كلية التربية

كلية معتمدة من الهيئة القومية لضمان جودة التعليم

إدارة: البحوث والنشر العلمي (المجلة العلمية)

=======

حجم تأثير الاختبارات الإحصائية المعلمية واللامعلمية المستخدمة في رسائل الماجستير بكلية التربية بجامعة الملك سعود

إعـــــداد

فيصل أحمد العبدالقادر

جامعة الملك سعود – كلية التربية – قسم علم النفس

﴿ المجلد السادس والثلاثون – العدد الرابع – أبريل ٢٠٢٠م ﴾

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

الملخص:

هدفت الدراسة الحالية إلى التعرف على أهم الطرق الإحصائية لحساب حجم التأثير للاختبارات الإحصائية بنوعيها المعلمية واللامعلمية والمستخدمة في رسائل الماجستير بكلية التربية بجامعة الملك سعود، والتعرف على واقع استخدام حجم التأثير، ولتحقيق هدف الدراسة؛ قام الباحث بتفريغ البيانات في نماذج معدة لكل اختبار إحصائي (تحليل التباين الأحادي، مربع كاي، معامل ارتباط بيرسون، الفروق بين المتوسطات، اختبار مان ويتني، ومعامل الإنحدار المتعدد)، واستخدم الباحث التكرارات والنسب المؤيوة، ومعادلات حجم التأثير وقيم مربع كاي وذلك للإجابة على تساؤلات الدراسة التالية: ما الاختبارات الإحصائية المعلمية واللامعلمية المستخدمة في رسائل الماجستير بكلية التربية بجامعة الملك سعود ؟ ما واقع استخدام حجم التأثير في الاختبارات الإحصائية المعلمية واللامعلمية المستخدمة في رسائل الماجستير بكلية التربية بجامعة الملك سعود ؟ ما أحجام التأثير للاختبارات الإحصائية المعلمية واللامعلمية المستخدمة في رسائل الماجستير بكلية التربية بجامعة الملك سعود؟، واستخدم الباحث المنهج الوصفي المسحي للإجابة على هذه التساؤلات الدراسة، حيث اشتملت عينة الدراسة على ١٥٤ رسالة ماجستير، وتوصلت الدراسة إلى النتائج التالية: الأساليب المعلمية من أكثر الأساليب الإحصائية المستخدمة في كلية التربية بجامعة الملك سعود بتكرار بلغ ٥٥٩ استخداماً ويمثل ذلك ٩٣% تقريباً، وكان اختبار ت من أكثرها شيوعاً بتكرار ٢٥٦ استخداماً مما يمثل نسبة ٤٣% تقريباً، وأيضاً توصلت الدراسة أن استخدام حجم التأثير كان منخفض بشكل كبير حيث تمثل نسبته ١.٣% فقط وبتكرار ٢ مرة، وأوصت الدراسة بضرورة حساب حجم التأثير بالطرق المناسبة في البحوث والدرسات العلمية وعدم الإكتفاء فقط بدلالة الاختبارات الإحصائية.

الكلمات المفتاحية: الأساليب الإحصائية، الدلالة الإحصائية، حجم التأثير كوهين، الدلالة العملية

Abstract

The present study aimed at identifying important statistical procedures used for calculating effect size of parametric and nonparametric tests in master theses registered at Faculty of Education-King Saud University from 1995 to 2008. Moreover, it explored the current usage of effect size and whether this varies according to the methods used (i.e. descriptive or experimental) or the sample sizes (i.e. small, medium, or large. The researcher tried to answer the following questions:

- What are the parametric and nonparametric tests used in master theses registered at faculty of Education King Saud University?
- What is the current usage of effect size in parametric and nonparametric tests used in master theses registered at Faculty of Education King Saud University?
- What are the effect sizes used in parametric and nonparametric tests used in master

theses registered at Faculty of Education King Saud University?

The study results were as follows:

1- The most statistical procedures used at Faculty of Education-King Saud University were

parametric tests (frequency=559, 93.48 %), and the t-test was the most common among

them (frequency=256,42.81%).

2- Effect size was rarely used at Faculty of Education-King Saud University (frequency=2,1.3%).

3-Effectsize for the statistical procedures used at Faculty of Education King Saud University Ranged from small to medium.

The researcher concluded with the following recommendations:

- Teaching nonparametric tests and integrating them into courses for postgraduate students.
- Encouraging researchers to calculate effect sizes for the statistical procedures used for testing their hypotheses.
- Integrating effect size into courses for post graduate students.
- In addition to the traditional use of statistical significance, the researcher encourages the inclusion of effect size in the results of published studies.

المقدمة:

يحتل علم الإحصاء أهمية خاصة بين العلوم المختلفة سواء كانت طبيعية أو إنسانية أو اجتماعية وذلك لأنه يتعامل مع الحاجات الإنسانية والعلمية في مختلف مستوياتها وبأساليب متنوعة بدءاً من البساطة التي تصف معلومات أولية (بيانات إحصائية) وصولاً إلى أساليب أكثر تعقيداً وتطوراً لتحليل البيانات ذات الصلة بوصف وتفسير الظواهر والأحداث وعلاقتها مع بعضها والتنبؤ بها والسيطرة عليها في ضوء القوانين التي تحكمها (الجادري ، ٢٠٠٣).

ويعرف علم الإحصاء بأنه ذلك الفرع من العلوم الذي يختص بالطرق العلمية لجمع البيانات و تنظيمها و تلخيصها و عرضها وتحليلها وذلك للوصول إلى نتائج مقبولة تؤدي إلى اتخاذ قرارات سليمه (بري وهندي والحسيني ، ١٩٩٨) ، ويوجد نوعان من الإحصاء ، أولهما ما يسمى بالإحصاء المعلمي) البارامتري (parametric) والذي يهتم بالبيانات الرقمية الحقيقية للمتغيرات لدى عينات كبيرة الحجم ممثلة للمجتمع الأصلي الذي سحبت منه ، كما تتمتع باعتدالية توزيعها (أبو زيد ، ٢٠٠٢) ، وثانيهما الإحصاء اللامعلمي (اللابارامتري) Nonparametric وهو الإحصاء الذي يتحرر من القيود المسبقة لشكل التوزيع التكراري وحجم العينة وتصلح لمستويات القياس التي لا يصلح لها الإحصاء المعلمي وخاصة المقاييس التي تعتمد على تصنيف الأفراد أو ترتيبهم تنازلياً أو تصاعدياً ولكن الإحصاء اللامعلمي أقل دقة من الإحصاء المعلمي (منسي ، ١٩٨٩).

ومن الملاحظ أن بعض الباحثين يعتمدون في تقرير نتائجهم على الدلالة الإحصائية دون محاولة الكشف عن مقدار العلاقة القائمة بين المتغيرين ، وتصبح هناك مغالاة في تفسير النتائج اعتماداً على مستوى الدلالة على الرغم من أنه ربما لا تكون لها قيمة من الناحية التطبيقية أو العملية ، ولذلك فإذا وجد الباحث أن القيمة دالة إحصائياً ، فمعنى ذلك أن المتغير المستقل له تأثير على المتغير التابع ، ولكنه لا يدل على حجم التأثير أو درجة العلاقة بين المتغيرين . وربما كانت الدلالة الإحصائية هنا لا تعني وجود علاقة قوية بين المتغيرين (الشربيني ، ١٩٩٥).

وفي دراسة (بابطين ،٢٠٠٢) عن مشكلات الدلالة الإحصائية في البحث التربوي والحلول البديلة لها ، توصل إلى أن أهم مشكلات اختبارات الدلالة الإحصائية في البحث التربوي هي : استخدام نتائج الدلالة الإحصائية كتفسير لأهمية النتائج و لتأثير المعالجة ، واستخدام قيمة الإحتمالية (p) لتقدير حجم التأثير وكتفسير لاحتمالية الفرضية الصفرية نفسها ، ولتقويم إعادة النتيجة ، إضافة إلى التحيز Biasing للنتائج الدالة إحصائياً ، وصياغة التساؤلات البحثية والفرضيات ، والطبيعة الثنائية لنتائج اختبار الدلالة الإحصائية .كما توصلت الدراسة إلى أن أهم المفاهيم والأساليب الإحصائية التي يمكن أن تقدم حلولاً لمشكلات الدلالة الإحصائية في البحث التربوي هي : تحليل القوة الإحصائية Power analysis، وتقدير حجم التأثير ، بينما

يرى كوهين بأن فكرة حجم التأثير Effect Size تعتمد على الفروق بين المتوسطات Mean Differences وذلك باستخدام الانحراف المعياري SDبين المتغيرات المستقلة من جهة والمتغيرات التابعة من جهة أخرى عن طريق استخراج حجم تباين المتغير التابع الذي يمكن تفسيره عن طريق المتغير المستقل كما هو الحال في تحليل التباين. (أبو هاشم ، ٢٠٠٥) ، ويدل حجم التأثير على مدى تأثير المتغير المستقل على المتغير التابع موضع الاهتمام ، وهو الدلالة العملية للنتائج (مراد ،٢٠٠٠).

وكانت الدلالة العملية لحجم التأثير محل اهتمام كثير من الباحثين حيث أكد كل من (Welge,p.1990;Galaraza,etal.,1993;Olejnik,&Algina,2000;Haase,2001): (Roberts,&Henson,2002; Thompson , 2002 ;على أهمية حجم التأثير واعتباره من المحكات الأساسية لاختبار الدلالة الاحصائية بقصد الاستفادة من نتائجها بصورة عملية.، هذا وقد أكدت لجنة الاستدلال الاحصائى التابعة للرابطة الأمريكية لعلم النفس APA على ضرورة حساب حجم التأثير في البحوث المنشورة في دورياتها العلمية ، ووصف البحوث التي تستخدم هذا الأسلوب بالبحوث العلمية الجيدة. (Bird, 2002)

مشكلة الدراسة

إن التركيز في معظم الدراسات والبحوث النفسية والتربوية في البلاد العربية على الدلالة الإحصائية ، احتمالية رفض الفرضية الصفرية باستخدام الأساليب الإحصائية المناسبة وإغفال حجم التأثير(مدى قدرة الباحث على استخدام النتائج المتحصل عليها تفسيراً أو تطبيقاً) أدى إلى التقليل من قيمة نتائج الدراسات وأحياناً إلى تضليلها فالدلالة الإحصائية شرط ضروري لقبول الفرض أو رفضه ولكنه ليس كافياً في الحكم على قيمة تأثير المتغير المستقل على المتغير التابع، ومن هنا تبلورت مشكلة الدراسة في التعرف على واقع استخدام حجم تأثير الاختبارات الإحصائية المعلمية واللامعلمية في رسائل الماجستير بكلية التربية بجامعة الملك سعود.

أسئلة الدراسة

سعت هذه الدراسة إلى الإجابة عن التساؤلات التالية:

- ما الاختبارات الإحصائية المعلمية واللامعلمية المستخدمة في رسائل الماجستير بكلية التربية بجامعة الملك سعود ؟
- ما واقع استخدام حجم التأثير في الاختبارات الإحصائية المعلمية واللامعلمية المستخدمة في رسائل الماجستير بكلية التربية بجامعة الملك سعود ؟
- ما أحجام التأثير للاختبارات الإحصائية المعلمية واللامعلمية المستخدمة في رسائل الماجستير بكلية التربية بجامعة الملك سعود؟

أهداف الدراسة :

تسعى هذه الدراسة إلى تحقيق الأهداف التالية:

١. الكشف عن الاختبارات الإحصائية المعلمية واللامعلمية المستخدمة في رسائل الماجستير بكلية التربية بجامعة الملك سعود

٢. التعرف على أحجام التأثير المستخدمة في الاختبارات الإحصائية المعلمية واللامعلمية في رسائل الماجستير بكلية التربية بجامعة الملك سعود.

٣. القيام بحساب أحجام التأثير للاختبارات الإحصائية المعلمية واللامعلمية المستخدمة في رسائل الماجستير.

