدالتها في قياس قدرات الطلبة، وهذا ما أكدته دراسة القواس (2017) التي رأت أن اختبارات القبول لا تراعي معايير الاختبار الجيد.

علاوة على ما سبق أُجريت مقابلة مع مجموعة من الطلبة المتقدمين للقبول في كلية الطب، بجامعة إقليم سبأ، حول جودة اختبارات مادة الأحياء، وتوصلت نتيجة المقابلة إلى أن مشكلات الاختبارات تتركز في أن زمن الإجابة لا يتناسب مع بعض الطلبة، وغموض بعض فقرات الاختبار، كما أن هناك فقرات تحتمل إجابتين.

مما سبق؛ أضحت اختبارات القبول معيارًا أساسيًا لاختيار الطلبة المؤهلين خصوصًا في التخصصات

الطبية التي تتطلب مهارات ومعارف متقدمة؛ لذا وجد الباحثان الحاجة الماسة لتقويم اختبارات القبول لمادة الأحياء للمتقدمين لكلية الطب في جامعة إقليم سبأ في ضوء معايير الاختبار الجيد.

وانطلاقًا مما سبق تتمحور مشكلة الدراسة حول السؤال الرئيس الآتي:

ما درجة توافر معايير الاختبار الجيد في اختبارات قبول مادة الأحياء للمتقدمين لكلية الطب في جامعة إقليم سبأ للأعوام (2020-2024)؟

ويتفرع من هذا السؤال الأسئلة الفرعية الآتية:

1. ما معايير الاختبار الجيد التي ينبغي أن تتوافر في اختبار قبول مادة الأحياء للمتقدمين لكلية الطب بالجامعات اليمنية؟
2. ما مستوى توافر معيار مواصفات الورقة الاختبارية ووضوح التعليمات في اختبارات القبول لمادة الأحياء للمتقدمين لكلية الطب، في جامعة إقليم سبأ للأعوام (2020-2024)؟
3. ما مستوى توافر المعايير الفنية للاختبار الجيد في أسئلة اختبار قبول مادة الأحياء للمتقدمين لكلية الطب في جامعة إقليم سبأ للأعوام (2020-2024)؟
4. ما مستوى توافر معايير الاختيار من متعدد في أسئلة اختبار قبول مادة الأحياء للمتقدمين لكلية الطب في جامعة إقليم سبأ للأعوام (2020-2024)؟
5. ما مستوى مراعاة اختبارات القبول لمادة الأحياء في كلية الطب بجامعة إقليم سبأ للأعوام (2020-2024) لمعيار المستوى المعرفي للأسئلة؟

أهداف الدراسة:

تهدف الدراسة الحالية إلى:

1. تحديد معايير الاختبار الجيد التي ينبغي أن تتوافر في اختبار قبول مادة الأحياء للمتقدمين لكلية الطب بالجامعات اليمنية.
2. التعرف على مستوى توافر معيار مواصفات الورقة الاختبارية ووضوح التعليمات في اختبارات القبول لمادة الأحياء للمتقدمين لكلية الطب في جامعة إقليم سبأ للأعوام (2020-2024).
3. الكشف عن مستوى توافر المعايير الفنية للاختبار الجيد في أسئلة اختبار قبول مادة الأحياء للمتقدمين لكلية الطب في جامعة إقليم سبأ للأعوام (2020-2024).
4. التعرف على مستوى توافر معايير الاختيار من متعدد في أسئلة اختبار قبول مادة الأحياء للمتقدمين لكلية الطب في جامعة إقليم سبأ للأعوام (2020-2024).
5. تحديد مستوى مراعاة اختبارات القبول لمادة الأحياء للمستوى المعرفي في كلية الطب بجامعة إقليم سبأ للأعوام (2020-2024).

أهمية الدراسة:

تكمن أهمية الدراسة في الآتي:

1. قد تفيد القائمين على بناء اختبارات القبول بالجامعة في تعريفهم بمعايير جودة اختبارات القبول لمادة الأحياء للمتقدمين لكلية الطب.
2. قد تفيد مُعدي اختبارات القبول من أعضاء هيئة التدريس في الجامعة في تحسين اختباراتهم وتطويرها.
3. قد تلفت نظر المسؤولين في الجامعة بعمل دورات تدريبية لأعضاء هيئة التدريس المكلفين بإعداد اختبارات القبول حول مهارات بناء الاختبارات.
4. قد تحفز المسؤولين في الجامعة على تقويم اختبارات القبول في كليات أخرى بجامعة إقليم سبأ.

حدود الدراسة:

تحددت الدراسة بتقويم اختبارات القبول لمادة الأحياء للمتقدمين لكلية الطب جامعة إقليم سبأ للأعوام (2020-2024) وذلك خلال الفصل الدراسي الثاني للعام الدراسي (2024/2025).

مصطلحات الدراسة:

التقويم:

يعرف التقويم بأنه: «عملية إصدار حكم على مدى تحقيق الأهداف المرجوة» (اليعقوبي، 2013، 34)

ويُعرف التقويم إجرائيًا بأنه: إصدار حكم حول جودة اختبارات قبول مادة الأحياء للمتقدمين بكلية الطب في جامعة إقليم سبأ في ضوء معايير الاختبار الجيد.

معايير الاختبار الجيد:

تُعرف معايير الاختبار الجيد بأنها: «عدد من الشروط التي يجب توافرها في الاختبار التحصيلي ليكون اختبارًا موضوعيًا جيدًا يؤدي الغرض الذي وضع من أجله على الوجه الأكمل» (داؤود، 2014، 364)

وتعرف معايير الاختبار الجيد في هذه الدراسة إجرائيًا بأنها: مجموعة من المحكات التي تم الاستفادة منها من الأدبيات السابقة، واستخدامها في تحليل أسئلة اختبارات القبول في كلية الطب لمادة الأحياء في جامعة إقليم سبأ.

اختبارات القبول:

يعرف الاختبار بأنه: «مجموعة من المثيرات (أسئلة شفوية أو كتابية أو صور أو رسوم) أعدت لتقيس بطريقة كمية أو كيفية سلوكًا ما، والاختبار يعطي درجة ما أو قيمة ما أو رتبة ما للمفحوص، ويمكن أن

يكون الاختبار مجموعة من الأسئلة أو جهازًا معينًا» (عبيدات وآخرون، 2015، 158).

وتُعرف اختبارات القبول إجرائيًا بأنها: مجموعة من الأسئلة التي تغطي المفاهيم الموجودة في كتب الأحياء للمرحلة الثانوية للقسم العلمي، وتقدم للطلبة المتقدمين للالتحاق بكلية الطب في جامعة إقليم سبأ للأعوام (2020-2024).

الإطار النظري:

تمهيد:

تعد الاختبارات أداة أساسية في تقويم العملية التعليمية، وقياس نواتج التعلم لدى الطلبة، وتتطلب فاعلية الأداة توافر خصائص سيكومترية دقيقة، تضمن صدقها وثباتها وموضوعيتها. وتم التركيز هنا على الاختبارات وأهميتها، ومعايير الاختبار الجيد، ونمط أسئلة الاختيار من متعدد.

الاختبارات:

للاختبارات دور كبير في قياس نواتج المعرفة لدى الطلبة؛ حيث إن نتائج الاختبارات تقدم معلومات واقعية وشبه واقعية يعتمد عليها التعليم في اتخاذ القرارات ورسم الخطط والسياسات التعليمية المناسبة للطلبة (الناقة، 2016).

أهمية الاختبارات:

هناك العديد من الفوائد التي تعود على الطلبة عند توجيه الأسئلة، لخصها عبد الهادي (2002) في الآتي:

- وضع الخطط العلاجية للطلبة الضعاف.
- تعطي فكرة واضحة عن قدرات وإمكانيات الطلبة ومستوى نشاطهم.
- جذب انتباه الطلبة نحو المادة التعليمية وزيادة تفاعل الطلبة في المواقف التعليمية.
- تلعب دورًا رئيسيًا في توفر مقاييس جيدة لكثير من مخرجات التعلم.
- من خلال نتائج الاختبارات يتم التعديل في مستوى الأساليب وتقنيات التدريس.
- تساعد الأسئلة في تقييم مستوى الطلبة.

