الجمعية الأردنية للعلوم التربوية، المجلة التربوية الأردنية، المجلد العاشر، العدد الأول، ملحق (2)، 2025
DOI: https://doi.org/10.46515/jaes.v10i1.1407

The Reality of the Practical Significance and the Power of Non-Parametric Statistical Tests Used in Doctoral Dissertations and Educational Master's theses at the University of Jordan

Hala Ibraheem Al-Haj*
Prof. Jehad Mohammed Anati**

Received 15/6/2023 Accepted 22/7/2023

Abstract:

This study aimed to study the reality of the practical significance and power of non-parametric statistical tests (Kruskal Wallis,Mann Whitney,Wilcoxon) used in doctoral dissertations and educational masters theses at the university of Jordan within the time period (2017-2021),and the total number of non-parametric statistical tests (124) statistical Tests,to achieve this,the content of theses and dissertations was analyzed for the period covered by the study,and the statistical values necessary to calculate the practical significance and power of statistical tests were recorded in the analysis card prepared for that,and the practical significance and power of statistical tests for each test was calculated using the appropriate statistical formulas,The results showed that the percentage of researchers associating practical significance values with statistical significance was 0%,and the researchers neglected to calculate them in all non-parametric statistical tests., and when talking about the power of the statistical tests included in the study ,the percentage of statistically significant hypotheses that were associated with a small test power was (25%) ,while the percentage of statistically significant hypotheses that were associated with a medium test power was (15%) and finally (60%) of the hypotheses were associated with large power test.

Keywords: practical significance, power of test, statistical significance, non-parametric statistical test.

* Jordan\ halaa-hasan2019@yahoo.com
** https://orcid.org/0000-0001-8513-7679
School of Educational Sciences\ The University of Jordan\ Jordan\ jehad54321@yahoo.co.uk

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

واقع الدلالة العملية وقوة الاختبارات الإحصائية اللامعلمية المستخدمة في أطروحات الدكتوراة ورسائل الماجستير التربوية في الجامعة الأردنية

هالة ابراهيم الحاج *

أ.د. جهاد محمد العناتي **

ملخص:

هدفت هذه الدراسة إلى دراسة واقع الدلالة العملية وقوة الاختبارات الإحصائية للاختبارات الإحصائية اللامعلمية (كروسكال واليس، مان ويتني، ويلككسون) المستخدمة في أطروحات الدكتوراة ورسائل الماجستير التربوية في الجامعة الأردنية ضمن الفترة الزمنية (2017–2021). وقد بلغ إجمالي عدد الاختبارات الإحصائية اللامعلمية (124) اختبارا إحصائيا. ولتحقيق ذلك تم تحليل محتوى الرسائل والأطروحات للفترة التي شملتها الدراسة وتسجيل القيم الإحصائية اللازمة لحساب الدلالة العملية وقوة الاختبارات الإحصائية في بطاقة التحليل المعدة لذلك، وحساب الدلالة العملية الخاصة بكل اختبار فضلاً عن حساب قوة الاختبارات الإحصائية المتضمنة في الدراسة باستخدام برمجية حاسوبية مناسبة، وأظهرت النتائج أن نسبة ما قرنه الباحثون لقيم الدلالة العملية مع الدلالة الإحصائية 0%، وأهمل الباحثون حسابها في جميع الاختبارات الإحصائية. وعند الحديث عن قوة الاختبارات الإحصائية المتضمنة في الدراسة فقد بلغت نسبة الفرضيات الدالة إحصائيا والتي اقترنت بقوة اختبار صغيرة بنسبة (25%) في حين بلغت نسبة الفرضيات الدالة إحصائيا والتي اقترنت بقوة اختبار متوسطة (15%) وأخيرا اقترن ما نسبته(60%) من الفرضيات الدالة إحصائيا بقوة اختبار كبيرة.

الكلمات المفتاحية: الدلالة العملية، الدلالة الإحصائية، قوة الاختبار الإحصائي، الاختبار الإحصائي اللامعلمي.

* الأردن/ halaa-hasan2019@yahoo.com

** كلية العلوم التربوية/ الجامعة الأردنية/ الأردن/ jehad54321@yahoo.co.uk

المقدمة:

تعد العملية البحثية عملية منظمة تتم بهدف التوصل إلى حلول للمشكلات، أو تفسير الظواهر والإجابة عن التساؤلات، ويتم فيها توظيف طرق استقصائية متعارف عليها في العملية البحثية في مجال معين، بهدف الوصول إلى معرفة جديدة، ومن ثم قرارات يمكن تعميمها، ومن أجل ذلك يقوم الباحث بجمع البيانات لغايات التحليل ومن ثم التفسير، وهنا لا بد من ذكر أن عملية التحليل والتفسير للبيانات تعتمد على مدى إتقان الباحث للأساليب الإحصائية المناسبة لعملية التحليل للوصول إلى نتائج ذات موثوقية وليست مضللة (Alkilani & Alshareefen،2011، ALnajar،2005).

ولتنفيذ ذلك يسير الباحث وفق خطوات علمية واضحة ومنظمة هدفها اختبار تعميمات تتعلق بالمجتمع اعتمادا على البيانات التي تم جمعها من العينة، فيقوم الباحث بداية بصياغة الفرضيات والتي تتخذ عادة نوعين: الأول الفرضية الصفرية، وهي التي يتم فحصها باستخدام الاختبار الإحصائي وتتم صياغتها بالاعتماد على مفاهيم إجرائية، إذ يتم جمع البيانات حولها ومن ثم فحصها باستخدام اختبار إحصائي مناسب. والثاني الفرضية البديلة، وهي الفرضية التي تُعد البديل للفرضية الصفرية، ولا يتم فحصها باستخدام الاختبار الإحصائي ويتم قبولها في حال رفض الفرضية الصفرية (Al_Shareefen،2017).