أهمية الدراسة:

تكمن أهمية الدراسة فيما يلي:

- الدراسة الحالية تعد من الدراسات القليلة التي تناولت البحث عن مدى استخدام الباحثين لحجم التأثير في اختباراتهم الإحصائية المعلمية واللامعلمية على مستوى رسائل الماجستير بالمملكة العربية السعودية.
- تعريف الباحثين بأهمية حساب حجم التأثير وعدم الاعتمادعلى الدلالة الإحصائية فقط ، مما يسهم بتحسين جودة الدراسات والبحوث في الوطن العربي، وزيادة التطبيقات العملية لنتائج البحوث في المجال النفسي والتربوي.

حدود الدراسة:

تتمثل محددات الدراسة بما يلي:

- **الحدود الموضوعية:** ركزت الدراسة الحالية على التعرف على واقع استخدام الباحثين لحجم تأثير الاختبارات الإحصائية المعلمية واللامعلمية.
- **الحدود المكانية:** اقتصرت الدراسة الحالية على رسائل الماجستير المقدمة من طلاب وطالبات كلية التربية بجامعة الملك سعود بمنطقة الرياض في التخصصات التالية (علم النفس – المناهج – الوسائل – الإدارة التربوية – التربية البدنية – التربية – التربية الخاصة).
- **الحدود الزمانية:** اقتصرت الدراسة الحالية على رسائل الماجستير خلال الفترة مابين ١٤١٤ هـ – ١٤٢٨ هـ.

مصطلحات الدراسة:

١. حجم التأثير Effect Size: هو عبارة عن قيمة كمية تشير إلى درجة العلاقة بين متغيرات الدراسة ضمن مجتمع محدد مسبقاً هو مجتمع الدراسة، بغض النظر عن الأسلوب الإحصائي المستخدم للتعبير عن ذلك الأثر (نصار، ٢٠٠٢) ويعرفه هاس Haase بأنة مقدار التباين في المتغير التابع الذي يمكن تفسيره من خلال متغير مستقل يرتبط أو يؤثر فيه .(النجار، ٢٠٠٥)

٢. الإحصاء المعلمي البارامتري : Parametric Statistics : أي الطرق الإحصائية التي تتطلب تحقق افتراضات معينة حول المجتمع الذي تسحب منه العينة ومن هذه الافتراضات أن يكون توزيع المجتمع طبيعياً ، وتجانس التباين ، والعينات العشوائية . ومن الأمثلة على الطرق الإحصائية المعلمية اختبار (ت) المستخدم للمقارنة بين المتوسط الحسابي لمجموعتين . واختبار (ف) المستخدم في المقارنة بين المتوسط الحسابي لعدة مجموعات (مراد ، ٢٠٠٠).

٣. الإحصاء اللامعلمي اللابارامتري : non– Parametric Statistics : وهي نوع من الطرق الإحصائية الاستدلالية التي يمكن باستخدامها التوصل إلى استنتاجات بشأن المجتمع في ضوء العينة بغض النظر عن نوع التوزيع النظري لذلك المجتمع أو الطريقة التي تستخدم في اختيار العينة ومن الأمثلة على الطرق الإحصائية اللامعلمية إختبار مان وتني – ولكوكسن – وكندال (توفيق ، ١٩٨٥).

الإطار النظري:

الإحصاء المعلمى Parametric Statistics هو أحد أنواع الأساليب الإحصائية الاستدلالية Inferential Statistics التي تهتم بالكشف والاستدلال على المجتمع اعتماداً على ما توافر من بيانات لدى الباحث خاصة بالعينة المأخوذة من هذا المجتمع ، كما تتناول أساليب اتخاذ القرارات الإحصائية .وعند إستخدام الإحصاء المعلمى فلابد من التنبه إلى أنه يستخدم في حالة العينات الكبيرة التي يشترط فيها توفر المعلومات عن مجتمعاتها (معلمات الأصل) مثل : أن يكون توزيع البيانات توزيعاً إعتدالياً ، تجانس التباين ، العينات العشوائية ، خطية العلاقة ، واستقلال العينات ، وغيرها ، ويستخدم فقط مع البيانات التي تكون عددية حقيقية ، أى مع البيانات التي تكون من نوع النسبة أو المسافة (الدردير ، ٢٠٠٦)

ومن أشهر الأساليب الإحصائية المعلمية اختبار ت t–test والذي يقوم على افتراض أن العينة المسحوبة من مجتمع إحصائي توزيعه معتدل أو قريب من الاعتدال . بناءً على ذلك يستطيع الباحث استخدام توزيع t لإجراء اختبارات الفروض وإيجاد حدود الثقة لمتوسط المجتمع أو الفرق بين متوسطين عندما تكون أحجام العينات صغيرة وتباينات المجتمعات مجهولة . أما إذا كان توزيع المجتمع الأصلى الذي سحبت منه العينة بعيداً عن الاعتدال بدرجة كبيرة فمن الأفضل استخدام الإختبارات اللامعلميه لإجراء اختبارات الفروض (تشاو ، ١٩٩٠) وفي أحيان كثيرة نحتاج إلى اختبار تساوى عدة مجتمعات طبيعية و ليس اثنين فقط لهذا تصبح طريقة اختبار t مطولة و تستغرق وقتاً أكثر حيث يجب علينا اختبار كل مجتمعين على حدة ، ومن ثم الاستنتاج فيما إذا كانت متوسطات المجتمعات متساوية أم لا (سليم و أبو حويج ، ٢٠٠٤) لهذا اقترح السير رونلاد فيشر Fisher أسلوباً إحصائياً يمكنه عقد هذه المقارنات في أن واحد وأطلق عليه تحليل التباين Analysis of Variance (ANOVA) . وقد كان لـبيرت Burt الريادة في تطبيق هذا الأسلوب في العلوم النفسية و التربوية وهناك أشكال لتحليل التباين تتوقف على عدد المتغيرات المستقلة و التابعة (الشربينى ، ١٩٩٥).

ومن الأساليب الإحصائية المعلمية مفهوم الانحدار Regression والذي يدل على مدى تأثر قيم المتغير التابع بالتغير الذي يطرأ على قيم المتغير المستقل ويستخدم أيضاً في التنبؤ بقيم المتغير التابع نتيجة للتغير الحاصل في المتغير المستقل. وللتمييز بين الارتباط والانحدار بشكل أساسي فالارتباط يبين قوة العلاقة بين متغيرين أما الانحدار فيدرس إمكانية التنبؤ من المتغير المستقل بالمتغير التابع . (أبو زيد ، ٢٠٠٥) . أما في تحليل الانحدار المتعدد فهو يمكن الباحث من تحليل العلاقات بين متغير تابع (لا بد أن يكون كمياً) ومتغيرين مستقلين أو أكثر ، والتنبؤ بقيم المتغير التابع بمعلومية قيم المتغيرات المستقلة . وبالطبع يكون التنبؤ باستخدام المتغيرات المستقلة مجتمعة أفضل من التنبؤ باستخدام أي منها على حدة بشرط أن يكون الارتباط بين هذه المتغيرات المستقلة منخفضاً وأرتباط كل منها بالمتغير التابع مرتفعاً.

أما الاختبارات اللامعلمية فهي تستخدم في الحالات التي لا يكون فيها نوع التوزيع الاحتمالي للمجتمع الأصل الذي سحبت منه العينة معروفاً أو في حالة عدم استيفاء شرط كون التوزيع النظري للمجتمع طبيعياً . وهناك العديد من الاختبارات اللامعلمية التي تستخدم في التحقق من صحة الفروض الإحصائية التي لا تتأثر بشكل التوزيع للمجتمع الأصل ولا بضرورة الاختيار العشوائي للعينة المستخدمة في البحث. ومن هذه الاختبارات اللامعلمية ما توصل إليه كارل بيرسون عام ١٩٠٠ حيث توصل لأسلوب مربع كاى لاختبار الفروض وهو طريقة لتحديد ما إذا كانت الفروق بين التكرارات النظرية والملاحظة في أي عدد من الأقسام يمكن إرجاعها منطقياً إلى اختلافات في اختيار العينات (سميث ،١٩٨٥) . ويرتبط توزيع χ^2 بدرجة كبيرة بعدد درجات الحرية . ويعتبر هذا التوزيع متصلاً ذا قمة واحدة ، كما هو الحال مع التوزيع

الطبيعي وتوزيع t ، ولكنهما يختلفان عن توزيع χ^2 في أنهما متماثلان حول وسطهما الحسابي μ . ويعتبر توزيع χ^2 توزيعاً ملتوياً جهة اليمين و خصوصاً عندما تكون درجات الحرية df صغيرة (عودة ، القاضي ، ٢٠٠٢) .

ومن ضمن الاختبارات اللامعلمية ما قام به كل من عالمي الرياضيات الروسيين كولموجروف وسميرنوف وذلك بتطوير اختبار يمكن أن يستخدمه الباحث في التحقق من حسن المطابقة بين توزيعيين أحدهما توزيع تجريبي لدرجات عينة و الآخر توزيع نظري محدد . ونعنى بذلك التحقق مما إذا كانت درجات التوزيع التجريبي تعد بمثابة عينة مستمدة من المجتمع الذي تتوزع فيه الدرجات بحسب التوزيع النظري المحدد. ويعتمد هذا الاختبار على المقارنة بين التوزيع التكراري المتجمع الذي يمكن الحصول عليه من التوزيع النظري. والتوزيع التكراري المتجمع التجريبي وذلك لتحديد أكبر اختلاف بينهما واختبار ما إذا كان هذا الاختلاف يمكن عزوه إلى الصدفة (علام ، ١٩٩٣)

ومن أكثر الاختبارات اللامعلمية استخداماً اختبار " مان – ويتني" لعينتين مستقلتين وخاصة في البحوث والدراسات التي يكون فيها المتغير التابع من النوع الرتبي أو يكون من النوع الفتري . ويريد الباحث تجنب الافتراضات المتعلقة بالاختبار المعلمي المقابل أي اختبار لعينتين مستقلتين كاعتدال توزيع المجتمعين أو تجانس التباين فيهما. الافتراض الوحيد لهذا الاختبار هو عدم وجود رتب مكررة في البيانات (ماضي ، عثمان ، ٢٠٠٥).

ويُعد الإحصاء المعلمى أدق وأكثر كفاءة من الإحصاء اللامعلمى ، كما أنه أكثر حساسية لخصائص البيانات التي يتم جمعها ، كما أنه يوفر فرصة ضئيلة لحدوث الخطأ من النوع الأول Type I error والخطأ من النوع الثانى Type II error ، ويؤخذ على الاختبارات المعلمية بأنها أكثر صعوبة عند حسابها ، بالإضافة إلى محدودية نوع البيانات التي يمكن اختبارها بواسطة تلك الاختبارات و تستغرق وقتاً وجهداً في تطبيقها ولكن مع توفر البرامج الإحصائية مثل SPSS أمكن التغلب على هذه السلبيات.

حجم التأثير Effect Size

المتتبع لكثير من البحوث والدراسات يجد بأن بعض الباحثين يعتمدون في تقرير نتائجهم على الدلالة الإحصائية دون محاولة الكشف عن مقدار العلاقة القائمة بين المتغيرين، وتصبح هناك مغالاة في تفسير النتائج اعتماداً على مستوى الدلالة على الرغم من أنه ربما لا تكون لها قيمة من الناحية التطبيقية أو العملية، ولذلك فإذا وجد الباحث أن القيمة دالة إحصائياً، فمعنى ذلك أن المتغير المستقل له تأثير على المتغير التابع ، ولكنه لا يدل على حجم التأثير أو درجة العلاقة بين المتغيرين . وربما كانت الدلالة الإحصائية هنا لا تعني وجود علاقة قوية بين المتغيرين (الشربيني ، ١٩٩٥).

ويعد مفهوم حجم التأثير Effect Size من الأساليب التي يتم من خلالها معرفة حجم الفرق أو حجم العلاقة بين متغيرين أو أكثر. ويسمى أحيانا الدلالة العملية Practical Significance ويطلق عليه أيضا مقاييس قوة التأثير Strength of effect Measures، وقوة الترابط Association . (أبو حطب وصادق ، ١٩٩١).