معايير الاختبار الجيد:

الاختبار الجيد يجب أن يتميز بعدة صفات أساسية؛ لضمان دقته وفعاليته في قياس ما وضع لأجله، وقد أشار كل من (اسعادي وعزي، 2015؛ الصلاحي، 2008؛ عبد الرحمن،2011) أن معايير

الاختبار الجيد تتمثل في الآتي:

1. الصدق: يعتبر الاختبار صادقًا إذا كان يقيس ما وضع لأجله، أي أن تكون الأسئلة مرتبطة بالمحتوى والأهداف التعليمية.

2. الثبات: وهو أن يعطي الاختبار نتائج متسقة عند تطبيقه في أوقات مختلفة أو مع مجموعات مختلفة من الأفراد تحت نفس الظروف.

3. الموضوعية: ويقصد بها أن يلتزم واضعها ومصححها بالحياد عند تصميمها وفي أثناء تصحيحها.

4. الشمولية: بمعنى أن الاختبار يغطي جميع جوانب المحتوى التعليمي وأهدافه دون التركيز على جانب وإهمال الجوانب الأخرى.

5. التمييز: هو مراعاة الفروق الفردية بين الطلبة.

6. وضوح صياغة الاختبار: وضوح عبارات الاختبار لتكون مناسبة لحصيلة الطالب اللغوية، وأن يعكس السؤال تفسيرًا واحدًا فقط.

7. تجنب المطابقة مع أسئلة الكتاب المدرسي.

8. النواحي الفنية: جودة تنظيم الورقة الاختبارية وإخراجها ووضوح طباعتها، وخلوها من الأخطاء اللغوية أو الإملائية، وتحديد البيانات الأساسية ووضوح التعليمات عليها، وكذلك تحديد الدرجة المخصصة للاختبار ككل ودرجة كل سؤال على حدة.

ويضيف الباحثان ألا تقتصر أسئلة الاختبار على قياس المستويات المعرفية الدنيا، وإنما ينبغي أن تقيس أسئلة الاختبار المستويات المعرفية العليا، مثل: التحليل، والتركيب، والتقويم.

الخصائص السيكومترية للاختبار الجيد:

يهتم الباحثون في مجالات العلوم النفسية والتربوية والاجتماعية بالحصول على معلومات وبيانات دقيقة حول الظواهر السلوكية، وفي سبيل الحصول على تلك البيانات أو المعلومات يستخدم عددا من الأدوات السيكولوجية منها: الاختبارات والمقاييس وقوائم الملاحظة وقوائم التقدير الذاتي ومقاييس التقدير، وعندما نستخدم اختبارًا من أجل الحصول على معلومات تساعدنا في اتخاذ قرار ما، فإننا نواجه مشكلة أساسية تتعلق باختيارنا للاختبار الذي يمكن أن يفيد في اتخاذ القرار، وقد يكون هناك عدة بدائل مطروحة يمكن استخدامها؛ لذلك فإن هناك أمورًا وأسساً تؤخذ في الاعتبار عند تقويم جودة الاختبار هي:

- الشروط العامة للاختبار منها: الشمول، والموضوعية وغيرها.

- الشروط التجريبية للاختبار وهي: الثبات والصدق وغيرها.

والشروط التجريبية هي التي يطلق عليها الخصائص السيكومترية للاختبار (الصدق، الثبات) كما جاءت عند العديد من الباحثين الذين كتبوا في هذا المجال مثل (النبهان، 2013) و (مجيد، 2014).

ويعرف مراد وسليمان (2005) الخصائص السيكومترية بأنها: «الخصائص الضرورية والمتعلقة بالصدق والثبات والمعايير والتي يتم حسابها بعد تجريب الاختبار على عينة ممثلة للمجتمع» (ص. 350).

الصدق:

هو مدى قياس الاختبار للاستعداد أو الخاصية أو السمة التي وضع لأجلها، وهو صفة أساسية ينبغي توافرها في الاختبار، مالم فقد الاختبار قيمته كوسيلة لقياس الاستعداد أو الخاصية المقصود به قياسها (الشربيني وآخرون، 2010).

وأشار كلٌ من (أبو علام، 2011؛ اسعادي وعزي، 2015؛ البناء، 2025؛ مجيد،2014؛ Pal, 2020; Sivakumar, 2019) إلى أن للصدق عدة أنواع منها:

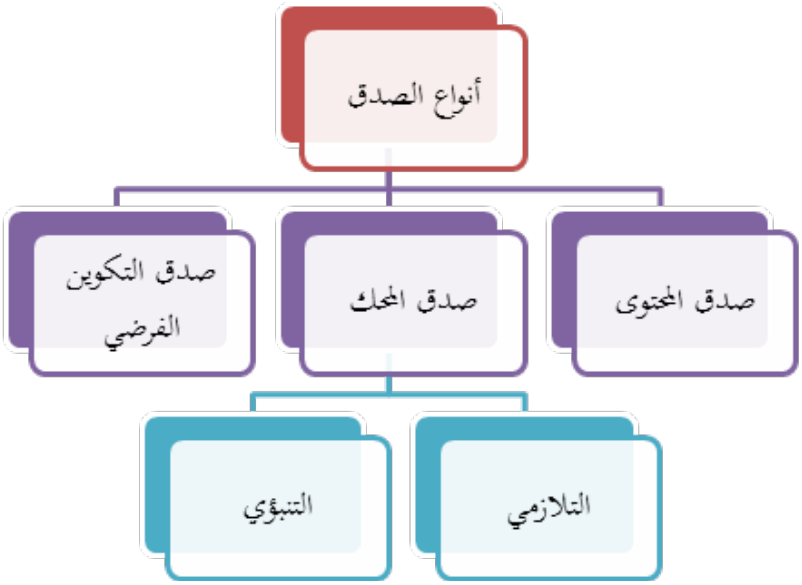
1. **صدق المحتوى:** هو مدى تمثيل الاختبار التحصيلي تمثيلًا صادقًا لمختلف أهداف وأجزاء المادة التعليمية.
2. **صدق التكوين الفرضي (المفهوم):** يشير إلى مدى قياس الاختبار لمفهوم أو خاصية معينة، فالاختبار الذي صمم لقياس المفاهيم الصحية ينبغي أن تكون فقراته متجانسة وتقيس المفاهيم الصحية، ولا تقيس شيئاً آخر، ومن أهم أساليبه ما يأتي:

أ. الصدق التمييزي (مقارنة المجموعتين العليا والسفلى): يكون من خلال تقدير صدق كل فقرة وكل مجال وأيضا قياس الصدق الكلي للاختبار؛ وذلك باستخدام الاختبار التائي لعينتين مستقلتين للمقارنة بين المجموعتين فإذا كانت قيمة (ت) دالة إحصائيًا فهذا يعني أن الاختبار يتمتع بصدق جيد.

ب. صدق الاتساق الداخلي: يعطي صورة عن مدى التناسق الموجود بين الفقرات الموجودة في الاختبار، ومدى اتساق هذه الفقرات مع الاختبار، ويتم التأكد من توافر صدق الاتساق الداخلي باستخدام معامل ارتباط (بيرسون) عن طريق حساب معامل الارتباط بين درجة كل فقرة مع الدرجة الكلية للاختبار.

3. **الصدق المحكي:** يستخدم من خلال تطبيق الاختبار الذي تم إعداده واختبار آخر سبق التحقق من صدقه على نفس العينة، ومن ثم إيجاد معامل الارتباط بين الاختبارين، وإذا طبق الاختباران معًا سمي بالصدق المحكي التلازمي، وإذا طبق الاختبار المحك بعد فترة زمنية من تطبيق الاختبار سمي بالصدق المحكي التنبؤي.

والشكل (1) يوضح أساليب تقدير صدق الاختبار.



شكل (1) أساليب تقدير صدق الاختبار (البناء،2024)

الثبـــات:

يقصد بالثبات دقة المقياس أو اتساقه، وعرف أبو علام (2011) الثبات بأنه: «حصول نفس الفرد على نفس الدرجة أو درجة قريبة منها في نفس الاختبار عند تطبيقه أكثر من مرة» (ص. 463).

