

أثر استخدام البديل "لا شيء مما سبق" على بعض الخصائص السيكومترية لأسئلة الاختيار من متعدد

د. ديانا فهمي حمّاد^(١)

(١) أستاذ القياس والتقويم التربوي والنفسي المساعد - كلية التربية - جامعة أم القرى، dfhammad@uqu.edu.sa

(قدم للنشر بتاريخ ٢٨/٠٨/٢٠١٩م - قبل للنشر بتاريخ ٠١/١٠/٢٠١٩م)

ملخص الدراسة:

هدف البحث الحالي إلى دراسة أثر وجود البديل "لا شيء مما سبق" على صعوبة السؤال، وعلى تمييزه، وعلى ثبات الاختبار، من خلال بناء نموذجين متناظرين في المحتوى لاختبار تحصيلي تكوّن من (٢٥) سؤالاً من نوع الاختيار من متعدد، كانت البدائل الأربعة حقيقية في النموذج الأول، أما النموذج الثاني فتم بناؤه باستبدال الإجابة الصحيحة في (١٣) سؤالاً، واستبدال أحد المشتتات في باقي الأسئلة بالبديل "لا شيء مما سبق"؛ وبتطبيق النموذجين على عينة من (٢٤٠) طالبة من طالبات الجامعة، خرج البحث بأن وجود البديل "لا شيء مما سبق" يؤدي إلى زيادة صعوبة أسئلة الاختبار بصورة دالة، بينما لا يؤثر وجوده على تمييز السؤال ولا على ثبات الاختبار.

الكلمات المفتاحية: اختبار الاختيار من متعدد، البديل "لا شيء مما سبق"، صعوبة السؤال، تمييز السؤال، ثبات الاختبار.

للاستشهاد من البحث

حمّاد، ديانا فهمي (٢٠٢١). أثر استخدام البديل "لا شيء مما سبق" على بعض الخصائص السيكومترية لأسئلة الاختيار من متعدد، مجلة جامعة أم القرى للعلوم التربوية والنفسية، المجلد (١٣)، العدد (٤)



The effect of using "none of the above" option on psychometric properties of multiple choice questions

Dr. Dayana Fahmi Hammad ¹

1-Assistant Professor Evaluation & Measurement - Umm Al-Qura University, dfhammad@uqu.edu.sa

Abstract:

The purpose of the research is to investigate using "none of the above" (NOTA) option on psychometric properties of multiple choice tests (MC): Item difficulty, item discrimination and test reliability. Using two test forms symmetrical in content, consisted of (25) MC items with four options, the first form items had real options, the second form was built by replacing the key option in (13) items, and one of the distractors in the (12) items left, by NOTA option. These forms were applied on (240) individuals. The research results indicate that the existence of the NOTA option make the test items more difficult, while its presence does not affect item discrimination nor test reliability.

Keywords: Multiple choice test, None of the above option, Item difficulty, Item discrimination, Test reliability.

How to cite this paper:

Hammad, D (2021). The effect of using "none of the above" option on psychometric properties of multiple choice questions, Journal of Umm Al-Qura University for Educational and Psychological Sciences, Vol (13), No (4)

مقدمة:

يشيع استخدام أسئلة الاختبار من متعدد في التقييم الصفي لسهولة تصحيحها وموضوعيتها (Wainer & Thissen, 1993; Brimi, 2011; Butler, 2018; D'Sa, Alharbi, & Visbal-Dionaldo, 2018) بالإضافة إلى أنها تغطي عددًا أكبر من الحقائق التحصيلية مقارنة بباقي أنواع الأسئلة (Bacon, 2003)، مما يجعلها النوع المفضل في حال كون عدد الطلاب كبيراً، ورغم الانتقادات الموجهة لهذا النوع من الأسئلة، والتي تتركز في وجود نسبة تخمين لا يمكن التخلص منها (أبوعلام، 2005)، بالإضافة إلى أنها غالباً ما تركز على قياس المستويات المعرفية الدنيا (أبوعلام، 2005؛ Walsh & Seldomridge, 2006)، إلا أن هذه العيوب يمكن التغلب عليها، فقد وجد (Downing, 2003) أنه من النادر أن يلجأ الطلاب إلى التخمين العشوائي، بل يقومون بتخمين مبني على تحصيل جزئي يحقق نوعاً من صدق القياس، كما أنه من الممكن صياغة أسئلة تقيس المستويات المعرفية العليا، إلا أنها تتطلب من واضع الأسئلة بذل جهد أكبر (أبوعلام، 2005)، لذلك كله لا تزال أسئلة الاختبار من متعدد هي الأجدود مقارنة بأسئلة الصواب والخطأ، والأكثر شيوعاً في قياس التعلم (Dibattista, Sinnige-Egger, & Fortuna, 2014; Kiat, Ong, and Ganesan, 2017; Butler, 2018) يسهم في تثبيت التعلم (Butler, 2018)؛ فكان من الضرورة ظهور توجه بحثي نحو الاهتمام بوضع إرشادات لصياغة هذا النوع من الأسئلة وصياغة بدائله، والتحقق من جودة هذه الإرشادات وأثرها على الخصائص السيكومترية للاختبار.

على مدى السنوات اهتمت العديد من الدراسات بتقديم إرشادات لصياغة أسئلة الاختبار من متعدد، وتؤكد هذه الإرشادات على أن صياغة بدائل فعالة يعد من الموضوعات الهامة والأساسية لبناء أسئلة تتمتع بخصائص سيكومترية عالية الجودة؛ أما على أرض الواقع وبرغم وجود هذه الإرشادات، إلا أن صياغة بدائل حقيقية جيدة يعد مهمة صعبة، لوجود عدد من الاشتراطات التي يجب أن يحافظ عليها واضع الاختبار، ومع زيادة عدد الأسئلة تصبح هذه المهمة أكثر صعوبة، فيلجأ لاستخدام البديل "لا شيء مما سبق" للحفاظ على عدد ثابت للبدائل في اختبار الاختيار من متعدد (Rich & Johanson, 1990)، ورغم أن أدبيات بناء الاختبارات تشير إلى أن البدائل يجب أن تكون بدائل حقيقية تخلو من البديل "كل ما سبق" و"لا شيء مما سبق"، على سبيل المثال (النبهان، 2004؛ أبوعلام، 2005؛ مراد وسليمان، 2012؛ ميرنز وليمان، 2016)، إلا أن هناك دلائل خرجت بها بعض الأبحاث التجريبية حول استخدام البديل "لا شيء مما سبق"، تشير إلى أن وجوده يُحسِّن الخصائص السيكومترية للأسئلة التحصيلية مع وجود اختلاف في هذه النتائج حول ماهي الخصائص السيكومترية التي تخضع للتحسن ومدى تحسنها (Rich & Johanson, 1990; Frary, 1991; Knowles & Welch, 1992) ولحسم الخلاف حول استخدام البديل "لا شيء مما سبق" كان من المهم بحث أثر استخدامه على الخصائص السيكومترية للاختبارات التحصيلية.