ويعّرف (Cohen,1992) حجم التأثير على أنه " درجة تواجد ظاهرة ما في مجتمع محدد أو هو درجة احتمالية أن تكون الفرضية الصفرية خاطئة ، بينما يعّرفه (نصار ،٢٠٠٢) على أنه " قيمة كمية تشير إلى درجة العلاقة بين متغيرات الدراسة ضمن مجتمع محدد مسبقاً هو مجتمع الدراسة ، بغض النظر عن الأسلوب الإحصائى المستخدم للتعبير عن ذلك التأثير " . و يعّرفه هاس Hays بأنه " مقدار التباين في المتغير التابع الذي يمكن تفسيره من خلال متغير مستقل يرتبط به أو يؤثر فيه " (النجار ، ٢٠٠٥).

ويشير حجم التأثير إلى مدى قوة أو ضعف الفرض الصفري . وعندما تتوافق البيانات (التي تم الحصول عليها من خلال التجريب) تماماً مع الفرض الصفري ، تصبح القيمة المناظرة لحجم التأثير = صفر . وتبتعد قيمة حجم التأثير عن الصفر ، عندما ينتج عن التجربة بيانات غير متوافقة مع الفرض الصفري(Snyder & Lawson,1993) ويرى Robey,2004 بأن حجم التأثير يشير إلى مدى قابلية تفسير المتغير التابع في ضوء المتغير المستقل .

ويلخص منصور (١٩٩٧) الآراء المختلفة حول مفهوم حجم التأثير في النقاط التالية:

أ. أن أساليب حجم التأثير هي الوجة المكمل للاختبارات الإحصائية بمستويات دلالاتها المختلفة . فالأولى تتحدث عن حجم التأثيرات وقوة الترابط بصرف النظر عن الثقة التي توضع في هذا الحجم. والثانية تعبرعن الثقة في النتائج بصرف النظر عن حجمها.

ب. أن أساليب حجم التأثير لا تتأثر بحجم العينة على حين نجد أن مستويات الدلالة الإحصائية للاختبارات الإحصائية تتأثر بحجم العينة.

ت. أن معامل الارتباط هو حلقة الوصل بين الأسلوبين الدلالة الإحصائية وحجم التأثير، فحجم المعامل بصرف النظر عن حجم العينة المبني عليها هو حجم تأثير وقوة ترابط أما دلالاتها إحصائياً فتتوقف على علاقة هذا الحجم بمقدار حجم العينة.

ث. أن معظم اختبارات الدلالة الإحصائية يمكن تحويلها إلى مقاييس لحجم التأثير.

ولقد أدى قياس الدلالة الإحصائية إلى تفكير الباحثين في معنى النتائج التي يتم التوصل إليها ، وأن الاقتصار على استخدام قياس الدلالة الإحصائية في اتخاذ القرارات ينجم عنه الكثير من المشكلات ، فالاعتماد البسيط عليها و زيادة استخدام تحليل القوة وتقديرات حجم التأثير قد يسهم في التحسن الكلي لجودة البحوث . كما أن الباحثين الأكثر خبرة بفوائد و عيوب الدلالة الإحصائية وحجم التأثير وتحليل القوة هم من المحتمل أن يكونوا أكثر قدرة على صياغة تفسيرات مفيدة حول جدوى نتائج بحوثهم (Huston,1993) .

ومن الملاحظ أن بعض الباحثين يعتمدون في تقرير نتائجهم على الدلالة الإحصائية دون محاولة الكشف عن مقدار العلاقة القائمة بين المتغيرين ، وتصبح هناك مغالاة في تفسير النتائج اعتماداً على مستوى الدلالة على الرغم من أنه ربما لا تكون لها قيمة من الناحية النتطبيقية أو العملية . ولذلك ، إذا وجد الباحث أن القيمة دالة إحصائياً ، فمعنى ذلك أن المتغير المستقل له تأثير على المتغير التابع ، و لكنه لا يدل على حجم التأثير أو درجة العلاقة بين المتغيرين ، وربما كانت الدلالة الإحصائية هنا لا تعنى وجود علاقة قوية بين المتغيرين (الشربينى ، ١٩٩٥).

ويمكن أن تكون الدلالة الإحصائية مضللة ، وأنه من المهم الرجوع إلى حجم التأثير في تقويم نتائج أية تجربة ، حيث تخبرنا أحجام التأثير عن مقدار التأثير في حين أن الدلالة الإحصائية لا تخبرنا عن ذلك (Wright,2003) .

ويعتمد الكثير من الباحثين على الاجراء الاحصائى التقليدى المتمثل في المقارنة بين المجموعة التجريبية (E) ونظيرتها الضابطة (C) عند فحص فعالية بعض المعالجات التطبيقية . ويعد حجم التأثير (ES) – أو متوسط الفروق المعيارية Standardized Mean Differences – بمثابة مؤشر لدرجة التأثير بغض النظر عن أداة القياس المستخدمة . ولأنه بارامتر طبيعى Natural Parameter ، يفيد ذلك المؤشر في تحقيق الدمج بين العديد من النتائج المستمدة من تجارب مختلفة. (Bayarri & Mayoral 2002) .

وخلال العقود الماضية ، ازداد عدد الأبحاث المنشورة التي تنتقد بدورها استخدامات الاختبارات الإحصائية ، وقد ظهر ذلك الاهتمام واضحا في الكثير من العلوم ، وبالأخص علم النفس (Anderson ,Burnham & Thompson,2000) . وفي نفس الوقت أكد عدد من الباحثين على أهمية الإشارة إلى حجوم التأثير و تفسيرها عند صياغة نتائج و بيانات الدراسة المراد نشرها في الدوريات العلمية (Huberty,2002).

واستجابة للمقالات النقدية المنشورة بخصوص اختبارات الدلالة الإحصائية قام مجلس إدارة الشئون العلمية برابطة علم النفس الأمريكي (APA) بتشكيل لجنة الاستدلال الاحصائى التي نشرت تقريرها (Wilkinson ,1999) حيث اشتمل على بعض التوصيات المرتبطة بالمناقشة الحالية لكيفية حساب أحجام التأثير، ويؤكد الاصدار الخامس من دليل الرابطة الأمريكية لعلم النفس (APA 2001) على أهمية كتابة التقارير بشأن حجم التأثير:

"لكى يفهم القارىء تماماً أهمية نتائج الدراسة العلمية لابد من احتوائها على بعض مؤشرات حجم التأثير أو قوة العلاقة في تلك النتائج . ويمكن تقدير حجم ذلك التأثير أو قوة هذه العلاقة بواسطة العديد من الطرق والأساليب الاحصائية . والقاعدة الأساسية هي عدم الاكتفاء بتزويد القارىء بمعلومات عن الدلالة الاحصائية فقط ، بل يجب تقديم المزيد من المعلومات التي تساهم بدورها في قياس حجم التأثير أو العلاقة الملحوظة "

وحالياً تتطلب سياسات التحرير لـ ١٧ من الدوريات العلمية التي تنشرها رابطة APA كتابة تقارير بشأن حجوم التأثير بما فيها دورية Educational and Pyschological Measurement وغيرها من الدوريات الأخرى التي تنشرها رابطة الارشاد النفسى الامريكية American Counseling Association و الرابطة الامريكية للأبحاث التربوية American Educational Research Association ومجلس الأطفال الموهوبين Counsil for Exceptional Children. (Bird,2002)

ولقد أورد هوستن Huston عدداً من المميزات لمقاييس الدلالة العملية (حجم التأثير) (محمود ٢٠٠٣) :

- تعطي تقديراً لدرجة وجود الظاهرة في المجتمع وذلك من خلال قيمة على مقياس متصل يبدأ من صفر حقيقي (يعني غياب الظاهرة) ويمتد ليقدم قيمة لحجم التأثير لهذه الظاهرة .
- تعطي قيمة محددة تشكل محكاً لوصف حجم التأثير بأنه صغير أو متوسط أو كبير
- يمكن استخدام نتائجها كمقياس كمي للمقارنة بين دراستين.
- يمكن استخدامها في تحليل القوة لتقدير حجم العينة اللازمة لدراسة ما.

كما أضاف كل من باركر وهاجان 2007 (Parker and Hagan) مميزات أخرى لحجم التأثير:

١. مقياس موضوعى لقوة التدخل Intervention Strength و ذلك في حالة صدق وقوة التصميم التجريبى.
٢. مؤشر له تدريج مستمر , Continuous Scale وبالتالى يدعم القرارات التدريجية في المعالجة الاحصائية .
٣. مقياس دقيق خصوصاً في حالة النتائج الصغيرة و غير الواضحة.
٤. ملخص موضوعى في حالة عدم الاتفاق على الأحكام البصرية بشأن الرسوم البيانية.
٥. طريقة للمقارنة بين النجاحات المختلفة للتدخل عبر أبحاث دراسة الحالة ، سواء كان التحليل أنياً أو بعدياً.
٦. مقياس معتمد في أبحاث دراسة الحالة بالمقارنةً مع التقاليد البحثية الأخرى.
٧. طريقة فعالة لتوثيق النتائج.
٨. مقياس جيد لاستقلالية النتائج ، خصوصاً عند استخدام فترات الثقة .

اتجه بعض الباحثين في الإحصاء التربوي والنفسي إلى صياغة مؤشرات إحصائية خاصة بحساب الدلالة العملية وذلك بقصد تعزيز القرار الإحصائي وتحسين من جودة البحوث، ولقد قام ليمور وروتو Elmore and Rotou (2001) مؤخراً بتسمية ٦١ من أحجام التأثير في الدراسات السلوكية المختلفة .

وتم توصيف الكثير من مؤشرات حجم التأثير. و مع ذلك في التاريخ الحديث لنظريات الإحصاء بُذلت جهود أكبر في الحصول على قيم P values دقيقة مما تم تخصيصه بشأن تطوير مؤشرات حجم التأثير . وهكذا ظلت هناك العديد من الإجراءات الإحصائية التي لم تساهم في تحديد مؤشر قياسى أو نموذجى لحجم التأثير (كما هو الحال في الكثير من الاجراءات الخالية من التوزيعات الحرة Distribution-free أو تلك اللابارامترية) (Rosnow and Rosenthal and Rubin ,2000) .

ويتفق كل من ماكسويل و ديلانى (Maxwell & Delaney 2004) و(أبو هاشم ٢٠٠٥) أن هناك نوعين من مؤشرات حجم التأثير هما :

أ. المؤشرات التي تدل على مقدار التأثير في الدراسات التي تبحث الفروق وتهتم هذه المؤشرات بالتأثير الذي يحدثه المتغير المستقل في المتغير التابع ، وتستخدم الاختبارات الإحصائية التي تبحث بدورها الفروق بين متوسطات درجات المجموعات . و تتصف بأن لها بنية مشتركة تتمثل في التغير نتيجة الفروض الخاصة بالتباين في ضوء الخطأ المعيارى . ومن أمثلتها : { مؤشر كوهن f^2 (Cohen,1988) ، مؤشر جلاس Δ (Glass,1976) ، مؤشر هدجز g (Hedges,1981) ، مؤشر كوهن d (Cohen,1988) }.

ب. المؤشرات التي تدل على مقدار التأثير في الدراسات الارتباطية و التنبؤية على الرغم من وجود أنواع مختلفة من معاملات الارتباط نتيجة اختلاف نوع المعامل المستخدم تبعاً لطبيعة البيانات ، إلا أن معامل ارتباط بيرسون هو الأكثر استخداماً في الدراسات التربوية والنفسية ، ويستخدم لمعرفة العلاقة الخطية بين متغيرين متصلين وحدوده بين (+١، -١) ، والاشارة التي تسبق قيم هذا المعامل تدل على اتجاه العلاقة و ليس قوة العلاقة . و لها بنية مشتركة تتمثل في التغيرات الناتجة عن معادلة R^2التي قد يشار إليها أحياناً بمصطلح مُقدِرات نسبة التباين أو المقاييس المرتكزة على التباين . ومن أمثلتها : { مؤشر مربع إيتا لبيرسون η^2 (Pearson, 1905) ، مؤشر مربع أوميغا لهاس ω^2 (Hays,1963) ،مؤشر مربع أبيلسون لكيلي ε^2 (Kelley,1935) } .

