

تقويم اسئلة الاختبارات النهائية لمواد الدراسات الاجتماعية للصف العاشر في ضوء مستويات بلوم المعرفية وقواعد صياغة اسئلة الاختيار من متعدد

إعداد/ بدر بن حمد بن حميد البادي

طالب في مرحلة الدكتوراه، كلية التربية بالجامعة الإسلامية العالمية بماليزيا

E-mail: bader39390@hotmail.com

إشراف/ أ.د/ داود عبدالمك الحداوي

رئيس مكتب وكيل الجامعة ، التدويل والصناعة والعلاقات المجتمعية ، الجامعة الإسلامية العالمية
(ماليزيا)

ملخص البحث:

يهدف هذا البحث الى تقويم اسئلة الاختيار من متعدد في اختبارات مواد الدراسات الاجتماعية للصف العاشر خلال الخمس السنوات من (٢٠١٥ إلى ٢٠٢٠)، حيث تم رصد وتحديد المستويات المعرفية (التذكر، والفهم، والتطبيق، والتحليل، والتركيب، والتقويم) لأسئلة الاختيار من متعدد، وقد أظهرت نتائج البحث أن المستويات المعرفية الدنيا (التذكر، والفهم، والتطبيق) قد حصلت على أعلى نسبة بين جميع المستويات، ولاسيما مستوى التذكر الذي جاء في المرتبة الأولى، بينما نجد غياب للأسئلة في مستوى التركيب والتقويم، كذلك تم تطوير أداة لتقويم أسئلة الاختيار من متعدد، تكونت من (٢٧) قاعدة أو معيار للتقويم، موزعة ضمن ثلاثة محاور، المحور الاول خاص بمتن السؤال، والمحور الثاني خاص بالمموهات، والمحور الثالث خاص بالإخراج الفني، حيث تم حساب تكرار كل معيار في كل سنة دراسة وترتيبها ورصدها، ويتضح من نتائج البحث أن ابرز القواعد التي لم يتم الالتزام بها من قبل معدي اسئلة الاختيار من متعدد في اختبارات مواد الدراسات الاجتماعية كانت: وجود النفي أو عباراته في متن السؤال وأيضا غياب الوضوح في صياغة متن السؤال والصياغة بعبارات اقل، وفي المحور الثاني نجد قاعدة ترتيب البدائل ترتيب منطقي، وترتيب البدائل عموديا من ابرز القواعد التي لم يتم الالتزام بها، وفي المحور الثالث نجد قاعدة وضع العناصر المشتركة للبدائل في متن السؤال هي من ابرز القواعد التي تمت مخالفتها.

الكلمات المفتاحية:

تقويم الاسئلة، اختبارات مواد الدراسات الاجتماعية، جودة الاسئلة.

Evaluation of Final Examinations Questions of Social Studies of Grade Ten in Light of Bloom's Cognitive Taxonomy and Principles of Writing Multiple Choice Questions

Badar Hamed Humaid Al Badi*

PhD student in College of Education, Department of Educational Psychology,
International Islamic University (Malaysia), (IIUM)

Supervised by Prof. Dr. DAWOOD ABDULMALEK YAHYA AL-HIDABI

Head of office of the deputy rector, internationalisation and industry and community
relations, International Islamic University (Malaysia)

Abstract

This study aims at evaluating multiple choice questions in social studies subjects examinations, grade ten, in the period (2015-2020). The cognitive levels (remembering, understanding, applying, analyzing, synthesizing, and evaluating) of multiple choice questions were observed and identified. Results showed that the lower cognitive levels (i.e. remembering, understanding, and applying) had the highest percentage particularly the remembering level, which came first. There was obvious lack of analyzing and evaluating questions. An instrument was developed to evaluate multiple choice questions. It consisted of (27) criteria of evaluation explained in three themes; the content of the question, distractors, and layout. Each criterion's frequency of every year was calculated, ordered and observed. The results showed that the most prominent criteria not abided by the writers of multiple choice questions in the social studies subjects were the negative statements in the wording of the question and also the rule of clarity of the questions with the minimum statements required. For the second theme, the rules of ordering options logically and the rule of vertical ordering were obviously not abided by. For the third theme, the rule of including the elements in common of the options in the question itself was the most prominent rule not abided by.

Key words: question evaluation, examinations of social studies subjects, quality of questions.

*ADDRESS FOR CORRESPONDENCE: Badar Hamed Humaid Al Badi,
International Islamic University (Malaysia), Muscat, Oman. E-mail address:
bader39390@hotmail.com / Tel: +968 92858085.

المقدمة:

ترتبط العملية التعليمية ارتباطاً وثيقاً بالاختبارات، وعلى ذلك فإن البيانات والنتائج التي تقدمها الاختبارات باختلاف أنواعها، تسهم بشكل كبير في مساعدة أركان العملية التعليمية كافة، للوصول للأهداف التي حُطت لتحقيقها، فهي تساعد المعلم على الاطلاع ومعرفة المعلومات والمعارف والمهارات التي يمتلكها الطلبة، وهو ما يمكنه من إعداد الخطط التي تتلاءم مع تلك المعارف والمهارات والقيم، وأيضاً تحديد الاستراتيجيات التدريسية الفعالة التي يمكنه أن يستخدمها وتتلاءم مع الطلبة على اختلاف مستوياتهم. وهي أيضاً تقدم لإدارات المدارس وأولياء الأمور معلومات حول مستوى الطلبة والأبناء، وتكشف قدرتهم على تذكر الأفكار والمعارف وتحليلها وتنظيمها ومدى امتلاكهم مهارة حل المشكلات (عتوم، ٢٠١٤)، ونظراً لاعتبار الاختبارات أكثر وسائل التقويم شيوعاً، فلا بد من توافر معايير لإصدار الاحكام واتخاذ القرارات المناسبة حول جودة الاسئلة التي تُألفها، وعلى ذلك سعت العديد من الدراسات إلى تحليل اختبارات عدة؛ وفي مراحل تعليمية مختلفة، ومنها دراسة موهنوت (Mohnot, 2006) التي توصلت إلى أن أكثر من (٩٠%) من الاسئلة في الشهادة الهندية للتعليم الثانوي (ICSE) هي في مستوى المعرفة والفهم، كذلك دراسة سيفلاي وسيركان وورهان وصالح (Sevilay, Serkan, Orhan & Salih, 2003)، والتي توصلت إلى أن (٩٦%) من الاسئلة كانت ضمن المستويات المعرفية الدنيا، وغيرها العديد من الدراسات التي توصلت إلى نتائج متقاربة، ومن بينها دراسة الشقيرات (٢٠٢٠)، ودراسة الشاهر والحمداني (٢٠١٦)، والتي كان من أبرز نتائجها: عظم التركيز على المستويات العقلية الدنيا وأهملت المستويات العقلية العليا (التحليل، والتركيب، والتقويم). واختبارات مواد الدراسات الاجتماعية (الدراسات الاجتماعية، وتاريخ عمان) في المعاهد الإسلامية بحاجة إلى التحليل والتقويم للكشف عن جوانب القصور والنقص التي قد تعترضها ومدى شمولها على المستويات المعرفية المختلفة.

ونظر لكون اسئلة الاختيار من متعدد أكثر أنواع الاسئلة الموضوعية استخداماً، ولاسيما في المؤسسات التعليمية كما يُشير هالدينا وداوننج (Haladyna and Downing, 2004)، فإن تركيز الدراسة الحالية سيكون على اسئلة الاختيار من متعدد في الاختبارات النهائية لمواد الدراسات الاجتماعية، ومدى تنوعها حسب تصنيف بلوم للمهارات المعرفية الذي يعد الأكثر شيوعاً، حيث قام بنيامين بلوم بإنشاء تصنيف بلوم خلال الخمسينيات من القرن المنصرم؛ لتصنيف مستويات مهارات التفكير المطلوبة في المواقف الصفية، حيث تكون التصنيف من ستة مستويات من الإدراك مرتبة ترتيباً هرمياً؛ هذه المستويات هي: المعرفة والفهم والتطبيق والتحليل والتركيب والتقويم، حيث يجب على المعلم ومُعد الامتحان اعتماد المستويات المعرفية الكاملة عند كتابة الاختبار، فلا بد من تنوع مستويات الاسئلة وعدم التركيز على المستويات الدنيا (المعرفة والفهم والتطبيق)، بل يجب أن تشمل الاسئلة المستويات العليا (Alzu'bi, 2014). ويرى كيلاي (Kelly, 2014) أن هناك اسباباً عدة قد تجعل الاختبار من إعداد المعلم يقتصر على المستويات المعرفية الدنيا والتي من بينها: توقعات المعلم المنخفضة حول مستويات الطلبة وبالتالي حول قدراتهم، والامر الآخر هو الصعوبة التي يواجهها المعلم في صياغة الاسئلة في

المستويات العليا والوقت الذي يستغرقه في إعدادها، لذلك لابد من تقويم أسئلة الاختبارات لضمان التنوع والتوزيع العادل للأسئلة وفق مستويات بلوم المعرفية.