مشكلة البحث:

تعددت أهداف الأبحاث حول أثر استخدام البديل "لا شيء مما سبق" على الخصائص السيكومترية لأسئلة الاختبار من متعدد، فبعض الأبحاث هدفت لدراسة أثر استخدام البديل على صعوبة السؤال وتمييزه (Frary, 1991; Knowles & Welch, 1992; Dochy, Moerkerke, Corte, & Segers, 2001; Dibattista, Sinnige-Egger, &

(Rich & Johanson, 1990; Atalmış & Kingston, 2017; Pachai, DiBattista, & Kim, 2015; Fortuna, 2014) وبعضها أضاف الثبات أيضا (Rich & Johanson, 1990; Atalmış & Kingston, 2017) وكما اختلفت أهداف الأبحاث، فقد اختلفت أيضا نتائجها، فبعضها خرج بأن استخدام البديل "لا شيء مما سبق" يزيد من صعوبة السؤال ومن تمييزه، وبعضها أشارت نتائجه بأنه لا يؤثر على الصعوبة ولا على التمييز، والبعض الآخر أشارت نتائجه إلى أنه يخفض من قدرة السؤال التمييزية.

وعليه فإن مشكلة البحث الحالية تتحدد في الإجابة على السؤال الآتي:

هل يوجد أثر لاستخدام البديل "لا شيء مما سبق" على صعوبة وتمييز وثبات أسئلة الاختيار من متعدد؟

ويتفرع من السؤال الرئيس التساؤلات الآتية:

١. هل توجد فروق دالة إحصائية بين متوسطي معاملات صعوبة أسئلة النموذجين تعود لاستخدام البديل "لا شيء مما سبق"؟

٢. هل توجد فروق دالة إحصائية بين متوسطي معاملات تمييز أسئلة النموذجين تعود لاستخدام البديل "لا شيء مما سبق"؟

٣. هل توجد فروق دالة إحصائية بين متوسطي معاملات ثبات النموذجين تعود لاستخدام البديل "لا شيء مما سبق"؟

أهداف البحث:

١. الكشف عن أثر استخدام البديل "لا شيء مما سبق" على معاملات صعوبة أسئلة الاختيار من متعدد.

٢. الكشف عن أثر استخدام البديل "لا شيء مما سبق" على معاملات تمييز أسئلة الاختيار من متعدد.

٣. الكشف عن أثر استخدام البديل "لا شيء مما سبق" على ثبات أسئلة الاختيار من متعدد.

أهمية البحث:

تتضح أهمية البحث فيما يلي:

١. تعد أسئلة الاختيار من متعدد شائعة الاستخدام للتقويم الصففي في مختلف مراحل التعليم، وبعد التحقق من جودة صياغة هذا النوع من الأسئلة، أمرًا هامًا لتجنب أخطاء القياس الناتجة عن سوء الصياغة، ونتائج البحث الحالي قد تساعد في تحسين جودة إرشادات صياغة هذا النوع من الأسئلة، وتسهم إيجابا في تحقيق صدق القياس.

٢. نتائج البحث الحالي قد تسهم في التحقق من أثر استخدام أحد البدائل، التي يلجأ لها واضع الاختبار بهدف بناء نماذج متكافئة من الاختبار، أو بهدف مساواة عدد البدائل بين الأسئلة، أو لزيادة عدد البدائل وتقليل نسبة التخمين، وقد تسهم نتائج البحث الحالي في ترجيح كفة النتائج المتباينة حول أثر استخدام البديل "لا شيء مما سبق" على صعوبة وتمييز وثبات أسئلة الاختيار من متعدد.

٣. الدراسات العربية الحديثة في هذا المجال قليلة والبحث الحالي يسهم في زيادة التراكم المعرفي فيه.

التعريف بالمصطلحات:

• أسئلة الاختيار من متعدد:

تعرفها الباحثة بأنها: أحد أنواع الأسئلة موضوعية التصحيح والتي تتميز بانخفاض نسبة تخمينها مقارنة بأسئلة الصواب والخطأ، ومرونتها في تغطية أنواع متباينة من الحقائق مقارنة بأسئلة المزاجية.

يتكون سؤال الاختيار من متعدد من مقدمة تسمى متن أو نص السؤال Stem، يليها خيارات يتراوح عددها من ٣-٥ خيارات تسمى بدائل، والبديل الصحيح هو الإجابة، أما البدائل الخاطئة فتسمى مشتتات أو موهات Distractors (مراد وسليمان، ٢٠١٢).

• صعوبة السؤال Item Difficulty:

هو نسبة الأفراد الذين أجابوا إجابة صحيحة على السؤال، ويحسب بقسمة عدد المفحوصين الذين أجابوا إجابة صحيحة على السؤال، على العدد الكلي للمفحوصين (Sowdani, Jaber & Thomas, 2018)، فإذا كان تصحيح السؤال ثنائياً (١، ٠) فإن معامل الصعوبة (P_i) يساوي المتوسط الحسابي لدرجات المفحوصين على السؤال.

• تمييز السؤال Item Discrimination:

"التمييز هو مؤشر لقدرة السؤال على فصل الأفراد الأعلى أداءً على الاختبار، عن الأفراد الأقل أداءً على الاختبار" (Scialfa,)، وفي هذا البحث يُعبّر عن معامل التمييز (Discrimination Index (DI) بمعامل الارتباط الثنائي الأصيل (Point-Biserial Coefficient).

• ثبات الاختبار Test Reliability:

يُعبّر ثبات الاختبار عن درجة التجانس بين نتائج تطبيق قياسين للفرد لنفس السمة أو السلوك، ومن الطرق الأكثر شيوعاً في تقدير ثبات الاختبار حساب معامل الاتساق الداخلي، والذي يعكس مدى ارتباط استجابات الأفراد على السؤال مع درجة الاختبار الكلية (النبهان، ٢٠٠٤)، وفي هذا البحث يُعبّر عن ثبات الاختبار بمعامل ألفا (α) لكرونباخ.

الإطار النظري:

• أسئلة الاختيار من متعدد:

تعد أسئلة الاختيار من متعدد من أنماط الأسئلة الأكثر شيوعاً في اختبارات الطلاب التحصيلية، ولأن وجود أخطاء في الصياغة يؤدي إلى مشكلات في صدق القياس، فقد تضمنت أدبيات القياس التربوي إرشادات صياغة كامل أسئلة الاختيار من متعدد وإرشادات صياغة تفصيلية لمتن السؤال ولبدائله، وبعد الرجوع إلى (Haladyna, Downing, & Rodriguez, 2002)؛ النبهان،

٢٠٠٤؛ Frey, Petersen, Edwards, Pedrotti, & Peyton, ٢٠١٢؛ مراد وسليمان، ٢٠١٢؛ ميرنز وليمان، (٢٠١٦)، يمكن إجمالها فيما يلي:

• إرشادات صياغة كامل سؤال الاختيار من متعدد:

- ١) أن تغطي الأسئلة حقائق وأهداف هامة.
- ٢) صياغة السؤال باستخدام كلمات مناسبة لثقافة وعمر المستجيبين.
- ٣) تجنب النقل الحرفي من مرجع المقرر.
- ٤) صياغة كل سؤال باستقلالية عن غيره.
- ٥) أن تكون هناك إجابة صحيحة واحدة.
- ٦) الاختصار في الصياغة بما لا يخل المعنى.
- ٧) توضع الأسئلة من نفس النوع بتتابع.
- ٨) يوضع السؤال كاملاً في صفحة واحدة.
- ٩) ترقيم جميع الأسئلة.
- ١٠) تجنب وجود نمط للإجابة.
- ١١) طباعة الأسئلة بوضوح والابتعاد عن كتابتها بخط اليد.
- ١٢) وضع درجة السؤال.
- ١٣) وضوح التعليمات.