مؤشر كوهين Cohen's d

يعّرف Cohen مؤشر حجم التأثير d على أنه " الفرق بين متوسطي مجموعتين M1 و M2 مقسوماً على الانحراف المعياري σ داخل المجموعتين " و يرى كوهن أنه يمكن استخدام الانحراف المعياري لأى مجموعة عندما تكون التباينات لكل منهما متجانسة . (Sanchez and Martinez and Chacon 2003)

وينتشر استخدام مؤشر Cohen's d بشكل خاص في الأبحاث التجريبية و شبه التجريبية حيث يكون الفرق في تأثير المعالجة التجريبية هاماً بالنسبة لمتوسطات المجموعات (سواء تجريبية أو ضابطة) . ويشتمل متوسط الفروق المعيارية على مجموعة من الإحصاءات المرتبطة ببعضها البعض ، وبالتالي هناك صعوبة في تقديم تلك الإحصاءات نتيجة الاختلافات في الرموز التي يستخدمها الباحثون وربما يعكس نقص المعيارية Standardization في الأبحاث ذلك الاهتمام الحالي بتلك الإحصاءات (Mcgrath & Meyer,2006)، ويختلف حساب حجم التأثير باختلاف الإحصائي المستخدم في فحص الفرضية الصفرية . وقد أوجد كوهن Cohen طرقاً مختلفة لحساب حجم التأثير لـ الإحصائي ت (t) والإحصائي ف (f) (إبراهيم ، ١٩٩٨) .

ومن الأمثلة المعروفة للباحثين في مجال التربية و علم النفس حالة العينتين المستقلتين لاختبار (ت) مثل التصميم التجريبي الذي يتضمن مجموعة ضابطة Control Group وأخرى تجريبية Experimental Group .

مربع ايتا ^{2} Eta

تعتبر نسبة الارتباط (Eta^2) أو مربع ايتا من المؤشرات التي تقيس حجم تأثير المعالجة ويستخدم هذا المعامل في حالة وجود علاقة منحنية بين المتغيرين أى أن الخط الأكثر ملائمة للبيانات ليس خطاً مستقيماً . ويفترض مربع آيتا أن خط الانحدار يمر من خلال كل متوسط من متوسطات المعالجة ، وهذا الافتراض صحيح عندما يتم معالجة البيانات كمجتمع ، ولكن عندما يتم التعامل مع البيانات كعينة من مجتمع ، فأن نسبة الارتباط تعتبر متحيزة ، ذلك لأن متوسطات العينات خاضعة لما يسمى بخطأ المعاينة ، و هنا احتمال أن لا يمر الخط بالضبط من خلال متوسطات العينات (**المنيزل** ، ٢٠٠٠) .

ومربع ايتا يمثل النسبة من التباين الكلي للمتغير التابع في العينات موضع البحث التي ترجع إلى أثر المتغير المستقل . وتساوى قيمته في تحليل التباين الأحادي للبيانات المستقلة بأنه مجموع المربعات بين المجموعات إلى مجموع المربعات الكلية (Hasse & Solomon and waechter,1982) .

وقد ذكرت الصائغ (١٩٩٦) في دراستها مميزات مربع ايتا منها :

١- يعتبر مقياساً مقبولاً و خصوصاً في البحوث التربوية لحساب الدلالة العملية أو حجم التأثير والمسمى بالتباين المفسر ، ويكون حسابها بدون صعوبة من جدول ANOVA الكامل في حالة اختبار تحليل التباين الأحادي .

٢- أن معظم الإحصائيين يوافقون على أن مربع ايتا يعطى معلومات أكثر دقة من قيمة الفرق الإحصائي حيث تأخذ في الاعتبار حجم العينة .

٣- تظهر أهميته وفائدته عند مراجعة الدراسات السابقة باستخدام ما يسمى بأسلوب ما وراء التحليل أو التحليل البعدي وهي طريقة كمية لمراجعة الدراسات السابقة في أى مجال من المجالات بطريقة تكاملية ، وبأقل قدر من التحيز سواء كان هذا التحيز مراجعه ذاتية الباحث الذي يقوم بالمراجعة أو تحيز يرجع للدراسات السابقة موضع المراجعة.

٤- يبقى لقيمة مربع آيتا دور كبير في تقييم الدلالة العملية وفي بيان درجة التنبؤ الفعلي للدراسة حتى لو كانت قيمتها منخفضة للغاية .

٥- أن قيمة مربع آيتا تساوى قيمة معامل التحديد R^2 المستخرج من نموذج تحليل الانحدار الذي يتم استخدامه عند تحليل الاختلاف في تصميم يشبه تحليل ANOVA

٦- يشكل مربع آيتا رابطة قوية بين التحليل التقليدي للتباين من ناحية وبين الانحدار المتعدد من ناحية أخرى .

٧- يستخدام مربع آيتا في تفسير أهمية البحث وفي قياس الفروق التي يظهرها التحليل الإحصائي وقد أصبح استخدامه شائعاً خاصة في مجال العلوم الاجتماعية نظراً للتشابه مع معامل التحديد R^2 الأكثر فائدة و الناتج من تحليل الانحدار.

مربع أوميجا ω^2

يشير الكيلانى والشريفين (٢٠٠٧) إلى مربع اوميجا (ω^2) على أنه النسبة من التباين الكلي التي يمكن تفسيرها في المتغير التابع في المجتمع الأصلي الذي اختيرت منه العينة ، ويعتبر معلمياً وينتمى إلى الإحصاء الاستدلالي أي عكس مربع آيتا الذي ينتمى بدوره إلى الإحصاء الوصفي .

ويشير المنيزل (٢٠٠٠) إلى أن حساب قيمة مربع أوميجا تعتمد فيما إذا كان المتغير المستقل الذي نتعامل معه عشوائي أو ثابت، وفي حالة المتغير العشوائي ، يمكن تعريف مربع أوميجا على أنها نسبة تباين المعالجة إلى مجموع تباين المعالجة والخطأ ، وبما أننا لا نستطيع قياس تباين المعالجة وتباين الخطأ بشكل مباشر ، فأننا نقدر ذلك من خلال العينات ، لذلك فأن قيمة مربع أوميجا المتوقعة هي بمثابة قيمة تقريبية.

مؤشر Hedges' g

يعتبر g بمثابة مقياس استدلالي ، و يمكن حسابه من خلال الجذر التربيعي لمتوسط مربع الخطأ MSE الناتج عن تحليل التباين ANOVA (الذي يختبر الفروق بين المجموعتين المستقلتين) . و تعود تسمية هذا المؤشر إلى جلاس Glass التي تعتبر رائدة التحليل البعدي. ويمكن حساب قيمة g من خلال قيمة اختبار ت t-test للفروق بين المجموعتين المستقلتين (في حالة عدم تساوى عدد المفحوصين في المجموعتين).

مربع ابيسلون ε^2:

أوجد هذا المؤشر كيلى (Kelly,1935) ، ثم عدله كوهين (Cohen,1965) ، ومما يجدر ذكره أن استخدامات مربع ابيسلون كمؤشر للدلالة العملية غير شائع الاستخدام في البحوث النفسية و التربوية ، و ذلك لكونه غير مألوف لكثير من الباحثين.

وتوجد مجموعة أخرى من مؤشرات حجم التأثير كما أوردها كل من (حجيمات ،١٩٩٥ ، ابراهيم ،١٩٩٨، محمود ، ٢٠٠٣، سكران ،٢٠٠٦ ، النجار ٢٠٠٥)، ومنها حجم التأثير لإختبار كا^2 ، حيث يستخدم لحساب حجم التأثير (W) من خلال حالة اختبار كا^2 لفحص جودة المطابقة، كما يمكن حساب مؤشر حجم التأثير لاختبار كا^2 من خلال معامل التوافق الذي يعد بدوره من أكثر المقاييس استخداماً لقياس الارتباط في جداول التوافق، كما يمكن كذلك حساب مؤشر حجم التأثير لاختبار كا^2 باستخدام معامل فاى ϕ ، ويمكن حساب مؤشر حجم التأثير لاختبار كا^2 باستخدام معامل كرامر.

حجم التأثير r^2 لمعامل ارتباط بيرسون

تستخدم في الدراسات الارتباطية في بحث العلاقة بين متغيرين أو أكثر ، ومن أشهر معاملات الارتباط المستخدمة معامل ارتباط بيرسون (r) حيث يعتبر مؤشراً لوجود علاقة خطية بين المتغيرين و تأخذ قيم بين (+١ و −١) بما فيها (−١ و +١) ، أما مؤشر حجم التأثير فيساوى مربع معامل ارتباط بيرسون (R^2) (أى معامل التحديد) و هو يمثل نسبة التباين في أحد المتغيرين الذي يمكن عزوه إلى تباين المتغير الآخر، ومنها **حجم التأثير f^2** الخاص بتحليل الانحدار المتعدد Multiple Regression حيث يتم دراسة العلاقة بين مجموعة من المتغيرات التي غالباً ما تكون متغيرات، ،ايضاً **حجم التأثير q** للفرق بين معاملي الارتباط ، حيث يتم اختبار الفرق بين معاملي الارتباط لمجتمعين من خلال معاملي الارتباط لعينتين مسحوبتين من هذين المجتمعين ، وهو يقوم على افتراض التوزيع الطبيعي لكلا المجتمعين وتجانس التباين . وبما أن توزيع r غير متساوي الوحدات فأن الفرق بين r1 و r2 لا يمكن استخدامه كمؤشر لحجم التأثير ، فنلجأ لتحويل قيم r إلى Fisher z (أي تحويلها من مستوى رتبى إلى مستوى فئوي).

وهناك **أيضاً حجم التأثير D^2** لاختبار تحليل التباين المتعدد MANOVA ، و فيه يتم فحص متوسطات المجموعات لجميع المتغيرات ما اذا كانت متساوية، حجم التأثير من خلال معامل الارتباط الثنائي الأصيل.

ويمكن تقدير حجم التأثير المتوقع من خلال ثلاث طرق (نصار ، ٢٠٠٢) :

١) مراجعة الأدب المتخصص وذلك من خلال مراجعة الدراسات السابقة التي تمت حول موضوع البحث ومن خلال استعراض نتائج تلك الأبحاث يستطيع الباحث بناء توقع حول مقدار حجم التأثير خاصةً إذا أشارت تلك الدراسات مباشرةً إلى قيمة حجم التأثير التي ظهرت من خلال بياناتها .

٢) تحديد مقدار الفارق الدال عملياً من وجهة نظر الباحث و ذلك بين متوسطي المجموعتين التجريبية والضابطة في حالة وجود عينتين أو الفارق الدال عملياً بين القياسين القبلي والبعدي في حالة اختبار "ت " للعينات المترابطة أو الفارق الدال عملياً بين متوسط العينة ومتوسط مجتمع معين في حالة اختبار" ت " للعينة الواحدة أو مقدار الترابط الدال عملياً بين تباين المتغير المستقل وتباين المتغير التابع في حالة تحليل التباين . ومن العوامل التي تلعب دوراً في تحديد الباحث لمقدار الأثر المرغوب به هي دعم نظرية معينة أو إثبات فعالية برنامج معين في تحقيق أهداف محددة وكذلك كلفة هذا البرنامج حيث كلما ازدادت الكلفة المادية للبرنامج المقترح من قبل الباحثين كلما زادت الحاجة إلى اثبات الأهمية أو الفائدة العملية لهذا البرنامج .

٣) استخدام المعايير التي أشار إليها كوهين ففي حالة اختبار " ت " أشار كوهين إلى أن حجم التأثير يعتبر صغيراً عند (٠,٢٠) ومتوسطاً عند (٠,٥٠) وكبيراً عند (٠,٨٠) ومن خلال ذلك يستطيع الباحث أن يحدد حجم التأثير الذي يسعى إليه في ضوء قوة النظرية التي يستند عليها البحث وقوة العلاقة بين متغيرات الدراسة .