مشكلة الدراسة:

تقوم العملية التعليمية على التخطيط وتحديد الأهداف واختيار الاستراتيجيات والأنشطة والوسائل المناسبة لتحقيق الأهداف، وفي جميع تلك المراحل لابد من عملية التقويم، لاسيما وأنها بوصلة التوجيه لتحقيق الأهداف المرجوة بالشكل الصحيح، وأبرز أدوات التقويم في العملية التعليمية هي الاختبارات التحصيلية، والتي تُعرف على أنها إجراء منظم لقياس سمة ما من خلال عينة من السلوك (هنداوي، ٢٠١٠)، ولأهميتها الكبيرة في العملية التعليمية فلا بد وأن تتمتع بالموضوعية والصدق والثبات والبراعة في الإعداد والإخراج، وعلى ذلك فنحن بحاجة لمعايير لأجل إصدار الاحكام واتخاذ القرارات المناسبة حول جودة أسئلة الاختبارات ومدى تنوع مستويات الاسئلة فيها، والدراسة الحالية تسلط الضوء على تقويم أسئلة الاختيار من متعدد في اختبارات مواد الدراسات الاجتماعية للمعاهد الإسلامية للصف العاشر بسلطنة عمان في الأعوام الدراسية الممتدة بين (٢٠١٥ إلى ٢٠٢٠)، وفق معايير الاختبار التحصيلي الجيد وتحديد مستوى الاسئلة في الاختبارات وفق تصنيف بلوم للأهداف المعرفية.

أسئلة الدراسة:

جاءت الدراسة الحالية للإجابة على الاسئلة التالية:

- ١- ما مدى توافر معايير الجودة للمستويات المعرفية (التذكر، والفهم، والتطبيق، والتحليل، والتركيب، والتقويم) في اسئلة اختبارات مواد الدراسات الاجتماعية للصف العاشر؟
- ٢- ما معايير جودة الاسئلة لاختبارات مواد الدراسات الاجتماعية التي يمكن التقويم على أساسها؟
- ٣- ما مدى تطابق بناء اسئلة الاختيار من متعدد في اختبار مواد الدراسات الاجتماعية للصف العاشر في المعاهد الإسلامية مع المعايير التربوية الجيدة للاختبار التحصيلي؟

أهمية الدراسة:

تكمن أهمية الدراسة الحالية في:

- ١- الكشف عن مدى توافر معايير الجودة للمستويات المعرفية (التذكر، والفهم، والتطبيق، والتحليل، والتركيب، والتقويم) في اسئلة اختبارات مواد الدراسات الاجتماعية للصف العاشر.
- ٢- الكشف عن مدى تطابق اسئلة الاختيار من متعدد في اختبار مواد الدراسات الاجتماعية للصف العاشر في المعاهد الإسلامية مع المعايير التربوية للاختبار الجيد.

٣- الإسهام في تحديد نقاط القوة والضعف في اسئلة الاختيار من متعدد في اختبار مواد الدراسات الاجتماعية للصف العاشر في المعاهد الإسلامية.

أهداف الدراسة:

تهدف الدراسة الحالية إلى:

١. التعرف على مدى توافر معايير الجودة للمستويات المعرفية (التذكر، والفهم، والتطبيق، والتحليل، والتركيب، والتقويم) في اسئلة اختبارات مواد الاجتماعيات للصف العاشر.
٢. التعرف على مدى تطابق اسئلة الاختيار من متعدد في اختبار مواد الاجتماعيات للصف العاشر في المعاهد الإسلامية مع المعايير التربوية للاختبار الجيد.
٣. التعرف على أبرز نقاط الاختلاف التي تجعل اسئلة الاختيار من متعدد في اختبار مواد الاجتماعيات للصف العاشر في المعاهد الإسلامية لا تتطابق مع المعايير التربوية للاختبار الجيد وبالتالي توجيه القائمين على إعداد الاختبارات لتعديلها.

حدود الدراسة:

تحدد نتائج الدراسة الحالية بالمحددات الآتية:

- الحدود الموضوعية: اسئلة الاختيار من متعدد في اختبارات مواد الدراسات الاجتماعية للصف العاشر.
- الحدود المكانية: تقتصر الدراسة على اختبارات مواد الدراسات الاجتماعية للصف العاشر في المعاهد الإسلامية التابعة لمركز السلطان قابوس العالي للثقافة والعلوم التابعة لديوان البلاط السلطاني في سلطنة عمان.
- الحدود الزمنية: اختبارات الأعوام الدراسة من عام ٢٠١٥/٢٠١٦ إلى العام الدراسي ٢٠١٩/٢٠٢٠م.

مصطلحات الدراسة:

- الاختبارات التحصيلية: هي إجراء منظم لقياس مقدار التعلم الذي حققه الطلبة في مادة دراسية أو مجال تعليمي معين (عتوم، ٢٠١٤).
- فقرات الاختيار من متعدد: وهي نوع من الاسئلة يتكون من جزئين: المتن وهو محل السؤال حيث تصاغ المشكلة، والجزء الآخر يعرف بالخيارات أو البدائل، وهي التي تعطى للمفحوص مع السؤال

- وتشكل إجابات محتمة للسؤال الوارد في المتن، ومن بينها بديل صحيح وباقي البدائل هي عبارة عن موهات ويطلب من المفحوص اختيار الإجابة الصحيحة من بين تلك البدائل (الزغيات، ٢٠١٤).
- **تصنيف بلوم:** هو تصنيف وضعه بنيامي بلوم للأهداف التعليمية، وهو يتألف من أهداف معرفية وأخرى وجدانية، وثالثة نفس حركية، وهي تتدرج من المستويات الدنيا إلى المستويات العليا، ومستويات بلوم للأهداف المعرفية تتكون من ستة مستويات هي: التذكر، والفهم، والتطبيق، والتحليل، والتركيب، والتقويم (الشقيرات، ٢٠٢٠).
- **معايير الاختبار التربوي الجيد (جودة الاختبارات):** هي مجموعة من الموصفات والشروط العلمية التي تحدد ما ينبغي أن يكون عليه كل سؤال من أسئلة الاختبارات، وفق معايير خاصة بالمتن وأخرى خاصة بالبدائل وثالثة خاصة بالإخراج الفني، علاوة على معايير تصنيف مستويات المعرفة للسؤال (التذكر، والفهم، والتطبيق، والتحليل، والتركيب، والتقويم)، والتي تؤدي جميعها إلى أسئلة تتصف بالجودة وبالتالي تقدم معلومات وبيانات جيدة عن مستويات الطلبة (الناقة، ٢٠١٦).
- **أسئلة الامتحان النهائي إجرائيا:** هي الفقرات التي قام معلمو مواد الدراسات الاجتماعية ببنائها لقياس تحصيل الطلاب في مواد الدراسات الاجتماعية.

الإطار النظري:

تهتم العملية التعليمية بتنمية شخصية المتعلم في كافة الجوانب ولاسيما المعرفية، ولذلك تضع قائمة شاملة للسلوك المراد الوصول إليه، وتحدد الأهداف السلوكية التي يجب تحقيقها وتتدرج فيها من البسيط إلى المعقد، لذلك وضع بلوم تصنيف هرمي للمجال المعرفي، يتكون هذا الشكل الهرمي من ست مستويات رئيسة (الفرجي، ٢٠١٩؛ مغامس، ٢٠١٨؛ جيوسي، ٢٠١٦)، وهي:

- ١- **المعرفة (Knowledge):** وهي تعد قاعدة هرم التصنيف في المجال المعرفي، وتتمثل في التذكر الذي هو وحدة بناء المعرفة، كتذكر الحقائق والتواريخ والأشخاص والقوانين، والتي تقدم للطلاب في المراحل التعليمية المختلفة، ومن الأفعال التي تستخدم في هذا المستوى: يُعَدُّ، يُعرَف، يَذكر.
- ٢- **الفهم أو الاستيعاب (Comprehension):** وهو يتمثل في قدرة المتعلم على إدراك المادة الدراسية وفهم الحقائق وتفسير البيانات، ومن الأفعال التي تستخدم في هذا المستوى: يُميز، يعطي مثال، يقارن.
- ٣- **التطبيق (Application):** وهو يتمثل في القدرة المتعلم على استخدام ما تعلمه في مواقف جديدة، ومن الأفعال التي تستخدم في هذا المستوى: يحسب، يرسم.
- ٤- **التحليل (Analysis):** وهو يُشير إلى قدرة المتعلم على تحليل المادة إلى مكوناتها الجزئية، ومعرفة الأجزاء والعناصر والأسس والقواعد التنظيمية المستخدمة، ومن الأفعال التي تستخدم في هذا المستوى: يَجزئ، يَفرق، يَحدد العناصر.
- ٥- **التركيب (Synthesis):** ويقصد به قدرة المتعلم على وضع الأجزاء معا لتكوين جديد، كإعداد موضوع جديد أو مشروع بحثي، ومن الأفعال المستخدمة في هذا المستوى: يَصف، يَؤلف، يَجمع، يَخلص.

٦- **التقويم (Evaluation):** ويتمثل في قدرة المتعلم على الحكم على قيمة الأشياء، كالحكم على عبارة أو قصة أو إعداد تقرير يحكم فيه على موضوع بناءً على معايير محددة، ومن الأفعال المستخدمة في هذا المستوى: يُقوم، يقارن، يدعم.

