كلية التربية
كلية معتمدة من الهيئة القومية لضمان جودة التعليم
إدارة: البحوث والنشر العلمي (المجلة العلمية)
=======

استراتيجيات أداء الاختبار وتأثيرها على مستوى الأداء والخصائص السيكومترية للاختبار وبنوده

إعـداد
د/ ربيع عبده أحمد رشوان
أستاذ علم النفس التربوي المساعد
كلية التربية بقنا –جامعة جنوب الوادي
أستاذ مشارك –كلية التربية –جامعة القصيم

﴿ المجلد الثالث والثلاثين– العدد السادس– أغسطس ٢٠١٧م ﴾
http://www.aun.edu.eg/faculty_education/arabic

ملخص:

هدف البحث الحالي إلى دراسة استراتيجيات أداء الاختبار لدى الطلاب والتعرف على أكثر هذه الاستراتيجيات استخداماً أثناء الموقف الاختبار وتأثيرها على درجات الطلاب في الاختبار، كذلك التعرف على تأثير استخدام هذه الاستراتيجيات على صعوبة وتمييز بنود الاختبار وصدقه وثباته، ولتحقيق هذه الأهداف تم بناء مقياس لاستراتيجيات أداء الاختبار يتضمن خمس استراتيجيات تمثلت في استراتيجية تخطيط وتنظيم وقت الإجابة واستراتيجية تحليل الأسئلة واستراتيجية تنظيم وتخطيط الإجابة واستراتيجية الضبط والسيطرة واستراتيجية ما بعد الاختبار، وتم كذلك بناء اختبار تحصيلي في مقرر القياس والتقويم، وتكونت عينة الدراسة من ١٨١ طالب من طلاب الدبلوم العام في التربية، وباستخدام الأساليب الإحصائية المناسبة توصل البحث للعديد من النتائج منها:

- تستخدم استراتيجيات أداء الاختبار من قبل الطلاب عينة الدراسة بدرجة متوسطة ماعدا استراتيجية الضبط والسيطرة فكانت بدرجة ضعيفة.
- تتأثر درجات الطلاب في الاختبار التحصيلي بدرجة استخدامهم لاستراتيجيات أداء الاختبار، حيث تزيد الدرجات بزيادة استخدام تلك الاستراتيجيات.
- يمكن التنبؤ بدرجات الطلاب في الاختبار التحصيلي من خلال درجاتهم في استراتيجيات أداء الاختبار وكانت أكثر الاستراتيجيات اسهاماً في التنبؤ هما استراتيجيتي الضبط والسيطرة وتحليل أسئلة الاختبار.
- تزيد معاملات صعوبة البنود (تزداد سهولتها) بزيادة استخدام تلك الاستراتيجيات بينما تقل معاملات تميز البنود.
- يتحسن الصدق التنبؤي والصدق التلازمي للاختبار التحصيلي بزيادة استخدام استراتيجيات أداء الاختبار، وكذلك يتحسن الثبات، لكن تأثر الثبات باستخدام استراتيجيات أداء الاختبار أقل من تأثر الصدق باستخدام هذه الاستراتيجيات.

الكلمات المفتاحية: استراتيجيات أداء الاختبار، الخصائص السيكومترية للاختبار، الخصائص السيكومترية لبنود الاختبار، الصدق، الثبات، الصعوبة، التمييز

Abstract:

The aim of this research is to examine students' test taking strategies, identify the most frequently used during the test situation, and their effect on test scores, Also, the effect of using these strategies on the difficulty and discrimination of the test items and its validity and reliability, To achieve these objectives, a measure of test taking strategies has been developed, including five strategies, such as response time planning, question analysis, response planning, control and management, post-test strategy, And achievement test in the measurement and evaluation course was also developed,The sample consisted of 181 students of the general diploma in education, the research reached many results, including:

- Test taking strategies used by students in a medium degree, except the control and management strategy used in poor degree.
- The scores of students in the test affected by the degree to which they use the test taking strategies, where the use of these strategies increase, the scores increase.
- Students' scores in the test can be predicted through their test taking strategies, The most important strategies for predicting are the control and management and the questions analysis.
- Item difficulty coefficients increased by the increased use of these strategies, while the item discrimination decreased.
- Improved test predictive and congenital validity by increased of the test taking strategies, as well as improved reliably, but the effect of test taking strategies on reliably is less than the effect of this strategies on test validity.

Key Words: test taking strategies, test psychometric properties, items psychometric properties, validity, reliably, difficulty and discrimination

مقدمة:

الفهم العميق للفروق الفردية بين المتعلمين يلقى مزيداً من الاهتمام بين التربويين بصفة عامة وبين المتخصصين في علم النفس التربوي بصفة خاصة، وفي الآونة الأخيرة تزايد الاهتمام بدراسة الكيفية التي يؤدى بها الطلاب في الاختبارات المعرفية المختلفة وبصفة خاصة الاختبارات التحصيلية، بافتراض أن الاختلافات في طريقة الأداء في الاختبار تعد مصدراً من مصادر التباين في نتائج هذه الاختبارات بجانب السمة موضوع القياس؛ فإذا كانت الاختبارات من أهم وسائل تقييم الأداء الأكاديمي للطالب والتعرف على قدراته وإمكانياته المختلفة، فيجب أن تتمتع بقدر كبير من الصدق والثبات، ومن هذا المنطلق يولي التربويون بصفة عامة اهتماماً كبيراً بالكشف عن العوامل المؤثرة على أداء الطالب في الموقف الاختباري، والتي يمكن أن تؤثر على صدق تعبير درجات الاختبار عن السمات المقاسة.

كذلك في حالة استخدام الاختبارات المعرفية في الأغراض البحثية فإن أداء الطالب في الموقف الاختباري قد لا يترتب عليه قرارات تخص الطالب ومستقبله الأكاديمي والمهني، وهو ما يتسبب في عدم توافر دافعية أداء الاختبار لدى العديد من الطلاب، والذي ينتج عنه عدم الجدية في الاجابة وعدم استخدام الاستراتيجيات المناسبة في أداء الاختبار وعدم بذلهم للجهد الكافي في الوصول إلى الإجابة الصحيحة واعتماد البعض منهم على التخمين العشوائي؛ وفي النهاية لا تتسم نتائج القياس بالدقة المطلوبة ولا تعبر الدرجات بصدق عن السمة المقاسة، وبالتالي تكون القرارات المترتبة على نتائج تلك البحوث غير دقيقة (Wolf, Smith & Birnbaum, 1995, 349; Parham, 1996, 3; Volante, 2006, 131; Xu & Wu, 2012, 176; Cui & Roberts, 2013, 41; Pohl, Sudkamp, Hardt, Carstensen & Weinert, 2016, 4؛)

ويشير Reise and Flannery (1996, 10) إلى أن اهتمام علماء علم النفس التربوي والقياس النفسي بقضية جودة الأداء في الموقف الاختباري ترجع لثلاثة دوافع رئيسية تتمثل في الحاجة العامة للتأكد من صدق نتائج القياس وقابلية النتائج للتفسير، وفي تحديد السلوكيات الشاذة للممتحنين في الموقف الاختباري والتي من المحتمل أن تنتج من نقص استخدام استراتيجيات أداء الاختبار أو استخدام استراتيجيات غير مناسبة كالغش والتخمين العشوائي أو عدم الرغبة في إبداء الاستجابة أو من البناء السيء للاختبار، وأخيراً في تحديد العمليات النفسية المسؤولة عن الخطأ في الاستجابة بخلاف مستوى السمة موضوع الاهتمام، وهذه الأمور تؤثر كثيراً على جودة نتائج القياس في ظل القياس جماعي المرجع.

فالتوجهات المعاصرة في القياس تحاول تحديد مصادر التباين في درجات الأداء على أدوات القياس بخلاف مستوى السمة موضوع القياس والتي منها خصائص الممتحنين وذلك من منطلق أن تحييد أثر هذه المصادر يسهم في زيادة جودة القياس وصدق تعبير الدرجات عل المستوى الحقيقي للسمة المقاسة (Takallou, Vahdany, Araghi & Tabrizi, 2015, 119).

فإذا كانت نتائج الاختبار أهم محكات المفاضلة بين الأفراد، فيجب أن نتساءل هل تعكس درجات الاختبار المعارف والمهارات المتعلقة بالمحتوى المستهدف فقط، وهل تعبر بصدق عن المستوى الحقيقي للسمة موضوع القياس؟ وهل هناك عوامل أخرى تؤثر في درجات الاختبار؟ وتكمن الإجابة في أن الأداء على الاختبار يتأثر بالعديد من المتغيرات المعرفية والنفسية بجانب ما وضع الاختبار لقياسه، مثل المحتوى، وقلق الاختبار والاتجاهات نحو موضوع الاختبار والاتجاهات نحو الاختبارات بصفة عامة، واستراتيجيات الأداء في الاختبار، فعلى الرغم من أهمية أن تتحدد درجة الفرد في أداة القياس بمحتوى الأداة ومستوى الصعوبة وكفاءة الفرد إلا أنه توجد العديد من العوامل الأخرى التي تؤثر على أداء الفرد في الموقف الاختباري (Ellis & Ryan, 2003; Dodeen, 2008; 2009; Wise & DeMars, 2010; Bicak, 2013; Banks & Eaton, 2014; Takallou, Vahdany, Araghi & Tabrizi, 2016; Stenlund, Eklof & Lyren, 2016).

ويكاد يكون هناك اتفاق على أن جودة الأداء على أدوات القياس المختلفة تحدد مدى تعبير الدرجات فيها عن المستوى الحقيقي للفرد في السمة موضوع القياس (Blerkom, 2009; Cui & Roberts, 2013; Wise & Kingsbury, 2016; Lon & Iliescu, 2017; Wolf et al., 1995).

ولذا فقد أكدت الجمعية الأمريكية لعلم النفس في عام ٢٠٠٤ في معايير الاختبارات النفسية والتربوية على ضرورة توافر دوافع قوية لأداء المشاركين للاختبار (Barry & Finney, 2009, 11)؛ فالاختبارات المعرفية بصفة عامه تعد من اختبارات الأداء الأقصى والتي تحتاج من المختبر ومؤدي الاختبار بذل أقصى جهد يضمن تعبير الدرجات عن المستوى الحقيقي للسمة المقاسة (Oakland, Callueng & Harris, 2012, 140)؛ وتعد استراتيجيات أداء الاختبار من أهم مؤشرات الأداء الجيد والجاد في الاختبار (Waskiewicz, 2011; Penk, Pohlmann & Roppelt, 2014; Robie, Mueller & Campion, 2001; Xu & Wu, 2012; Arsal, 2014).

ويؤكد Wise, Wise and Bhola (2006, 66) على أن هذه القضية تؤثر على جودة أدوات القياس حتى في حالة استخدام النظرية المعاصرة في تطويرها وتدريج بنودها؛ وهو ما أكدته نتائج دراسة بلال الشواقفة (٢٠١١) حيث اختلفت معالم بنود الاختبار والمقدرة باستخدام النموذج اللوغاريتمي ثلاثي البارامتر باختلاف مستوى مهارات الحكمة الاختبارية؛ والبحث الحالي يحاول التعرف على تأثير درجة استخدام استراتيجيات أداء الاختبار على مستوى الأداء في الاختبار والخصائص السيكومترية لكل من الاختبار وبنوده.

مشكلة البحث:

تتأثر مصداقية تعبير الدرجات عن المستوى الحقيقي للأفراد في السمة المقاسة باثنين من العوامل الجوهرية، يتعلق الأول منهم بالبناء الجيد لأداة القياس والثاني بأداء المشاركين، والعامل الأول يمكن التحكم فيه من قبل واضع الاختبار من خلال اتباع الإجراءات العلمية لبناء أدوات القياس، أما الثاني فيتوقف على العديد من المتغيرات والتي منها المعتقدات الذاتية للممتَحنين والاستعداد للاختبار والتنافسية والقرارات المترتبة على نتائج الاختبار، والجهد المبذول ووقت الاستجابة واستراتيجيات الأداء في الموقف الاختباري، والسياق الذي يحدث فيه الاختبار ومعتقدات الممتَحنين عن أهمية الأداء الجيد في الاختبار وأن درجاتهم يتوقف عليها قرارات ذات أهمية بالنسبة لهم، ومثل هذه القضايا والتي تتناول العوامل التي تؤثر في جودة أو مطابقة أداء الفرد Person Fit نادراً ما تلقى الاهتمام الكافي سواء من الناحية البحثية أو الناحية العملية (Reise & Flannery, 1996, 12; Wise, Ma, Kingsbury & Hauser, 2010, 15; Lon & Iliescu, 2017, 1).

فافتراض أن أدوات القياس بصفة عامة تعمل بطريقة جيدة باختلاف خصائص من تطبق عليهم هو افتراض غير كافي ولذا يجب الحذر عند استخدام أدوات القياس في سياقات قد لا تمثل أهمية بالنسبة لمن تطبق عليهم؛ ويعبر Barry and Wise et al. (2006, 67, 73) Finney (2009, 3-4); عن هذه القضية بصورة أكثر توضيحا بأن عدم وجود حافز لدى مؤدي الاختبار لكي يبذل أقصى جهد ممكن في الإجابة ويستخدم استراتيجيات مناسبة لتحقيق ذلك، يتسبب في العديد من النتائج السلبية المهددة لدقة تعبير درجات الاختبار عن المستوى الحقيقي للسمة موضوع القياس؛ ويشير Amer (2007, 14) إلى أن التوجهات الحديثة في مجال القياس النفسي تؤكد على ضرورة أن تتضمن إجراءات تقنين الاختبارات على معلومات عن كيفية أداء الأفراد على الاختبار وربط تلك المعلومات بمحتوى الاختبار ومستوى الأداء فيه.

بل وذهبت بعض الأدبيات إلى أن الاختبار التحصيلي يوفر نتائج عن نمطين مستقلين من المعرفة الأول وهو المعرفة بموضوع الاختبار والثاني وهو المعرفة بكيفية أداء الاختبار، فامتلاك الطالب للمعرفة المطلوبة في الاختبار ومألوفيته بها لا يضمن الأداء الجيد فيه (Orth, 1995, 4; Dodeen, 2009, 36)؛ ويشير Guion (1998, 478) إلى ضرورة وضع استراتيجيات أداء الاختبار على أجندة البحث العلمي في مجال القياس والتقويم باعتبارها متغير غير مقاس يؤثر في دقة التنبؤات حول السمة المقاسة ويتوسط علاقتها بالمتغيرات الأخرى، وربما يكون السبب في ظهور الفروق بين المجموعات في تلك السمات.

وفي محاولة تحديد الأنماط السلوكية التي تتسبب في الاستجابات غير المناسبة على أدوات القياس أكدت الدراسات أن هذه الأسباب منها عدم الاستجابة بجدية للبنود الأخيرة، والأداء غير المناسب على البنود المتقدمة نظراً لعدم الألفة بنوعية الأسئلة، والغش في الاستجابة، وضعف استخدام استراتيجيات أداء الاختبار (Rudner, Bracey & Skaggs, 1996, 92)؛ وعلى الرغم من أهمية الكشف عن مدى مطابقة الأفراد وجديتهم في الأداء على الاختبار إلا أن البحوث التطبيقية لهذه النتائج البحثية والتأكيدات النظرية في مواقف اختبارية عملية نادرة للغاية (Cui & Roberts, 2013, 40).

فعلى الرغم من التأكيدات السابقة إلا أن الباحث الحالي وفي حدود اطلاعه لم يتوصل إلى دراسات تناولت تأثير استراتيجيات الأداء في الموقف الاختباري باعتبارها مؤشر لدافعية الفرد للمشاركة في الاختبار وتعبر عن مستوى الجهد المبذول للوصول لأفضل وأقصى أداء على الخصائص السيكومترية للاختبار والبنود بصورة صريحة.

ففي البيئة العربية تركزت معظم الجهود البحثية في الكشف عن العوامل المؤثرة على الخصائص السيكومترية للاختبار والبنود في دراسة تأثير عدد البدائل في اختبارات الاختيار من متعدد وموقع المشتت الجيد مثل دراسات (جابر عبدالله، ٢٠١٤، نعمان الموسوي، ١٩٩٧؛ مايسة أبومسلم، ٢٠١٦) أو صياغة المفردات وبدائل الإجابة بطريقة غير جيدة مثل دراسة (إبراهيم يعقوب وباسل أبوفودة، ٢٠١٢؛ بندر الأمير، ٢٠١٦)؛ ووجود أكثر من بديل صحيح مثل دراسة (عصام الجبة، ٢٠١٢)؛ وهنا تساءل يطرح نفسه مفاده هل وجود درجات غير جيدة

ولا تعبر عن المستوى الحقيقي للسمة لدى بعض الأفراد والمترتب على عدم جدية المشاركين في الإجابة يهدد صدق بعض النتائج التي توصلت لها الدراسات السابقة؟ وما هو تأثير هذه الدرجات على الخصائص السيكومترية لأدوات القياس؟ وهل لو تم استبعاد الأفراد غير الجيدين في الأداء على الاختبار تتحسن الخصائص السيكومترية للاختبار وبنوده؟.

فقد لاحظ الباحث شكوى العديد من طلاب الدراسات العليا ومن يستخدمون اختبارات الأداء الأقصى من وجود بعض المشكلات في الخصائص السيكومترية لهذه الأدوات على الرغم من أن العديد منها اختبارات سبق تقنينها والتأكد من كفاءتها السيكومترية، وتبين من مراجعة الدراسات السابقة المرتبطة بموضوع البحث الحالي أن هناك ندرة في الدراسات التي تناولت تأثير بعض المتغيرات النفسية المرتبطة بالموقف الاختباري على تباين الخصائص السيكومترية لأداة القياس، فدراسة هدى العريفي (٢٠٠٨) أكدت نتائجها أن موضع الضبط يؤثر على الخصائص السيكومترية للاختبار بينما لا تختلف تلك الخصائص باختلاف الفاعلية الذاتية، وأشارت الدراسة إلى أن عدم جدية الطلاب في الإجابة عن الاختبار ربما أسهم في عدم وضوح الفروق في خصائصه السيكومترية الراجعة لاختلاف الفاعلية الذاتية؛ ودراسة فريح العطوي (٢٠١٣) أكدت نتائجها على تأثير المخاطرة على الخصائص السيكومترية للاختبار التحصيلي فبزيادة المخاطرة تحدث زيادة مصطنعة في معامل الثبات وصعوبة البنود والتمييز، ودالة معلومات الاختبار.

وأكدت نتائج دراسة بلال الشواقفة (٢٠١١) تأثير تعليم الحكمة الاختبارية وبصفة خاصة التخمين الذكي على خصائص فقرات اختبارات الاختيار من متعدد حيث كانت فقرات الاختبار أسهل وأقل تمييزاً بالنسبة لطلاب المجموعة التجريبية عنها لدى طلاب المجموعة الضابطة، وتتعارض نتائج الدراسة السابقة مع نتائج دراسة باسل أبوفودة (٢٠٠٨) والتي أكدت نتائجها عدم تأثير تعليم الطلاب لحكمة الاختبار على ثبات الاختبار والخصائص السيكومترية للفقرات (صعوبة وتمييز).

كذلك أشارت نتائج دراسة Wise et al. (2006) إلى أن جودة الخصائص السيكومترية للاختبار تتوقف على جدية الأفراد في الإجابة؛ بينما أشارت دراسة Rogers

(1984) إلى عدم اختلاف الخصائص السيكومترية (صدق، ثبات) لأداة القياس باختلاف تعليمات الاستجابة والتي تمثلت في ضرورة بذل الجهد الكافي للوصول للاستجابة الصحيحة ممثلا فيما يمكن اعتباره باستراتيجية مراجعة الإجابة واستراتيجية إدارة الوقت هذا بالنسبة للمجموعة التجريبية بينما تمثلت التعليمات بالنسبة للمجموعة الضابطة في أن الإجابة الصحيحة هي أول إجابة تخطر بذهن الفرد وأنه يجب الانتهاء من الإجابة في أقصر وقت ممكن؛ ودراسة Robie et al. (2001) أشارت نتائجها إلى عدم اختلاف الخصائص السيكومترية لبنود الاختبارات المعرفية (الصعوبة والتمييز) باختلاف دافعية أداء الاختبار، كذلك أشارت نتائج الدراسة إلى عدم اختلاف الخصائص السيكومترية للاختبار (التمييز والثبات، وتباين الخطأ، والصدق البنائي باستخدام التحليل العاملي الاستكشافي والتوكيدي) باختلاف مستوى دافعية أداء الاختبار؛ وباستخدام بيانات مولدة عن طريق المحاكاة أكدت نتائج دراسة Clause (1996) أن استبعاد الاستجابات غير الجادة لم يؤثر كثيراً في ثبات الدرجات وصدقها.

وأشارت نتائج دراسة Maspons and Llabre (1983) إلى تأثير نقص استراتيجيات أداء الاختبار لدى بعض المشاركين على خصائصه السيكومترية حيث أن التدريب على هذه الاستراتيجيات يسهم في تحسين الصدق التنبؤي للاختبار ولكن لا يؤثر في ثباته؛ وأشارت نتائج دراسة Rudner et al. (1996) إلى ان استبعاد الأفراد منخفضي استخدام استراتيجيات أداء الاختبار يسهم في دقة تقدير السمة؛ وأشارت نتائج دراسة Westfall and D'Costa (1987) إلى أن استبعاد الأفراد غير المطابقين باستخدام نموذج راش وهم الذين يؤدون بطريقة غير متسقة على الاختبار أسهم في زيادة دقة تقدير السمة وقدرتها التنبئية؛ وأشارت نتائج دراسة Wise et al. (2006) إلى أن استبعاد الطلاب غير الجادين في الإجابة على الاختبار نتج عنه نقصان في الاتساق الداخلي للدرجات كمؤشر لثبات الدرجات وكذلك نقص في الخطأ المعياري المترتب على نقص الانحراف المعياري للدرجات وزيادة في الصدق التقاربي لدرجات الاختبار.

ومن العرض السابق يتضح أن هناك تعارض بين نتائج الدراسات السابقة فيما يتعلق بتأثير التباين في استخدام استراتيجيات أداء الاختبار ومعدل الجهد المبذول في الوصول لأفضل

أداء في الاختبار على خصائصه السيكومترية وفقراته والذي يرجع في جزء كبير منه إلى تباين الاستراتيجيات موضوع الاهتمام في هذه الدراسات والخلط بينها وبين الحكمة الاختبارية في كثير من الدراسات.

ومن ناحية أخرى وبمراجعة الدراسات التي تناولت تأثير استراتيجيات أداء الاختبار على مستوى الأداء في الاختبار؛ اتضح للباحث أنه هذه الدراسات توصلت إلى نتائج متناقضة في بعض الأحيان، والناتج من عدم وضوح مفهوم استراتيجيات أداء الاختبار في الكثير من الدراسات السابقة واعتباره مرادف لمهارات الحكمة الاختبارية بمفهومها الشائع في بعض الأحيان فمهارات الحكمة الاختبارية وكما يذكر Xu and Wu (2012, 176) تناسب فقط أسئلة الاختيار من متعدد؛ كذلك هذا التعارض ناتج لعدم التمييز بين الاستراتيجيات العامة لأداء الاختبارات والاستراتيجيات الخاصة التي تناسب أنواع معينة من الاختبارات كاختبارات الاختيار من متعدد أو اختبارات الكفاءة اللغوية.

فاستراتيجيات الأداء في الاختبار لا تتضمن فقط السلوكيات الصريحة للطالب أثناء الموقف الاختباري ولكن أكثر من ذلك تتضمن القرارات المعرفية والأفكار الإيجابية وضبط التأثير السلبي للقلق (Holzer, Madaus, Bary & Kehle, 2009, 46; Zhang, Liu, Zhao & Xie, 2011, 165)؛ فبملاحظة أداء الطلاب في الاختبارات نجد أن بعض الطلاب ممن لديهم المعرفة الجيدة بأسئلة الاختبار واستعدوا جيداً له، قد لا يحققون المستويات المتوقعة منهم والذي يرجع إلى تأثر أداء الفرد في الاختبار بالعديد من المتغيرات النفسية والمعرفية والشخصية بجانب السمة موضوع القياس (Dodeen, 2015, 109)؛ وهو ما يجب أخذه بعين الاعتبار عند تفسير نتائج الاختبارات حتى لا تبنى عليها قرارات غير دقيقة (إياد حمادنة، ٢٠١١، ٢٩٧).