• إرشادات صياغة متن السؤال وصياغة البدائل:

- ١) صياغة متن السؤال بحيث يتضمن مشكلة واضحة.
- ٢) تجنب استخدام النفي.
- ٣) صياغة المتن في جملة تامة.
- ٤) تجنب الإشارات النحوية واللغوية بين البدائل ومتن السؤال.
- ٥) تجنب تكرار الكلمات بين المتن والبديل الصحيح.
- ٦) تجنب استخدام البديل "كل ما سبق"، ومثلها البديل (أ و ب) صحيحان .
- ٧) تجنب استخدام البديل "لا شيء مما سبق".
- ٨) صياغة جميع البدائل بحيث تكون إجابات محتملة.
- ٩) تجنب استخدام الكلمات التي تشير إلى الإجابة الصحيحة مثل (دائماً، أبداً).
- ١٠) تجنب استخدام الكلمات الغامضة مثل (عادة، غالباً).
- ١١) يجب أن تكون البدائل متجانسة في المحتوى.

- (١٢) أن تصاع البدائل متقاربة الأطوال.
 (١٣) أن تكون البدائل دائماً أقصر من المتن.
 (١٤) أن يتراوح عدد البدائل من ٣ - ٥ بدائل.
 (١٥) أن تكون البدائل مستقلة عن بعضها.

• **مزايا وعيوب أسئلة الاختيار من متعدد:**

تعد موضوعية التصحيح وسهولته من المزايا الرئيسة التي تجعل هذا النوع من الأسئلة شائع الاستخدام، بالإضافة إلى انخفاض نسبة التخمين فيه مقارنة بأسئلة الصواب والخطأ (أبوعلام، ٢٠٠٥؛ مخائيل، ٢٠١٦؛ ميرنز وليمان، ٢٠١٦)، وتشير نتائج دراسة (Mccoubrie, 2004) إلى أنها تمتاز بجودة خصائصها السيكومترية فنباتها وصدقها عاليين، كما تمتاز بقدرتها على تغطية مجال واسع من المحتوى المعرفي، وتنبأ وترتبط بالأداء التحصيلي الكلي للطالب.

إلا أنه يُعاب على هذا النوع من الأسئلة، أنه رغم جميع الإرشادات المتبعة لصياغته تظل هناك نسبة من التخمين تؤثر على صدق الإجابة، كما أن إعدادها يستهلك وقتاً طويلاً، بالإضافة إلى أنها تتطلب مهارة عالية لصياغة مشتتات فعالة ذات تمييز مرتفع (أبوعلام، ٢٠٠٥؛ مخائيل، ٢٠١٦؛ ميرنز وليمان، ٢٠١٦).

• **معامل صعوبة السؤال (Item Difficulty Value (Pi):**

تحسب قيمة معامل الصعوبة (P_i) بقسمة عدد المفحوصين الذين أجابوا إجابة صحيحة على السؤال على العدد الكلي للمفحوصين، وعندما يكون تصحيح السؤال ثنائياً (١، ٠) فإن معامل الصعوبة يساوي المتوسط الحسابي لدرجات المفحوصين على السؤال، وتنحصر قيمته في المدى (صفر $> P_i > ١$) (النبهان، ٢٠٠٤)، فإذا كانت ($\sum X_i$) هي مجموع الإجابات الصحيحة على السؤال (i)، وعدد الأفراد المختبرين (N)، فإن معامل الصعوبة (P_i) يحسب بالمعادلة:

$$P_i = \frac{\sum X_i}{N}$$

وانخفاض قيمة معامل الصعوبة المحسوب بالمعادلة السابقة يشير إلى صعوبة السؤال، وارتفاعها يشير إلى سهولته.

• **معامل تمييز السؤال (Discrimination Index (DI):**

يشير معامل التمييز إلى الدرجة التي يميز فيها السؤال بين المفحوصين في السمة التي يقيسها الاختبار، وتتراوح قيمته بين (± ١)؛ وفي الاختبارات التحصيلية، يعبر معامل تمييز السؤال عن مدى فاعلية السؤال في التمييز بين الطلاب مرتفعي القدرة التحصيلية ومنخفضيها (Atalmış & Kingston, 2017)، ويعد السؤال جيداً كلما ارتفع معامل تمييزه (Varm, 2003)، وفي هذا البحث سيعبر عنه بمعامل الارتباط الثنائي الأصيل (Point-Biserial Coefficient)، ويحسب معامل التمييز معبراً عنه بمعامل الارتباط الثنائي الأصيل (DI_{PB}) بالمعادلة:

$$DI_{PB} = \frac{M_i - M}{S} \sqrt{\frac{P_i}{1 - P_i}}$$

حيث:

M : متوسط درجات الطلاب الذين اختاروا البديل S : الانحراف المعياري لدرجات الطلاب في كامل الاختبار.

M_i : متوسط درجات الطلاب في كامل الاختبار. P : نسبة الطلاب الذين اختاروا البديل إلى العدد الكلي (معامل الصعوبة).

• دلالة قيم معامل الصعوبة ومعامل التمييز

يشير (Sowdani, Jaber & Thomas, 2018) إلى أن قيمة معامل الصعوبة المقبولة تكون في المدى $(0,15 > P_i > 0,85)$ ، أما معامل التمييز فحده المقبول أن يكون في المدى $(DI > 0,1)$ ، ويتفق (Scialfa et al., 2001) معهم في حد معامل التمييز، بينما يرون أن مدى الصعوبة المقبول يكون في المدى $(0,25 > P_i > 0,75)$ ، أما (Varm, 2003) فيرى أن الأسئلة التي يكون معامل تمييزها أكبر من $(0,3)$ تعد عالية التمييز، وفي المدى من $(0,1)$ إلى $(0,3)$ يعد مقبولاً، أما التي ينخفض معامل تمييزها عن $(0,1)$ فتكون ضعيفة.

وفي البحث الحالي سيتم الإشارة إلى الأسئلة التي يكون معامل صعوبتها في المدى $(0,15 > P_i > 0,85)$ بأنها مقبولة وغير ذلك تكون خارج الحد المقبول، أما بالنسبة للتمييز سيشار إلى الأسئلة التي يكون معامل تمييزها أقل من $(0,1)$ بأنها ضعيفة، وفي المدى من $(0,1) \geq DI \geq 0,3$ تكون مقبولة، وأعلى من $(0,3)$ ذات تمييز عال.

• ثبات الاختبار

يُعبّر ثبات الاختبار عن درجة التجانس بين نتائج قياسين لصفة أو سلوك ما، ونظرياً هو النسبة بين التباين في الدرجة الملاحظة والتباين في الدرجة الحقيقية؛ ولتقدير الثبات عملياً يشيع استخدام معامل ألفا (α) الذي طور معادلته كرونباخ (النبهان، ٢٠٠٤)، وتعتمد قيمة (α) على النسبة بين مجموع تباينات الأسئلة $(\sum S_i^2)$ إلى تباين كامل الاختبار (S^2) ، فإذا كان لدينا مقياس عدد أسئلته (n) ، فإن معادلة حساب ثباته هي:

$$\alpha = \frac{n}{n - 1} \left(1 - \frac{\sum S_i^2}{S^2} \right)$$

الدراسات السابقة:

هدفت دراسة (Rich & Johanson 1990) إلى تحديد أثر استخدام البديل "لا شيء مما سبق" على صعوبة وتمييز السؤال وثبات اختبار، من خلال إعادة حساب هذه المعاملات لثلاثة دراسات استخدمت البديل "لا شيء مما سبق" فيما لا يقل عن (٢٥%) من أسئلة الاختبار، وكان البديل هو إجابة الصحيحة في (٢٥%) من الأسئلة التي تضمنت وجوده، وخرجت الدراسة بأن استخدام البديل يزيد صعوبة السؤال وتمييزه، ولا يؤثر في ثبات الاختبار.