الدراسات السابقة:

تناولت دراسة سميثون (Smithson, 2001) أحجام التأثير من خلال مربع معامل الارتباط R^2 والذي يستخدم في تحليل الانحدار المتعدد وبخاصة في التوزيعات اللامركزية مثل (F) و (X^2) وترى ضرورة وضع قيمة R في الاعتبار عند استخدام التنبؤات من خلال معاملات الانحدار وانتهت إلى أن قيمة حجم التأثير متوسطة بالإضافة إلى وجود علاقة قوية بين حجم التأثير وفترات الثقة باستخدام التوزيعات اللامركزية.

وهدفت دراسة نصار (٢٠٠٢) إلى تقديم مفهوم حجم الأثر كأسلوب إحصائي مكمل لفحص الفرضيات الإحصائية وتضمنت تحليلاً رقمياً لبرهنة أن حجم الأثر أقل تأثراً بحجم العينة من الأساليب الإحصائية التي تستخدم لفحص الفرضيات. بالإضافة إلى بعض المؤشرات الاحصائية التي تستخدم للدلالة على قيمة حجم الأثر في حالة استخدام بعض الأساليب وبالتحديد اختبار (t) سواء في العينات المترابطة أو المستقلة واختبار (F) في حالة تحليل التباين الأحادي ومعامل ارتباط بيرسون وأخيراً الأنحدار الخطي البسيط ، وقدمت الدراسة نموذجاً حول كيفية وصف مقدار حجم الأثر في الدراسات المنشورة في المجلات العربية المتخصصة وبالتحديد في مجلة جامعة الملك سعود . وأظهرت نتائجها أنه يمكن وصف ما نسبته ٦٠% من قيم حجم الأثر والمستخرجة للأساليب الإحصائية المنشورة خلال خمس سنوات على أنه حجم أثر صغير أو متوسط.

وتهدف دراسة روبين وروسنثال (Rubin & Rosenthal , 2003) إلى اقتراح مُقدِر بسيط لحجم التأثير Simple Size Effect Estimate (يتم حسابه من خلال حجم العينة N ، وقيمة p value) وذلك بغرض استخدامه في : أ) دراسات التحليل البعدى حيث يقرر الباحث نفس أحجام العينة وقيم P values فقط ، ب) في الحالات التي لايتوافر فيها مؤشرلحجم التأثير متفق عليه بشكل عام ، ج) في الحالات التي تكون فيها مؤشرات حجم التأثير التي تم حسابها بشكل مباشر مضللة Misleading . ويُطلق على مؤشر حجم التأثير المستخدم في الدراسة مصطلح R المكافئة R equivalent وذلك لأنها تتساوى إلى حد كبير مع الارتباط النقطي ذي الحدين بين مؤشر المعالجة Treatment Indicator والناتج ذي التوزيع الاعتدالي Normaly Distributed Outcome في التجارب البحثية التي بها اثنتين من المعالجات Two –treatment Experiments حيث يوجد N/2 من الوحدات في كل مجموعة ، إلى جانب قيمة p value . ويؤكد المؤلفان على أنه رغم قابلية المكافئة (R) للحساب إلا أنها ليست فعالة في كل الأحوال . لقد كان الهدف من تصميم مؤشر R equivalent لتلائم المواقف التالية :-

١) عندما يكون البديل المتاح هو عدم وجود حجم تأثير على الاطلاق (في حالة معرفة أحجام العينات و قيم p values فقط.
٢) استخدام اجراءات لا معلمية ليس لها مؤشرات أحجام تأثير متفق عليها .
٣) عندما تكون أحجام العينات صغيرة جداً أو تكون البيانات غير اعتدالية مما يجعل الحساب المباشر لأحجام التأثير أكثر تضليلاً من قيمته الحقيقية .

وقد هدفت دراسة محمود (٢٠٠٣) إلى التعرف على واقع الدلالة الإحصائية ، والدلالة العملية وقوة الاختبار للاختبارات الإحصائية المستخدمة في رسائل ماجستير كلية العلوم التربوية . وبالتحديد ، فقد حاولت الدراسة الإجابة على الأسئلة التالية :

١) ما مستوى الدلالة الإحصائية المستخدم في فحص الفرضيات الصفرية في هذه الدراسات ؟
٢) ما نسبة الاختبارات الإحصائية التي اقترنت دلالتها الإحصائية بدلالة عملية ؟
٣) ما مستوى قوة الاختبارات الإحصائية المستخدمة في فحص الفرضيات الصفرية ؟
٤) ما نسبة الفرضيات التي اقترنت دلالتها الإحصائية باختبارات ذات قوة عالية؟ . وقد أظهرت النتائج ميل معظم الباحثين لاستخدام مستوى الدلالة ٠,٠٥ ، وقد بلغت نسبة الاختبارات الدالة إحصائياً (٤٥,٢ %) من مجموع الاختبارات ، تمتع منها (١٧,٨ %) بدلالة عملية كبيرة و(١١,٧ %) منها بدلالة عملية متوسطة و(٧١,٣ %) بدلالة عملية صغيرة . وقد أظهر اختبار مربع كاى عدم استقلالية الدلالة العملية عن الدلالة الإحصائية.

وهدفت دراسة روبي (Robey, 2004) إلى :

(أ) مراجعة حجم التأثير في العلوم الاجتماعية.

(ب) توضيح أهمية حجم التأثير في بحوث علم النفس العيادي.

(ج) وضع معيار أو محك يستند إليه عند قبول أو رفض حجم التأثير. وأظهرت النتائج وجود أحجام تأثير بنسب متفاوتة في بحوث علم النفس العيادي التي تستخدم تحليل التباين الأحادي أو الثنائي أو التصميمات التجريبية ذات المجموعة التجريبية مقابل الضابطة.

وأجرى كل من اجينيس وبيتي (Aguinis and Beaty , 2005) مراجعة للدراسات عبر ٣٠ عاماً (١٩٦٩–١٩٩٨) في الدوريات العلمية التالية:

- Journal of Applied Psychology
- Academy of Management Journal
- Personal Psychology

حيث أنها من أكثر الدوريات العلمية تأثيراً في مجال علم النفس التطبيقي وعلوم الإدارة فضلاً عن ذلك تمتاز تلك الدوريات بوضع سياسات تحرير صارمة تلزم الباحثين بإتباع طرق بحثية و منهجية قياسية . وهكذا نلاحظ أن أحجام التأثير في تلك الدوريات محايدة . وبمعنى آخر نتيجة لسياسات التحرير و تقرير نتائج الأبحاث المستخدمة في تلك الدوريات من المتوقع أن تكون أحجام التأثير المنشورة فيها أكبر من تلك المنشورة في الدوريات الأخرى في مجال علم النفس التطبيقي والعلوم الإدارية من أجل قياس أحجام تأثير العوامل المعدلة Moderating Effects للمتغيرات الفئوية التي يتم قياسها بواسطة معادلة الانحدار المتعدد . وقد كان متوسط حجم التأثير، والذي يرمز إليه بـ f^2 مساوياً لـ(٠٠٢.) فقط ، و لكن مع ذلك أظهرت (٧٢ %) من اختبارات العامل المعدل Moderators قوة إحصائية مقدارها (٨٠.) أو أكثر مما أدى بدوره إلى التنبؤ بوجود حجم صغير لهذا التأثير . وتشير النتائج إلى أن هناك حاجة إلى التخلص من القيم التي تعمل بدورها على وجود انحياز منخفض في حجم التأثير الملحوظ ، وبالتالي يمكن تعريف أحجام التأثير ذات الصلة بالعلاقات بين المتغيرات بشكل أكثر دقة . وطالما كان للتأثير دلالة قوية ينصح المؤلفون الباحثون في مجال العلوم السلوكية وعلوم الإدارة بإجراء تحليل القوة وتخطيط التصميمات البحثية المستقبلية التي ترتكز على أحجام التأثير الصغيرة و لكنها الأكثر واقعية .

يطرح كل من ماك جراث و مير (mcgrath & Meyer ,2006) الاستخدام المتنامي لأحجام التأثير في دراسات الحالة و التحليلات البعدية تساؤلات جديدة بشأن الاستدلال الإحصائي . وقد يؤثر اختيار مؤشر حجم التأثير على تفسير نتائج الدراسات ويوضح الباحثان بالأدلة العملية ذلك التأثير من خلال التركيز على اثنين من أكثر مقاييس حجم التأثير استخداما : معامل الارتباط ، وفرق المتوسطات المعيارية (مثل Cohen's d ، و Hedges's g) ، والتي يتم استخدامها في حالة وجود متغير ثنائى Dichotomous و آخر كمي Quantitative . ورغم قابلية استخدام تلك المؤشرات بنفس الأسلوب من الناحية العملية ، فأن الفروق في حساسية تلك المؤشرات أو تباين المتغير الثنائي قد تؤدى إلى تغيير النتائج بشأن حجم التأثير وذلك في ضوء الاحصاء المستخدم . ولأنه لا توجد احصاء جيد ومناسب بشكل عام متفق عليه لدى جميع الباحثين ، فإنه يجب الاهتمام بمعدل الفترة الأساسية عند صياغة الاستدلالات الصحيحة و تبرير اختيار الاحصاء الأساسي لقياس حجم التأثير .

وهدفت دراسة ويلكوكس (Wilcox, 2006) إلى استخدام الطرق البيانية لقياس حجم التأثير كبدائل لمدخل كوهين وتوصلت إلى أن حجم التأثير يلعب دوراً أساسياً في البحث النفسي والتربوي، وباستخدام الطرق البيانية Graphical Methods كبدائل لمدخل كوهين لا توجد فروق كبيرة وبخاصة في بعض الاختبارات اللامعلمية مثل مان – ويتني ، وهذه التمثيلات البيانية تتغلب على عيوب مدخل كوهين في تحديد حجم التأثير. ويرى ويلكوكس أن الهدف الأساسي في مجال الأبحاث التطبيقية ليس مجرد ايجاد الفرق بين مجموعتين من المشاركين ، ولكن الشيء الأكثر أهمية هو فهم كيفية و مدى الاختلاف بين هاتين المجموعتين . وقد صاغ Cohen مدخلاً بحثياً عاما ومنطقياً لـ علاج تلك المسألة . وقد اعتمدت استراتيجيته في القياس على المنظور البيانى . وبشكل أكثر تحديداً عندما نفحص التوزيعين Distributions المرتبطة بـ تلك المجموعتين ، افترض Cohen وجود تأثير كبير إذا كان الفرق بين هذين التوزيعين (ويشار إليه من خلال التغير Shift في الموضع) " يُرى بالعين المجردة " . وفي نهاية ألأمر توصل Cohen إلى أنه بالنسبة لـ التوزيعات الطبيعية التي بينها تباين مشترك ، يكون الفرق d صغيراً إذا كان مساويا لـ(2,) وكبيراً إذا كان مساوياً لـ(8,) وحالياً هناك مسألة بحثية شائكة بشأن الفرق d الذي ربما يصبح صغيراً عندما يكون حجم التأثير كبيراً (في ضوء المنظور البياني لـ Cohen) . وهنا يهدف ويلكوكسن إلى توضيح الطريقة البيانية المستخدمة في علاج تلك المشكلة ، إلى جانب غيرها من الطرق البيانية التي قد تستخدم في إضافة المزيد من الأفكار بشأن قياس حجم التأثير .