ويتخذ الباحث قرارا بقبول الفرضية الصفرية أو رفضها بعد فحصها باستخدام اختبار إحصائي محدد، وتجدر الإشارة إلى أن قرار الباحث برفض الفرضية الصفرية أو قبولها لا يمكن أن يكون يقينيا، فهو يعبر عن احتمالية صواب هذا القرار، ومن هنا فإن هنالك نوعين من الخطأ يمكن أن يقع الباحث بأحدهما: الخطأ من النوع الأول، وهو احتمال رفض الباحث الفرضية الصفرية وهي في واقع الأمر صحيحة، ويطلق على القيمة القصوى لاحتمالية وقوع الباحث في هذا النوع من الخطأ مستوى الدلالة الإحصائية (α)، والخطأ من النوع الثاني، وهو احتمال قبول الباحث الفرضية الصفرية وهي في واقع الأمر خاطئة(β)، وبناء عليه أيا كان القرار هناك درجة من الخطأ ولا ينبغي أن ينظر إلى اختبار الفرضيات على أنه برهان لها، وهنالك عديد من العوامل التي تؤثر في قرار الباحث برفض الفرضية الصفرية أو قبولها من مثل: مستوى الدلالة الإحصائية، وحجم العينة، وحجم الأثر، وقوة الاختبار الإحصائي(Abuzainah & Alkilani & Alshareefen،2011،Albatsh ،Cohen،1988).

ومع أن القرار المتعلق برفض الفرضية الصفرية أو قبولها يتضمن أن للمتغير المستقل أثرا على المتغير التابع، بمعنى أن حجم الأثر لا يعود للصدفة (Konstantopulos,2008)، إلا أنه لا يقدم معلومات أخرى من مثل كون هذا الأثر حقيقيا، وإذا كان حقيقيا فكم حجمه، وهل هذا الأثر كبير بدرجة كافية للوصول إلى نتائج موثوقة (Kirk,1996)، فاختبار الدلالة الإحصائية لا يُعد مقياسا موثوقا لقياس أثر المعالجة وذلك لاعتماده على حجم العينة التي كلما كانت بازدياد كانت احتمالية رفض الفرضية الصفرية أكبر، الأمر الذي يتطلب من الباحث للوصول إلى نتائج ذات معنى استخدام ما يسمى بالدلالة العملية (Practical Significance) (Hasse,2001).

وذكر ويلكنسون (Wilkinson,1992) أن هناك عديداً من الانتقادات الموجهة للدلالة الإحصائية منها أن كثيراً من الدراسات التي تم نشرها والتي تضمنت دلالة إحصائية كبيرة كان مستوى القوة فيها متدنيا، كما وأن بعض الدراسات يمكن أن تكون ذات دلالة إحصائية ولكنها ليست ذات أهمية فعلية وليس لها قيمة تطبيقية في المجتمع.

وأضاف الشريفين(Al_Shareefen,2017) أن هناك عديداً من العوامل التي لها تأثير في الدلالة الإحصائية من مثل: حجم العينة، ومستوى الدلالة الإحصائية، وحجم الأثر، وقوة الاختبار المستخدم، فإذا قام الباحث بزيادة حجم العينة أو زيادة مستوى الدلالة "ألفا" فقد يحصل على نتائج ذات دلالة إحصائية، كما قد يقبل الباحث الفرضية الصفرية على الرغم من وجود أثر للمعالجة ويعود السبب في هذا إلى ضعف قوة الاختبار الإحصائي الذي تم استخدامه للكشف عن الفروق، ومما يجدر ذكره أن الأبحاث التي اعتمد فيها الباحثون على الدلالة العملية لاتخاذ القرار أبحاث نادرة.

مما سبق تتضح أهمية أن يقرن الباحث نتائج الدلالة الإحصائية بنتائج الدلالة العملية، التي توفر للقارئ دليلا آخر يثبت أن أثر المتغير المستقل في المتغير التابع حقيقي وكبير بدرجة كافية للوصول إلى نتائج موثوقة، ومن هنا كان لا بد من البحث في الدلالة العملية للاختبارات الإحصائية المستخدمة في أطروحات الدكتوراة ورسائل الماجستير التربوية بشكل عام، وللاختبارات الإحصائية اللامعلمية بشكل خاص، لأن مراجعة الدراسات السابقة في كلية العلوم التربوية في الجامعة الأردنية أظهرت بشكل أولي وجود فجوة في البحث فيها على العكس من الاختبارات الإحصائية المعلمية التي نالت نصيبا من اهتمام الباحثين أردنيا وعربيا.

وقد أشار هيوستن (Huston,1993) لأهمية فحص الدلالة العملية التي تزودنا بقيمة

تقديرية لمقدار وجود الظاهرة في المجتمع، ويتم من خلالها وصف حجم الأثر (صغير أو متوسط أو كبير). كما ويمكن الاعتماد عليها بوصفها قيمة كمية للمقارنة بين نتائج دراستين. وقد تحدث كيرك (Kirk,1996) عن فئتين لتصنيف مؤشرات الدلالة العملية هما: مؤشرات قوة العلاقة، ومؤشرات حجم الأثر.كما تحدث أرون وأرون (Aron & Aron,1997)عن فئتين لتصنيفها هما: مقاييس حجم الأثر، ومقاييس قوة العلاقة.

ولكل اختبار إحصائي يستخدم لفحص الفرضية الصفرية التي يتبناها الباحث مقياس حجم أثر خاص به، وحيث إن الاهتمام في هذا البحث كان خاصا بالاختبارات الإحصائية اللامعلمية (كروسكال واليس Kruskal Wallis ، مان ويتني Mann Whitney ، ويلككسون Wilcoxon)، لذا سوف يقتصر العرض على ذكر القوانين الخاصة بهذه الاختبارات. فيتم حساب حجم الأثر لاختبار ويلككسون من خلال المعادلة الآتية:

$$r = \frac{z}{\sqrt{n}} \quad \text{(1)}.....$$

حيث: r: حجم الأثر.Z: قيمة الإحصائي المستخدم.n: حجم العينة.

وصنف كوهن (Cohen,1988)حجم الأثر وفق الفترات الآتية: إذا كانت 0.10≤r<0.30يكون حجم الأثر صغيرا، وإذا كانت 0.30≤r<0.50يكون حجم الأثر متوسطا، وإذا كانت r≥0.50يكون حجم الأثر كبيرا.

ويتم حساب حجم الأثر لاختبار مان ويتني من خلال المعادلة الآتية:

$$\eta^2 = \frac{Z^2}{n^2} \quad \text{(2)}.....$$

حيث: Z^2: قيمة الإحصائي مان ويتني. n : حجم العينة.

