كلية التربية

كلية معتمدة من الهيئة القومية لضمان جودة التعليم

إدارة: البحوث والنشر العلمي (المجلة العلمية)

=======

واقع مؤشرات الدلالة الإحصائية وحجم الأثر في رسائل الماجستير المجازة بكلية التربية بالجامعة العالمية الإسلامية الماليزية خلال الفترة (١٩٩٣-٢٠١٧)

إعـــداد

إبراهيم بن سعيد بن حميد الوهيبي -

طالب دكتوراه بالجامعة الإسلامية العالمية الماليزية

أ.د/ داود بن عبدالملك الحدابي -

الجامعة الإسلامية العالمية الماليزية

د/حسين بن علي بن طالب الخروصي-

جامعة السلطان قابوس

﴿ المجلد الخامس والثلاثون – العدد الحادى عشر – نوفمبر ٢٠١٩م ﴾

http://www.aun.edu.eg/faculty_education/arabic

الملخص:

هدفت الدراسة إلى بحث واقع استخدام مؤشرات الدلالة الإحصائية وحجم الأثر في رسائل الماجستير المجازة بكلية التربية بالجامعة العالمية الإسلامية الماليزية خلال الفترة (١٩٩٣-٢٠١٧)، ولتحقيق أهداف الدراسة فقد تم اختيار عينة عشوائية منتظمة بلغت (١٠٢) رسالة ماجستير، بنسبة (٢٥.٨%) من مجتمع الدراسة والبالغ عدده (٣٩٦) رسالة ماجستير وقام الباحث بتصميم أداة لجمع البيانات الخاصة بالاختبارات الإحصائية التي تضمنتها هذه الدراسة، وهي اختبارات (ت) بأنواعها الثلاثة، واختبارات التباين بأنواعها الثلاثة (الأحادي Anova، والثنائي Univariate، والمتعدد Manova) ومعامل ارتباط بيرسون، ومعامل الانحدار، وخلصت الدراسة إلى جملة من النتائج أهمها: بلغت نسبة الاختبارات التي تم إيجاد حجم الأثر لها في رسائل الماجستير (١.٨١%) وهي نسبة ضئيلة جدا، حيث تم إيجاد حجم الأثر لـ(٢٠) اختبارا فقط من إجمالي عدد الإختبارات الإحصائية المستخدمة والتي بلغت (١١٠٨) اختبارا، كما أن متوسط حجم الأثر لاختبارات (ت)، واختبارات التباين فقد جاء منخفضا، وبالنسبة لمتوسط حجم الأثر لمعامل الانحدار ومعامل ارتباط بيرسون فقد كان عاليا، وأظهرت النتائج أن مستويات حجم الأثر لكل من اختبارات (ت) واختبارات تحليل التباين ومعامل ارتباط بيرسون اتفقت مع نتائج الدلالة الإحصائية، وكانت بينهما علاقة دالة إحصائيا، بعكس مستويات حجم الأثر لمعاملات الانحدار التي جاءت مخالفة لنتائج الدلالة الإحصائية، وبعلاقة غير دالة إحصائيا

يهدف البحث العلمي إلى تزويد الحقول المعرفية بالحقائق العلمية الجديدة في العلوم التطبيقية التي تخدم الواقع الحالي، أو تحدد التوجهات المستقبلية، وابتكار الحلول الممكنة والمناسبة للمشكلات البحثية والقضايا الحضارية المتجددة بطرق علمية، ومن أهم مراحل البحوث العلمية وخاصة البحوث الكمية هي مرحلة التحليل الإحصائي للبيانات، وتهدف هذه المرحلة إلى استخدام الأساليب الإحصائية المختلفة في تنظيم وترتيب وتبويب البيانات التي تم تجميعها من عينة البحث، وعرضها بشكل منظم باستخدام الجداول والرسوم أو المخططات البيانية المختلفة، ومن ثم معالجتها وتحليلها وتحديد جوانب التوازن فيها، وكذلك التعرف على المعلومات المتناقضة فيها، كالحدود العليا أو الحدود الدنيا في هذه البيانات، وتوضيح العلاقات والارتباطات، ومن أهم مراحل التحليل الإحصائي هو استخدام اختبارات الدلالة الإحصائية لإيجاد الفروقات بين المتوسطات، واستخلاص النتائج، للوصول إلى التفسيرات المنطقية المناسبة، وتقدير إمكانية تعميمها، وحيث أن اختبارات الدلالة الإحصائية تعاني من بعض القصور في تفسير نتائج الدراسات والبحوث؛ فقلد نادت العديد من الدراسات (حسن،٢٠٠٨؛ نصار،٢٠٠٦)، بأهمية استخدام أساليب إحصائية أخرى تساند الدلالة الإحصائية، ويمكن الوثوق بها في تفسير النتائج، ومن أهم هذه الأساليب الإحصائية ما يسمى بمؤشرات حجم الأثرEffect Size(حسن،٢٠٠٨)، ويُعرف حجم الأثر بمقدار التباين في المتغير التابع الذي يمكن تفسيره من خلال متغير مستقل يرتبط به أو يؤثر فيه (Howitt & Cramer,2008)، ونظرا لأهمية تضمين مؤشرات حجم الأثر في الأبحاث التربوية والنفسية، لحل إشكاليات مؤشرات الدلالة الإحصائية؛ فلقد دأبت الجمعية الأمريكية لعلم النفس American Psychological Association(APA) في دليلها للنشر العلمي منذ الطبعة الرابعة عام ١٩٩٤م(APA,1994) إلى الطبعة السادسة عام ٢٠١٠م (APA,2010)، على التوصية بضرورة تضمين مؤشرات حجم الأثر في الأبحاث التربوية والنفسية كمقاييس مكملة لاختبارات الدلالة الإحصائية، وكذلك توصي جمعية الأبحاث التربوية الأمريكية (AERA) في معايير الإبلاغ الخاصة بها بأن تتضمن النتائج الإحصائية بعضا من مقاييس حجم التأثير(Ellis,2010)، ولهذا أصبحت المجلات المتخصصة في العلوم التربوية وعلم النفس تشترط على الباحثين استخدام مؤشرات حجم الأثر، وعدم اكتفاء بالدلالة الإحصائية فقط في اختبار فرضيات أبحاثهم (Arnemann,2003).

لذلك قام العديد من الباحثين بتحليل العديد من الأبحاث المنشورة في المجلات العلمية المحكمة ودراسات(الماجستير والدكتوراه)،بهدف الوقوف على واقع استخدام الباحثون لمؤشرات حجم الأثر، ومن هذه الدراسات(الصائغ،١٩٩٦؛ الحداد،٢٠٠٦؛ محمود،٢٠٠٣؛إبراهيم،٢٠٠٠؛الجودة،٢٠٠٤؛البارقي،٢٠١٢؛بابطين،٢٠٠٢؛الشريفين،٢٠١٧؛ نصار،٢٠٠٦؛النجار،٢٠٠٥؛الشمراني،٢٠١٢؛ بخارى،٢٠١٦؛ سلامة،٢٠٠٤؛ باهي،٢٠١٠؛ الشمراني،٢٠٠٨؛عصر،٢٠٠٣؛حمزة،٢٠١٦؛محمد،٢٠١٣؛أبو جراد،٢٠١٣؛ الدرابيع،٢٠٠٣؛ عبدالفتاح،٢٠١١؛ النجار،٢٠٠٦؛ Gentry &Cohen,1962; Plucker,1997; Peters,2009; Levine,Asada, & CAPRARO&CAPRARO,2002; Hess,2003; Carpenter,2009; Fidalgo, Alavi, & Amirian,2014).

وتعتبر الأطروحات والرسائل العلمية إحدى المؤشرات الرئيسة للحكم على واقع البحث العلمي في أي بلد، وتمثل مصدر رئيسي للإسهام في التنمية البشرية (بخيت، والشيخ، ٢٠٠٤)، كما أن الأطروحات والرسائل العلمية تتميز بخصائص تجعلها من أبرز مصادر المعلومة في عالم الاتصال العلمي، من حيث توافر الابتكار فيها، وتخضع لمراحل تحكيم متعددة (الحلاق، ٢٠٠٨)، ويرى الذبياني (٢٠١٥) بأن "الرسائل والأطروحات العلمية في الكليات التربوية، من أنماط البحث التربوي الذي بدوره يسهم في إثراء المعرفة التربوية والتعليمية، ودراسة مشكلاتها، وتصويب مسارها" (ص١)، لذلك فإن مراجعة هذه الاطروحات والرسائل يشّكل جزءًا مهما من مراجعة البحث التربوي في أي بلد، للوقوف على خصائصها ومجالاتها البحثية، وتحليل وتقويم توجهاتها المنهجية، والأساليب الإحصائية المستخدمة فيها، لبناء تصور عن واقع هذه الاطروحات والرسائل، وتحديد جوانب القصور فيها، والحكم على مدى دقة نتائجها، وعلى الرغم من كثرة الدراسات التي أُجريت بهدف مراجعة البحوث التربوية من جوانب مختلفة، (الذبياني،٢٠١٥؛ الثبيتي،٢٠١٥؛ المعثم،١٤٢٩هـ)، إلا أن رسائل الماجستير المجازة من كلية التربية بالجامعة العالمية الإسلامية بماليزيا لم يحظى بمثل هذه الدراسات – على حد علم الباحث– لذا تسعى هذه الدراسة إلى تحليل واقع استخدام مؤشرات حجم الأثر في رسائل الماجستير المجازة من كلية التربية بالجامعة العالمية الإسلامية بماليزيا.

مشكلة الدراسة: تسعى هذه الدراسة إلى الإجابة على الأسئلة الآتية:

١- ما نسبة اقتران الدلالة الإحصائية بمؤشرات حجم الأثر في نتائج رسائل الماجستير المجازة بكلية التربية بالجامعة الإسلامية العالمية الماليزية؟
٢- ما مستوى مؤشرات حجم الأثر في رسائل الماجستير المجازة بكلية التربية بالجامعة الإسلامية العالمية الماليزية؟
٣- هل تختلف مؤشرات حجم الأثر باختلاف نتائج الدلالة الإحصائية في رسائل الماجستير المجازة بكلية التربية بالجامعة الإسلامية العالمية الماليزية؟
٤- هل توجد علاقة ذات دلالة إحصائية بين كلٍ من الدلالة الإحصائية ومؤشرات حجم الأثر في رسائل الماجستير المجازة بكلية التربية بالجامعة الإسلامية العالمية الماليزية؟

أهداف الدراسة: تسعى هذه الدراسة إلى تحقيق الأهداف الآتية:

١. تحديد نسب اقتران مؤشرات حجم الأثر بالدلالة الإحصائية في رسائل الماجستير المجازة بكلية التربية بالجامعة الإسلامية العالمية الماليزية.