فقد أكدت العديد من نتائج الدراسات السابقة على أن من يستخدم استراتيجيات أداء الاختبار تزداد احتمالية نجاحه في الاختبار وأن هذه الاستراتيجيات تعد من العوامل المميزة بين

مرتفعي ومنخفضي الأداء (Hong, Sas & Sas, 2006; Lancaster, Schumaker, Lancaster & Deshler, 2009; Dodeen, 2015; Takallou et al., 2016)؛ وأن التدريب على هذه الاستراتيجيات يحسن دافعية أداء الاختبار ويخفض من قلق الاختبار (Maspons & Llabre, 1983; Orth, 1995; Ellis & Ryan, 2003) إلا أن غالبية هذه الدراسات تناولت الدرجة الكلية لاستراتيجيات أداء الاختبار ولم تحاول تحديد الأهمية النسبية لهذه الاستراتيجيات؛ هذا من جهة ومن جهة أخرى تناولت بعض هذه الدراسات استراتيجيات أداء الاختبار كمرادف لمهارات الحكمة الاختبارية وفي اختبارات الاختيار من متعدد بصفة خاصة.

فدراسة مجدي الشحات (٢٠٠٧) أكدت نتائجها فاعلية التدريب على مهارات الحكمة الاختبارية على تحسين تحصيل الطلاب وخفض قلق الاختبار، كذلك توصلت دراسة Holzer et al. (2009) إلى أن نقص استراتيجيات أداء الاختبار يمكن اعتباره من أسباب تدني تحصيل الطلاب ذوي صعوبات التعلم، والذي يظهر في نقص قدرتهم على إظهار معارفهم الحقيقية في الاختبار والبرهنة على الاتقان.

ولكن هناك بعض النتائج البحثية التي لم تدعم النتائج السابقة وتتعارض معها؛ فقد انتهت دارسة Arsal (2014) إلى أن استراتيجيات أداء الاختبار تربطها علاقة ضعيفة بقلق الاختبار؛ وانتهت نتائج دراسة Zhang et al. (2011) إلى أن استراتيجيات الأداء في الاختبار تربطها علاقات ضعيفة بالأداء في اختبارات اللغة الإنجليزية؛ وانتهت دراسة محمد عبدالوهاب (٢٠٠٧) إلى أن التدريب على مهارات الحكمة الاختبارية لا يسهم كثيراً في خفض قلق الاختبار؛ والبحث الحالي يحاول التعرف على اسهامات استراتيجيات أداء الاختبار في التنبؤ بدرجات الطلاب في الاختبار التحصيلي؛ وتتحدد مشكلة البحث في محاولة الإجابة عن الأسئلة التالية:

١- ما استراتيجيات أداء الاختبار الأكثر استخداماً لدى الطلاب عينة البحث الحالي؟

٢- هل توجد فروق دالة إحصائياً في درجات الطلاب عينة البحث الحالي في الاختبار التحصيلي راجعة لاختلاف درجة استخدامهم لاستراتيجيات أداء الاختبار؟

٣- هل يمكن التنبؤ بدرجات الطلاب عينة البحث الحالي في الاختبار التحصيلي من خلال درجاتهم في استراتيجيات أداء الاختبار؟

٤- هل تختلف الخصائص السيكومترية لبنود الاختبار التحصيلي (الصعوبة، التمييز) باختلاف استراتيجيات أداء الاختبار لدى الطلاب عينة البحث الحالي؟

٥- هل تختلف الخصائص السيكومترية للاختبار التحصيلي (الصدق، الثبات) باختلاف استراتيجيات أداء الاختبار لدى الطلاب عينة البحث الحالي؟

مصطلحات البحث:

- استراتيجيات أداء الاختبار:

تشير إلى المهارات والفنيات المتعلقة بكيفية الأداء في الاختبار والتي تمكن الفرد من إدارة وقت الاختبار والسيطرة على مشاعره السلبية وعرض وتنظيم إجاباته في الاختبار بالطريقة التي تتيح له الحصول على أعلى درجة ممكنة بما يتفق مع مستواه في السمة موضوع القياس، ويمكن تطبيقها في كافة أنواع الاختبارات ولا تقتصر على اختبارات الاختيار من متعدد، وتتحدد إجرائياً بالدرجة التي يحصل عليها الطالب في المقياس المستخدم في البحث الحالي وأبعاده الفرعية.

- الاختبار التحصيلي:

يرى "براون" أن الاختبار التحصيلي أداة لقياس التعلم الذي يحدث نتيجة الخبرة في موقف تعلم يتم التحكم فيه مثل الصف الدراسي أو البرنامج التدريبي (صلاح علام، ٢٠١٢، ١٢٢-١٢٣) ؛ ويقصد به في البحث الحالي الاختبار التحصيلي لمقرر القياس والتقويم الذي يقوم الباحث بتدريسه لطلاب الدبلوم العام في التربية بجامعة القصيم بالمملكة العربية السعودية.

- مستوى الأداء في الاختبار التحصيلي:

ويُعرف إجرائياً بالدرجة التي يحصل عليها الطالب في الاختبار التحصيلي المستخدم في البحث الحالي.

- الخصائص السيكومترية للاختبار وبنوده:

وتشير إلى الكفاءة السيكومترية للاختبار التحصيلي المستخدم في البحث الحالي متمثلة في ثبات الاختبار باستخدام معادلة كيودر وريتشاردسون KR-20 وصدق الاختبار التحصيلي باستخدام الصدق التلازمي والصدق التنبؤي، وبالنسبة لبنود الاختبار يقصد بها معاملات صعوبة البنود ومعاملات التمييز كما تقاس بالطرق المعتادة المستخدمة لهذا الغرض.

أهداف البحث:

يهدف البحث الحالي إلى التعرف على:

١- استراتيجيات أداء الاختبار الأكثر استخداماً، ودرجة استخدام الطلاب لهذه الاستراتيجيات.

٢- مدى اختلاف درجات الطلاب في الاختبار التحصيلي باختلاف درجة استخدامهم لاستراتيجيات أداء الاختبار.

٣- استراتيجيات أداء الاختبار التي تسهم في التنبؤ بدرجات الطلاب في الاختبار التحصيلي وأكثر هذه الاستراتيجيات تأثيراً على مستوى الأداء في الاختبار.

٤- مدى اختلاف الخصائص السيكومترية لبنود الاختبار التحصيلي (الصعوبة، التمييز) باختلاف استخدام الطلاب لاستراتيجيات أداء الاختبار.

٥- مدى اختلاف الخصائص السيكومترية للاختبار التحصيلي (الصدق، الثبات) باختلاف استخدام الطلاب لاستراتيجيات أداء الاختبار.

أهمية البحث:

من الممكن وأن تسهم نتائج البحث الحالي في:

- توجيه مستخدمي اختبارات الأداء الأقصى ومطوري هذه الاختبارات لضرورة توفير الظروف التي تضمن بذل المشاركين لأقصى جهد ممكن في الوصول للإجابة الصحيحة، فقد تسهم النتائج في توضيح تأثير التباين في استخدام هذه الاستراتيجيات على الخصائص السيكومترية للاختبار وبنوده.
- توجيه من يستخدمون الاختبارات المعرفية لأغراض بحثية لأهمية وجود دوافع قوية لدى المشاركين لأداء الاختبار، وأنه يمكن تدريب المشاركين على استراتيجيات أداء الاختبار قبل تطبيق هذه الاختبارات للحد من تأثير التباين بينهم في هذه الاستراتيجيات على الخصائص السيكومترية للاختبار وبنوده.
- توضيح الأهمية النسبية لاستراتيجيات أداء الاختبار من حيث تأثيرها على أداء الطلاب في الاختبار التحصيلي وهو ما يمكن الاستفادة منه في تحسين أداء الطلاب بالتدريب على هذه الاستراتيجيات.
- لفت نظر المهتمين بتطوير المناهج وطرق التدريس لأهمية هذه الاستراتيجيات وأهمية تضمينها في مقررات الإعداد العام، وضرورة أن تؤخذ بعين الاعتبار من قبل القائمين بالتدريس.

الإطار النظري والدراسات السابقة:

أولاً: استراتيجيات أداء الاختبار: Test Taking Strategies

فضل الباحث ترجمة هذا المصطلح على انه استراتيجيات أداء الاختبار بدلاً من استراتيجيات أخذ الاختبار أو استراتيجيات تناول الاختبار، وذلك لأن هذه الترجمة توضح المقصود بالمفهوم بصورة أدق حيث تصف ما يقوم به مؤدي الاختبار قبل وأثناء وبعد الاختبار، وهو ما يتفق مع ما يؤكد عليه Dodeen, Abdelfatah and Alshumrani (2014, 1) حيث تعبر هذه الاستراتيجيات عن مهارات الطالب المعرفية التي تتيح له الأداء في الاختبار بطريقة مناسبة ومعرفة ما عليه فعلة قبل وأثناء وبعد الاختبار، وتمكنه من ترجمة ما تعلمه من معارف ومهارات والاستفادة منها في الإجابة على أسئلة الاختبار وتسهم إيجابياً في كفاءة الأداء في الاختبار وبالتالي الإنجاز الأكاديمي.

ويعرفها Kim and Goetz (1993, 212) بأنها "العمليات المعرفية وما وراء المعرفية التي يعتمد عليها الفرد أثناء قراءة أسئلة الاختبار والإجابة عنها"؛ وتعرفها Parham (1996,

(3 بأنها "أي فنيات أو قواعد أو إجراءات تستخدم لتفسير أسئلة الاختبار والإجابة عنها، ويمكن استخدامها في العديد من المواقف الاختبارية باختلاف أنماط الاختبارات"؛ ويعرفها Nguyen (2003, 37, 38) بأنها "مهارات ومعارف الفرد المتعلقة بكيفية الأداء في الاختبار والتي تمكنه من عرض وتنظيم إجاباته بما يتفق مع مستوى قدراته، ويمكن تطبيقها في كافة أنواع الاختبارات"؛ ويعرفها Dodeen (2015, 108) بأنها "القدرات المعرفية التي تمكن الفرد من التعامل مع أي موقف اختباري بطريقة جيدة ومعرفة ما الذي يجب فعله أثناء الاختبار"؛ ويعرفها Stenlund et al. (2016, 5) بأنها "مجموعة من الإجراءات التي تستخدم لفهم أسئلة الاختبار والإجابة عنها بطريقة صحيحة".

والتعريفات السابقة لاستراتيجيات أداء الاختبار تؤكد على عمومية هذه الاستراتيجيات ومناسبتها لكافة أنواع الاختبارات، كذلك تؤكد على أن هذه الاستراتيجيات تساعد الفرد في فهم أسئلة الاختبار وعرض وتنظيم إجابته عليها، بما يتيح له تحقيق درجة مناسبة تعبر عن مستواه في السمة موضوع القياس.

التداخل بين مفهوم استراتيجيات أداء الاختبار ومفهوم الحكمة الاختبارية:

لاحظ الباحث من خلال الاطلاع على البحوث والدراسات المتاحة في البيئة العربية الاهتمام بمصطلح الحكمة الاختبارية بشكل واسع ولم يتوصل لأي دراسة تناولت مفهوم استراتيجيات أداء الاختبار، وعلى الرغم من اعتبار أن مصطلح الحكمة الاختبارية مرادف لمصطلح استراتيجيات أداء الاختبار في العديد من الدراسات السابقة (Maspons & Llabre, 1983; Orth, 1995; Eme & Udoh, 2011) إلا أنه يمكن اعتبار استراتيجيات أداء الاختبار جزء من الحكمة الاختبارية بمعناها الشامل وليس معناها التقليدي الشائع في الدراسات ذات الصلة؛ أما الحكمة الاختبارية بمعناه الشائع في الدراسات السابقة وكما يشير Stenlund et al. (2016, 5) فهي تناسب بدرجة كبيرة اختبارات الاختيار من متعدد وليس لها صفة العمومية.

فالحكمة الاختبارية في معظم الدراسات السابقة عرفت بأنها قدرة الفرد على الاستفادة من خصائص ومواصفات الاختبار والموقف الاختباري للحصول على درجة مرتفعة نتيجة الإفادة من

خبرة التعامل مع الاختبارات وأنها قدرة الطالب على تخمين ما يهدف إليه واضع الاختبار وفحص ورقة الامتحان للبحث عن مفاتيح إجابة إضافية تسهم في الوصول للإجابة الصحيحة (أحمد عودة، ١٩٨٩، ١٢٨؛ إيثار الخصاونة، ١٩٩٤، ٥؛ مجدي الشحات، ٢٠٠٧، ٤؛ باسل أبوفودة، ٢٠٠٨، ٢٧؛ ديانا حماد، ٢٠١٠، ٣٠٢؛ بلال الشواقفة، ٢٠١١، ٢٠؛ إبراهيم يعقوب وباسل أبوفودة، ٢٠١٢، ٤٣٨؛ Amer, 2007, 3).

وهنا يجب أن نتساءل هل تحرى مؤدي الاختبار للثغرات الممكنة في الاختبار والاستفادة منها ومن أخطاء بناء الاختبار في تحسين درجته بغض النظر عن مستوى السمة يعتبر حكمة أو حنكة اختبارية؟.

فإذا كانت الحكمة الاختبارية هي مجرد محاولة الطالب الاستفادة من أخطاء القياس المتمثلة في الأدلة والمفاتيح الممكن الاستفادة منها في الإجابة على بعض أسئلة الاختبار لكان من الأولى توجيه الاهتمام لتدريب واضعي الاختبارات على تحري الدقة في بناء الاختبارات وعدم وجود مثل هذه الأخطاء؛ فعلى الرغم من اعتبار أن الحكمة الاختبارية مصدراً من مصادر الفروق في الأداء على الاختبار، وأن اللجوء إليها يعتبر حقيقة ثابته، إلا أن هناك العديد من المهتمين أكدوا على أن أصحاب الحكمة الاختبارية يحصلون على درجات أعلى غير مستحقة، خاصة في الاختبارات الموضوعية؛ وهو ما يتضح من مفهوم ميلمان وبيشوب وإبل Millman, Bishop and Ebel في ١٩٦٥ عن الحكمة الاختبارية والذي يعد هو المفهوم الأساسي المعتمد عليه في معظم الدراسات السابقة المتعلقة بالحكمة الاختبارية حيث يؤكد هذا المفهوم على أن الحكمة الاختبارية تعني مقدرة المفحوص على الاستفادة من الموقف الاختباري وخصائص الاختبار والصياغات غير الجيدة للأسئلة وما تحوية من مؤشرات تدل على الإجابة الصحيحة في الحصول على درجات أعلى دون الإتقان الفعلي للمحتوى (ديانا حماد، ٢٠١٠، ٢٩٧؛ فاطمة مطلك، ٢٠٠٩، ٥٦٦، ٥٧١؛ سامي العنزى، ٢٠١٤، ٣٥٠؛ Rogers & Harley, 1999, 345; Chittooran & Dorothy, 2001, 11; Xu & Wu, 2012, 176).

فقد توصلت دراسة Ghafournia and Afghari (2014, 91) إلى أن مهارات الحكمة الاختبارية كما يعبر عنها في تلك الدراسة باستراتيجيات أداء الاختبار تفسر حوالي ١٨% من التباين في درجات طلاب الدراسات العليا في اختبار الفهم القرائي في اللغة الإنجليزية بجانب الكفاءة اللغوية، وهو ما تم اعتباره دليلاً على عدم صدق درجات الاختبار في التعبير عن الكفاءة اللغوية، وأشارت الدراسة إلى أن استراتيجيات أداء الاختبار يمكن اعتبارها من أهم مصادر خطأ القياس التي لا يمكن تجاهلها في تفسير نتائج القياس؛ بل وفسرت دراسة ديانا حماد (٢٠١٠) عدم وجود علاقات دالة إحصائياً بين الحكمة الاختبارية والتحصيل بأن ذلك راجع لجودة بناء الاختبار وفق نموذج راش وهو ما أسهم في الحد من تأثير الحكمة الاختبارية، وهو ما يؤكد صراحة على أن الحكمة الاختبارية أمر يجب الحد من تأثيره على درجات الأفراد في الاختبار؛ على عكس ما هو شائع في البحوث والدراسات ذات الصلة؛ وأشارت لنفس النتائج تقريبا دراسة Rogers and Harley (1999) في تفسيرها لمميزات الاعتماد على ثلاثة بدائل بدلاً من أربعة بدائل في أسئلة الاختيار من متعدد بأن الاعتماد على ثلاثة بدائل يقلل من تأثير الحكمة الاختبارية.

وتوصلت لنفس النتيجة السابقة تقريباً دراسة سعود العنزي (٢٠١٤) بأنه ليست هناك علاقة بين مهارات الحكمة الاختبارية والتحصيل في الاختبارات التي يُعدها المركز الوطني للقياس والتقويم بالمملكة العربية السعودية وأن طبيعة هذه الاختبارات وطريقة البناء المحكم لها لا تتيح الفرصة للطلاب للاستفادة من خصائص الاختبار وصياغاته للتعرف على الإجابة الصحيحة دون معرفة.

ويتأكد من ذلك أن الحكمة الاختبارية قد تسهم في أن يحصل بعض الطلاب على درجات غير مستحقة نتيجة للاستفادة من بعض الأخطاء المرتبطة ببناء الاختبار والتخمين الذكي، وهو ما يترتب عليه عدم صدق الدرجات في التعبير عن المستوى الحقيقي للسمة.

بل وتذكر فاطمة مطلك (٢٠٠٩، ٥٦٩) أن اكتساب استراتيجيات الحكمة الاختبارية من شأنه إتاحة الفرصة للتلاميذ أن يجيبوا عن بعض الأسئلة الامتحانية التي ليس لديهم معرفة عنها؛ وهو ما يتنافى مع فكرة صدق الاختبار والعدالة في تقدير السمة والحد من مشكلة تحيز الاختبار؛ وكذلك تتضح المشكلة في مفهوم الحكمة الاختبارية من التعريف الذي وضعه جيب Gibb في عام ١٩٦٤ بأن الحكمة الاختبارية هي قدرة المفحوص على الإجابة الصحيحة عن بعض المفردات اعتماداً على أدلة خارجية من أجل الحصول على درجة أعلى دون معرفة حقيقية بموضوع الاختبار (زين ردادي، ٢٠٠١، ٥)؛ ويذكر إياد حمادنة (٢٠١١، ٢٩٣) أن عدم معرفة باني الاختبار بالأسس الصحيحة لبناء الاختبار يسهم في زيادة تأثير استراتيجيات حكمة الاختبار، حيث يستفيد بعض الطلاب في هذه الحالة من الثغرات في بناء الأسئلة في معرفة الإجابة الصحيحة.

وبالتالي الدرجة الإضافية التي يحصل عليها أصحاب الحكمة الاختبارية بجانب الدرجات المتوقعة والراجعة لمستوى السمة الفعلي لا تعكس تحصيلاً أفضل وإنما تعتبر من أنوع خطأ القياس؛ وهو ما دعا الاتحاد الكندي للمعلمين في ١٩٩٩ إلى اعتبار أن مهارات أداء الاختبار كمرادف للحكمة الاختبارية من المهارات غير الأخلاقية والتي يجب تدريب الطلاب على تجنب استخدامها (Volante, 2006, 133)؛ ويري Banks and Eaton (2014, 212) أنه يجب الانتقال من ثقافة الامتحان إلى ثقافة التقويم والتي تؤكد على مفهوم التقويم الأصيل لمعارف الطالب ومهاراته وحتى نصل إلى هذه المرحلة يجب أن تكون مهارات واستراتيجيات أداء الاختبار جزء مكمل للممارسات التدريسية للحد من التحيز الناتج عن الفروق في هذه الاستراتيجيات.

ووجهة النظر السابقة تؤكد الخطر الذي يسببه اللجوء إلي الحكمة الاختبارية على صدق درجات الاختبار في التعبير عن المستوى الحقيقي للسمة موضوع القياس، وهناك وجهة نظر أخرى على عكس وجهة النظر السابقة تؤكد على الخطر الذي يسببه غياب مثل هذه الاستراتيجيات على مستوى أداء الفرد في الاختبار حيث يسبب النقص في هذه الاستراتيجيات تدني في مستوى درجات بعض الطلاب وعدم تعبيرها عن المستوى الحقيقي للسمة في بعض

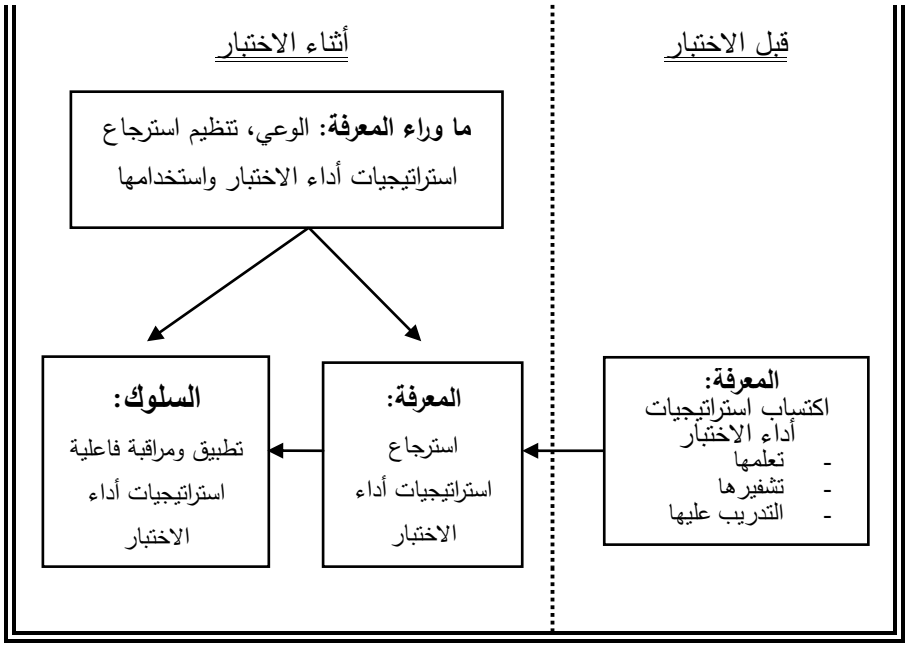
الأحيان، وذلك لسببين يتمثل الأول في ظهور فروق بين الممتحنين لا ترتبط بالفروق الحقيقية في السمة وإنما ترتبط بالفروق في الاستراتيجيات والمهارات التي يلجأ لها بعض الأفراد أثناء الأداء في الاختبار هذا من جهة ومن جهة أخرى يتمثل السبب الثاني في أن عدم توافر هذه الاستراتيجيات لدى الفرد قد يعوق إظهاره لمستوى معرفته الحقيقية في الاختبار وعدم قدرته على البرهنة على الإتقان.