وقام Frary (1991) بدراسة لتقويم استخدام البديل "لا شيء مما سبق"، بتحليل بيانات (١٠١٩) سؤالاً من نوع الاختيار من متعدد، استخرجت من (٢٠) اختباراً طبق على أكثر من (١٠٠) طالب، في (٨) مجالات أكاديمية، (٢٩٥) من أسئلة الاختيار من متعدد تضمن البديل "لا شيء مما سبق"، و(٧٢٤) سؤالاً عادياً، خرجت الدراسة بأن الأسئلة التي تتضمن البديل "لا شيء مما سبق" كانت أكثر صعوبة من الأسئلة العادية، لكن لها نفس التمييز، والأسئلة التي كان البديل "لا شيء مما سبق" هو الإجابة الصحيحة كانت أكثر صعوبة من الأسئلة التي كان فيها البديل مشتتاً، ولم توجد فروق بين معاملات التمييز في الحالتين (استخدام البديل إجابةً صحيحةً، أو مشتتاً)، وكان الطلاب يميلون لاختيار البديل عندما يكون مشتتاً، ويميل الطلاب الأقل تحصيلاً لتفادي اختيار المشتت "لا شيء مما سبق" عندما لا يعرفون الإجابة، وبناء عليه أوصى الباحث باستخدام البديل لأن نتائج البحث تشير إلى أن استخدامه يرتبط إيجاباً بجودة قياس التحصيل الصفي.

وهدف Knowles & Welch (1992) إلى دراسة أثر استخدام البديل "لا شيء مما سبق" على صعوبة وتمييز السؤال، من خلال تطبيق تحليل بعدي لحجوم الأثر في عينة من (١٢) بحثاً، تناولت أثر استخدام البديل على صعوبة السؤال، وعينة من (٧) أبحاث، لدراسة أثر استخدام البديل على تمييز السؤال؛ فكانت حجوم الأثر المحسوبة ضئيلة جداً بلغت (-٠,١٧، للصعوبة و ٠,٠١، للتمييز)، وتشير هذه النتائج إلى أن استخدام البديل "لا شيء مما سبق"، لا يؤثر على جودة خصائص السؤال المرتبطة بالصعوبة والتمييز. وفي دراسة (Dochy et al. 2001) تم تحليل إجابات (١٦٧) طالبا استجابوا على اختبار فيه (١٢) مسألة رياضية على شكل اختيار من متعدد تضمن البديل "لا شيء مما سبق"، وقد طلب من الطلاب اختيار البديل الصحيح مع كتابة الحل التفصيلي، وتشير النتائج إلى أن الطلاب يميلون إلى اختيار البديل "لا شيء مما سبق" عندما لا يعرفون الإجابة، كما تدعم النتائج أن وجود البديل يجعل السؤال أكثر صعوبة وتمييزاً، خاصة إذا كان السؤال يتطلب حسابات رياضية.

وباستخدام تحليل المحتوى، قام Haladyna, Downing & Rodriguez (2002) ببناء تصنيف لإرشادات جودة صياغة أسئلة الاختبارات للمهتمين بالتقييم الصفي، هذه الإرشادات منها (٣١) إرشاداً تختص بأسئلة الاختيار من متعدد، تم استخلاصها اعتماداً على مصدرين من الأدلة: إجماع الباحثين في (٢٧) مرجعاً في الاختبارات التربوية، والمصدر الثاني نتائج (٢٧) بحثاً تم نشرها منذ عام ١٩٩٠؛ وفيما يخص البديل "لا شيء مما سبق" فقد أظهر تحليل محتوى المراجع انقسام هذه المراجع حوله فنسبة المراجع التي أشارت إليه وكانت مؤيدة لاستخدامه بلغت (٤٤%)، ونسبة المراجع المعارضة لاستخدامه (٤٨%)؛ وأوصى الباحثون بأن البديل "لا شيء مما سبق" يظل بديلاً مستخدماً من قبل واضع الأسئلة، ونظراً لتباين نتائج الأبحاث حوله فإنه لا ينصح باستخدامه من قبل واضع الأسئلة المبتدئ.

وهدفت دراسة (Frey et al. (2005 لاستخلاص قواعد صياغة الأسئلة الموضوعية بمسح (٢٠) مرجعا في التقييم التربوي، وقاموا ببناء جدول لقواعد الصياغة يضم (٤٠) قاعدة للأسئلة الموضوعية، وسجلوا فيه عدد المراجع التي دعمت القاعدة وما إذا ورد دعم للقاعدة بأبحاث العلمية في المراجع العشرين، وكان من ضمن القواعد عدم استخدام البديل "لا شيء مما سبق" في أسئلة الاختيار من متعدد وكان عدد المراجع التي أشارت للقاعدة (١٥) مرجعاً، ومرجعين دعماً استخدام البديل لزيادة صعوبة السؤال، كما وأن هذه القاعدة لم تشر المراجع إلى أنه تم دعمها بأبحاث تجريبية.

وأجرى (Odegard & Koen (2007) دراسة لتحديد أثر إضافة البديل "لا شيء مما سبق" على تحسين أداء عينة من (٣٢) طالبا جامعيًا في الاختبار، فتم صياغة (٣) نماذج من أسئلة الاختيار من متعدد، في النموذج الأول أربع بدائل حقيقية أحدها هو الإجابة الصحيحة، وفي النموذج الثاني تم إضافة البديل "لا شيء مما سبق" بديلا خامسا مشتتا، وفي النموذج الأخير تم استبدال الإجابة الصحيحة في النموذج الثاني بإجابة خاطئة ليصبح الخيار "لا شيء مما سبق" هو الإجابة الصحيحة؛ احتوى الاختبار على (٧٢) سؤالاً على قطعة سبق قراءتها، و(٧٢) سؤالاً على قطعة لم يسبق قراءتها. هذه الأسئلة كان ربعها من النموذج الأول ذي الأربع بدائل الحقيقية، وربعها الثاني من النموذج الثاني ذي البدائل الخمسة والتي فيها البديل "لا شيء مما سبق" مشتتا، والربع الثالث كان من النموذج الثالث ذي الخمسة بدائل التي يمثل فيها البديل "لا شيء مما سبق" إجابة صحيحة، وجميع الأرباع الثلاثة من الأسئلة سبق للطلاب الاختبار فيها، أما الربع الأخير فتكوّن من أسئلة جديدة لم يسبق للطلاب الاختبار فيها، واعتبرها الباحثان مجموعة أسئلة ضابطة، وخرج الباحثان بأن وجود البديل "لا شيء مما سبق" يخفض أداء الطلاب في الاختبار اللاحق، حيث يميل الطلاب إلى تجنب اختياره حتى في الأسئلة التي سبق لهم الاختبار فيها.

وطبق (Dibattista, Sinnige-Egger & Fortuna (2014) دراسة على عينة من (٢٠١) طالبًا من طلاب الجامعة، وتكونت أداة الدراسة الأساسية من اختبار اختيار من متعدد في المعلومات العامة به (٤٠) سؤالاً ذي أربع بدائل، بنيت وفق إرشادات جودة الصياغة، اشتق من هذه الأداة أربع نماذج إضافية للاختبار، في النموذج الأول تم استبدال الخيار الصحيح بالبديل "لا شيء مما سبق"، وفي باقي النماذج الثلاث استبدلت المشتتات بالبديل "لا شيء مما سبق"، وخرجت الدراسة بأن وجود البديل "لا شيء مما سبق" مكان الإجابة الصحيحة يرفع صعوبة السؤال، ويخفض قدرته التمييزية.