٢. التعرف على واقع استخدام مؤشرات حجم الأثر في رسائل الماجستير المجازة بكلية التربية بالجامعة الإسلامية العالمية الماليزية.

٣. تحديد نسب الاختلاف بين نتائج الدلالة الإحصائية ومؤشرات حجم الأثر في رسائل الماجستير المجازة بكلية التربية بالجامعة الإسلامية العالمية الماليزية.

٤. معرفة العلاقة بين الدلالة الإحصائية ومؤشرات حجم الأثر في رسائل الماجستير المجازة بكلية التربية بالجامعة الإسلامية العالمية الماليزية.

أهمية الدراسة:

١- توضح هذه الدراسة أهمية معرفة مقدار تأثير المتغيرات المستقلة في المتغيرات التابعة، ومقدار العلاقة بين المتغيرات.

٢- قد تفيد هذه الدراسة في رفد الجامعة الإسلامية العالمية الماليزية بمعلومات كمية عن واقع استخدام مؤشرات حجم الأثر في الدراسات التربوية بكلية التربية، مما قد يساعد في توجيه برامج الدراسات العليا فيها نحو الأفضل.

٣- قد تفيد كلية التربية بالجامعة الإسلامية العالمية الماليزية في معرفة مستوى طلبتها في الدراسات العليا في استخدام الأساليب الإحصائية في دراساتهم، كما قد تسهم هده الدراسة في تقييم مقررات الإحصاء في برامج الدراسات العليا في ضوء أثر هذه المقررات في التفسير الصحيح لنتائج دراسات الماجستير التي يعدها طلبتهم، وقد يؤدي ذلك إلى تطوير هذه المقررات، من خلال تضمينها مؤشرات حساب حجم الأثر.

حدود الدراسة:

اقتصرت هذه الدراسة على رسائل الماجستير بكلية التربية بالجامعة الإسلامية العالمية الماليزية، والمتوفرة في مركز مصادر التعلم، بكلية التربية، والمجازة في الفترة (١٩٩٣م –٢٠١٧م)، والتي استخدمت تصاميم كمية واستخدمت الإحصاء الاستدلالي في الإجابة عن أسئلتها، وبالتحديد الاختبارات المعلمية الآتية (اختبارات ت T–Test، اختبارات تحليل التباين (الأحادي ، المتعدد) ، معامل ارتباط بيرسون، معامل الانحدار).

مصطلحات الدراسة:

الفرضية الإحصائية Statistical Hypothesis: يعرفها طعمة(٢٠١١) بأنها "إدعاء أو تصريح قد يكون (صحيحا أو خاطئاً) حول معلمة أو أكثر لمجتمع إحصائي أو عدد من المجتمعات الإحصائية" ص(١٩).

الدلالة الإحصائية Statistical Significance: يعرفها كارفر carver(1978:381) بأنها: الشي القليل أو المعدوم، وهي احتمال وجود علاقة بين اثنين أو أكثر من المتغيرات ناتجة عن شيء آخر غير الصدفة (Kenton,2019)، ويعرفها الباحث إجرائيا بنتيجة قيمة p المستخدمة للحكم على الاختبارات الإحصائية المستخدمة في رسائل الماجستير المجازة بكلية التربية بالجامعة الإسلامية العالمية الماليزية، وتصنف نتيجة الدلالة الإحصائية إلى مستويين: (دالة) إذا كانت قيمة p أقل من (٠.٠٥)، و (غير دالة) إذا كانت أكبر من (٠.٠٥).

حجم الأثر Effect Size: يعرفه أما بلانت Pallant(2010) بأنه مجموعة من المعادلات الرياضية المستخدمة لإيجاد الفروق بين المتوسطات الحسابية لمستويات المتغير المستقل، أو مقدار التباين الكلي الذي يحدثه المتغير المستقل في المتغير التابع، والتعريف الإجرائي لحجم الأثر هو القيمة أو القيم التي يتم الحصول عليها من خلال استخدام معادلات حجم الأثر لكل اختبار إحصائي من الاختبارات الإحصائية الموضحة في حدود هذه الدراسة.

الإطار النظري والدراسات السابقة:

نظرا لقصور اختبارات الدلالة الإحصائية في تفسير النتائج، وُجدت العديد من الأساليب الإحصائية التي تدعم الدلالة الإحصائية وتكون بمثابة الجانب المكمل لها، ومن بين تلك الأساليب الإحصائية ما يسمى بحجم الأثر أو بالدلالة العملية، التي كان له دور في تحرير الباحثين من ربقة التشبث بثنائية الإجابة على الأسئلة البحثية، حيث يمكنهم من الإجابة عن مقدار حجم الأثر الذي تحدثه المتغيرات المستقلة في المتغيرات التابعة، بدلا من الاقتصار على وجود الفروق أم لا، فعندما تكون الفرضية الصفرية صحيحة.

مفهوم حجم الأثر: ظهر مصطلح حجم الأثر عام ١٩٠١م على يد العالم كارل بيرسونKarl Pearson ، إلا أنه لم يكن واضح بصورة عملية إلا بعد أن عرفه رونالد فيشر Ronald Fisher عام ١٩٢٥ حيث اقترح أن الباحثين يكملون اختبار الدلالة الإحصائية في تحليل التباين مع نسبة الارتباط أو eta ، والتي تقيس قوة الارتباط بين المتغيرات المستقلة والتابعة (Kirk,1996) ولا يرتبط حجم الأثر بالدلالة الإحصائية حيث أن النتائج التي لها دلالة إحصائية ليست بالضرورة أن يكون لها حجم أثر معتبر(نصار، ٢٠١٧)، ويهدف حجم الأثر إلى توضيح الأهمية العملية والفعلية للنتائج بغض النظر عن كونها دالة إحصائيا أم لا، وذلك من خلال حساب حجم الفرق بين مجموعتي الدراسة، وقياس الفعالية، لذلك يُعرّف حجم الأثر بقوة العلاقة بين المتغير/المتغيرات المستقلة والمتغير/المتغيرات التابعة، أو حجم الفروق بين مستويات المتغيرات المستقلة في المتغير التابع (George, Robert, & Jeffrey,2006).

وعلى الرغم من أن مؤشرات حجم الأثر ظهرت منذ أربعينيات القرن الماضي(Huberty, 2002) إلا أنه لا يزال استخدامها يُعد محدودا، في مختلف المجالات، فيندر التطرق إلى شرح هذه المؤشرات في معظم الكتب الإحصائية، إلا في الكتب الإحصائية المتخصصة في التحليل البعدي Meta analysis، وكذلك البرامج الإحصائية الحاسوبية لا تقوم بحساب أغلب هذه المؤشرات(Coe,2002)، لذلك أكدت جمعية علم النفس الأمريكية (APA) في إصداراتها منذ إصدارها الرابع (١٩٩٤) على ضرورة تضمين مؤشرات حجم الأثر في الدراسات والأبحاث، وعلى نحو مماثل أوصت جمعية الأبحاث التربوية الأمريكية (AERA) في معايير الإبلاغ الخاصة بها، بضرورة أن يكون الإبلاغ عن النتائج الإحصائية مصحوبًا

بحجم التأثير و"التفسير الكيفي للتأثير" (AERA 2006: P.5). وتتابعت بعد ذلك العديد من المجلات العلمية الأجنبية –بلغت عدد (23) – في اشتراط تضمين مؤشرات حجم الأثر في الأبحاث المقدمة لها، (Vacha-Haase, & Thompson, 2004; Pedersen,2003)، ولم يتوصل الباحث إلى عدد المجلات العلمية العربية التي اتبعت سياسة ضرورة الإبلاغ عن مؤشرات حجم الأثر في البحوث المقدمة لها.

- **مؤشرات حجم الأثر:** توجد مؤشرات عديدة لحساب حجم الأثر أو أحجام التأثير، قد ذكرا (Vacha-Haase, & Thompson, 2004,p:473)، بأنها تصل إلى (٤٠) مقياسا، بل أوصلها (Kirk, 2003) إلى (٧٠) مقياسا، واشتملت على مسميات مختلفة كنسب الأرجحية، والمخاطر النسبية،...إلخ، وحيث أن معظم مؤشرات حجم الأثر تهتم بإجراء مقارنات أو قياس العلاقات، فإنه يمكن تصنيف معظمها إلى مؤشرات حجم الأثر للاختبارات الإحصائية المعلمية، مؤشرات حجم الأثر للاختبارات الإحصائية اللا معلمية، مؤشرات حجم الأثر للبحوث الطبية

Vacha-Haase, & Thompson, 2004; Kirk, 2003; Ellis,2010;Ialongo,2016) Lakens,2013 ; العبيدي، ٢٠١٩)، وحيث أن هذه الدراسة سوف تقتصر بدراسة مؤشرات حجم الأثر للاختبارات الإحصائية المعلمية فقط (اختبارات "ت" –تحليل التباين–معامل ارتباط بيرسون–تحليل الانحدار البسيط)، فإن الباحث سوف يقتصر بتوضيح مؤشرات حجم الأثر لهذه الاختبارات فقط.

مؤشرات حجم الأثر للاختبارات الإحصائية المعلمية: يتم تصنيف هذه المؤشرات إلى مقاييس تهتم بقياس الفروق بين المجموعات، وأخرى تبحث عن قوة العلاقات بين المتغيرات، وفيما يلي توضيح لهذه المقاييس.

