متطلبات الاستخدام الفعال لاختبارات الخيارات المتعددة ضمن وسائل التقويم في العلوم الاجتماعية

إعـــداد

د. خالد أحمد مصطفى حجر
الأستاذ المساعد بقسم الخدمة الاجتماعية
كلية العلوم الاجتماعية
جامعة أم القرى – مكة المكرمة

ملخص:

تهدف هذه الدراسة إلى لفت نظر الباحثين والأكاديميين في مختلف مجالات العلوم الاجتماعية إلى أهمية التقويم في تلك العلوم والذي يعتبر من أهم جوانب آليات التعرف على مدى ما حققته تلك العلوم من رسالتها نحو تزويد الطالب بالفهم العميق للعلاقات الاجتماعية التي تحدث بين أفراد المجتمع وتزويده بالفكر الناقد البناء والمعارف والمهارات التي تمكنه من المساهمة في خدمة مجتمعه بأقصى قدر ممكن من الفاعلية. وتأمل الدراسة أن يجد موضوع التقويم في العلوم الاجتماعية ما يستحقه من اهتمام بعد أن ظلت أهم آلياته متروكة بصورة شبه كاملة للآخرين من المجالات العلمية الأخرى على الرغم من ارتباطها العضوي بجميع أركان ومقومات المنهج.

وفي المجال الأكاديمي يستند التقويم (evaluation) على مرتكزات من نوع خاص، تأتي على رأسها وسائل التقييم (assessment) التي تتمثل في الاختبارات بشتى أشكالها. وفي هذا الإطار تتناول الدراسة جانبا من أبرز تلك الوسائل السائدة حاليا في مختلف العلوم الاجتماعية يتمثل في اختبارات الخيارات المتعددة (multiple-choice tests) للتعريف بماهيتها وطبيعتها وإيجابياتها وسلبياتها وشروط بنائها بشكل سليم، كما توضح من خلال استطلاع تم إجراؤه ضمن مقتضياتها نظرة طلاب العلوم الاجتماعية لهذا النوع من الاختبارات وأبرز إيجابياتها وسلبياتها من وجهة نظرهم والتي تبرر ارتفاع شعبيتها لديهم مقارنة بوسائل التقييم الأخرى.

وتسعى الدراسة إلى الخروج بمقترحات تساعد على الاستخدام الأمثل لهذه الاختبارات، تتضمن شروط البناء السليم لبنودها بعد قيامها

بتحديد أبرز مخالفات تلك الشروط من خلال تحليل عينة من الاختبارات المبنية على هذا النمط، وبناء تصور بديل يتضمن ضرورة التعرف على ما ينبغي تقويمه وقياسه من الحصيلة المعرفية لطالب العلوم الاجتماعية واستخدام أكثر من أسلوب واحد في ذلك التقويم وبناء الأسئلة القادرة على قياس التفكير الناقد والمعارف والمهارات التطبيقية والتحليلية التي اكتسبها.

المفاهيم الأساسية: التقويم ؛ التقييم؛ اختبارات الخيارات المتعددة.

Abstract:

The primary objective of this study is to draw the attention of researchers and academicians in the various branches of the social sciences to the importance of evaluation, being the only mechanism that enables them to assess the learning outcomes of their respective fields. Means of evaluation in the social sciences have so far been the concern of researchers in academic fields other than the social sciences, although it is a process that directly relates to all curriculum aspects and reflects on the level of objectives achievement; it is time – the study anticipates - that social scientists took over this responsibility instead of others.

Along this line, this study concerns itself with assessment as a fundamental tool of evaluation, and assessment through the use of multiple choice tests in particular. It shows the main features, advantages and disadvantages of multiple-choice tests, and explains the rules for building them to become a useful vehicle for authentic, performance-based assessment. A survey was carried out among a random sample of the students studying social sciences to show what pros and cons they mostly think about in such tests, and to what extent they favor them to other types of tests for assessment.

The study puts forward some suggestions to improve efficiency of multiple-choice tests in assessing students learning in the social sciences, including proper item-building of these tests, clear understanding of what aspect of learning to assess, and employment of more than one assessment format to assess critical thinking and actual achievement level attained by the student.

تقديـــــم:

تشهد أنظمة التعليم العالي على مستوى العالم تغيرات كمية وكيفية كبيرة ومتسارعة نتيجة لما يعتري عالم اليوم من تحول وتعقيد ومستجدات اقتصادية واجتماعية وسياسية وتقنية انتظمت كافة أوجه الحياة . وكان لا بد للتعليم العالي أن يستجيب لتلك المستجدات ويتفاعل معها مؤثرا ومتأثراً، وخاصة في مجال تقويم الحصيلة المعرفية الذي لا يسعى فقط للتعرف على ما تم جمعه من بيانات ومعلومات متفرقة ، بل لمعرفة ما قد تحقق وما يمكن أن يتحقق مستقبلاً من معارف ومهارات وتوجهات وقيم.

إن من أهداف تدريس العلوم الاجتماعية إعداد الخريجين المؤهلين بالمهارات والنظرة الناقدة ومهارات الاتصال والمرونة في الفكر والنظرة الموضوعية للظروف الاقتصادية والاجتماعية في إطار مستجدات الواقع المعاش وآمال المستقبل المرجوة، والقدرة على التعامل مع ظروف التغير الاجتماعي المتسارع بعلمية وموضوعية ورؤية متعمقة وفهم للروئ والأفكار الأخرى. ولتأهيل برامج العلوم الاجتماعية حتى تتمكن من تحقيق هذه الأهداف لا بد من إعادة النظر في بنائها ومحتواها لإعداد الخريج ذو الفكر الناقد والمزود بمهارات الاتصال والقادر على المساهمة الإيجابية في حل المشكلات. ويعتبر التقويم واحد من أهم جوانب هذا الإعداد،خاصة وأنه يؤثر تأثيرا مباشرا على ما تم تحصيله وعلى وسائل الحصول عليه.

غير أن مجال التقويم في العلوم الاجتماعية لم يحظ بما يستحقه من اهتمام على الرغم من أهميته. فمن المعروف أن التقويم لا يسعى إلى التعرف على حصيلة الطالب المعرفية فحسب، بل يعتبر من الوسائل الضرورية لتطوير المنهج المدروس لتأثيره المباشر على مختلف جوانب

المنهج الأخرى. والاختبارات هي الوسائل لذلك التقويم – يقوى بقوتها ويضعف بضعفها، مما يستدعي إيلاءها التفاتة خاصة لتدعيم مواطن قوتها ومعالجة أسباب ضعفها.

ومن تلك الوسائل الاختبارات الموضوعية وعلى رأسها اختبارات الخيارات المتعددة التي تشهد إقبالا غير مسبوق على استخدامها في مختلف مجالات العلوم الاجتماعية المعاصرة في مؤسسات التعليم العالي نظرا لما تتميز به من خصائص عملية مقارنة بغيرها من وسائل التقويم مثل الموضوعية ودقة الإجابات وعدم تأثرها بقدرات الطلاب اللغوية وأساليبهم في الكتابة وسهولة تصحيحها ورصدها آليا وتوفيرها للوقت والجهد.

وفي المقابل نجد أن هذا النوع من الاختبارات يعاني من سلبية جوهرية بتركيزه على اختبار معرفة الطالب الاسترجاعية وبعض ما لديه من مهارات معرفية وضعف مقدرته على التعرف على ما لديه من قدرات تطبيقية وتحليلية وتفكير ناقد ومرونة في التعامل مع القضايا المعقدة التي يتسم بها الواقع الاجتماعي الذي يتوقع أن يتعامل معه بعد تخرجه.

إن الاقتصار على استخدام اختبارات الخيارات المتعددة في التقويم في مجال العلوم الاجتماعية لا يؤدي إلى تقويم ناقص فحسب، بل يساهم أيضا في تدني معارف الطلاب، حيث كشفت هذه الدراسة ودراسات أخرى أن الطلاب يميلون نحو القراءة السريعة أو الحفظ وتخطي المواضيع الصعبة عندما يعرفون أن اختبار الخيارات المتعددة هو وحده الذي سيستخدم في تقويمهم، بينما يكون استعدادهم أكثر عمقاً وشمولاً وترابطا في حالة توقع استخدام أساليب أخرى في التقويم كالاختبارات المفتوحة والمقالية والعملية.

والقضية هنا ليست مجرد النظرة إلى الاختبارات الموضوعية كآلية

مناسبة أم لا لتقويم المستوى في العلوم الاجتماعية بقدر ما هي قضية معرفية هامة تعكس ما هنالك من رؤى حول ما هو واقع من شعور بدور العلوم الاجتماعية في صناعة المستقبل وما هو مأمول من مقدرات وكفاءات توفرها لذلك الهدف. فالتقويم هو أحد العناصر الهامة للمنهج، وهو لا يساعد في معرفة ما تحقق من أهداف المادة فحسب، بل يساعد في وضع وسائل تحقيق تلك الأهداف وفي تطوير محتواها وفعالياتها. إن الاهتمام بالعلوم الاجتماعية والسعي نحو تطويرها والتصدي لأزماتها المتتابعة لا يكتمل إلا بتحسس جميع مواطن العلل والتي يمثل التقويم المتبع حاليا واحدا منها لم يحظ بما يستحقه من اهتمام.

موضوع الدراسة:

تدخل عملية تقويم الطلاب ضمن أهم القضايا التعليمية في كافة "مجالات وبمختلف مستوياتها. ويوضح تاريخ تطور هذه العملية الاهتمام المستمر بدقة وكفاءة الاختيارات المستخدمة في هذه العملية. وعلى مستوى التعليم العالي – وخاصة في مجال العلوم الاجتماعية – لم يجد أمر التقويم ما يستحقه من اهتمام من الباحثين والأكاديميين المختصين في تلك المجالات باعتباره من صميم اختصاص العاملين في الحقل التربوي . هذا الضعف الواضح في الاهتمام بالبحوث التي تهتم بقضايا التقويم في العلوم الاجتماعية انعكس سلبا على فرص منحه ما يستحقه من اهتمام كأحد مقومات المنهج، حيث يقوم المختصون في تلك التخصصات بمهامهم التدريسية دون أن يهمهم من أمر التقويم سوى وضع اختبارات من ثنايا المنهج الذي يدرسونه بالشكل الذي يرونه وبالأسلوب الذي يراه كل منهم مناسبا، على الرغم من أنه يمثل

أحد المداخل الرئيسية لتطوير عناصر المنهج الأخرى المتمثلة في أهدافه ومحتواه وفعالياته.