طرق حساب معامل الثبات:

تشير أدبيات القياس والتقويم إلى عدة طرق لتقدير معامل الثبات منها: تطبيق صورتين متكافئتين للاختبار؛ لتقدير مدى الاتساق، تطبيق نفس الاختبار مرتين بفاصل زمني؛ لتقدير استقرار الدرجات، الجمع بين الطريقتين السابقتين؛ لتقدير كل من الاستقرار والتكافؤ، تجزئة الاختبار إلى قسمين متكافئين لتقدير الاتساق الداخلي، وتحليل البنية الداخلية للاختبار لمعرفة تجانس الفقرات (أبو علام، 2011). وفيما يأتي توضيح لطرق قياس ثبات الاختبار (البناء، 2025؛ Rossiter, 2011):

1. معامل التكافؤ: يتطلب هذا النوع إعداد نموذجين من الاختبار متكافئين بالمحتوى، وفقرات الاختبار، وطريقة الإجابة والتصحيح، والزمن المخصص للإجابة، وصعوبة فقرات الاختبار، ومن ثم تطبيقهما على أفراد العينة.

2. طريقة إعادة الاختبار (الاستقرار): يستخدم هذا النوع من الثبات عند تطبيق الاختبار على عينة، ومن إعادة تطبيق الاختبار نفسه وفي ظروف مماثلة؛ بعد أسبوعين إلى ثلاثة أسابيع على العينة نفسها، ومن ثم حساب الارتباط بين التطبيقين.

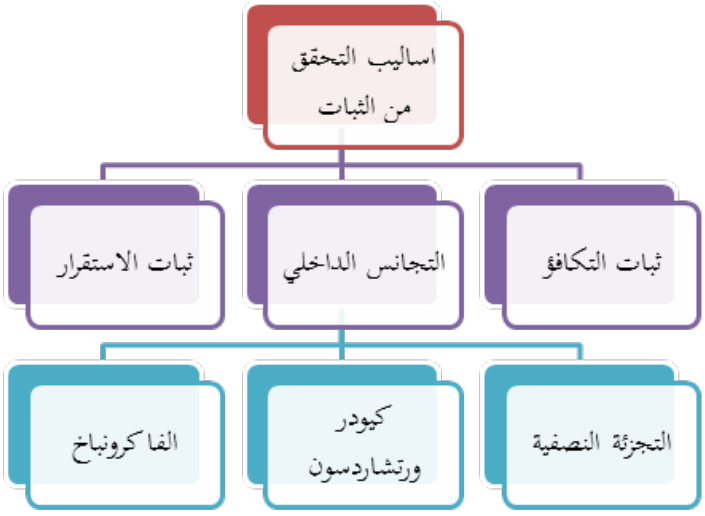
3. التجانس الداخلي: له ثلاثة أساليب هي ما يأتي:

أ. **معادلة كيودر ريتشاردسون**: تستخدم في الاختبارات الموضوعية التي تكون درجات فقراتها (0) أو (1). وتعتمد على حساب الارتباطات بين درجات مجموعة الثبات على جميع الفقرات الداخلة بالاختبار.

ب. **طريقة التجزئة النصفية**: يستخدم من خلال تجزئة الاختبار إلى قسمين (فردي، زوجي) أو (علوي، سفلي) ومن ثم إيجاد معامل الارتباط بين القسمين، ومع ذلك لا يفضل استخدامه في الاختبارات لتفاوت صعوبة فقرات الاختبار.

ج. **معامل ألفا كرونباخ**: يستخدم (ألفا كرونباخ) في قياس التجانس الداخلي في أي أداة تعتمد على أسلوب (ليكرت) ويمكن استخدامه في الاختبارات المقالية دون الموضوعية.

والشكل (2) يوضح أساليب تقدير الثبات.



شكل (2) أساليب التحقق من ثبات الاختبار (البناء، 2024)

أسئلة الاختيار من متعدد:

أشار (كوسدار وآخرون Kocdar et al., 2016) إلى أن فقرات الاختيار من متعدد تتكون من جزأين: أصل أو مقدمة الفقرة وإبدال متعددة تقدم إجابات ممكنة للمشكلة التي عرضت في الأصل، وأن الفقرة الجيدة يجب أن ترتبط بتحقيق هدف معين يرمي الاختبار إلى قياسه، وتستخدم فقرات الاختيار من متعدد لقياس أنواع المعارف من مصطلحات وحقائق ومفاهيم وطرائق، كما أنها تصلح لقياس الفهم وعمليات التعلم المركبة فهي قادرة على قياس عمليات عقلية من جميع المستويات.

ويمكن أن ننوه هنا إلى أن الاختبارات العالمية واختبارات الجامعات التي تقدم للطلبة تركز على اختبار الاختيار من متعدد عند تبني الاختبارات الموضوعية؛ وذلك لما تمتاز به من مزايا تميزها عن بقية أنواع الاختبارات الموضوعية.

مزايا فقرات الاختيار من متعدد:

تمتاز اختبارات الاختيار من متعدد بعدد من الصفات أهمها: تدني نسبة التخمين مقارنة بأسئلة الصواب والخطأ، إلى أنها صادقة وثابته بدرجة أكبر من بقية أنواع الاختبارات الموضوعية، ويمكن استخدامها لقياس المستويات العقلية العليا بجانب الدنيا، وتساعد في تشخيص أخطاء الطلبة أو سوء الفهم لديهم من خلال استجاباتهم للإبدال الخاطئة، كما أنها تعود الطلبة على الحكم الصائب والموازنة وتمييز الأفضل، فضلًا عن إمكانية تحليل نتائجها إحصائيًا بسهولة، ومن السهل تصحيح هذا النوع من الاختبارات لا سيما إذا استعمل المصحح مفتاح الإجابة المثقب (البناء، 2024؛ مجيد، 2014).

عيوبها:

أشار النبهان (2013) أنه على الرغم من مميزات اختبار الاختيار من متعدد، فإن لها سلبيات، ومن أهم تلك السلبيات: صعوبة إعدادها فهي تحتاج إلى مهارة بجانب الوقت أكثر من أي اختبار آخر.

الدراسات السابقة:

تم البحث عن الدراسات السابقة ذات العلاقة بموضوع الدراسة الحالي، ومن أهمها ما يأتي:

دراسة أبو جامع والأشقر (2021) التي هدفت إلى تقويم أسئلة اختبارات الأحياء للصف الثاني عشر في فلسطين في ضوء عمليات العلم من عام (2018 إلى 2020). وقد اتبعت الباحثتان أسلوب تحليل المحتوى بوصفه أحد أساليب المنهج الوصفي، وشملت عينة الدراسة اختبارات الأحياء للصف الثاني عشر للأعوام (2018–2020). وصممت أداة التحليل الخاصة بعمليات العلم، واستخدمت التكرارات والنسب المئوية لتحليل البيانات، وتوصلت الدراسة إلى أن مهارة الاستدلال حصلت على أعلى نسبة من المهارات الأساسية، بينما حصلت مهارة العلاقات الزمانية والمكانية على أقل نسبة، أما مهارات العلم التكاملية فقد حصلت مهارة تفسير البيانات على أعلى نسبة، بينما مهارة ضبط المتغيرات حصلت على

أقل نسبة.

وأجريت دراسة همام (2018) بهدف تقويم اختبارات طلاب الشعب العلمية بكلية التربية جامعة المنيا في ضوء معايير جودة الاختبارات، وقد اتبع الباحث أسلوب تحليل المحتوى أحد أساليب المنهج الوصفي، وتكونت العينة من (60) ورقة اختبارية للعام الدراسي (2016 – 2017). وصُممت استمارة تحليل مكونة من ثلاثة معايير، واستخدمت التكرارات والنسب المئوية والمتوسطات والانحرافات المعيارية لتحليل البيانات، وأظهرت نتائج الدراسة أن هناك قصورًا واضحًا في الاختبارات التي تعد من قبل أعضاء هيئة التدريس للشعب العلمية الثلاث بكلية التربية جامعة المنيا، سواء من حيث الشكل أو المضمون أو نوعية الأسئلة، وخاصة ما يتعلق بالمضمون؛ حيث جاءت أغلب الأسئلة عند المستويات الدنيا من التفكير وأغفلت المستويات العليا من التفكير.