وفي هذا الصدد يؤكد السيد أبوهاشم (٢٠٠٨، ٢١٥) أنه يجب ألا يفهم أبداً أن امتلاك مهارات الحكمة الاختبارية وحده يكون كافياً لكي ينجح الطالب دون استعداد جيد للاختبار فهذه المهارات تساعد في الحصول على أقصى درجة تسمح بها معلومات الفرد ودرجة استعداداه للاختبار؛ وفي مراجعة زين ردادي (٢٠٠١، ١١) للدراسات التي تناولت الحكمة الاختبارية أشار إلى أن بعض الدراسات أكدت على أن فاعلية استخدام الحكمة الاختبارية مرتبط بالمعرفة موضوع الاختبار وعلق الباحث على هذه النتيجة بأنها متعارضه مع مفهوم الحكمة الاختبارية والذي يؤكد على أنها مستقلة عن محتوى الاختبار وفسر ذلك بأن الطالب المتفوق هو الذي يستطيع الاستفادة من هذه المهارات أكثر من غيره، وهو ما يؤكد مرة أخرى على أن مهارات الحكمة الاختبارية ليست مجرد الاستفادة من قدرات الطالب ومهاراته في الحصول على درجات غير مستحقة، وأن هذه المهارات غير مستقلة عن المعرفة بمحتوى الاختبار؛ ويتفق ذلك مع ما أكد عليه Rogers and Harley (1999, 235) بأن فاعلية الحكمة الاختبارية مرتبطة بالمعرفة الجزئية بالإجابة المطلوبة للبند في حالة عدم التيقن من الإجابة.

استراتيجيات أداء الاختبار والحكمة الاختبارية بمعناه الشامل:

وفي محاولة لتحديث مفهوم الحكمة الاختبارية ليدل على البنية المقصودة بمعناها الشامل اقترح Nguyen (2003, 25-28) النموذج التالي الذي يوضح بصورة قاطعة علاقة استراتيجيات أداء الاختبار بالحكمة الاختبارية بمعناه الشامل وليس المعنى الشائع في الدراسات السابقة والذي يتضح من الشكل التالي:

الحكمة الاختبارية



شكل (١): استراتيجيات أداء الاختبار كمكون من مكونات الحكمة الاختبارية (Nguyen, 2003, 27)

وفي هذا النموذج تعتبر استراتيجيات أداء الاختبار مكون من مكونات الحكمة الاختبارية بمعناها الشامل والذي يصف توظيف الممتحن لكل جهوده للوصول لأقصى أداء ممكن، وفي مرحلة ما قبل الاختبار يؤكد النموذج أهمية تعليم هذه الاستراتيجيات والتدريب على استخدامها أثناء الاختبار.

أما أثناء الاختبار فتظهر استراتيجيات أداء الاختبار في قراءة الطالب لكل الخيارات قبل الإجابة عن أسئلة الاختيار من متعدد، وتقرير الإجابة الصحيحة قبل قراءة الخيارات المعروضة واستبعاد الإجابات الخاطئة أثناء الاختيار وقراءة جميع الأسئلة قبل البدء في الإجابة، وكذلك تكوين توقعات عن الأسئلة المحتملة والتدريب على الإجابة، والبحث عن الإجابات الإضافية، ويميز الكلمات المفتاحية ويقوم بعمل المخططات والرسومات التوضيحية إن كان ذلك ممكناً (Holzer et al., 2009, 44; Ghafournia & Afghari, 2014, 90).

كذلك من يستخدمون استراتيجيات أداء الاختبار يقومون بقراءة التعليمات والتخطيط وفهم السؤال جيداً وتأخير الإجابة عن الأسئلة الصعبة أو غير المفهومة (Arsal, 2014, 307)؛ فمن أهم خصائص هذه الاستراتيجيات الوعي والهدف، بمعنى وعي مؤدي الاختبار بها وبكيفية تطبيقها والهدف من تطبيقها (Takallou et al., 2016, 178)؛ ويؤكد زين ردادي (٢٠٠١، ٢٥) في تفسير الفروق بين الطلاب العاديين والمتفوقين في مهارات الحكمة الاختبارية بأن الطلاب المتفوقين يتميزون بالمهارات المتعلقة بالتعامل مع ورقة الإجابة حيث أنهم لا يُضيعون الكثير من الوقت في قراءة الأسئلة ويركزون جهودهم في كيفية عرض المعلومة التي لديهم بطريقة منظمة ومرتبة ومنطقية، وهو ما يؤكد أن استراتيجيات أداء الاختبار ترتبط أكثر بطريقة الأداء.

ويتفق الباحث الحالي مع وجهة النظر التي ترى أن استراتيجيات أداء الاختبار من أهم المتغيرات التي يساعد امتلاكها على إظهار الطالب لدرجته الحقيقية في السمة وذلك من خلال بذل الجهد الكافي للوصول للإجابة الصحيحة وعرض الإجابات بالطريقة التي تؤكد وتبرهن معرفته بالمطلوب من أسئلة الاختبار، فهي تساعد الطالب على الوصول لأقصى درجة ممكنه تأهله لها معارفه ومهاراته المرتبطة بمحتوى الاختبار؛ ولذا يؤكد Bicak (2013, 279) أننا بحاجة للتركيز على هذه الاستراتيجيات لتحسين الأداء في الاختبارات.

ويمكن اعتبار استراتيجيات أداء الاختبار تمثل الجانب الإجرائي للحكمة الاختبارية، ولا تعنى حصول الطالب على درجات غير مستحقة أو أن درجة الطالب لا تعبر عن المستوى الحقيقي للسمة المقاسة، بل على العكس من ذلك فهذه الاستراتيجيات تساعد الطالب على إظهار مستواه الحقيقي في السمة وذلك بمساعدته على تنظيم الإجابة والبرهنة على امتلاكه للمعلومات والمهارات المطلوبة والتحليل العميق لما هو مطلوب في السؤال وخاصة الأسئلة المركبة والجديدة والحد من التأثير السلبي لقلق الاختبار، فهي بمثابة التخطيط للأداء الجيد في الاختبار، وتطبيق مهارات ما وراء المعرفة لتحقيق الفرد لأهدافه التي يسعى لتحقيقها من الاختبار.

أهمية استراتيجيات أداء الاختبار:

أكدت معظم الدراسات التي تناولت علاقة استراتيجيات أداء الاختبار بمستوى ونتائج الاختبار إلى تأكيد العلاقة الإيجابية بينهما (Orth, 1995; Parham, 1996; Hong et al., 2006; Eme & Udoh, 2011)؛ ولذا فهناك تأكيدات على أهمية تدريس هذه الاستراتيجيات للمتعلمين وتدريبهم على استخدامها حتى يتم اتقانها وتطبيقها بدقة أثناء الموقف الاختباري (Blerkom, 2009, 174)؛ ولا يعتبر امتلاك الطلاب لهذه الاستراتيجيات شرطاً كافياً لاستخدامها وتطبيقها في الموقف الاختباري وإنما يتوقف كذلك على المتغيرات الدافعية والتي من أهمها الفاعلية الذاتية (Bicak, 2013, 280)؛ بل ويقترح Banks and Eaton (2014, 212) أنه يجب أن تكون مهارات واستراتيجيات أداء الاختبار جزء مكمل للمنهج الدراسي حتى تصبح بمثابة عادات يلجأ لها الفرد أثناء الموقف الاختباري.

وهذه الاستراتيجيات منها ما يتميز بإمكانية تطبيقها في مدى واسع من الموضوعات والاختبارات والشروط والإجراءات بل ويمكن أن يستفيد منها الفرد في حياته العملية كالقدرة على إدارة الوقت بفاعلية ووضع الأهداف والتخطيط والعمل بجد ودقة وسرعة وإيجاد الأدلة للبرهنة على صحة الأفكار والمعلومات، وهناك من يعتبر أن امتلاك الطالب لهذه الاستراتيجيات والمهارات لا يقل أهمية عن امتلاكه للمعرفة والمهارات المطلوبة للإجابة عن أسئلة الاختبار نظراً لدورها في مساعدة الطالب على الأداء الجيد في الاختبار وبالتالي تحقيق مستوى مرتفع من الإنجاز الأكاديمي(Dodeen et al., 2014, 1,3)؛ كذلك ترتبط استراتيجيات أداء الاختبار بالتعامل مع الانفعالات السلبية والتي منها القلق في الموقف الاختباري والذي ينخفض بالتدريب على استراتيجيات أداء الاختبار (Hong et al., 2006, 145).

ويؤكد على ذلك Dodeen (2009, 32) بأن تلك الاستراتيجيات تساعد الطالب من خلال الاستخدام الجيد لمعارفه ووقت الاختبار والجهد المبذول في الإجابة على أسئلة الاختبار وفي نفس الوقت تحسن اتجاهات الطلاب نحو الاختبارات ونحو التعلم بصفة عامة؛ وكذلك يشير Dodeen (2008, 410) إلى أن تلك الاستراتيجيات تميز الطلاب الذين يستطيعون الأداء على الاختبار بطريقة أفضل من زملائهم من نفس مستوى القدرة فهؤلاء الطلاب يمتلكون استراتيجيات أو مهارات تساعدهم على الأداء الجيد في الاختبار، حيث يتميز هؤلاء بمعرفة ما يجب فعله قبل وأثناء وبعد الاختبار.

ويشير Orth (1995, 6) ; Banks and Eaton (2014, 207) إلى أن نقص استراتيجيات أداء الاختبار يحرم الطالب من الحصول على درجات مستحقة في الاختبار وهو ما يترتب على فقدانهم للطريقة السليمة في الإجابة على الاختبار والتعامل مع أسئلته؛ فالأداء الجيد على الاختبارات يحتاج من المشاركين الجدية في الأداء وبذل الجهد الكافي للإجابة عن أسئلته وفي حالة عدم توافر ذلك فإن درجات الطلاب في الاختبار لا تعكس مستواهم الحقيقي في السمة موضوع القياس.

استراتيجيات أداء الاختبار في البحث الحالي:

اعتمدت العديد من الدراسات الأجنبية في تصنيفها لاستراتيجيات أداء الاختبار على تصنيف مهارات الحكمية الاختبارية كما أشار إليه ميلمان وبيشوب وإبل Millman, Bishop and Ebel في ١٩٦٥؛ وظهرت العديد من استراتيجيات أداء الاختبار في تصنيف مهارات الحكمة الاختبارية كاستراتيجيات التعامل مع ورقة الأسئلة والتعامل مع ورقة الإجابة والمراجعة وإدارة وقت الاختبار والتعامل مع صيغ الأسئلة المختلفة ووضع خطة الحل (زين ردادي، ٢٠٠١، ٦–٧؛ مجدي الشحات، ٢٠٠٧، ٢٠؛ ديانا حماد، ٢٠١٠، ٣٠٢–٣٠٤؛ فاطمة مطلك، ٢٠٠٩، ٥٧٢–٥٧٣ ؛ السيد أبوهاشم، ٢٠٠٨، ٢٢٦–٢٢٧؛ سعود العنزي، ٢٠١٤، ٣٤٢).

وينظر لاستراتيجيات أداء الاختبار باعتبارها الاستراتيجيات التي تمكن الطالب من عمل خطة للإجابة عن الاختبار وتنظيم الوقت بحكمة والاستعداد العقلي للموقف الاختباري وقراءة التعليمات والتوجيهات بدقة وقراءة أسئلة الاختبار بعناية وتذكر المعلومات التي تم استذكارها واستبعاد المعلومات المشتتة والقراءة العامة لأسئلة الاختبار لتحديد الأسئلة المعروف اجابتها وتأخير الإجابة عن الأسئلة غير المعروفة والتخمين الذكي لإجابة الأسئلة غير المعروفة والمراجعة النهائية لاكتشاف الأسئلة التي لم يجاب عنها أو أُجيب عنها إجابات خاطئة (Lancaster et al., 2009, 166)؛ وأشارت دراسة Bicak (2013) إلى أن استراتيجيات أداء الاختبار تتمثل في استراتيجية تحليل البنود واستراتيجية تخطيط الوقت، واستراتيجية توقع الاختيار، واستراتيجية ما بعد الاختبار، والاستراتيجيات الاجتماعية.

وصنفت دراسة إياد حمادنه (٢٠١١، ٢٩٤) استراتيجيات حكمة الاختبار في ثلاثة مجالات رئيسية تتمثل في استراتيجيات تستخدم قبل البدء في الإجابة ومنها قراءة جميع الأسئلة والبدء بالسؤال الأسهل، واستراتيجيات تستخدم خلال الإجابة على الاختبار ومنها استغلال الوقت وتجنب أخطاء التخمين واستراتيجيات تستخدم بعد الانتهاء من الإجابة ومنها إعادة قراءة جميع الأسئلة للتأكد من فهمها بشكل صحيح وكذلك مراجعة الإجابات؛ بينما افترضت دراسة Dodeen (2008); Arsal (2014) أن استراتيجيات أداء الاختبار تتمثل في مهارات ما قبل الاختبار، وإدارة الوقت، وأثناء الاختبار، وبعد الاختبار.

وباستقراء الدراسات السابقة والاستراتيجيات التي تم افتراضها سواء عند استخدام مفهوم الحكمة الاختبارية أو استراتيجيات أداء الاختبار وبصفة خاصة الاستراتيجيات التي يتم تطبيقها فعلا أثناء الاختبار، يتبنى البحث الحالي مجموعة من الاستراتيجيات التي أكدت عليها الدراسات السابقة والتي تناسب المواقف الاختبارية المختلفة بغض النظر عن نوعية أسئلة الاختبار، حيث أن هناك من يفترض أن معظم الاستراتيجيات التي فرضت تحت مظلة الحكمة الاختبارية بمفهومها التقليدي استراتيجيات تناسب أسئلة الاختيار من متعدد فقط وليست استراتيجيات عامة (Crocker, 2006, 116; Takallou et al., 2015, 120) وهذه الاستراتيجيات هي:

١- استراتيجية تخطيط وتنظيم وقت الإجابة:

وهذه الاستراتيجية أكدت عليها معظم نماذج الحكمة الاختبارية تحت مسمى مهارات استخدام الوقت والتي تتمثل في البدء في الإجابة بأسرع وقت ممكن مع الحرص على الاتقان ووضع جدول زمنى للإجابة والإجابة عن الأسئلة السهلة أولاً وتأجيل الإجابة عن الأسئلة الصعبة مع محاولة استغلال ما يتبقى من وقت الاختبار في المراجعة (إياد حمادنة، ٢٠١١، ٢٩٣؛ عفاف وادي، ٢٠١٣، ٣٠٠)؛ وفي دراسة Hong et al. (2006) ودراسة Stenlund et al. (2016, 10) تم تسمية استراتيجية تخطيط وتنظيم وقت الإجابة باستراتيجية التنظيم البنائي Structural Organization وتضمنت استراتيجيات تنظيم الوقت وتوزيع الوقت في ضوء الصعوبة وطول الإجابة وترتيب الإجابة حسب مستوى الصعوبة؛ وفي حالة الأسئلة المقالية يفيد جداً في تنظيم وقت الإجابة أخذ الدرجات المخصصة لكل سؤال بعين الاعتبار (Blerkom, 2009, 170).

وتظهر هذه الاستراتيجية في تحديد وقت الإجابة لكل سؤال حسب الدرجة المخصصة له وعدم تعدية الوقت المخصص للسؤال بقدر الإمكان وتجنب إضاعة الوقت في إجابة سؤال سهل أو درجته أقل وترك وقت للمراجعة وعدم ترك الاختبار قبل انتهاء الوقت (زين ردادي، ٢٠٠١، ٨)؛ وتم في بعض الدراسات اعتبار المراجعة استراتيجية منفصلة لاستراتيجيات أداء الاختبار أو الحكمة الاختبارية كما هو في دراسة السيد أبوهاشم (٢٠٠٨، ٢٢٧) ولكن في أغلب الدراسات تم اعتبارها ضمن استراتيجية إدارة وتنظيم وقت الاختبار.

٢- استراتيجية تخطيط وتنظيم الإجابة:

وهذه الاستراتيجية من أهم استراتيجيات الأداء في الاختبارات المقالية والتي تحتاج لتنظيم أفضل للإجابة ووضوح ومنطقية فيما يتم كتابته وهنا قد يحدد الفرد في البداية النقاط الرئيسية للإجابة ثم يبدأ في تفصيل تلك النقاط (Blerkom, 2009, 170)؛ كذلك هناك بعض الاستراتيجيات التي ظهرت في نماذج الحكمة الاختبارية وتقترب في المعني والهدف من هذه الاستراتيجية مثل استراتيجية تجنب الخطأ والتي يتم فيها تركيز الانتباه أثناء الإجابة لتجنب الخطأ وذلك بالانتباه للتعليمات وقراءة الأسئلة جيداً واستيضاح الغموض ومراجعة الإجابات.

ويسمي السيد أبوهاشم (٢٠٠٨، ٢٢٦) هذه الاستراتيجية بالتعامل مع ورقة الإجابة والتي تتمثل في التفكير في الإجابة لبعض ثواني وتجهيزها ذهنياً قبل كتابتها وتقديم أفضل إجابة ممكنة والاهتمام بعرض عناصر الإجابة في تسلسل منطقي وتجنب التورية والشطب أثناء الإجابة وتخصيص صفحات معينة لكل سؤال حسب كم الإجابة.

وتظهر هذه الاستراتيجية في محاولة بداية إجابة كل سؤال في صفحة جديدة وكتابة رقم السؤال ورأس السؤال أعلى الصفحة ووضع خطاً تحته وإجابة السؤال في حدود المطلوب وتحديد الأفكار الرئيسية المطلوبة في السؤال ووضع خطوط تحت الكلمات المهمة في الإجابة، وتحديد الأفكار الرئيسية وإبرازها والاهتمام بطريقة عرض الإجابة (زين ردادي، ٢٠٠١، ٨).

٣- استراتيجية تحليل الأسئلة:

وسميت هذه الاستراتيجية في بعض الدراسات بمهارات التعامل مع ورقة الأسئلة كما في دراسة السيد أبوهاشم (٢٠٠٨، ٢٢٦) والتي تتمثل في قراءة التعليمات بعناية وفهمها جيداً، وتكوين فكرة عامة عن جميع الأسئلة بقراءتها قبل البدء في الإجابة وتحديد الكلمات الهامة في السؤال وإعادة قراءة السؤال أكثر من مرة للتأكد من المطلوب.

وكذلك تظهر هذه الاستراتيجية في الانتباه لأي تعليمات في السؤال مثل أجب باختصار أو بالتفصيل ومعرفة المطلوب بدقة وعدم الإجابة عن أسئلة غير مطلوبة أو ترك أسئلة مطلوبة (زين ردادي، ٢٠٠١، ٨)؛ ففي هذه الاستراتيجية يقوم الطلاب بوضع خطوط أو دوائر حول المعلومات المهمة في السؤال، وربما يسجل بعض المعلومات الإضافية حول السؤال قبل الإجابة عليه لفهمه جيداً، وتسجيل مفاتيح للإجابة عن السؤال أثناء قراءة السؤال، ووضع علامات على الأسئلة التي لا يستطيع الإجابة عليها، (Kim & Goetz, 1993, 212).

٤ – استراتيجية الضبط والسيطرة:

وهذه الاستراتيجية ظهرت في معظم الدراسات التي تناولت استراتيجيات ومهارات الأداء في الموقف الاختباري وكذلك الدراسات التي تناولت مهارات الحكمة الاختبارية وسميت في بعض الأحيان باستراتيجية ضبط قلق الاختبار (Chittooran & Doroyhy, 2001, 17)؛ وفي بعض الأحيان سميت باستراتيجية خفض قلق الاختبار وزيادة الثقة (Stenlund et al., 2016, 5).

وتظهر في قدرة الفرد على تغيير وضبط أفكاره السلبية ومعتقداته واتجاهاته عن الكفاءة والسلوك ويمكن تحسين هذه الاستراتيجية عن طريق الإعداد الجيد للاختبار وزيادة دافعية أداء الاختبار والتنظيم الذاتي؛ فالقدر القليل من قلق الاختبار قد يعمل على زيادة دافعية الطلاب وتشجيعهم على بذل الجهد بينما زيادة قلق الاختبار قد يشتت الطالب ويعوق أداؤه الجيد في الاختبار وإذا زاد قلق الاختبار بدرجة كبيرة فإن ذلك يعمل على خفض تقدير الذات ونقص معدل التحصيل والفشل في كثير من الأحيان والاتجاهات السلبية نحو الدراسة والمشاعر السلبية والتعصب ولا يتوقف ذلك على عمر معين(Banks & Eaton, 2014, 9).

ويضيف Volante (2006, 134-135) أن هذه الاستراتيجية تسهم في تحسين صدق الاختبار الناتج عن تحسن الأداء في الاختبار، وهذه الاستراتيجية يجب الاهتمام بتدريب الطلاب عليها؛ وأكدت دراسة Nguyen (2003) على هذه الاستراتيجية كاستراتيجية لأداء الاختبار تحت مسمى الضبط والسيطرة والذي يعبر عن قدرة الممتحن على إبعاد أي مؤثرات سلبية قد تؤثر على جودة الأداء في الاختبار، ويسمى السيد أبوهاشم (٢٠٠٨، ٢٢٥) هذه الاستراتيجية بمهارات الضبط النفسي وتتضمن مهارات استخدام أساليب الاسترخاء والتفكير الإيجابي والقدرة على إحكام الذات عند تناول الاختبار والتعبير عن المشاعر والانفعالات من خلال الحديث للزملاء عن مشاعر القلق تجاه الاختبار لكي يتم التخلص منها.

٥- استراتيجية ما بعد الاختبار:

وهنا يحاول الفرد الاستفادة من تقييم أداؤه بعد الاختبار لتحسين طرق الاستعداد للاختبارات المقبلة بناءً على نتيجة الأداء في الاختبار الحالي والتغذية المرتدة عن جودة أدائه فيه (Stenlund et al., 2016, 5).

وهنا يكون للمراقبة الذاتية كاستراتيجية ضمن استراتيجيات ما وراء المعرفة أهمية كبيرة في تحسين الأداء والاستعداد للاختبارات القادمة بصفة خاصة، حيث تسهم في تحليل الطالب لاستراتيجياته المستخدمة في أداء الاختبار وكيفية الإعداد للاختبار واستراتيجيات الاستذكار التي يعتمد عليها، وهو ما يسهم بالتالي في زيادة وعي الفرد باستراتيجيات أداء الاختبار والتخطيط لاستخدامها في الاختبارات المستقبلية وهو ما يمكنه من التغلب على العقبات التي يواجهها أثناء الأداء في الاختبار (Hong et al., 2006, 145).

وهذا التصنيف لاستراتيجيات أداء الاختبار يتفق مع تأكيد العديد من الدراسات السابقة بأن أفضل تصنيفات استراتيجيات أداء الاختبار هو الذي يتم فيه الاهتمام بالاستراتيجيات المستخدمة قبل البدء في الإجابة والاستراتيجيات المستخدمة أثناء الإجابة والاستراتيجيات المستخدمة بعد الانتهاء من الاستجابة (إياد حمادنة، ٢٠١١، ٢٩٩؛ Amer, 2007, 4-5).

ثانياً: الخصائص السيكومترية للاختبار وبنوده:

يشير جابر عبدالله (٢٠١٤، ١٢٩)؛ هدى العريفي (٢٠٠٨، ٥)؛ إبراهيم يعقوب وباسل أبوفودة، (٢٠١٢، ٤٢٧)؛ عصام الجبة (٢٠١٢، ١٥١)؛ مايسة أبومسلم (٢٠١٦، ١٦٠) أن أهم الخصائص السيكومترية للاختبار تتمثل في معاملات الصعوبة ومعاملات التمييز لمفرداته وثبات وصدق الاختبار.

- **صعوبة المفردة** Item Difficulty:

يشير صلاح علام (٢٠١٢، ١١٣) إلى أن صعوبة المفردة يرتبط بعدد الإجابات الصحيحة على المفردة وأنه كلما زاد هذا العدد زادت سهولة المفردة، ونسبة الطلاب الذين أجابوا إجابة صحيحة على المفردة إلى العدد الكلي للطلاب هو ما يعبر عن معامل أو مؤشر الصعوبة Difficulty Index، وبالتالي الزيادة في معامل الصعوبة يعبر عن زيادة سهولة المفردة.