الاختبارات التحصيلية (Achievement Test):

عرفت الاختبارات على أنها أداة قياس يتم إعدادها وفق خطوات تتضمن العديد من الإجراءات، بغرض تحديد درجة امتلاك الأفراد لسمة أو قدرة معينة من خلال الإجابة على عينة من المثيرات التي تمثل السمة أو القدرة المرغوب قياسها (عودة، ٢٠١٠). وهي تلعب دوراً كبيراً في تثبيت التعلم للمتعلم وتشخيص جوانب القوة والضعف لديه، وهو ما يعطي الفرصة لتعزيز نقاط القوة وتعديل جوانب الضعف، وتصنيف المتعلمين لتنويع الأنشطة والأساليب التدريسية التي تتلاءم مع مختلف مستوياتهم (الحسن، ٢٠١٢، والشمراني، ٢٠١٢، والحياني، ٢٠١٠)، كذلك تسهم الاختبارات التحصيلية في عملية رفع مستوى طموح المعلم وزيادة الجهد الذي يبذله لرفع جودة العملية التعليمية؛ من خلال الملاحظة والمتابعة للتحسن والتغيير الذي يطرأ على مستويات الطلبة (الطاهر ومحمد، ٢٠١٣).

والاختبارات التحصيلية من أكثر أدوات التقويم شيوعاً في المؤسسات التربوية في العالم، وعلى ذلك تعددت أنواعها، فمنها الاختبارات اللفظية وغير اللفظية، واختبارات معيارية المحك وأخرى مرجعية المحك، وأيضاً اختبارات قوة وأخرى اختبارات سرعة، كذلك هناك اختبارات شفوية واختبارات التحريرية واختبارات أدائية (العملية)، لكن الأكثر شيوعاً واستخداماً في المؤسسات التعليمية هي الاختبارات الشفوية والتحريرية والعملية، وفي هذه الدراسة سيتم التركيز على الاختبارات التحريرية والتي هي بدورها تنقسم إلى نوعين، النوع الأول المقالي وهو النوع الذي يطلب فيه من الطلبة توليد الإجابة ولذلك تعرف باسم اختبارات الاستدعاء، والنوع الأخرى وهو ذو الإجابة المصاغة، بحيث يطلب من الطلبة فقط انتقاء الإجابة من بين الإجابات التي أعطيت له، وهي بدورها تنقسم إلى أربعة أنواع: اختبارات التكملة القصيرة، واختبارات الصوب والخطأ، واختبارات المزاوجة أو المطابقة، واختبارات الاختيار من متعدد (الطاهر ومحمد، ٢٠١٣؛ والشمراني، ٢٠١٢).

وبالمقارنة بين أنواع الاسئلة الموضوعية، نجد أن فقرات الاختبار من متعدد الأكثر شيوعاً واستخداماً والأكثر مرونة، وبالتالي تستخدم لقياس مدى استيعاب الحقائق العلمية وتطبيقاتها، وله أيضاً العديد من المميزات الأخرى كما أشارت إلى ذلك دراسات السابقة (العنزي، ٢٠١٧؛ وأبو فودة، ٢٠١٤؛ Dehnad, Nasser & Hosseini, 2014)، وهي: استخدامها لقياس مستويات مختلفة من الأهداف التربوية، وقدرتها على شمولية المحتوى الدراسي بشكل كبير، علاوة على تمتعها بسهولة وموضوعية التصحيح، وعدم تأثرها بقدرة الطلبة اللفظية والمساعدة على تشخيص أخطاء الطلبة ونقاط ضعفهم وتحليل النتائج بشكل سهل وسريع.

خطوات بناء الاختبار التحصيلي: تمر عملية إعداد الاختبار التحصيلي بعدد من الخطوات إشارات إليها

عدد من الدراسات (عتوم، ٢٠١٤؛ والحياني، ٢٠١١) وهي:

- ١- تحديد الغرض من الاختبار: لاسيما وأن للاختبارات أنواع عدة، فلا بد من تحديد نوع الاختبار لتحديد نوع الفقرات.
- ٢- تحديد الأهداف: حيث يتم تحديد الأهداف السلوكية (نواتج التعليم) المراد تحقيقها من تدريس المادة الدراسية.
- ٣- تحليل المحتوى: وذلك لضمان تحقيق الشمولية والتوازن في تغطية جميع موضوعات المادة الدراسية.
- ٤- إعداد جدول الموصفات: والذي هو عبارة عن مخطط تفصيلي، يحدد محتوى الاختبار ويربط محتوى المادة الدراسية بالأهداف التعليمية، ويحدد الأوزان النسبية لكل موضوع من موضوعات المحتوى التعليمي، والأوزان النسبية للأهداف التعليمية في مستوياتها المختلفة.
- ٥- كتابة أسئلة الاختبار: حيث تكتب فقرات الاختبار بصورة أولية، ثم تراجع لحين صياغتها بصورة نهائية.
- ٦- تحليل الاختبار: حيث تتم هذه الخطوة بعد الانتهاء من تطبيق الاختبار؛ لحساب قيم مؤشرات النزعة المركزية والتشتت، للكشف عن مدى جودة الفقرات لاستبعادها أو التعديل عليها.

الدراسات السابقة:

نظر لأهمية الاختبارات التحصيلية يجب العناية والاهتمام الجيد في إعدادها، ولاسيما وأن صدق المعلومات المستسقة منها يتوقف على البناء الجيد لها، وعلى ذلك فنحن بحاجة إلى معايير لأجل إصدار الاحكام واتخاذ القرارات المناسبة حول جودة أسئلة اختبارات مواد الاجتماعيات وفق معايير جيدة ومحددة، وقد أجريت العديد من الدراسات هدفت لتحليل الاختبارات التحصيلية وفي مراحل تعليمية مختلفة، والتي من بينها دراسة الشقيرات (٢٠٢٠) والتي سعت لتحليل أسئلة امتحانات الثانوية العامة في الأردن، لمادة تاريخ الأردن في ضوء مستويات المجال المعرفي حسب تصنيف بلوم، وباستخدام المنهج الوصفي التحليلي قام الباحث بتحليل (٣٦٢) سؤال، وتوصلت الدراسة إلى ارتفاع الأسئلة في مستوى التذكر والفهم، حيث بلغت نسبة أسئلة التذكر حوالي (٥٢,٥%)، في حين بلغت نسبة أسئلة الفهم حوالي (٤١,٤%)، كذلك توصلت الدراسة إلى انعدام الأسئلة في مستوى التركيب والتقييم، وتوصي الدراسة بضرورة التركيز على الأسئلة التي تقيس المستويات المعرفية العليا.

وفي عام (٢٠١٨) وباستخدام المنهج الوصفي التحليلي أجرى سلمان وطلافة دراسة سعت للكشف عن مدى تمثيل أسئلة الاختبارات التحصيلية النهائية، لمستويات الأهداف المعرفية لدى معلمي مبحث العلوم الحياتية للصف الأول الثانوي العلمي في جنوب الأردن، وباستخدام تصنيف بلوم للأهداف المعرفية، وقائمة محكات مألّفه من (٣٥) محك للحكم على جودة الاختبارات التحصيلية، قام الباحثين بتحليل (١٨٤٢) فقرة اختبارية، وخلصت الدراسة إلى أن أسئلة التذكر والفهم جاءت في المرتبة الأولى بين أسئلة الاختبارات النهائية، وهي تشكل نسبة (٩١,٤%) من المجموع الكلي للأسئلة، بينما شكلت أسئلة التطبيق والتحليل ما نسبته (٨,٥%)، في حين أنه لم ترد أسئلة في مستوى التركيب والتقييم، وقد أوصت الدراسة بضرورة عقد دورات تدريبية في كيفية إعداد الاختبارات التحصيلية للمعلمين.

وعلى المستوى الجامعي قامت المزوغي (٢٠١٨) بدراسة باستخدام المنهج الوصفي التحليلي، للتحقق من تغطية أسئلة الامتحانات النهائية للأهداف المعرفية بمستوياتها المختلفة وفق تصنيف بلوم، وذلك باستخدام قائمة تضم (٥) معايير وهي: كتابة البيانات الرئيسية، ومعيار عدم وجود أسئلة اختيارية، ومعيار تنوع الأسئلة بالمادة الواحدة، ومعيار بيان درجة كل سؤال، ومعيار الخط الواضح والطباعة الجيدة، وبطاقة تحليل والتي تم تطبيقها على (٥٠٩) سؤال من أسئلة الامتحانات النهائية لقسم علم النفس بكلية الآداب للعام الجامعي (٢٠١٦-٢٠١٧)، وقد خلصت الدراسة إلى أن مستوى التذكر قد جاء في المرتبة الأولى بعدد (٤٣٣) سؤال، وجاء مستوى الفهم في المرتبة الثانية بواقع (٧١) سؤال، وخلت جميع الاختبارات من أسئلة تقيس مستوى التركيب وسؤال واحد فقط في مستوى التحليل ومستوى التقويم، وتوصي الدراسة بضرورة الاهتمام بالأسئلة المثيرة للتفكير، وتوزيع الاسئلة وفق مستويات بلوم المعرفية لخلق نوع من التوازن.