وصنف كوهن (Cohen,1988) حجم الأثر وفق الفترات الآتية: إذا كانت0.10< η^2 0.30يكون حجم الأثر صغيرا، وإذا كانت 0.30< η^2 0.50يكون حجم الأثر متوسطا، وإذا كانت η^2 ≥ 0.50حجم الأثر كبيرا

ويتم حساب حجم الأثر لاختبار كروسكال واليس من خلال المعادلة الآتية:

$$\eta^2 = \frac{H-K+1}{n-k} \quad \text{(3)}.....$$

حيث: H: قيمة الإحصائي كروسكال واليس.K: عدد المجموعات.n : حجم العينة.

وصنف كوهن (Cohen,1988) حجم الأثر وفق الفترات الآتية: إذا كانت0.01 < η^2

0.06يكون حجم الأثر صغيرا، وإذا كانت0.06< $\eta^2 \geq$ 0.14يكون حجم الأثر متوسطا، وإذا كانت $\eta^2 \leq$ 0.14يكون حجم الأثر كبيرا.

وعند الحديث عن قوة الاختبار الإحصائي فهناك ما يشير إلى اتفاق الباحثين على أهمية المعلومات الناتجة عن قوة الاختبارات الإحصائية حيث إن قيمة هذه المعلومات تتأتى من تدليلها على دقة القرار البحثي وصدقه فيما يتعلق برفض الفرض الصفري الخاطئ من جهة كما ويُعد إشارة على الخطأ من النوع الثاني من جهة أخرى وافترض كوهن بأن من الواجب ألا تقل قوة الاختبار الإحصائي لفحص الفرض الصفري عن 0.80 (Park,2008).

وهناك عديد من العوامل المؤثرة في قوة الاختبار الإحصائي وأهمها:

1. حجم الأثر: تزداد قوة الاختبار الإحصائي بزيادة حجم الأثر إذ إنه بزيادة حجم الأثر تزداد قيمة التباين المفسر.
2. حجم العينة: تزداد قوة الاختبار الإحصائي بزيادة حجم العينة إذ إنه بزيادة حجم العينة يزداد تمثيل العينة للمجتمع، الأمر الذي يؤدي إلى زيادة قوة الاختبار في الكشف عن الفروق.
3. مستوى الدلالة الإحصائية المستخدم لاختبار الفرضيات إذ إنه بزيادة مستوى الدلالة الإحصائية ألفا يقل الخطأ من النوع الثاني الأمر الذي يؤدي إلى زيادة قوة الاختبار الإحصائي(Aron & Aron,1997).

وهناك مبررات عديدة تجعل الباحثين مهتمين بحساب قوة الاختبار الإحصائي ومنها أنه:

1. يساعد الباحث على اتخاذ القرارات فيما يتعلق بزيادة حجم العينة.
2. يقدم للباحث تفسيرا لرفض الفرضية الصفرية.
3. إذا كانت نتائج الاختبار ذات دلالة إحصائية وكان حجم الأثر كبيرا تزداد احتمالية الحصول على نتائج ذات دلالة عملية وفي المقابل إذا كان حجم الأثر صغيرا فهذا يعني أنها ذات دلالة إحصائية بسبب عوامل مغايرة.
4. إذا كانت نتائج الاختبار غير دالة إحصائيا وقوة الاختبار صغيرة فهذا يشير إلى عدم صحة فرضية البحث أو عدم مناسبة حجم العينة وفي المقابل إذا كانت قوة الاختبار مرتفعة فهناك احتمال كبير لعدم صحة الفرضية (Huston,1993).

ويمكن حساب قوة الاختبار الإحصائي باستخدام المعادلات الرياضية أو باستخدام جداول خاصة بقوة الاختبارات الإحصائية أو من خلال برامج حاسوبية تم إعدادها خصيصا لحساب قوة

الاختبارات الإحصائية من مثل برنامج (Statistical Power Analysis) الذي تم استخدامه من قبل الباحثين لتحقيق هدف الدراسة.

مشكلة الدراسة وأسئلتها:

جاءت الدراسة الحالية لاستكمال ما قام به الباحثون من دراسة لواقع الدلالة العملية وقوة الاختبارات الإحصائية لعديد من الدراسات في مجالات مختلفة، إلا أن الاهتمام في هذه الدراسة كان منصبا بشكل خاص على الاختبارات الإحصائية اللامعلمية (كروسكال واليس Kruskal Wallis ، مان ويتني Mann Whitney ، ويلككسون Wilcoxon). والتي تضمنتها رسائل الماجستير وأطروحات الدكتوراة في الفترة الزمنية (2017–2021)، بهدف تقييم واقع الدلالة العملية وقوة الاختبارات الإحصائية للاختبارات الإحصائية اللامعلمية بالتزامن مع حساب الدلالة الإحصائية للوصول إلى قرارات أكثر دقة ومصداقية. وقد هدفت بالتحديد إلى الإجابة عن الأسئلة الآتية:

1. ما واقع الدلالة العملية للاختبارات الإحصائية اللامعلمية المستخدمة في رسائل الماجستير وأطروحات الدكتوراة التربوية في الجامعة الأردنية في الفترة الزمنية (2017–2021)؟
2. هل توجد فروق ذات دلالة إحصائية بين تكرارات مستويات الدلالة العملية(دلالة صغيرة، دلالة متوسطة، دلالة كبيرة) للاختبارات الإحصائية اللامعلمية المستخدمة في رسائل الماجستير وأطروحات الدكتوراة التربوية في الجامعة الأردنية في الفترة الزمنية (2017–2021)؟
3. ما واقع قوة الاختبارات الإحصائية اللامعلمية المستخدمة في رسائل الماجستير وأطروحات الدكتوراة التربوية في الجامعة الأردنية في الفترة الزمنية (2017–2021)؟
4. هل توجد علاقة ارتباطية دالة إحصائيا عند مستوى الدلالة(0.05) بين قيم ثبات الأدوات المستخدمة لجمع البيانات لاختبار الفرضيات الإحصائية اللامعلمية وقيم الدلالة العملية لها في رسائل الماجستير وأطروحات الدكتوراة التربوية في الجامعة الأردنية في الفترة الزمنية (2017–2021)؟