- **تمييز المفردة** Item Discrimination:

يعبر تمييز المفردة عن قدرة المفردة على التمييز بين الطلاب ذوي المستويات المختلفة في السمة، وتتراوح معاملات تمييز المفردات ما بين (−١، +١) ويعتبر تمييز المفردة منخفضاً إذا قل عن ٠.٢ وعندما يساوي معامل التمييز صفراً دل ذلك على انعدام قدرة المفرد على التمييز وعندما يساوي معامل التمييز الواحد الصحيح تكون المفردة مميزة تماماً (صلاح علام، ٢٠١٢، ١١٣)؛ ومن الناحية العملية يصعب جداً الوصول إلى معاملات تمييز تامة، وإذا كانت معاملات التمييز بإشارة سالبة فهذا يعنى أن المفردة تمييز بطريقة عكسية وهو ما يعبر عن خلل فيها ويجب حذفها من الاختبار (جابر عبدالله، ٢٠١٤، ١٤٢؛ Orth, 1995, 73).

ويتم استخدام معامل الارتباط الثنائي الأصيل في حساب معامل تمييز المفردة، وذلك بحساب معامل الارتباط بين درجات الطلاب مرتفعي ومنخفضي الأداء في الاختبار على المفردة ودرجاتهم الكلية على الاختبار، وهو ما يشير إلى ما إذا كان ذوي القدرة الأعلى قد أجابوا إجابة صحيحة على المفردة أكثر من ذوي القدرة الأدنى (Liao, Hatrak & Yu, 2010, 6)، ويمكن اعتباره كذلك بمثابة الفرق بين معاملي صعوبة المفردة لدى مجموعتي طلاب الإرباعي الأعلى والأدنى في الدرجات الكلية على الاختبار (عصام الجبة، ٢٠١٢، ١٥١، ١٥٦).

- **صدق الاختبار**: Test Validity

يؤكد صلاح علام (٢٠٠٦، ٣٩) أن الصدق يعبر عن مدى استفادتنا من درجات الطلاب في اختبار أو مقياس نفسي أو تربوي في اتخاذ قرارات مناسبة تتعلق بموقف ما بقدر كبير من الثقة؛ ويعبر الصدق عن قياس الاختبار لما وضع لقياسه بالفعل، وبصفة عامة يصنف الصدق إلى صدق المحتوى والصدق المرتبط بالمحكات وصدق البناء (Domino & Domino, 2006, 52)؛ ويذكر Colliver, Conlee and Verhulst (2012, 367) أن هذا المفهوم للصدق المتعلق بأن الاختبار الصادق هو الذي يقيس ما وضع لقياسه بالفعل يعتبر المفهوم التقليدي للصدق والذي يؤكد على أن الصدق خاصية للاختبار ومدى دقته في قياس البنية المستهدفة، وفي الوقت الحالي وفي ضوء التوجيهات الصادرة من الجمعية الأمريكية لعلم النفس APA حدثت تحولات في المفهوم ليعبر عن صدق التنبؤات المبنية على الدرجات وتفسيراتها ومدى الاستفادة منها من الناحية التطبيقية، ولذلك أصبح المفهوم الأكثر مناسبة في الوقت الحالي صدق درجات الاختبار وليس صدق الاختبار.

وحالياً تشير معايير الاختبارات الصادرة عن الهيئات العالمية المرتبطة بالقياس التربوي مثل الجمعية الأمريكية لعلم النفس والجمعية الأمريكية للبحوث التربوية والمجلس الوطني

الأمريكي للقياس التربوي إلى أن هناك خمسة مصادر للبرهنة على الصدق تتمثل في الأدلة المعتمدة على محتوى الاختبار والأدلة المعتمدة على الاتساق الداخلي والأدلة المعتمدة على العلاقة بمتغيرات أخرى والأدلة المعتمدة على عمليات الاستجابة والأدلة المعتمدة على توابع عملية الاختبار والأخيرة تشير إلى التأكد من الصدق في ضوء الاستفادة من نتائج عملية القياس وما يترتب عليها من قرارات مرتبطة بقدر كبير من الثقة (Johnson & Arnold, 2004, 267; Cizek, Bowen & Church, 2010, 2).

وفي البحث الحالي تم الاعتماد على الصدق التلازمي والذي يعبر عن الارتباط بين درجات الطلاب في الاختبار ودرجاتهم في محك ملازم له والذي يتمثل في درجاتهم في الاختبار النهائي لنفس المقرر، وعلى الصدق التنبؤي والمتمثل في علاقة درجات الطلاب في الاختبار بدرجاتهم في باقي المقررات كما يعبر عنها بمعدلاتهم الأكاديمية في دراسة تأثير استراتيجيات أداء الاختبار على صدق الاختبار، ويمكن اعتبار الصدق التلازمي والصدق التنبؤي ضمن الأدلة المعتمدة على علاقة درجات الاختبار بمتغيرات أخرى وكذلك ضمن الأدلة المعتمدة على توابع الاختبار في تقويم صدق الاختبار.

- **ثبات الاختبار**: Test Reliability

يشير صلاح علام (٢٠٠٦، ٨٨، ٨٩) إلى أن ثبات الاختبار يعبر عن اتساق درجات الاختبارات والمقاييس لمجموعة معينة من الأفراد، ويكون باتساق الدرجات عبر الزمن أو اتساق صيغ مختلفة من نفس الاختبار أو اتساق بين مفردات الاختبار أو الاتساق عبر أفراد مختبرين ومصححين مختلفين، وهي خاصية للدرجات أو النتائج وليست للاختبار نفسه؛ وفي ضوء فكرة أن ثبات الاختبار هو عبارة عن نسبة التباين الحقيقي للتباين الكلي للدرجات (Blerkom, 2009, 204) فإن ثبات الاختبار يزداد كلما اقتربت الدرجات الملاحظة للأفراد من درجاتهم الحقيقية في السمة.

ويشير Blerkom, (2009, 48) إلى أن ثبات الاتساق الداخلي يعد من أفضل طرق التأكد من ثبات الاختبارات مقارنة بثبات إعادة التطبيق أو الصور المتكافئة، ويعبر الثبات هنا

عن الاتساق بين البنود في قياس نفس البنية المستهدف قياسها، ويمكن التأكد من الثبات في هذه الحالة باستخدام التجزئة النصفية كبديل للصور المتكافئة أو باستخدام أحد معادلات الاتساق الداخلي كمعادلة كيودر وريتشاردسون KR-20 في حالة البنود ثنائية الإجابة أو الصورة العامة لهذه المعادلة باستخدام معامل ثبات ألفا كرونباخ، وفي البحث الحالي تم الاعتماد على معادلة كيودر وريتشاردسون KR-20 في دراسة تأثير استراتيجيات أداء الاختبار على ثباته حيث أن جميع أسئلة الاختبار التحصيلي المستخدم في البحث الحالي ثنائية الاستجابة.

- **الخطأ المعياري للقياس:** Standard Error of Measurement (SEM)

وهو من المؤشرات العامة لجودة الاختبار ويعبر عن الانحراف المعياري للدرجات الممثلة لدرجة الخطأ في الدرجات الملاحظة للأفراد وكلما انخفضت قيمة الخطأ المعياري كلما كانت الدرجات الملاحظة تعبر عن الدرجة الحقيقية بدرجة أكبر من الدقة؛ ويذكر Blerkom (2009, 204) أن هذا الخطأ ممكن يستدل عليه بمعلومية الانحراف المعياري للدرجات الملاحظة وثبات الاختبار من خلال العلاقة التالية:

معادلة (١)

$$SEM = SD \times \sqrt{1 - Reliability} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

وبالتالي كلما قل الانحراف المعياري للدرجات وزاد ثبات الاختبار كلما قل الخطأ المعياري للقياس، فالخطأ المعياري للقياس يعبر عن مقدار الاختلاف الممكن في درجات الاختبار والراجع لتأثير المتغيرات العشوائية، فهو متوسط الفروق بين درجات الأفراد في الاختبار ومتوسط درجاتهم عند الأداء على الاختبار عدة مرات دون تحسن فعلي في القدرة (Liao et al., 2010, 4).

منهج البحث وإجراءاته:

منهج البحث:

استخدم في الإجابة عن أسئلة البحث الحالي والتأكد من مدى صحة فروضه المنهج الوصفي بالاعتماد على العديد من الأساليب الإحصائية التي تم توضيحها عند الإجابة عن كل سؤال ومنها (المتوسطات والانحرافات المعيارية، واختبار "ت" لعينة واحدة، ومعاملات الارتباط وتحليل الانحدار، وتحليل التباين أحادي الاتجاه، وتحليل التباين للقياسات المتكررة).

عينة البحث:

تكونت العينة الأساسية في البحث الحالي من ١٨١ طالب من الطلاب الذكور بالدبلوم العام في التربية بالمقر الرئيسي في جامعة القصيم بالمملكة العربية السعودية في الفصل الدراسي الأول من العام الجامعي ١٤٣٧/ ١٤٣٨هـ، ٢٠١٦/٢٠١٧م، تتراوح أعمارهم من ٢٢ إلى ٢٦ سنة بمتوسط عمر قدره ٢٤.٢ سنة وانحراف معياري ١.٩ سنة، من تخصصات مختلفة في مرحلة البكالوريوس.

أدوات البحث:

١- مقياس استراتيجيات أداء الاختبار:

تمت الاستفادة من الإطار النظري للبحث الحالي وكذلك المقاييس التي تم إعدادها في بعض الدراسات السابقة في تحديد أبعاد المقياس الحالي وخاصة دراسة (عفاف وادي، ٢٠١٣) والتي افترضت أن مهارات الحكمة الاختبارية تتضمن عدد من المهارات منها مهارة إدارة وقت الاختبار ومهارة تجنب الخطأ ومهارة التعامل مع ورقة الأسئلة ومهارة التعامل مع ورقة الإجابة ومهارة المراجعة؛ والمقياس الذي تم إعداده في دراسة Bicak (2013) والذي تضمن استراتيجيات أداء الاختبار المتمثلة في استراتيجية تحليل البنود واستراتيجية تخطيط الوقت، واستراتيجية توقع الاختيار، واستراتيجية ما بعد الاختبار، ومن المقياس الذي تم إعداده في دراسة Dodeen (2008) والذي اعتبر أن استراتيجيات أداء الاختبار تتمثل في مهارات ما قبل الاختبار، وإدارة الوقت، وأثناء الاختبار، وبعد الاختبار؛ وفي ضوء ذلك تم إعداد المقياس الحالي والذي تضمن في صورته الأولية ٤٩ عبارة موزعة على ٥ أبعاد فرعية.

وتم تعديل صياغة عبارات المقياس أكثر من مرة حتى وصلت لشكلها النهائي بعد عرض المقياس في صورته الأولية على مجموعة من الزملاء بقسم علم النفس بكلية التربية جامعة القصيم، وتتم الاستجابة لعبارات المقياس الحالي في ضوء تدريج خماسي الاستجابة يبدأ بلا

تنطبق علي أبداً، وينتهي بتنطبق علي دائماً لتقابل الدرجات (١، ٢، ٣، ٤، ٥) على الترتيب مع مراعات اتجاه العبارة؛ والجدول التالي يوضح توزيع عبارات المقياس على استراتيجيات أداء الاختبار موضوع البحث الحالي:

جدول (١):

عبارات مقياس استراتيجيات أداء الاختبار وتوزيعها على الأبعاد المختلفة للمقياس

استراتيجيات أداء الاختبار	العبارات			أقل درجة	أعلى درجة
	العدد	الموجبة	السلبية		
تخطيط وتنظيم وقت الاختبار	١٠	١-١٠	×××	١٠	٥٠
تحليل أسئلة الاختبار	١٣	١١، ١٢، ١٤-٢٣	١٣	١٣	٦٥
تخطيط وتنظيم الإجابة	١٣	٢٤-٣٦	×××	١٣	٦٥
الضبط والسيطرة	٨	٣٧-٤١	٤٢، ٤٣، ٤٤	٨	٤٠
ما بعد الاختبار	٥	٤٥-٤٩	×××	٥	٢٥
المجموع	٤٩	٤٥	٤	٤٩	٢٤٥

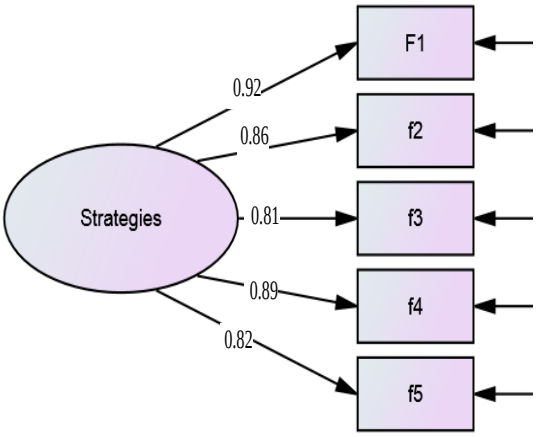
وبعد التأكد من صدق المحتوي عن طريق عرض المقياس على مجموعة من المحكمين تم تطبيقه على ٣٥ طالباً من طلاب البكالوريوس بقسمي التربية البدنية والتربية الخاصة، وتم التأكد من الصدق البنائي باستخدام التحليل العاملي التوكيدي بالنسبة للأبعاد للتأكد من تشبعها على عامل كامن واحد هو استراتيجيات أداء الاختبار، فبعد أن تم تحديد النموذج البنائي للمقياس تم استخدام برنامج Amos v.21.0 في تقدير معالم النموذج ومؤشرات جودة المطابقة له باستخدام طريقة الاحتمال الأقصى فكانت كما هي موضحة بجدول (٢):

جدول (٢):

مؤشرات جودة المطابقة للنموذج البنائي لمقياس استراتيجيات أداء الاختبار

المؤشر			الحد الأقصى
χ^2	مربع كاي	٦.١٨١	عدم دلالة مربع كاي أو النسبة بين مربع كاي ودرجات الحرية أقل من ٢
Df	درجات الحرية	٥	
χ^2/df	النسبة بين مربع كاي ودرجات الحرية	١.٢٣٦	
GFI	مؤشر جودة المطابقة	٠.٩٣٩	تزداد جودة المطابقة كلما اقترب المؤشر من الواحد الصحيح
AGFI	مؤشر جودة المطابقة المعدل	٠.٩٨٤	
CFI	مؤشر المطابقة المقارن	٠.٩٩١	
IFI	مؤشر المطابقة المتزايد	٠.٩٩٢	
NFI	مؤشر المطابقة المعياري	٠.٩٥٨	
RMSEA	جذر متوسط مربعات خطأ التقريب	٠.٠٢٨	تزداد جودة المطابقة باقتراب هذا المؤشر من الصفر

يتضح من جدول (٢) أن النموذج المفترض للمقياس يطابق بدرجة كبيرة بيانات العينة وهو ما يؤكد تشبع أبعاد المقياس على متغير كامن واحد هو استراتيجيات أداء الاختبار حيث كانت مربع كاي غير دالة إحصائياً وكانت كذلك النسبة بين مربع كاي ودرجات الحرية أقل من (٢) وكانت قيم (مؤشر حسن المطابقة GFI ومؤشر حسن المطابقة المعدل AGFI ومؤشر المطابقة المقارن CFI ومؤشر المطابقة المعياري NFI ومؤشر المطابقة المتزايد IFI) جميعها قيم مرتفعة وقريبة من الحد الأقصى لهذه المؤشرات (واحد صحيح)، وكذلك مؤشر جذر متوسط مربعات خطأ التقريب RMSEA كان أقل من ٠.٠٥ وهو ما يؤكد الصدق البنائي للمقياس؛ والشكل التالي يوضح النموذج البنائي المفترض للمقياس وتشبعات العوامل:



شكل (٢): النموذج البنائي لمقياس استراتيجيات أداء الاختبار

يتضح من الشكل السابق أن تشبع استراتيجية تخطيط وتنظيم وقت الاختبار على المتغير الكامن كان مساوياً ٠.٩٢، وتشبع استراتيجية تحليل أسئلة الاختبار مساوياً ٠.٨٦، وتشبع استراتيجية تخطيط وتنظيم الإجابة مساوياً ٠.٨١، وتشبع استراتيجية الضبط والسيطرة مساوياً ٠.٩٦، وأخيراً كان تشبع استراتيجية ما بعد الاختبار مساوياً ٠.٨٢، وهو ما يؤكد أن أبعاد المقياس لها تشبعات مقبولة على المتغير الكامن.

وتم التأكد من ثبات درجات المقياس وأبعاده الفرعية باستخدام معاملات ثبات ألفا كرونباخ فكانت كما هو موضح بجدول (٣):

جدول (٣):

معاملات ثبات مقياس استراتيجيات أداء الاختبار وأبعاده الفرعية

تخطيط وتنظيم وقت الاختبار	تحليل أسئلة الاختبار	تخطيط وتنظيم الاجابة	الضبط والسيطرة	ما بعد الاختبار	المقياس ككل
٠.٨٧٣	٠.٨٢٠	٠.٧٨٤	٠.٧٤٩	٠.٧٩١	٠.٨٥٥

ومن جدول (٣) يتضح أن للمقياس وأبعاده الفرعية معاملات ثبات مقبولة، ومما سبق يتأكد تمتع المقياس بمؤشرات سيكومترية (صدق، ثبات) جيدة، وهو ما يؤكد صلاحيته للاستخدام في البحث الحالي.

٢- اختبار تحصيلي في مقرر القياس والتقويم:

خطوات إعداد الاختبار التحصيلي:

- تم تحديد الموضوعات التي سيتم تضمينها في الاختبار وهي الموضوعات المحددة في توصيف مقرر القياس والتقويم الذي يقوم الباحث بتدريسه لطلاب الدبلوم العام في التربية في الفصل الدراسي الأول من العام الجامعي ٢٠١٦/ ٢٠١٧م، ١٤٣٧/١٤٣٨هـ، والتي تم تصنيفها في خمسة مجالات رئيسية هي (مفاهيم القياس والتقييم والتقويم وخصائص كل منها والفرق بينها، ومستويات القياس وأنواع التقويم وخصائص التقويم الجيد، الأسئلة الموضوعية والمقالية وأنواعها وخصائصها وشروطها، الاختبارات النفسية وتصنيفاتها وشروطها، بعض الأساليب الإحصائية).
- تم تحديد الأهمية النسبية لكل مجال فرعي في ضوء مساحته وأهميته في المقرر.
- تم تحديد الأهداف السلوكية التي سيتم قياسها كنواتج مستهدف تحقيقها من خلال دراسة المقرر والتي تمثلت في المستويات المعرفية الثلاثة الأولى من تصنيف بلوم وهي التذكر والفهم والتطبيق، وتم تحديد الأهمية النسبية لكل هدف حسب الأهداف التي يسعى المقرر لتحقيقها.
- تم إعداد جدول مواصفات الاختبار التحصيلي على أنه يتكون من ٤٠ سؤال منها أسئلة مقالية قصيرة وأسئلة موضوعية (اختيار من متعدد، صواب وخطأ) كما هو موضح في جدول (٤):

جدول (٤):

جدول مواصفات الاختبار التحصيلي في مقرر مبادئ القياس والتقويم

الهدف / المجال	التذكر ٣٠%	الفهم ٥٠%	التطبيق ٢٠%	المجموع
المفاهيم الأساسية ١٥%	٠.١٥×٠.٣٠×٤٠ = ١.٨ (سؤالين)	٠.١٥×٠.٥٠×٤٠ = ٣ (٣ أسئلة)	٠.١٥×٠.٢٠×٤٠ =١.٢ (سؤال واحد)	٦
مستويات القياس وأنواع التقويم ١٥%	٠.١٥×٠.٣٠×٤٠ = ١.٨ (سؤالين)	٠.١٥×٠.٥٠×٤٠ = ٣ (٣ أسئلة)	٠.١٥×٠.٢٠×٤٠ =١.٢ (سؤال واحد)	٦
الأسئلة الموضوعية والمقالية ٢٥%	٠.٢٥×٠.٣٠×٤٠ = ٣ (٣ أسئلة)	٠.٢٥×٠.٥٠×٤٠ = ٥ (٥ أسئلة)	٠.٢٥×٠.٢٠×٤٠ = ٢ (سؤالين)	١٠
الاختبارات وتصنيفاتها وشروطها ٢٠%	٠.٢٠×٠.٣٠×٤٠ = ٢.٤ (سؤالين)	٠.٢٠×٠.٥٠×٤٠ = ٤ (٤ أسئلة)	٠.٢٠×٠.٢٠×٤٠ =١.٦ (سؤالين)	٨
المبادئ الإحصائية ٢٥%	٠.٢٥×٠.٣٠×٤٠ = ٣ (٣ أسئلة)	٠.٢٥×٠.٥٠×٤٠ = ٥ (٥ أسئلة)	٠.٢٥×٠.٢٠×٤٠ = ٢ (سؤالين)	١٠
المجموع	١٢	٢٠	٨	٤٠

- بعد تكوين جدول مواصفات الاختبار وكتابة الأسئلة تم عرض الاختبار على الزملاء ممن يقومون بتدريس المقرر ولديهم خلفية بالقياس والتقويم وكانت آرائهم متفقة بشأن انتماء كل سؤال للمجال والهدف الذي حدده الباحث مع بعض التعديلات في صياغة بعض الأسئلة، ومما سبق يكون قد تم التحقق من صدق المحتوى للاختبار التحصيلي في ضوء آراء المحكمين وجدول مواصفات الاختبار.

- كذلك تم تطبيق الاختبار على مجموعة من طلاب البكالوريوس بقسم التربية البدنية وقسم التربية الخاصة بلغ عددهم ٣٥ طالباً وتم حساب معامل ثبات كيودر وريتشاردسون للاختبار فكان معامل الثبات مساوياً ٠.٧٤٦ وكذلك كان معامل ثبات التجزئة النصفية مساوياً ٠.٧٣٨ وبذلك يكون تم التأكد من ثبات الاختبار التحصيلي، تم كذلك التأكد من الاتساق الداخلي لبنود الاختبار بحساب معاملات الارتباط بين درجات بنود الاختبار والدرجة الكلية للاختبار بعد حذف درجة البند منها، فكانت معاملات الارتباط كما هو موضح بجدول (٥):

جدول (٥):

معاملات الارتباط بين درجات بنود الاختبار التحصيلي والدرجة الكلية بعد حذف درجة البند منها

م	الارتباط	م	الارتباط	م	الارتباط	م	الارتباط	م	الارتباط
١	٠.٤٤٦**	٩	٠.٤٥٩**	١٧	٠.٤٢١**	٢٥	٠.٣٧٤*	٣٣	٠.٦٢٥**
٢	٠.٤٣٨**	١٠	٠.٤٨٩**	١٨	٠.٥٩٦**	٢٦	٠.٤٨٢**	٣٤	٠.٦٣٢**
٣	٠.٤٧٤**	١١	٠.٦٤٤**	١٩	٠.٥٣٦**	٢٧	٠.٥٩٦**	٣٥	٠.٧٨٢**
٤	٠.٥٢٧**	١٢	٠.٤٣٦**	٢٠	٠.٣٣٨*	٢٨	٠.٧٧٠**	٣٦	٠.٥٤٥**
٥	٠.٥٢٠**	١٣	٠.٧٢٤**	٢١	٠.٤٧١**	٢٩	٠.٤٢٤**	٣٧	٠.٧١٨**
٦	٠.٤٣٨**	١٤	٠.٣٨٢*	٢٢	٠.٧١٢**	٣٠	٠.٥٤٤**	٣٨	٠.٥٣٩**
٧	٠.٤٣٧**	١٥	٠.٣٦٦*	٢٣	٠.٥٩١**	٣١	٠.٦٠٩**	٣٩	٠.٤٨٢**
٨	٠.٤٦٦**	١٦	٠.٤٦٨**	٢٤	٠.٤٩٦**	٣٢	٠.٦٦٣**	٤٠	٠.٥٣٩**

القيمة الجدولية لمعامل الارتباط لحجم عينة ٣٥ ومستوي دلالة ٠.٠٥ و ٠.٠١ تساوي ٠.٣٢٤٦ و ٠.٤١٨٢ على الترتيب

تم كذلك التأكد من الاتساق الداخلي للمجالات الفرعية للاختبار بحساب معاملات الارتباط بين درجات تلك المجالات وكذلك معاملات الارتباط بين درجات تلك المجالات والدرجة الكلية للاختبار بعد حذف درجة المجال منها، فكانت معاملات الارتباط كما هو موضح بجدول (٦):

جدول (٦):

معاملات الارتباط بين درجات المجالات الفرعية للاختبار التحصيلي وبينها وبين الدرجة الكلية بعد حذف درجة المجال

الاختبار التحصيلي	المفاهيم الأساسية	مستويات القياس وأنواع التقويم	الأسئلة الموضوعية والمقالية	الاختبارات وتصنيفاتها وشروطها	المبادئ الإحصائية
مستويات القياس وأنواع التقويم	٠.٧١٢**				
الأسئلة الموضوعية والمقالية	٠.٧٢٨**	٠.٥٠٩**			
الاختبارات وتصنيفاتها وشروطها	٠.٧٤٥**	٠.٦٩٧**	٠.٥٨٦**		
المبادئ الإحصائية	٠.٦٥٢**	٠.٤٦٤**	٠.٦٢٢**	٠.٧٤٣**	
الدرجة الكلية للاختبار	٠.٨٣٠**	٠.٨٥٩**	٠.٧٤٨**	٠.٧٦٠**	٠.٧٩٩**

القيمة الجدولية لمعامل الارتباط لحجم عينة ٣٥ وعند مستوي دلالة ٠.٠١ تساوي ٠.٤١٨٢

ومما سبق يتأكد تماسك بنود الاختبار وكذلك الاتساق والتماسك بين مجالاته الفرعية، والاختبار في صورته النهائية يتكون من ٤٠ سؤالاً والإجابة الخطأ تقابل الدرجة صفر بينما الإجابة الصحيحة تقابل الدرجة (١) وبالتالي تكون أقصى درجة يمكن الحصول عليها في الاختبار هي الدرجة ٤٠، وفي ضوء الخطوات السابق تتأكد الكفاءة السيكومترية للاختبار وإمكانية استخدامه في البحث الحالي.