ومن خلال استعراض الباحث للدراسات السابقة وجد :-

- أن معظم الدراسات وخاصة الأجنبية تناولت أهمية حجم التأثير وأثره في دقة نتائج القرارات الإحصائية وأوصت بضرورة عدم الاعتماد فقط على الدلالة الإحصائية دون الدلالة العملية .
- اقتصرت الدراسات السابقة على عدد محدود من الاختبارات الإحصائية حيث نجد أن الدراسات أوجدت حجم التأثير المرتبط بالاختبارات الإحصائية (F , T , R , χ^2) فقط مثل دراسة (نصار ، ٢٠٠٢ و محمود ، ٢٠٠٣ و النجار ، ٢٠٠٥ ، Smithson (2001, أما الدراسة الحالية فسوف تتناول حجم التأثير للاختبارات الإحصائية المعلمية (اختبار ت ، وتحليل التباين ، والانحدار ، ومعامل ارتباط بيرسون) والاختبارات الإحصائية اللامعلمية (كا^2 ، إختبار مان – وتني) .
- هنالك خلط في بعض الدراسات العربية فيما يتعلق بالعلاقة بين حجم التأثير والدلالة الإحصائية حيث نجد بأن تلك الدراسات قامت بحساب حجم التأثير للاختبارات الإحصائية الغير دالة ، حيث يرى (الضوي ، ٢٠٠٦) أنه من الأخطاء الشائعة التي يجب أن يلتفت إليها الباحثون حساب حجم التأثير المناظر لقيم الاختبار الإحصائي بطريقة عفوية سواء أكانت قيم الاختبار الإحصائي دالة أم غير دالة ، والصواب فقط هو حساب أو تقدير حجم التأثير فقط في حالة القبول بالفرض البديل ورفض الفرض الصفري . والنقطة المهمة أن حجم التأثير يكون في صورته المثالية عندما يستخدم الباحث الاختبار الإحصائي المناسب لتحليل بياناته أي عندما تفي بياناته بالافتراضات الأساسية التي يستند إليها الاختبار بالإضافة إلى قبول الفرض البديل ورفض الفرض الصفري .
- عدم تناول الدراسات السابقة أثر المنهج المستخدم على حجم التأثير وهذا ما سوف تسعى الدراسة الحالية إلى إيجاده .
- قلة من الدراسات السابقة تناولت أثر حجم العينة على حجم التأثير مثل دراسة (Bayarri and Mayoral 2002) والتي أظهرت وجود تأثير لحجم العينة على دلالة أحجام التأثير . بالإضافة إلى دراسة الصائغ والتي أظهرت بأن حجم العينة كان مناسباً عند حجم التأثير الصغير . بينما دراسة نصار أظهرت بأن حجم التأثير أقل تأثراً بحجم العينة . وفي المقابل نجد بأن دراسة النجار ٢٠٠٥ أظهرت بأن زيادة حجم العينة قد يكون لة تأثير سالب على حجم التأثير .
- أن معظم الدراسات أكدت على انخفاض قيمة حجم التأثير في البحوث التربوية والنفسية لبعض الاختبارات الإحصائية وخاصة الدراسات العربية ، مثل (نصار ، ٢٠٠٢ ومحمود ، ٢٠٠٣ والنجار ، ٢٠٠٥) .

منهج الدراسة :

أعتمد الباحث في هذه الدراسة على المنهج المسحي الوصفي والذي يتم من خلاله جمع معلومات وبيانات عن ظاهرة ما أو حادث ما وذلك بقصد التعرف على الظاهرة التي ندرسها وتحديد الوضع الحالي لها والتعرف على جوانب القوة والضعف فيها من اجل معرفة مدى صلاحية هذا الوضع أو مدى الحاجة لإحداث تغيرات جزئية أو اساسية فيه.

عينة الدراسة:

تكونت عينة الدراسة من رسائل الماجستير بكلية التربية بأقسامه المختلفة (علم النفس – المناهج – الوسائل – الإدارة التربوية – التربية البدنية – التربية – التربية الخاصة) والتي استخدمت المنهج الوصفي والتجريبي في دراساتها ، حيث يبلغ العدد الإجمالي لرسائل الماجستير بكلية التربية بعد استبعاد قسم الثقافة الإسلامية لعدم استخدام الأساليب الإحصائية في دراساته(٩٣٣) رسالة خلال الفترة ما بين ١٤١٤– ١٤٢٨ هـ ، ولكون المجتمع صغير نسبياً فلقد تم اختيار ٢٠% من إجمالي هذا العدد، وتم اختيار العينة بالطريقة العشوائية الطبقية.

الأساليب الاحصائية:

أستخدم الباحث في دراسته التكرارات والنسب المئوية ومعادلات حجم التأثير وفقاً للأسلوب الإحصائي المستخدم في كل دراسة (وجد الباحث بأن الأساليب الإحصائية المستخدمة في عينة الدراسة محصورة في الأساليب التالية : اختبار (ت) لعينتين مستقلتين أو مترابطتين – تحليل التباين – الانحدار المتعدد – معامل ارتباط بيرسون – اختبار كا^{2} – اختبار مان – وتني) وقيمة كا^{2} ، عن طريق برنامج حزمة البرامج الأحيائية للعلوم الاجتماعية (SPSS).

نتائج الدراسة وتفسيرها ومناقشتها:

هدفت الدراسة إلى التعرف على أهم الطرق الإحصائية لحساب حجم التأثير للاختبارات الإحصائية المعلمية واللامعلمية المستخدمة في رسائل الماجستير بكلية التربية بجامعة الملك سعود للفترة مابين (١٤١٤ هـ – ١٤٢٨ هـ) والتعرف أيضاً على واقع استخدام حجم التأثير.

ولتحقيق الهدف من الدراسة قام الباحث بتفريغ البيانات في النماذج المعدة لكل اختبار إحصائي (تحليل التباين ، مربع كاي ، معامل ارتباط بيرسون ، الفروق بين المتوسطات ، مان وتني ، والانحدار المتعدد) ومن ثم تم استخدام التكرارات والنسب المئوية ومعادلات حجم التأثير وقيم مربع كا^{2} وذلك للإجابة على تساؤلات الدراسة التالية :-

١-ما الاختبارات الإحصائية المعلمية واللامعلمية المستخدمة في رسائل الماجستير بكلية التربية بجامعة الملك سعود ؟

٢-ما واقع استخدام حجم التأثير في الاختبارات الإحصائية المعلمية واللامعلمية المستخدمة في رسائل الماجستير بكلية التربية بجامعة الملك سعود ؟

٣-ما أحجام التأثير للاختبارات الإحصائية المعلمية واللامعلميه المستخدمة في رسائل الماجستير بكلية التربية بجامعة الملك سعود ؟

نتائج الدراسة:

نتائج التساؤل الأول:

ما الاختبارات الإحصائية المعلمية واللامعلمية المستخدمة في رسائل الماجستير بكلية التربية بجامعة الملك سعود؟

وللإجابة على هذا السؤال استخدم الباحث التكرارات والنسب المئوية على النحو التالي:-

جدول رقم (١)

التكرارات والنسب المئوية للاختبارات الإحصائية المعلمية واللامعلمية

م	نوع الاختبار	التكرارات	النسبة المئوية
١	المعلمية	٥٥٩	٩٣,٤٨
٢	اللامعلمية	٣٩	٦,٥٢
المجمـــــوع		٥٩٨	%١٠٠

يتضح من الجدول رقم (١) بأن الاختبارات المعلمية أكثر استخداما من نظيرتها اللامعلمية في كلية التربية بجامعة الملك سعود حيث تكرر استخدام الاختبارات المعلمية (٥٥٩) مرة بنسبة وقدرها ٩٣,٤٨ % بينما تكرر استخدام الاختبارات اللامعلمية بواقع (٣٩) مرة بنسبة مئوية تقدر بـ ٦,٥٢ % .

جدول رقم (٢)

التكرارات والنسب المئوية للاختبارات الإحصائية

م	الاختبار الإحصائي المستخدم	التكرارات	النسب المئوية
١	اختبار (ت)	٢٥٦	٤٢.٨١
٢	تحليل التباين	٢٢١	٣٦.٩٦
٣	معامل ارتباط بيرسون	٦٩	١١.٥٤
٤	الانحدار المتعدد	١٣	٢.١٧
٥	مربع كاي	٢١	٣.٥١
٦	مان وتني	١٨	٣.٠١
المجمـــوع		٥٩٨	%١٠٠

يتضح من الجدول رقم (٢) بأنه تم استخدام (٥٩٨) اختبارا إحصائياً في كلية التربية بجامعة الملك سعود ، وقد كان أكثر الأساليب الإحصائية استخداما هو اختبار (ت) حيث أستخدم (٢٥٦) مرة بنسبة ٤٢,٨١% من مجموع الاختبارات الإحصائية المستخدمة في الكلية . ويلي هذا الأسلوب الإحصائي اختبار تحليل التباين (ف) حيث أستخدم (٢٢١) مرة بنسبة وقدرها ٣٦,٩٦ % من مجموع الاختبارات الإحصائية المستخدمة في الكلية .

ثم يلي ذلك اختبار معامل الارتباط لبيرسون وقد أستخدم (٦٩) مرة بنسبة ١١,٥٤% من مجموع الاختبارات الإحصائية المستخدمة في الكلية .

ثم يلي ذلك اختبار كا^{2} حيث أستخدم (٢١) مرة بنسبة ٣,٥١ % من مجموع الاختبارات الإحصائية المستخدمة في الكلية .

ثم يليه اختبار مان وتني وقد أستخدم (١٨) مرة بنسبة ٣,٠١ % من مجموع الاختبارات الإحصائية المستخدمة في الكلية .

ثم يليه اختبار الانحدار المتعدد وقد أستخدم (١٣) مرة بنسبة ٢,١٧ % من مجموع الاختبارات الإحصائية المستخدمة في الكلية .

وهذا يتفق مع ما توصلت إلية دراسة كل من (العجلان ، ١٩٩٠ و الصياد ، ١٩٨٨) حيث أظهرت نتائجهما بأن الاختبار الإحصائي (ت) يعد من أكثر الاختبارات الإحصائية استخداما ، بينما دراسة (النجار،١٩٩١) أشارت إلى أن اختبار كا^{2} أكثر استخداما في كلية التربية بجامعة أم القرى .

ويرى هارول Harwell بأن هنالك عدة أسباب محتملة وراء ندرة استخدام الاختبارات اللامعلمية في مجال البحوث النفسية و التربوية و الاجتماعية وهي : (الضوي ، ٢٠٠٦)

أ- تضمين برامج الحاسب الآلى المعروفة مثل Minitab و SPSS بعدد صغير جداً من البدائل اللامعلمية .

ب- استمرار الاعتقاد لدى كثير من الباحثين التربويين أن الاختبارات اللامعلمية أقل قوة وأقل قبولاً مقارنةً بنظيراتها المعلمية .

ج – عدم وعى الباحثين في المجالات التربوية والنفسية والاجتماعية بالبدائل اللامعلمية المتاحة للإستخدام في التصميمات التجريبية المعقدة، و كيفية إنجاز تحليل البيانات باستخدام البرامج المتوفرة، فالعديد منهم يعتقدون بشكل واضح أن البدائل مرتبطة بالبيانات المستقاة من تصميمات تجريبية بسيطة جداً .

بالإضافة إلى أن الكثير من طلاب الدراسات العليا ليس لديهم معرفه شامله بالبدائل اللامعلمية وذلك لأنهم لم يدرسوها بصوره مفصله كمتطلب في مقرر الإحصاء .

نتائج التساؤل الثاني:

ما واقع استخدام حجم التأثير في الاختبارات الإحصائية المعلمية واللامعلمية المستخدمة في رسائل الماجستير بكلية التربية بجامعة الملك سعود ؟

وللإجابة على هذا التساؤل أستخدم الباحث التكرارات والنسب المئوية على النحو التالي

جدول رقم (٣)

التكرارات والنسب المئوية لواقع استخدام حجم التأثير بكل قسم من أقسام كلية التربية

م	القسم	التكرارات	النسب المئوية
١	علم النفس	٠	٠
٢	قسم التربية	٠	٠
٣	وسائل وتكنولوجيا التعليم	٠	٠
٤	مناهج وطرق تدريس	٠	٠
٥	الإدارة التربوية	٠	٠
٦	التربية البدنية	٢	١,٣
٧	التربية الخاصة	٠	٠
	المجمـــوع	٢	١,٣

يتضح من الجدول رقم (٣) بأنه تم استخدام حجم التأثير في رسالتين فقط من مجموع عينة الدراسة والبالغة ١٥٤ رسالة في كلية التربية بجامعة الملك سعود أي بنسبة وقدرها ١,٣ % وهذه تعد نسبة ضئيلة جداً و هذا بسبب عدم معرفة الباحثين على مستوى الماجستير بمدلول حجم التأثير نتيجة عدم تضمين تدريس طرق حساب حجم التأثير في المقررات الدراسية وخاصة لطلاب الدراسات العليا .