ولعل أبرز وجوه ذلك القصور في الاهتمام بالتقويم في العلوم الاجتماعية هو النظرة القاصرة إليه باعتباره مجرد "مناسبة" تقليدية تعقد بصورة آلية يتم من خلالها اختتام فعاليات المقرر. إن التعامل مع التقويم كـــآخر فعاليات المنهج أو آخر فعاليات جزء منه لهو دليل قاطع على أن التقويم لا يوظف بصورة صحيحة في عناصر المنهج الأخرى ولا يساهم في زيادة حصيلة الطالب، ويندر – أو يكاد أن ينعدم – اهتمام مدرسي العلوم الاجتماعية بمراجعة نتائج الاختبارات مع طلابهم سواء كانت خلال الفصل الدراسي أو في نهايته. ففي كثير من الأحيان تستخدم هذه الاختبارات للتقويم النهائي مما يجعل التقويم يبدو وكأنه منفصل عن التعلم – مجرد طريقة لرقابة الأداء وليست وسيلة لتحسينه. ولكن عندما تستخدم هذه الاختبارات لغرض التقويم المستمر فإنها توفر وسيلة فعالة من وسائل التدريس والتعلم. وفي ظل تدني اهتمام النظام العلمي بالتقويم كوسيلة من وسائل تطوير مناهجه فليس من المستغرب أن ينعدم الاهتمام بوسائله – أي الاختبارات – التي أصبحت أمرا خاصا بكل مدرس يتعامل معه كيفما شاء.

لقد أدى هذا الإهمال الواضح لقضايا التقويم في العلوم الاجتماعية إلى ترسيخ ثقافة الاختبارات التقليدية والنوعية الضحلة من الأسئلة التي أضعفت من فرصة امتلاك الطالب لرصيد حقيقي من مهارات التفكير الناقد والمرونة والعقلانية المطلوبة للتعامل الفعال مع الواقع الاجتماعي عندما يواجهونه بعد إكمال دراستهم. وقد ظلت تلك الأنماط التقليدية من وسائل

التقويم قاصرة عن تلبية الطموحات في التطوير العلمي والخروج بالمعرفة الاجتماعية من قاعات الدراسة إلى آفاق الواقع والتمكن من تطبيقها بكفاءة واقتدار.

ويلاحظ في الآونة الأخيرة ارتفاع نسبة استخدام الاختبارات الموضوعية في العلوم الاجتماعية وخاصة اختبارات الخيارات المتعددة (multiple-choice tests) لما تتمتع به من مزايا تجعل "مهمة" التقويم أمرا سهلا وسريعا، في ظل أعداد متنامية من الطلاب الملتحقين بدراسة العلوم الاجتماعية في شتى مجالاتها. وقد نما هذا التركيز الزائد على الاختبارات الموضوعية على حساب وسائل التقويم الأخرى كالاختبارات المقالية أو ما تسمى باختبارات الاستجابة الحرة (free- response tests) والتي تتمتع بقدرة كبيرة على قياس مستويات أعلى من المعرفة مقارنة بغيرها.

إن استخدام الاختبارات الموضوعية في العلوم الاجتماعية هو أمر ضروري في بعض الأحيان ولكنه غير كاف بمفرده لقياس الحصيلة المعرفية بكافة مستوياتها. كما أن غالبية الاختبارات الموضوعية المستخدمة تتضمن مخالفات كبيرة في شروط البناء الجيد المطلوب لحقيق فاعليتها، وتتطلب تدريبا وتوجيها في كيفية إعدادها. وينطبق هذا القول على اختبارات الاستجابة الحرة أيضا، إلا أنه من الضروري التعامل مع هذه الاختبارات في إطار منظومة فكرية قائمة على خصائص النظام العلمي قيد النظر.

وتسعى هذه الدراسة لإلقاء المزيد من الضوء على مشكلة التقويم في العلوم الاجتماعية واستكشاف أبرز معالمها من خلال التعرف على مشكلة الاقتصار على استخدام الاختبارات الموضوعية ممثلة في اختبارات

الخيارات المتعددة التي تتناول الدراسة أبرز إيجابياتها وسلبياتها وكيفية بنائها وشروط دقتها، وأهمية دمجها مع الوسائل التقويمية الأخرى المتمثلة في اختبارات الاستجابات الحرة لتحقيق أغراض تقويم الحصيلة المعرفية على أعلى مستوياتها. وينصب تركيز هذه الدراسة على اختبارات الخيارات المتعددة لأسباب نظرية تتمثل في هيمنتها على ساحة وسائل التقويم وتعدد مشكلاتها البنائية، وأسباب عملية تتمثل في ضيق الحيز الزمني والمكاني الممنوح لهذه الدراسة والذي لا يتيح المجال لتناول أشكال الاختبارات الأخرى بقدر كاف من التفصيل .

أهداف الدراسة وتساؤلاتها:

تهدف هذه الدراسة إلى لفت نظر الباحثين في مختلف مجالات العلوم الاجتماعية إلى أهمية التقويم في تلك العلوم، حيث ظل متروكا بصورة شبه كاملة للآخرين من المجالات العلمية الأخرى كعلم التربية مثلا، ما عدا قلة من المهتمين بمسألة القياس والتقويم في بعض مجالات العلوم الإنسانية الأخرى كعلم النفس وعلوم اللغات مثلا، علما بأن التقويم هو مسألة ترتبط ارتباطا عضويا بالمنهج الذي يتم تقويمه وبجميع أركانه ومقوماته.

وفي هذا الإطار تتناول الدراسة جانبا من أبرز أركان التقويم يتمثل في إحدى أهم آلياته وهي الاختبارات، حيث تركز فيها على اختبارات الخيارات المتعددة باعتبارها أهم أشكال الاختبارات الموضوعية السائدة في عمليات التقويم فتحاول توضيح ماهيتها وإيجابياتها وسلبياتها وما ينبغي

الالتزام به أو تجنبه من أجل بنائها بشكل سليم يحقق فاعليتها كوسيلة تقويمية صادقة.

كما تهدف الدراسة أيضا إلى وضع تصور يساعد على الاستخدام الأمثل لهذه الاختبارات في تقويم الأداء في العلوم الاجتماعية يوضح شروط بناء بنودها وكيفية دمجها مع وسائل تقويمية أخرى للتمكن من قياس ما اكتسبه الطالب من فهم وتطبيق وتحليل وتركيب وتقويم بدلا من مجرد اختبار مقدراته على التذكر واستدعاء المعلومات.

وعلى ضوء أهداف هذه الدراسة جاءت تساؤلاتها على النحو التالي:

١- ما هي اختبارات الخيارات المتعددة وما أبرز سلبياتها وإيجابياتها وخاصة فيما يتعلق بقياس الحصيلة المعرفية في مجالات العلوم الاجتماعية؟

٢- كيف ينظر طلاب العلوم الاجتماعية لاختبارات الخيارات المتعددة كإحدى وسائل التقويم، وما أبرز إيجابياتها وسلبياتها من وجهة نظرهم؟ وهل توجد فروق ذات دلاله إحصائية بين طلاب العلوم الاجتماعية حول سلبيات وإيجابيات اختبارات الخيارات المتعددة حسب الجنس أو المستوى الدراسي؟

٣- ما أهم التوجيهات للبناء السليم لبنود اختبارات الخيارات المتعددة وأهم المخالفات لتلك التوجيهات في الاختبارات التي تعقد لطلاب العلوم الاجتماعية؟

٤- كيف يمكن توظيف اختبارات الخيارات المتعددة ضمن وسائل تقييمية أخرى للتمكن من تقييم الحصيلة المعرفية في العلوم الاجتماعية على أعلى مستوياتها الفكرية والمهارية ؟

منهجية الدراسة:

وتعد هذه الدراسة من الدراسات الاستطلاعية التي تسعى إلى الكشف عن مختلف جوانب المشكلة قيد الدراسة من خلال وصف أبرز جوانبها وتحليلها في إطار مقاربة عبر- منهجية (interdisciplinary approach) تتعامل مع العلوم الاجتماعية كمجال واحد لدراسة الظاهرة الاجتماعية من جوانب مختلفة ولكنها تشترك جميعها فيما يتعلق بقضايا التقويم ومستوياته وأهدافه ووسائله. وقد استخدم في البحث أسلوب التعدد المنهجي (methodological triangulation) في تحديد مصادر البيانات وجمعها وتحليلها. ويتضمن هذا الأسلوب استخدام أكثر من مصدر للبيانات وأكثر من طريقة واحدة لجمع البيانات و أكثر من مقاربة أو نظرية لتحليلها (الدامغ، ١٩٩٦؛ Denzin, 1970) وهو أسلوب يساعد في جمع البيانات بطريقة شاملة ومتكاملة ويستقطب مختلف وجهات النظر في تحليلها وبالتالي المساهمة في تأكيد مصداقيتها.

طبقت هذه الدراسة بين طلاب وطالبات كلية العلوم الاجتماعية بجامعة أم القرى خلال الفصل الدراسي الأول من العام الدراسي ٢٠٠٦/٢٠٠٧ واستخدمت فيها عدة أساليب لجمع البيانات تضمنت تحليل محتوى بعض اختبارات الخيارات المتعددة واستبياناً وزع على عينة من طلاب وطالبات قسم الخدمة الاجتماعية بالكلية للتعرف على آرائهم حول بعض جوانب اختبارات الخيارات المتعددة. هذا بالإضافة إلى ملاحظات الباحث القائمة على التأمل والاستبطان من خلال تجربته في هذا المجال مع

مجتمع الدراسة، والاطلاع على ذخيرة وافرة من الأدبيات المتعلقة بموضوع الدراسة.

أولا: اختبارات الخيارات المتعددة وقياس الحصيلة المعرفية:

يتكون السؤال في اختبار الخيارات المتعددة من جزءين: الجذر (stem) الذي يتضمن السؤال أو الحالة أو الموقف المستفسر عنه، ومجموعة خيارات تمثل البدائل المحتملة للإجابة (distractors) والتي تكون من بينها إجابة واحدة متفق على أنها هي الإجابة الصحيحة، أو أكثر من إجابة واحدة صحيحة تتفاوت في درجة صحتها. وعادة ما يتراوح عدد هذه الخيارات بين ثلاثة وأربعة ينبغي أن تكون متعلقة جميعها بمعلومات لها درجة ما من القبول للإجابة على السؤال المطروح. والفكرة المضمنة في خيار ما إما أن تكون حقيقة (صواب) أو ليست بحقيقة (خطأ) أو على مدرج متصل من القبول بالمقارنة مع الحقائق المضمنة في الخيارات الأخرى المطروحة لنفس السؤال. وبناء على ما تتضمنه صيغة السؤال فإن اختيار الخيار الذي يتضمن الحقيقة يكون هو الإجابة الصحيحة ، ويكون ذلك الخيار مطابقا لمفتاح الإجابات الصحيحة المعد سلفا.