نتائج البحث:

أولاً: نتائج الإجابة عن السؤال الأول:

ينص السؤال الأول للبحث الحالي على "ما استراتيجيات أداء الاختبار الأكثر استخداماً لدى الطلاب عينة البحث الحالي؟"

وللإجابة عن هذا السؤال تم استخدام اختبار "ت" لمقارنة متوسط درجات عينة مع متوسط فرضي وذلك للتعرف على درجة استخدام كل استراتيجية من استراتيجيات أداء الاختبار وذلك بافتراض أن المتوسط الفرضي لكل استراتيجية يساوي (٣× عدد العبارات) فكانت النتائج كما هي موضحة بجدول (٧):

جدول (٧):

دلالة الفروق بين المتوسط الفرضي والتجريبي لاستراتيجيات أداء الاختبار

استراتيجيات أداء الاختبار	المتوسط الفرضي	المتوسط التجريبي	الانحراف المعياري	قيمة "ت" ودلالتها	درجة الاستخدام	مستوى الاستخدام
تخطيط وتنظيم وقت الاجابة	٣٠.٠٠	٣١.٧٨٥	٧.٧٥٣	٣.٠٩٧**	%٦٣.٥٧٠	متوسط
تحليل أسئلة الاختبار	٣٩.٠٠	٣٨.٧٣٥	١٢.٢٩٣	٠.٢٩٠	%٥٩.٥٩٢	متوسط
تخطيط وتنظيم الاجابة	٣٩.٠٠	٤٠.٢٥٩	١٠.٧٣٠	١.٥٧٩	%٦١.٩٣٧	متوسط
الضبط والسيطرة	٢٤.٠٠	٢٢.٥٥٣	٨.١٣٨	٢.٣٩٣*	%٥٦.٣٨٣	ضعيف
ما بعد الاختبار	١٥.٠٠	١٦.٠٢٢	٤.٣٨٩	٣.١٣٣**	%٦٤.٠٨٨	متوسط
الدرجة الكلية	١٤٧.٠٠	١٤٩.٣٥٤	٣٩.٦٣٥	٠.٧٩٩	%٦٠.٩٦١	متوسط

يتضح من جدول (٧) السابق أنه:

- استراتيجيات أداء الاختبار ككل جاءت متحققة بدرجة متوسطة حيث لا توجد فروق دالة إحصائياً بين المتوسط التجريبي والمتوسط الفرضي للدرجات الكلية، أما بشأن الاستراتيجيات الفرعية لأداء الاختبار فجاءت استراتيجية تخطيط وتنظيم وقت الاجابة أعلى الاستراتيجيات من حيث درجة الاستخدام يليها استراتيجية ما بعد الاختبار، ثم استراتيجية تخطيط وتنظيم الإجابة فاستراتيجية تحليل أسئلة الاختبار، أما في حالة استراتيجية الضبط والسيطرة فجاءت الفروق دالة إحصائياً عند مستوى ٠.٠٥ بين المتوسط التجريبي والمتوسط الفرضي لصالح المتوسط الفرضي وبالتالي تكون هذه الاستراتيجية هي أقل الاستراتيجيات من حيث درجة الاستخدام.

وعدم وجود مستويات مرتفعة من استراتيجيات أداء الاختبار لدى الطلاب عينة البحث الحالي على الرغم من أنهم من طلاب الدبلوم العام أي مروا بالعديد من المواقف الاختبارية حتى الانتهاء من مرحلة البكالوريوس، يفسر في ضوء عدم تلقي هؤلاء الطلاب لتدريبات متخصصة على هذه الاستراتيجيات فحتى يتقن الطلاب هذه الاستراتيجيات وتصبح جزء من سلوكهم الاعتيادي في الموقف الاختباري يجب أن يتم تدريب الطالب وتدريسه هذه الاستراتيجيات بشكل مباشر.

وهذا التفسير يتماشى مع ما يؤكد عليه Xu and Wu (2012) من أهمية التدريب على هذه الاستراتيجيات في مواقف تشابه الموقف الاختباري الفعلي؛ فلكي يصبح الطالب أكثر ألفة بهذه الاستراتيجيات وبكيفية تطبيقها بنجاح يجب أن تتاح للطالب الفرصة للتدريب الكافي عليها (Pour-Mohammadi & Abidin, 2011)؛ بل ويدعوا البعض إلى ضرورة أن تصبح هذه الاستراتيجيات جزء من المنهج الدراسي فهي مهارات يحتاج إتقانها إلى تدريب وتعليم واهتمام من قبل المعلمين في جميع المراحل الدراسية ومثل هذه الجهود منخفضة جداً لحد غير متوقع، وأكدت نتائج دراسة (Takallou et al., 2015) على أن اتجاهات طلاب الجامعة نحو تعلم هذه الاستراتيجيات كان مرتفع بل وكانت هناك مطالبة من أفراد عينة الدراسة بضرورة أن تكون هناك تدريبات كافية علي هذه الاستراتيجيات والتي تساعد الطالب في معرفة ما الذي عليه فعلية أثناء وبعد الاختبار؛ وبالتالي يمكن تفسير انخفاض مستوى هذه الاستراتيجيات لدى الطلاب عينة البحث الحالي في ضوء عدم الاهتمام بتعليمها وتدريب الطلاب عليها.

كذلك يمكن تفسير عدم ارتفاع مستوى استراتيجيات أداء الاختبار لدى الطلاب عينة البحث الحالي في ضوء انخفاض دافعية أداء الاختبار لديهم هذا من جهة ومن جهة أخرى انخفاض دافعية الإنجاز والرغبة في الاتقان بصفة عامة والمرتبط بعدم وجود فرص للتوظيف والتحاق الطالب بالدبلوم رغبة في الحصول على الشهادة بغض النظر على التعلم والاتقان، ففي هذه الحالة لا يحاول الطالب بذل الجهد الكافي للوصول للدرجة المرتفعة، وتساعد نظم التقويم المتبعة وتركيز معظم الاختبارات على الأسئلة الموضوعية في حرمان الطالب من تعلم وممارسة هذه الاستراتيجيات بشكل فعال.

وتتفق النتائج الحالية مع نتائج دراسة Hong et al. (2006) والتي توصلت إلى أن استراتيجية تنظيم وقت الإجابة من أعلى الاستراتيجيات استخداماً في الموقف الاختباري؛ وتتفق كذلك مع نتائج دراسة Zhang et al. (2011) التي أكدت أن مستوى استراتيجيات أداء الاختبار لدى طلاب الجامعة الصنين متوسط؛ ومع نتائج دراسة Dodeen et al. (2014) والتي أكدت نتائجها على تدني مستوى مهارات أداء الاختبار لدى طلاب الجامعة ماعدا استراتيجية قبل الاختبار كانت في المدى المتوسط؛ ومع نتائج دراسة فاطمة مطلك (٢٠٠٩) والتي أكدت على أن مستوى الدرجة الكلية للحكمة الاختبارية لدى الطلاب ذوي المستوى العادي في التحصيل مستوى متوسط، وهي نفس النتيجة التي توصل إليها زين ردادي (٢٠٠١)؛ وتتفق النتائج الحالية مع نتائج دراسة سعود العنزي (٢٠١٤) والتي أكدت أن مهارات خطة الحل والتي تقترب في المعني من استراتيجية تخطيط وتنظيم الإجابة في البحث الحالي مستواها متوسط بينما تختلف معها في أن مهارات إدارة الوقت والتعامل مع الأسئلة كانت ضعيفة بينما في البحث الحالي كانت هذه الاستراتيجيات متوسطة.

ومع نتائج دراسة أحمد عودة (١٩٨٩) والتي أكدت على أن استراتيجيات حكمة الاختبار يجب أن يتم تعليمها بعكس ما هو مفترض أن هذه الاستراتيجيات يتم اكتسابها تلقائياً من خلال خبرة المرور بالاختبارات والألفة بالاختبارات؛ بينما تختلف مع نتائج دراسة Dodeen (2009) والتي أكدت نتائجها أن مستوى استراتيجيات أداء الاختبار لدى طلاب جامعة الإمارات العربية المتحدة مستوى مرتفع؛ ومع نتائج دراسة إياد حمادنة (٢٠١١) والتي أكدت على أن مهارات حكمة الاختبار تستخدم بدرجة عالية لدى طلاب الجامعة الأردنيين، ولكن على الرغم من أن استراتيجية ما بعد الاختبار في البحث الحالي جاءت بدرجة متوسطة إلا أنها تحتل مرتبة أعلى من حيث درجة الاستخدام مقارنة بباقي الاستراتيجيات كما في دراسة "إياد حمادنة".

وتختلف مع نتائج دراسة Bicak (2013) والتي أشارت إلى أن استراتيجيات أداء الاختبار المتمثلة في استراتيجية تحليل البنود واستراتيجية تخطيط وتنظيم الوقت، واستراتيجية توقع الاختيار، واستراتيجية ما بعد الاختبار، كانت بدرجة ضعيفة.

ثانياً: نتائج الإجابة عن السؤال الثاني:

ينص السؤال الثاني للبحث الحالي على "هل توجد فروق دالة إحصائياً في درجات الطلاب عينة البحث الحالي في الاختبار التحصيلي راجعة لاختلاف درجة استخدامهم لاستراتيجيات الأداء في الاختبار؟"

للإجابة عن هذا السؤال تم استخدام تحليل التباين أحادي الاتجاه One Way ANOVA في الكشف عن دلالة الفروق بين درجات الطلاب في الاختبار التحصيلي، وذلك في المقارنة بين درجات الطلاب ككل، ودرجات الطلاب بعد استبعاد أدني ١٠% في استراتيجيات أداء الاختبار، ودرجات الطلاب بعد استبعاد أدني ٢٠% في استراتيجيات أداء الاختبار، ودرجات الطلاب بعد استبعاد أدني ٣٠% في استراتيجيات أداء الاختبار، فكانت النتائج كما هو موضح في التالي، مع ملاحظة أنه تم تفسير نتائج هذا السؤال ضمن تفسير نتائج السؤال الثالث للارتباط بين نتائج السؤالين منعا للتكرار:

جدول (٨):

الإحصاءات الوصفية لدرجات الطلاب في الاختبار التحصيلي في ضوء درجة استخدام استراتيجيات أداء الاختبار

المجموعات	العدد	المتوسط	الانحراف المعياري
الأفراد ككل	١٨١	٢٥.١٠٥	٩.٢٤٩
بعد استبعاد أدنى ١٠ % في استراتيجيات أداء الاختبار	١٦١	٢٥.٧٨٣	٩.٣١٠
بعد استبعاد أدنى ٢٠ % في استراتيجيات أداء الاختبار	١٤١	٢٧.١١٣	٨.٥١٥
بعد استبعاد أدنى ٣٠ % في استراتيجيات أداء الاختبار	١٢٢	٢٨.٩٩٢	٧.٥٥٥

جدول (٩):

دلالة الفروق في درجات الطلاب في الاختبار التحصيلي في ضوء درجة استخدام استراتيجيات أداء الاختبار

مصدر التباين	مجموع المربعات	درجات الحرية	التباين	النسبة الفائية
بين المجموعات	١٢٤٤.٨٤٠	٣	٤١٤.٩٤٧	٥.٣٨٤**
داخل المجموعات	٤٦٣٢٣.٥٧٣	٦٠١	٧٧.٠٧٧	
الكلي	٤٧٥٦٨.٤١٣	٦٠٤		

يتضح من جدول (٩) وجود فروق دالة إحصائياً عند مستوى ٠.٠١ في درجات الطلاب في الاختبار التحصيلي راجعة لدرجة استخدام استراتيجيات أداء الاختبار، وتم استخدام اختبار أقل فرق دال LSD في المقارنات البعدية بين المجموعات في درجات الاختبار التحصيلي في ضوء درجة استخدام استراتيجيات أداء الاختبار فكانت كما هي موضحة في جدول (١٠):

جدول (١٠):

المقارنات البعدية بين المجموعات في درجات الاختبار التحصيلي في ضوء درجة استخدام استراتيجيات أداء الاختبار

المجموعات	الأفراد ككل (م= ٢٥.١٠٥)	بعد استبعاد أدنى ١٠ % (م= ٢٥.٧٨٣)	بعد استبعاد أدنى ٢٠ % (م= ٢٧.١١٤)
بعد استبعاد أدنى ١٠ % (م= ٢٥.٧٨٣)	٠.٦٧٨		
بعد استبعاد أدنى ٢٠ % (م= ٢٧.١١٤)	٢.٠٠٩*	١.٣٣١	
بعد استبعاد أدنى ٣٠ % (م= ٢٨.٩٩٢)	٣.٨٨٧**	٣.٢٠٩**	١.٨٧٨

يتضح من جدول (١٠) أنه توجد فروق دالة إحصائياً عند مستوى ٠.٠٥ في درجات الطلاب في الاختبار التحصيلي بين مجموعة الطلاب ككل وبين مجموعة الطلاب بعد استبعاد أدنى ٢٠% في استراتيجيات أداء الاختبار لصالح مجموعة الطلاب بعد الاستبعاد، كذلك توجد فروق دالة إحصائياً عند مستوى ٠.٠١ بين مجموعة الطلاب ككل وبين مجموعة الطلاب بعد استبعاد أدنى ٣٠% في استراتيجيات أداء الاختبار لصالح مجموعة الطلاب بعد الاستبعاد، وتوجد فروق دالة إحصائياً عند مستوى ٠.٠١ بين مجموعة الطلاب بعد استبعاد أدنى ٣٠% ومجموعة الطلاب بعد استبعاد أدنى ١٠ % في استراتيجيات أداء الاختبار لصالح مجموعة الطلاب بعد استبعاد أدنى ٣٠ %، والنتائج هنا تؤكد في مجملها أن تدني استخدام الطلاب لاستراتيجيات أداء الاختبار ينتج عنه تدني في درجاتهم في الاختبار، وأن الأداء على الاختبار يرتبط بعلاقة كبيرة باستخدام استراتيجيات أداء الاختبار.

ثالثاً: نتائج الإجابة عن السؤال الثالث:

ينص السؤال الثالث للبحث الحالي على "هل يمكن التنبؤ بدرجات الطلاب عينة البحث الحالي في الاختبار التحصيلي من خلال درجاتهم في استراتيجيات الأداء الاختبار؟"

تمثل الهدف من هذا السؤال في التعرف على أكثر استراتيجيات أداء الاختبار تأثيراً وإسهاماً في التنبؤ بدرجات الطلاب في الاختبار، للإجابة عن هذا السؤال تم استخدام تحليل الانحدار المتعدد بطريقة التحليل المتتابع Stepwise باعتبار أن استراتيجيات الأداء في الاختبار متغيرات مستقلة ونتائج الأداء في الاختبار التحصيلي متغير تابع؛ حيث يساعد هذا الأسلوب الإحصائي في وصف التأثيرات المباشرة للمتغيرات المستقلة في المتغيرات التابعة في صورة نموذج خطي تعرف معادلته باسم معادلة الانحدار الخطي المتعدد، وفي البداية تم عرض معاملات ارتباط بيرسون التي تؤكد دلالة العلاقة بين درجات الطلاب في الاختبار ودرجاتهم في استراتيجيات أداء الاختبار كما هو موضح في جدول (١١):

جدول (١١):

دلالة العلاقة بين أداء الطلاب في الاختبار التحصيلي واستراتيجيات أداء الاختبار

استراتيجيات أداء الاختبار	تخطيط وتنظيم وقت الإجابة	تحليل أسئلة الاختبار	تخطيط وتنظيم الإجابة	الضبط والسيطرة	ما بعد الاختبار	الدرجة الكلية
معامل الارتباط	**٠.٥١٠	**٠.٥٦٥	**٠.٤٦٣	**٠.٦٠٤	**٠.٤١٢	**٠.٥٧٠

يتضح من جدول (١١) أنه توجد علاقات ارتباطية موجبة دالة إحصائياً بين درجات الطلاب في الاختبار التحصيلي واستراتيجيات أداء الاختبار، وكانت نتائج تحليل الانحدار المتعدد للتنبؤ بدرجات الطلاب في الاختبار التحصيلي من خلال استراتيجيات أدائهم للاختبار كما هي موضحة في جدول (١٢):

جدول (١٢):

تحليل الانحدار المتعدد للتنبؤ بدرجات الاختبار التحصيلي من خلال استراتيجيات أداء الاختبار

استراتيجيات أداء الاختبار	معاملات الانحدار	
ثابت الانحدار	B	٤.٢٩٣
تحليل أسئلة الاختبار	B	*٠.١٦٣
	Beta	٠.٢١٦
الضبط والسيطرة	B	**٠.٤٨٨
	Beta	٠.٤٢٨
معامل الارتباط المتعدد R	٠.٤١٧	
معامل التحديد R2	٠.١٧٤	
النسبة الفائية ودرجات الحرية لتحليل تباين الانحدار المتعدد	ف =٥٤.٦٩١**، درجات الحرية (٢، ١٧٨)	

ومن جدول (١٢) يتضح أن:

تسهم استراتيجيات الضبط والسيطرة وتحليل أسئلة الاختبار على الترتيب إيجابياً في التنبؤ بدرجات الطلاب في الاختبار، وكانت النسبة الفائية لتحليل تباين الانحدار المتعدد مساوية (٢، ١٧٨) وهي دالة إحصائيا عند مستوى دلالة ٠.٠١ مما يؤكد أهمية هذه الاستراتيجيات وإسهامها في التنبؤ بدرجات الطلاب في الاختبار، وكان معامل الارتباط المتعدد مساوياً ٠.٤١٧، وكان معامل التحديد مساوياً ٠.١٧٤ مما يعني أن نسبة التباين المفسر في درجات الطلاب والتي ترجع لتأثير هذه الاستراتيجيات يساوي ١٧.٤٪ ويمكن صياغة معادلة التنبؤ كالتالي:

درجات الاختبار التحصيلي = ٤.٢٩٣ + ٠.١٦٣× استراتيجية تحليل أسئلة الاختبار + ٠.٤٨٨× استراتيجية الضبط والسيطرة

ويمكن تفسير ارتفاع مستوى الأداء في الاختبار لدى مرتفعي استخدام استراتيجيات أداء الاختبار بأن ارتفاع مستوى هذه الاستراتيجيات يصحبه زيادة في دافعية الطلاب واهتمامهم بالأداء الجيد في الاختبار والذي ينعكس إيجابياً على تركيز الطالب وجهده المبذول في محاولة الوصول للإجابة الصحيحة ويقلل من قلق الاختبار وتمكن الطالب من الضبط الانفعالي والمعرفي أثناء الموقف الاختباري وهو ما تؤكد عليه دراسات (Maspons & Llabre, 1983; Orth, 1995; Ellis & Ryan, 2003; Lancaster et al., 2009 Takallou et al., 2015)؛ بأن التدريب على هذه الاستراتيجيات يحسن دافعية أداء الاختبار ويقلل من قلق الاختبار؛ فمن أهم مميزات أصحاب استراتيجيات أداء الاختبار تعاملهم بفاعلية مع الانفعالات السلبية والتي منها القلق في الموقف الاختباري (Hong et al., 2006).

كذلك يمكن تفسير زيادة مستوى الأداء لدى الطلاب مرتفعي استخدام استراتيجيات أداء الاختبار في ضوء أن هؤلاء الطلاب يستطيعون التعامل مع الموقف الاختباري بنوع من الفاعلية، فالموقف الاختباري يعتبر موقف ضغط وتوتر للعديد من الطلاب، وتمكن استراتيجية السيطرة والضبط وكذلك الإدارة الجيدة لوقت الاختبار من التغلب على مثل هذه الضغوط، كذلك يستفيد الطالب من أدائه في الاختبارات السابقة في الاستعداد للاختبارات المقبلة وهو ما يسهم في زيادة شعوره بالفاعلية والتمكن ويولد لدية اعتقادات إيجابية عن قدرته على الأداء الجيد في الموقف الاختباري.

فمن خصائص مرتفعي استخدام استراتيجيات أداء الاختبار والتي تميزهم عن غيرهم أنهم يتمكنون من البرهنة على الاتقان والفهم وتنظيم الإجابة بالشكل الذي يعطيهم فرصه أكبر للحصول على درجات أعلى كذلك يستطيعون استغلال وقت الاختبار بشكل أفضل ولا يقضون وقتاً في التفكير في نتائج الأداء والذي قد يسهم في تشتيت تركيزهم وإنما يركزون على كيفية الأداء بشكل أفضل وحتى في حالة الأسئلة الصعبة والغامضة فإن تحليل أسئلة الاختبار يمكنهم من الإجابة ولو بشكل جزئي، كما أن استراتيجية ما بعد الاختبار تتيح للطالب التقدم للاختبار ولديه شعور بالارتياح والثقة وتمكنه من الاستعداد الجيد للاختبار فمن يستخدم هذه الاستراتيجية يتمكن من تحليل أخطائه في الاختبارات السابقة وحاول تجنب تلك الأخطاء في الاختبارات اللاحقة وقد يسهم ذلك في زيادة توجه الطالب لعزو نجاحة أو فشلة للعوامل الداخلية القبالة للضبط والتحكم مثل الجهد أكثر من العوامل الخارجية غير القابلة للتحكم كالظروف وصعوبة المقررات؛ فاستخدام الطلاب لاستراتيجيات أداء الاختبار يزيد من وعي الطالب بأدائه Pour-Mohammadi and Abidin (2011) والذي ينعكس على تنظيم الطالب للإجابة وتجنبه للخطأ قدر الإمكان وكذلك استغلاله لوقت الاختبار الاستغلال الأمثل وهو ما يفسر ارتفاع معدل درجات الطلاب بزيادة استخدام هذه الاستراتيجيات.

وعدم ظهور تأثيرات دالة لاستراتيجية تخطيط وتنظيم وقت الاجابة وتنظيم وتخطيط الإجابة واستراتيجية ما بعد الاختبار في التنبؤ بمستوى الأداء في الاختبار على الرغم من دلالة العلاقة بين هذه الاستراتيجيات ودرجات الطلاب في الاختبار يمكن تفسيره بأن هذه الاستراتيجيات لها تأثيرات غير مباشرة في درجات الطلاب في الاختبار حيث أن هذه الاستراتيجيات تؤثر في استراتيجيتي تحليل أسئلة الاختبار والضبط والسيطرة.