أهداف الدراسة:

هدفت هذه الدراسة إلى تقييم واقع الدلالة العملية وقوة الاختبارات الإحصائية للاختبارات الإحصائية اللامعلمية المستخدمة في أطروحات الدكتوراة ورسائل الماجستير التربوية في الجامعة

الأردنية للسنوات (2017–2021) من خلال حساب حجم الأثر والتوزيع الإحصائي لمستويات الدلالة العملية(دلالة صغيرة، دلالة متوسطة، دلالة كبيرة) فضلاً عن حساب قوة الاختبار الإحصائي الخاص بكل اختبار متضمن في الدراسة، كما وهدفت الدراسة إلى البحث في العلاقة الارتباطية بين قيم ثبات الأدوات المستخدمة لجمع البيانات لاختبار الفرضيات الإحصائية اللامعلمية وقيم الدلالة العملية لها في رسائل الماجستير وأطروحات الدكتوراة التربوية في الجامعة الأردنية في الفترة الزمنية (2017–2021).

أهمية الدراسة:

إن اعتماد الباحثين قيم الدلالة الإحصائية منفردة لا يكفي ما لم يتم النظر إلى مؤشرات الدلالة العملية لاتخاذ قرارات ذات دقة ومصداقية تمكن من فهم النتائج التي تتضمنها الدراسات وتفسيرها. وحيث أن مفهوم الدلالة الإحصائية وكذلك مفهوم الدلالة العملية من المفاهيم التي حظيت باهتمام كثير من الباحثين للوصول إلى قرارات ذات دقة ومصداقية لتفسير الظواهر الاجتماعية والإنسانية، ونظرا لأهمية الدلالة العملية والدلالة الإحصائية وقوة الاختبارات الإحصائية، والانتشار الكبير للدراسات والأبحاث تأتي أهمية هذه الدراسة من حيث أنها:

- تُعد دراسة وصفية تبين مدى اهتمام الباحثين بالدلالة العملية وقوة الاختبارات الإحصائية اللامعلمية المستخدمة في أطروحات الدكتوراة ورسائل الماجستير التربوية في الجامعة الاردنية ضمن الفترة الزمنية (2017–2021).
- تقدم معلومات حول حالة الدلالة العملية وقوة الاختبارات الإحصائية اللامعلمية المستخدمة في أطروحات الدكتوراة ورسائل الماجستير التربوية في الجامعة الاردنية ضمن الفترة الزمنية (2017–2021).
- تقدم تقييماً لرسائل الماجستير وأطروحات الدكتوراة التربوية في الجامعة الاردنية والتي استخدمت الاختبارات الإحصائية اللامعلمية ضمن الفترة الزمنية (2017–2021) فيما يتعلق بالدلالة الإحصائية والدلالة العملية وقوة الاختبارات الإحصائية اللامعلمية المستخدمة الأمر الذي يعطي تصورا عن مدى دقة النتائج التي تم الحصول عليها ومصداقيتها والتي يمكن تقديمها لإدارة مكتبة الجامعة الاردنية ، الأمر الذي يمكنهم من عدم قبول الأبحاث والدراسات الخاصة بهم ما لم يتم حساب الدلالة العملية المرافقة لنتيجة الدلالة الإحصائية وكذلك حساب قوة الاختبارات الإحصائية المستخدمة للحصول على نتائج ذات قيمة عملية.

- حيث أن العملية البحثية عملية تراكمية، فغالبا ما يلجأ الباحثون إلى الأبحاث السابقة للحصول على مشكلة بحثية أو الاستفادة من نتائج الأبحاث السابقة في العملية البحثية، ومن هنا تُعد هذه الدراسة دليلاً للباحثين يبين مدى أهمية اقران الدلالة الإحصائية بالدلالة العملية كمؤشر على مصداقية النتائج، الأمر الذي يمكنهم من اتخاذ القرار بالرجوع إلى الدراسات والتي تم فيها حساب الدلالة العملية إلى جانب الدلالة الإحصائية.

التعريفات الاصطلاحية:

كان الاهتمام في هذه الدراسة البحثية منصبا على دراسة واقع الدلالة العملية وقوة الاختبارات الاحصائية المستخدمة في أطروحات الدكتوراة ورسائل الماجستير التربوية في الجامعة الاردنية والتي استخدمت الاختبارات الاحصائية اللامعلمية (كروسكال واليس Kruskal Wallis، مان ويتني Mann Whitney ، ويلككسون Wilcoxon) ضمن الفترة الزمنية 2017–2021، ومن هنا فإن التعريفات الاصطلاحية الخاصة بالدراسة هي:

1. الدلالة العملية: نسبة التباين في المتغير التابع الذي يمكن تفسيره من خلال المتغير المستقل.(Hass,2001).
2. الدلالة الإحصائية: أقصى احتمال لأن يرفض الباحث الفرضية الصفرية وهي صحيحة.(Alshareefen ،2017)
3. الاختبار الإحصائي اللامعلمي: مجموعة من الطرق الإحصائية البديلة عن الطرق الإحصائية المعلمية والتي تستخدم في حالات عدم تحقق الافتراضات حول المجتمع الذي تسحب منه العينة أو في حالة البيانات الاسمية والرتبية.(Allam،2005).
4. قوة الاختبار الإحصائي: احتمال رفض الفرض الصفري وهو في الواقع خاطئ.(2007، Albatsh & Abuzainah).

حدود الدراسة:

اقتصرت الدراسة على رسائل الماجستير وأطروحات الدكتوراة التربوية في الجامعة الأردنية في الفترة الزمنية (2017–2021) والتي استخدمت الاختبارات الإحصائية اللامعلمية (كروسكال واليس Kruskal Wallis ، مان ويتني Mann Whitney ، ويلككسون Wilcoxon).