وتستخدم اختبارات الخيارات المتعددة بصورة واسعة اليوم وسط مختلف المجالات العلمية. وليس بمستغرب أن يشيع استخدامها بصفة خاصة في مجال العلوم الطبيعية نظرا للطبيعة المعرفية والمنهجية المميزة لها والتي تتسم بالموضوعية والدقة والثبات النسبي لظواهرها مما يسر من استخدامها بكفاءة عالية في تقييم الطلاب والتمييز بين مستوياتهم بدرجة عالية من الصدق. غير أن ما يعاب على هذا الشكل من أشكال التقويم تركيزه على

عمليات استدعاء المعلومات والتي يمكن أن تقوم في بعض الأحيان على البداهة والتخمين إذا لم يراعى الحذر في تصميم بنودها. إن مجرد استدعاء معلومة صحيحة مطروحة أمام الطالب لا يربط بالكفاءة ولا يعني توفرها لديه، وعليه فإن هذا النوع من الاختبارات لا يمكنه بمفرده تقويم إستراتيجية التقصي أو صياغة المشكلة أو المهارات الذاتية المتوفرة لدى طالب العلوم الاجتماعية.

ومن عيوب اختبارات الخيارات المتعددة أنها تضر ببعض الطلاب أكثر من الآخرين، فقد وجد "لين" وآخرون (Linn et.al, 1991) أن بعض الطلاب يجد صعوبة في اللغة وفي التمييز ما بين الخيارات المتشابهة في معانيها، ويعني ذلك أنها غير صادقة لأنها تقيس درجة فهم الطالب للغة بدلا من فهمه للمادة موضوع السؤال. كما يشيرون أيضا إلى أن النظرة البديلة لمفهوم الصدق يجب أن تأخذ في الاعتبار ما يترتب عليه الاختبار، ولكي يكون الاختبار صادقا لا بد من معرفة تأثير التقييم على أسلوب الأساتذة والطلاب في التفكير في أهداف المقرر التعليمية: هل هذا النوع من التقييم يؤدي إلى أنشطة بناءة بين الطلاب تساهم في تحقيق الأهداف المرجوة؟

وهناك نظرة تقليدية ترى أن اختبارات الخيارات المتعددة تمتاز بدرجة عالية من الثبات لأنها لا تمنح المصحح فرصة للتحيز أو التفضيل. ولكن "ويجنز" (Wiggins, 1993) برى غير ذلك ، فهو يرى أن المصحح عادة ما يكون هو واضع الاختبار نفسه ويمكنه أن يجانب الموضوعية عند تصميمه له. كذلك ترى "جبس" (Gipps, 1996) أن فكرة القياس الدقيق غير المتأثر باعتبارات ذاتية هي فكرة غير صحيحة لأن التقييم يتأثر بدوافع

وأفكار من يقوم بعملية التقييم. ويرى "ويجنز" (Wiggins, 1993) أن على المختبرين مسؤولية إثبات صدق نتائج الاختبارات بدراسة المؤثرات على إجابات الطلاب والأسباب التي دفعتهم للإجابة بالطريقة التي أجابوا بها – وهو أمر نادر الحدوث بعد اختبارات التقييم النهائية.

كما يشير آخرون إلى أن درجة الثبات في اختبارات الخيارات المتعددة ، والتي تعتبر إحدى مزايا هذه الاختبارات، مبالغ فيها بعض الشيء وينظر إليها على حساب الصدق (Gipps, 1994; Linn et al., 1991; Moss, 1992)، علماً بأن الثبات في الاختبار فهو شرط ضروري ولكنه غير كاف لتحقيق الصدق (Moss, 1992) . كذلك يترتب على الاعتماد على اختبارات الخيارات المتعددة لوحدها دون غيرها في تقويم الطلاب عدم منح الطلاب فرص تنمية مهارات الاتصال في المواد التي يدرسونها. فكفاءة الاتصال تتطلب مجموعة من المهارات كالمهارات اللغوية والمعرفية واستخدام أسلوب مناسب لموقف التفاعل مع مراعاة أسلوب اللغة والمفردات ومستوى المخاطب. ويتعلم الأفراد كفاءة التخاطب من خلال المعايشة في تفاعلات ذات معنى مع من حولهم. وبالتالي فإذا أردنا أن نمحو أمية الطالب الاجتماعية فيجب إعطاؤه الفرص للدخول في تعاملات واتصالات اجتماعية حقيقية تمكنه من بناء المعاني واستيعاب أهداف البرامج وغاياتها الاجتماعية التي ينتظر منه التعامل معها في المستقبل.

وكثيرا ما يثار التساؤل حول إمكانية التعميم حول مستوى تحصيل الطلاب بناء على نتائج هذه الاختبارات وحول ما إذا كانت النجاحات التي تحققت فيها يمكن أن تتحقق أيضا على أرض الواقع. وحتى عند اختيار

الإجابة الصحيحة يظل التساؤل قائما حول درجة استيعاب الطالب للمعاني والمفاهيم لأن اختبارات الخيارات المتعددة لا تخبرنا عن ذلك. وقد لاحظت "فلور" (Flower, 1994, p. 10) أن هناك علاقة غير مؤكدة بين ما تقيسه الاختبارات المعيارية وما يفعله الناس. وكما تمت الإشارة إليه من قبل فإن المعلومات غير المترابطة التي تهتم بها تلك الاختبارات بعد أن تنتزعها من أطرها تتلاشى سريعا من الذاكرة لأنها دون معنى ولا تنتمي إلى الإطار المعرفي للطالب، وبالتالي فهي معرفة يصعب تعميمها لأنها تنسى سريعا . (Resnick & Resnick, 1992)

وهناك سلبيات أخرى عديدة تشتهر بها اختبارات الخيارات المتعددة دون غيرها تؤثر سلبا على كفاءتها. ومن خلال تجربة الباحث في هذا المجال وملاحظاته الناقدة لهذه للاختبارات لاحظ العديد من السلبيات التي ترتبط بطبيعة الاختبار أو ضعف إعداده. فالعديد من السلبيات قد تظهر من الوهلة الأولى وكأنها ترتبط في المقام الأول بواضعي الاختبارات أكثر من ارتباطها بالاختبارات نفسها، مثل ضعف محتوى البنود وسذاجة الخيارات واعتمادها على بعضها البعض أو تداخلها وكتابتها بلا مسافات كافية فيما بينها توفيرا للمساحة في ورقة الاختبار، إلا أن طبيعة تلك الاختبارات هي التي تؤدي بهم في كثير من الأحيان إلى ارتكاب تلك الأخطاء. فطبيعة التساؤلات في المجال الاجتماعي لا تتسق مع أسلوب البحث عن الحقائق المنتزعة من إطارها الذي تنتهجه اختبارات الخيارات المتعددة مما يؤدي إلى اختبارات ضعيفة وضحلة المحتوى إذا بنيت بهذا الأسلوب.

لقد شاع استخدام اختبارات الخيارات المتعددة في العلوم الاجتماعية بصورة لافتة. والمفارقة هنا تكمن في ازدياد اللجوء لاستخدام هذا النوع من الاختبارات لطلاب المستويات المتقدمة من الدراسة الجامعية في مختلف مجالات العلوم الاجتماعية. ويمكن القول بأن مثل هذه الاختبارات بفردها لا تعطي نتائج دقيقة لمستوى الفكر الناقد أو مهارات الممارسة والاتصال لدى الطالب ، خاصة عندما يكون إعدادها ضعيفا – وهو أمر شائع بصورة كبيرة في مثل هذه الاختبارات دون غيرها.

وكثيرا ما يكون لحجم الطلاب الدارسين للمقرر دور كبير في تحديد نوعية الاختبار. فالأعداد الكبيرة للطلاب داخل المجموعة الواحدة وتعدد المجموعات للمقرر الواحد تدفع بمدرس المقرر إلى اختيار أسلوب الخيارات المتعددة لإمكانية تصحيحه آليا أو بمساعدة الغير باستخدام مفتاح الإجابات السحيحة وسهولة رصد الدرجات وتوفير الوقت والجهد مقارنة بأساليب [illegible] الأخرى كالمقالية وغيرها. وفي مثل هذه الحالات تكون مثل هذه [illegible] التي تمتاز بسهولة الإدارة والرصد أكثر جاذبية من غيرها.

ولا ننكر أن اختبارات الخيارات المتعددة تساعد في قياس أشياء هامة مثل تذكر المعلومات وحفظ التعاريف وفهم المفاهيم الأساسية، وهي جزء من متطلبات ضرورية لتكوين الأساس للفهم الواسع والتفكير المتعمق، ولكنها لا تؤدي إلى ذلك تلقائيا . وحتى عند تعلم أبسط الحقائق يحتاج الطالب لإيجاد "نظرية" مصغرة لنفسه يتمكن من خلالها استيعاب تلك الحقائق لأن حفظ المعلومات لا يتم في غياب المعرفة المنظمة وإلا زالت عن الذاكرة سريعا.

ولا يكفي الحفظ والتذكر وحده لتنمية الفكر الاجتماعي لأن الطالب في حاجة إلى المحاولة والتدرب على التطبيق والتعبير عن فهمه بأسلوبه الخاص أولا مهما كان به من أخطاء. فالظواهر الاجتماعية ليست كالظواهر الطبيعية في مستوى دقتها وانتظامها والعوامل المؤثرة فيها، مما يزيد من أهمية الممارسة لطالب العلوم الاجتماعية لربط المعرفة النظرية بأرض الواقع واستقراء الوقائع واستنباط الآراء والأفكار. إن اختبارات الخيارات المتعددة لا تطلب من الطالب استنباط معنى أو إنتاج فكرة، بل اختيار إجابة ما وكتابة الرقم أو الحرف الذي يشير إليها باستدعائها من ذاكرته أو استنتاجها من خلال صياغة البدائل وفي بعض الأحيان تأتي الإجابة الصائبة إما تخمينا أو خبط عشواء.