وتتفق النتائج الحالية مع معظم نتائج الدراسات السابقة والتي أكدت دلالة العلاقة بين استراتيجيات أداء الاختبار والأداء في الاختبار وأن من يستخدم هذه الاستراتيجيات تزداد احتمالية نجاحه الأكاديمي في الاختبار مثل دراسة Arsal (2014) والتي أشارت نتائجها إلى التأثير الإيجابي لاستراتيجيات أداء الاختبار في خفض قلق الاختبار والذي ينعكس إيجابياً على الأداء في الاختبار؛ ودراسة Zhang et al. (2011); Hong et al. (2006) والتي أكدت نتائجها أن استراتيجيات أداء الاختبار تعد من أهم العوامل المميزة بين منخفضي ومرتفعي الأداء

في الاختبار؛ وكذلك تتفق مع تأكيد (Famularo, 2007, 29) بأن استخدام هذه الاستراتيجيات يرتبط إيجابياً بدرجات الاختبارات؛ وتأكيد Chittooran and Dorothy (2001) بأن الفرصة للنجاح في الأداء على الاختبار تزداد بزيادة استخدام استراتيجيات أداء الاختبار والتي تتمثل في الألفة بالاختبار والاستعداد للاختبار والحكمة الاختبارية وضبط قلق الاختبار؛ وتتفق كذلك مع نتائج دراسة عفاف وادي (٢٠١٣) والتي أشارت نتائجها لإمكانية التنبؤ بدرجات تحصيل الطلاب من خلال مهارات الحكمة الاختبارية لديهم؛ وتتفق مع نتائج دراسة Stenlund et al. (2016) والتي أكدت على دلالة الفروق بين منخفضي ومرتفعي التحصيل في استراتيجيات أداء الاختبار لصالح مرتفعي التحصيل.

وتتفق كثيراً مع نتائج دراسة Wise et al. (2006) والتي أشارت إلى أن استبعاد الأفراد غير الجادين في الإجابة عن الاختبار وهم منخفضي دافعية أداء الاختبار يزيد من معدل الدرجات ما بين (٠.٢٥) إلى (٠.٥) انحراف معياري؛ ونتائج دراسة Wise and DeMars (2010) والتي أكدت أن استبعاد الأفراد منخفضي دافعية أداء الاختبار يصحبه زيادة في متوسط الدرجات؛ والنتائج الحالية تؤكدها نتائج العديد من الدراسات التي أشارت إلى جدوى التدريب على استراتيجيات أداء الاختبار أو مهارات الحكمة الاختبارية في رفع مستوى الأداء على الاختبار، وخفض قلق الاختبار والذي ينعكس إيجابياً على مستوى الأداء في الاختبار (Holzer et al., 2009; Dodeen, 2015; Takallou et al., 2016) ودراسات (أحمد عودة، ١٩٨٩؛ إيثار الخصاونة، ١٩٩٤)؛ وكذلك تتفق النتائج الحالية مع نتائج دراسة Ellis and Ryan (2003) والتي أشارت إلى أن تدريب الطلاب على استراتيجيات ومهارات أداء الاختبار يزيد من متوسط الدرجات في الاختبار بمعدل ما بين (٠.٢٥) إلى (١) انحراف معياري؛ وكذلك تتفق مع نتائج دراسة محمد عبدالوهاب (٢٠٠٧) والتي أكدت على أن مهارات حكمة الاختبار تسهم بنحو ١٤% في درجات التحصيل وكذلك استفادة الطلاب في المجموعة التجريبية من التدريب على مهارات الحكمة الاختبارية في رفع مستوى التحصيل؛ وتتفق كذلك مع نتائج دراسة Kim and Goetz (1993) التي أكدت على دلالة العلاقة بين استخدام استراتيجيات أداء الاختبار بمستوى الأداء.

بينما تختلف النتائج الحالية مع نتائج دراسة محمد عبدالوهاب (٢٠٠٧) في أن مهارات حكمة الاختبار تسهم بنحو ٥% فقط في قلق الاختبار وكذلك تختلف معها في أن التدريب على مهارات الحكمة الاختبارية لم يسهم في خفض قلق الاختبار؛ كذلك لا تتفق مع نتائج دراسة سامي العنزي (٢٠١٤) والتي توصلت إلى عدم دلالة الفروق في مهارات الحكمة الاختبارية باختلاف مستوى التحصيل؛ كذلك تختلف مع نتائج دراسة سعود العنزي (٢٠١٤) والتي توصلت لعدم اختلاف مهارات الحكمة الاختبارية باختلاف مستوى التحصيل؛ وتختلف مع نتائج دراسة ديانا حماد (٢٠١٠) والتي أكدت عدم دلالة العلاقة بين مهارات الحكمة الاختبارية والتحصيل الدراسي؛ ولا تتفق مع نتائج دراسة Zhang et al. (2011) والتي أشارت إلى أن استراتيجيات الأداء في الاختبار لدى الطلاب الصينين تربطها علاقات ضعيفة بالأداء في اختبارات اللغة الإنجليزية؛ والاختلاف مع بعض نتائج الدراسات السابقة وخاصة التي تناولت الحكمة الاختبارية يؤكد على أن المفهوم الشائع لمهارات الحكمة الاختباري يختلف عن استراتيجيات أداء الاختبار بمفهومها المستخدم في البحث الحالي.

رابعاً: نتائج الإجابة عن السؤال الرابع:

ينص السؤال الرابع للبحث الحالي على "هل تختلف الخصائص السيكومترية لبنود الاختبار (الصعوبة والتمييز) باختلاف درجة استخدام استراتيجيات أداء الاختبار لدى الطلاب عينة البحث الحالي؟"

١- صعوبة البنود:

للتعرف على مدى اختلاف صعوبات بنود الاختبار التحصيلي باختلاف درجة استخدام استراتيجيات أداء الاختبار لدى الطلاب عينة البحث الحالي، تم حساب معاملات صعوبة البنود في حالة العينة الكلية، وكذلك حساب معاملات الصعوبة في حالة استبعاد أدنى ١٠%، ٢٠%، ٣٠% في استراتيجيات أداء الاختبار على الترتيب، فكانت كما هي موضحة بجدول (١٣):

جدول (١٣):

معاملات صعوبة البنود في ضوء استخدام استراتيجيات أداء الاختبار

العينة الكلية (ن= ١٨١)				بعد استبعاد أدنى ١٠% (ن= ١٦١)				بعد استبعاد أدنى ٢٠ % (ن=١٤١)				بعد استبعاد أدنى ٣٠% (ن= ١٢٢)			
م	صعوبة	م	صعوبة	م	صعوبة	م	صعوبة	م	صعوبة	م	صعوبة	م	صعوبة	م	صعوبة
١	٠.٧٢٤	٢١	٠.٤٥٣	١	٠.٧٥٢	٢١	٠.٤٧٢	١	٠.٧٥٢	٢١	٠.٥٠٤	١	٠.٧٩٥	٢١	٠.٥٢٥
٢	٠.٦٥٧	٢٢	٠.٤٦٤	٢	٠.٦٨٣	٢٢	٠.٤٧٢	٢	٠.٧٢٣	٢٢	٠.٤٩٦	٢	٠.٧٧٩	٢٢	٠.٥٣٣
٣	٠.٦٠٨	٢٣	٠.٤٨٦	٣	٠.٦٤٠	٢٣	٠.٤٩٧	٣	٠.٦٨٨	٢٣	٠.٥١٨	٣	٠.٧٤٦	٢٣	٠.٥٥٤
٤	٠.٥٨٠	٢٤	٠.٤٧٠	٤	٠.٦٠٢	٢٤	٠.٤٨٤	٤	٠.٦٥٢	٢٤	٠.٥٠٤	٤	٠.٧٢١	٢٤	٠.٥٣٣
٥	٠.٥٦٩	٢٥	٠.٤٧٠	٥	٠.٥٩٠	٢٥	٠.٤٨٤	٥	٠.٦٤٥	٢٥	٠.٥٠٤	٥	٠.٦٩٧	٢٥	٠.٥٢٥
٦	٠.٥٥٨	٢٦	٠.٤٤٨	٦	٠.٥٨٤	٢٦	٠.٤٤٧	٦	٠.٦٣١	٢٦	٠.٤٦٨	٦	٠.٦٨٩	٢٦	٠.٤٨٤
٧	٠.٤٩٧	٢٧	٠.٤٥٩	٧	٠.٥٣٤	٢٧	٠.٤٩٧	٧	٠.٥٦٠	٢٧	٠.٥٢٥	٧	٠.٦٠٧	٢٧	٠.٥٥٧
٨	٠.٤٧٥	٢٨	٠.٤٣٦	٨	٠.٤٩١	٢٨	٠.٤٤١	٨	٠.٥١٨	٢٨	٠.٤٦٨	٨	٠.٥٦٦	٢٨	٠.٥٠٠
٩	٠.٤٩٧	٢٩	٠.٤٥٣	٩	٠.٥٢٨	٢٩	٠.٤٩١	٩	٠.٥٧٤	٢٩	٠.٥١١	٩	٠.٦١٥	٢٩	٠.٥٤٩

العينة الكلية (ن= ١٨١)				بعد استبعاد أدنى ١٠% (ن= ١٦١)				بعد استبعاد أدنى ٢٠ % (ن=١٤١)				بعد استبعاد أدنى ٣٠% (ن= ١٢٢)			
١٠	٠.٤٥٣	٣٠	٠.٤٣٦	١٠	٠.٤٦٠	٣٠	٠.٤٦٠	١٠	٠.٤٩٦	٣٠	٠.٤٨٩	١٠	٠.٥٢٥	٣٠	٠.٥٢٥
١١	٠.٤٤٨	٣١	٠.٤٣٦	١١	٠.٤٩١	٣١	٠.٤٤١	١١	٠.٥٠٤	٣١	٠.٤٧٥	١١	٠.٥٣٣	٣١	٠.٥١٦
١٢	٠.٤٧٠	٣٢	٠.٥١٩	١٢	٠.٥٠٣	٣٢	٠.٥٥٣	١٢	٠.٥٣٩	٣٢	٠.٥٦٠	١٢	٠.٥٤٩	٣٢	٠.٦٠٧
١٣	٠.٤٥٩	٣٣	٠.٥٠٣	١٣	٠.٤٦٦	٣٣	٠.٥١٦	١٣	٠.٤٨٢	٣٣	٠.٥٥٣	١٣	٠.٥٢٥	٣٣	٠.٥٨٢
١٤	٠.٤٢٠	٣٤	٠.٤٧٥	١٤	٠.٤٢٩	٣٤	٠.٥٠٣	١٤	٠.٤٥٤	٣٤	٠.٥٣٢	١٤	٠.٤٨٤	٣٤	٠.٥٧٤
١٥	٠.٤٦٤	٣٥	٠.٤٦٤	١٥	٠.٥٠٣	٣٥	٠.٤٨٤	١٥	٠.٥٠٤	٣٥	٠.٤٨٩	١٥	٠.٥٣٣	٣٥	٠.٥١٦
١٦	٠.٤٥٩	٣٦	٠.٤٧٠	١٦	٠.٤٧٨	٣٦	٠.٤٩٧	١٦	٠.٤٩٦	٣٦	٠.٥١٨	١٦	٠.٥٠٨	٣٦	٠.٥٥٧
١٧	٠.٤٣١	٣٧	٠.٥١٤	١٧	٠.٤٦٠	٣٧	٠.٥١٨	١٧	٠.٤٨٩	٣٧	٠.٥٣٢	١٧	٠.٤٩٢	٣٧	٠.٥٤٩
١٨	٠.٤٣٦	٣٨	٠.٤٧٥	١٨	٠.٤٥٣	٣٨	٠.٥٠٩	١٨	٠.٤٧٥	٣٨	٠.٥٣٢	١٨	٠.٤٧٥	٣٨	٠.٥٥٧
١٩	٠.٥١٤	٣٩	٠.٥٠٨	١٩	٠.٥٥٣	٣٩	٠.٥٢٢	١٩	٠.٥٧٤	٣٩	٠.٥٤٦	١٩	٠.٦٣١	٣٩	٠.٥٥٧
٢٠	٠.٤٦٤	٤٠	٠.٥٥٢	٢٠	٠.٤٧٢	٤٠	٠.٥٧١	٢٠	٠.٥١١	٤٠	٠.٥٨٩	٢٠	٠.٥٥٧	٤٠	٠.٦٢٣

بعد حساب معاملات صعوبة البنود في الأربع حالات السابقة، ونظراً لأن معاملات الصعوبة في جوهرها نسب مئوية وتنتمي إلى مستوى القياس الرتبي تم تحويل معاملات الصعوبة إلى درجات معيارية وللتغلب على مشكلة الدرجات السالبة تم تحويل الدرجات المعيارية إلى معاملات دلتا المقابلة (Δ) حيث أن ($\Delta = 13 + 4z$) وتمثل z الدرجة المعيارية المقابلة لكل معامل صعوبة، بمتوسط ١٣ وانحراف معياري ٤ (Domino & Domino, 2006, 30) وتمتد قيمة معامل الصعوبة دلتا ما بين (٦) في حالة المفردات الصعبة جداً و(٢٠) في حالة المفردات السهلة جداً وتساوي ١٣ في حالة المفردات متوسطة الصعوبة والسهولة (Liao et al., 2010, 4)؛ ثم تم استخدام تحليل التباين للقياسات المتكررة للكشف عن دلالة الفروق بين معاملات صعوبة البنود دلتا في الحالات الأربعة؛ فكانت النتائج كما هي موضحة في التالي:

جدول (١٤):

الإحصاءات الوصفية لمعاملات صعوبة البنود دلتا في ضوء درجة استخدام استراتيجيات أداء الاختبار

المجموعات	المتوسط	الانحراف المعياري
الأفراد ككل	١١.٠٢١	٣.٣٣٠
بعد استبعاد أدنى ١٠ % في استراتيجيات أداء الاختبار	١٢.١٤٥	٣.٥١٩
بعد استبعاد أدنى ٢٠ % في استراتيجيات أداء الاختبار	١٣.٥١٩	٣.٦٧٠
بعد استبعاد أدنى ٣٠ % في استراتيجيات أداء الاختبار	١٥.٣١٥	٤.٢١٢

جدول (١٥):

دلالة الفروق في صعوبات البنود دلتا في ضوء درجة استخدام استراتيجيات أداء الاختبار (باستخدام تحليل التباين للقياسات المتكررة)

مصدر التباين	مجموع المربعات	درجات الحرية	التباين	النسبة الفائية
درجة استخدام الاستراتيجيات	٤١١.٠٦٠	٣	٢٧٥.٤٩٨	**٢٧٦.٢١١
الخطأ	٥٨.٠٤٠	١١٧	٠.٩٩٧	

يتضح من جدول (١٥) أنه توجد فروق دالة إحصائياً عند مستوى ثقة ٠.٠١ في معاملات صعوبة البنود في الحالات الأربعة التي تمت المقارنة بينها، مما يعني أنه توجد فروق في معاملات صعوبة البنود راجعة لاختلاف استخدام استراتيجيات أداء الاختبار، وتم استخدام اختبار أقل فرق دال LSD في المقارنات البعدية بين الحالات الأربعة في معاملات الصعوبة دلتا فكانت النتائج كما هي موضحة في جدول (١٦):

جدول (١٦):

المقارنات البعدية لمعاملات صعوبة البنود في ضوء استخدام استراتيجيات أداء الاختبار

المجموعات	الأفراد ككل (م= ١١.٠٢١)	بعد استبعاد أدنى ١٠ % (م= ١٢.١٤٥)	بعد استبعاد أدنى ٢٠ % (م= ١٣.٥١٩)
بعد استبعاد أدنى ١٠ % (م= ١٢.١٤٥)	**١.١٢٤		
بعد استبعاد أدنى ٢٠ % (م= ١٣.٥١٩)	**٢.٤٩٨	**١.٣٧٤	
بعد استبعاد أدنى ٣٠ % (م= ١٥.٣١٥)	**٤.٢٩٤	**٣.١٧٠	**١.٧٩٦

يتضح من جدول (١٦) أنه توجد فروق دالة إحصائياً عند مستوى ثقة ٠.٠١ في معاملات صعوبة البنود المحولة (دلتا) حيث تزداد سهولة البند في المجموعة التي تم فيها استبعاد أدنى ٣٠% من الأفراد في استخدام استراتيجيات أداء الاختبار، بينما كانت أعلى المجموعات في معاملات الصعوبة هي مجموعة الأفراد ككل دون استبعاد متدني استخدام استراتيجيات أداء الاختبار.

٢- تمييز البنود:

للتعرف على مدى اختلاف تمييز بنود الاختبار التحصيلي باختلاف درجة استخدام استراتيجيات أداء الاختبار لدى الطلاب عينة البحث الحالي، تم حساب معاملات تمييز البنود في حالة العينة الكلية، وكذلك في حالة استبعاد أدنى ١٠%، ٢٠%، ٣٠%، وفي كل حالة تم تحديد أعلى وأدني ٢٧% من الطلاب في درجات الاختبار التحصيلي وتم حساب معامل الارتباط بين درجاتهم على بنود الاختبار والدرجة الكلية في الاختبار باعتباره مؤشر لتمييز البنود، فكانت كما هي موضحة بجدول (١٧):

جدول (١٧):

معاملات تمييز البنود في ضوء استخدام استراتيجيات أداء الاختبار

قبل الاستبعاد (ن= ٤٩ لكل مجموعة)				بعد استبعاد أدنى ١٠% (ن= ٤٤ لكل مجموعة)				بعد استبعاد أدنى ٢٠ % (ن=٣٨)				بعد استبعاد دنى ٣٠% (ن= ٣٣ لكل مجموعة)			
م	تمييز	م	تمييز	م	تمييز	م	تمييز	م	تمييز	م	تمييز	م	تمييز	م	تمييز
١	٠.٤٤٢	٢١	٠.٢٢٤	١	٠.٣٩٨	٢١	٠.٢٩٦	١	٠.٤٤٩	٢١	٠.١٠٥	١	٠.٣٩٤	٢١	٠.٠٣٠
٢	٠.٥٦٣	٢٢	٠.٣٤٦	٢	٠.٥٨٢	٢٢	٠.١٥٩	٢	٠.٥١٩	٢٢	٠.٣٤٠	٢	٠.٥٣٨	٢٢	٠.٢٧٣
٣	٠.٤٩٢	٢٣	٠.٣٦٦	٣	٠.٥٠٦	٢٣	٠.٣٦٢	٣	٠.٥٤٢	٢٣	٠.٣٦٦	٣	٠.٤٥٨	٢٣	٠.٣٩٤
٤	٠.٥٠٨	٢٤	٠.٤٨٨	٤	٠.٤٧٨	٢٤	٠.٣٣٩	٤	٠.٤٥٧	٢٤	٠.٤٧١	٤	٠.٢٩٩	٢٤	٠.٥١٥
٥	٠.٥٧٨	٢٥	٠.٣٢٥	٥	٠.٥٣١	٢٥	٠.٤٩٩	٥	٠.٥٤٢	٢٥	٠.٢٣٦	٥	٠.٥١٠	٢٥	٠.٢١١
٦	٠.٥٥٠	٢٦	٠.٢٨٥	٦	٠.٥٠٢	٢٦	٠.٢٤٩	٦	٠.٥٠٥	٢٦	٠.٢٣٦	٦	٠.٣٥٥	٢٦	٠.٢٤٢
٧	٠.٤٠٨	٢٧	٠.٣٤٥	٧	٠.٣١٧	٢٧	٠.٢٢٦	٧	٠.٣١٤	٢٧	٠.٢٠٩	٧	٠.١٥١	٢٧	٠.٢١١
٨	٠.٣٦٦	٢٨	٠.٢٤٦	٨	٠.٣٨٤	٢٨	٠.٢٧١	٨	٠.٤٢١	٢٨	٠.١٠٦	٨	٠.٣٠٨	٢٨	٠.٠٦٠
٩	٠.٤٠٦	٢٩	٠.٣٠٧	٩	٠.٣١٦	٢٩	٠.٢٠٥	٩	٠.٢٨٨	٢٩	٠.٢٣٥	٩	٠.٠٦٣	٢٩	٠.٢١١

قبل الاستبعاد (ن= ٤٩ لكل مجموعة)				بعد استبعاد أدنى ١٠% (ن= ٤٤ لكل مجموعة)				بعد استبعاد أدنى ٢٠ % (ن=٣٨)				بعد استبعاد دنى ٣٠% (ن= ٣٣ لكل مجموعة)			
١٠	٠.٣٠٦	٣٠	٠.٢٨٩	١٠	٠.٣١٧	٣٠	٠.٢٤٩	١٠	٠.٢٦٣	٣٠	٠.٢٠٩	١٠	٠.٢٧١	٣٠	٠.٠٩٠
١١	٠.٣٨٦	٣١	٠.٤٠٨	١١	٠.٣١٧	٣١	٠.٢٠٥	١١	٠.٢٨٨	٣١	٠.٤١٨	١١	٠.٢٤٣	٣١	٠.٣٣١
١٢	٠.٢٤٤	٣٢	٠.٥١١	١٢	٠.٢٢٧	٣٢	٠.٤٠٧	١٢	٠.٢١٢	٣٢	٠.٤٨٨	١٢	٠.١٨٣	٣٢	٠.٤١٠
١٣	٠.٣٠٩	٣٣	٠.٤٣٢	١٣	٠.٣١٨	٣٣	٠.٤٩٠	١٣	٠.٢٣٦	٣٣	٠.٤٨٨	١٣	٠.١٥١	٣٣	٠.٥٢٧
١٤	٠.٢٣٠	٣٤	٠.٣٢٥	١٤	٠.٢٢٨	٣٤	٠.٤٦٤	١٤	٠.٢٣٦	٣٤	٠.٣١٤	١٤	٠.١٢١	٣٤	٠.٢١٢
١٥	٠.٢٢٦	٣٥	٠.٤٦٧	١٥	٠.١٣٧	٣٥	٠.٣١٦	١٥	٠.٢١٠	٣٥	٠.٥٢٦	١٥	٠.١٢٠	٣٥	٠.٥٤٤
١٦	٠.١٢٢	٣٦	٠.٤٠٧	١٦	٠.١١٣	٣٦	٠.٤٧٥	١٦	٠.١٥٧	٣٦	٠.٤٤٦	١٦	٠.١٨١	٣٦	٠.٤٨٧
١٧	٠.٢٠٣	٣٧	٠.٢٠٣	١٧	٠.١٨١	٣٧	٠.٣٨٥	١٧	٠.١٨٣	٣٧	٠.٢٦٣	١٧	٠.١٨١	٣٧	٠.٣٠٢
١٨	٠.١٦٢	٣٨	٠.٣٦٦	١٨	٠.١٨٢	٣٨	٠.٢٢٧	١٨	٠.١٣١	٣٨	٠.٢٨٨	١٨	٠.١٥١	٣٨	٠.٣٠١
١٩	٠.٣٦٦	٣٩	٠.٢٠٣	١٩	٠.٣١٦	٣٩	٠.٢٧٢	١٩	٠.٢٦٣	٣٩	٠.١٨٦	١٩	٠.٠٩٧	٣٩	٠.٢٧٥
٢٠	٠.٢٤٦	٤٠	٠.٣٩١	٢٠	٠.٢٩٦	٤٠	٠.١٨٤	٢٠	٠.٢٦٢	٤٠	٠.٤١٠	٢٠	٠.٠٩٠	٤٠	٠.٤٣٨

ثم تم استخدام تحليل التباين للقياسات المتكررة للكشف عن دلالة الفروق بين معاملات تمييز البنود في الحالات الأربع؛ فكانت النتائج كما هي موضحة في التالي:

جدول (١٨):

الإحصاءات الوصفية لمعاملات تمييز البنود في ضوء درجة استخدام استراتيجيات أداء الاختبار

المجموعات	المتوسط	الانحراف المعياري
الأفراد ككل	٠.٣٥١	٠.١١٥
بعد استبعاد أدنى ١٠ % في استراتيجيات أداء الاختبار	٠.٣٢٣	٠.١٢١
بعد استبعاد أدنى ٢٠ % في استراتيجيات أداء الاختبار	٠.٣٢١	٠.١٣١
بعد استبعاد أدنى ٣٠ % في استراتيجيات أداء الاختبار	٠.٢٧٣	٠.١٤٩

جدول (١٩):

دلالة الفروق في معاملات تمييز البنود في ضوء درجة استخدام استراتيجيات أداء الاختبار

مصدر التباين	مجموع المربعات	درجات الحرية	التباين	النسبة الفائية
درجة استخدام الاستراتيجيات	٠.١٢٥	٣	٠.٠٤٢	**٨.٨٣٠
الخطأ	٠.٥٥٤	١١٧	٠.٠٠٥	

يتضح من جدول (١٩) وجود فروق دالة إحصائياً عند مستوى ٠.٠١ في معاملات تمييز البنود راجعة لاختلاف استخدام استراتيجيات أداء الاختبار، وتم استخدام اختبار أقل فرق دال LSD في المقارنات البعدية بين معاملات التمييز في ضوء درجة استخدام استراتيجيات أداء الاختبار فكانت كما هي موضحة في جدول (٢٠):

جدول (٢٠):

المقارنات البعدية لمعاملات تمييز البنود في ضوء استخدام استراتيجيات أداء الاختبار

المجموعات	الأفراد ككل (م= ٠.٣٥١)	بعد استبعاد أدنى ١٠ % (م= ٠.٣٢٣)	بعد استبعاد أدنى ٢٠ % (م= ٠.٣٢١)
بعد استبعاد أدنى ١٠ % (م= ٠.٣٢٣)	٠.٠٢٨		
بعد استبعاد أدنى ٢٠ % (م= ٠.٣٢١)	**٠.٠٣١	٠.٠٠٣	
بعد استبعاد أدنى ٣٠ % (م= ٠.٢٧٣)	**٠.٠٧٨	*٠.٠٥٠	**٠.٠٤٧

يتضح من جدول (٢٠) أن معاملات تمييز البنود تزداد في المجموعتين قبل استبعاد أي من الأفراد وبعد استبعاد ١٠% فقط من الأفراد بينما تقل معاملات التمييز للبنود في المجموعة التي تم استبعاد أدنى ٢٠% من الأفراد في استخدام استراتيجيات أداء الاختبار وكانت أقل معاملات التمييز في حالة المجموعة التي تم استبعاد أقل ٣٠% من الأفراد المتدنين في استخدام استراتيجيات أداء الاختبار.