الدراسات السابقة:

حظي موضوع الدلالة العملية وقوة الاختبارات الاحصائية اهتماما كبيرا من قبل الباحثين

على مر السنوات، فقد تنوعت الدراسات العربية وكذلك الأجنبية والتي درست هذا الموضوع ومن هذه الدراسات:

قام كوهن (Cohen,1962) بدراسة تضمنت مسحا شاملا لمجلة علم النفس الاجتماعي والشواذ المنشورة عام 1960. في هذه الدراسة قام كوهن بحساب قوة الاختبارات الإحصائية عند كل مستوى من مستويات حجم التأثير، وذلك في (80) دراسة تجريبية، وقد كان معدل القوة المحسوبة لجميع هذه الاختبارات هو 0.17، 0.46، 0.89 لحجم التأثير الصغير والمتوسط والكبير على التوالي. كما وجد الباحث بأن الوسيط لقوة الاختبارات الإحصائية المستخدمة هو 0.46 ولعل كوهن لجأ لاستخدام الوسيط لتجنب القيم المتطرفة والتي قد يكون لها تأثير في قيمة المتوسط الحسابي.

ونشر روسي (Rossi,1990) دراسة بعنوان القوة الإحصائية للبحث النفسي، إذ قام بتحليل قوة الاختبارات الإحصائية المستخدمة في مجلة علم النفس المنشورة عام (1982)، فضلاً عن مجلة علم النفس العيادي والإرشادي ومجلة علم النفس الاجتماعي، إذ اشتملت دراسته على (40000) اختبار إحصائي منشور في (1500) دراسة وكان معدل القوة لهذه الاختبارات 0.26 للتأثيرات الصغيرة، 0.64 للتأثيرات المتوسطة، 0.85 للتأثيرات الكبيرة.

وأجرى دانييل (Daniel,1993) دراسة تناولت قوة الاختبارات الإحصائية لمجلة الدراسة في التربية الموسيقية ما بين عامي (1987–1991) وقد تكونت عينة الدراسة من (255) اختباراً إحصائياً، وقد أشارت النتائج إلى أن متوسط قوة الاختبارات لحجم التأثير الصغير 0.13، ولحجم التأثير المتوسط 0.64، أما حجم التأثير الكبير فقد كانت قوة الاختبار الإحصائي له 0.97.

وأجرى محمود (Mahmoud,2003) دراسة تم فيها بحث واقع الدلالة العملية الخاصة برسائل الماجستير التربوية في جامعة مؤتة، وقد أشارت النتائج إلى أن نسبة الاختبارات التي تم استخدامها والتي كانت ذات دلالة إحصائية هي (45.2%)، وقد اقترن (17.8%) منها بدلالة عملية كبيرة و(11.17%) منها بدلالة عملية متوسطة و(71.03%) منها بدلالة عملية صغيرة.

كما بينت دراسة أجراها القضاة (Al-Qudah,2016) تناولت درجة التوافق بين نتيجة الدلالة الإحصائية والدلالة العملية الخاصة برسائل الماجستير التربوية في جامعة مؤتة، أن (1.7%) فقط من الاختبارات الإحصائية والبالغ عددها (1433) اختبارا إحصائيا تم إتباعها بمؤشرات الدلالة العملية، كما أُشير إلى أن (75.9%) من الاختبارات الإحصائية اقترنت بدلالة

عملية صغيرة، في حين (24.1%) اقترنت بدلالة عملية متوسطة وكبيرة.

كما قام الشريفين (Al_Shareefen,2017) بتقييم الدلالة الإحصائية والدلالة العملية وقوة الاختبارات الإحصائية لجميع الدراسات المنشورة في المجلة الأردنية في العلوم التربوية التي استخدمت التصاميم الكمية وتكونت عينة الدراسة من (1363) اختباراً إحصائياً وتم حساب كل من الدلالة العملية وقوة الاختبارات الإحصائية من جداول كوهن وقد تبين أن (75,79٪) من الفرضيات اقترنت بدلالة عملية صغيرة و (10,86٪) بدلالة عملية متوسطة كما تبين أن (54,86٪) من الفرضيات استقلت فيها الدلالة الإحصائية عن كل من الدلالة العملية وقوة الاختبار الإحصائي.

وقام المالكي (Al-Maliki,2020) بدراسة هدفت إلى تقدير قوة الاختبارات الإحصائية المعلمية في البحوث التربوية للمجلات الخليجية العربية على فترتين زمنيتين يفصل بينهما عشرون عاما (1997/1996– 2017/2016) بهدف معرفة التغير الذي طرأ على مستوى القوة الإحصائية في المجلات التربوية ومعرفة واقع استخدام تحليل القوة الإحصائية من قبل الباحثين واشتملت العينة على (79) دراسة تضمنت (875) اختبارا إحصائيا وتوصلت الدراسة إلى ضعف استخدام تحليل القوة الإحصائية في البحوث التربوية وإلى انخفاض في القوة الإحصائية وفق مستويات الدلالة العملية بين الفترتين (0,97 /0,79/ 0,24)، (0,20/0,70/0,93) على التوالي.

التعقيب على الدراسات السابقة:

كان الاهتمام في الدراسات السابقة منصبا على دراسة واقع الدلالة العملية وقوة الاختبارات الإحصائية المعلمية المستخدمة من قبل الباحثين في عديد من الدراسات الخاصة بهم، وفيما يتعلق بالاختبارات الإحصائية اللامعلمية فلم يكن هناك اهتمام خاص بها ، ومن خلال مراجعة التوصيات في الأبحاث السابقة والتي كانت تدعو الباحثين لإجراء مزيد من عمليات التقييم المتعلقة بواقع الدلالة العملية وقوة الاختبارات الإحصائية للأبحاث المنشورة في عديد من المجالات، تضمنت هذه الدراسة دراسة واقع الدلالة العملية وقوة الاختبارات الإحصائية بشكل خاص للاختبارات الإحصائية اللامعلمية (كروسكال واليس Kruskal Wallis ، مان ويتني Mann Whitney ، ويلككسون Wilcoxon).

منهج الدراسة:

تم استخدام المنهج الوصفي، إذ استخدم التحليل النوعي الوثائقي للدراسات الكمية من رسائل

الماجستير وأطروحات الدكتوراة التربوية في الجامعة الأردنية ضمن الفترة الزمنية (2017-2021)، وتم تحليلها باستعمال بطاقة تحليل تم تصميمها لتناسب الغرض، وتوفر القيم الكمية لمدخلات الصيغ الإحصائية المختلفة لحساب حجم الأثر (Effect Size) في الاختبارات الإحصائية اللامعلمية، ومن ثم وصف الظاهرة المدروسة وهي الدلالة العملية في الاختبارات الإحصائية اللامعلمية في رسائل الماجستير وأطروحات الدكتوراة التربوية في الجامعة الأردنية ضمن الفترة الزمنية (2017-2021) بالطرق المناسبة فضلاً عن الاستعانة بالبرامج الحاسوبية لحساب قوة الاختبارات الإحصائية.