مؤشرات قياس الاختلافات بين المجموعات: تستخدم هذه المؤشرات لقياس أثر المعالجات على متغيرات ثنائية، كأثر طريقة تدريس على التحصيل الدراسي مقارنة بطريقة التدريس التقليدية، أو أثر برنامج إرشادي في تخفيض قلق الامتحانات، أو مقارنة بين نوعين من الأدوية في علاج مرض معين، ولقد اقترح كوهينCohen عام ١٩٦٢م معادلة لحساب الفرق بين مجموعتين، بأن نقوم بطرح متوسط المجموعة الأولى من المجموعة الثانية (M1 – M2) وتقسيم النتيجة على الانحراف المعياري (SD) الخاص بمجتمع الدراسة الذي تم أخذ العينات منه، كما في المعادلة (١) (Cohen,1988;Howitt,Cramer,2008; Ellis,2010):

$$\mathbf{d} = \frac{\mathbf{m_1} - \mathbf{m_2}}{\sigma} \quad \textbf{المعادلة(١)}$$

حيث m_1: متوسط المجموعة الأولى، m_2: متوسط المجموعة الثانية، ð: الانحراف المعياري لمجتمع الدراسة.

مجموعة مقاييس نسبة التباين المفسر (عائلة r):

ذكر سنايدر ولاوسن (Syyder&Lawson,1993 **المشار إليهما في نصار** ، ٢٠٠٦) عددا من المؤشرات التي تعتمد على نسبة التباين في المتغير التابع الذي يفسره المتغير أو المتيرات المستقلة، ومن هذه المؤشرات مؤشر مربع ايتا (η^2) الذي يعتمد أقدم مؤشر واقترحه العالم فيشر (Fisher) عام ١٩٣٢، وبعده بحوالي (٣) سنوات أي عام ١٩٣٥ قدم كيلي (Kelly) مؤشر مربع ايبسون epsilon squared (ε^2)، كما قدم كوهين (Cohen) عام ١٩٧٣ مؤشر مربع ايتا الجزئي (η_P^2)، وفي عام ١٩٦٣ اقترح هايز (Hays) مؤشر مربع أوميجا omega squared (ω^2)، وغيرها من المؤشرات، وحيث أن مؤشرات نسبة التباين المفسر كثيرة ويطول البحث في حصرها وشرح تفاصيلها، لذا فإن الباحث يقتصر على المؤشرات التي تخدم أهداف هذه الدراسة فقط، كما يلي:

أ) مؤشر مربع ايتا(η^2) eta^2 : نظرا لكون مؤشرات – كوهين د ، ودلتا Glass و Hedges's g – لها معاييرها الخاصة بمدى تجانس مجتمع الدراسة، وأحجام العينات؛ فإن الكثير من الباحثين (المالكي، ٢٠١٨، البارقي، ٢٠١٢، الصائغ، ١٩٩٦) يفضلون استخدام نسبة الارتباط (مربع ايتا) (η^2)eta^2 لإيجاد حجم الأثر لاختبارات (ت) سواء كان اختبار (ت) لعينتين مستقلتين أو اختبار (ت) لعينتين مترابطتين، وذلك لسهولة استخدامه، واعتباره المؤشر الأكثر أهمية في البحوث التربوية، والأكثر صدقا في التنبؤ (الصائغ، ١٩٩٦)، وعدم اعتماده على الانحرافات المعيارية لمجموعتي الدراسة أو أحجام العينات، حيث يقتصر فقط على استخدام قيمة اختبار (ت) كما في المعادلة (٢) الآتية (Tomczak & Tomczak,2014؛ المالكي، ٢٠١٨):

$$\eta^2 = \frac{t^2}{(t^2 + df)} \qquad \text{المعادلة(2)}$$

حيث: df: درجات الحرية لاختبار (ت) في حالة عينتين مستقلتين (ن١+ن٢ –٢)، وفي حالة اختبار (ت) لعينتين مترابطتين (عدد أفراد العينة – ١).

وتتراح قيمة نسبة الارتباط (مربع ايتا) (η^2)eta^2 بين (٠–١)، وتفسر بنسبة التباين بين متوسطي مجموعتي الدراسة، وتدل القيمة (٠) على انعدام الفرق بين المتوسطين، أما القيمة (١) فتدل على الاختلاف التام بين متوسطي مجموعتي الدراسة (المالكي، ٢٠١٨: ٦٥):

ويُشير (المالكي ٢٠١٨) إلى أنه يمكن تحويل قيمة(η^2) إلى مؤشر كوهين (d) باستخدام المعادلة (٣):

$$d = \frac{2\sqrt{\eta 2}}{\sqrt{1-\eta 2}} \qquad \text{المعادلة(3)}$$

ونظرا لكون مربع ايتا (η^2)eta^2 يحدد نسبة التباين في المتغير التابع الذي يسببه المتغير المستقل، فإنه يتم استخدامه أيضا في إيجاد حجم الأثر لتحليل التباين (ANOVA)، بأنواعه المختلفة الأحادي (One way ANOVA)، أو في اتجاهين (Two way ANOVA)، أو في ثلاثة اتجاهات، ففي حالة تحليل التباين الأحادي فإن مربع إيتا (η^2) يساوي حاصل قسمة مجموع مربعات التباين الكلي على مجموع مربعات التباين بين المجموعات، كما في المعادلة (٤) (Tomczak & Tomczak,2014، العبيدي، ٢٠١٩):

$$\eta^2 = \frac{SS_b}{SS_t} \qquad \text{المعادلة}(4)$$

حيث: SS_b: مجموع التباين بين المجموعات (تباين المتغير المستقل)، SS_t: مجموع التباين الكلي

كما يمكن إيجاد قيمة حجم الأثر لتحليل التباين المتعدد (Manova) وذلك باستخدام المعادلة (ward,2002):

$$\eta^2 = 1 - \text{Wilks } Lambda \qquad \text{المعادلة}(5)$$

وحيث أن مربع ايتا (η^2) من المفاهيم الإحصائية التي شاع استخدامها في الدراسات والبحوث الإنسانية والإجتماعية المختلفة؛ فلقد لحقه ما لحق الدلالة الإحصائية من سوء فهم وتفسير خاطئ، فلقد فسره بعض الباحثين مثل تفسيرهم لمعامل الانحدار المتعدد (R^2) كمقياس لتحديد التباين المفسر عن طريق المعالجة الإحصائية، فعلى الرغم من التشابه في طريقة إيجاد كل منهما إلا أن ذلك لا يعني أنهما لهما نفس الغرض الإحصائي، فمربع ايتا (η^2) يعتمد على نموذج تحليل التباين، و(R^2) يعتمد على معامل الانحدار المتعدد، كما أن بعض الباحثين يفضل استخدام مؤشر مربع ايتا (η^2) بحجة أنه لا يعتمد على حجم العينة، إلا أنهم قد وقعوا في مخالطة إحصائية حيث أنهم لم يدركوا بأن اختبار (ف) يعتمد على حجم العينة، وكذلك يمثل حجم العينة أحد مكونات حساب قيم (SS_b) مجموع التباين بين المجموعات (Murray,Dosser,1987، المشار إليهما في كل من: الحداد، ٢٠٠٦، و إبراهيم، ٢٠٠٠)، ولإيجاد حجم الأثر من خلال مؤشر مربع ايتا (η^2) لتحليل التباين؛ فلقد أوضح كوهين cohen(1988) معادلات تحويل نسبة التباين (η^2) إلى حجم أثر (f) والعكس، كما يلي:

$$\eta^2 = \frac{f^2}{1+f^2} \qquad \text{المعادلة}(6)$$

$$f = \sqrt{\frac{\eta^2}{1-\eta^2}} \qquad \text{المعادلة}(7)$$

ب) معامل الارتباط: يعد معامل الارتباط هو أهم مقياس يستخدم في مجموعة حجم الأثر لقياس نسبة التباين المفسر أو قوة العلاقة بين المتغيرات سواء كانت متغيرات ثنائية أو مستمرة، فيمكن قياس حجم تأثير أي علاقة من خلال حساب معامل الارتباط بأنواعه المختلفة حسب طبيعة المتغيرات، مثل معامل ارتباط بيرسون (r)، معامل ارتباط سبيرمان (rs)، معامل ارتباط كندال توو Kendall's Tau (τ)، معامل ارتباط بوينت بيرسال Point–biserial (rpb)، معامل ارتباط فاي phi (φ)، ومعامل ارتباط جوودمان لامدا Goodman and Kruskal's lambda (λ)، وغيرها من معاملات الارتباط المستخدمة في قياس العلاقة بين المتغيرات(Ellis,2010).

ويرى بعض الباحثين (Cohen ,1988) بأن حجم الأثر لمعاملات الارتباط يتم حسابه من خلال معامل التحديد والذي يساوي تربيع قيمة معامل الارتباط (R^2)، ويمثل نسبة التباين الذي يُحدثه المتغير المستقل في المتغير التابع.