وسعت دراسة القواس (2017) إلى تقويم اختبارات القبول لمادة الرياضيات بجامعة إب في ضوء معايير الاختبار الجيد، واتبع الباحث أسلوب تحليل المحتوى أحد أساليب المنهج الوصفي، وتكونت العينة من (241) سؤالاً من أسئلة اختبارات القبول لمادة الرياضيات للأعوام من (2010 – 2015). وتم بناء استمارة تحليل مكونة من ثمانية معايير، واستخدمت التكرارات والنسب المئوية لتحليل البيانات، وكشفت نتائج الدراسة أن الاختبارات لا تراعي معيار المواصفات الفنية للورقة الاختبارية؛ حيث أهملت بعض المعلومات الأساسية، كما أنها تخلو من الأخطاء العلمية، فضلًا عن أن الاختبارات لا تراعي معيار توفر التعليمات، بالإضافة إلى أن أسئلة الاختبارات ركزت على قياس المستويات العقلية الدنيا.

وأجريت دراسة الناقة (2016) بهدف تقويم أسئلة اختبارات الكيمياء للصف الثاني عشر خلال السنوات من (2007 إلى 2015) في فلسطين، وفق معايير الجودة المتمثلة في: (الشمولية، المطابقة، النواحي الفنية، المستويات المعرفية). واتبع الباحث أسلوب تحليل المحتوى، وتم بناء أداة لتقويم الأسئلة، ولتحليل بيانات الدراسة استخدمت التكرارات والنسب المئوية، وأسفرت الدراسة عن تصنيف الناحية الفنية حصل على أعلى نسبة، بينما حصلت المطابقة على أقل نسبة، أما المستويات المعرفية فقد اقتصرت معظم الأسئلة على المستويات الدنيا دون المستويات العليا من التفكير.

وأجرى الصلاحي (2008) دراسته بهدف تقويم الاختبارات النهائية في التخصصات الإنسانية للفصل الدراسي الأول من العام الجامعي (2004–2005) في جامعة صنعاء، وفق معايير الاختبارات الجيدة ، واتبع الباحث أسلوب تحليل المحتوى، وتكونت العينة من (86) ورقة اختبارية، وتم بناء أداة لتقويم الأسئلة، واستخدمت التكرارات والنسب المئوية لتحليل بيانات الدراسة، وبينت نتائج الدراسة أن بعض الأوراق الاختبارية أغفلت العديد من الجوانب الفنية المهمة، وأن معظم الأوراق الاختبارية تركز على الاختبارات المقالية والموضوعية، كما أهملت مراعاة شمول الاختبارات لجميع المستويات المعرفية؛ حيث ركزت على المستويات المعرفية الدنيا وخاصة التذكر ثم الفهم.

مما سبق يتضح أن جميع الدراسات السابقة تناولت تقويم الاختبارات بصورة عامة، واتفقت الدراسة الحالية مع دراسة القواس (2017) التي تناولت تقويم اختبارات القبول لمادة الرياضيات، كما اتفقت

الدراسة الحالية مع الدراسات السابقة في استخدامها أسلوب تحليل المحتوى القائم على المنهج الوصفي، وكذلك استخدامها لاستمارة تحليل المحتوى لتحليل أسئلة الاختبارات. وتميزت الدراسة الحالية عن الدراسات السابقة كونها من الدراسات المحلية النادرة التي تناولت تقويم اختبارات القبول لمادة الأحياء للمتقدمين لكلية الطب في جامعة إقليم سبأ. واستفادت الدراسة الحالية من الدراسات السابقة في بناء قائمة معايير الاختبار الجيد.

منهجية الدراسة وإجراءاتها:

أولًا: منهج الدراسة:

تم استخدام المنهج الوصفي من خلال استخدام أسلوب تحليل المحتوى الذي يُعد الأنسب لتحقيق هدف الدراسة، والإجابة عن أسئلتها.

ثانيًا: مجتمع الدراسة وعينتها:

تكونت عينة الدراسة من مجتمع الدراسة نفسه المتمثلة في جميع اختبارات القبول لمادة الأحياء بكلية الطب بجامعة إقليم سبأ لخمس سنوات للفترة من (2020 إلى 2024) المتوافرة في كلية الطب بجامعة إقليم سبأ، البالغ عددها خمسة نماذج وكل نموذج مكون من (25) سؤالًا؛ أي أن مجموع الأسئلة التي تم تحليلها (125) سؤالًا من نوع الاختيار من متعدد.

ثالثًا: أداتا الدراسة:

لتحقيق هدف الدراسة أعد الباحثان قائمة معايير الاختبار الجيد واستمارة تحليل المحتوى، وقد مر بناءهما بالخطوات الآتية:

أولًا: قائمة معايير الاختبار الجيد:

تم بناء قائمة معايير الاختبار الجيد من خلال الخطوات الآتية:

1. الهدف من القائمة:

تهدف قائمة المعايير إلى تحديد معايير الاختبار الجيد اللازم توافرها في أسئلة اختبارات القبول لمادة الأحياء للمتقدمين لكلية الطب في جامعة إقليم سبأ.

2. بناء القائمة في صورتها الأولية:

تم بناء قائمة معايير اختبارات القبول وذلك من خلال الاطلاع والاستفادة من العديد من الأدبيات والدراسات ذات العلاقة (الصلاحي، 2008، القواس، 2017؛ همام، 2018). وتكونت القائمة في صورتها الأولية من (31) مؤشرًا موزعة على أربعة معايير رئيسة، وهي على النحو الآتي:

1. معيار مواصفات ورقة الاختبار الجيد ووضوح التعليمات (المعايير والمواصفات للورقة الاختبارية) ويشمل (6) معايير (مؤشرات) فرعية.
2. المعايير الفنية للاختبار الجيد، ويشمل (7) معايير (مؤشرات) فرعية.
3. معايير الاختيار من متعدد، ويشمل (12) معيارًا (مؤشرًا) فرعيا.

4. معيار المستوى المعرفي للأسئلة، ويشمل (6) معايير (مؤشرات) فرعية.

3. التحقق من صدق الأداة:

تم عرض قائمة المعايير بصورتها الأولية على (7) من أعضاء هيئة التدريس المتخصصين في قسم المناهج وطرق التدريس بجامعة إقليم سبأ؛ لإبداء الرأي في مدى ملاءمتها وشمولها، وتم أخذ ملاحظاتهم بعين الاعتبار. وتم إجراء بعض التعديلات في صياغة بعض المعايير، وإضافة مؤشرين، أحدهما في المعايير الفنية للاختبار، والآخر في معايير الاختيار من متعدد، دون إجراء حذف أي منها.

4. القائمة في صورتها النهائية:

بعد مراجعة ملاحظات المحكمين وتعديل الصورة الأولية، أصبحت القائمة النهائية مكوّنة من (32) مؤشرًا فرعيًا موزعًا على أربعة معايير رئيسة.

ثانيًا: استمارة تحليل المحتوى:

تم بناء الاستمارة وفق الخطوات الآتية:

1. الهدف من استمارة تحليل المحتوى:

تهدف الاستمارة إلى تحليل أسئلة اختبارات القبول لمادة الأحياء للمتقدمين لكلية الطب في جامعة إقليم سبأ وفقًا لمعايير الاختبار الجيد.

2. بناء استمارة تحليل المحتوى:

تم تحويل قائمة المعايير في صورتها النهائية إلى استمارة تحليل المحتوى؛ حيث اشتملت على مهارات الاختبار الجيد ومؤشراتها الفرعية.