ثانيا: وجهات نظر الطلاب:

لا غرو أن تجد اختبارات الخيارات المتعددة شعبية كبيرة وسط الطلاب لما توفره من سهولة في الاستعداد والتنفيذ. وتأكدت هذه الشعبية من خلال نتائج استطلاع تم إجراؤه على عينة عشوائية من طلاب وطالبات قسم الخدمة الاجتماعية بكلية العلوم الاجتماعية بجامعة أم القرى شملت (١٧٠) طالبا وطالبة للتعرف على آرائهم حول بعض جوانب اختبارات الخيارات المتعددة ومدى تفضيلهم لها على أشكال الاختبارات الأخرى. وقد تم توزيع العينة كما هو موضح في الجدول رقم (١) أدناه:

جدول رقم (١) عينة استطلاع آراء الطلاب والطالبات حول اختبارات الخيارات المتعددة

المستوى	طـــــلاب		طالبـــات		المجمـــوع	
	العدد	% من العينة	العدد	% من العينة	العدد	%
المستوى الأول	٢٠	%١١,٨	٣٠	%١٧,٧	٥٠	%٢٩,٤
المستوى الثالث	٢٠	%١١,٨	٣٠	%١٧,٧	٥٠	%٢٩,٤
المستوى الخامس	١٥	%٨,٨	٢٠	%١١,٧	٣٥	%٢٠,٦
المستوى السابع	١٥	%٨,٨	٢٠	%١١,٧	٣٥	%٢٠,٦
المجموع	٧٠	%٤١,٢	١٠٠	%٥٨,٨	١٧٠	%١٠٠

وكما توضح بيانات الجدول فقد تم اختيار العينة من طلاب المستويات الفردية لأن الدراسة أجريت خلال الفصل الأول من العام الدراسي والذي يكون فيه الطلاب جميعهم في المستويات الأول والثالث والسابع والخامس، ما عدا فئة قليلة من الطلاب غير الملتزمين بخطة الدراسة لأسباب مختلفة. ومثل الطلاب (٤١,٢%) من مجموع العينة بينما مثلت الطالبات (٥٨,٨%) لأن عدد الطالبات بالقسم يفوق عدد الطلاب، لذا روعي هذا الأمر عند اختيار العينة.

وتم توزيع استبيان اشتمل على جزءين: الجزء الأول للبيانات الأولية عن أفراد العينة والتي شملت الجنس والمستوى الدراسي والمعدل التراكمي وطريقة الالتحاق بالقسم، والجزء الثاني تضمن مجموعة من البنود التي تستفسر عن سلبيات وإيجابيات اختبارات الخيارات المتعددة من وجهة نظر الطلاب وضعت على مقياس "ليكرت" الخماسي (Likert, 1932).

وجاءت نتائج الاستطلاع على النحو التالي:

(أ) : اتضح أن الشكل المفضل للاختبارات لغالبية الطلاب والطالبات هو اختبارات الخيارات المتعددة حيث أشار غالبية أفراد العينة إلى تفضيلهم لهذا الشكل على الأشكال الأخرى كما يتضح من بيانات الجدول رقم (٢) أدناه.

جدول رقم (٢) الشكل المفضل للاختبارات

نوع الاختبار	التكرار	النسبة المئوية	النسبة المئوية الحقيقية	النسبة المئوية التراكمية
الشفهي	٧	٤,١	٤,١	٤,١
المقالي	٣٥	٢٠,٦	٢٠,٦	٢٤,٧
اختبار الخيارات المتعددة	١٢٨	٧٥,٣	٧٥,٣	١٠٠
المجموع	١٧٠	١٠٠	١٠٠	

يوضح الجدول التكراري أعلاه أن الغالبية العظمى من الطلاب والطالبات يفضلون اختبارات الخيارات المتعددة على غيرها من أنواع الاختبارات كالشفهية والمقالية حيث أشار (٧٥,٣%) منهم إلى ذلك. ومن الطبيعي أن يكون الدافع الأول للطلاب لاختيار نوع ما من الاختبارات هو الرغبة في النجاح وليس قياس المستوى، حيث يرى الطلاب والطالبات أن هناك العديد من المزايا في اختبارات الخيارات المتعددة والتي تيسر لهم أمر النجاح.

(ب): تضمن الاستبيان الذي وزع على الطلاب الاستفسار عن آرائهم حول أهم إيجابيات اختبارات الخيارات المتعددة. وباستخدام الإحصاء الوصفي لبنود هذا المحور وترتيبها تنازليا اتضح أن أبرزها يتمثل في ما هو وارد في الجدول رقم (٣) أدناه.

جدول رقم (٣) أبرز إيجابيات اختبارات الخيارات المتعددة من وجهة نظر الطلاب

البنـــــود	العدد	الحد الأدنى	الحد الأعلى	الوسط الحسابي	الانحراف المعياري
تستبعد تأثير المقدرة على التعبير والكتابة والخط على الدرجة	170	1	5	4.33	.90
تساعد على النجاح لسهولتها	170	1	5	4.25	1.03
درجة التوتر فيها أقل مقارنة بالاختبارات الأخرى	170	1	5	4.15	1.02
لا تحتاج لوقت طويل للمذاكرة والإعداد لها	170	1	5	4.08	1.20
تساعد على تجنب التحيز في التصحيح	170	1	5	3.95	1.19
رصد درجاتها يمتاز بالدقة	170	1	5	3.74	1.02
لا تحتاج لوقت طويل لأدائها	170	1	5	3.73	1.10

وكما هو واضح من بيانات الجدول فقد جاء على رأس قائمة الميزات ميزة استبعاد اختبار الخيارات المتعددة لتأثير المقدرة على التعبير والكتابة والخط على الدرجة، حيث بلغ متوسطها الحسابي (٤,٣٣) بانحراف معياري بلغ (٠,٩)، تليها ميزة سهولة هذا النوع من الاختبارات لاعتماده على الحفظ والاسترجاع في المقام الأول. كذلك تضمنت قائمة الإيجابيات مجموعة من الاعتبارات مثل انخفاض درجة التوتر فيه مقارنة مع غيره وعدم الحاجة للمذاكرة لساعات طويلة للاستعداد له أو لأدائه، والثقة العالية في التصحيح ورصد الدرجات حيث تراوح الوسط الحسابي لهذه البنود ما بين (٤,٢٥) و (٣,٧٣) بانحرافات معيارية متقاربة.

(ج): تضمن الاستبيان الذي وزع على الطلاب الاستفسار عن آرائهم حول أهم سلبيات اختبارات الخيارات المتعددة. وباستخدام الإحصاء الوصفي لبنود هذا المحور وترتيبها تنازليا اتضح أن أبرزها يتمثل في ما هو وارد في الجدول رقم (٤) أدناه.

جدول رقم (٤) أبرز سلبيات اختبارات الخيارات المتعددة من وجهة نظر الطلاب

البنـــــود	العدد	الحد الأدنى	الحد الأعلى	الوسط الحسابي	الانحراف المعياري
لا تعطيني الفرصة للتعبير عما أعرفه	170	1	5	3.88	1.15
الغش سهل في هذا النوع من الاختبارات	170	1	5	3.82	1.31
أشعر بأنها لا تقيس مستواي في المقرر بصورة دقيقة	170	1	5	3.75	1.23
تقلل من هيبة الاختبارات	170	1	5	3.50	1.18
تقلل من دافعية الطالب للاجتهاد	170	1	5	3.45	1.29
يمكن النجاح فيها عن طريق التخمين والخبط العشوائي	170	1	5	3.42	1.36
تحتاج إلى تدرب الطالب على استخدامها	170	1	5	2.86	1.43

جاءت سلبية عدم إعطاء اختبار الخيارات المتعددة الفرصة للطالب للتعبير عما يعرفه كأبرز السلبيات، وكادت سلبية سهولة الغش أن تتساوى معها في الترتيب حيث بلغ المتوسط الحسابي (٣,٨٨) و (٣,٨٢) لكل منهما على التوالي، بانحراف معياري بلغ (١,١٥) للسلبية الأولى و (١,٣١) للسلبية الثانية. أما بقية السلبيات فقد تضمنت شعور الطلاب بعدم دقتها في قياس مستوياتهم العلمية وتقليلها من هيبة الاختبارات ومن دافعية الطالب للاجتهاد وإمكانية النجاح فيها عن طريق الغش والاختيار العشوائي. أما سلبية احتياج الطالب إلى التدرب على استخدامها فقد جاءت في ذيل القائمة لتعود الطلاب على أدائها ومنذ مراحل تعليمية سبقت المرحلة الجامعية.

(د) : تم استخدام اختبار "ت" لتحليل التباين بين الطلاب والطالبات للتعرف على دلالات الفروق بينهم بسبب اختلاف الجنس حول إيجابيات وسلبيات اختبارات الخيارات المتعددة. ولم يوضح الاختبار أي فروق بين الجنسين حيث لم يصل مستوى الدلالة إلى ما يساوي أو أقل من (٠,٠٥) لأي منهما.

(هـ) : لم تكن هناك فروق ذات دلالة إحصائية بين الطلاب والطالبات حول إيجابيات اختبارات الخيارات المتعددة حسب المستوى الدراسي ، ولكن اتضح وجود فروق ذات دلالة إحصائية بينهم حول السلبيات تتضح من بيانات الجدول رقم (٥) أدناه.

جدول رقم (٥) دلالات الفروق حول سلبيات وإيجابيات اختبارات الخيارات المتعددة حسب المستوى الدراسي

		مجموع المربعات	درجة الحرية	متوسط المربعات	قيمة "ف"	مستوى الدلالة
الإيجابيات	بين المجموعات	.184	3	6.134	.330	.804
	داخل المجموعات	30.848	166	.186		
	المجموع	31.032	169			
السلبيات	بين المجموعات	11.235	3	3.745	20.146	.000
	داخل المجموعات	30.857	166	.186		
	المجموع	42.091	169			

تم استخدام اختبار "ف" لتحليل التباين حسب المستوى الدراسي واتضح منه عدم وجود فروق حول محور الإيجابيات حيث لم تزد قيمة "ف" عن (٠,٣٣). أما بالنسبة للسلبيات فقد اتضح وجود فروق حولها بصورة واضحة حيث بلغت قيمة "ف" (٢٠,١٤٦) وقل مستوى الدلالة كثيرا عن (٠,٠٥). وباستخدام اختبار "توكي" (Tukey) البعدي (post hoc multiple comparisons) اتضح أن مصدر الفروق يتمثل في الاختلاف الواضح في وجهات النظر بين طلاب وطالبات المستوى الأول مع طلاب وطالبات المستويات الأخرى – الثالث والسابع والخامس. وربما يعود السبب إلى حداثة عهد طلاب المستوى الأول باختبارات الخيارات المتعددة على المستوى الجامعي حيث تم إجراء هذا الاستطلاع خلال الفصل الأول من دراستهم الجامعية.

وتؤكد هذه النتائج ما توصلت إليه مختلف الدراسات عن سلبيات وإيجابيات اختبارات الخيارات المتعددة كما تدعم فرضية شعبيتها بين طلاب العلوم الاجتماعية وتفضيلهم لها على سائر أشكال الاختبارات الأخرى لما يمكن أن توفره لهم من سهولة للنجاح في مقرراتهم المدروسة.

صحيح أن الطلاب يركزون على إيجابيات هذه الاختبارات عند تفضيلهم لها، دون أن يتأثر هذا التفضيل بسلبياتها التي تتركز أساسا في طبيعتها ومقدرتها على التقويم الشامل ومشكلات إعدادها وصباغة بنودها، لأنها لا تؤثر في نتائجهم أو تقلل من إمكانية نجاحهم بصورة مباشرة. بل أن بعضهم يرى أن بعض السلبيات – كالغش وإمكانية التخمين – من العوامل التي يمكن أن توظف لتحقيق النجاح.