كذلك قام الهدور (٢٠١٧) بدراسة لتحليل فقرات اسئلة اختبارات الشهادة الثانوية العامة في الرياضيات بالجمهورية اليمنية، للكشف عن المستويات المعرفية التي تمثلها وفق التصنيفات الحديثة للأهداف التعليمية، حيث تكونت عينة الدراسة من فقرات الاختبارات النهائية لمادة الرياضيات بلغ عددها (٨٢٣) فقرة للأعوام الممتدة من العام (٢٠١٠/٢٠١١) وإلى العام الدراسي (٢٠١٤/٢٠١٥)، وأظهرت الدراسة أن فقرات الاختبارات قد ركزت على المستويات الدنيا (التذكر والفهم والتطبيق) من مستويات تصنيف بلوم المعدل، وهو ما يمثل نسبة (٩٢%)، بينما لم تتجاوز النسبة (٨%) في المستويات العليا (التحليل والتقويم والإبداع).

كذلك قام الناقة (٢٠١٦) بتحليل (٩٠) سؤال من نوع الاختيار من متعدد في اختبارات مبحث الكيمياء للصف الثاني عشر في الفترة (٢٠٠٧-٢٠١٥)، وباستخدام قائمة معايير تضمنت ثلاثة محاور: محور الشمولية ومحور المطابقة ومحور النواحي الفنية، بالإضافة لتصنيف الاسئلة وفق المستويات المعرفية (التذكر، والفهم، والتطبيق، والتحليل، والتركيب، والتقويم)، وتوصلت الدراسة الى أن محور الناحية الفنية حصل على أعلى نسبة من التصنيفات الثلاثة والترتيب الأول بينها حصل محور المطابقة على أقل نسبة، كذلك حصل مستوى التذكر على أعلى نسبة، وفي مقابل ذلك حصل مستوى التقويم على أقل نسبة في توزيع اسئلة الاختبارات وفق تصنيف بلوم المعرفي.

كما أجرى الشاهر والحمداني (٢٠١٦) دراسة باستخدام المنهج الوصفي التحليلي، سعت لتقويم أسئلة الاختبارات النهائية لمادة النحو في كلية العلوم الإسلامية في ضوء مستويات بلوم للمجال المعرفي، حيث شملت الدراسة على تحليل (٩٤) سؤال، وتوصلت الدراسة إلى أن مستوى الفهم قد حصل على أعلى تكرار بنسبة (٤٣%)، في حين حصل مستوى التقويم على أقل معدل في التكرارات وهو بنسبة (١%)، وتوصي الدراسة بضرورة التركيز في اسئلة الاختبارات النهائية على الاسئلة التي تحفز التفكير والنقد والاستنتاج والابداع والتقويم.

و كذلك أجرى الزعبي (Alzu'bi, 2014) دراسة هدفت لتحليل أسئلة اللغة الإنجليزية لامتحانات الشهادة الثانوية الأردنية من خلال المستويات المعرفية لبلوم، وأوضحت نتائج الدراسة أن النسبة المئوية للمستويات الثلاثة الأولى (الفهم والمعرفة والتطبيق) هي (٦٩,٦%)، ولكن النسبة المئوية للمستويات الثلاثة الأخيرة (التحليل والتركيب والتقويم) هي (٣٠,٤%). وقد خلصت الدراسة أن أسئلة اللغة الإنجليزية المدرجة في الامتحانات

الثانوية العامة تؤكد على مستويات التفكير منخفضة الترتيب، ويوصي الباحث بضرورة تنوع مستويات الأسئلة والاهتمام بالأسئلة ذات المستويات العليا.

وفي بنجلاديش أجرى حسن ونعمي وبلقيس (Hasan, Naomee & Bilkis, 2013) دراسة باستخدام المنهج الوصفي، هدفت لتحليل اسئلة الاختبارات في العلوم الاجتماعية لامتحانات شهادة الثانوية العامة للأعوام من (٢٠٠٩) إلى (٢٠١٢)، حيث تم تحليل اسئلة الاختبارات وفق تصنيف بلوم المعدل، وكان من أبرز النتائج التي توصلت إليها الدراسة أن اسئلة الاختبارات في العام الدراسي (٢٠١٠/٢٠٠٩) تم التركيز بشكل كبير على مستوى المعرفة والفهم في الاسئلة الموضوعية، وعلى مستوى الفهم والتحليل في الاسئلة المقالية، بينما تم التركيز على مستوى المعرفة والفهم والتطبيق والتحليل في الاسئلة الموضوعية والمقالية في العام الدراسي (٢٠١٢/٢٠١١)، وتم تجاهل مستوى التقويم والإبداع، وتوصي الدراسة بضرورة تنوع الاسئلة وفق تصنيف بلوم المعدل عند كتابة الاسئلة الموضوعية والمقالية.

وأجريت دراسات أخرى للكشف عن مدى الالتزام بقواعد صياغة فقرات الاختيار من متعدد في الامتحانات، والتي من بينها دراسة عتوم (٢٠١٤)، والتي أجريت باستخدام استبانة مكونة من (٣٠) فقرة موزعة على أربعة مجالات وهي: تعليمات الاختبار، وكتابة فقرات الاختبار النوعية والكمية، والطباعة، والإخراج، وبتطبيقها على (١١٦) اختبار تحصيلي توصلت الدراسة إلى: أن الاختبارات التحصيلية من إعداد المعلمين لا تتفق مع المعايير الجيدة للاختبارات التحصيلية، وهناك فروق دالة إحصائية ولصالح المعلمات في معايير الاختبار الجيد، كذلك أجرى ظاظا (٢٠١٢) دراسة سعت للكشف عن مدى مخالفة قواعد صياغة فقرة الاختيار من متعدد في امتحان الثانوية العامة في الأردن، حيث تم مراجعة (٢٩١٠) فقرة من فقرات الاختيار من متعدد، في اختبارات الأعوام من (٢٠٠٧ إلى ٢٠١٠) باستخدام قائمة مكونة من (١٨) معيار للكشف عن مخالفة الفقرة لقواعد صياغتها، وظهرت نتائج الدراسة أن (٤١٦) فقرة تضمنت مخالفة واحدة على الأقل، في حين وجدت (٧٨٢) فقرة مخالفة، وأن أكثر قاعد تم انتهاكها هي: ترتيب البدائل عموديا، ثم قاعدة ترتيب البدائل منطقيا أو عدديا أو أبجديا.

التعليق على الدراسات السابقة:

- اعتمدت أغلب الدراسات تصنيف بلوم للمجال المعرفي وشمول الأسئلة.
- استعملت الدراسات السابقة المنهج الوصفي في تحليل الأسئلة ويتفق هذا مع الدراسة الحالية.
- تتفق الدراسة الحالية مع ما ذهبت إليه الدراسات السابقة باستخدام النسب المئوية.
- اختلفت الدراسات في المرحلة الدراسية التي حللت أسئلتها، وتباينت كذلك في المواد التي أجري تحليل لأسئلتها.

إجراءات الدراسة:

منهجية الدراسة:

استخدم الباحث المنهج الوصفي التحليلي في هذه الدراسة، وهو أحد المناهج المستخدمة في دراسة مضمون مادة علمية أو وسيلة اتصال مكتوبة أو مسموعة وذلك بأخذ عينة من المادة المزمع دراستها وتصنيفها وتحليلها كميًا وكيفيًا، بناءً على قواعد وأسس محددة للوصول إلى تفسيرات واستنتاجات موضوعية (المزوعي، ٢٠١٨).

مجتمع الدراسة وعينتها:

تكون مجتمع الدراسة من الاختبارات النهائية (الفصل الدراسي الأول، والفصل الدراسي الثاني) لمواد الدراسات الاجتماعية بالمعاهد الإسلامية التابعة لمركز السلطان قابوس العالي للثقافة والعلوم للأعوام الدراسية الخمسة الأخيرة من (٢٠١٥) إلى (٢٠٢٠)، وقد بلغ مجموع أسئلة الامتحانات في اختبائي الدراسات الاجتماعية وتاريخ عمان مجتمعة حوالي (١٨) سؤال، وكل سؤال يتكون من (٦) فقرات من نوع الاختيار من متعدد، باستثناء اختبارات العام الدراسي (٢٠٢٠) حيث أن سؤال الاختيار من متعدد في اختبارات الفصل الدراسي الأول تكون من (١٢) مفردة اختيارية، ولم يكن هناك اختبارات لنهاية الفصل الدراسي الثاني بسبب جائحة كورونا كوفيد-١٩، وعلى ذلك يكون مجموع الفقرات الاختبارية التي تم مراجعتها وتطبيق أداة الدراسة عليها حوالي (١٢٠) سؤال من نوع الاختيار من متعدد، والجدول (١) يوضح توزيع الأسئلة على المواد الدراسية.