مجتمع الدراسة:

تكون مجتمع الدراسة وعينتها من جميع رسائل الماجستير وأطروحات الدكتوراة التربوية والبالغ عددها (20) دراسة احتوت (5) رسائل ماجستير و (15) أطروحة دكتوراة ضمن الفترة الزمنية (2017-2021) وقد أُستخدمت فيها الاختبارات الإحصائية اللامعلمية (كروسكال واليس Kruskal Wallis، مان ويتني Mann Whitney ، ويلككسون Wilcoxon) ، ويعرض الجدول(1) أعداد الفرضيات الإحصائية التي تم حساب الدلالة العملية لها، ونسبها المئوية تبعا لنوع الاختبار المستخدم

الجدول(1): عدد الفرضيات الإحصائية تبعا لنوع الاختبار المستخدم

نوع الاختبار	عدد الفرضيات	النسبة
مان ويتنيMann Whitney	41	%33
ويلككسونWilcoxon	38	%31
كروسكال واليسKruskal Wallis	45	%36
الكلي	124	100

أداة الدراسة:

في ضوء أهداف الدراسة تم إعداد بطاقة تحليل تتضمن جداول خاصة بكل اختبار إحصائي لامعلمي، إذ تضمن كل جدول خلايا تم تعبئتها بالمعلومات اللازمة لحساب الدلالة العملية باستخدام القانون الخاص بكل اختبار إحصائي.

إجراءات الدراسة

لتحقيق الهدف من الدراسة تم تنفيذ الإجراءات الآتية:

1. حصر جميع أطروحات الدكتوراة ورسائل الماجستير التربوية في الجامعة الأردنية للسنوات (2017-2021) من خلال كشف تم الحصول عليه من مكتبة الجامعة الأردنية يحتوي

أسماء هذه الرسائل والأطروحات.

2. قراءة الرسائل والأطروحات وتسجيل قيم الدلالة الإحصائية الخاصة بكل اختبار إحصائي لامعلمي وكذلك القيم الخاصة بحساب مؤشرات الدلالة العملية وتفريغها في جداول إكسل خاصة أعدت لذلك، فقد تم إعداد جدول خاص بكل اختبار إحصائي لامعلمي منفردا، وبلغ عدد الجداول ثلاثة جداول لكل من اختبار مان ويتني، واختبار ويلككسون، واختبار كروسكال واليس.
3. حساب مؤشرات الدلالة العملية لكل اختبار إحصائي لامعلمي باستخدام المعادلة الخاصة به.
4. تصنيف قيم مؤشرات الدلالة العملية بحسب ما أشار إليه كوهن إلى ثلاثة مستويات هي: دلالة عملية صغيرة ودلالة عملية متوسطة ودلالة عملية كبيرة.
5. الاستعانة بالبرمجية الحاسوبية (Statistical Power Analysis) لحساب قوة الاختبارات الإحصائية اللامعلمية المتضمنة في الدراسة.
6. تصنيف قيم قوة الاختبارات الإحصائية اللامعلمية بحسب ما أشار إليه كوهن إلى ثلاثة مستويات هي: قوة اختبار صغيرة وقوة اختبار متوسطة وقوة اختبار كبيرة.
7. إستخدام اختبار مربع كاي لحسن المطابقة للكشف عن وجود فروق ذات دلالة إحصائية بين مستويات الدلالة العملية (الصغيرة، المتوسطة، الكبيرة) في الاختبارات الإحصائية اللامعلمية المستخدمة في رسائل الماجستير وأطروحات الدكتوراة التربوية في الجامعة الأردنية للسنوات (2017-2021)
8. استخدام معامل ارتباط بيرسون للبحث في العلاقة الترابطية بين الدلالة العملية ومعامل الثبات الخاص بأدوات جمع البيانات في كل من رسائل الماجستير وأطروحات الدكتوراة التربوية في الجامعة الأردنية ضمن الفترة الزمنية (2017-2021).

الأساليب الإحصائية المستخدمة:

لوصف واقع استعمال الدلالة العملية في الاختبارات الإحصائية اللامعلمية المستخدمة في رسائل الماجستير وأطروحات الدكتوراة التربوية في الجامعة الأردنية للسنوات (2017-2021) تم حساب التكرارات والنسب المئوية للفرضيات الدالة إحصائيا وتلك غير الدالة إحصائيا، ومن ثم تم حساب التكرارات والنسب المئوية ضمن كل مستوى من مستويات الدلالة العملية (صغير، متوسط، كبير)، وتم توظيف إحصائي مربع كاي لحسن المطابقة للكشف عن وجود فروق ذات دلالة

إحصائية بين مستويات الدلالة العملية، وأخيرا تم إستخدام معامل ارتباط بيرسون للبحث في العلاقة الترابطية بين الدلالة العملية ومعامل الثبات الخاص بأدوات جمع البيانات في كل من رسائل الماجستير وأطروحات الدكتوراة التربوية وكذلك لوصف واقع قوة الاختبارات الإحصائية اللامعلمية المستخدمة في رسائل الماجستير وأطروحات الدكتوراة التربوية في الجامعة الأردنية للسنوات (2017–2021) تم حساب التكرارات والنسبة المئوية للفرضيات الإحصائية المتضمنة في الدراسة، ومن ثم تم حساب التكرارات والنسبة المئوية ضمن كل مستوى من مستويات قوة الاختبار الإحصائي (صغير، متوسط، كبير).