3. التحقق من ثبات التحليل:

تم حساب ثبات التحليل؛ حيث تم استخدام الثبات عبر الزمن من خلال إعادة التحليل لأسئلة الاختبار بعد أسبوعين من التحليل الأول، وتم حساب نسبة الاتفاق بين التحليلين باستخدام معادلة (هوليستي) حيث بلغت نسبة الاتفاق (93%) وهي نسبة عالية للثبات، مما يدل على ثبات التحليل.

ثالثًا: خطوات التحليل:

تم إجراء التحليل وفقًا للخطوات الآتية:

1. **تحديد الهدف من التحليل:** هدفت عملية التحليل إلى تحليل أسئلة اختبارات القبول لمادة الأحياء للمتقدمين لكلية الطب في جامعة إقليم سبأ، وفقًا لمعايير الاختبار الجيد.
2. **تحديد عينة التحليل:** حيث شملت عينة التحليل جميع أسئلة اختبارات القبول لمادة الأحياء للمتقدمين لكلية الطب في جامعة إقليم سبأ للأعوام (2020– 2024).
3. **تحديد فئات التحليل:** تم اعتماد معايير الاختبار الجيد كفئات تحليل رئيسة، كما تم اعتماد المؤشرات الدالة على هذه المعايير فئات التحليل الفرعية.
4. **تحديد وحدة التحليل:** تم اعتبار كل فقرة من فقرات أسئلة الاختبار وحدة للتحليل؛ كون ذلك مناسب لعملية التحليل.

5. **صدق القائمة:** تم التحقق من صدق القائمة من خلال عرضها على مجموعة من المحكمين وفقًا لما تمت الإشارة إلى ذلك مسبقًا.

6. **صدق التحليل:** تم عرض نموذج من الاختبارات التي خضعت للتحليل على مجموعة من المتخصصين في المناهج وطرائق تدريس العلوم والقياس والتقويم، وقد أجمعوا على صلاحية التحليل.

7. **ثبات التحليل:** تم التحقق من ثبات التحليل من خلال استخدام الثبات عبر الزمن.

رابعًا: إجراءات التطبيق:

أجريت الدراسة وفقًا للخطوات الآتية:

1. استخراج الأوراق الاختبارية للأعوام (2020-2024) المتوافرة في كلية الطب بجامعة إقليم سبأ.
2. قراءة النماذج قراءة دقيقة لتحليل محتوى الأوراق في ضوء معيار مواصفات ورقة الاختبار الجيد والتعليمات، ورصد التكرارات في استمارة التحليل.
3. قراءة كل سؤال قراءة دقيقة لتحليل أسئلة الاختبار في ضوء المعايير الفنية للاختبار الجيد، ورصد التكرارات في استمارة التحليل.
4. تحليل محتوى الاختبارات في ضوء معايير الاختيار من متعدد، من خلال قراءة كل سؤال وفحصه؛ للتأكد من وجود هذه المعايير من عدمها ورصد التكرارات في استمارة التحليل.
5. تحليل الاختبارات في ضوء معيار المستوى المعرفي الذي يقيسه كل سؤال من خلال قراءة كل سؤال قراءة دقيقة، ثم رصد التكرارات في استمارة التحليل.
6. حساب النسبة المئوية لكل التكرارات في استمارة التحليل؛ حتى يسهل تفسيرها.

أساليب المعالجة الإحصائية:

تم استخدام الأساليب الإحصائية الآتية: معادلة (هوليستي) لحساب الثبات، والتكرارات والنسب المئوية لتحليل النتائج.

عرض نتائج الدراسة ومناقشتها:

أولًا: النتائج المتعلقة بالإجابة عن السؤال الأول:

نص السؤال الأول على: **ما معايير جودة الاختبار التي ينبغي أن تتوافر في اختبار قبول مادة الأحياء للمتقدمين لكلية الطب؟**

للإجابة عن هذا السؤال، تم الاطلاع والاستفادة من الأدبيات التربوية والدراسات ذات الصلة العلمية بموضوع الدراسة، وتم إعداد قائمة بمعايير جودة الاختبار اللازم توافرها في **اختبارات القبول لمادة الأحياء بكلية الطب في الجامعات اليمنية**، ومن ثم التحقق من الصدق الظاهري لها، وتكونت القائمة في صورتها النهائية من (32) مؤشرًا فرعيًا وزعت على أربعة معايير رئيسة، والجدول (1) يوضح ذلك.

جدول (1) معايير الاختبار الجيد اللازم توافرها في اختبارات القبول لمادة الأحياء بكلية الطب في الجامعات اليمنية

المعيار الرئيس	م	المعيار الفرعي
مواصفات ورقة الاختبار الجيد ووضوح التعليمات	1	تضمين ورقة الاختبار للعبارات الإرشادية الخاصة بذيل الصفحة، مثل (يتبع، أو تمت الأسئلة)
	2	ترقيم أسئلة الاختبار بشكل متسلسل دون تكرار رقم السؤال
	3	تضمين ورقة الاختبار البيانات الخاصة برأس الصفحة، وهي (اسم الجامعة، شعار الجامعة، الكلية، العام، اسم المقرر)
	4	ترقيم أوراق الاختبار بشكل صحيح، تحديد الدرجة المخصصة للسؤال.
	5	تحديد عدد صفحات الاختبار.
المعايير الفنية للاختبار الجيد	1	صياغة الأسئلة بلغة سليمة.
	2	فقرات الأسئلة خالية من الأخطاء النحوية.
	3	مقدمة السؤال خالية من الأخطاء النحوية.
	4	الأسئلة خالية من الأخطاء الإملائية.
	5	الأسئلة خالية من الأخطاء العلمية.
	6	ترك مسافة بين كل سؤال وآخر.
	7	توافر جودة الطباعة للاختبار.
	8	بنط الخط المستخدم في الاختبار موحد في جميع الفقرات.

المجلد(8) العدد(2) ديسمبر 2025م
https://doi.org/10.54582/TSJ.2.2.126

المعيار الرئيس	م	المعيار الفرعي
معايير الاختيار من متعدد	1	مقدمه الفقرات مناسبة من حيث طولها.
	2	البدائل مرتبطة مع مقدمة الفقرة.
	3	إبراز النفي بالسؤال بوضع خط تحت النفي.
	4	خلو بدائل الإجابة من أي تلميح للإجابة الصحيحة مثل الارتباطات اللفظية بين أصل السؤال والإجابة الصحيحة.
	5	تساوي البدائل في جميع الفقرات.
	6	ترقيم الفقرات وأبجدة البدائل.
	7	تجنب استخدام بقدر الإمكان عبارات، مثل: جميع ما ذكر، أو ليس مما سبق، أو عبارة (أ) و (ب) في البدائل.
	8	تجنب تكرار مفاهيم ضمن البدائل.
	9	تجانس البدائل بحيث لا تحتمل التخمين.
	10	مقدمة السؤال تعكس مشكلة مهمة.
	11	توزيع البدائل الصحيحة بشكل عشوائي.
	12	شمول الاختبار للوحدات المختارة.
	13	أسئلة الاختبار لا تحتمل أكثر من معنى.
المستوى المعرفي للأسئلة	1	شمول الأسئلة لمستوى التذكر.
	2	شمول الأسئلة لمستوى الفهم.
	3	شمول الأسئلة لمستوى التطبيق.
	4	شمول الأسئلة لمستوى التحليل.
	5	شمول الأسئلة لمستوى التركيب.
	6	شمول الأسئلة لمستوى التقويم.

يتضح من الجدول (1) ما يأتي:

- المعيار الأول: مواصفات ورقة الاختبار الجيد ووضوح التعليمات واشتملت على (5) معايير فرعية.
- المعيار الثاني: المعايير الفنية للاختبار الجيد واشتملت على (7) معايير فرعية.
- المعيار الثالث: معايير الاختيار من متعدد واشتملت على (13) معيارًا فرعيا.

- المعيار الرابع: معيار المستوى المعرفي للأسئلة واشتملت على (6) معايير فرعية.