جدول (١) يوضح توزيع الأسئلة على المواد الدراسية				
المادة الدراسية		الدراسات الاجتماعية		تاريخ عمان
بيانات الأسئلة والفقرات	عدد الأسئلة	فقرات الاختيار من متعدد	عدد الأسئلة	فقرات الاختيار من متعدد
الفصل الدراسي الأول	٥	٣٦	٥	٣٦
الفصل الدراسي الثاني	٤	٢٤	٤	٢٤
الإجمالي	٩	٦٠	٩	٦٠

أدوات الدراسة:

أداة الكشف عن انتهاك قواعد صياغة فقرة الاختيار من متعدد لقواعد صياغتها (ملحق ١) تم تطويرها من خلال مراجعة عدد من الدراسات التي تناولت الاختبارات الموضوعية ومواصفاتها وقواعد صياغة فقرة الاختيار من متعدد (المزوعي، ٢٠١٨؛ والناقعة، ٢٠١٦؛ وعوم، ٢٠١٤؛ وظاظا، ٢٠١٢، و Haladyna, Downing, 2002)، وتم فرز (٢٧) قاعدة ملائمة لتحقيق أهداف الدراسة الحالية، ووضعت القواعد في شكل عبارات، كما تم استخدام بطاقة تحليل تتضمن مستويات تصنيف بلوم وهي: التذكر، والفهم، والتطبيق، والتحليل، والتركيب والتقويم؛ حيث تم تصنيف أسئلة الامتحان النهائي في مواد الدراسات الاجتماعية على شكل تكرارات ونسب مئوية.

صدق الأداة:

تم التحقق من دلالات صدق الأداة بعرضها على عدد من الخبراء (ملحق ٢)، وطلب إليهم إصدار الحكم على شمولية القواعد لأبعاد فقرة الاختيار من متعدد، ودرجة ارتباطها بالاختيار من متعدد، وقد تم الاتفاق على مناسبة كل العبارات والبالغ عددها (٢٧)، وبطاقة التحليل المتضمنة لمستويات أهداف المجال المعرفي حسب تصنيف بلوم له شهرته عالمياً، وتعد الأداة صادقة ولا تحتاج إلى تحكيم.

ثبات الأداة:

للتحقق من ثبات إجراءات التحليل وأدوات الدراسة، قام الباحث بتحليل عينة من أسئلة الاختيار من متعدد وعددها (٢٤) سؤال وهي تمثل نسبة (٢٠%) من إجمالي أسئلة الاختيار من متعدد، وأيضاً قام بثلاثة محللين مستقلين يعملون منفردين بتحليل نفس الأسئلة، ولإيجاد قيمة معامل ثبات التحليل واتفاق المحللين تم استخدام معادلة هولستي (سلمان وطلافة، ٢٠١٨) الآتية:

$$\text{معامل الثبات} = \text{عدد حالات التي اتفق عليها} / \text{عدد الحالات الكلية} * 100\%$$

والجدول (٢) يوضح قيم معاملات الثبات:

جدول (٢) نسبة الاتفاق بين المحللين

المحلل	الباحث	الأول	الثاني	الثالث
الأول	٩١,٦%	-	٩١,٦%	٩١,٦%
الثاني	٨٣,٣%	٩١,٦%	-	٩١,٦%
الثالث	٩١,٦%	٩١,٦%	٩١,٦%	-

$$\text{ثم تم إيجاد متوسط نسبة الاتفاق} = (91,6 + 91,6 + 91,6 + 91,6 + 91,6 + 91,6 + 91,6 + 83,3 + 91,6) / 9 = 90,6\%$$

ومن ذلك نلاحظ أن متوسط معامل الثبات بين المحكمين قد بلغ (٩٠,٦%) وهذا مؤشر جيد لصلاح الأداة.

عرض ومناقشة نتائج الدراسة:

- الإجابة على السؤال الأول من أسئلة الدراسة : ما مدى توافر معايير الجودة للمستويات المعرفية (التذكر، والفهم، والتطبيق، والتحليل، والتركيب، والتقويم) في أسئلة اختبارات مواد الدراسات الاجتماعية للصف العاشر؟

وللإجابة على هذا السؤال تم حساب تكرار كل معيار في كل سنة، وترتيبها ورصد مجموع كل تصنيف من مستويات بلوم المعرفية، والجدول (٣) يوضح مجموع التكرارات والنسب المئوية لأسئلة اختبارات مادة الدراسات الاجتماعية وتاريخ عمان وفق السنة الدراسية والفصل الدراسي

جدول (٣) مجموع التكرارات والنسب المئوية لأسئلة اختبارات مواد الدراسات الاجتماعية وفق السنة الدراسية والفصل الدراسي

الترتيب	النسبة	المجموع	العام الدراسي									المستوى	المادة
			2019/2020	2018/2019		2017/2018		2016/2017		2015/2016			
			الفصل الأول	الفصل الثاني	الفصل الأول	الفصل الثاني	الفصل الأول	الفصل الثاني	الفصل الأول	الفصل الثاني			
1	%60	29	5	-	-	6	2	2	5	4	5	التذكر	الدراسات الاجتماعية
4	%8	4	1	-	-	0	0	2	1	0	0	الفهم	
3	%10	5	3	-	-	0	0	0	0	2	0	التطبيق	
2	%21	10	3	-	-	0	4	2	0	0	1	التحليل	
0	.	0	0	-	-	0	0	0	0	0	0	التركيب	
0	.	0	0	-	-	0	0	0	0	0	0	التقويم	
0	%100	48	12	-	-	6	6	6	6	6	6	المجموع	
1	%73	35	7	5	4	5	3	-	-	5	6	التذكر	تاريخ عمان
3	%6	3	2	0	1	0	0	-	-	0	0	الفهم	
2	%15	7	3	1	0	1	1	-	-	1	0	التطبيق	
3	%6	3	0	0	1	0	2	-	-	0	0	التحليل	
0	.	0	0	0	0	0	0	-	-	0	0	التركيب	
0	.	0	0	0	0	0	0	-	-	0	0	التقويم	
0	%100	48	12	6	6	6	6	-	-	6	6	المجموع	
(-) أسئلة تم تحليلها في العينة الاستطلاعية.													

ويلاحظ من الجدول (٣) أن مستوى التذكر قد حصل على أعلى نسبة (٦٠%) في أسئلة اختبارات مادة الدراسات الاجتماعية، وعلى نسبة (٧٣%) في أسئلة اختبارات مادة تاريخ عمان، وبذلك يكون قد جاء في المرتبة الأولى، وهذا يتفق مع دراسة الشقيرات (٢٠٢٠)، ودراسة سلمان والطلافة (٢٠١٨)، ودراسة المزوغي (٢٠١٨)، ودراسة الهدور (٢٠١٧)، ودراسة الزعبي (Alzu'bi, 2014)، وقد يعزى اهتمام واضعي الاسئلة بمستوى التذكر إلى طبيعة موضوعات مادة الدراسات الاجتماعية وتاريخ عمان التي تتضمن كم كبير من الحقائق والمعلومات التي يجب على الطالب تذكرها، ويلاحظ أيضا أن مستوى التحليل في أسئلة اختبارات الدراسات الاجتماعية قد جاء في المرتبة الثانية بنسبة (٢١%)، وجاء في المرتبة الثالثة في اختبارات مادة تاريخ عمان بنسبة (٦%)، بينما جاء مستوى التطبيق بنسبة (١٥%) في المرتبة الثانية في أسئلة اختبارات مادة تاريخ عمان، وفي المرتبة الثالثة بنسبة (١٠%) في اختبارات أسئلة الدراسات الاجتماعية، كذلك نجد أن مستوى الفهم قد جاء في المرتبة الرابعة في أسئلة اختبارات الدراسات الاجتماعية بنسبة (٨%)، والمرتبة الثالثة مكرر في أسئلة اختبارات مادة تاريخ عمان بنسبة (٦%)، بينما نلاحظ أنه لا توجد أسئلة في مستوي التركيب والتقويم، وهذا يتفق

مع نتائج دراسة الشقيرات (٢٠٢٠)، ودراسة سلمان والطلافة (٢٠١٨)، ودراسة حسن ونعومي وبلقيس (Hasan, Naomee & Bilkis, 2013).

- الإجابة على السؤال الثاني: ما معايير جودة الاسئلة لاختبارات مواد الدراسات الاجتماعية التي يمكن التقويم على أساسها؟

قد تم تطوير قائمة معايير تقويم اسئلة الاختيار من متعدد بالرجوع للأدب التربوي والدراسات السابقة (المزوعي، ٢٠١٨؛ والناقعة، ٢٠١٦؛ وعتوم، ٢٠١٤؛ وظاظا، ٢٠١٢، و (Haladyna, Downing & Rodriguez, 2002))، وتألفت قائمة المعايير من ثلاثة محاور: المحور الأول خاص بالمتن، وهو مكون من (١٠) قواعد، والمحور الثاني خاص بالمموهات أو البدائل وهو أيضا مكون من (١٠) قواعد، والمحور الثالثة الإخراج الفني، وهو مكون من (٧) قواعد، وقد تألفت القائمة في شكلها النهائي من (٢٧) قاعدة (الملحق "١").