نتائج الدراسة ومناقشتها:

أولا: الإجابة عن السؤال الأول ونصه " ما واقع الدلالة العملية للاختبارات الإحصائية اللامعلمية المستخدمة في رسائل الماجستير وأطروحات الدكتوراة التربوية في الجامعة الأردنية في الفترة الزمنية (2017–2021)؟

تم حساب حجم الأثر كمؤشر على الدلالة العملية لكل من الفرضيات التي تم فحصها باستخدام الاختبارات الإحصائية اللامعلمية (كروسكال واليس Kruskal Wallis، مان ويتني Mann Whitney، ويلككسون Wilcoxon)، وكانت قيم حجم الأثر الخاصة بكل اختبار إحصائي لامعلمي مصنفة وفق الفئات التي وضعها كوهن (Cohen,1988) كما في الجداول (2، 3، 4):

الجدول (2) التكرارات والنسب المئوية لكل فئة من فئات الدلالة العملية والدلالة الإحصائية الخاصة باختبار مان ويتني Mann Whitney

مستويات الدلالة العملية	الدلالة الإحصائية						حساب الدلالة العملية من قبل الباحث					
	دال		غير دال		الكلي		ذكرها		لم يذكرها		الكلي	
	التكرار	النسبة	التكرار	النسبة	التكرار	النسبة	التكرار	النسبة	التكرار	النسبة	التكرار	النسبة
صغير(0.1 0.3)	11	%100	30	%100	41	%100	0	%0	41	%100	41	%100
متوسط(0.4–0.5)	0	%0	0	%0	0	%0	0	%0	0	%0	0	%0
كبير(0.6– 1)	0	%0	0	%0	0	%0	0	%0	0	%0	0	%0
الكلي	11	%26	30	%74	41	%100	0	%0	41	%100	41	%100

يتبين من الجدول (2) أن العدد الكلي للفرضيات الإحصائية التي تم اختبارها باستخدام الإحصائي مان ويتني Mann Whitney (41) فرضية، تضمنت (11) فرضية دالة إحصائيا و(30) فرضية ليست ذات دلالة إحصائية، ولم يذكر الباحثون في أي منها مؤشرات الدلالة

العملية ومستواها، وبعد أن تم حساب الدلالة العملية لها أظهرت النتائج أن جميع الفرضيات ذات الدلالة الإحصائية والبالغ عددها (11) فرضية اقترنت بدلالة عملية صغيرة، وبنسبة (100%). وأن جميع الفرضيات غير الدالة إحصائيا وعددها (30) فرضية اقترنت بدلالة عملية صغيرة، وبنسبة (100%).

الجدول (3). التكرارات والنسب المئوية لكل فئة من فئات الدلالة العملية والدلالة الإحصائية الخاصة باختبار كروسكال واليس Kruskal Wallis

مستويات الدلالة العملية	الدلالة الإحصائية						حساب الدلالة العملية من قبل الباحث					
	دال		غير دال		الكلي		ذكرها		لم يذكرها		الكلي	
	التكرار	النسبة	التكرار	النسبة	التكرار	النسبة	التكرار	النسبة	التكرار	النسبة	التكرار	النسبة
صغير(0.01- 0.06)	0	%0	10	%29	10	%22	0	%0	10	%22	10	%22
متوسط(0.07-0.14)	0	%0	14	%40	14	%32	0	%0	14	%32	14	%32
كبير(0.15- 1)	10	%100	11	%31	21	%46	0	%0	21	%46	21	%46
الكلي	10	%23	35	%77	45	%100	0	%0	45	%100	45	%100

يتبين من الجدول (3) أن العدد الكلي للفرضيات الإحصائية التي تم اختبارها باستخدام اختبار كروسكال واليس Kruskal Wallis (45) فرضية، تضمنت (10) فرضيات ذات دلالة إحصائية، فضلاً عن ذلك كانت هناك (35) فرضية غير دالة إحصائيا، ولم تذكر مؤشرات الدلالة العملية لها ومستواها، وبعد حساب الدلالة العملية لها أظهرت النتائج أن جميع الفرضيات ذات الدلالة الإحصائية والبالغ عددها (10) فرضيات اقترنت بدلالة عملية كبيرة أي ما نسبته (100%)، في حين أظهرت النتائج أن توزيع مستويات الدلالة العملية للفرضيات غير الدالة إحصائيا كانت 29% صغيرة،40% متوسطة، و31% كبيرة.

الجدول (4). التكرارات والنسب المئوية لكل فئة من فئات الدلالة العملية والدلالة الإحصائية الخاصة باختبار ويلككسون Wilcoxon

مستويات الدلالة العملية	الدلالة الإحصائية						حساب الدلالة العملية من قبل الباحث					
	دال		غير دال		الكلي		ذكرها		لم يذكرها		الكلي	
	التكرار	النسبة	التكرار	النسبة	التكرار	النسبة	التكرار	النسبة	التكرار	النسبة	التكرار	النسبة
صغير(0.1- 0.3)	0	0	0	0	0	0	0	%0	0	%0	0	%0
متوسط(0.4-0.5)	1	%3	6	%55	7	%18	0	%0	7	%18	7	%18
كبير(0.6- 1)	26	%97	5	%45	31	%82	0	%0	31	%82	31	%82
الكلي	27	%71	11	%29	38	%100	0	%0	38	%100	38	%100

يتبين من الجدول (4) أن العدد الكلي للفرضيات الإحصائية التي تم اختبارها باستخدام

الإحصائي ويلككسون Wilcoxon (38) فرضية، تضمنت (27) فرضية ذات دلالة إحصائية، في حين بلغ عدد الفرضيات غير الدالة إحصائياً (11) فرضية، ولم يذكر الباحثون في أي منها مؤشرات الدلالة العملية ومستواها، وبعد أن تم حساب الدلالة العملية لها أظهرت النتائج أن هناك فرضية واحدة ذات دلالة إحصائية اقترنت بدلالة عملية متوسطة، أي ما نسبته (3%) في حين اقترنت باقي الفرضيات ذات الدلالة الإحصائية والبالغ عددها (26) فرضية بدلالة عملية كبيرة أي ما نسبته (97%). في حين أظهرت النتائج أن توزيع مستويات الدلالة العملية للفرضيات غير الدالة إحصائيا توزعت في الفئتين المتوسطة بنسبة55%، والكبيرة بنسبة45%.