ثانيًا: النتائج المتعلقة بالإجابة عن السؤال الثاني:

نص السؤال الثاني على: **ما مستوى توافر معيار مواصفات الورقة الاختبارية ووضوح التعليمات في اختبارات القبول لمادة الأحياء للمتقدمين لكلية الطب في جامعة إقليم سبأ للأعوام (2020–2024)؟**

للإجابة عن هذا السؤال، تم رصد التكرارات والنسب المئوية للمعايير الفرعية لمعيار مواصفات الورقة الاختبارية ووضوح التعليمات المتوافرة في اختبارات القبول، لمادة الأحياء للمتقدمين لكلية الطب في جامعة إقليم سبأ للأعوام (2020–2024). ويوضح الجدول (2) ذلك.

جدول (2): التكرارات والنسب المئوية لمعايير مواصفات الورقة الاختبارية

م	المؤشرات	التكرار		النسبة المئوية	
		متوفر	غير متوفر	متوفر	غير متوفر
1	تضمين ورقة الاختبار للعبارات الإرشادية الخاصة بذيل الصفحة، مثل: يتبع، أو تمت الأسئلة.	0	5	0%	100%
2	ترقيم فقرات أسئلة الاختبار بشكل متسلسل دون تكرار رقم السؤال	5	0	100%	0%
3	تضمين ورقة الاختبار البيانات الخاصة برأس الصفحة، وهي: اسم الجامعة، شعار الجامعة، الكلية، العام، اسم المقرر.	5	0	100%	0%
4	ترقيم أوراق الاختبار بشكل صحيح.	5	0	100%	0%
5	تحديد عدد صفحات الاختبار.	3	2	60%	40%

يتضح من الجدول (2) أن غالبية المؤشرات الخاصة بورقة الاختبار في الخمسة النماذج قد تحققت بنسبة (100%) مثل: وجود العبارات الإرشادية، ترقيم الأسئلة بشكل متسلسل، إدراج البيانات الخاصة بالجامعة، وترقيم أوراق الاختبار بشكل صحيح.

بينما أظهر معيار تحديد عدد صفحات الاختبار بعض القصور؛ حيث تحقق بنسبة (60%) فقط، في حين أن (40%) من الأوراق لم يحدد فيها عدد الصفحات.

وبالرجوع إلى الجدول السابق نلاحظ أن بعض المعايير متوفرة بنسبة (100%) مثل: «ترقيم فقرات

أسئلة الاختبار بشكل متسلسل»، و»تضمين البيانات الخاصة برأس الصفحة»، و»ترقيم أوراق الاختبار بشكل صحيح»، هذه النتائج تشير إلى وجود اهتمام بالجوانب التنظيمية الأساسية للورقة الاختبارية، أما بالنسبة للمؤشر: «تحديد عدد صفحات الاختبار»، فتوافرت بنسبة (60%)؛ وهذه النتيجة تسلط الضوء على الحاجة لتحسينات في تصميم الورقة الاختبارية لضمان وضوح التعليمات وتوجيه الطلبة بشكل أفضل في أثناء الاختبار. ومن ناحية أخرى يظهر الجدول ضعفاً كبيراً في توافر مؤشرات أخرى، على سبيل المثال: المؤشر «تضمين ورقة الاختبار للعبارات الإرشادية الخاصة بذيل الصفحة، مثل: يتبع، أو تمت الأسئلة.».

ثالثًا: النتائج المتعلقة بالإجابة عن السؤال الثالث:

نص السؤال الثالث على: **ما مستوى توافر المعايير الفنية للاختبار الجيد في أسئلة اختبار قبول مادة الأحياء للمتقدمين لكلية الطب في جامعة إقليم سبأ للأعوام (2020–2024)؟**

للإجابة عن هذا السؤال، تم رصد التكرارات والنسب المئوية للمعايير الفنية للاختبار الجيد المتوافرة في أسئلة اختبار قبول مادة الأحياء للمتقدمين لكلية الطب في جامعة إقليم سبأ للأعوام (2020-2024). والجدول (3) يوضح ذلك.

جدول (3) التكرارات والنسب المئوية للمعايير الفنية للاختبار الجيد

م	المؤشرات	التكرار		النسبة المئوية	
		متوفر	غير متوفر	متوفر	غير متوفر
1	صياغة الأسئلة بلغة سليمة.	125	0	100%	0%
2	فقرات الأسئلة خالية من الأخطاء النحوية.	125	0	100%	0%
3	مقدمة السؤال خالية من الأخطاء النحوية	0	5	0%	100%
4	الأسئلة خالية من الأخطاء الإملائية.	62	63	50%	50%
5	الأسئلة خالية من الأخطاء العلمية.	121	4	97%	3%
6	ترك مسافة بين كل الفقرات/البدائل.	0	125	0%	100%
7	توافر جودة الطباعة للاختبار.	5	0	100%	0%
8	بنط الخط المستخدم في الاختبار موحد في جميع الفقرات.	50	75	40%	60%

يُظهر الجدول (3) أن جميع أسئلة الاختبار كانت مصاغة بلغة عربية سليمة وخالية من الأخطاء اللغوية بنسبة (100%). وجميع فقرات الأسئلة كانت خالية تمامًا من الأخطاء النحوية بنسبة (100%). بينما جميع مقدمات الأسئلة في الخمسة النماذج احتوت على أخطاء نحوية بنسبة (100%). والأسئلة خالية من الأخطاء الإملائية متوفر بنسبة (50%) و (63) تكرارًا؛ أي أن نصف هذه الأسئلة تقريبًا كان بها أخطاء إملائية، والنصف الآخر كان خاليًا منها، بينما الأسئلة خالية من الأخطاء العلمية متوفرة بنسبة (97%) و (121) تكرارًا وغير متوفر بنسبة (3%) و(4) تكرارات؛ حيث أن الغالبية العظمى من الأسئلة كانت دقيقة من الناحية العلمية مع وجود عدد قليل جدًا من الأخطاء.

وكانت جودة طباعة أوراق الاختبار في الخمسة النماذج ممتازة ومتوافرة بنسبة (100%). بينما بنط الخط المستخدم في الاختبار موحد في جميع الفقرات كان متوفر بنسبة (40%) و (50) تكرارًا، وغير متوفر بنسبة (60%) و (75) تكرارًا.

وبالرجوع إلى الجدول السابق نلاحظ أن المؤشرات الآتية: «صياغة الأسئلة بلغة سليمة»، و»فقرات الأسئلة خالية من الأخطاء النحوية»، و»توافر جودة الطباعة» حصلت على نسبة (100%). بينما المؤشر: «الأسئلة خالية من الأخطاء العلمية» متوافرة بنسبة (97%) وهي نسبة عالية جدًا؛ حيث ظهرت الأخطاء العلمية في أربعة فقرات منها (هرمون الباراثومين) ولا يوجد هرمون بهذا المسمى والأصح علميًا هو (هرمون الباراثورمون) والخطأ الثاني والثالث في الفقرة التي تنص على: «بعض المحاور العصبية تكون مغطاة بمادة دهنية بروتينية تسمى: أ. غلاف الميلانين، ب. غلاف اللاميلينية ج. غلاف ليفي، د. غلاف عضلي ليفي». فهناك خطآن علميان في البدائل (غلاف الميلانين و غلاف اللاميلينية)؛ حيث إن المصطلح الصحيح هو (غلاف المايلين) أما غلاف الميلانين هذا خطأ، فالميلانين هو صبغة موجودة في الجلد والشعر والعينين وليس له علاقة بالمحاور العصبية، وكذلك غلاف اللاميلينية غير صحيح علميًا، والصحيح هو المحاور العصبية غير المايلينية، والخطأ الرابع والأخير كان الفقرة التي تنص على: «تحاط بغشاء واحد وتعمل على إنتاج البروتين: أ. الفجوات ب. الرايبوسومات ج. الجسم المركزي د. الليسوسومات». حيث اشتمل السؤال على خاصيتين هما (الغشاء وبناء البروتين)؛ فالرايبوسومات تنتج البروتين لكن غير محاطة بغشاء وبقية البدائل محاطة بغشاء لكنها لا تنتج البروتين، لكن هناك جهوداً جيدة نوعًا لضمان السلامة اللغوية والعلمية للأسئلة. ومع ذلك يكشف الجدول عن تحديات في جوانب فنية أخرى، فالمؤشر «خلو الأسئلة من الأخطاء الإملائية» متوافر بنسبة (50%) فقط، وهي نسبة متوسطة، مما يشير إلى الحاجة للتحسين في التدقيق الإملائي، بينما المؤشر «توحيد بنط الخط المستخدم في الاختبار» متوافر بنسبة (40%) وهي نسبة منخفضة، مما يؤثر أيضاً على الجودة الفنية والتنسيق العام للورقة الاختبارية، في حين المؤشرات الآتية: «مقدمة الأسئلة من الأخطاء النحوية»، و»ترك مسافة بين كل الفقرات/البدائل» ليست متوافرة على الاطلاق؛ حيث بلغت نسبة توافرها (0%)؛ وظهرت الأخطاء النحوية في كلمة (اختار) والأصح هو (اختر) وهذا يؤثر سلباً في تنسيق الاختبار وقابليته للقراءة.