جـ) تحليل الانحدار: معامل التحديد (R^2) الناتج من تحليل الانحدار يعتبر مؤشر غير مصحح أو متخيز للأعلى(Ellis,2010)، ويوضح كوهين (cohen(1988: 412 معادلة حساب حجم الأثر لتحليل الانحدار كالآتي:

$$F^2 = \frac{R^2}{1 - R^2} \qquad \text{المعادلة(8)}$$

وبناء على معيار كوهين (cohen(1988، فإن مؤشرات حجم الأثر للإختبارات الإحصائية التي تضمنتها هذه الدراسة يتم تفسيرها وفق الجدول (١) الآتي:

الجدول (١): مستويات حجم الأثر للاحتبارات الإحصائية

م	الاختبار الاحصائي	المؤشر	مستوى حجم الأثر		
			منخفض	متوسط	عالي
١	اختبارات (ت)	d	$0 \le d < 0.50$	$0.5 \le d < 0.80$	$d \ge$
٤	اختبارات تحليل التباين	f	$0 \le f < 0.25$	$0.25 \le f < 0.40$	$f \ge 0.40$
٦	معامل الارتباط	R^2	$0 \le R^2 < 0.09$	$0.09 \le R^2 < 0.25$	$R^2 \ge$
٧	تحليل الانحدار	F	$0 \le F^2 < 0.15$	$0.15 \le F^2 < 0.35$	$F^2 \ge$

الدراسات السابقة:

اطلع الباحث مجموعة من الدراسات التي تبحث في العلاقة بين الدلالة الإحصائية، وحجم الأثر، وفيما يلي عرضا لبعض هذه الدراسات مرتبة ترتيبا تنازليا حسب سنة النشر:

أجرى المالكي(٢٠١٨) دراسة تطبيقية على بحوث المجلات التربوية بالجامعات السعودية بعنوان "واقع الدلالة الإحصائية والدلالة العملية وقوة الاختبار الإحصائي، خلال الفترة من ١٤٣٦هـ – ١٤٣٨هـ"، وهدفت إلى التعرف على الفروق بين المجلات في استخدامها لمؤشرات الدلالة العملية والقوة الإحصائية، في الاختبارات الإحصائية الآتية: اختبارات ت بأنواعها الثلاثة، واختبار تحليل التباين الأحادي فقط، كما هدفت الدراسة إلى الكشف عن أثر حجم العينة ونوع المنهج المستخدم في الدلالة الإحصائية وكل من الدلالة العملية وقوة الاختبار الإحصائي، وتحديد نسبة الاتفاق بين هذه المؤشرات الإحصائية، وتكونت عينة الدراسة من (١٣٣) بحثا، اشملت على (١٣٥٥) اختبارا إحصائيا، وخلصت الدراسة إلى جملة من النتائج أهمها ضعف استخدام مؤشرات الدلالة العملية والقوة الإحصائية في البحوث التربوية، وتدني الدلالة العملية للاختبارات الإحصائية المستخدمة، ووجود تأثير إيجابي لحجم العينة له على نتائج الدلالة الإحصائية، وسلبي على مستويات الدلالة العملية، وأوصت الدراسة بأهمية استخدام مؤشرات الدلالة العملية كمؤشرات مكملة للدلالة الإحصائية، في البحوث والدراسات، للحكم على أهمية نتائجها.

وهدفت دراسة القضاة(٢٠١٦) إلى بحث درجة التوافق بين نتيجة الدلالة الإحصائية والدلالة العملية في دراسات الماجستير المنجزة بكلية التربية بجامعة مؤتة لتحديد نسبة استخدام مؤشرات الدلالة العملية كمقاييس مكملة للدلالة الإحصائية، وتكونت عينة الدراسة من (١٢٠) رسالة ماجستير، واشتملت على (١٤٣٣) اختبار إحصائي من الاختبارات الإحصائية التي شملتها الدراسة وهي اختبار (ت) للعينات المستقلة، واختبار (ف) تحليل التباين، وتوصلت النتائج إلى أن نسبة ضئيلة جدا (١.٧%) من الاختبارات الإحصائية التي تم حساب الدلالة العملية لها إلى جانب الدلالة الإحصائية، وتم حساب الدلالة العملية لجميع الاختبارات الإحصائية وتوصلت النتائج إلى أن أكثر من ثلاثة أرباعها كان لها دلالة عملية صغيرة، و(٢٤.١%) فقط كانت دلالتها العملية متوسطة وكبيرة، وأن (٧٥.٤%) من نتائج الدلالة الإحصائية توافقت مع نتائج الدلالة العملية، حيث أن (٥٣.٢%) من الاختبارات الغير دالة إحصائيا كانت دلالتها العملية صغيرة، وأن (٢٢.٢%) من الاختبارات الدالة إحصائيا كانت دلالتها العملية متوسطة وكبيرة.

وقام البارقي(٢٠١٢) بدراسة هدفت إلى بحث واقع استخدام الدلالة العملية مقارنة بالدلالة الإحصائية في مجلة جامعة أم القرى للاختبارات الإحصائية الآتية: اختبارات (ت)، واختبارات تحليل التباين (ف)، ومعامل ارتباط بيرسون وتحليل الانحدار، واختبار مربع كاي للاستقلالية، وتكونت عينة الدراسة من (٦١) بحثا كميا في مجلة جامعة أم القرى للعلوم التربوية في الفترة المحددة، وشملت هذه البحوث (٢٢٢٣) اختبارا إحصائيا، وتوصلت النتائج إلى أن أكثر من ثلثي الاختبارات الإحصائية المستخدمة كانت لها دلالة إحصائية عند مستوى (٠.٠٥) على الأقل، إلا أن نسبة الاختبارات التي تم حساب الدلالة العملية لها بلغت (٠.٧٦%) وهي نسبة متدنية جدا، كما أن أكثر من نصف الاختبارات الإحصائية كانت لها دلالة عملية متدنية، وأوصت الدراسة بإعطاء الدلالة العملية مزيدا من الاهتمام لا سيما في البحوث العربية.

وهدفت دراسة الحداد(٢٠٠٦) إلى معرفة واقع الدلالة الإحصائية والدلالة العملية وقوة الاختبارات الإحصائية المستخدمة في بحوث مجلة الهيئة القومية للبحث العلمي في الجماهيرية العظمى، خلال الفترة (١٩٩٧-٢٠٠٣)، وتكونت عينة الدراسة من (٤٠) بحثا كميا، تضمنت (٦٢٣) اختبارا إحصائيا من الاختبارات الإحصائية التي شملتها الدراسة وهي اختبار تحليل التباين الأحادي واختبار ت لعينتين مستقلتين ومعامل ارتباط بيرسون وتحليل الانحدار، واتفقت النتائج مع نتائج المالكي(٢٠١٨) والبارقي(٢٠١٢) في تدني نسبة استخدام الدلالة العملية كمؤشر مكمل للدلالة الإحصائية، حيث بلغت النسبة في هذه الدراسة (٣.٩%) فقط، وكذلك تدني الدلالة العملية لـ(٦١.٣%) من الفرضيات، وبالنسبة لقوة الاختبار فقد أظهرت النتائج أن (٥٨.٧%) من إجمالي الفرضيات تمتعت بقوة اختبار كبيرة تراوحت بين (٠.٦- فما فوق).

وقام الجودة(٢٠٠٤) بإجراء دراسة بعنوان الدلالة الإحصائية والدلالة العملية وقوة الاختبار الإحصائي للأبحاث المنشورة في مجلة أبحاث اليرموك سلسلة العلوم الإنسانية والاجتماعية لمدة (٢٦) سنة خلال الفترة ١٩٨٥-٢٠٠١، وشملت العينة (٥٤) بحثا، بإجمالي (٨٨٩) اختبارا إحصائيا من الاختبارات المحددة في دراسته وهي اختبارات ت واختبارات (ف) تحليل التباين واختبار z، واختبار مربع كاي للاستقلالية، وخلصت الدراسة إلى جملة من النتائج من أهمها أن ما يقارب من نصف الفرضيات (٤٨.٦%)، استخدمت مستوى دلالة ٠.٠٥، وأن البحوث التي استخدمت الدلالة العملية قليلة جدا حيث بلغت النسبة (٨.١%) فقط، كما ان ما يقارب من ثلثي الفرضيات الدالة إحصائيا (٦٢.٢%) ذات دلالة عملية ضعيفة، في حين كانت أن (٦٠.٤%) منها تمتعت بقوة اختبار عالية.

وفي الأردن قام محمود(٢٠٠٣) بنفس دراسة إبراهيم (٢٠٠٠) لكن على مجتمع مختلف، حيث كان مجتمع محمود (٢٠٠٣) عبارة عن رسائل الماجستير المجازة بكلية العلوم التربوية بجامعة مؤته، خلال الفترة من ١٩٩٥-٢٠٠٠، وشملت العينة (٥٥) رسالة ماجستير اتبعت الإحصاء التحليلي في فحص فرضياتها، وبلغت عدد الاختبارات الإحصائية المستخدمة (١١٦٧) اختبارا، وأشارت النتائج إلى أن (٤٥.٣%) من الاختبارات كانت ذات دلالة إحصائية، إلا أن الدلالة العملية لأغلبية الاختبارات (٨٣.٨%) كانت في المستوى المتدني، وبالنسبة لقوة الاختبار فأوضحت النتائج أن(٤٨.٢%)من الاختبارات حظيت بقوة إحصائية كبيرة.

كما قام الدرابيع(٢٠٠٣) بدراسة هدفة إلى التعرف على واقع الدلالات الإحصائية والعملية وقوة الاختبارات الإحصائية المستخدمة في بحوث مجلة مؤتة للبحوث والدراسات سلسلة العلوم الإنسانية والاجتماعية، خلال الفترة من ١٩٩٦-٢٠٠٠، وتكونت عينة الدراسة من (٤٨) بحثا، اشتملت على (٣٠١) اختبارا إحصائيا لنفس الاختبارات الإحصائية التي فحصها إبراهيم(٢٠٠٠) في دراسته، ولم تختلف نتائج البحوث عن نتائج الرسائل العلمية

(الدكتوراه والماجستير) حيث أوضحت نتائج الدرابيع(٢٠٠٣) عدم اهتمام الباحثين باستخدام مؤشرات الدلالة العملية كمؤشرات مكملة للدلالة الإحصائية في بحوثهم، كما أن الدلالة العملية لجميع الاختبارات الإحصائية تراوحت بين مستوى المتوسط والضعيف، وعدم وجود استقلالية للدلالة الإحصائية عن كل من الدلالة العملية وقوة الاختبار لأغلب الاختبارات الإحصائية.

وبهدف دراسة واقع الدلالة الإحصائية والدلالة العملية وقوة الاختبار للاختبارات الإحصائية المستخدمة قام إبراهيم (٢٠٠٠) بإجراء دراسته على بحوث مجلة دراسات العلوم التربوية بالجامعة الأردنية في الأعوام ١٩٩٧–١٩٩٨، وتكونت عينته من (٤٨) بحثا التي استخدمت تصاميم كمية، شملت (٦٢٣) اختبارا إحصائيا، من نوع (r,t,f_1,f_2)، وتوصلت النتائج إلى أن (٥٢.٧%) من الفرضيات ذات الدلالة الإحصائية كانت دلالتها العملية ضعيفة، وبلغ معدل القوة لجميع الاخبارات الإحصائية (٠.٥٩)، وتمتعت (٩٠.٥%) من الفرضيات الدالة إحصائيا بقوة اختبار كبيرة، كما أشارت النتائج إلى أن الباحثين لم يتحيزوا لنشر النتائج ذات الدلالة الإحصائية فقط، حيث أن كلا النتيجتين ذات أهمية على المستوى العلمي، وكعادة جميع الدراسات السابقة أوضحت نتائج إبراهيم(٢٠٠٠) إلى ضعف استخدام مؤشرات الدلالة العملية في فحص الاختبارات الإحصائية، واكتفاء الباحثون بالدلالة الإحصائية فقط.