ولاحظ الباحث أثناء جمع البيانات من عينة الدراسة بأن هنالك ٣ رسائل من قسم المناهج وطرق التدريس ورسالة واحده من قسم الإدارة التربوية استخدمت حجم التأثير مع العلم بأن نتائجها غير داله إحصائياً وهذا عكس ما أكد عليه المتخصصون بأن حجم التأثير هو المكمل للدلالة الإحصائية وليس بديلاً عنه (منصور ، ١٩٩٧ و Mclean and Eronest , 1997 و Lane , 1999) . ويرى (الضوي ، ٢٠٠٦) أنه من الأخطاء الشائعة التي يجب أن يلتفت إليها الباحثون حساب حجم التأثير المناظرة لقيم الاختبار الإحصائي بطريقة عفوية سواء أكانت قيم الاختبار الإحصائي دالة أم غير دالة ، والصواب فقط هو حساب أو تقدير حجم التأثير فقط في حالة القبول بالفرض البديل ورفض الفرض الصفري . والنقطة المهمة أن حجم التأثير يكون في صورته المثالية عندما يستخدم الباحث الاختبار الإحصائي المناسب لتحليل بياناته أي عندما تفي بياناته بالافتراضات الأساسية التي يستند إليها الاختبار بالإضافة إلى قبول الفرض البديل ورفض الفرض الصفري.

ويرى شيفر (Shaver, 1992) أن اختبارات الدلالة الإحصائية لا تشير إلى احتمالية خطأ أو صحة الفرضية الصفرية وإنما تعطي للباحث معلومات تتعلق باحتمال الحصول على نتيجة معينة بفرق دال إحصائياً يحدده الفرض الصفري ، وأن كثيراً من الباحثين يقدمون آراء حول رفض الفرضية الصفرية بناء على نتيجة معينة ذات دلالة إحصائية. ويتطلب ذلك ضرورة تشجيع الباحثين للتقليل من الاعتماد على الدلالة الإحصائية في تحليلاتهم وتفسيراتهم ولا ينبغي نشر الأبحاث التي تعتمد على الدلالة الإحصائية فقط ولابد من احتواء التحليلات ونتائج الدراسات على أحجام التأثير.

نتائج التساؤل الثالث:-

ما أحجام التأثير للاختبارات الإحصائية المعلمية واللامعلمية المستخدمة في رسائل الماجستير بكلية التربية بجامعة الملك سعود ؟

وللإجابة على هذا التساؤل استخدم الباحث معادلات حجم التأثير المناسبة لكل اختبار إحصائي والموضحة سابقاً في فصل إجراءات الدراسة ووفقاً للمعطيات المتوفرة في عينة الدراسة .

أولاً :- متوسط أحجام التأثير لاختبار (ت)

أ) لعينتين مستقلتين

جدول رقم (٤)

متوسط أحجام التأثير لاختبار (ت) في حال عينتين مستقلتين

	حجم العينة	أقل قيمة	أعلى قيمة	المتوسط	الانحراف المعياري
حجم التأثير	٢٣٣	-	,٩٧	,٤٤١١	,٢٤٥٥٣

ب) لعينتين مترابطتين

جدول رقم (٥)

متوسط أحجام التأثير لاختبار (ت) في حال عينتين مترابطتين

	حجم العينة	أقل قيمة	أعلى قيمة	المتوسط	الانحراف المعياري
حجم التأثير	٢٣	,٠٥	,٨٥	,٣٤٠٨	,٢٣٨٩٥

يتضح من الجدولين رقم (٤ و ٥) بأن متوسط حجم التأثير للاختبار الإحصائي (ت) لعينتين مستقلتين ولعينتين مترابطتين (,٤٤١١ ، ,٣٤٠٨) على الترتيب وهذا حجم تأثير متوسط وفقاً لتصنيف كوهن لأحجام التأثير<

وهذا يتفق مع دراسة نصار (٢٠٠٢) والتي أظهرت أن قيمة حجم الأثر لاختبار (t) سواء في العينات المترابطة أو المستقلة واختبار (ف) في حالة تحليل التباين الأحادي ومعامل ارتباط بيرسون وأخيراً الانحدار الخطي البسيط إلى أنه حجم أثر صغير أو متوسط.

بينما توصلت دراسة كل من (الصياد ، ١٩٨٨ النجار ،١٩٩١ إبراهيم ١٩٩٧) إلى أن حجم التأثير المصاحب للدلالة الإحصائية ضعيف جداً بينما أشارت دراسة (الصائغ ١٩٩٦) إلى أن حجم التأثير للأحصائي (ت) كان كبيراً حيث بلغ ,٨٠ .

ثانياً :- متوسط أحجام التأثير لاختبار (الانحدار المتعدد)

جدول رقم (٦)

متوسط أحجام التأثير لاختبار (الانحدار المتعدد)

	حجم العينة	أقل قيمة	أعلى قيمة	المتوسط	الانحراف المعياري
حجم التأثير	١٣	,٠٢	,٩٥	,٣٨٧٦	,٣٢٤٤٥

يتضح من الجدول رقم (٦) بأن متوسط حجم التأثير لاختبار الانحدار المتعدد بلغت (,٣٨٧٦) وهذا حجم تأثير متوسط وفقاً لتصنيف كوهن لأحجام التأثير، وهذا يتفق مع دراسة سميثون (Smithson, 2001) حيث أظهرت نتائجها أن قيمة حجم التأثير متوسطة، بينما دراسة كل من اجينيس وبيتي (Aguinis and Beaty , 2005) أظهرت بأن متوسط حجم التأثير مساوياً لـ(002.) فقط وهذا بناءً على تصنيف كوهن لأحجام التأثير يعتبر صغير جداً .

ثالثاً :- متوسط أحجام التأثير لاختبار (تحليل التباين)

يتضح من الجدول رقم (٧) بأن متوسط حجم التأثير لاختبار تحليل التباين بلغ (,٠٨٥٤) وهذا حجم تأثير صغير جداً وفقاً لتصنيف كوهن لأحجام التأثير، وهذا يتفق مع دراسة كل من النجار ، ١٩٩١ و الصائغ ١٩٩٦ ونصار ٢٠٠٢ و Haase and waechter and Solomon, 1982 والتي أظهرت أن قيمة حجم الأثر لاختبار (ف) كان صغيراً .

جدول رقم (٧)

متوسط أحجام التأثير لاختبار (تحليل التباين)

	حجم العينة	أقل قيمة	أعلى قيمة	المتوسط	الانحراف المعياري
حجم التأثير	٢٢١	-	,٦٠	,٠٨٥٤	,١٠٩١٠

بينما أظهرت دراسة إبراهيم (١٩٩٧) والتي هدفت إلى التعرف على واقع الدلالة الإحصائية والعملية و قوة الاختبار الإحصائي في رسائل الماجستير في مجال علم النفس التربوي في الجامعة الأردنية ما بين عامي (١٩٩٠-١٩٩٥) وأشارت نتائجها إلى أن قيمة حجم التأثير لاختبار (ف) كان متوسط ، ويعود سبب ضعف قيمة حجم التأثير المستخرجة إلى ضعف في تصميم هذه الدراسات.

رابعاً :- متوسط أحجام التأثير لاختبار (معامل ارتباط بيرسون)

يتضح من الجدول رقم (٨) بأن متوسط حجم التأثير لاختبار معامل ارتباط بيرسون بلغ (,٢٦٧٦) وهذا حجم تأثير متوسط وفقاً لتصنيف كوهن لأحجام التأثير.

جدول رقم (٨)

متوسط أحجام التأثير لاختبار (معامل ارتباط بيرسون)

	حجم العينة	أقل قيمة	أعلى قيمة	المتوسط	الإنحراف المعياري
حجم التأثير	٦٩	,٠١	,٧٦	,٢٦٧٦	,٢٣٦٨

وتتفق نتيجة الدراسة الحالية مع ما توصلت إلية دراسة نصار (٢٠٠٢) والتي أظهرت بأن قيمة حجم التأثير لمعامل ارتباط بيرسون كان عند مستوى صغير ومتوسط.

بينما دراسة كل من إبراهيم (١٩٩٧) ومحمود(٢٠٠٣) أظهرت بأن قيمة حجم التأثير لمعامل إرتباط بيرسون ضعيفة وفقاً لتصنيف كوهن لأحجام التأثير .

خامساً :- متوسط أحجام التأثير لاختبار (مان وتني)

يتضح من الجدول رقم (٩) بأن متوسط حجم التأثير لاختبار مان – وتني بلغ (,٦٦٩١) وهذا حجم تأثير متوسط وفقاً لتصنيف كوهن لأحجام التأثير.

جدول رقم (٩)

متوسط أحجام التأثير لاختبار (مان – وتني)

	حجم العينة	أقل قيمة	أعلى قيمة	المتوسط	الانحراف المعياري
حجم التأثير	١٨	,٠٢	١	,٦٦٩١	,٣٢٠٨

سادساً :- متوسط أحجام التأثير لاختبار (كاي تربيع)

يتضح من الجدول رقم (٩) بأن متوسط حجم التأثير لاختبار كاي تربيع بلغ (,٢٢٥١) وهذا حجم تأثير متوسط وفقاً لتصنيف كوهن لأحجام التأثير .

جدول رقم (٩)
متوسط أحجام التأثير لاختبار (كاى تربيع)

	حجم العينة	أقل قيمة	أعلى قيمة	المتوسط	الانحراف المعياري
حجم التأثير	٢١	,١٢	,٤٣	,٢٢٥١	,٠٩٨٤

وتتفق نتيجة الدراسة الحالية مع ما توصلت إلية دراسة إبراهيم (١٩٩٧) والتي أظهرت بأن قيمة حجم التأثير لاختبار كاى تربيع كان عند مستوى متوسط .

بينما أظهرت دراسة النجار(١٩٩١ ، ٢٠٠٥) أن حجم التأثير لاختبار كاى تربيع المصاحب للدلالة الإحصائية كان عند مستوى ضعيف جداً بنسبة ٤٥,٨ % من مجموع الاختبارات الإحصائية المستخدمة لاختبار كاى تربيع .

وفي المقابل أظهرت دراسة محمود (٢٠٠٣) أن حجم التأثير لاختبار كاى تربيع والمصاحب للدلالة الإحصائية كان عند مستوى كبير بنسبة ٤٨,٨٤ % .

وعموماً فإن متوسط أحجام التأثير للاختبارات الإحصائية المستخدمة في كلية التربية بجامعة الملك سعود والدالة إحصائياً ذات أحجام تأثير متوسطة .

وتتفق هذه الدراسة مع دراسة تشيس وتشيس (Chase & Chase, 1976) في أن قيمة حجم التأثير والمصاحبة للدلالة الإحصائية كانت عند مستوى متوسط.

توصيات الدراسة

من خلال النتائج التي توصلت لها هذه الدراسة؛ هناك عدد من التوصيات التي يمكن إجازها في النقاط التالية:

- ضرورة تضمين تدريس البدائل اللامعلمية في المقررات الدراسية وخاصة لطلاب الدراسات العليا .
- حث الباحثين عل ضرورة حساب حجم التأثير الخاصة بالاختبارات الإحصائية لفروضهم العلمية وعدم الاكتفاء فقط بالدلالة الإحصائية ، وتضمين ذلك في ملخصات دراساتهم .
- ضرورة تضمين تدريس طرق حساب حجم التأثير في المقررات الدراسية وخاصة لطلاب الدراسات العليا .
- ضرورة عدم نشر الأبحاث التي تعتمد على الدلالة الإحصائية فقط ولا بد من احتواء نتائج الدراسات على أحجام التأثير.