- الإجابة على السؤال الثالث: ما مدى تطابق بناء أسئلة الاختيار من متعدد في اختبار مواد الدراسات الاجتماعية للصف العاشر في المعاهد الإسلامية مع المعايير التربوية الجيدة للاختبار التحصيلي؟

وللإجابة على السؤال تم حساب تكرار كل معيار في كل فصل دراسي ضمن السنوات الدراسية ضمن إطار الدراسة، وهي على النحو الآتي:

أولاً: محور متن السؤال: حيث أن الجدول (٤) يبين مجموع التكرارات والنسب المئوية والترتيب للمعايير الخاصة بمتن السؤال.

جدول (٤) مجموع التكرارات والنسب المئوية والترتيب للمعايير الخاصة بمتن السؤال												
الترتيب	النسبة	المجموع	العام الدراسي									
			2019/2020		2018/2019		2017/2018		2016/2017		2015/2016	
			١ ف	٢ ف	١ ف	٢ ف	١ ف	٢ ف	١ ف	٢ ف	١ ف	٢ ف
-	-	0	0	0	0	0	0	0	0	0	0	1
-	-	0	0	0	0	0	0	0	0	0	0	2
-	-	0	0	0	0	0	0	0	0	0	0	3
1	%35	6	0	0	0	0	1	2	1	1	1	4
3	%18	3	0	0	0	0	0	0	2	0	1	5
-	-	0	0	0	0	0	0	0	0	0	0	6
4	%12	2	0	0	0	0	2	0	0	0	0	7
-	-	0	0	0	0	0	0	0	0	0	0	8
4	%12	2	0	0	0	0	0	1	1	0	0	9
2	%24	4	0	0	0	3	0	1	0	0	0	10
-	-	0	0	0	0	0	0	0	0	0	0	1
-	-	0	0	0	0	0	0	0	0	0	0	2
-	-	0	0	0	0	0	0	0	0	0	0	3
3	%25	3	0	1	0	0	0	0	2	0	0	4
-	-	0	0	0	0	0	0	0	0	0	0	5
-	-	0	0	0	0	0	0	0	0	0	0	6
-	-	0	0	0	0	0	0	0	0	0	0	7
-	-	0	0	0	0	0	0	0	0	0	0	8
1	%42	5	0	3	0	0	0	0	1	1	0	9
2	%33	4	0	3	0	0	0	0	1	0	0	10
٢: الفصل الدراسي الثاني									١: الفصل الدراسي الأول.			

وبلاحظ من الجدول (٤) أن العبارة رقم (٤) والتي تنص على: يخلو من النفي، مثل: ما عدا أو لا والنفي المركب، قد جاءت في المرتبة الأولى في اسئلة اختبارات الدراسات الاجتماعية بنسبة (٣٥%)، بينما احتلت المرتبة الثالثة في اسئلة اختبارات مادة تاريخ عمان بنسبة قد بلغت (٢٥%)، في حين نجد أن العبارة رقم (١٠) والتي تنص على: (عبر في المتن عن الفكرة الرئيسة للسؤال بوضوح)، قد جاءت في المرتبة الأولى بنسبة (٤٢%) في اسئلة مادة تاريخ عمان وهي بذلك تحتل المرتبة الأولى، بينما نجدها تحتل المرتبة الثانية في اسئلة مادة الدراسات الاجتماعية بنسبة (٢٤%)، وجاءت العبارة رقم (٥) والتي تنص على: (يُعبّر المتن عن فكرة واضحة وليس رأي) في المرتبة الثالثة بنسبة (١٨%) في اسئلة مادة الدراسات الاجتماعية، بينما جاءت العبارة رقم (٧) والتي تنص على: (تم بناء المتن على محتوى عام للتعليم) في المرتبة الرابعة بنسبة (١٢%)، وكذلك

جاءت العبارة رقم (٩) والتي تنص على: (عُبر عن متن السؤال بعدد قليل من العبارات) في المرتبة الرابعة مع العبارة (٧) في اسئلة اختبارات مادة الدراسات الاجتماعية، في حين نجد أن العبارة رقم (٩) قد احتلت المرتبة الأولى في اسئلة تاريخ عمان.

والجدول (٥) يوضح عبارات المحور الثاني الخاصة ببدائل الاختبار أو ما يعرف بالموهات.

جدول (٥) مجموع التكرارات والنسب المئوية والترتيب للمعايير الخاصة ببدائل الاختبار												
الترتيب	النسبة	المجموع	العام الدراسي									
			2019/2020	2018/2019		2017/2018		2016/2017		2015/2016		العبارة
			ف ١	ف ٢	ف ١	ف ٢	ف ١	ف ٢	ف ١	ف ٢	ف ١	
1	%90	35	8	0	0	5	4	4	6	2	6	11
2	%28	11	5	0	0	1	4	0	0	1	0	12
-	-	0	0	0	0	0	0	0	0	0	0	13
3	%21	8	0	0	0	0	1	4	1	0	2	14
4	%3	1	0	0	0	0	0	0	1	0	0	15
-	-	1	0	0	0	1	0	0	0	0	0	16
4	%3	1	0	0	0	0	0	0	0	0	1	17
-	-	0	0	0	0	0	0	0	0	0	0	18
-	-	0	0	0	0	0	0	0	0	0	0	19
-	-	0	0	0	0	0	0	0	0	0	0	20
1	%70	38	11	6	5	2	3	0	6	5	0	11
2	%28	15	12	0	0	2	0	0	0	1	0	12
-	-	0	0	0	0	0	0	0	0	0	0	13
3	%2	1	0	0	1	0	0	0	0	0	0	14
-	-	0	0	0	0	0	0	0	0	0	0	15
-	-	0	0	0	0	0	0	0	0	0	0	16
-	-	0	0	0	0	0	0	0	0	0	0	17
-	-	0	0	0	0	0	0	0	0	0	0	18
-	-	0	0	0	0	0	0	0	0	0	0	19
-	-	0	0	0	0	0	0	0	0	0	0	20
ف٢: الفصل الدراسي الثاني			ف١: الفصل الدراسي الأول.									

ويلاحظ من الجدول (٥) أن العبارة رقم (١١) والتي تنص على: رتبت البدائل ترتيباً منطقياً، هجائياً أو زمنياً أو عددياً ، قد جاءت في المرتبة الأولى في اسئلة اختبارات مواد الدراسات الاجتماعية، حيث أنها تمثل

نسبة (٩٠%) في أسئلة اختيارات مادة الدراسات الاجتماعية، ونسبة (٧٠%) في أسئلة اختبارات مادة تاريخ عمان، وهذا يتفق مع نتائج دراسة ظاظا (٢٠١٢)، كذلك نلاحظ أن العبارة رقم (١٢) والتي تنص على: (رتبت البدائل عموديا) قد جاءت في المرتبة الثانية في ترتيب القواعد التي تم انتهاكها، حيث أنها تمثل نسبة (٢٨%) في أسئلة اختبارات مادة الدراسات الاجتماعية وأيضا مادة تاريخ عمان، وقد يعود ذلك للرغبة في تقليص عدد أوراق الاختبار، كاستراتيجية لتوفير الأوراق وأيضا قد يكون بهدف دفع الطالب للاستجابة الجيدة على أسئلة الاختبار والبعد به عن الخوف من كثر عدد أوراق أسئلة الاختبار، وفي المرتبة الثالثة نلاحظ أن العبارة رقم (١٤) والتي تنص على (يوجد توازن في طول صياغة البدائل)، حيث أنها جاءت بنسبة (٢١%) في أسئلة مادة الدراسات الاجتماعية، في حين أنها شكلت ما نسبته (٢%) في أسئلة اختبارات مادة تاريخ عمان.

والجدول (٦) يوضح عبارات المحور الثالث الخاصة بالإخراج الفني.

جدول (٦) مجموع التكرارات والنسب المئوية والترتيب للمعايير الخاصة بالإخراج الفني												
المادة	العبارة	العام الدراسي										
		2019/2020		2018/2019		2017/2018		2016/2017		2015/2016		
		ف ١	ف ٢	ف ١	ف ٢	ف ١	ف ٢	ف ١	ف ٢	ف ١	ف ٢	
الدراسات الاجتماعية	21	0	0	0	0	0	0	1	0	0	0	2
	22	0	0	0	0	0	0	1	0	1	0	1
	23	0	0	0	0	0	0	1	0	0	0	2
	24	0	0	0	0	0	0	0	0	0	0	-
	25	0	0	0	0	1	0	1	0	0	0	1
	26	0	0	0	0	0	0	0	0	0	0	-
	27	0	0	0	0	0	0	0	0	0	0	-
تاريخ عمان	21	0	0	0	0	0	0	0	0	0	0	-
	22	0	0	1	0	0	0	0	0	0	0	1
	23	0	0	0	0	0	0	0	0	0	0	-
	24	0	0	0	0	0	0	0	0	0	0	-
	25	0	0	0	0	0	0	0	0	0	0	-
	26	0	0	0	0	0	0	0	0	0	0	-
	27	0	0	0	0	0	0	0	0	0	0	-
١: الفصل الدراسي الأول.		٢: الفصل الدراسي الثاني										

ويلاحظ من الجدول (٦) أن العبارة رقم (٢٢) والتي تنص على: وضعت العناصر المشتركة للبدائل في متن السؤال، قد جاءت في المرتبة الأولى في أسئلة اختبارات مواد الدراسات الاجتماعية، حيث أنها تمثل نسبة (٣٣%) في أسئلة اختبارات مادة الدراسات الاجتماعية، ونسبة (١٠٠%) في أسئلة اختبارات مادة تاريخ عمان، وكذلك نلاحظ أن العبارة رقم (٢٥) والتي تنص على: توجد إجابة صحيحة واحدة وقطعية بين البدائل المعطاة، أيضاً قد احتلت المرتبة الأولى في أسئلة اختبارات مادة الدراسات الاجتماعية بنسبة (٣٣%)، ونلاحظ أيضاً أن العبارة رقم (٢١) والعبارة رقم (٢٣) قد جاءت كل منهما في المرتبة الثانية في أسئلة اختبارات مادة الدراسات الاجتماعية، حيث أن العبارة رقم (٢١) تنص على: تشابه الصياغة اللغوية للبدائل من حيث: التذكير أو التأنيث أو الجمع أو الأفراد، والعبارة رقم (٢٣) تنص على: يوجد انسجام لمحتوى متن السؤال لغوياً مع البدائل المعطاة.