ومن خلال استعراض هذه الدراسات تبين بأن كل الدراسات السابقة استخدمت معيار كوهين لتفسير الدلالة العملية: صغير إذا كانت أقل من (٠.٥)، وكبير إذا بلغ (٠.٨)، ومتوسط بينهما، كما اتفقت كل الدراسات على قلة استخدام حجم الأثر كمؤشر مكمل للدلالة الإحصائية، وكذلك على تدني مستوى الدلالة العملية للفرضيات والاختبارات الإحصائية.

منهجية الدراسة:

تتبع هذه الدراسة المنهج الوصفي التحليلي وذلك لكونه يمثل أداة بحثية قيمة تسهم في مجموعة واسعة من الدراسات، سواء كانت الدراسات الوصفية أو السببية، حيث يعتمد على دراسة الظواهر المعنية بشكل كامل كما هي موجودة في الواقع، وتحديد السمات الأكثر ظهورا فيها ووصفها وصفا دقيقا، وتحديد التدابير ذات الصلة التي تمثل هذه السمات (Loeb & et al.,2017)، وتحليلها للوصول إلى التفسير الأنسب لها.

وباستخدام المنهج الوصفي التحليلي يستطيع الباحث تحليل النتائج التي توصلت إليها رسائل الماجستير المجازة بكلية التربية بالجامعة الإسلامية العالمية الماليزية، من خلال دراسة وتحليل الاختبارات الإحصائية المستخدمة في هذه الرسائل، وقيم الدلالة الإحصائية، واستخدامها في إيجاد حجم الأثر لها، والتوصل إلى العلاقات التي تربط بين الدلالات الإحصائية وأحجام الأثر.

مجتمع الدراسة وعينتها:

يتكون مجتمع الدراسة من جميع رسائل الماجستير المجازة بكلية التربية بالجامعة الإسلامية العالمية الماليزية، منذ عام ١٩٩٣م وحتى عام ٢٠١٧م، والمتوفرة في مركز مصادر التعلم بالكلية، والتي تمت كتابتها باللغتين العربية والإنجليزية، وقد اقتصر الباحث على رسائل الماجستير الموجودة في مركز مصادر التعلم بالكلية وذلك نظرا لسهولة الوصول إليها لكونها موجودة في مكان محدد ومرقمة ترقيما عدديا، بعكس تلك الموجودة في المكتبة الرئيسة –دار الحكمة– بالجامعة، كما أن مركز مصادر التعلم يحتوي على جميع رسائل الماجستير التي تمت إجازتها خلال الفترة المحددة سلفا –على حد قول موظف المركز–، وبلغ مجتمع الدراسة (٣٩٦) رسالة ماجستير، أما عينة الدراسة فمن خلال جدول تحديد حجم العينة للعالمان Krejcie & Morgan(1970) ، فإن أقل عينة مناسبة لهذه الدراسة تبلغ (١٩٦) رسالة، لذلك أخذ الباحث (٥٠%) من المجتمع لتمثل عينة الدراسة، وتم استخدام طريقة المعاينة العشوائية المنتظمة، حيث تم الاطلاع على جميع الرسائل التي تحمل الرقم الفردي، وبلغت عينة الدراسة (١٩٨) رسالة ماجستير، وبعد الإطلاع على جميع رسائل الماجستير التي مثلت عينة الدراسة، تم استبعاد (٩٦) رسالة نظرا لكونها اتبعت المنهج النوعي أو أنها استخدمت الإحصاء الوصفي فقط في تحليل بياناتها –وهذا السبب الأكثر تكرارا في الرسائل المستبعدة–، أو أنها استخدمت إحصاء إستدلالي خارج حدود هذه الدراسة، لذا فإن العينة الفعلية لهذه الدراسة والتي اتبعت المنهج الكمي في تحليل بياناتها بلغت (١٠١) رسالة ماجستير، وتمثل (٢٥.٥١%) من مجتمع الدراسة، وهي نسبة مقبولة تربويا.

أداة الدراسة: لتحقيق أهداف الدراسة؛ قام الباحث بالاطلاع على مجموعة من الدراسات السابقة للاستفادة منها في صياغة أداة الدراسة، والتي شملت على بيانات خاصة لكل اختبار من الاختبارات الإحصائية، والتي يتم استخدامها في المعادلات التي اعتمدها الباحث في حساب حجم الأثر للاختبارات الإحصائية المحددة في هذه الدراسة.

المعالجة الإحصائية:

للإجابة على أسئلة الدراسة؛ تم استخدام الأساليب الإحصائية الآتية من خلال برنامجي EXCEL و SPSS:

- للإجابة على السؤالين الأول والثاني من الدراسة تم استخدام الإحصاءات الوصفية: التكرارات والنسب المئوية، المتوسطات الحسابية، الانحرافات المعيارية.
- للإجابة على السؤال الثالث تم استخدام مؤشرات (F^2 ، f ، d, μ^2)؛ لحساب حجم الأثر للاختبارات الاحصائية المحددة في الدراسة.
- للإجابة على السؤال الرابع تم استخدام اختبار كاي تربيع χ^2 للاستقلالية، ومؤشر (W) لكوهن؛ لحساب حجم الأثر لاختبارات كاي تربيع χ^2، وكذلك معامل ارتباط بوينت بايسيريال Point Bi-serial Coefficient؛ لإيجاد قوة العلاقة بين متغيرين اسمي و رتبي.

نتائج الدراسة: السؤال الأول: ما نسبة اقتران الدلالة الإحصائية بمؤشرات حجم الأثر في نتائج رسائل الماجستير المجازة بكلية التربية بالجامعة الإسلامية العالمية الماليزية؟

للإجابة على هذا السؤال فقد تم عد جميع الاختبارات الإحصائية التي تضمنتها هذه الدراسة، وتحديد أي من هذه الاختبارات تم حساب حجم الأثر لها، والجدول (٢) يوضح هذه النتائج.

الجدول(٢) أعداد الاختبارات الإحصائية التي تم حساب حجم الأثر لها

الاختبار الإحصائي	العدد	اقتصرت على الدلالة الإحصائية فقط		استخدمت حجم أثر	
		العدد	النسبة	العدد	النسبة
إجمالي اختبارات ت	431	422	97.91	9	2.09
إجمالي اختبارات تحليل	225	216	96.00	9	4.00
معامل ارتباط بيرسون	397	395	99.50	2	0.50
معامل الانحدار	55	55	100.00	0	0.00
إجمالي الاختبارات	1108	1088	98.19	20	1.81

يوضح الجدول(٢) جملة من النتائج أهمها:

- بلغت عدد الاختبارات الإحصائية المستخدمة في عينة الدراسة (١١٠٨) اختبار، حيث كانت أكثر الاختبارات المستخدمة هي اختبارات (ت) بأنواعها الثلاثة والتي بلغت (٤٣١) بنسبة (٣٨.٩%) من إجمالي الاختبارات، وجاء معامل ارتباط بيرسون في المرتبة الثانية من حيث الاستخدام حيث تم استخدامه (٣٩٧) مرة بنسبة (٣٥.٨%)، أما معامل الانحدار فلم يتستخدمه طلبة الماجستير إلا قليلا، حيث تم استخدامه (٥٥) مرة فقط، بنسبة لم تتجاوز الـ(٥%)، كما أن تحليل التباين الثنائي والثلاثي لم يُستخدم البتة، مما يُشير إلى أن طلبة الماجستير يميلون إلى استخدام الاختبارات الإحصائية الإستدلالية البسيطة.
- بلغت نسبة الاختبارات الإحصائية المستخدمة التي تم إيجاد حجم الأثر لها(١.٨١%) فقط، وهي نسبة ضئيلة جدا، حيث لم يستخدم طلبة الماجستير أي مؤشر من مؤشرات حجم الأثر لاختبارات (ت) لعينة واحدة ولعينتين مترابطتين، وكذلك لمعامل الانحدار، أما إجمالي اختبارات (ت) فنسبة (٢.٠٩%) فقط منها استخدمت مؤشرات حجم الأثر، وكان مصدر هذه النسبة من استخدام حجم الأثر لاختبار (ت) لعينتين مستقلتين فقط، وعلى الرغم من سهولة إيجاد حجم الأثر لمعامل ارتباط بيرسون حيث أنه مجرد تربيع قيمة معامل الارتباط، إلا أنه لم يتم الإبلاغ عنه إلا مرتين فقط وبنسبة (٠.٥%) فقط، أما إجمالي اختبارات تحليل التباين فبلغت نسبة الإبلاغ عن حجم الأثر لها (٤%) فقط.

- مما يشير إلى غياب شبة تام لمؤشرات حجم الأثر في رسائل الماجستير، وهذا مؤشر خطير جدا، حيث أن الاقتصار على الدلالة الإحصائية في وصف النتائج لا يعطي قيمة حقيقية لأهمية هذه النتائج وجودتها، كما أن استخدام الدلالة الإحصائية فقط وعدم إسنادها بمؤشرات أخرى كمؤشرات حجم الأثر قد يُعيق تقدم العلوم وتطورها، حيث أنها قد تؤدي إلى الاهتمام ببرامج معينة أو التركيز على حلول دون أخرى نظرا لكون نتائجها دالة إحصائيا، على الرغم من أنها ليست ذات قيمة عملية وليس لها تأثير في الواقع، وإنما جاءت دلالتها بسبب حجم العينة التي أُستخدمت لفحص تلك البرامج أو الحلول، وفي المقابل قد يتم إهمال برامج أخرى ذات جودة وقيمة عملية وذات أهمية بسبب عدم دلالتها الإحصائية.
- تتفق هذه النتيجة مع نتائج مجموعة كبيرة من الدراسات (المالكي،٢٠١٨؛ القضاة،٢٠١٦؛ البارقي، ٢٠١٢؛ الحداد، ٢٠٠٦؛ جودة، ٢٠٠٤؛ محمود، ٢٠٠٣؛ الدرابيع، ٢٠٠٣؛ إبراهيم، ٢٠٠٠) والتي أشارت إلى ضعف استخدام مؤشرات حجم الأثر كمؤشرات مكملة للدلالة الإحصائية.