وهدفت دراسة (Pachai, DiBattista & Kim (2015) إلى تحديد أثر استخدام البديل "لا شيء مما سبق" على خصائص اختبار اختيار من متعدد رباعي الخيارات، وقد تم تضمينه إجابةً صحيحةً في بعض نماذج أداة الدراسة التحصيلية، ومشتتًا في نماذج أخرى، وبتطبيق أداة الدراسة على عينة بلغت (٣٦٩٧) طالبًا في مقرر المدخل إلى علم النفس، خرج الباحثون بأن تضمين الخيار "لا شيء مما سبق" يزيد من صعوبة السؤال إذا استبدل بالإجابة الصحيحة، ويخفض من قدرته التمييزية، ولا يؤثر على الصعوبة إذا استبدل بمشتت عال الجودة.

كما هدفت دراسة الأمير (٢٠١٦) إلى تحديد الأخطاء الشائعة في بناء مفردات اختبار الاختيار من متعدد لدى معلمي الرياضيات في الصف الثاني متوسط، وتحديد الاختلافات في الخصائص السيكومترية للاختبار الجيد وغير الجيد وهي معاملات الصعوبة والتمييز وفعالية المشتتات وصدق وثبات الاختبار؛ وأشارت المعلومات المجموعة من (٥١) اختبار تحصيلي، إلى أن نسبة استخدام الخيار "لا

شيء مما سبق" و"كل ما سبق" (معا) في عينة الأسئلة بلغت (١,٢ %)، ولم يفصل بين نسب استخدام البديلين؛ وتطبيق اختبارين أحدهم جيد الصياغة والأخر غير جيد الصياغة على (٤٥٢) طالبا من طلاب الصف الثاني المتوسط، توصلت الدراسة إلى وجود فروق في جميع المعاملات الإحصائية لصالح الاختبار جيد الصياغة.

بينما هدفت دراسة [Atalmış & Kingston \(2017\)](#) إلى اختبار أثر عدد البدائل وأثر استخدام البديل "لا شيء" مما سبق" على معاملات صعوبة وتمييز وثبات السؤال باستخدام اختبار في مقرر رياضيات، وتطبيق أداة تحصيلية على (٥٨٥) طالبًا من طلاب الصف السابع والثامن، خرجت الدراسة بأن لا توجد فروق في صعوبة السؤال ولا في تمييزه ولا في معامل الثبات بين الأسئلة ذات الأربع بدائل بنوعيتها: ذات البدائل الحقيقية والتي تتضمن البديل "لا شيء" مما سبق".

تعليق على الدراسات السابقة:

تعددت الجوانب التي تناولتها الدراسات بالبحث، فهناك دراسات تناولت قواعد صياغة أسئلة الاختبار من متعدد وبدائله ([Haladyna, Downing, & Rodriguez, 2002](#)؛ [Frey et al., 2005](#)؛ الأمير، ٢٠١٦)، من هذه الأبحاث دراستين تناولتا قواعد صياغة أسئلة الاختبار من متعدد وبدائله ومن ضمنها وجود البديل "لا شيء" مما سبق" من خلال تحليل مضمون مراجع لقواعد صياغة هذا النوع من الأسئلة، ([Haladyna, Downing, & Rodriguez, 2002](#)؛ [Frey et al., 2005](#)) وأقرت الدراستان بأنه رغم وجود نسبة غالبية من المراجع أوصت بعدم اللجوء إليه إلا أنه لا يزال مستخدمًا، وأشارت دراسة ([Frey et al., 2005](#)) بأن التوصية بعدم استخدامه لم يتم دعمها بأبحاث امبيريقية؛ أما دراسة (الأمير، ٢٠١٦) فهدفت إلى تحديد الأخطاء الشائعة في بناء مفردات اختبار الاختيار من متعدد وأشارت إلى استخدام البديل "لا شيء" مما سبق" ضمن تلك الأخطاء.

وهناك أبحاث هدفت لدراسة أثر وجود البديل على متغيري صعوبة السؤال وتمييزه ([Frary, 1991](#)؛ [Knowles & Welch, 1992](#)؛ [Dochy et al., 2001](#)؛ [Dibattista, Sinnige-Egger, & Fortuna, 2014](#)؛ [Pachai, DiBattista, & Kim, 2015](#))، ودراستين تعرضتا لمتغير ثبات الاختبار بالإضافة إلى صعوبة وتمييز السؤال ([Rich & Johanson, 1990](#)؛ [Atalmış & Kingston, 2017](#))، وكانت نتائج هذه الدراسات متباينة، فقد أشارت أغلب الدراسات إلى أن وجود البديل "لا شيء" مما سبق" يؤدي إلى زيادة صعوبة السؤال ([Rich & Johanson, 1990](#)؛ [Frary, 1991](#)؛ [Dochy et al., 2001](#)؛ [Dibattista, Sinnige-Egger, & Fortuna, 2014](#)؛ [Pachai, DiBattista, & Kim, 2015](#))، أما بحثي ([Knowles & Welch, 1992](#)؛ [Atalmış & Kingston, 2017](#)) فأشارا إلى أنه لا يؤثر على الصعوبة. وفيما يرتبط بأثر البديل على تمييز السؤال فقد أشارت دراستين إلى أنه يخفف من تمييز السؤال بأن وجود البديل يزيد من قدرة السؤال التمييزية ([Dochy et al., 2001](#))، ودراسة خرجت بعدم وجود أثر للبديل على تمييز السؤال ([Dibattista, Sinnige-Egger, & Fortuna, 2014](#)). واتفقت الدراستان اللتان بحثتا أثر البديل على ثبات الاختبار، على عدم وجود أثر للبديل "لا شيء" مما سبق" على الثبات ([Rich & Johanson, 1990](#)؛ [Atalmış & Kingston, 2017](#)).

وهناك دراسة واحدة تعرضت لأثر وجود البديل على أداء الطلاب (Odegard & Koen, 2007) وخرجت بأن وجود البديل "لا شيء مما سبق" يخفض أداء الطلاب.

فروض البحث:

- (١) لا توجد فروق ذات دلالة إحصائية بين متوسطي معاملات صعوبة الأسئلة تعود لاستخدام البديل "لا شيء مما سبق".
- (٢) لا توجد فروق ذات دلالة إحصائية بين متوسطي معاملات تمييز الأسئلة تعود لاستخدام البديل "لا شيء مما سبق".
- (٣) لا توجد فروق ذات دلالة إحصائية بين معاملي ثبات نموذجي الأسئلة تعود إلى استخدام البديل "لا شيء مما سبق".

منهجية البحث:

في هذا البحث تم استخدام المنهج الوصفي السببي المقارن لاختبار فروض الدراسة، وهو المنهج الأنسب والذي يُعنى بوصف ودراسة متغيرات البحث كما هي على أرض الواقع.

عينة البحث:

قامت الباحثة بتطبيق الاختبار التحصيلي على عينة من طالبات المستوى الثامن بكلية التربية بجامعة أم القرى بلغ عددهن (٢٤٠) طالبة، أجابت (١١٦) طالبة على النموذج الأول ذي البدائل الحقيقية، وباقي العينة البالغ عددها (١٢٤) طالبة، أجبن على النموذج الثاني ذي البديل "لا شيء مما سبق".