رابعًا: النتائج المتعلقة بالإجابة عن السؤال الرابع:

نص السؤال الرابع على: **ما مستوى توافر معايير الاختيار من متعدد في أسئلة اختبار قبول مادة الأحياء للمتقدمين لكلية الطب في جامعة إقليم سبأ للأعوام (2020–2024)؟**

للإجابة عن هذا السؤال، تم رصد التكرارات والنسب المئوية للمعايير الفرعية لمعايير الاختيار من متعدد المتوافرة في أسئلة اختبار قبول مادة الأحياء للمتقدمين لكلية الطب بجامعة إقليم سبأ للأعوام (2020–2024). والجدول (4) يوضح ذلك.

جدول (4) التكرارات والنسب المئوية لمعايير الاختيار من متعدد

م	المؤشر	التكرار		النسبة المئوية	
		متوفر	غير متوفر	متوفر	غير متوفر
1	مقدمه الفقرات مناسبة من حيث طولها.	125	0	100%	0%
2	البدائل مرتبطة مع مقدمة الفقرة.	125	0	100%	0%
3	إبراز النفي بالسؤال بوضع خط تحت النفي.	0	5	0%	% 100
4	خلو بدائل الإجابة من أي تلميح للإجابة الصحيحة، مثل: الارتباطات اللفظية بين أصل السؤال والإجابة الصحيحة.	125	0	100%	0%
5	تساوي البدائل في جميع الفقرات.	125	0	100%	0%
6	ترقيم الفقرات وأبجدة البدائل	125	0	100%	0%
7	تجنب استخدام عبارات، مثل: جميع ما ذكر، أو ليس مما سبق، أو عبارة (أ) و (ب) في البدائل.	89	36	71%	29%
8	تجنب تكرار مفاهيم ضمن البدائل	125	0	100%	0%
9	تجانس البدائل بحيث لا تحتمل التخمين.	125	0	100%	0%

م	المؤشر	التكرار		النسبة المئوية	
		متوفر	غير متوفر	متوفر	غير متوفر
10	مقدمة السؤال تعكس مشكلة مهمة.	125	0	100%	0%
11	توزيع البدائل الصحيحة بشكل عشوائي.	4	1	80%	20%
12	شمول الاختبار للوحدات المختارة.	125	0	100%	0%
13	أسئلة الاختبار لا تحتمل أكثر من معنى.	125	0	100%	0%

يتبين من الجدول (4) أن جميع الفقرات كانت مناسبة من حيث الطول بنسبة (100%). وجاءت البدائل مرتبطة بمقدمة الفقرة بنسبة (100%). كما أن جميع الفقرات خلت من أي تلميح للإجابة الصحيحة، وحققت تساوي البدائل وترقيمها بنسبة (100%) إلا أن معيار إبراز النفي بوضع خط تحته لم يتحقق إطلاقاً (100% غير متوفر). كما ظهر قصور في معيار تجنب العبارات العامة، مثل: «جميع ما ذكر» حيث تحقق بنسبة (71%) فقط، أما توزيع البدائل بشكل عشوائي فقد تحقق بنسبة (80%) فقط، في حين أن بقية المؤشرات مثل تجنب التكرار وتجانس البدائل ومقدمة السؤال التي تعكس مشكلة مهمة وشمولية الاختبار قد تحققت كاملة بنسبة (100%).

وبالرجوع إلى الجدول السابق نلاحظ أن المؤشرات رقم (1، 2، 4، 5، 6، 8، 9، 10، 12، 13) حصلت على نسبة توافر (100%) وبنسبة عالية، وتوافر المؤشر رقم (7) «تجنب استخدام عبارات مثل جميع ما ذكر، أو ليس مما سبق، أو عبارة أ و ب في البدائل» بنسبة (89%). في حين توافر المؤشر رقم (11) «توزيع البدائل الصحيحة بشكل عشوائي» بنسبة (80%). في حين توافر المؤشر رقم (3) «إبراز النفي بالسؤال بوضع خط تحت النفي» توافر بنسبة (0%) ويعزى غياب إبراز النفي إلى ضعف إدراك واضعي الأسئلة لأهمية هذا المعيار في وضوح الصياغة ودقة الفهم؛ حيث إن إغفال تمييز النفي يعد من الأخطاء الشائعة في بناء أسئلة الاختيار من متعدد ويظهر عادة نتيجة الاعتماد على خبرات شخصية في إعداد الأسئلة دون الرجوع إلى الأدلة الإرشادية أو المعايير التربوية المتخصصة؛ مما يؤدي إلى ضعف الالتزام بالمعايير الفنية للاختبار الجيد. وهذه النتيجة تشير إلى أن بناء أسئلة الاختيار من متعدد تتبع إلى حد كبير معايير أسئلة الاختيار من متعدد، ومع ذلك هناك بعض المؤشرات التي تحتاج إلى اهتمام منها: إبراز النفي بالسؤال بوضع خط تحت النفي، ويعد هذا المعيار مهم؛ لتجنب التباس الطلبة، وتجنب استخدام عبارات مثل جميع ما ذكر، أو ليس مما سبق، أو عبارة أ و ب في البدائل، التي استخدمت بشكل متكرر في اختبار القبول لمادة الأحياء، وأخيراً قلة توزيع البدائل الصحيحة بشكل عشوائي، مما قد يساعد الطلبة في تخمين البديل الصحيح.

خامسًا: النتائج المتعلقة بالإجابة عن السؤال الخامس:

نص السؤال الخامس على: **ما مستوى مراعاة اختبارات القبول لمادة الأحياء في كلية الطب بجامعة إقليم سبأ للأعوام (2020–2024) لمعيار المستوى المعرفي للأسئلة؟**

للإجابة عن هذا السؤال، تم رصد التكرارات والنسب المئوية للمعايير الفرعية لمعيار المستوى المعرفي للأسئلة في اختبار قبول مادة الأحياء للمتقدمين لكلية الطب بجامعة إقليم سبأ للأعوام (2020-2024). والجدول (5) يوضح ذلك.

جدول (5) التكرارات والنسب المئوية لمعيار المستوى المعرفي للأسئلة

م	المعيار	التكرار	النسبة المئوية
1	شمول الأسئلة لمستوى التذكر.	88	70%
2	شمول الأسئلة لمستوى الفهم.	33	26%
3	شمول الأسئلة لمستوى التطبيق.	2	2%
4	شمول الأسئلة لمستوى التحليل.	2	2%
5	شمول الأسئلة لمستوى التركيب.	0	0%
6	شمول الأسئلة لمستوى التقويم.	0	0%
المجموع		**125**	**100%**

يتضح من الجدول (6) أن معظم الأسئلة ركزت على المستوى الأدنى من المجال المعرفي وهو التذكر بنسبة (70%) وبعدد (88) سؤالاً، تلاه مستوى الفهم بنسبة (26%) وبعدد (33) سؤالاً، بينما جاء مستوى التطبيق والتحليل بنسبة متدنية جداً بلغت (2%) لكل منهما وبعدد سؤالين فقط، أما المستويات العليا (التركيب والتقويم) فقد غابت تماماً عن الاختبارات بنسبة (0%).