السؤال الثاني: ما مستوى مؤشرات حجم الأثر في رسائل الماجستير المجازة بكلية التربية بالجامعة الإسلامية العالمية الماليزية؟

للإجابة على هذا السؤال تم حساب المتوسطات الحسابية والانحرافات المعيارية لقيم حجم الأثر لكل الاختبارات الإحصائية، وذلك لمعرفة الواقع الفعلي لمستويات حجم الأثر، وتم تصنيفها حسب معيار كوهين cohen ، والجدول(٣) يوضح هذه النتائج.

الجدول(٣): المتوسطات الحسابية والانحرافات المعيارية لقيم حجم الأثر لكل الاختبارات الإحصائية، ومستوها حسب معيار كوهين cohen

الاختبار الإحصائي	العدد	حجم الأثر		
		المتوسط الحسابي	الانحراف المعياري	المستوى
إجمالي اختبارات ت	431	0.37	0.60	منخفض
إجمالي اختبارات تحليل [illegible]	225	0.23	0.24	منخفض
معامل ارتباط بيرسون	397	0.26	0.25	عالي
معامل الانحدار	55	2.11	5.63	عالي

يبين الجدول(٣) أن المتوسط الحسابي لحجم الأثر لكل من معامل الانحدار ومعامل ارتباط بيرسون بلغا (٢.١١)، و (٠.٢٦) على التوالي، وتمثل مستويات عالية حسب تصنين كوهين، أما بالنسبة لمستويات حجم الأثر لإجمالي اختبارات (ت) وإجمالي اختبارات التباين فقد جاءت منخفضة، وحيث أن دراسات الماجستير أغلبها لم تستخدم مؤشرات حجم الأثر، وأن نتائجها جاءت بمستوى حجم أثر منخفض، لذا فإنه يصعب الاعتماد عليها في بناء القرارات التربوية، كما أن الاستفادة منها في الميدان التربوي يحتاج إلى تأني وحذر.

هذا وقد تفاوتت نتائج الدراسات السابقة في مستويات حجم الأثر للاختبارات الإحصائية المستخدمة، حيث أن نتائج هذه الدراسة تتفق مع نتائج دراسة (المالكي،٢٠١٨) في مستوى حجم الأثر لاختبارات (ت)، وتخالفها في نتائج اختبارات التباين، كما اتفقت نتائج هذه الدراسة مع نتائج دراسات (القضاة، ٢٠١٦؛ البارقي، ٢٠١٢؛ الحداد، ٢٠٠٦؛ محمود، ٢٠٠٣؛ الدرابيع، ٢٠٠٣؛ إبراهيم،٢٠٠٠) والتي أشارت انخفاض مستوى حجم الأثر للاختبارات الإحصائية المستخدمة.

السؤال الثالث: هل تختلف مؤشرات حجم الأثر باختلاف نتائج الدلالة الإحصائية في رسائل الماجستير المجازة بكلية التربية بالجامعة الإسلامية العالمية الماليزية؟

للإجابة على هذا السؤال فقد تم تصنيف نتائج مؤشرات حجم الأثر للاختبارات الإحصائية، إلى مستوياتها الثلاثة (منخفض، متوسط، عالي)، وكما تم أيضا تحديد نتيجة الدلالة الإحصائية (دال، غير دال) لجميع الاختبارات الإحصائية، وتم عرض نسب الاتفاق والاختلاف بين مستويات حجم الأثر، ومستويي الدلالة الإحصائية، حيث أن النتائج غير الدالة إحصائيا يفترض أن يكون لها حجم أثر منخفض، أما النتائج الدلالة إحصائيا فيكون حجم أثرها متوسط أو عالي، والجدول(٤) يوضح هذه النتائج.

الجدول(٤): نسب الاتفاق والاختلاف بين مستويات حجم الأثر، ومستويي الدلالة الإحصائية للاختبارات الإحصائية

الاختبار الإحصائي	حجم الأثر	الدلالة الإحصائية				المجموع	
		دالة إحصائيا		غير دالة			
		العدد	النسبة	العدد	النسبة	العدد	النسبة%
اختبارات ت	منخفض	114	26.45	252	58.47	366	84.92
	متوسط	18	4.18	11	2.55	29	6.73
	عالي	35	8.12	1	0.23	36	8.35
	المجموع	167	38.75	264	61.25	431	100
اختبارات تحليل التباين (Anova– Univariate –) Manova	منخفض	25	11.11	135	60.00	160	71.11
	متوسط	22	9.78	14	6.22	36	16.00
	عالي	18	8.00	11	4.89	29	12.89
	المجموع	65	28.89	160	71.11	225	100
معامل ارتباط بيرسون	منخفض	58	14.61	87	21.91	145	36.52
	متوسط	76	19.14	2	0.50	78	19.65
	عالي	174	43.83	0	0.00	174	43.83
	المجموع	308	77.58	89	22.42	397	100
معامل الانحدار	منخفض	24	43.64	8	14.55	32	58.18
	متوسط	6	10.91	1	1.82	7	12.73
	عالي	9	16.36	7	12.73	16	29.09
	المجموع	39	70.91	16	29.09	55	100

*التظليل يشير إلى تكرار ونسب الاتفاق بين الدلالة الإحصائية وحجم الأثر

تشير نتائج الجدول(٤) إلى أن أغلب اختبارات (ت) المستخدمة في رسائل الماجستير كان حجم أثرها منخفض (٨٤.٩٢%)، وكذلك أغلب اختبارات التباين (٧١.١١%)، كما أن أكثر من نصف معاملات الانحدار المستخدمة في رسائل الماجستير لها حجم أثر منخفضا، إلا أن المتوسط الحسابي لحجم أثر لاختبارات معامل الانحدار كان عاليا، وذلك لوجود قيم متطرفة له بلغ بعضها (١٩)، و(١١.٥)، أما حجم الأثر لمعامل ارتباط بيرسون فتوزعت على المستويات الثلاثة، إلا أن النسبة الأكبر كانت لحجم الاثر العالي.

كما يتضح من الجدول(٤) أن مستويات حجم الأثر لكل من اختبارات (ت) واختبارات تحليل التباين ومعامل ارتباط بيرسون المستخدمة في رسائل الماجستير اتفقت مع نتائج الدلالة الإحصائية، حيث أن أغلب النتائج التي لها حجم أثر منخفض كانت غير دالة إحصائيا، وأكثر النتائج ذات حجم أثر متوسط أو عالي لها دلالة إحصائية، واتفقت هذه النتائج مع نتائج دراسات (المالكي، ٢٠١٨؛ القضاة، ٢٠١٦)، كما اختلفت هذه النتائج مع نتائج دراسات (إبراهيم،٢٠٠٠؛ الجودة،٢٠٠٤) التي أشارت إلى أن النسبة الأكبر من الفرضيات الدلالة إحصائيا اقترنت بحجم أثر منخفض، كما اختلفت مع دراسة الحداد(٢٠٠٦) والتي توصلت إلى أن (٦٥.١%) من الفرضيات الدالة إحصائيا اقترنت بحجم أثر منخفض، أما مستويات حجم الأثر لمعاملات الانحدار فلم تتفق مع نتائج الدلالة الإحصائية، حيث أن النسبة الأكبر للمعاملات ذات حجم أثر منخفض، اقترنت بالدلالة الإحصائية، مما يُشير إلى احتمالية أن يكون الطلبة الذين استخدموا معامل الانحدار في فحص فرضياتهم الصفرية قد وقعوا في الخطأ من النوع الثاني وهو قبول الفرضية الصفرية وهي في الواقع خاطئة.

السؤال الرابع: هل توجد علاقة ذات دلالة إحصائية بين كلٍ من الدلالة الإحصائية ومؤشرات حجم الأثر في رسائل الماجستير المجازة بكلية التربية بالجامعة الإسلامية العالمية الماليزية؟

للكشف عن العلاقة بين الدلالة الإحصائية ومستويات الدلالة العملية في كل اختبار من الإختبارات الإحصائية المحددة في هذه الدراسة؛ قام الباحث باستخدام اختبار مربع كاي X^2 ، ومؤشر (w) لكوهين Cohen لمعرفة الدلالة العملية لاختبار مربع كاي، والجدول(٥) يوضح نتائج هذا الاختبار.

الجدول(٥): نتائج اختبار مربع كاي X^2 للعلاقة بين الدلالة الإحصائية ومستويات حجم الأثر في اختبارات (ت) واختبارات تحليل التباين ومعامل ارتباط بيرسون والانحدار المتعدد

الاختبار	قيمة X2	درجات	مستوى	حجم الأثر	مستوى
اختبارات ت	67.418	2	0.000	0.76	عالي
اختبارات تحليل	47.438	2	0.000	0.85	عالي
معامل ارتباط	185.708	2	0.000	1.14	عالي
معامل الانحدار	2.670	2	0.263	0.52	عالي

يتضح من نتائج مربع كاي X^2 بالجدول (٥) وجود علاقة ذات دلالة إحصائية بين الدلالة الإحصائية ومستويات حجم الأثر لكل من اختبارات (ت) واختبارات تحليل التباين ومعامل ارتباط بيرسون في رسائل الماجستير بكلية التربية بالجامعة الإسلامية العالمية الماليزية، كما يبين الجدول(٥) نتائج مؤشر w لقياس مقدار حجم الأثر لاختبار مربع كاي X^2 لكل من اختبارات (ت) واختبارات تحليل التباين ومعامل ارتباط بيرسون، وكلها توضح أن حجم الأثر عالي لكل هذه الاختبارات، مما يؤكد عدم استقلالية الدلالة الإحصائية عن حجم الأثر لهذه الاختبارات الإحصائية، وكذلك يوجد تطابق بين نتائج الدلالة الإحصائية لاختبار مربع كاي X^2 ونتائج حجم الأثر، وهذا ما أكدته نتائج السؤال الثالث من هذه الدراسة، كما أشارت نتائج اختبار مربع كاي X^2 عدم وجود علاقة دالة إحصائيا بين الدلالة الإحصائية ومستويات حجم الأثر لاختبار معامل الانحدار، حيث بلغت قيمته (٢.٦٧) وهي ليست ذات دلالة إحصائية عند مستوى دلالة أقل من (٠.٠٥)، وهو ما يشير إلى عدم تطابق نتائج الدلالة الإحصائية لها مع نتائج مستويات حجم الأثر، وهو أيضا أكدته نتائج السؤال الثالث من هذه الدراسة حيث أوضحت أن نسبة كبيرة من المعاملات غير الدالة إحصائية كانت ذات حجم أثر منخفض.