أداة البحث:

استخدمت الباحثة اختباراً تحصيلياً من إعداد المعلم، أعدته في مقرر الاختبارات والمقاييس، تكوّن من (٢٥) سؤالاً اختبارياً من نوع الاختيار من متعدد ذي الأربع بدائل، وكانت البدائل في النموذج الأول، بدائل حقيقية؛ أما النموذج الثاني فتكون من نفس أسئلة النموذج الأول بنفس ترتيبها، مع استبدال البديل الأخير بالبديل "لا شيء مما سبق"، ماعداً إذا كان البديل المستبدل هو الإجابة الصحيحة فيرحل البديل الأخير مكان الإجابة الصحيحة المستبدلة؛ فحصلت الباحثة على نموذجين متكافئين في عدد الأسئلة وترتيبها ومحتواها، وكان الفرق بين النموذجين أن النموذج (١) جميع بدائله حقيقية، أما النموذج (٢) فجميع أسئلته تضمنت البديل "لا شيء مما سبق" بديلاً صحيحاً في (١٣) سؤالاً، ومشتتاً في (١٢) سؤالاً.

ثم أجرت تحليلاً لأسئلة الاختبار فحسبت معاملات الصعوبة، ومعاملات التمييز معبراً عنها بمعامل الارتباط الثنائي الأصيل وذلك لكل سؤال، والجدول (١) يعرض إحصائيات أسئلة النموذجين:

جدول ١: إحصائيات تحليل أسئلة نموذجي أداة البحث

نموذج ٢ (ذو البديل "لا شيء مما سبق")					نموذج ١ (ذو البدائل الحقيقية)				
رقم السؤال	عدد الأفراد	عدد الإجابات الصحيحة	معامل الصعوبة	معامل التمييز	رقم السؤال	عدد الأفراد	عدد الإجابات الصحيحة	معامل الصعوبة	معامل التمييز
١	١٢٤	٧٤	٠,٦	٠,٥٣٧	١	١١٦	٩٤	٠,٨١	٠,٤٧
٢	١٢٤	٥٧	٠,٤٦	٠,٦٩٤	٢	١١٦	٨٤	٠,٧٢	٠,٥١١
٣	١٢٤	٨١	٠,٦٥	٠,٣٤٧	٣	١١٦	٧٢	٠,٦٢	٠,٤٠٦
٤	١٢٤	٥٩	٠,٤٨	٠,٥٠٨	٤	١١٦	٨٧	٠,٧٥	٠,٤٥٥
٥	١٢٤	٢١	٠,١٧	٠,٣٣٩	٥	١١٦	٦٤	٠,٥٥	٠,٠٦٨-
٦	١٢٤	٧٢	٠,٥٨	٠,٤٢٥	٦	١١٦	٦٤	٠,٥٥	٠,٥٠١
٧	١٢٤	٧٤	٠,٦	٠,٣٠٤	٧	١١٦	٦٤	٠,٥٥	٠,٤٧٦
٨	١٢٤	٢٦	٠,٢١	٠,٤١٤	٨	١١٦	٥٦	٠,٤٨	٠,٣٥٥
٩	١٢٤	٧٩	٠,٦٤	٠,٢١٢	٩	١١٦	٨٩	٠,٧٧	٠,٣٢٥
١٠	١٢٤	٧٤	٠,٦	٠,٤٩٢	١٠	١١٦	٧٤	٠,٦٤	٠,٣٧٣
١١	١٢٤	٦٤	٠,٥٢	٠,٢٨١	١١	١١٦	٤٩	٠,٤٢	٠,٢٤٧
١٢	١٢٤	٥٠	٠,٤	٠,٦١٢	١٢	١١٦	٦٤	٠,٥٥	٠,٤١٩
١٣	١٢٤	٦٤	٠,٥٢	٠,٤٨٣	١٣	١١٦	١٠٥	٠,٩١	٠,٣٢
١٤	١٢٤	٩٧	٠,٧٨	٠,١٦٣	١٤	١١٦	٩٠	٠,٧٨	٠,٢٠٨
١٥	١٢٤	٢٩	٠,٢٣	٠,٥٤٤	١٥	١١٦	٥٢	٠,٤٥	٠,٤٩١
١٦	١٢٤	٥٤	٠,٤٤	٠,٥١١	١٦	١١٦	٦٦	٠,٥٧	٠,٣٨٤
١٧	١٢٤	٤٥	٠,٣٦	٠,٧٠٩	١٧	١١٦	٩٢	٠,٧٩	٠,٤٤
١٨	١٢٤	٣٣	٠,٢٧	٠,٦٣٣	١٨	١١٦	٤٠	٠,٣٤	٠,٥٣
١٩	١٢٤	٨٩	٠,٧٢	٠,١١٦	١٩	١١٦	٨٨	٠,٧٦	٠,٥٠٣

نموذج ٢ (ذو البديل "لا شيء مما سبق")					نموذج ١ (ذو البدائل الحقيقية)				
رقم السؤال	عدد الأفراد	عدد الإجابات الصحيحة	معامل الصعوبة	معامل التمييز	رقم السؤال	عدد الأفراد	عدد الإجابات الصحيحة	معامل الصعوبة	معامل التمييز
٢٠	١٢٤	١١٢	٠,٩	٠,٠٧	٢٠	١١٦	١٠٨	٠,٩٣	٠,٠٣٣-
٢١	١٢٤	٩١	٠,٧٣	٠,١٠٥-	٢١	١١٦	٨٢	٠,٧١	٠,٠٦٢
٢٢	١٢٤	٧٤	٠,٦	٠,٢١٨	٢٢	١١٦	٧٦	٠,٦٦	٠,٣٣٣
٢٣	١٢٤	٥٠	٠,٤	٠,٥٨٤	٢٣	١١٦	٣٦	٠,٣١	٠,٣٤٧
٢٤	١٢٤	٧٥	٠,٦	٠,٣٢٦	٢٤	١١٦	٩٥	٠,٨٢	٠,٢٨٩
٢٥	١٢٤	٢٩	٠,٢٣	٠,٠٤٧	٢٥	١١٦	٢١	٠,١٨	٠,٠١٥-

يعرض الجدول السابق رقم (١) أن معظم الأسئلة في النموذج (١) ذات معامل صعوبة مقبول، ماعدا سؤالين رقمهما: (١٣، ٢٠) معامل صعوبتهما غير مقبول، أي أقل من (٠,١٥) أو أعلى من (٠,٨٥)؛ وللنموذج (٢) هناك سؤال واحد فقط معامل صعوبته خارج الحد المقبول ورقمه (٢٠).

وفيما يخص معامل التمييز، فالنموذج الأول يتضمن (٤) أسئلة معاملات تمييزها أقل من (٠,١)، وأرقامها: (٥، ٢٠، ٢١، ٢٥)، أما باقي الأسئلة فهناك سؤالين مقبولة التمييز أرقامها (١٤، ٢٤)، والتسعة عشر سؤالاً الباقية تمييزها عالي.

وفي النموذج الثاني، كانت هناك (٣) أسئلة فقط ذات معاملات تمييز منخفضة، وهي الأسئلة (٢٠، ٢١، ٢٥)؛ أما الأسئلة مقبولة التمييز فأرقامها (٩، ١١، ١٤، ١٩، ٢٢) وباقي الأسئلة تمييزها عال.

صدق وثبات الأداة:

تم التحقق من صدق الأداة الارتباطي من خلال حساب معامل الارتباط الثنائي الأصيل والذي يعد حالة خاصة من معامل ارتباط بيرسون بين السؤال والدرجة الكلية، وقد سبق مناقشته في تحليل أسئلة نموذجي الأداة؛ وتم حساب ثبات الأداة عند اختبار الفرض الثالث لأنه أحد متغيرات البحث الحالي.

نتائج الدراسة:

اختبار الفرض الأول ونصه: لا توجد فروق دالة بين معاملات صعوبة الأسئلة تعود لاستخدام البديل "لا شيء مما سبق".