ملخص النتائج:

تم تحليل أسئلة الاختيار من متعدد في اختبارات مواد الدراسات الاجتماعية (الدراسات الاجتماعية وتاريخ عمان) وفق مستويات بلوم المعرفية، حيث تم تحليل أسئلة بعدد (٩٦) سؤال، وتم التوصل إلى أن الأسئلة في معظمها قد تركزت في المستويات الثلاثة الأولى من تصنيف بلوم للأهداف المعرفية وهي: التذكر والفهم والتطبيق، مع غياب الأسئلة في مستويات التركيب والتقويم، كذلك تم بناء استبانة قائمة معايير تقييم أسئلة الاختيار من متعدد ضمن (٢٧) معيار موزعة على ثلاثة محاور، أولها قواعد خاصة بالمتن، وهو يتضمن (١٠) معايير، والمحور الثاني خاص بالموهات أو البدائل، وهو يتألف من (١٠) معايير، والمحور الثالث خاص بالإخراج الفني، وهو يتكون من (٧) معايير، وكانت أبرز القواعد التي لم يتم الالتزام بها في محور متن السؤال كانت: كثرة استخدام النفي وماعدا، وإيضاً التعبير عن فكرة السؤال بعدد قليل من الكلمات وبصورة واضحة، بينما في محور الموهات نجد أن قاعدة أو معيار رتب البدائل ترتيباً منطقياً، هجائياً أو زمنياً أو عددياً وكذلك ترتيب البدائل عمودياً، بينما نجد المعيار وضع العناصر المشتركة للبدائل في متن السؤال كانت من أبرز القواعد التي لم يراعها معدي الاختبارات في محور الإخراج الفني.

التوصيات:

في ضوء النتائج التي تم التوصل إليها والتي تشير إلى اقتصار معظم الأسئلة التي يعدها معلمو مواد الدراسات الاجتماعية على المستويات المعرفية الدنيا (التذكر، والفهم، والتطبيق) دون المستويات العليا، وهو ما قد يحرم الطلاب من تنمية قدراتهم العقلية العليا وممارسة أنواع التفكير المختلفة، وبالتالي هو امر قد يخل بتحقيق الأهداف المرجوة من العملية التعليمية والتي تسعى في كافة مراحلها للنهوض بإمكانية الطلاب وقدراتهم العقلية في كافة الجوانب، وعلى ضوء ذلك يوصي الباحث بالتالي:

- إعادة النظر في الطريقة والمنهجية التي يتم من خلالها وضع أسئلة الاختيار من متعدد في مواد الدراسات الاجتماعية، مع مراعاة تضمين أسئلة في المستويات المعرفية المختلفة.
- ضرورة الاهتمام ومراعاة القواعد والمعايير عند بناء أسئلة الاختيار من متعدد، من حيث الاهتمام بقواعد صياغة متن السؤال وأيضاً القواعد الخاصة بالموهات والإخراج الفني للأسئلة.

- عقد دورات تدريبية متخصصة لمعلمي مواد الدراسات الاجتماعية في مجال إعداد وبناء الاختبارات وفق المعايير المحددة لذلك، وطرق صياغة أسئلة في المستويات العقلية العليا.
- الاستفادة من قائمة المعايير التي تم استخدامها في هذه الدراسة، لتفادي الملاحظات في أسئلة الاختيار من متعدد في الاختبارات القادمة.

المراجع

أولاً: المراجع العربية.

- أبو فودة، باسل خميس سالم. (٢٠١٤). أثر إعادة ترتيب بدائل الاستجابة في صعوبة فقرة الاختيار من متعدد. *دراسات عربية في التربية وعلم النفس: رابطة التربويين العرب*، (٥٣): ٢٦٥ - ٢٨٧. مسترجع من <http://search.mandumah.com/Record/654093>
- الحسن، مراد حسن صالح (٢٠٠٣). أثر التغذية الراجعة المقدمة بعد أداء الاختبارات الصفية على التحصيل في الرياضيات لطلاب الصف السابع الأساسي في محافظة قلقيلية (رسالة ماجستير غير منشورة). جامعة النجاح الوطنية، فلسطين.
- الحضرمي، أحمد بن سعيد (٢٠٠٨). تصور مقترح لتطوير إدارة معاهد السلطان قابوس للعلوم الإسلامية بسلطنة عمان في ضوء متطلبات الإدارة الإلكترونية (رسالة ماجستير غير منشورة) معهد البحوث والدراسات العربية، القاهرة، جمهورية مصر العربية.
- جيوسي، مجدي (٢٠١٦)، مدى تحقيق أسئلة الامتحانات النهائية في جامعة فلسطين التقنية للأهداف التعليمية تبعاً لهرم بلوم في ضوء معايير الورقة الامتحانية الجيدة، *المجلة الدولية للتربية المتخصصة*، الجمعية الأردنية لعلم النفس، ٥ (٨).
- الزغيلات، أحمد عبد الحافظ (٢٠١٤). أثر عدد البدائل على دقة تقديرات القدرة للأفراد و دالة المعلومات لاختبار اختيار من متعدد وفق النموذج ثنائي المعلم. *مجلة عالم التربية*. المؤسسة العربية للاستشارات العلمية وتنمية الموارد البشرية، ١٥ (١٢): ٢٥٥ - ٢٧٩.
- سلمان، مفلح حسين؛ وطلافة، فؤاد (٢٠١٨). مدى تمثيل أسئلة الاختبارات التحصيلية النهائية لمستويات الأهداف المعرفية لدى معلمي مبحث العلوم الحياتية في إقليم جنوب الأردن. *دراسات العلوم التربوية*، ٤٥ (١): ٢١٥ - ٢٣١.
- الشقيرات، غازي عبدالله (٢٠٢٠). تحليل أسئلة امتحانات الثانوية العامة في الأردن لمادة تاريخ الأردن في ضوء مستويات المجال المعرفي حسب تصنيف بلوم "دراسة تحليلية لمادة تاريخ الأردن"، *مجلة العلوم التربوية والنفسية*، ٤ (١٥): ١٤٠ - ١٤١.
- الشاهر، تحسين علي؛ والحمداني، انتظار جواد (٢٠١٦). تقويم أسئلة الاختبارات النهائية لمادة النحو في كلية العلوم الإسلامية في ضوء مستويات بلوم المعرفية. *مجلة كلية التربية*، الجامعة المستنصرية، (٢)، ٦١٩ - ٦٥٠.