وللتأكيد على نتجية اختبار مربع كاي X^2 ، تم استخدام معامل ارتباط بوينت بايسيريال Point Bi–serial Coefficient؛ لإيجاد قوة العلاقة بين المتغير الاسمي(الدلالة الإحصائية) والمتغير الكمي (قيم حجم الأثر)، وذلك لكل الاختبارات الإحصائية المستخدمة في هذه الدراسة، والجدول(٦) يوضح هذه النتائج.

الجدول(٦): نتائج معامل ارتباط بوينت بايسيريال Point Bi-serial Coefficient لإيجاد قوة العلاقة بين الدلالة الإحصائية وقيم حجم الأثر

الاختبار الإحصائي	معامل ارتباط Point Bi -serial Coefficient	مستوى الدلالةsig	معامل التحديد R2	مستوى حجم الأثر لمعامل الارتباط
اختبارات ت	0.415	0.000	0.172	متوسط
اختبارات تحليل التباين	0.414	0.000	0.171	متوسط
معامل ارتباط بيرسون	0.544	0.000	0.296	عالي
معامل الانحدار	-0.163	0.235	0.027	منخفض

يتضح من الجدول(٦) وجود علاقة ارتباطية متوسطة ودالة إحصائيا عند مستوى دلالة أقل من (٠.٠٥) بين الدلالة الإحصائية وقيم حجم الأثر، لكل من اختبارات (ت) واختبارات تحليل التباين ومعامل ارتباط بيرسون، كما أن حجم الأثر لهذه العلاقات جاء بمستوى متوسط لاختبارات (ت) واختبارات تحليل التباين، وعالي لمعامل ارتباط بيرسون، أما علاقة الدلالة الإحصائية بقيم حجم الأثرلمعاملات الانحدار فجاءت ضعيفة وليست ذات دلالة إحصائية، وهذه النتائج تؤكد على نتائج اختبار مربع كاي X^2 في الجدول(٧) السابق، هذا واتفقت هذه نتائج علاقة الدلالة الإحصائية بمستويات حجم الأثر مع نتائج دراسات (المالكي، ٢٠١٨؛ الدرابيع، ٢٠٠٣؛ الجودة، ٢٠٠٤)

التوصيات والمقترحات:

بناء على النتائج السابقة يوصي الباحث بالآتي:

١- إعادة النظر في معايير إجازة رسائل الماجستير بكلية التربية بالجامعة الإسلامية العالمية الماليزية، بحيث يُشترط تزويد كل اختبار إحصائي يتم استخدامه في تحليل النتائج بمؤشرات حجم الأثر.

٢- التأكيد على ضرورة تدريس مؤشرات حجم الأثر في مقررات الإحصاء لطلبة الماجستير بكلية التربية بالجامعة الإسلامية العالمية الماليزية.

٣- ضرورة تعيين مشرف إحصائي لطلبة الماجستير بكلية التربية بالجامعة الإسلامية العالمية الماليزية، لتوجيههم إلى استخدام الأساليب الإحصائية المناسبة لأسئلتهم البحثية.

٤- اقتراح إنشاء مركز الاستشارات الإحصائية يتبع مركز الدراسات العليا بالجامعة الإسلامية العالمية الماليزية، يسعى إلى نشر الثقافة الإحصائية للباحثين من الهيئة التدريسية والطلبة، وتوجيههم ومساعدتهم في الجوانب الإحصائية، من خلال الدورات التدريبية، والنشرات الدورية، والاستشارات الإحصائية، كما يكون من اختصاصات هذا المركز اعتماد اطروحات الدكتوراه ورسائل الماجستير التي يعدها الطلبة من مختلف الكليات بالجامعة الإسلامية العالمية الماليزية، بحيث يعتمد الأساليب الإحصائية التي يقترحها الطالب في مخطط دراسته قبل مناقشة المخطط، والمرحلة الثانية يعتمد التحليل الإحصائي النهائي للدراسة.

٥- اقتراح إجراء دراسة مماثلة لبحث واقع استخدام مؤشرات حجم الأثر للأساليب الإحصائية غير المعلمية، في رسائل الماجستير المجازة بكلية التربية بالجامعة الإسلامية العالمية الماليزية.

٦- اقتراح إجراء دراسات مماثلة لتحليل اطروحات الدكتوراه المجازة بكلية التربية بالجامعة الإسلامية العالمية الماليزية.

المراجع:

إبراهيم، إبراهيم رشاد إبراهيم الحاج(٢٠٠٠). واقع الدلالة الإحصائية والدلالة العملية وقوة الاختبارات الإحصائية المستخدمة في بحوث مجلة دراسات العلوم التربوية في الأعوام (١٩٩٧-١٩٩٨) في الجامعة الأردنية. رسالة ماجستير غير منشورة، كلية الدراسات العليا، الجامعة الأردنية، الأردن.

أبو جراد، حمدي يونس(٢٠١٣). قوة الاختبارات الإحصائية وحجم الأثر في البحوث التربوية المنشورة في مجلة جامعة القدس المفتوحة للأبحاث والدراسات. مجلة العلوم التربوية والنفسية، ١٤(٢)، ٣٤٩-٣٦٨. البحرين.

بابطين، عادل أحمد(٢٠٠٢). مشكلات الدلالة الإحصائية في البحث التربوي وحلول بديلة. رسالة ماجستير غير منشورة، كلية التربية بجامعة أم القرى، السعودية.

البارقي، طلال هيازع (٢٠١٢). واقع الدلالة الإحصائية والدلالة العملية للبحوث المنشورة في مجلة جامعة أم القرى للعلوم التربوية والاجتماعية والإنسانية، في المدة ١٤٢٥-١٤٣٠هـ. رسالة ماجستير غير منشورة، كلية التربية بجامعة أم القرى، السعودية.

باهي، مصطفى حسين إبراهيم (٢٠١٠). العلاقة بين الدلالة الاحصائية وحجم التأثير في البحوث التربوية والنفسية، مستقبل إعداد المعلم في كليات التربية وجهود الجمعيات العلمية في عمليات التطوير بالعالم العربي. مج٢، ٤١٥-٤٤٤، كلية التربية، جامعة حلوان، مصر.

بخارى، ماجد عبدالفتاح (٢٠١٦). أثر انتهاك افتراض تجانس التباين على قيم مربع ابتا ومربع أوميجا كمؤشرات لفحص الدلالة العملية في تحليل التباين الأحادي. دراسات عربية في التربية وعلم النفس، ع ٧٦.

بخيت، صلاح الدين والشيخ، فضل المولى (٢٠٠٤). رسائل الماجستير والدكتواره في علم النفس بالجامعات السودانية في ربع قرن (١٩٧٧-٢٠٠٣م)، (دراسة توثيقية تحليلية). مجلة الدراسات التربوية (١٣)، ٧٧-١١٣.

الثبيتي، علي حامد (٢٠٠٨). تصاميم البحوث العلمية ودورها في صدق نتائج الدراسات التربوية. رسالة الخليج العربى، ٢٩(١٠٨)، السعودية.

الجودة، ماجد محمود شريف(٢٠٠٤). الدلالة الإحصائية والدلالة العملية وقوة الاختبار للأبحاث المنشورة في مجلة أبحاث اليرموك سلسلة العلوم الإنسانية والاجتماعية (١٩٨٥-٢٠٠١). رسالة ماجستير غير منشورة، كلية الدراسات العليا، جامعة اليرموك، الأردن.

الحداد، سعدة أحمد محمد (٢٠٠٦). قوة الاختبارات الإحصائية وواقع الدلالة الإحصائية والدلالة العملية في بحوث مجلة الهيئة القومية للبحث العلمي في الجماهيرية العظمى. رسالة ماجستير غير منشورة، كلية الدراسات العليا، جامعة أم درمان، السودان.

حسن، عبدالمنعم أحمد(٢٠٠٨). أوجه القصور في استخدام مؤشرات الدلالة العملية في البحوث التربوية والنفسية، دراسات في المناهج وطرق التدريس. ع١٣٤، ١٤-٣٩.

الحلاق، علي سامي علي(٢٠٠٨). دراسة تحليلية لأطروحات مناهج اللغة العربية وطرائق تدريسها في جامعة عمان العربية للدراسات العليا من عام 2003 إلى عام 2007م. مجلة اتحاد الجامعات العربية، ع ٥١، الأردن، ٤٤٥-٤٨٣.

حمزة، إبراهيم حمزة، وإسماعيل، فخر الدين الحاج(٢٠١٦). أثر حجم العينة والانحراف المعياري على قوة الاختبار T للتوزيع الأسي. مجلة الدراسات العليا –جامعة النيلين، ٦(٢٢).

الدرابيع، ماهر يونس(٢٠٠٣). واقع الدلالات الإحصائية والعملية وقوة الاختبارات الإحصائية المستخدمة في بحوث مجلة مؤتة للبحوث والدراسات – سلسلة العلوم الإنسانية والاجتماعية. مؤتة للبحوث والدراسات – العلوم الانسانية والاجتماعية ، ١٨(١)، الاردن.

الذبياني، عبدالله فالح (٢٠١٥). توجيات أطروحات الدكتوراه بقسم الإدارة التربوية والتخطيط بكلية التربية بجامعة أم القرى. رسالة ماجستير غير منشورة، كلية التربية، جامعة أم القرى، مكة المكرمة.

سلامة، حسن علي حسن (٢٠٠٤). الدلالة الإحصائية والدلالة العلمية في البحوث التربوية. المجلة التربوية، ج٢٠، ٣-١٤، جامعة سوهاج – كلية التربية، مصر.