تم تطبيق اختبار (t) لمجموعتين مستقلتين على معاملات صعوبة الأسئلة، والجدول (٢) يوضح قيمة الإحصاء (t) ودلالاتها.

جدول ٢: نتائج اختبار (t) بين متوسطي معاملات صعوبة أسئلة نموذجي أداة الدراسة

رقم النموذج	متوسط معاملات الصعوبة	الانحراف المعياري	عدد الأسئلة	درجات الحرية	قيمة إحصاء (t)	مستوى الدلالة
نموذج ١	٠,٦٢٤٨	٠,١٩٠١٧٨	٢٥	٤٨	٢,١٧٥	*٠,٠٣٥
نموذج ٢	٠,٥٠٧٦	٠,١٩٠٨٨١	٢٥			

* دال عند ($\alpha = 0,05$)

من جدول (٢) يتضح أن متوسط صعوبة أسئلة النموذج الأول ذي البدائل الحقيقية، قد بلغت القيمة (٠,٦٢٤٨)، ومتوسط صعوبة أسئلة النموذج الثاني المتضمن البديل "لا شيء مما سبق"، بلغت القيمة (٠,٥٠٧٦)، وهذا يشير إلى أن النموذج الثاني متوسط الصعوبة حيث تقترب قيمة متوسط الصعوبة من القيمة (٠,٥)، بينما يشير متوسط صعوبة أسئلة النموذج الأول إلى أنه أكثر سهولة من النموذج الثاني؛ كما تشير قيمة إحصاء (t) إلى أن الفروق بين النموذجين في متوسط صعوبة الأسئلة هي فروق دالة؛ وهذا يعني أن النموذج الذي يتضمن البديل "لا شيء مما سبق" أصعب بصورة دالة من النموذج ذي البدائل الحقيقية.

اختبار الفرض الثاني ونصه: لا توجد فروق دالة بين معاملات تمييز الأسئلة تعود لاستخدام البديل "لا شيء مما سبق".

لاختبار الفرض تم تطبيق اختبار (t) لمجموعتين مستقلتين على متوسط معاملات التمييز للنموذجين، والجدول (٣) يوضح قيمة الإحصاء (t) ودلالاتها.

جدول ٣: نتائج اختبار (t) بين متوسطي معاملات تمييز أسئلة نموذجي أداة الدراسة

رقم النموذج	متوسط معاملات التمييز	الانحراف المعياري	عدد الأسئلة	درجات الحرية	قيمة إحصاء (t)	مستوى الدلالة
نموذج ١	٠,٣٣٣١٦	٠,١٧٦٨٠٤	٢٥	٤٨	٠,٨١٥-	٠,٤١٩
نموذج ٢	٠,٣٧٨٥٦	٠,٢١٥٢٨٩	٢٥			

يوضح الجدول السابق رقم (٣) أن متوسط معاملات التمييز للنموذج الثاني أعلى من متوسط معاملات التمييز للنموذج الأول، إلا أن هذه الفروق بين النموذجين في متوسط معاملات التمييز لم تكن دالة، أي أن وجود البديل "لا شيء مما سبق" لا يؤثر في تمييز السؤال بصورة دالة.

اختبار الفرض الثالث ونصه: لا توجد فروق دالة بين معاملي ثبات نموذجي الأسئلة تعود إلى استخدام البديل "لا شيء مما سبق".

لاختبار دلالة الفرق بين معاملي ألفا لثبات النموذجين لعينتين مستقلتين، تم استخدام معادلة اختبار مربع كاي (χ^2) رقم (٢١)، والتي طورها فلْد وزميليه (Feldt, Woodruff & Salih, 1987)، وأتمتها ديدنخوفن وموش (Diedenhofen & Musch, 2016)، وجدول (٤) يوضح قيمة إحصاءة (χ^2):

جدول ٤: نتائج اختبار مربع كاي (χ^2) لدلالة الفرق بين معاملي ثبات النموذجين

رقم النموذج	معامل الثبات ألفا	عدد الأسئلة	حجم العينة	درجات الحرية	قيمة إحصاءة (χ^2)	مستوى الدلالة
نموذج ١	٠,٨٠٤	٢٥	١١٦	١	١,٢٠٠١	٠,٢٧٣٣
نموذج ٢	٠,٨٤١	٢٥	١٢٤	١	١,٢٠٠١	٠,٢٧٣٣

يتضح من الجدول السابق رقم (٤) ارتفاع معاملات ثبات النموذجين حيث بلغت القيم (٠,٨٠٤) و(٠,٨٤١) للنموذجين (١) و(٢) على الترتيب؛ كما تشير قيمة إحصاءة (χ^2) إلى عدم وجود فروق دالة بين معاملي ثبات النموذجين.

مناقشة النتائج والتوصيات:

خرج البحث الحالي بأن وجود البديل "لا شيء مما سبق" يزيد من صعوبة السؤال ولا يؤثر في تمييزه ولا في ثبات الاختبار؛ ففيما يختص بأثر البديل على صعوبة السؤال، فإن هذه النتائج تتسق مع أغلب الأبحاث التي تم عرضها (Rich & Johanson, 1990؛ Frary, 1991؛ Dochy et al., 2001؛ Pachai, Dibattista, Sinnige-Egger, & Fortuna, 2014؛ DiBattista, & Kim, 2015)، ولا تتسق مع بحثي (Knowles & Welch, 1992؛ Atalmış & Kingston, 2017) اللذين خرجت دراستهما بأن وجود البديل لا يؤثر على الصعوبة؛ وفيما يخص معامل التمييز، فإن نتائج البحث الحالي تتسق مع نتائج بحث (Dibattista, Sinnige-Egger, & Fortuna, 2014) في أن وجود البديل لا يؤثر على التمييز، ولا تتفق مع نتيجة بحث (Dochy et al., 2001) والذين خرجوا بأن وجود البديل أدى إلى ارتفاع القدرة التمييزية للأسئلة؛ وتتسق نتائج البحث الحالي مع بحثي (Rich & Johanson, 1990؛ Atalmış & Kingston, 2017) في أن وجود البديل لا يؤثر على ثبات الاختبار.

وبناء على ما سبق وعلى الصعيد المرتبط بالخصائص السيكومترية للمقياس وأسئلته، فإن النتائج قد تشير إلى أن البديل "لا شيء مما سبق" يمثل خيارًا متاحًا لاستخدام معدي المقياس، حيث أن استخدام البديل لا يغير بصورة جوهرية من خصائص المقياس المرتبطة بتمييز السؤال وبثبات الاختبار بين نماذج المقياس متكافئة المحتوى، بل يُحسِّن صعوبة أسئلة النموذج المتضمن لهذا البديل، فمتوسط قيم صعوبة أسئلة أداة البحث للنموذج المتضمن البديل "لا شيء مما سبق" تقترب من القيمة (٠,٥)، بينما متوسط صعوبة أسئلة النموذج الذي يتضمن بدائل حقيقية يزيد عن القيمة (٠,٥) مما يشير إلى سهولتها؛ ورغم أن الفروق في التمييز والثبات لم تكن دالة، إلا أنه من

الملاحظ أن قيم هذه المعاملات للنموذج الثاني -الذي يتضمن البديل- أعلى من قيمها للنموذج الأول، وإن لم تصل الفروق بينهما إلى مستوى الدلالة.