ويمكن تفسير النتائج السابقة والخاصة بزيادة سهولة البنود والانخفاض في معاملات التمييز في ضوء خصائص ذوي الاستخدام المنخفض لاستراتيجيات أداء الاختبار حيث يلجأ هؤلاء إلى التخمين العشوائي في حالة البنود الصعبة بينما أصحاب استراتيجيات الأداء يحاولون الاستفادة من معارفهم الجزئية والتامة في الوصول إلى الإجابة الصحيحة وهو ما جعل البنود تبدو أكثر سهولة لهم، وكذلك باستبعاد هؤلاء الأفراد منخفضي استخدام استراتيجيات أداء الاختبار أصبحت الإجابة على البنود أكثر تجانساً مما أسهم في انخفاض تمييز البنود، وهو ما يدل على أن التباين الكبير في استخدام استراتيجيات أداء الاختبار يصحبه زيادة مصطنعة في معاملات صعوبة البنود وفي معاملات تمييزها وهو ما يستدعي الحذر عند تقنين الاختبارات المعرفية في مثل هذه الحالات، وهو ما يتفق مع ما يؤكد عليه Wolf et al. (1995, 350) ففي حالة عدم بذل الممتحنين للجهد الكافي في الإجابة على بنود الاختبار واستخدام الاستراتيجيات المناسبة للوصول للحل الصحيح والبرهنة على امتلاك المعرفة اللازمة يحدث تشويه لمعاملات صعوبة البنود وقدرتها على التمييز.

ويمكن تفسير النتائج الحالية كذلك بأن استبعاد الطلاب منخفضي استخدام استراتيجيات أداء الاختبار يصحبه دقة في تقدير السمة المقاسة، وبالتالي يتناسب أداء الفرد على بنود الاختبار مع درجته الكلية في الاختبار، ففي هذه الحالة يكون الأداء على البنود وبالتالي الدرجة الكلية دالة للمستوى الحقيقي للسمة، وهو ما تؤكد عليه نتائج دراسة باسل أبوفودة (٢٠٠٨) حيث كان هناك زيادة في دقة تقدير صعوبة وتمييز البنود متمثلة في تدني الأخطاء المعيارية في تقديرهما بعد تدريب الطلاب على مهارات الحكمة الاختبارية؛ وفي ضوء فكرة أنه كلما اقتربت صعوبة البنود من المستوى المتوسط تكون أكثر تمييزاً يتأكد أن تقدير معاملات التمييز بعد استبعاد منخفضي استخدام استراتيجيات أداء الاختبار أصبح أكثر دقة، وهو ما يؤكد فكرة أن التباين في استخدام استراتيجيات أداء الاختبار ينتج عنه تضخيم وهمي في معاملات الصعوبة ومعاملات تمييز البنود.

وتتفق النتائج الحالية مع نتائج دراسة فريح العطوي (٢٠١٣) والتي أكدت على أن ارتفاع سمة المخاطرة لدى الممتحنين يصحبه زيادة مصطنعة في الثبات والصعوبة وتمييز البنود ودالة معلومات الاختبار؛ ويتفق مع نتائج دراسة بلال الشواقفة (٢٠١١) والتي أكدت على انخفاض صعوبة البنود وانخفاض معاملات تمييزها لدى طلاب المجموعة التجريبية الذين تم تدريبهم على الحكمة الاختبارية؛ ويتفق مع نتائج دراسة إيثار الخصاونة (١٩٩٤) التي أكدت على زيادة سهولة البنود بعد تدريب الطلاب على مهارات الحكمة الاختبارية وتختلف معها في أن تلك الدراسة أشارت إلى ارتفاع معاملات تمييز البنود بعد التدريب بينما في البحث الحالي استبعاد الأفراد منخفضي استخدام استراتيجيات أداء الاختبار نتج عنه انخفاض في معاملات التمييز؛ ولكنها تختلف مع نتائج دراسة Robie et al. (2001) والتي أكدت على عدم اختلاف الخصائص السيكومترية للبنود (الصعوبة، التمييز) باختلاف دافعية أداء الاختبار.

خامساً: نتائج الإجابة عن السؤال الخامس:

ينص السؤال الخامس للبحث الحالي على "هل تختلف الخصائص السيكومترية للاختبار باختلاف استراتيجيات أداء الاختبار لدى الطلاب عينة البحث الحالي؟"

١- ثبات الاختبار، والخطأ المعياري للقياس:

للتعرف على تأثير درجة استخدام استراتيجيات أداء الاختبار على ثبات الاختبار تم حساب ثبات الاختبار باستخدام معادلة KR-20 في حالة كل الطلاب وكذلك بعد استبعاد أدني ١٠%، ٢٠%، ٣٠% في درجة استخدام استراتيجيات أداء الاختبار، وتم استخدام الاحصائي W في مقارنة معاملات ثبات الاختبار في الحالات الأربع (إبراهيم يعقوب وباسل أبوفودة، ٢٠١٢، ٤٣٦)؛ والذي يتبع توزيع F بدرجات حرية (ن١ −١، ن٢ − ١) والذي يحسب من المعادلة $W= 1-\alpha_1 / 1-\alpha_2$ حيث $\alpha1, \alpha2$ هما معاملي ثبات الاختبار المراد المقارنة بينهما ون١، ن٢ هما عدد الممتحنين في كل مجموعة، وللتعرف على الخطأ المعياري للقياس تم استخدام المعادلة رقم (٢١)، فكانت النتائج كما هي موضحة في التالي:

جدول (٢١):

معاملات ثبات الاختبار والخطأ المعياري للقياس في ضوء استخدام استراتيجيات أداء الاختبار

معاملات الثبات	معامل الثبات	الانحراف المعياري	الخطأ المعياري للقياس
العينة ككل	٠.٧٢٦	٩.٢٤٩	٤.٨٤١
بعد استبعاد أدنى ١٠%	٠.٧٦١	٩.٣١٠	٤.٥٥١
بعد استبعاد أدنى ٢٠%	٠.٧٨٢	٨.٥١٥	٣.٩٧٦
بعد استبعاد أدنى ٣٠%	٠.٨٢٢	٧.٥٥٥	٣.١٨٧

ويتضح من جدول (٢١) أن الخطأ المعياري للقياس يقل باستبعاد الطلاب منخفضي استخدام استراتيجيات أداء الاختبار، وهو ما يترتب على زيادة الثبات ونقص الانحراف المعياري، والجدول التالي يوضح دلالة الفروق بين معاملات ثبات الاختبار في الحالات المختلفة:

جدول (٢٢):

دلالة الفروق في معاملات ثبات الاختبار في ضوء استخدام استراتيجيات أداء الاختبار

معاملات الثبات	العينة ككل (٠.٧٢٦)	بعد استبعاد أدنى ١٠% (٠.٧٦١)	بعد استبعاد أدنى ٢٠% (٠.٧٨٢)
بعد استبعاد أدنى ١٠% (٠.٧٦١)	١.١٤٦		
بعد استبعاد أدنى ٢٠% (٠.٧٨٢)	١.٢٥٧	١.٠٩٦	
بعد استبعاد أدنى ٣٠% (٠.٨٢٢)	*١.٥٣٩	١.٣٤٣	١.٢٢٥

يتضح من جدول (٢٢):

- أن معامل ثبات الاختبار التحصيلي يزداد باستبعاد الطلاب منخفضي استخدام استراتيجيات أداء الاختبار، ولم تكن الفروق بين معاملات الثبات في الحالات المختلفة فروق دالة إحصائياً، ماعدا الفرق بين معاملي الثبات في حالة الدرجات قبل الاستبعاد وبعد استبعاد ٣٠% من الطلاب في ضوء استخدام استراتيجيات أداء الاختبار حيث كانت الفروق في هذه الحالة دالة إحصائياً عند مستوى ٠.٠٥.

٢- صدق الاختبار:

تم الكشف عن تأثير استخدام استراتيجيات أداء الاختبار على صدق الاختبار باستخدام الصدق التلازمي عن طريق حساب معاملات الارتباط بين درجات الطلاب في الاختبار التحصيلي ودرجاتهم في الاختبار النهائي لمقرر القياس والتقويم وذلك بالنسبة للأفراد ككل وللأفراد بعد استبعاد أدنى ١٠%، ٢٠%، ٣٠% في استخدام استراتيجيات أداء الاختبار وباستخدام الصدق التنبؤي بحساب معاملات الارتباط بين درجات الطلاب في الاختبار ومعدلاتهم الأكاديمية في الفصل الدراسي الأول في جميع المقررات.

وللمقارنة بين معاملات الارتباط في الحالات الأربع تم استخدام التحويل اللوغاريتمي لتحويل كل معامل ارتباط إلى مقابلة اللوغاريتمي حتى تقترب معاملات الارتباط من التوزيع الطبيعي (Steiger, 1980, 246)؛ وتستخدم معادلة الدرجة الحرجة للمقارنة بين معاملات الارتباط اللوغاريتمية من المعادلة:

$$r' = (0.5)\log_e\left[\frac{1+r}{1-r}\right]$$

$$z = \frac{r_1' - r_2'}{\sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}}}$$

حيث أن r هي معامل الارتباط و r' هي القيمة اللوغاريتمية المقابلة لمعامل الارتباط من جداول فيشر، و n_1, n_2 تمثل عدد الممتحنين في كل مجموعة، فكانت النتائج كما هي موضحة بجدول (٢٣):

جدول (٢٣):

الصدق التلازمي والصدق التنبؤي لدرجات الطلاب في الاختبار التحصيلي في ضوء استخدام استراتيجيات أداء الاختبار

المجموعات	الصدق التلازمي				الصدق التنبؤي			
	معامل الارتباط	دلالة الفروق في معامل الارتباط			معامل الارتباط	دلالة الفروق في معامل الارتباط		
		(١)	(٢)	(٣)		(١)	(٢)	(٣)
العينة ككل (١)	٠.٦١٥				٠.٤٦٨			
بعد استبعاد أدنى ١٠% (٢)	٠.٦٥٨	٠.٦٥٩			٠.٥٢٢	٠.٦٤٩		
بعد استبعاد أدنى ٢٠% (٣)	٠.٧٣٦	*١.٩٨٤	١.٣١٣		٠.٦١٤	١.٨٢٥	١.١٦٧	
بعد استبعاد أدنى ٣٠% (٤)	٠.٧٨٢	**٢.٨١٢	*٢.١٥٠	٠.٨٦٣	٠.٦٥٠	*٢.٢٥٥	١.٦١٥	٠.٤٧٩

يتضح من جدول (٢٣) أن:

- هناك زيادة في صدق درجات الاختبار كما يعبر عنها بعلاقة درجات الاختبار بدرجات الاختبار النهائي للمقرر (الصدق التلازمي) وعلاقة درجات الاختبار التحصيلي بمعدل الطالب الأكاديمي (الصدق التنبؤي)، حيث أن استبعاد الطلاب منخفضي استخدام استراتيجيات أداء الاختبار ينتج عنه زيادة دالة في علاقة درجات الاختبار التحصيلي بدرجات الاختبار النهائي للمقرر ومعدل الطالب الأكاديمي.

ويمكن تفسير النتائج السابقة في ضوء فكرة مدى تعبير الدرجات في الاختبار عن المستوى الحقيقي للسمة فعدم استخدام الطلاب لاستراتيجيات أداء الاختبار والذي قد يرجع لنقص دافعية أداء الاختبار أو لعدم المعرفة بهذه الاستراتيجيات يؤدى إلى عدم قدرة الطالب على البرهنة على امتلاكه للمعرفة المطلوبة في الاختبار ونقص في الدرجة المستحقة، وهو ما ينتج عنه عدم تعبير الدرجات في الاختبار بصدق عن المستوى الحقيقي للسمة، وباستبعاد هؤلاء الأفراد يزداد صدق تعبير الدرجات عن المستوى الحقيقي للسمة وهو ما يصحبه زيادة في علاقة درجات الاختبار بغيرها من درجات المتغيرات المرتبطة.

وهو ما يتفق مع ما يؤكد عليه Barry and Finney (2009) ; Wise et al. (2006) ; Wise and DeMars (2010); Pohl et al. (2016); Lon and Iliescu (2017) حيث يؤثر عدم بذل الجهد الكافي في أداء الاختبار وعدم استخدام استراتيجيات مناسبة لتحقيق ذلك على صدق الاختبار فاستخدام استراتيجيات أداء الاختبار يُمكن الطالب من البرهنة على معارفه ومهاراته وإظهار مستواه الحقيقي وهو ما يزيد من صدق الدرجات، وتتفق النتائج السابقة مع نتائج دراسة Westfall and D'Costa (1987) والتي أكدت على أن استبعاد الأفراد غير المطابقين يزيد من الصدق التنبؤي لدرجات الاختبار؛ ومع نتائج دراسة Waskiewicz (2011) والتي أكدت أن تنقية درجات الاختبار باستبعاد الطلاب الذين لا يؤدون بجدية على الاختبار يسهم في زيادة دقة درجات الاختبار في التعبير عن مستوى السمة المقاسة.

ومن جهة أخرى يمكن تفسير تأثر صدق وثبات درجات الاختبار بمستوى استخدام الأفراد لاستراتيجيات أداء الاختبار بأن التباين في هذه الاستراتيجيات يمكن اعتباره مصدراً للفروق بين الأفراد في درجات الاختبار بجانب مستوى السمة الحقيقي، وبالتالي استبعاد الأفراد منخفضي استراتيجيات أداء الاختبار يسهم في تحييد أثر الفروق بين الطلاب والراجع لتباين درجة استخدام هذه الاستراتيجيات وتكون الفروق بينهم راجعة لمستوى السمة المقاسة وهو ما يزيد من ثبات الدرجات وصدقها في التعبير عن مستوى السمة، وهذا التفسير يؤيده تأكيد البعض على أن سلوكيات الفرد في الموقف الاختباري والتي من أهمها استخدام استراتيجيات أداء الاختبار تعد من مصادر التباين غير المرغوبة (تباين الخطأ) في الدرجات بجانب التباين الحقيقي (Stenlund et al., 2016; Xu & Wu, 2012; Famularo, 2007; Wise et al., 2010; Penk et al., 2014).

وفي ضوء ذلك يفسر تحسن صدق الاختبار وثباته بأن استبعاد الأفراد منخفضي استخدام استراتيجيات أداء الاختبار يسهم في الحد من التباين غير المرغوب خاصة وأن الموقف الاختباري في البحث الحالي لا يتوقف عليه أي قرارات تخص الطلاب مما أسهم في زيادة نسبة منخفضي استخدام هذه الاستراتيجيات؛ وكذلك يؤكد Crocker (2006, 124) على أن صدق التنبؤات حول السمة المقاسة يتأثر كثيراً بتباين الطلاب في استخدام استراتيجيات أداء الاختبار؛ ولذلك فإن تنقية درجات الاختبار في ضوء مستوى استخدام استراتيجيات أداء الاختبار يسهم في زيادة صدق تعبير درجات الاختبار عن السمة المقاسة (Parham, 1996, 3)؛ وهو ما يتفق مع ما يؤكد عليه Dodeen (2015, 109) بأن استراتيجيات أداء الاختبار تسهم في تحسين صدقه والمترتب على أن تلك الاستراتيجيات تسهم في إظهار المستوى الحقيقي للفرد في السمة هذا من ناحية ومن ناحية أخرى تقلل من تحيز الدرجات للمرتفعين في هذه الاستراتيجيات.

وكذلك تتفق النتائج الحالية مع ما أشار إليه Ghafournia and Afghari (2014, 91) بأن التدريب على استراتيجيات أداء الاختبار يُمكن أن يسهم في زيادة صدق الاختبار وثبات درجاته، حيث أن التدريب على هذه الاستراتيجيات يؤدي إلى الحد من التباين الراجع لاختلاف درجة استخدامها، ويتفق مع ما أكد عليه شيبرد Shepard في ١٩٩٠ بأن التدريب على هذه الاستراتيجيات يسهم في تحقيق التكافؤ بين الأفراد في الأداء على الاختبار (In: Orth, 1995, 5)؛ وهو ما يسهم في الحد من التأثير السلبي للتباين في استخدام تلك الاستراتيجيات على دقة نتائج القياس (Blerkom, 2009, 169).

فبالنسبة لثبات الاختبار التحصيلي في البحث الحالي وتأثره بدرجة استخدام استراتيجيات أداء الاختبار فيلاحظ أن هناك تحسن في معاملات ثبات الاختبار وخاصة في حالة استبعاد أكبر قدر ممكن من منخفضي استخدام هذه الاستراتيجيات وهو ما يمكن تفسره بأن استبعاد هؤلاء الأفراد يصاحبه نقص في تباين الخطأ وبالتالي تزداد نسبة التباين الحقيقي للتباين الكلي للدرجات وهو ما يسهم في زيادة ثبات الاختبار، فمن مسلمات النظرية الكلاسيكية في القياس أن التباين في الدرجات الملاحظة يتكون من التباين الحقيقي وتباين الخطأ، وتباين الخطأ يمكن أن يكون له ثلاثة مكونات يتعلق أحدهما بالفرد والأخر بتفاعل الفرد مع البند وهى من الأخطاء المنتظمة بالإضافة للخطأ العشوائي مجهول المصدر وهو من الأخطاء غير المنتظمة (Westfall & D'Costa, 1987, 3)؛ وبالتالي يمكن تفسير التأثير الإيجابي لاستبعاد الطلاب منخفضي استخدام استراتيجيات أداء الاختبار على ثباته بأن استبعادهم يحد من تباين الخطأ الراجع للفرد وتفاعله مع البنود؛ وتتفق النتائج الحالية مع افتراضات النظرية التقليدية في القياس والمتعلقة بأن الأخطاء المنتظمة لا تؤثر كثيراً على ثبات الدرجات، ولكنها تؤثر أكثر على صدق الدرجات (فريح العطوي، ٢٠١٣، ٢٦)؛ حيث اتضح من النتائج المتعلقة بالثبات بأن هناك تحسن في الثبات لكن الفروق بين معاملات الثبات لم تصل لمستوى الدلالة الإحصائية في غالبية الحالات.

ومن ناحية أخرى يمكن تفسير النتائج الحالية في ضوء أن الأفراد منخفضي استراتيجيات أداء الاختبار قد يعتمدون على التخمين العشوائي في الإجابة عن أسئلة الاختبار وبالتالي استبعاد هؤلاء يسهم في زيادة صدق الاختبار وثباته، حيث أكدت نتائج دراسة باسل أبوفودة (٢٠٠٨) على أن الاعتماد على التخمين في الإجابة يؤثر سلباً على صدق وثبات الاختبار؛ وكذلك يشير Pohl et al. (2016, 4) إلى أن التخمين العشوائي للاستجابة والناتج عن نقص دافعية أداء الاختبار ونقص الحكمة الاختبارية ينتج عنه مشكلات كبيرة تهدد المؤشرات السيكومترية للبند وصدق وثبات الاختبار؛ ويؤكد Robie et al. (2001, 178) على نفس الأمر تقريباً بأن انخفاض دافعية أداء الاختبار وبالتالي ما يترتب عليها من جهد مبذول في محاولة الوصول لأقصى أداء ممكن يؤثر سلباً على الخصائص السيكومترية لأداة القياس.

وتتفق النتائج الحالية مع نتائج دراسة Wise et al. (2006); Wise and DeMars (2010) في أن انخفاض دافعية أداء الاختبار لدى بعض المشاركين ينتج عنها نقص غير حقيقي في معاملات الارتباط بين درجات الاختبار والمحكات الخارجية التي يفترض أن ترتبط بالسمة موضوع القياس؛ وتختلف معها في أن ذلك ينتج عنه زيادة مصطنعة في الثبات حيث أن البحث الحالي أكدت نتائجه أن انخفاض استخدام استراتيجيات أداء الاختبار والذي قد ينتج من انخفاض دافعية أداء الاختبار ينتج عنه نقص في ثبات الاختبار، فنقص استخدام استراتيجيات أداء الاختبار قد ينتج عنه احياناً استجابات عشوائية وعدم قدره على السيطرة على الانفعالات المصاحبة للموقف الاختباري والذي ينتج عنه تدني في مستوى الدرجات وعدم تجانس في أداء الافراد على بنود الاختبار وبالتالي قد ترجع نسبة كبيرة من التباين في الدرجات إلى تباين غير حقيقي لا يعكس التباين في السمة وهو ما يؤدى الي انخفاض ثبات الاختبار.

وفي ضوء نتائج البحث الحالي الخاصة بصعوبة البنود وتمييزها والتي اتضح منها أن استبعاد الأفراد منخفضي استخدام استراتيجيات أداء الاختبار نتج عنه زيادة في سهولة البنود ونقص في معاملات التمييز، والذي قد ينتج عنه نقص في التباين الكلي للدرجات وفي الوقت نفسه يقل تباين الخطأ ويكون النقص في تباين الدرجات الحقيقية محدود مقارنة بالنقص في التباين الكلي للدرجات وهو ما يسهم في زيادة معامل ثبات الاختبار، وقد يبرر التفسير السابق أنه على الرغم من تحسن معاملات الثبات إلا ان الفروق بين معاملات الثبات في الحالات الأربع للبحث الحالي كانت غير دالة في أغلب الأحيان.