المقترحات:

من خلال ماتم التوصل إليه من خلال هذه الدراسة؛ فإن الباحث يوصي بما يلي:

- القيام بدراسة مماثلة على الدوريات العلمية للاختبارات الإحصائية الشائعة .

إصدار دليل لطلاب وطالبات الدراسات العليا (دليل إحصائي) يحتوي على الأساليب الإحصائية المعلمية واللامعلمية وطرق حساب أحجام التأثير .**المراجع**

أولا: المراجع العربية:

ابراهيم، ألطاف (١٩٩٨) . واقع الدلالات الإحصائية والعملية وقوة الاختبارات الإحصائية المستخدمة في رسائل ماجستير علم النفس التربوى فى الجامعة الأردنية. رسالة ماجستير ، الأردن ، الجامعة الأردنية .

أبو حطب ، فؤاد و صادق ، آمال (١٩٩١) . مناهج البحث وطرق التحليل الإحصائي في العلوم النفسية والتربوية والاجتماعية٠ القاهرة ، الأنجلو المصرية.

أبو زيد ، مدحت (٢٠٠٢) . الإحصاء في العلوم السلوكية الوصف الإحصائي (١) تبويب البيانات وتمثيلها بيانيا ونزعتها المركزية. الإسكندرية ، دار المعرفة الجامعية.

أبو هاشم ، السيد (٢٠٠٥) . مؤشرات التحليل البعدي لبحوث فعالية الذات في ضوء نظرية باندورا. مركز بحوث كلية التربية بجامعة الملك سعود ، العدد ٢٣٨. الرياض .

بابطين ، عادل (٢٠٠٢) . مشكلات الدلالة الإحصائية في البحث التربوي وحلول بديلة ، رسالة ماجستير ، جامعة أم القرى، مكة المكرمة .

بري ، عدنان وهندي، محمود وراضي، الحسيني (١٩٩٨) . اساسيات طرق التحليل الإحصائية ، الرياض ، جامعة الملك سعود .

تشاو ، لنكولن (١٩٩٠) . الإحصاء فى الإدارة. ترجمة عبد المرضى عزام ، الرياض . دار المريخ للنشر .

توفيق ، عبد الجبار (١٩٨٥) . التحليل الإحصائي في البحوث التربوية والنفسية و الاجتماعية – الطرق اللا معلمية ٠ الكويت ، مؤسسة الكويت للتقدم العلمي .

الجادري، عدنان (٢٠٠٣) . الإحصاء الوصفي في العلوم التربوية ، عمان: دار المسيرة للنشر والتوزيع والطباعة .

حجيمات، تحسين (١٩٩٥) . دراسة واقع الدلالة العملية وقوة الأختبارات الإحصائية المستخدمة في رسائل الماجستير في مجال الإرشاد والتوجية ، رسالة ماجستير ، الأردن ، الجامعة الأردنية .

الدردير ، عبد المنعم (٢٠٠٦) . الإحصاء البارامترى و اللابارامترى فى اختبار فروض البحوث النفسية و التربوية و الإجتماعية . القاهرة . عالم الكتب.

سليم ، صباح و أبو حويج ، مها (٢٠٠٤) . مقدمة فى الإحتمالات و الإحصاء ، عمان . دار قنديل للنشر .

سميث ، ج ملتون (١٩٨٥) . الدليل إلى الإحصاء فى التربية و علم النفس . ترجمة ابراهيم بسيونى عميرة ، القاهرة ، دار المعارف .

الشربينى ، زكريا (١٩٩٥) . الإحصاء وتصميم التجارب في البحوث النفسية والتربوية والإجتماعية ، القاهرة ، الأنجلو المصرية .

الصائغ ، ابتسام (١٩٩٦) . الدلالة الإحصائية والدلالة العملية لاختبار (ت) و(ف) دراسة تحليلية تقويمية من خلال رسائل الماجستير التي قدمت في كلية التربية بجامعة أم القرى بمكة المكرمة ، رسالة ماجستير ، مكة المكرمة، جامعة أم القرى.

الصياد ، عبد العاطي (١٩٨٨) . الدلالة العملية وحجم العينة المصاحبتين للدلالة الإحصائية لاختبار (ت) في البحث التربوي والنفسي العربي – بحوث مؤتمر البحث التربوي الواقع والمستقبل. المجلد الثاني ، القاهرة ، ص ١٩٩ –٢٣٠ .

الضوي ، محسوب (٢٠٠٦) . الإحصاء الاستدلالى المتقدم فى التربية و علم النفس ، القاهرة ، الأنجلو المصرية .

العجلان ، فتحية محمد (١٩٩٠) . دراسة تقويمية للأساليب الإحصائية المستخدمة في رسائل الماجستير بكلية التربية بجامعة أم القرى . مكة المكرمة . جامعة أم القرى .

علام ، صلاح الدين (١٩٩٣) . الأساليب الإحصائية الاستدلالية البارامترية و اللابارامترية فى تحليل بيانات البحوث النفسية و التربوية ، القاهرة ، دار الفكر العربى .

عودة ، أحمد / القاضى ، منصور (٢٠٠٢) . الإحصاء الوصفى و الاستدلالى ، الكويت ، مكتبة الفلاح للنشر و التوزيع .

ماضى ، محمد و عثمان ، ماجد (٢٠٠٥) . الإحصاء فى التربية و علم النفس مع استخدام SPSS و , Minitabالطبعة الثانية ، دبى ، دار القلم للنشر و التوزيع .

محمود ، فريال (٢٠٠٣) . واقع الدلالة الاحصائية والدلالة العملية وقوة الاختبار للاختبارات الاحصائية المستخدمة في رسائل الماجستير الصادرة عن كلية العلوم التربوية في جامعة مؤتة ، رسالة ماجستير ، الأردن ، جامعة مؤتة .

مراد ، صلاح (٢٠٠٠) . الأساليب الإحصائية في العلوم النفسية والتربوية والاجتماعية ، القاهرة ، الانجلو المصرية .

منسي ، محمود (١٩٨٩) . الإحصاء والقياس في التربية وعلم النفس. الإسكندرية، دار المعرفة الجامعية .

منصور، رشدي فام (١٩٩٧م) . حجم التأثير – الوجه المكمل للدلالة الإحصائية، المجلة المصرية للدراسات النفسية ، العدد (١٦) ، ٥٧-٧٥ .

المنيزل ، عبدالله (٢٠٠٠) . الإحصاء الاستدلالى وتطبيقاته فى الحاسوب بإستخدام الرزم الإحصائية , SPSSعمان ، دار وائل للنشر .

النجار ، عبد الله (٢٠٠٥) . الدلالة الإحصائية والدلالة العملية لاختبار كا٢ في البحوث الإدارية المنشورة (دراسة تقويمية). المجلة العربية للعلوم الإدارية . (١٢) ، (٢) ، ١٦٩-١٩٣ .

النجار، عبد الله (١٩٩١) . دراسة تقويمية مقارنة للأساليب الإحصائية التي استخدمت في تحليل البيانات في رسائل الماجستير في كل من كلية التربية جامعة أم القرى بمكة المكرمة وكلية التربية بجامعة الملك سعود بالرياض ، رسالة ماجستير ، مكة المكرمة ، كلية التربية جامعة أم القرى .

نصار، يحيى (٢٠٠٢). حجم الأثر كأسلوب إحصائي مكمل لفحص الفرضيات الإحصائية ، مركز بحوث كلية التربية – الرياض ،جامعة الملك سعود، العدد ١٧٦.

ثانياً المراجع الأجنبية:

Anderson, D , Burnham,K , and Thompson, W (2000). Null Hypothesis Testing:Problems ,Prevalance, and an Alternative. Journal of Wildlife Management, 64 , 912-923 .

Aguinis, H and Beaty , J. and Boik , R and Pierce , D (2005). Effect Size and Power in Assessing Moderating Effects of Categorial Variables Using Multiple Regression :A 30-Year Review. Journal of Applied Psychology, 91 , 1,94-107 .

Bayarri, M and Mayoral, A (2002). Bayesian Analysis and Testing for Comparison of Effect Sizes. Journal of Statistical Planning and Inference, 103 , 225-243 .

Bird, K (2002). Confidence Intervals for Effect Sizes in Analysis of Variance. Educational and Psychological Measurement, 62 , 2, 197-226 .

Chase, L. and chase, R (1976) A Statistical Power Analysis of Applied Psychological Research, Journal of Applied psychology, 61: 234-237 .

Elmore, P and Rotou, O (2001).A Primer on Basic Effect Size Concepts .Paper Produced at the Annual Meeting of the American Educational Research Association .

Hashgwhzyse. R , waechter, D. and Solomon, G. (1982). How Significant is a Significant Difference? Average Effect Size of Research in Counseling Psychology, Journal of Counseling Psychology, 29, 58-65 .

Huberty, C (2002). A History of Effect Size Indices. Educational and Psychological Measurement, 62 , 2, 227-240 .

Huston, H (1993). Meaning Fullness, Statistical Significance, Effect Size and Power Analysis: A General Discussion with Implications for MANOVA, Paper presented at the annual meeting of the mid – south Educational Research Association , 22 nd ,New Orleans, ERIC Document Reprcuctive Service. No ED 364608 .

Lane, G (1999). Show my the Magnitude! The Consequences of Overemphasis on Null Hypothesis Significance Testing. Paper presented at the annual meeting of the mid – south educational research association (al, November 16-19, 1999). (Eric document Reproductive service no. ED 436557).

Maxwell , S and Delaney , H (2004) . Designing experiments and analyzing data : A model comparisonperspective (2 nd ed) . Mahwah , NJ : Lawrence Erlbaum .

McGrath, R and Meyer , G (2006). When Effect Sizes Disagree : The Case of r and d. Psychological Methods, 11 , 4,386-401.

Mclean, J and Eronest, J (1997). Has Testing for Statistical Significance outlived its usefulness? Paper presented at the annual meeting of the mid-south. Educational Research association (26, mephis, TN November. 12-14, 1997) (eric reproductive service NO. Ed 415265).

Olejnik, S and Algina, J (2000). Measures of Effect Size for Comparative Studies: Applications, Interpretations, and Limitations. Contemporary Educational Psychology, 25, 241-286 .

Parker, R and Hagan-Burke, S (2007). Useful Effect Size Interpretations for Single Case Research. Behavior Therapy, 83 , 95-105.

Roberts, M and Henson, R (2002). Correction for Bias in Estimating Effect Sizes. Educational and Psychological Measurement , 62 ,2, 241-253.

Robey, R (2004). Reporting Point and Interval Estimates of Effect-Size for Planned Contrasts : Fixed Within Effect Analyses of Variance. Journal of Fluency Disorders, 29 , 307-341.

Rosnow, R and Rosenthal , R and Rubin ,D (2000). Contrasts and Correlations in Effect Size Estimations. Psychological Science, 11 , 446-453.

Shaver, J (1992). What Statistical Significance Testing Is , and what it is not? Journal of experimental Education, 61, 4 , 293-316.

Smithson, M (2001). Correct Confidence Intervals for Various Regression Effect Sizes and Parameteters : the Importance of NonCentral Distributions in Computing Intervals. Educational and Psychological Measurement , 61 ,4, 605-632.

Snyder, p. and Lawson , S (1993). Evaluating Results using Corrected and Uncorrected Effect Size Estimates. Journal of Experimental Education, 61, 334-349.

Thompson, B (2002). Statistical ,practical, and clinical :how many kinds of significance do Counselors need to consider? Journal of Counseling and Development , 80,64- 71.

Welge , P (1990). Looking beyond Statistical Significance: Result Importance and Result Generalizability. Paper presented at the Annual Meeting of the American Psychological Society NO. Ed 320-965 .

Wilcox, R (2006). Graphical Methods for assessing Effect Size : some Alternatives to Cohen's d. Journal of Experimental Education, 74 (4) , 353-367.

Wright , D (2003) . Making Friends with your data : Improving how statistics are conducted and Reprovted . British Journal of Educational Psychology . 73, 123- 136 .