لقد فرضت اختبارات الخيارات المتعددة نفسها كوسيلة هامة من وسائل التقويم في العلوم الاجتماعية، لذا ينبغي النظر إلى أبرز مشكلاتها والسعي لحلها وتطوير هذه الوسيلة واستخدامها مع وسائل أخرى تسد ثغراتها والتكامل معها لتقويم حصيلة الطالب من كافة جوانبها.

ثالثا: قواعد بناء البنود وأهم أشكال مخالفاتها:

لعل أبرز أوجه قصور اختبارات الخيارات المتعددة هو صعوبة صياغة الأسئلة الجيدة واستخدام أسئلة ضعيفة البناء تؤدي إلى قياس خاطئ للحصيلة المعرفية في المجال المدروس. ولكن على الرغم من ذلك يمكن في العلوم الاجتماعية استخدام اختبارات الخيارات المتعددة لقياس بعض جوانب المحصلة التعليمية للمقرر قياسا جيدا والتعرف على مدى تحقيقه لبعض

أهدافه إذا أعدت بصورة سليمة والتزمت بشروط بناء البنود السليمة التي سيرد ذكرها فيما بعد لتوفير نتائج دقيقة وموضوعية وتحقق درجة عالية من الثبات.

ومن الضروري أن يكون الاختبار مكتوبا بصورة سليمة حتى يكون التقويم المبني عليه موضوعيا وصادقا. ولا توجد حتى الآن "قواعد عامة" متفق عليها لبنائها حيث أنها تعتمد في كل حالة على الخبرة والحكمة أكثر من البحث، فعلى سبيل المثال لم يجد كل من "كريهان" و "هلادينا" (Crehan & Haladyna, 1991, p. 183). أي دلائل تدعم الدعوة إلى وضع الجذر بصيغة سؤالية بدلا من شكل إكمال الجملة، كما قام "فيولاتو" (Violato 1991) بمقارنة أسلوب كتابة الجذر في شكل جملة كاملة مقارنة بأسلوب استكمال الجملة ولم يكتشف أي فروق بين الأسلوبين في درجة الصدق أو الصعوبة.

وفي دراستهم عن استخدام عبارة "لا شيء مما ورد أعلاه" باستخدام أسلوب القياس المتكرر ونماذج مختلفة من أسئلة الاختبار وجد "كريهان" وزملاؤه أن استخدام هذه العبارة يزيد من صعوبة الاختبار ولكنه لا يفيد في تفسير درجة التباين في الدرجات (Crehan, Haladyna & Brewer, 1993)، كما وجدوا في نفس الدراسة عدم وجود اختلاف في تفسير درجة التباين عند استخدام ثلاثة خيارات بدلا من أربعة خيارات للجذر الواحد، على الرغم من أن استخدام ثلاثة خيارات فقط يقلل من صعوبة الاختبار، كما أنه يوفر بعض الوقت في وضع وإدارة الاختبار ويساهم في كتابة بدائل أفضل.

ويشترط في السؤال الجيد أن يكون واضحا ومفهوما للجميع وأن تكون بدائله مبنية بطريقة لا تكشف عن الإجابة الصحيحة. وهناك العديد من المفكرين الذين قدموا أمثلة متنوعة للأسئلة الجيدة والأسئلة السيئة. فعلى سبيل المثال يورد "جرونلند" (Gronlund: 1993) في الفصل الثالث من كتابه عن تقويم الحصيلة العلمية قائمة تتكون من ثمان عشرة قاعدة لكتابة أسئلة الخيارات المتعددة. من ناحية أخرى قام "هالادينا" و "داوننج" (Haladyna and Downing :1989) بمراجعة ست وتسعين دراسة تضع ثلاثة وأربعين شرطا للبناء السليم للبدائل.

أما "أيكن" (Aiken: 1987) فقد راجع العديد من الأعمال في مجال التقويم وقدم أربعة عشر توجيها لبنائها، بينما يقدم "إيلسورث" وآخرون (Ellsworth et al.: 1990) اثنتا عشرة توصية جاءت ضمن أبرز التوصيات للبناء السليم لتلك الأسئلة وردت ضمن ثلاث وثلاثين مرجعا قاموا باستعراضها؛ وتتلخص توجيهات "إيلسورث"وزملائه في الآتي:

١- حدد سؤالا أو مشكلة أو موقفا واحدا في الجذر، وتجنب الأسئلة المركبة التي تنطوي على أكثر من سؤال.

٢- اكتب الجذر بلغة سليمة واضحة وتجنب التكرار والإضافات والمترادفات غير الضرورية.

٣- اكتب الجذور في جمل قصيرة حيث أنها أيسر استيعابا من الجمل الطويلة وأكثر دقة.

٤- اكتب الجذر بجمل مثبتة بقدر الإمكان، فمعرفة الطالب لما هو مثبت أفضل من معرفته لما هو غير ذلك. وتستخدم البنود المنفية في بعض الأحيان لسهولة توفيرها.

٥- وضح الكلمات التي تدل على النفي بتسويدها أو كتابتها بخط مائل أو وضع خط تحتها في حال استخدامها في الجذر لأن مثل هذه الكلمات يمكن أن تمر دون أن يلاحظها الجالس للاختبار.

٦- يجب التأكد من صحة الإجابة الصحيحة أو الأفضل لتحسين جودة الاختبار وتجنب الجدل حولها مع الطلاب.

٧- يجب أن تكون البدائل متسقة نحويا مع الجذر ومتوازية معه شكليا، لأن عدم ذلك يوفر تلميحات للطالب عن الإجابة الصحيحة ويمكنه من استبعاد البدائل غير المتسقة.

٨- ينبغي تجنب التلميحات اللفظية التي يمكن أن تساعد على استبعاد أو اعتماد بدائل معينة، ومن وسائل ذلك ما يأتي:

أ- تجنب استخدام كلمات متشابهة في الجذر وفي الإجابة الصحيحة لأن ذلك يجعل الإجابة التي تتضمن كلمات مشابهة لتلك الموجودة بالجذر أكثر جاذبية ليختارها الطالب الذي يجهل الإجابة الصحيحة.

ب- يجب ألا تكون الإجابة الصحيحة أكثر تفصيلا عن غيرها بصورة ملفتة وألا تحتوي على كلمات من كتاب المقرر أكثر من غيرها من البدائل.

ج- تجنب استخدام العبارات المطلقة في الجذور، فالطلاب الذين يستخدمون ذكاءهم في الاختبار أكثر من حصيلتهم المعرفية (test-wise students) يستبعدون البدائل التي تتضمن كلمات مثل "كل" أو "فقط" أو "مطلقا" أو "أبدا" لأن مثل هذه العبارات عادة ما تكون غير صحيحة.

د- تجنب استخدام العبارات التي تحمل نفس المعنى لأن الطالب سوف يستبعدها نظرا لوجود إجابة صحيحة واحدة فقط.

هـــ إذا تضمنت البدائل أزواجا من الإجابات فينبغي تجنب البناء الذي يمكن أن تستنتج منه الإجابة الصحيحة. (فإذا كانت الإجابة الصحيحة على سؤال ما هي (س) وأيضا (ص) فقد توضع البدائل على النحو التالي: ١- (س) و (ص) ، ٢- (س) و (ن) ،٣- (ج) و (ص) ،٤- (ك) و (ل). وسوف يميل الطالب الذي لا يعرف الإجابة الصحيحة إلى البديل الأول نظرا لأهمية كل من (س) و (ص) المضمنة في تكرارهما.

٩- يجب جعل البدائل جميعها مقبولة كإجابات محتملة.

١٠ تجنب استخدام صيغة "كل ما سبق" لأن الطالب يمكنه استبعاده لمعرفته لإجابة خاطئة واحدة فقط، مما يجعل إمكانية تخمين الإجابة الصحيحة من البديلين المتبقيين عالية.

١١- استخدم صيغة "لا شيء مما سبق" بحذر، لأن ذلك يمكن أن يقيس فقط المقدرة على تحديد الإجابات غير الصحيحة. وعادة ما يستخدم واضع الاختبار هذه الصيغة عند عدم مقدرته على صياغة بديل آخر يكون بنفس الدرجة من القبول كالبدائل الأخرى.

١٢- لا بد من الاهتمام بالقوانين العامة للترقيم والقواعد النحوية

مدى الالتزام بالتوجيهات وأهم المخالفات:

استخدم الباحث هذه التوجيهات في تحليل أسئلة تم اختيارها عشوائيا من اختبارات منتصف الفصل الدراسي ونهايته في مقررين يقدمان ضمن متطلبات كلية العلوم الاجتماعية بجامعة أم القرى لدرجة البكالوريوس بأقسامها الخمسة استخدمت فيها اختبارات الخيارات المتعددة (باستثناء التوجيهين الرابع والخامس الخاصين بـ "توضيح الكلمات التي تدل على النفي"، و "التأكد من صحة الإجابة" وذلك لعدم ورود جذور منفية ضمن الأسئلة المدروسة وارتباط التوجيه الآخر بمعد الاختبار أكثر من ارتباطه بالاختبار نفسه.

وكانت الاختبارات التي تم تقويمها قد عقدت لطلاب وطالبات المستوى الثالث خلال الفصل الدراسي الأول من العام الدراسي ٢٠٠٦/٢٠٠٧ في منتصف الفصل الدراسي وفي نهايته. وتم الحصول على نسخ من تلك الاختبارات بعد موافقة أساتذة تلك المقررات ومعرفتهم للغرض منها، بلغ عددها ثمانية اختبارات من المقررين وضعها ثمانية من مدرسي ومدرسات تلك المقررات للطلاب والطالبات وتضمنت أربعة اختبارات نصفية وأربعة اختبارات نهائية بواقع اختبارين لكل مقرر، وتم اختيار (١٠) أسئلة من كل اختبار للتحليل حيث بلغ عدد الأسئلة التي تم تحليلها (١٦٠) سؤالا.

كشفت عملية تقويم الاختبارات التي تم اختيارها عن وجود مخالفات عديدة للموجهات العشر الواردة أعلاه . وقد بلغ عدد المخالفات (٢٢٧) مخالفة، أي (١٤,٢%) من جملة المخالفات المحتملة التي تبلغ (١٦٠٠) مخالفة محتملة (١٦٠ سؤالا × ١٠ توجيهات لم تنفذ). الجدول رقم (١) أدناه يوضح تلك المخالفات مرتبة تنازليا حسب نسبة ورودها.