وتختلف النتائج الحالية مع نتائج دراسة Robie et al. (2001) والتي توصلت إلى أن اختلاف دافعية الطلاب لأداء الاختبار لا يؤثر على الخصائص السيكومترية للاختبار من حيث الصدق البنائي والثبات؛ وتتفق النتائج الحالية كثيراً مع نتائج دراسة Maspons and Llabre (1983) في أن مستوى استراتيجيات أداء الاختبار لا يؤثر كثيراً في ثبات الاختبار ولكنه يؤثر على صدقه؛ كذلك تتفق مع ما يشير إليه Clause (1996, 5) بأن التعرف على المستجيبين الذين لا تمثل درجاتهم في الاختبار مستواهم الحقيقي في السمة بناءً على محكات ومؤشرات دقيقة يسهم في دقة تقدير الخصائص السيكومترية للاختبار ودقة القرارات المترتبة على الأداء فيه.

التوصيات والمقترحات:

- ضرورة تدريس وتدريب الطلاب على استراتيجيات أداء الاختبار والذي يمكن أن يزيد من صدق الاختبار بتحييد أثر الفروق بين الطلاب في الاستراتيجيات على نتائج القياس.
- في حالة الدراسات التجريبية يجب الحذر عند استخدام اختبارات الأداء الأقصى في سياقات لا تمثل أهمية بالنسبة للمشاركين في التجربة، حيث أن التطبيق البعدي للاختبار قد يصاحبه زيادة في دافعية المشاركين لأداء الاختبار وهو ما قد يسهم في زيادة معدل بذل الجهد واستخدام استراتيجيات أفضل في أداء الاختبار، وبالتالي قد تكون الفروق بين القياسين القبلي والبعدي ترجع في جزءٍ كبيرٍ منها لاختلاف تلك المتغيرات في مرتي التطبيق وليس ناتج لتحسن في المتغيرات موضوع التجربة.
- ضرورة العمل على حفز المشاركين في البحوث والدراسات التربوية التي تعتمد على استخدام اختبارات الأداء الأقصى بصفة عامة لبذل الجهد ومحاولة الوصول لأقصى استجابة ممكنة، بالقدر الذي يضمن تعبير درجات القياس عن المستوى الحقيقي للسمة، لما لذلك من تأثير إيجابي على خصائص الاختبار السيكومترية وبنوده.
- استبعاد الحالات غير الجادة في الاستجابة عند تقنين أدوات البحوث لما لذلك من تأثير إيجابي على دقة تقدير الخصائص السيكومترية للأداة، فوجود مثل هؤلاء يؤدي إلى تضخيم وهمى في معاملات الصعوبة والتمييز، ويهدد صدق وثبات أدوات القياس.

وفي ضوء النتائج الحالية يمكن اقتراح اجراء البحوث التالية:

- تحري تأثير استخدام استراتيجيات أداء الاختبار على خصائصه السيكومترية باستخدام نماذج نظرية الاستجابة للمفردة.
- دراسة تأثير استخدام استراتيجيات أداء الاختبار على مؤشرات مطابقة الفرد وجودة أداؤه في الاختبار.
- الكشف عن المتغيرات النفسية والمعرفية التي تنبأ باستخدام الطلاب لاستراتيجيات أداء الاختبار.
- المقارنة بين استراتيجيات أداء الاختبار لدى شرائح عمرية مختلفة ومراحل دراسية مختلفة.

المراجع:

إبراهيم محمد يعقوب وباسل خميس أبوفودة (٢٠١٢). أثر مخالفة قواعد صياغة فقرات الاختيار من متعدد على الخصائص السيكومترية للاختبار وفقراته. ***مجلة جامعة دمشق***، ٢٨(١)، ٤١٩-٤٤٣.

أحمد سليمان عودة (١٩٨٩). أثر تعليم استراتيجية التخمين في اكتساب الطلبة بالمستوى الجامعي لمهارة حكمة الاختبار: دراسة تجريبية. ***المجلة التربوية***، جامعة الكويت، ٦(٢٠)، ١٢٣-١٣٧.

إياد محمد حمادنة (٢٠١١). درجة استخدام الطلبة بالمستوي الجامعي لاستراتيجيات حكمة الاختبار. ***مجلة المنارة للبحوث والدراسات***، الأردن، ١٧(١)، ٢٩١-٣١٢.

إيثار محمد الخصاونة (١٩٩٤). ***أثر برنامج تدريبي لحنكة الاختبار على توزيع علامات الاختبار لأداء الطلبة والخصائص الإحصائية للفقرات في مادتي اللغة العربية والرياضيات للصف السادس والصف التاسع***. رسالة ماجستير غير منشورة، كلية الدراسات العليا، الجامعة الأردنية، الأردن.

باسل خميس أبوفودة (٢٠٠٨). ***أثر تعلم حكمة الاختبار على افتراض أحادية البعد والتقديرات المختلفة لنظرية استجابة الفقرة***. رسالة دكتوراه غير منشورة، كلية التربية، جامعة اليرموك، الأردن.

بلال محمود الشواقفة (٢٠١١). ***أثر تعلم حكمة الاختبار على معالم الفقرات والاختبار عند مستويين من قدرات المفحوصين باستخدام النموذج ثلاث المعلمة***. رسالة دكتوراه غير منشورة، كلية التربية، جامعة اليرموك، الأردن.

بندر زيد سالم الأمير (٢٠١٦). الأخطاء الشائعة في بناء مفردات اختبار الاختيار من متعدد وتأثيرها على الخصائص السيكومترية للاختبارات التحصيلية. ***المجلة العربية للعلوم الاجتماعية***، المؤسسة العربية للاستشارات العلمية وتنمية الموارد البشرية، جمهورية مصر العربية، ٩(٤)، ١٣٧-١٧٥.

جابر محمد عبدالله عيسى (٢٠١٤). تأثير عدد بدائل الاستجابة لبنود الاختبار التحصيلي اختيار من متعدد على خصائصه السيكومترية. ***مجلة كلية التربية***، جامعة الأزهر، ١٥٧(١)، ١٢١-١٧٠.

ديانا فهمي علي حماد (٢٠١٠). علاقة الحكمة الاختبارية بالأداء الناتج من اختبار تحصيلي ذو اختيار من متعدد مبني وفق نموذج راش لدى طالبات كلية التربية للبنات الأقسام الأدبية بجامعة أم القرى. ***مجلة دراسات عربية في التربية وعلم النفس***، رابطة التربويين العرب، ٤(٤)، ٢٩٧-٣٣٨.

زين حسن ردادي (٢٠٠١). الحكمة الاختبارية وعلاقتها بنوع الطالب وتخصصه وتحصيله الدراسي. ***مجلة كلية التربية***، جامعة الزقازيق، ٣٩، ١-٣٤.

سامي عايد العنزي (٢٠١٤). علاقة الحكمة الاختبارية بتحصيل طلبة المرحلة الثانوية في دولة الكويت في اختبار تحصيلي مبني وفق نموذج راش. ***عالم التربية***، المؤسسة العربية للاستشارات العلمية وتنمية الموارد البشرية، جمهورية مصر العربية، ١٥(٤٧)، ٣٤٥-٣٦٩.

سعود بشير العنزي (٢٠١٤). مهارات الحكمة الاختبارية في ضوء متغيرات النوع والقدرة العامة والتحصيل الدراسي لدى طلاب السنة التحضيرية في جامعة الحدود

الشمالية. ***مجلة التربية***، كلية التربية، جامعة الأزهر، ١٦١(٢)، ٣٢٧-٣٦٤.

السيد محمد أبوهاشم (٢٠٠٨). النموذج البنائي التنبؤي لمهارات الدراسة والحكمة الاختبارية والتحصيل الدراسي لدى طلاب المرحلة الثانوية. ***مجلة كلية التربية***، جامعة المنصورة، ٦٨(١)، ٢١٠-٢٧٠.

صلاح الدين محمود علام (٢٠٠٦). ***القياس والتقويم التربوي والنفسي***. القاهرة: دار الفكر العربي.

صلاح الدين محمود علام (٢٠١٢). ***الاختبارات والمقاييس التربوية والنفسية***. عمان: دار الفكر.

عصام الدسوقي إسماعيل الجبة (٢٠١٢). الخصائص السيكومترية لمفردات اختيار من متعدد تتضمن أكثر من إجابة صحيحة. ***مجلة كلية التربية***، جامعة المنصورة، ٨١، ١٤١-١٧٤.

عفاف زياد وادي (٢٠١٣). مهارات الحكمة الاختبارية وعلاقتها بالتحصيل الدراسي لدى طلبة كلية التربية/ ابن الهيثم للعلوم الصرفة. ***مجلة البحوث التربوية والنفسية***، جامعة بغداد، العراق، ٣٧، ٢٩٥-٣٢٢.

فاطمة عباس مطلك (٢٠٠٩). قياس مستوى الحكمة الاختبارية لدى الطلبة المتميزين والمتميزات وأقرانهم الاعتياديين من المرحلة الإعدادية (دراسة مقارنة في مركز محافظة نينوي). ***مجلة كلية الآداب***، جامعة بغداد، العراق، ٩١، ٥٦٦-٥٩٨.

فريح محمد العطوي (٢٠١٣). ***أثر المخاطرة على الخصائص السيكومترية لفقرات الاختبار وقدرات الأفراد***. رسالة ماجستير غير منشورة، كلية التربية، جامعة اليرموك، الأردن.

مايسة فاضل أبومسلم (٢٠١٦). أثر كل من عدد البدائل وترتيبها لمفردات اختبار تحصيلي دراسي من نوع الاختيار من متعدد على الخصائص السيكومترية للاختبار ومفرداته. ***مجلة اتحاد الجامعات العربية للتربية وعلم النفس***، ١٤(٢)، ١٥٤-١٨٧.

مجدي محمد أحمد الشحات (٢٠٠٧). أثر التدريب على الحكمة الاختبارية على مستوى كل من قلق الاختبار والتحصيل الدراسي. ***مجلة كلية التربية***، جامعة بنها، ١٧(٦٩)، ١-٣٧.

محمد محمود محمد عبدالوهاب (٢٠٠٧). ***أثر برنامج تدريبي لمهارة حكمة الاختبار على مستوى الأداء التحصيلي وقلق الاختبار لدى عينة من طلبة كلية التربية بالمنيا***. رسالة ماجستير غير منشورة، كلية التربية، جامعة المنيا.

نعمان الموسوي (١٩٩٧). تحديد العدد الأمثل لخيارات الفقرة في اختبار الاختيار من متعدد في ضوء بعض خصائصه السيكومترية. ***المجلة التربوية***، جامعة الكويت، ١١(٤٤)، ١٩٩-٢٥٦.

هدى محمد عبدالله العريفي (٢٠٠٨). ***تباين الخصائص السيكومترية لاختبار متعدد البدائل بتباين عدد مفرداته وبعض الخصائص الدافعية للمفحوصين***. رسالة ماجستير غير منشورة، كلية التربية، جامعة أم القرى، المملكة العربية السعودية.

Amer, A. (2007). EFL/ESL Test-Wiseness and Test-taking Strategies. ***ERIC Number: ED497399.***

Arsal, Z. (2014). Relationships between Test Anxiety and Test-Taking Strategies. ***Faculty of Education Journal***, Abant İzzet Baysal University, Turkey, 14(2), 306-320.

Banks, T., & Eaton, I. (2014). Improving Test-Taking Performance of Secondary At-Risk Youth and Students with Disabilities. ***Preventing School Failure***, 58(4), 207–213. Doi: https://doi.org/10.1080/1045988X.2013.792765.

Barry, C., & Finney, S. (2009, October). ***Exploring Change in Test-Taking Motivation***. Paper Presented at Annual Meeting of the Northeastern Educational Research Association, Rocky Hill, New Jersey.

Bicak, B. (2013). Scale for Test Preparation and Test Taking Strategies. ***Educational Sciences: Theory & Practice***, 13(1), 279-289.

Blerkom, M. (2009). ***Measurement and Statistics for Teachers***. New York: Routledge, Taylor & Francis Group.

Chittooran, M., & Dorothy, M. (2001, April). ***Test-Taking Skills for Multiple-Choice Formats: Implications for School Psychologists.*** Annual Conference of the National Association of School Psychologists, Washington.

Cizek, G.; Bowen, D., & Church, K (2010). Sources of Validity Evidence for Educational and Psychological Tests: A Follow-Up Study. ***Educational and Psychological Measurement***, 70(5), 732-743. DOI: https://doi.org/10.1177/0013164410379323.

Clause, C. (1996). ***The Effects of Test Appropriateness on Reliability and Validity***. Master Degree Thesis, Michigan State University.

Colliver, J.; Conlee, M., & Verhulst, S. (2012). From Test Validity to Construct Validity and back?. ***Medical Education***, 46, 366–371. DOI: https://doi.org/10.1111/j.1365-2923.2011.04194.x.

Crocker, L. (2006). Preparing Examinees for Test Taking: Guidelines for Test Developers and Test Users. In: S., M., Downing & T., M., Haladyna (Eds.), ***Handbook of test development***, (Pp.115-128), New Jersey: Lawrence Erlbaum Associates, Publishers.

Cui, Y., & Roberts, R. (2013). Validating Student Score Inferences with Person-Fit Statistic and Verbal Reports: A Person-Fit Study for Cognitive Diagnostic Assessment. ***Educational Measurement: Issues and Practice***, 32 (1), 34-42. DOI: https://doi.org/10.1111/emip.12003.

Dodeen, H. (2008). Assessing Test-Taking Strategies of University Students: Developing a Scale and Estimating its Psychometric Indices. ***Assessment & Evaluation in Higher Education***, 33(4), 409-419. Doi: https://doi.org/10.1080/02602930701562874.

Dodeen, H. (2009). Test-Related Characteristics of UAEU Students: Test-Anxiety, Test-Taking Skills, Guessing, Attitudes toward Tests, and Cheating. ***Faculty of Education Journal,*** United Arab Emirates University, 26, 31-66.

Dodeen, H. (2015). Teaching Test-Taking Strategies: Importance and Techniques. ***Psychology Research***, 5(2), 108-113.

Dodeen, H.; Abdelfatah, F., & Alshumrani, S. (2014). Test-Taking Skills of Secondary Students: The Relationship with Motivation, Attitudes, Anxiety and Attitudes towards Tests. ***South African Journal of Education***, 34(2), 1-18.

Domino, M., & Domino, G. (2006). ***Psychological Testing: An Introduction***. 2nd ed. Cambridge University Press

Ellis, A., & Ryan, A. (2003). Race and Cognitive-Ability Test Performance: The Mediating Effects of Test Preparation, Test-Taking Strategy Use and Self-Efficacy. ***Journal of Applied Social Psychology***, 33(12), 2607-2629. DOI: https://doi.org/10.1111/j.1559-1816.2003.tb02783.x.

Eme, J., & Udoh, A. (2011). Some Psycho-Academic Variables as Determinants of Students Performance in English Studies in Akwa Ibom State, Nigeria. ***International Multidisciplinary Research Journal***, 5(2), 177-188.

Famularo, L. (2007). ***The Effect of Response Format and Test Taking Strategies on Item Difficulty: A Comparison of Stem-Equivalent Multiple-Choice and Constructed-Response Test Items***. PH. D. Dissertation, Lynch School of Education, Boston College.

Ghafournia, N., & Afghari, A. (2014). The Interaction among Using Test-Taking Strategies، Level of Language Proficiency, and Test Performance. ***Iranian Journal of Research in English Language Teaching***, 1(3), 83-94.

Guion, R. (1998). ***Assessment, Measurement, and Prediction for Personnel Decisions***. New Jersey: Erlbaum.

Holzer, M.; Madaus, J.; Bray, M., & Kehle, T. (2009). The Test-Taking Strategy Intervention for College Students with Learning Disabilities***.*** ***Learning Disabilities Research & Practice***, 24(1), 44–56.

Hong, E.; Sas, M., & Sas, j. (2006). Test-Taking Strategies of High and Low Mathematics Achievers. ***The Journal of Educational Research***, 99(3), 144-155. DOI: https://doi.org/10.1111/j.1540-5826.2008.01276.x.

Johnson, E., & Arnold, N. (2004). Validating an Alternate Assessment. ***Remedial and Special Education***, 25(5). 266-275. DOI: https://doi.org/10.1177/07419325040250050101

Kim, Y., & Goetz, E. (1993). Strategic Processing of Test Questions: The Test Marking Responses of College Students. ***Learning and individual Differences***, 5(3), 211-218. DOI: https://doi.org/10.1016/1041-6080(93)90003-B

Lancaster, P.; Schumaker, J.; Lancaster, S., & Deshler, D. (2009). Effects of a Computerized Program on Use of the Test Taking Strategy by Secondary Students with Disabilities. ***Learning Disability Quarterly***, 32(3), 165-179. DOI: https://doi.org/10.2307/27740366

Liao, C.; Hatrak, N., & Yu, F. (2010). Comparison of Content, Item Statistics and Test-Taker Performance for the Redesigned and Classic TOEIC Listening and Reading Test. ***Test of English for International Communication (TOEIC) Compendium***, 4, 1-9.

Lon, A., & Iliescu, D. (2017). The Measurement Equivalence of Personality Measures across High- and Low-Stake Test Taking Settings. ***Personality and Individual Differences***, 110, 1–6. DOI: https://doi.org/10.1016/j.paid.2017.01.008

Maspons, M., & Llabre, M. (1983, April). ***Effect of Pretraining Hispanic Students on Test-Taking Strategies on the Reliability and Predictive Validity of a Mathematics Predictor test***. Paper presented at the Annual Meeting of the American Educational Research Association, Montreal, Quebec.

Nguyen, H. (2003). ***Constructing a New Theoretical Framework for Test Wiseness and Developing the Knowledge of Test-Taking Strategies (KOTTS) Measure***. Master Degree Thesis, Michigan State University.

Oakland, T.; Callueng, C., & Harris, G. (2012). The Impact of Test-Taking Behaviors on WISC-IV Spanish Domain Scores in Its Standardization Sample. ***Journal of Psychoeducational Assessment***, 30, 139-147. DOI: https://doi.org/10.1177/0734282911423358

Orth, J. (1995). ***The Effect of Test Taking Strategies for Multiple-Choice Tests***. PH. D. Thesis, Graduate Faculty, School of Education, University of Pittsburgh, Pennsylvania.

Parham, S. (1996). ***The Relationships between Test-Taking Strategies and Cognitive Ability Test Performance***. PH. D. Thesis, Graduate College, Bowling Green State University, Ohio.

Penk, C.; Pohlmann, C., & Roppelt, A. (2014). The role of Test-Taking Motivation for Students' Performance in Low-Stakes Assessments: An Investigation of School-Track-Specific Differences. ***Large-scale Assessments in Education***, 2(5), 1-17. DOI: https://doi.org/10.1186/s40536-014-0005-4

Pohl, S.; Sudkamp, A.; Hardt, K., Carstensen, C., & Weinert, S. (2016) Testing Students with Special Educational Needs in Large-Scale Assessments-Psychometric Properties of Test Scores and Associations with Test Taking Behavior. ***Frontiers in Psychology***, 7(154), 1-14. DOI: https://doi.org/10.3389/fpsyg.2016.00154

Pour-Mohammadi, M., & Abidin, M. (2011). Attitudes towards Teaching and Learning Test-taking Strategies for Reading Comprehension Tests: The Case of Iranian EFL Undergraduates. ***Journal of Studies in Education***, 1(1), 1-12. DOI: https://doi.org/10.5296/jse.v1i1.1028

Reise, S., & Flannery, W. (1996). Assessing Person-Fit on Measures of Typical Performance. ***Applied Measurement in Education***, 9(1), 9-26.

Robie, C.; Mueller, L., & Campion, J. (2001). Effects of a Motivational Inducement on the Psychometric Properties of a Cognitive Ability Test. ***Journal of Business and Psychology***, 16(2), 177-189. DOI: 10.1023/A:1011153015413

Rogers, B. (1984, April). ***The Effects of Instructions on the Psychometric Properties of a survey Instrument***. Paper Presented at the Annual Meeting of the American Educational Research Association, New Orleans.

Rogers, W.; T., & Harley, D. (1999). An Empirical Comparison of Three- and Four-Choice Items and Tests: Susceptibility to Testwiseness and Internal Consistency Reliability. ***Educational and Psychological Measurement***, 59(2), 234-247. DOI: https://doi.org/10.1177/00131649921969820

Rudner, L.; Bracey, G., & Skaggs, G. (1996). The Use of a Person-Fit Statistic with One High-Quality Achievement Test. ***Applied Measurement in Education***, 9(1), 91-109. DOI: http://dx.doi.org/10.1207/s15324818ame0901_8

Steiger, J. (1980). Tests for Comparing Elements of a Correlation Matrix. ***Psychological Bulletin***, 87(2), 245-251. DOI: http://dx.doi.org/10.1037/0033-2909.87.2.245

Stenlund, T.; Eklof, H., & Lyren, P. (2017). Group Differences in Test-Taking Behavior: An Example from a High-Stakes Testing Program. ***Assessment in Education***, 24(1), 4-20, DOI: http://dx.doi.org/10.1080/0969594X.2016.1142935.

Takallou, F.; Vahdany, F.; Araghi, S., & Tabrizi, A. (2015). The Effect of Test Taking Strategy Instruction on Iranian High School Students' Performance on English Section of the University Entrance Examination and Their Attitude towards Using These Strategies. ***International Journal of Applied Linguistics & English Literature***, 4(6), 119-129. DOI: http://dx.doi.org/10.7575/aiac.ijalel.v.4n.6p.119

Takallou, F.; Vahdany, F.; Araghi, S., & Tabrizi, A. (2016). Test-Taking Strategy Instruction in the National University Entrance Examination (NUEE) English Section. ***International Journal of Asian Social Science***, 6(3), 177-191. DOI: 10.18488/journal.1/2016.6.3/1.3.177.191

Volante, L. (2006). Toward Appropriate Preparation for Standardized Achievement Testing. ***The Journal of Educational Thought***, 40(2), 129-144.

Waskiewicz, R. (2011). Pharmacy Students' Test-Taking Motivation-Effort on a Low-Stakes Standardized Test. ***American Journal of Pharmaceutical Education***, 75 (3), 1-8. DOI: https://doi.org/10.5688/ajpe75341

Westfall, P., & D'Costa, A. (1987). Improving Prediction by Correcting Test Scores for Person Disturbances Using the Rasch Model. ***ERIC Number: ED319768***

Wise, V.; Wise, S., & Bhola, D. (2006). The Generalizability of Motivation Filtering in Improving Test Score Validity. ***Educational Assessment***, 11(1), 65–83.

Wise, S., & DeMars, C. (2010). Examinee Non-Effort and the Validity of Program Assessment Results. ***Educational Assessment***, 15, 27–41. DOI: http://dx.doi.org/10.1080/10627191003673216

Wise, S.; Ma, L.; Kingsbury, G., & Hauser, C. (2010, May). ***An Investigation of the Relationship between Time of Testing and Test-Taking Effort***. Paper presented at the annual meeting of the National Council on Measurement in Education, Denver, Colorado.

Wise, S., & Kingsbury, G. (2016). Modeling Student Test-Taking Motivation in the Context of an Adaptive Achievement Test. ***Journal of Educational Measurement***, 53(1), 86-105. DOI: 10.1111/jedm.12102

Wolf, L.; Smith, J., & Birnbaum, M. (1995). Consequence of Performance, Test Motivation, and Mentally Taxing Items. ***Applied Measurement in Education***, 8(4), 341-351. DOI: http://dx.doi.org/10.1207/s15324818ame0804_4

Xu, T., & Wu, Z. (2012). Test-taking strategies for a high-stakes writing test: An exploratory study of 12 Chinese EFL learners. ***Assessing Writing***, 17(3), 174–190. DOI: https://doi.org/10.1016/j.asw.2012.03.001

Zhang, W.; Liu, M.; Zhao, S., & Xie, Q. (2011). English Test-Taking Strategy Use and Students' Test Performance. ***Asian English as Foreign Language Journal***, 13(2), 133-168.