الشريفين، نضال كمال محمد(٢٠١٧). ما وراء التحليل للأبحاث المنشورة في المجلة الأردنية في العلوم التربوية : الدلالة العملية وقوة الاختبار. مجلة اتحاد الجامعات العربية للتربية وعلم النفس،١٥(٣)، سوريا.

الشمراني، محمد موسى (٢٠٠٨). أثر اختلاف أحجام العينات على قيم مستويات الدلالة الإحصائية المحسوبة للاختبارات الإحصائية لتحليل التباين متعدد المتغيرات : دراسة مقارنة. مجلة التربية للبحوث التربوية والنفسية والاجتماعية، ١٣٨(٢)، جامعة الأزهر – كلية التربية، مصر.

الشمراني، محمد موسى (٢٠١٢). إستخدام مقاييس الدلالة العملية لحجم التأثير في الحكم على قياس أهمية نتائج البحوث العلمية. مجلة كلية التربية بالمنصورة، ٧٨(٣)، مصر.

الصائغ، ابتسام حسن مدني (١٩٩٦). الدلالة الإحصتائية والدلالة العملية لاختبار ت و ف: دراسة تحليلية تقويمية من خلال رسائل الماجستير التي قدمت في كلية التربية بجامعة أم القرى حتى عام ١٤١٥هـ. رسالة ماجستير غير منشورة، جامعة أم القرى، السعودية.

طعمة، حسن ياسين (٢٠١١). الاختبارات الإحصائية (أسس وتطبيقات). دار صفاء للنشر والتوزيع، عمّان، الأردن، ط١.

عبدالفتاح، فيصل أحمد (٢٠١١). قوة الاختبارات الاحصائية وحجم الأثر في البحوث التربوية والنفسية. مجلة جامعة الملك سعود – العلوم التربوية والدراسات الاسلامية، ٢٣(٢)، السعودية.

العبيدي، عبدالله أحمد خلف(٢٠١٩). حجم الأثر وتطبيقاته في منهجية البحوث التربوية والنفسية والطبية. دار وائل للنشر والتوزيع، الأردن، ط١.

عصر، رضا مسعد السعيد(٢٠٠٣). حجم الأثر: أساليب إحصائية لقياس الأهمية العملية لنتائج البحوث التربوية. المؤتمر العلمي الخامس عشر، مناهج التعليم والإعداد للحياة المعاصرة، مج٢، ٦٤٤-٦٧٤، الجمعية المصرية للمناهج وطرق التدريس، مصر.

القضاة، محمد عواد سالم (٢٠١٦). درجة التوافق بين نتيجة الدلالة الإحصائية والدلالة العملية في نتائج التحليلات الإحصائية في رسائل الماجستير في كلية العلوم التربوية في جامعة مؤتة. رسالة ماجستير غير منشورة، جامعة مؤتة، الأردن.

المالكي، فهد بن عبدالله بن عمر (٢٠١٨). واقع الدلالة الإحصائية والدلالة العملية وقوة الاختبار الإحصائي (دراسة تطبيقية على بحوث المجلات التربوية بالجامعات السعودية). رسالة دكتوراه غير منشورة بجامعة أم القرى، مكة المكرمة، السعودية.

محمد، محمد إبراهيم محمد (٢٠١٣). تحليل قوة الاختبار الإحصائي و علاقتها بمستوى الدلالة وحجم التأثير في البحوث التربوية. دراسات عربية في التربية وعلم النفس، ٣٧(٣)، ١٠٠-١٢٥، السعودية.

محمود، فريال محمود محمد الحاج (٢٠٠٣). واقع الدلالة الإحصائية والدلالة العملية وقوة الاختبار للاختبارات الإحصائية المستخدمة في رسائل الماجستير الصادرة عن كلية العلوم التربوية في جامعة مؤتة. رسالة ماجستير غير منشورة، كلية الدراسات العليا، جامعة مؤتة، الأردن.

المعثم، خالد بن عبدالله بن صالح (١٤٢٩هـ). توجيات أبحاث تعليم الرياضيات في الدراسات العليا بجامعات الممكة العربية السعودية (دراسة تحليلية لرسائل الماجستير والدكتوراه). اطروحة دكتوراه غير منشورة، كلية التربية، جامعة أم القرى، مكة المكرمة.

النجار، عبدالله بن عمر (٢٠٠٦). دراسة تحليلية لقوة الاختبار الإحصائي في البحوث الإدارية المنشورة. المجلة العلمية لجامعة الملك فيصل – العلوم الانسانية والادارية، ٧(٢)، السعودية.

النجار، عبدالله بن عمر (٢٠٠٦). دراسة تحليلية لقوة الاختبار الإحصائي في البحوث الإدارية المنشورة. المجلة العلمية لجامعة الملك فيصل – العلوم الانسانية والادارية، ٧(٢)، السعودية.

نصار، يحيى حياتي بكر(٢٠١٧). الدلالة الاكلينيكية للبحوث النفسية والتربوية المستخدمة للتصاميم التجريبية :دراسة تحليلية. مجلة الدراسات التربوية والنفسية. مج ١١(٢). جامعة السلطان قابوس. كلية التربية، سلطنة عمان.

نصار، يحيى حياتي(٢٠٠٦). استخدام حجم الأثر لفحص الدلالة العملية للنتائج في الدراسات الكمية. مجلة العلوم التربوية والنفسية، ٧(٢)، ٣٥–٥٩، البحرين.

AERA (2006), "Standards for reporting on empirical social science research in AERA publications," . Retrieved from https://www.aera.net/Portals/38/docs/12ERv35n6Standard4Report.pdf

APA (1994). Publication Manual of the American Psychological Association, 4th Edition. Washington, DC: American Psychological Association.

APA (2010). Publication Manual of the American Psychological Association, 6th Edition. Washington, DC: American Psychological Association.

CAPRARO,R.; CAPRARO, M.M (2002). TREATMENTS OF EFFECT SIZES AND STATISTICAL SIGNIFICANCE TESTS IN TEXTBOOKS. Educational and Psychological Measurement, 62(5), 771-782. DOI: 10.1177/001316402236877

Carver, R. P. (1978). The case against statistical significance testing. Harvard Educational Review, 48, 378-399.

Coe, Robert(2002). It's the Effect Size, Stupid :What effect size is and why it is important. Paper presented at the Annual Conference of the British Educational Research Association, University of Exeter, England, 12-14 September 2002.

Cohen, J(1988).Statistical Power Analysis for the Behavioral Sciences, 2nd ed., Hillsdale, NJ: Lawrence Erlbaum.

Cohen, J. (1962). The statistical power of abnormal-social psychological research: A review. The Journal of Abnormal and Social Psychology, 65(3), 145-153.

Ellis,Paul D.(2010). The Essential Guide to Effect Sizes Statistical Power, Meta-Analysis, and the Interpretation of Research Results. Cambridge University Press, New York, First published.

Fidalgo, A. M., Alavi, S. M., & Amirian, S. M. R. (2014). Strategies for testing statistical and practical significance in detecting DIF with logistic regression models. Language Testing, 31(4), 433–451. DOI: 10.1177/0265532214526748

Gentry,M., Peters, S. J.(2009). Effect Sizes in Gifted Education Research, Gifted Child Quarterly,53(3), 219–222.

George A. Morgan, Jeffrey A. Gliner, Robert J. Harmon (2006). Understanding and Evaluating Research in Applied and Clinical Settings. Lawrence Erlbaum Associates, Publashers Mahwah, New Jersey, London.

Hess, Melinda Rae, (2003). Effect sizes, significance tests, and confidence intervals: Assessing the influence and impact of research reporting protocol and practice. Graduate Theses and Dissertations. University of South Florida , A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy. Retrieved from http://scholarcommons.usf.edu/etd/1390

Howitt,D., & Cramer,D. (2008). Introduction to Statistics in psychology, Peason Education Limited, Edinburgh Gate, Harlow, Essex CM20 2JE, England, 4th ed.

Huberty, C.J. (2002). 'A history of effect size indices'. Educational and Psychological Measurement, 62(2), 227–240.

Ialongo, C. (2016). Understanding the effect size and its measures. Biochemia medica, 26(2), 150–63.

Kenton,W.(2019). Statistical Significance, Investopedia, Retrieved from https://www.investopedia.com/terms/s/statistically_significant.asp.

Kirk, R. E. (1996). Practical significance: A concept whose time has come. Educational and psychological measurement, 56(5), 746-759.

Kirk,R.E. (2003). "The importance of effect magnitude," in S.F. Davis (editor), Handbook of Research Methods in Experimental Psychology. Oxford, UK: Blackwell, 83-105.

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. Frontiers in psychology,4(863).

Levine, T. R., Asada, K. J., & Carpenter, C. (2009). Sample sizes and effect sizes are negatively correlated in meta-analyses: Evidence and implications of a publication bias against non-significant findings. Communication Monographs, 76(3), 286-302 .

Loeb, S., Dynarski, S., McFarland, D., Morris, P., Reardon, S., & Reber, S. (2017). Descriptive analysis in education: A guide for researchers. (NCEE 2017-4023). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Pallant,J.(2010).SPSS survival manual : a step by step guide to data a nalysis using SPSS. Maidenhead, Berkshire. U.K. ; New York, NY : Open University Press, c2005, 4th ed.

Pedersen, S.(2003). Effect Sizes and "What If" Analyses as Supplements to Statistical Significance Tests. Journal of Early Intervention, 25(4), 310-319.

Plucker, J. A. (1997). Debunking the Myth of the 'Highly Significant' Result: Effect Sizes in Gifted Education Research. ROEPER REVIEW, 20(2).

Tomczak, M., & Tomczak, E. (2014). The need to report effect size estimates revisited. An overview of some recommended measures of effect size. Trends in Sport Sciences, 21(1), 19-25 .

Vacha-Haase, T., & Thompson, B. (2004). How to estimate and interpret various effect sizes. Journal of Counseling Psychology, 51, 473- 481.

Ward,R.M.(2002). HIGHLY SIGNIFICANT FINDINGS IN PSYCHOLOGY: A POWER AND EFFECT SIZE SURVEY, Graduate Theses and Dissertations, A Dissertation Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy, UNIVERSITY OF RHODE ISLAND.