جدول رقم (١) عدد المخالفات في التوجيهات العشر ونسب ورودها

#	التوجيهـــــات	عدد المخالفات	النسبة المئوية	النسبة المئوية التراكمية
١	يجب أن تكون البدائل متسقة نحويا مع الجذر	٤٣	١٨,٩٤	١٨,٩٤
٢	اجعل البدائل جميعها مقبولة كإجابات محتملة	٤٣	١٨,٩٤	٣٧,٨٨
٣	اهتم بالقوانين العامة للترقيم والنحو	٤٢	١٨,٥٠	٥٦,٣٨
٤	اكتب الجذر بلغة سليمة وتجنب التكرار والمترادفات غير الضرورية.	٣١	١٣,٦٦	٧٠,٠٤
٥	تجنب التلميحات اللفظية التي قد تساعد على استبعاد أو اعتماد بدائل معينة	٢٣	١٠,١٣	٨٠,١٧
٦	اكتب الجذور في جمل قصيرة	٢٠	٨,٨١	٨٨,٩٨
٧	تجنب استخدام صيغة "كل ما سبق"	١٢	٥,٢٩	٩٤,٢٨
٨	تجنب استخدام صيغة "لا شيء مما سبق"	١١	٤,٨٥	٩٩,١٢
٩	حدد سؤالا أو موقفا واحدا في الجذر	١	٠,٤٤	٩٩,٥٦
١٠	اكتب الجذر بجمل مثبتة	١	٠,٤٤	١٠٠
	المجموع	٢٧٧	١٠٠	

يلاحظ من بيانات الجدول أعلاه أن مخالفة عدم اتساق البدائل نحويا مع الجذر وعدم التمكن من جعل البدائل جميعها مقبولة كإجابات محتملة هما أكثر المخالفات تكرارا حيث مثلت كل منهما (١٨,٩٤%) من مجموع المخالفات، وتقاربت معهما مخالفة عدم الاهتمام بالنحو وقوانين الترقيم التي بلغت (١٨,٥٠%) من مجموع المخالفات. وتأتي بعد ذلك مشكلات تدني مستوى اللغة والتكرار واستخدام المترادفات غير الضرورية وتوفير

التلميحات اللفظية التي قد تساعد على استبعاد أو اعتماد بدائل معينة مثل استخدام كلمات متشابهة في الجذر وفي الإجابة الصحيحة والتفصيل الزائد للإجابة الصحيحة واستخدام العبارات المطلقة في الجذور والعبارات التي تحمل معان متشابهة. ومثلت هاتان المخالفتان (٢٣,٧٩%) من مجموع المخالفات الواردة في الأسئلة المدروسة. أما بقية المخالفات فقد تراوحت ما بين استخدام صيغة "كل ما سبق" أو صيغة "لا شيء مما سبق" أو وضع جذر مركب يحتوي على أكثر من سؤال أو كتابة الجذور بعبارات منفية، ومثلت نسبة قليلة من المخالفات.

وشملت المخالفات (٤٠%) من أسئلة الاختبار النصفي في المقرر الأول و(٦٠%) من أسئلة الاختبار النصفي في المقرر الثاني بالنسبة للطلاب. كما بلغت نسبة المخالفات في الاختبار النهائي (٢٥%) من أسئلة المقرر الأول للطلاب و (٥٥%) من أسئلة المقرر الثاني.

أما بالنسبة للطالبات فقد شملت المخالفات (٦٠%) من أسئلة الاختبار النصفي في المقرر الأول و(٧٥%) من أسئلة الاختبار النصفي في المقرر الثاني، بينما بلغت نسبة المخالفات في الاختبار النهائي (٤٥%) من أسئلة المقرر الأول و (٦٠%) من أسئلة المقرر الثاني، وذلك كما هو موضح في الجدول رقم (٢) أدناه.

مهمة وضع اختباراتهما على الرغم من حداثة عهدهن بوضع مثل تلك الاختبارات.

وتتفاوت المخالفات التي وردت في الاختبارات المدروسة بين البسيطة التي يمكن تعديلها والكبيرة التي لا تعكس ضعفا في وضع الأسئلة فحسب، وإنما تشير أيضا إلى خلل في مفهوم التقويم وأهدافه وما ينبغي تقويمه في مثل هذه العلوم، وبالتالي خلل في عملية تدريسها نفسها باعتبار أن التدريس والتقويم عمليتان متكاملتان، وأن أولويات التقويم هي بمثابة المؤشرات لأولويات المقرر وأهدافه وما تم تحقيقه منها. إن تدني الكفاءة في تقويم وتدريس مثل هذه المقررات المدخلية (introductory courses) لا يحدث من وجود عجز في أعضاء هيئة التدريس المؤهلين لذلك فحسب، وإنما أيضا لعدم رغبة المؤهلين منهم في تدريسها على الرغم من إيمانهم بأهمية الدور الذي تلعبه تلك المقررات في بناء قاعدة صلبة يقوم عليها التخصص المدروس أو المجال الذي تدرس في إطاره. فبالإضافة إلى أهميتها في إرساء القاعدة الضرورية للانطلاق في رحاب مجال علمي ما أو استيفاء متطلبات الالتحاق بمجالات علمية أخري، فهي أيضا ضرورية للوصول إلى الآخرين الذين لم تكن لديهم الفكرة أو الرغبة في هذا المجال والتي قد تتغير عندما تقدم لهم صورة سليمة عنها من قبل المؤهلين لذلك.

وهناك أهمية أخرى للمقررات المدخلية في العلوم الاجتماعية تكمن في توفيرها لفوائد أخرى لدارسيها باعتبارهم من عامة أبناء المجتمع. فمفاهيم السلوك الإنساني والعرق ودناميات التفاعل الاجتماعي وقضايا البيئة ومشكلات الأسرة والتحول الاجتماعي على سبيل المثال تصب في صميم

اهتمامات كافة قطاعات المجتمع المعاصر، وهي قضايا تقدمها مقررات العلوم الاجتماعية بكافة مستواها وتتطلب مهارة وخبرة في تناولها وتقديها، خاصة لمن هم في أول عهدهم بهذا المجال. أضف إلى ذلك أن كثيرا من المؤلفات التي أنتجها غير المختصين في العلوم الاجتماعية تتطلب قراءة ناقدة توفرها العلوم الاجتماعية من خلال مناهجها التي تتميز بالشمول والعمق وبعد النظر، وبتزويدها للقارئ بما يمكنه من الفهم السليم للقضايا المطروحة واستيعابها من مختلف جوانبها. إن الأسلوب المتبع الذي يقوم على الاعتماد على طلاب الدراسات العليا والمتعاونين من خارج الأقسام العلمية ومساعدي التدريس لا يساعد على تحقيق تلك الأهداف، على الرغم من وجود الأسباب القوية التي تستوجب اللجوء إلى هذا الخيار في بعض الأحيان. فلا بد إذن من وجود أعلى المستويات الأكاديمية داخل قاعات الدراسة، ليس لتدريس من هم على أعتاب التخرج وطلاب الدراسات العليا فحسب، بل أيضا لتدريس المبتدئين والمستجدين في المقررات المدخلية والمبتدئة.

رابعا: شروط توظيف اختبارات الخيارات المتعددة لتقييم المستويات العليا للحصيلة المعرفية فى العلوم الاجتماعية :

يتطلب استخدام اختبارات الخيارات المتعددة في مجال العلوم الاجتماعية حذرا شديدا نظرا لما تتميز به الظواهر الاجتماعية من تداخل وتغير سريع، حيث يتطلب بناؤها جهدا ووقتا كبيرين لتخرج بالصورة المناسبة التي تؤهلها للقياس الصادق مع بقية أشكال القياس والتقويم الأخرى.

ولتحقيق ذلك فلا بد من الالتزام بشروط التقويم الفعال والتي تتضمن ثلاثة جوانب وهي: (أ) التعرف على ما ينبغي قياسه من الحصيلة المعرفية، و(ب) عدم الاقتصار على أسلوب واحد و (ج) استخدام أسئلة قياس التفكير الناقد.

(أ) التعرف على ما ينبغي قياسه من الحصيلة المعرفية

يتطلب المجال المعرفي في العلوم الاجتماعية أن يقوم الطلاب بمعالجة بيانات كثيرة ومعقدة والتعامل معها بمسؤولية ومهارة، وعليه فلابد من كتابة أسئلة اختباراتها بحيث تعبر عن مستوى تلك المسئولية والمهارة التي يتوقع من الطلاب ممارستها. ولا يتوقع من الطلاب الذين يجلسون لاختبارات لا تهتم إلا بمستواهم المعرفي الاسترجاعي فقط أن يمارسوا مهنتهم في شتى مجالات الممارسة المهنية والاجتماعية بالكفاءة المطلوبة لأنها تتطلب مستويات أعلى من الفكر والمعرفة والمهارة، وعليه فلا بد من أن يتطابق التقويم المتمثل في أسئلة الاختبار مع توقعات التعلم والمهارات المطلوب اكتسابها. وبظهور التيار السلوكي وهيمنته على التعليم أهدافا وممارسة ، تحول الاهتمام من التركيز على تقويم المعارف إلى التركيز على المهارات والقدرات والتمكن والتملك. ويعمل هذا التوجه على تقسيم المواد التعليمية إلى محتويات معرفية تسلط عليها قدرات ومهارات ظاهرة ملموسة محسوسة تمثل الناتج الفعلي للتعليم والتعلم ، إذ أن الغاية من التعليم لدى السلوكيين هي إدخال تعديل محسوس ملموس قابل للقياس على سلوك المتعلم بتزويده بمهارات وقدرات مقصودة في علاقة بمعارف محددة. ومنذ أواخر الأربعينيات وأوائل الخمسينيات من القرن المنصرم أولى بعض المهتمين بأمر التقويم من أولئك السلوكيين هذا الأمر

اهتماما كبيرا، وكان هدفهم هو توفير قاعدة يحددون ما تعلمه طلابهم على ضوئها. وكانت القاعدة التي يتم التقويم بناء عليها آنذاك محصورة فيما جمعه الطلاب من معلومات يستطيعون استرجاعها متى ما سئلوا عنها، علما بأن ذلك الأمر كان من أقل الأمور التي يسعى المقومون للتعرف على وجودها في حصيلة الطالب.

وهذا ما سعى "بنيامين بلوم" وزملاؤه (Bloom, 1956) لتأكيده في صنافتهم الشهيرة لأهداف التعليم والتي تضمنت ثلاث مجالات للأنشطة التعليمية وهي: المجال الإدراكي (Cognitive domain) والذي يتضمن المعرفة وتنمية الاتجاهات والمهارات المعرفية، والمجال الانفعالي (أو الوجداني) (Affective Domain)، والمجال النفسحركي (أو المهاري) (Psychomotor Domain). وقد قسم "بلوم" المجال الإدراكي إلى عدة تقسيمات تتدرج من أبسط أشكال السلوك إلى أكثرها تعقيدا وهي: المعرفة (التذكر)، والفهم، والتطبيق، والتحليل، والتركيب، والتقويم Bloom,1956: (26).