- الشمراني، صالح موسى، (٢٠١٢ / ١٤٣٣). أثر التفاعل بين نمط الاختبار وجنس المستجيب على ثبات الاختبار وبعض الخصائص السيكمومترية للفقرات (رسالة ماجستير غير منشورة)، كلية التربية، جامعة أم القرى، المملكة العربية السعودية.
 - الطاهر، طعيلي محمد، و محمد، قوارح. (٢٠١٣). معالجة نظرية لمفهوم الاختبارات التحصيلية وأنواعها بمجلة دراسات نفسية وتربوية: جامعة قاصدي مرباح - مخبر تطوير الممارسات النفسية والتربوية، (١٠): ١٧٣-٢٠٢، مسترجع من <http://search.mandumah.com/Record/493965>
 - ظاظا، حيدر (٢٠١٢). الكشف عن مدى انتهاك قواعد صياغة فقرة الاختيار من متعدد في أسئلة شهادة الدراسة الثانوية العامة في الأردن. المجلة الأردنية في العلوم التربوية، مجلد (٨)، العدد (١). ٩١-٨١.
 - عتوم، محمد صالح (٢٠١٤). تقييم الاختبارات التحصيلية من إعداد معلمي العلوم في محافظة جرش دراسة تحليلية لنتائج الطلبة للفصل الدراسي الثاني للعام ٢٠١٢/٢٠١٣ في مبحث العلوم، جرش للبحوث والدراسات، (١٥)، (٢)، ٣٦١-٣٧٥.
 - العنزي، محمد عويض (٢٠١٧). أثر شكل فقرة الاختبار من متعدد والصواب والخطأ على الخصائص السيكمومترية وفق نظرية استجابة الفقرة لاختبار مادة الحاسب الآلي للصف الأول ثانوي في مدينة تبوك (رسالة ماجستير غير منشورة)، كلية التربية والآداب، جامعة تبوك، المملكة العربية السعودية.
 - عودة، أحمد (٢٠١٠). القياس والتقويم في العملية التربوية، الأردن: دار الأمل للنشر والتوزيع.
 - الفرجي، محمد حسن (٢٠١٩). تقويم المهارات العقلية في اختبارات الثانوية العامة بالجمهورية اليمنية ما بين- ٢٠٠٩-٢٠١٤ - (أطروحة دكتوراه غير منشورة)، جامعة محمد الخامس بالرباط، كلية علوم التربية، المملكة المغربية.
 - محمد، محمد أبو العباس فضل المولى (٢٠١٧). تقويم مطابقة اختبارات مادة الرياضيات للصف الرابع مرحلة الأساس للمعايير العلمية للاختبارات (رسالة ماجستير غير منشورة)، جامعة السودان للعلوم والتكنولوجيا، كلية التربية. السودان.
 - المزوغي، ابتسام سالم (٢٠١٨). تقييم أسئلة الامتحانات المهنية في ضوء تصنيف بلوم للأهداف المعرفية بقسم علم النفس بكلية الآداب جامعة صبراتة "انموذجا"، مجلة جامعة صبراتة العلمية، (٣) يونيو: ٩٢-١٠٧.
 - مغامس، خولة فاضل (٢٠١٨). التباين بين التعرف والاستدعاء في قياس المعرفة في ضوء التمييز وصدق الفقرات لدى تلاميذ الصف السادس الابتدائي في مادة العلوم. مجلة كلية التربية الأساسية، ٢٤ (١٠٢): ٦٢٥-٦٤٨.
 - الناقة، صلاح أحمد عبد الهادي. (٢٠١٦). تقويم أسئلة اختبارات الكيمياء للصف الثاني عشر وفق معايير الجودة خلال السنوات من ٢٠٠٧ إلى ٢٠١٥. مجلة الجامعة الإسلامية للدراسات التربوية والنفسية: الجامعة الإسلامية بغزة - شئون البحث العلمي والدراسات العليا، ٢٤ (٣): ٦٤ - ٨٢.
- مسترجع من <http://search.mandumah.com/Record/761897>

- اللحاني، عفاف بنت راضي (٢٠١٠-١٤٣١). أثر بعض طرق تقدير الدرجات للمفردات على ثبات وصدق درجات اختبار تحصيلي في الرياضيات ذي الاختيار من متعدد لدى طالبات الصف الأول الثانوي بمكة المكرمة (رسالة ماجستير غير منشورة)، كلية التربية، جامعة أم القرى- مكة المكرمة، المملكة العربية السعودية.
 - الهدور، زيد أحمد ناصر (٢٠١٧). تحليل اسئلة اختبارات الشهادة الثانوية العامة لمادة الرياضيات بالجمهورية اليمنية في ضوء التصنيفات الحديثة للأهداف التعليمية، مجلة الجامع في الدراسات النفسية والعلوم التربوية، (٧)، ٢٣٢-٢٥٧.
 - هندوي، أسامة سعيد علي (٢٠١٠). أثر ثلاثة تصميمات لأنماط الاستجابة على الاختبارات الإلكترونية على معدل الأداء الفوري والمؤجل لطلاب الجامعة في الاختبار، مجلة العلوم التربوية – كلية الدراسات العليا للتربية، جامعة القاهرة، ١٨ (٣): ١٠٢-١٤٣.
- ثانياً: المراجع الأجنبية.

- Alzu'bi, Mohammad Akram. (2014). The Extend of Adaptation Bloom's Taxonomy of Cognitive Domain In English Questions Included in General Secondary Exams. **Advances in Language and Literary Studies**. ISSN: 2203-4714 Vol. 5 No. 2; April 2014. URL: <http://dx.doi.org/10.7575/aiall.v5n.2p.67>
- Dehnad, A. , Nasser, H. & Hosseini, A. (2014). A Comparison between Three- and Four-Option Multiple Choice Questions. **Social and Behavioral Sciences** 98 (2014) 398 – 403. <https://doi.org/10.1016/j.sbspro.2014.03.432>
- Haladyna, T.; Downing, S. & Rodriguez. (2002). A Review of Multiple-Choice Item-Writing Guidelines for Classroom Assessment. **APPLIED MEASUREMENT IN EDUCATION**, 15(3), 309–334.
- Hasan, M. Naomee, I and Bilkis, R. (2013). REFLECTION OF BLOOM'S REVISED TAXONOMY IN THE SOCIAL SCIENCE QUESTIONS OF SECONDARY SCHOOL CERTIFICATE EXAMINATION. **THE INTERNATIONAL JOURNAL OF SOCIAL SCIENCES**. (14). (1). ISSN: 2305-4557. P. 47-56.
- Kelly, M. (2014). **Bloom's Taxonomy in the Classroom**. . [Online] Available <http://712educators.about.com/od/testconstruction/p/bloomstaxonomy.htm>
- Mohnot, A. (2006). 'Seeing the Bigger Picture': Higher Order Cognition in the Indian Certificate of Secondary Education (ICSE) English Literature Examination. (Master Degree). Stanford University.

- Sevilay, K.; Serkan, S.; Orhan, K. & Salih Ç. (2003). Analysis of Turkish High-school Chemistry-examination questions according to Bloom's Taxonomy. **Chemistry Education: Research and Practice**. (4) 1. 25-30.

م	بنية الفقرة	معايير التصنيف	ينطبق	لا ينطبق
1	المتن	يخلو من الأخطاء الإملائية والنحوية		
2		يخلو من الكلمات المضللة للمستجيب		
3		لا توجه صياغة المتن الطالب للإجابة الصحيحة		
4		يخلو من النفي، مثل: ما عدا أو لا والنفي المركب		
5		صيغت مفردات المتن بعبارات واضحة		
6		يُعبّر المتن عن فكرة واضحة وليس رأي		
7		تم بناء المتن على محتوى عام للتعلم		
8		يستقل متن السؤال الواحد عن الأسئلة الأخرى		
9		عُبر عن متن السؤال بعدد قليل من العبارات		
10		عُبر في المتن عن الفكرة الرئيسة للسؤال بوضوح		
11	المموهات / البدائل	رتبت البدائل ترتيباً منطقياً، هجائياً أو زمنياً أو عددياً		
12		رتبت البدائل عمودياً		
13		تستقل البدائل عن بعضها البعض		
14		يوجد توازن في طول صياغة البدائل		
15		يوجد تجانس في محتوى البدائل المعطاة للسؤال		
16		تتساوى البدائل المعطاة للسؤال في جاذبيتها ظاهرياً		
17		لا تتضمن البدائل عبارة "لا شيء مما ذكر" كإجابة صحيحة		
18		لا تتضمن البدائل عبارة "جميع ما ذكر" كإجابة صحيحة		
19		لا تتضمن البدائل إيعاءات للإجابة الصحيحة		
20		تخلو البدائل من الألفاظ المحددة مثل: قطعاً أو دائماً أو فقط أو لا غير		
21	الإخراج الفني	تشابه الصياغة اللغوية للبدائل من حيث: التذكير أو التأنيث أو الجمع أو الأفراد		
22		وضعت العناصر المشتركة للبدائل في متن السؤال		
23		يوجد انسجام لمحتوى متن السؤال لغوياً مع البدائل المعطاة		
24		لا تمثل الإجابة الصحيحة البديل الأطول		
25		توجد إجابة صحيحة واحدة وقطعية بين البدائل المعطاة		
26		يتغير موقع الإجابة الصحيحة بين البدائل بأسلوب عشوائي		

٢٧	يوجد بديلاً غير واضح كإجابة صحيحة من البدائل المعطاة		
----	--	--	--

ملحق (٢) قائمة بأسماء المحكمين لأداة الدراسة*

م	الاسم	الوظيفة	الدرجة العلمية	جهة العمل
١	أحمد بن سعيد بن ناصر الحضرمي	مساعد مدير دائرة تنمية الموارد البشرية	دكتوراه	مركز السلطان قابوس العالي للثقافة والعلوم
٢	حكم بن سالم بن حكم الفارسي	مدير عام مساعد	دكتوراه	وزارة التربية والتعليم
٣	صبري حسن طراونة	أستاذ جامعي	أستاذ مشارك	جامعة مؤتة
٤	علي مهدي كاظم	أستاذ جامعي	أستاذ دكتور	جامعة السلطان قابوس
٥	نبيل بن محمد بن زاهر العبري	مشرف مواد الدراسات الاجتماعية	دكتوراه	مركز السلطان قابوس العالي للثقافة والعلوم
٦	ياسر أحمد عبدالله المسهلي	أستاذ جامعي	أستاذ مساعد	جامعة حجة

*مرتبة أجبدياً