أما على الصعيد المرتبط بالجانب المعرفي وبوظيفة الاختبار في تثبيت التعلم و/أو تعديله، من خلال وجود المعلومة الصحيحة ضمن بدائل السؤال أثناء الاختبار (Butler, 2018)، فإن المؤشرات الموجودة تشير إلى أن استخدامه يجب أن يتم بحذر، خاصة إذا كان واضع الأسئلة مبتدئاً، لأن الدراسات حول أثره على التعلم قليلة، مع وجود مؤشرات من دراسة (Odegard & Koen, 2007) تشير إلى أن الطلاب يميلون لتجنب البديل "لا شيء مما سبق" مما يجعله يؤثر سلباً على درجاتهم إذا كان هو الإجابة الصحيحة؛ كما أن وجود البديل إجابياً، يعني أن السؤال لا يقدم المعلومة الصحيحة التي تساعد في تثبيت و/أو تعديل التعلم.

المراجع:

- أبوعلام، رجاء (٢٠٠٥). *تقويم التعلم*. عمان: دار المسيرة.
- الأمير، بندر (٢٠١٦). الأخطاء الشائعة في بناء مفردات اختبار الاختيار من متعدد و تأثيرها على الخصائص السيكومترية للاختبارات التحصيلية. *المجلة العربية للعلوم الاجتماعية*، ٥(٩)، ١٣٩ - ١٧٥.
- مخائيل، انطانيوس (٢٠١٦). *القياس والتقويم في التربية الحديثة*. دمشق: جامعة دمشق.
- مراد، صلاح أحمد وسليمان، أمين علي (٢٠١٢). *الاختبارات والمقاييس في العلوم النفسية والتربوية: خطوات إعدادها وخصائصها*. القاهرة: دار الكتاب الحديث
- ميزنز، ويليام وليمان، إيرفين (٢٠١٦). *القياس والتقويم في التربية وعلم النفس* (ترجمة هشام الزبيدي، ط٤). العين: دار الكتاب الجامعي.
- النبهان، موسى (٢٠٠٤). *أساسيات القياس في العلوم السلوكية*. عمان: دار الشروق.
- Atalmış, E., & Kingston, N. (2017). Three, four, and none of the above options in multiple-choice items. *Turkish Journal of Education*, 6(4), 143-157. doi:10.19128/turje.333687
- Bacon, D. R. (2003). Assessing learning outcomes: A comparison of multiple-choice and short answer questions in a marketing context. *Journal of Marketing Education*, 25(1), 31- 36.
- Brimi, H. M. (2011). Reliability of Grading High School Work in English. *Practical Assessment, Research & Evaluation*, 16(17), 1-12.
- Butler, A. C. (2018). Multiple-Choice Testing in Education: Are the Best Practices for Assessment Also Good for Learning? *Journal of Applied Research in Memory and Cognition*, 7(3), 323-331. doi:10.1016/j.jarmac.2018.07.002
- Dibattista, D., Sinnige-Egger, J., & Fortuna, G. (2014). The “None of the Above” Option in Multiple-Choice Testing: An Experimental Study. *The Journal of Experimental Education*, 82(2), 168-183. doi:10.1080/00220973.2013.795127
- Diedenhofen, B., & Musch, J. (2016). cocron: A Web Interface and R Package for the Statistical Comparison of Cronbach’s Alpha Coefficients. *International Journal of Internet Science*, 11(1), 51-60. ISSN 1662 - 5544
- Dochy, F., Moerkerke, G., Corte, E. D., & Segers, M. (2001). The assessment of quantitative problem-solving skills with “none of the above”-items (NOTA items). *European Journal of Psychology of Education*, 16(2), 163-177. doi:10.1007/bf03173023
- Downing, S. M. (2003). Guessing on selected-response examination. *Medical Education*, 37(8), 670- 671
- D'Sa, J., Alharbi, M. F., & Visbal-Dionaldo, M. (2018). The Relationship between Item Difficulty and Non-functioning Distractors of Multiple Choice Questions. *International Journal of Nursing Education*, 10(3), 48-52. doi:10.5958/0974-9357.2018.00065.X

- Feldt, L. S., Woodruff, D. J., & Salih, F. A. (1987). Statistical Inference for Coefficient Alpha. *Applied Psychological Measurement*, 11(1), 93-103. doi:10.1177/014662168701100107
- Frary, R. B. (1991). The none-of-the-above option: An empirical study. *Applied Measurement in Education*, 4(2), 115–124. doi:10.1207/s15324818ame0402_2
- Frey, B. B., Petersen, S., Edwards, L. M., Pedrotti, J. T., & Peyton, V. (2005). Item-writing rules: Collective wisdom. *Teaching and Teacher Education*, 21(4), 357-364. doi:10.1016/j.tate.2005.01.008
- Haladyna, T. M., Downing, S. M., & Rodriguez, M. C. (2002). A Review of Multiple-Choice Item-Writing Guidelines for Classroom Assessment. *Applied Measurement in Education*, 15(3), 309-333. doi:10.1207/s15324818ame1503_5
- Kiat, J. E., Ong, A. R., & Ganesan, A. (2017). The influence of distractor strength and response order on MCQ responding. *Educational Psychology*, 38(3), 368-380. doi:10.1080/01443410.2017.1349877
- Knowles, S. L., & Welch, C. A. (1992). A Meta-Analytic Review of Item Discrimination and Difficulty in Multiple-Choice Items Using "None-Of-The-Above". *Educational and Psychological Measurement*, 52(3), 571-577. doi:10.1177/0013164492052003006
- Mccoubrie, P. (2004). Improving the fairness of multiple-choice questions: A literature review. *Medical Teacher*, 26(8), 709-712. doi:10.1080/01421590400013495
- Odegard, T. N., & Koen, J. D. (2007). “None of the above” as a correct and incorrect alternative on a multiple-choice test: Implications for the testing effect. *Memory*, 15(8), 873-885. doi:10.1080/09658210701746621
- Pachai, M. V., DiBattista, D., & Kim, J. A. (2015). A Systematic Assessment of ‘None of the Above’ on Multiple Choice Tests in a First Year Psychology Classroom. *Canadian Journal for the Scholarship of Teaching and Learning*, 6(3), 1-17. doi:10.5206/cjsotl-rcacea.2015.3.2
- Rich, C. E., & Johanson, G. A. (1990, April). An item-level analysis of “none of the above”. Paper presented at the annual meeting of the American Educational Research Association, Boston, MA. Retrieved May 28th, 2019, from <http://files.eric.ed.gov/fulltext/ED400299.pdf>
- Scialfa, C., Legare, C., Wenger, L., & Dingley, L. (2001). Difficulty and Discriminability of Introductory Psychology Test Items. *Teaching of Psychology*, 28(1), 11-15. doi:10.1207/s15328023top2801_03
- Sowdani, K., Jaber, H., & Thomas, A. (2018). Item Analysis as a Tool for Educational Assessment as Compared to Students, Evaluation to lectures. *Al Mustansiriyah Journal of Pharmaceutical Sciences*, 18(2), 105-113. doi:10.32947/ajps.18.02.0380
- [Varm, S. \(2003\). Preliminary Item Statistics Using Point-Biserial Correlation and P-Values. Educational Data System, Inc. Retrieved July 4, 2019, from: \[https://eddata.com/wp-content/uploads/2015/11/EDS_Point_Biserial.pdf\]\(https://eddata.com/wp-content/uploads/2015/11/EDS_Point_Biserial.pdf\)](#)
- Wainer, H., & Thissen, D. (1993). Combining multiple-choice and constructed-response test scores: Toward a Marxist theory of test construction. *Applied Measurement in Education*, 6(2), 103- 118.

Walsh, C. M., & Seldomridge, L. A. (2006). Critical thinking: Back to square two. *Nursing Education*, 45(6), 212- 219.