وقد رأى بلوم وزملاؤه في هذه الصنافة أكثر من مجرد وسيلة للمساعدة على القياس الفعال لما تم تحقيقه من أهداف تعليمية، حيث أنها تمثل لغة مشتركة حول أهداف التعلم تساعد على التواصل بين الأفراد والتخصصات المختلفة، وترسي قاعدة عامة لتحديد عمق وفعاليات وآليات تقويم المناهج عند بنائها. وقد وجدت هذه الصنافة رواجا كبيرا وترجمت إلى أكثر من ٢٢ لغة وما زالت تستخدم حتى يومنا هذا في مختلف المؤسسات التعليمية (Krathwohl,2002 :18). إلا أن الواقع الذي تعيشه العلوم الاجتماعية فيما

يتعلق بهذه الأهداف يجعلها مجرد ركام معرفي خارج نطاق التطبيق لما تعانيه هذه العلوم من إهمال للمهارات واختزال لأهداف التحليل والتركيب والتقويم في المحتوى المعرفي الذي يقدس الحفظ والاسترجاع.

(ب) عدم الاقتصار على أسلوب واحد

تعتبر الاختبارات الحرة كالاختبارات المقالية وغيرها أفضل من اختبارات الخيارات المتعددة في قياس مختلف جوانب التحصيل والكفاءة والمهارة (Kubiszyn & Borich 1996, p. 109). وكما يشير "ويلش" و "ساوندرز" (Welsh and Saunders, 1990, pp. 201-210) فإن اختبارات الخيارات المتعددة وعلى الرغم من إمكانية تصميمها بدرجة عالية من الجودة تضمن حفزها لعمليات فكرية متقدمة لدى الطالب وقياسها لمهارات الطالب، لكنه لا يمكن التحقق من خلالها من مدى معرفة الطالب لاستخدام تلك المهارات. ويعترف "باري" و "طومسون" (Barry and Thompson 1997, p. 305) بما توفره اختبارات الخيارات المتعددة من توفير للوقت والجهد للمصحح على حساب شمولية التقييم ويقدمان بعض المقترحات التي توفر وقتا أطول عند تصحيح الاختبارات الأفضل في تقويمها للطالب.

كثير من أسئلة مختلف مجالات العلوم الاجتماعية يمكن أن تكون أسئلة تطبيقية، خاصة بالنسبة لمجالات مثل علوم الاقتصاد والاجتماع والخدمة الاجتماعية الدراسات السكانية ومناهج البحث الاجتماعي. وما من شك في أن الأسئلة التطبيقية المصاغة جيدا تشجع الطالب على تأطير معارفه لأنه عليه أن يطبق مفهوما ما على أرض الواقع. ولكنها على أية حال لا تقيس بالضرورة أي

عمليات فكرية أعلى لأن مشكلة الطلاب لا تكمن في التعامل الآلي مع الرموز والمفاهيم ولكن مع الفهم التصوري الذي تتضمنه والذي لا يمكن قياسه بصورة صادقة إلا باستخدام أسئلة الاستجابة الحرة.

وتستخدم أسئلة الاستجابة الحرة لتقويم الطلاب من خلال المقالات والأسئلة القصيرة والعروض والتجارب والمقابلات والتقارير الشفهية والمشاريع البحثية، حيث يحاول الطالب من خلال هذه الاختبارات تصميم وبناء إجاباته بنفسه. ويؤخذ على هذا الشكل من الاختبارات أنه يستغرق وقتا أطول في تصحيحه مقارنة بالاختبارات الموضوعية وأنه أقل ثباتا منها وأقل تغطية لمحتوى المنهج.

ويقترح "وود" (wood, 1998) شكلا بديلا لاختبار الخيارات المتعددة يجمع ما بين سهولة التصحيح وقوة قياس اختبارات الاستجابة الحرة أطلق عليه اسم "الخيارات المتعددة المرتبطة" (Linked multiple choice). والفكرة هي أن يقوم مصمم الاختبار بوضع أسئلة مفتوحة تتضمن عناصر مغلقة. فمثلا يمكن سؤال الطالب عن تعريف معين لأحد المفاهيم وتطبيق ذلك المفهوم. فالتعريف مغلق على الرغم من أنه جزء من سؤال مفتوح.

وإحدى مشكلات هذا الأسلوب هي أن بعض الطلاب الذين ينقصهم الاستعداد للاختبار يقومون باختلاق الإجابة على السؤال المفتوح الذي يواجههم. ومثل هذه الإجابة يسهل التعرف عليها وتقييمها بناء على ذلك . وتصبح المهمة أكثر صعوبة عندما يحاول نفس الطالب الاستفادة من بيانات خيارات السؤال المرتبط لاختلاق إجابة تبدو وكأنها معقولة. وعلى أي حال

البراعة في التفكير الناقد يمكن تطويرها في إطار نظام علمي محدد . ويؤكد "ماكبيك" (McPeck, 1981) أن الذين يتمتعون بتفكير ناقد يقيمون الأوضاع على ضوء الخلفية العلمية والسياق والتشكك التأملي ويسلم بأنه من المستحيل تدريس التفكير الناقد بمعزل عن المهارات التي تدرس للطلاب ، كما يشير أيضا إلى أن مهارات التفكير الناقد تقتصر على مجالات معينة وجوانب ضيقة من جوانب التطبيق (McPeck, 1981: 27). من ناحيه أخرى يرى البعض أن التفكير الناقد يمكن أن يكون أمرا عاما كما يمكن أن يرتبط بسياق محدد (Bandman and Bandman, 1995).

إن التفكير الناقد في العلوم الاجتماعية أمر سياقي يتطلب معرفة خاصة ترتبط بمجال علمي محدد، وعليه فينبغي قياسه في إطار ذلك المجال المحدد، والشخص الذي يتمتع بالتفكير الناقد هو الذي يتمكن من استخدام ما تعلمه من مفاهيم محددة في مجال محدد لمجابهة مختلف بدائل الموقف الذي يجد نفسه فيه ضمن ذلك المجال. لقد تحدث الكثيرون عن طرق التدريس التي تساعد على تنمية الفكر الناقد ومنهم "كيس" (Case, 1994) الذي يرى أن تنميته في الدارسين تتطلب حوارا جادا بين الطالب والمدرس وتشجيعا مستمرا للطالب على الاستبطان والتأمل.

وهناك العديد من الأنشطة التي تساعد في تنمية التفكير الناقد مثل تشجيع الطلاب على الكتابة والنشر واستخدام مجموعات النقاش التي تستعرض حلولا ممكنة متنوعة لشتى المواقف والتوصل إلى نتائج وقرارات حولها، وطرح القضايا التي تتساءل حول مختلف البدائل مثل: "ماذا إذا..." أو "ماذا بعد" أو "ماذا أيضا"، وغير ذلك من التساؤلات التي تتطلب إعمال الفكر

والمنطق واستخدام المهارة في تطبيق المعارف للتوصل إلى الحلول التي تلتزم بأخلاقيات التدخل المهني في مجال معين.

ويتطلب بناء الاختبار القادر على قياس مستوى التفكير الناقد لدى الطالب أن يكون واضع الاختبار نفسه متمتعا بهذه الصفة بدرجة كافية ولديه الدراية بتصميم الاختبار وبناء بنوده بصورة جيدة وخالية من المخالفات. وعلى الرغم من أن الهدف الرئيس للاختبار هو التقويم، إلا أنه يمكن أن يكون أيضا أداة للتعلم.

ولا شك في أن بنود الاختبار تؤثر على تعلم الطالب. وعليه فإن تبرير البدائل يجعل منه تجربة تعليمية إيجابية ويساعد في نفس الوقت في الوصول إلى الهدف الذي وضع من أجله وهو التقويم. وينبغي أن يتضمن تبرير البدائل أسباب صحة أحدها وأسباب عدم صحة البدائل الأخرى. ومن المؤكد أن ذلك يزيد من ذخيرة الطالب المعرفية ويساعده في تنمية الاختيار القائم على الاختيار بين البدائل وليس استرجاع معلومة ما فحسب. وتقدم تلك المبررات للطلاب بعد أدائه وتصحيحه للاطلاع عليها.

وينبغي أن يكون السؤال على مستوى التطبيق أو أعلى من مستوى الإدراك حتى يكون قادرا على قياس مستوى التفكير الناقد لدى الطالب. لأن الحفظ والاسترجاع يمكن أن يتم بحكم العادة وليس بإعمال الفكر والتأمل، كما أنه يأتي في أدنى درجات الإدراك حسب صنافة "بلوم" (Bloom) المشار إليها أعلاه.

ومن شروط الأسئلة المؤهلة لقياس مستوى التفكير الناقد وضع بنودها بحيث تتطلب تفكيرا ذا منطلقات منطقية متعددة للتعرف على ما هو صحيح منها، ويعني ذلك أن يتمكن من التعامل مع المشكلة المطروحة من منطلقات متعددة وأن ينظر إليها بمناظير متنوعة. ويمكن التعرف على استيفاء السؤال على هذا الشرط من خلال التساؤل عما إذا كان من الضروري أن يكون الطالب ملما بأكثر من مفهوم واحد أو أكثر من منهج واحد أو أكثر من حقيقة واحدة للتمكن من الإجابة على السؤال المطروح. ولا تعتبر الإجابة على سؤال ما صالحة لقياس الفكر الناقد إذا كانت مضمنة في صفحة واحدة من الكتاب الدراسي لأنها لا تتطلب في هذه الحالة سوى الحفظ والاسترجاع.

إن البنود التي تتمتع بالقدرة العالية على تفسير التباين بين البدائل هي الأنسب لقياس التفكير الناقد. ولا بد من وضع بدائل تتميز بإمكانية أن يكون أي منها إجابة صحيحة، وعلى الطالب ذي الفكر الناقد التمكن من التمييز بينها واختيار الإجابة الصحيحة. وهذا يتطلب المقدرة على ترتيب البدائل المطروحة في شكل أولويات للتوصل إلى الإجابة الصحيحة، وهذه العملية تساعد أيضا في تنمية التفكير الناقد بالإضافة إلى الكشف عنه.

خاتمة:

وجهت العديد من الانتقادات لاختبارات الخيارات المتعددة باعتبارها منخفضة الكفاءة في قياس مستويات التحصيل والمهارة الفعلية للطلاب في مختلف مجالات العلوم الاجتماعية. ولا ننكر أن اختبارات الخيارات المتعددة تساعد في قياس أشياء هامة مثل تذكر المعلومات وحفظ التعاريف وفهم المفاهيم الأساسية، وهي جزء من متطلبات ضرورية لتكوين الأساس للفهم الواسع والتفكير المتعمق، ولكنها لا تؤدي إلى ذلك تلقائيا . وحتى عند تعلم أبسط الحقائق يحتاج الطالب لإيجاد "نظرية" مصغرة لنفسه يتمكن من خلالها استيعاب تلك الحقائق لأن حفظ المعلومات لا يتم في غياب المعرفة المنظمة وإلا زالت عن الذاكرة سريعا.