وبالرجوع إلى الجدول السابق نلاحظ أن أسئلة اختبار القبول لمادة الأحياء اشتملت على مستوى التذكر بنسبة (70%) وبدرجة عالية جدًا، يليه مستوى الفهم بنسبة (26%)، وبدرجة عالية، أما مستويي التطبيق والتحليل، فجاءت بنسبة (2%) لكل منهما، وبدرجة منخفضة، في حين أن مستويي التركيب والتقويم غير متوفرين على الإطلاق، فجاءت نسبة توافرهما بنسبة (0%) مما يدل على اعتماد كبير على المستويات الدنيا للمعرفة وإهمال شبه كامل للمستويات العليا في بناء أسئلة القبول. واتفقت هذه النتيجة مع نتيجة دراسة (القواس، 2017) ودراسة (الصلاحي، 2008) التي توصلت إلى أن الاختبارات لا تراعي معايير الاختبار الجيد.

وهذه النتيجة تشير بوضوح إلى أن اختبارات القبول لمادة الأحياء تركز بشكل أساسي على قياس قدرة الطلبة على استرجاع المعلومات وتذكرها، مع إهمال شبه كامل للمستويات المعرفية العليا التي تتطلب

التفكير النقدي، وحل المشكلات، والتحليل والتركيب، وهذا يثير تساؤلات حول مدى قدرة الاختبار على تقييم القدرات المعرفية الشاملة للمتقدمين لكلية الطب.

التوصيات:

بناءً على النتائج توصي الدراسة بالآتي:

1. عقد دورات تدريبية وورش عمل لأعضاء هيئة التدريس، لتدريبيهم على بناء الاختبارات وفقًا لمعايير الاختبار الجيد.
2. تضمين المستويات العليا (التطبيق، التحليل، التركيب، التقويم) في فقرات الاختبار بدلًا من التركيز على مستوى التذكر والفهم فقط.
3. التقييم المستمر للأسئلة للتأكد من وضوح الصياغة وخلوها من الأخطاء.
4. مراعاة معايير الجودة في تصميم الورقة الاختبارية ووضوحها وتنظيمها.

المقترحات:

في ضوء ما سبق تقترح الدراسة الآتي:

1. تقويم اختبارات القبول لمادتي الكيمياء واللغة الانجليزية بكلية الطب في جامعة إقليم سبأ في ضوء معايير الاختبار الجيد.
2. تقويم اختبارات القبول للتخصصات الأخرى (علوم الحاسوب) وفقًا لمعايير الاختبار الجيد.
3. تقويم اختبارات المقررات الدراسية في المرحلة الجامعية وفقًا لمعايير الاختبار الجيد.

المراجع:

أولًا: المراجع العربية:

1. أبو جامع، ليندا حرب، الأشقر، أحلام عدنان. (2021). تقويم أسئلة اختبارات الأحياء للصف الثاني عشر في ضوء عمليات العلم خلال السنوات من 2018 إلى 2020م. مجلة الجامعة الإسلامية للدراسات التربوية والنفسية، 29(4)، 78-106.

2. أبو علام، رجاء محمود. (2011). مناهج البحث في العلوم النفسية والتربوية (ط.5). دار النشر للجامعات.

3. اسعادي، فارس، وعزي، إيمان. (2015). مواصفات الاختبار التحصيلي الجيد. مجلة الدراسات والبحوث الاجتماعية، (11)، 189-200.

4. آغا، ناصر حسين، الشمالي، عيسى جاسم، العبيدلي، ناصر حسن، يوسف، حسين صالح ملا. (2023). تحليل محتوى امتحانات الثانوية العامة لمواد العلوم للصف الثاني عشر العلمي وفقًا لتصنيف بلوم المطور للمستويات المعرفية. المجلة التربوية، 37(45)، 53-84.

5. بركات، زياد. (2020). معايير الاختبار الجيد من وجهة نظر معلمي المرحلتين الأساسية والثانوية في المدارس الحكومية في محافظة طولكرم. مجلة الروانز، 4 (1)، 5-34.

6. البناء، مأمون. (2024). قياس وتقييم ثم تقويم تربوي ونفسي. دار وائل للنشر.

7. الجبلي، سوسن شاكر. (2005). أساسيات بناء الاختبارات والمقاييس النفسية والتربوية. مؤسسة علاء الدين للطباعة والتوزيع.

8. داؤود، أحمد عيسى. (2014). أصول التدريس «النظري والعملي» . دار يافا العلمية للنشر والتوزيع.

9. الرحيمي، حصة حمدان غبيش. (2020). تقويم الاختبارات المدرسية لمعلمات المرحلة الابتدائية في ضوء معايير الاختبار الجيد في مدارس إدارة التعليم بالمهد. المجلة العربية للنشر العلمي، 6 (62)، 294-343.

10. الشربيني، زكريا أحمد، إبراهيم، رزق سند، دمهنوري، رشاد صالح، مطحنة، السيد كمال. (2010). القياس النفسي. مكتبة الشقري.

11. الشيخ، تاج السر عبد الله؛ أخرس، نائل محمد؛ عبد المجيد، بثينة أحمد (2020). القياس والتقويم التربوي. مكتبة الرشد.

12. الصلاحي، عبد السلام. (2008). تقييم الاختبارات النهائية في جامعة صنعاء. مجلة الباحث الجامعي.

13. عبد الرحمن، أحمد محمد. (2011). تصميم الاختبارات أسس نظرية وتطبيقات عملية. دار أسامة للنشر والتوزيع.

14. عبد الهادي، نبيل. (2002). المدخل إلى القياس والتقويم التربوي واستخدامه في مجال التدريس الصفي (ط.2). دار وائل للطباعة والنشر.

15. عبيدات، ذوقان، عبد الحق، كايد، عدس، عبد الرحمن (2015). البحث العلمي مفهومة وأدواته وأساليبه (ط.17). دار الفكر للنشر والتوزيع.

16. القواس، محمد أحمد مرشد. (2017). تقويم اختبارات القبول لمادة الرياضيات بجامعة إب في ضوء معايير الاختبار الجيد. مجلة العلوم النفسية والتربوية، 5(1)، 257-278.

17. مجيد، سوسن شاكر. (2014). أسس بناء الاختبارات والمقاييس النفسية والتربوية (ط3). مركز ديبونو لتعليم التفكير للنشر والتوزيع.

18. مراد، صلاح أحمد، سليمان، أمين علي سليمان. (2005). الاختبارات والمقاييس في العلوم النفسية والتربوية خطوات إعدادها وخصائصها. دار الكتب الحديث.

19. الناقة، صلاح أحمد. (2016). تقويم أسئلة اختبارات الكيمياء للصف الثاني عشر وفق معايير الجودة خلال السنوات من 2007 إلى 2015. مجلة الجامعة الإسلامية للدراسات التربوية والنفسية، 24(3)، 64-82.

20. النبهان، موسى. (2013). أساسيات القياس في العلوم السلوكية. دار الشروق للنشر والتوزيع.

21. همام، عبد الرزاق سويلم. (2018). تقويم اختبارات طلاب الشعب العلمية بكلية التربية جامعة المنيا في ضوء معايير جودة الاختبارات. مجلة البحث في التربية وعلم النفس، 33(3)، 272-304.

22. اليعقوبي، حيدر. (2013). التقويم والقياس في العلوم التربوية والنفسية رؤيا تطبيقية. مركز المرتضى للتنمية الاجتماعية.

ثانيًا: المراجع الأجنبية:

1. Kocdar, S., Karadag, N., & Sahin, M. D. (2016). Analysis of the Difficulty and Discrimination Indices of Multiple-Choice Questions According to Cognitive Levels in an Open and Distance Learning Context. *The Turkish Online Journal of Educational Technology, 15*(4), 16-24.

2. Rossiter, J. R. (2011). *Measurement for the Social Sciences: The C-OAR-SE Method and Why It Must Replace Psychometrics*. Springer Science+Business Media.

3. Pal, K. (2020). Educational Measurement and Evaluation. USI Publications.

4. Sivakumar, A., & Moorthy, T. (2019). Measurement and Evaluation in Education. A.P.H. Publishing Corporation.