ويفرز الاعتماد على اختبارات الخيارات المتعددة في التقويم نوعية خاصة من المتعلمين غالبيتهم من الطلاب المهرة في هذا النوع من الاختبارات نتيجة للتجربة وتنمية مهارات واستراتيجيات الإجابة، دون أن يكونوا قد نموا أي مهارات في الفكر الناقد والتطبيق والتقويم. وعادة ما يعاني هؤلاء الطلاب من ضعف القدرة على توصيل أفكارهم للآخرين من خلال الكتابة، وقد ينمون توجهات معرفية ضحلة من خلال نظرتهم للمعرفة كمجموعة من الحقائق غير المترابطة يمكن الحصول عليها من خلال الحفظ والاسترجاع دونما تأمل في الفلسفة والأهداف والاستراتيجيات.

ولا يقلل ذلك من حقيقة إن اختبارات الخيارات المتعددة هي واحدة من تلك الأدوات التي تتميز بالسهولة في إدارتها ورصد درجاتها، على الرغم من أن

هذه الوسيلة لا تكفي لوحدها للتقويم الكامل والفعال لمستوى الطالب ودرجة مهارته وخصائصه الفكرية، وخاصة في مجالات العلوم الاجتماعية التي تتميز ظواهرها بالتداخل والتغير السريع. فهي تفيد كثيرا في التقويم الجيد لبعض مناطق المعرفة إذا صممت تصميما سليما.

إن تقديم أسئلة الخيارات المتعددة مستقلة عن بعضها البعض وغير مترابطة، وتكرار استخدامها بهذه الصورة في مجال العلوم الاجتماعية يؤدي إلى انتزاع هذه العلوم عن أطرها التصورية وفقدان الرؤية لما يربطها ببعضها البعض كمجالات متداخلة ومتفاعلة، فمشكلات المجتمع وقضاياه لا تأتي فرادى، بل يضمها نسق واحد ينبغي تناولها في إطاره.

إن استخدام اختبارات الخيارات المتعددة كالأسلوب الوحيد للتقييم يبعد هذه الحقيقة عن ذهن الطالب لأنه يتعامل مع الوقائع والحقائق منفصلة عن بعضها البعض في شكل معلومات مبعثرة هنا وهناك. ولا تتحقق أهداف التقويم الشامل لطالب العلوم الاجتماعية ما لم تستخدم فيه شتى وسائل التقويم الأخرى بجانب اختبارات الخيارات المتعددة والتي يجب أن تصاغ بنودها بصورة تساعد على قياس التفكير الناقد والمستويات العليا من الفكر والمعرفة.

المراجـــــــع

أولا: المراجع العربية:

١. أبو ناشي، منى سعيد (٢٠٠٧). **دراسة في القدرات العقلية – قدرة التفكير الناقد**. عمان: دار الجنادرية.

٢. الدامغ، سامي عبد العزيز (١٩٩٦). *التعدد المنهجي*: أنواعه ومدى ملاءمته للعلوم الاجتماعية. مجلة العلوم الاجتماعية، جامعة الكويت، مجلد (٢٤) عدد (٤) شتاء.

٣. الدوسري، ابراهيم مبارك (١٤٢١هـ). **الإطار المرجعي للتقويم التربوي**. الرياض: مكتب التربية العربي لدول الخليج.

٤. الكبيسي، عبد الواحد حميد (٢٠٠٧).**القياس والتقويم: تجديدات ومناقشات**.عمان: دار جرير للنشر والتوزيع.

٥. الكيلاني، عبد الله زيد (٢٠٠٨). **القياس والتقويم في التعلم والتعليم**. القاهرة: الشركة العربية المتحدة للتسويق والتوريدات.

٦. دعمس، مصطفى نمر (٢٠٠٨). **منهجية البحث العلمي في التربية والعلوم الاجتماعية**. عمان: دار غيداء.

٧. دلبو، فضيل (٢٠٠١). *قضايا منهجية في العلوم الاجتماعية*. الجزائر: ديوان المطبوعات الجامعية.

٨. حجر، خالد أحمد مصطفى (٢٠٠٣). **معايير شروط الموضوعية والصدق والثبات في البحث الكيفي**. مجلة جامعة أم القرى للعلوم التربوية والاجتماعية والإنسانية. المجلد (١٥)، العدد (٢)، يوليو.

٩. سعادة، جودت أحمد (١٩٨٥). استخدام الاختبارات ذات الاختيار المتعدد في التاريخ والجغرافيا. مجلة العلوم الاجتماعية، جامعة الكويت، مجلد (١٣) عدد (١)، ربيع.

١٠. شحادة، نعمان (٢٠٠٩). التعلم والتقويم الأكاديمي. عمان: دار صفاء.

١١. عبد الله، معتز سيد (٢٠٠٦).آليات الإبداع ومقوماته في العلوم الاجتماعية. القاهرة: دار غريب.

١٢. ملحم، سامي محمد (١٤٢٣هـ). القياس والتقويم في التربية وعلم النفس. عماد: دار الميسرة.

١٣. نصار، يحيى حياتي (٢٠٠٤). إدراكات طلبة جامعة الملك سعود لأساليب التقويم وللاختبار كأداة تعليمية وعلاقة ذلك بتخصصاتهم ومستوياتهم التحصيلية. المجلة التربوية، مجلد (١٨)، عدد (٧٠) ، مارس.

ثانيا: المراجع الأجنبية:

1. Aiken, L. R. (1987). ***Testing with multiple-choice items***. Journal of Research and Development, 20(4), 44-58.

2. Alfaro-LeFevre, R. (1995). ***Critical thinking in nursing: A practical approach***. Philadelphia: W.B. Saunders.

3. Bandman, E.L., & Bandman, B. (1995). ***Critical thinking in nursing.*** Norwalk, CN: Appleton and Lange. Blatz, C.V. (1989). Contextualism and critical thinking: Programmatic investigations. Educational Theory, 39(2), 107-119.

4. Barry, T. E., & Thompson, R. (1997). ***Some intriguing relationships in business teaching evaluations***. Journal of Education for Business, 72(5), 303-312.

5. Bloom, B.S. (Ed.), Engelhart, M.D., Furst, E.J., Hill, W.H., & Krathwohl, D.R. (1956). ***Taxonomy of educational objectives: The classification of educational goals.*** Handbook 1: Cognitive domain. New York: David McKay.

6. Case, B. (1994). ***Walking around the elephant: A critical thinking stategy for decision making.*** Journal of Continuing Education in Nursing, 25, 101-109.

7. Crehan, K., & Haladyna, T.M. (1991). ***The validity of two item-writing rules.*** Journal of Experimental Education, 59, 183-192.

8. Crehan, K., Haladyna, T.M., & Brewer, B.W. (1993). ***Use of an inclusive option and the optimal number of options for multiple-choice items.*** Educational and Psychological Measurement, 53, 241-247.

9. Ellsworth, R. A., Dunnell, P., & Duell, O. K. (1990). ***Multiple-choice test items: What are textbook authors telling teachers?*** Journal of Educational Research, 83(5), 289-293.

10. Ennis, R.H. (1993). ***Critical thinking assessment. Theory Into Practice***, 32, 180-186.

11. Facione, N.C., Facione, P.A., & Sanchez, C.A. (1994). ***Critical thinking as a measure of competent clinical judgment: The development of the California critical thinking dispositions inventory.*** Journal of Nursing Education, 33, 345-349.

12. Flower, L. (1994) ***The Construction of Negotiated Meaning: a social cognitive theory of writing*** (Carbondale and Edwardsville, Southern Illinois University Press).

13. Gipps, C. & Murphy, P. (1994) ***A Fair Test?*** (Buckingham, Open University Press.

14. Gronlund, N. E. (1993). ***How to make achievement texts and assessments***. Boston: Allyn and Bacon.

15. Haladyna, T. M., & Downing, S. M. (1989). ***A taxonomy of multiple-choice item-writing rules.*** Applied Measurement in Education, 2(1), 37-50.

16. Krathwohl, D. R. (2002). ***A Revision of Bloom's Taxonomy: An Overview. Theory into Practice*** 41 no4 212-18 Autumn.

17. Kubiszyn, T., & Borich, G. (1996). ***Educational testing and measurement: Classroom application and practice***. New York: Harper Collins.

18. Likert, Rensis (1932). ***"A Technique for the Measurement of Attitudes".*** Archives of Psychology 140: 1–55.

19. Linn, R., Baker, E. & Dunbar, S. (1991) ***Complex, performance-based assessment: expectations and validation criteria***, Educational Researcher, (20) 8, pp. 15-21.

20. McPeck, J.E. (1981). ***Critical thinking and education***. New York: St. Martin's Press.

21. McPeck, J.E. (1990). ***Teaching critical thinking***. New York: Routledge.

22. Morrison, s. and Free, K. W. (2001). ***Writing Multiple-Choice Test Items That Promote and Measure Critical Thinking***. Journal of Nursing Education 40 no1 17-24 January.

23. Moss, P. (1992) ***Shifting conceptions of validity in educational measurement: implications for performance assessment***, Review of Educational Research, 60 (3), pp. 229-258.

24. Patience, W. & Swartz, R. (1987). ***Essay Score Reliability***. Paper Presented at the Annual Meeting of the National Council on Measurement in Education, Washington, DC, April.

25. Paul, R.W. (1987). ***Dialogical thinking: Critical thought essential to the acquisition of rational knowledge and passions***. In J. Baron & R. Sternberg (Eds), Teaching thinking skills: Theory and practice. New York: W.H. Freeman.

26. Resnick, L. B. & Resnick, D. P. (1992) ***Assessing the thinking curriculum: new tools for educational reform***, in: B. Gifford & M. O'Connor, (eds) Changing Assessments: alternative views of aptitude, achievemer and instruction (London, Kluwer Academic Publishers).

27. Violato, C. (1991). ***Item difficulty and discrimination as a function of stem completeness***. Psychological Reports, 69, 739-743.

28. Welsh, A. L., & Saunders, P. (1990). ***Essay questions and tests.*** In P. Saunders & W. B. Walstad (Eds.), The principles of economics course: A handbook for instructors. New York: McGraw-Hill.

29. Wiggins, G. (1993) ***Assessing Student Performance: exploring the limits of testing*** (San Francisco, Exploring the Purpose and Limits of Testing).

30. Wood, W. (1998). ***Linked Multiple Choice Questions: The Tradeoff Between Measurement Accuracy and Grading Time***. Journal of Education for Business, Vol. 74, Issue 2.