وقد تبين من النتائج السابقة أن نسبة الفرضيات الإحصائية للاختبارات الإحصائية اللامعلمية التي اقترنت فيها الدلالة الإحصائية بدلالة عملية في رسائل الماجستير وأطروحات الدكتوراة التربوية في الجامعة الأردنية في الفترة الزمنية (2017–2021) هي 0%، أي أنه لم يتم إتباع الدلالة الإحصائية بالدلالة العملية من قبل الباحثين، إذ أن الباحثين اتخذوا القرارات المتعلقة بوجود أثر للمعالجة اعتمادا على قيمة الدلالة الإحصائية فقط، على الرغم من أن الدلالة الإحصائية غير كافية للاعتماد عليها لاتخاذ القرارات الموثوقة فيما يتعلق برفض الفرضية الصفرية بل يجب إتباعها بمؤشرات الدلالة العملية فضلاً عن ذلك فيما يتعلق باختبار مان ويتني، فإن جميع الفرضيات ذات الدلالة الإحصائية والبالغ عددها (11) فرضية اقترنت بدلالة عملية صغيرة، أي ما نسبته (100%) من الفرضيات الإحصائية ذات الدلالة الإحصائية الخاضعة لهذا الاختبار، الأمر الذي يثير عديداً من التساؤلات حول مدى مصداقية القرارات التي تم التوصل إليها بالاعتماد على قيمة الدلالة الإحصائية، أما فيما يتعلق باختبار كروسكال واليس، فقد بلغ عدد الفرضيات الدالة إحصائيا (10) فرضيات، أي ما نسبته (100%) من الفرضيات الدالة إحصائيا كانت جميعها ذات دلالة عملية كبيرة ، في حين أن هناك (11) فرضية أي ما نسبته (31%) من الفرضيات الخاضعة لهذا الاختبار غير دالة إحصائيا ولكنها ذات دلالة عملية كبيرة، الأمر الذي يستوجب التساؤل حول صحة اتخاذ القرارات المتعلقة بقبول الفرض الصفري في الوقت الذي تشير إليه النتائج إلى أن هناك تباينا في المتغير التابع يمكن تفسيره من خلال التباين في المتغير المستقل، وقد فشلت الاختبارات التي تم استخدامها في الكشف عن هذه الفروق، الأمر الذي يعزى لضعف الاختبار الإحصائي المستخدم أو ضعف التصميم الإحصائي الذي تم استخدامه لفحص الفرضية الصفرية.

ثانيا: الإجابة عن السؤال الثاني ونصه: "هل هناك فروق ذات دلالة إحصائية بين تكرارات مستويات الدلالة العملية للاختبارات الإحصائية اللامعلمية المستخدمة في رسائل الماجستير وأطروحات الدكتوراة التربوية في الجامعة الأردنية في الفترة الزمنية (2017-2021)؟" تم استخدام اختبار مربع كاي لحسن المطابقة وقد بلغت قيمة الاختبار (11.2) وكانت هذه القيمة ذات

دلالة إحصائية عند مستوى الدلالة (0.05) بدرجات حرية 2 الأمر الذي يشير إلى وجود فروق ذات دلالة إحصائية بين تكرارات مستويات الدلالة العملية المتوقعة والمشاهدة كما في الجدول(5) الأمر الذي يدلل على أن ما يتوقعه الباحث من تكرارات مستويات الدلالة العملية يختلف عما هو مشاهد فعليا، مما يؤكد ضرورة حساب مؤشرات الدلالة العملية من قبل الباحث.

الجدول (5) قيمة الاحصائي مربع كاي للكشف عن الفروق بين تكرارات مستويات الدلالة العملية المتوقعة والمشاهدة

قيمة الإحصائي مربع كاي	df	P value	alpha
11.2	2	0	0.05

ثالثا: الإجابة عن السؤال الثالث ونصه: "ما واقع قوة الاختبارات الإحصائية اللامعلمية المستخدمة في رسائل الماجستير وأطروحات الدكتوراة التربوية في الجامعة الأردنية في الفترة الزمنية (2017-2021)؟ قام الباحثان بحساب قوة الاختبارات الإحصائية اللامعلمية المتضمنة في الدراسة مصنفة تبعا لمستويات القوة التي وضعها كوهن كما في الجداول (6،7،8):

الجدول(6) تكرارات الفرضيات ونسبها المئوية الخاصة بالإحصائي مان ويتني Mann Whitney ضمن كل مستوى من مستويات قوة الاختبار الإحصائي

مستوى مؤشرات القوة	التكرار	النسبة المئوية
0-0.39	9	%82
0.40-0.59	2	%18
0.6-1	0	%0
الكلي	11	%100

تشير النتائج في الجدول (6) إلى أن 82% من الفرضيات التي تم اختبارها باستخدام الإحصائي مان ويتني Mann Whitney كانت ضمن المستوى الصغير من مستويات القوة و18% منها كانت ضمن المستوى المتوسط و0% منها كانت ضمن المستوى الكبير من مستويات القوة.

الجدول(7) تكرارات الفرضيات ونسبها المئوية الخاصة بالإحصائي كروسكال واليس Kruskal Wallis ضمن كل مستوى من مستويات

مستوى مؤشرات القوة	التكرار	النسبة المئوية
0.39-0	1	%10
0.59-0.40	2	%20
1-0.6	7	%70
الكلي	10	%100

تشير النتائج في الجدول (7) إلى أن 10% من الفرضيات التي تم اختبارها باستخدام الإحصائي كروسكال Kruskal واليس كانت ضمن المستوى الصغير من مستويات القوة و20% منها كانت ضمن المستوى المتوسط و70% منها ضمن المستوى الكبير من مستويات القوة.

الجدول(8) تكرارات الفرضيات ونسبها المئوية الخاصة بالإحصائي ويلككسون Wilcoxon ضمن كل مستوى من مستويات قوة الاختبار الإحصائي

مستوى مؤشرات القوة	التكرار	النسبة المئوية
0.39-0	2	%7
0.59-0.40	3	%11
1-0.6	22	%82
الكلي	27	%100

تشير النتائج في الجدول (8) إلى أن 7% من الفرضيات التي تم اختبارها باستخدام الإحصائي ويلككسون Wilcoxon كانت ضمن المستوى الصغير من مستويات القوة و11% منها كانت ضمن المستوى المتوسط و82% منها كانت ضمن المستوى الكبير من مستويات القوة.وبناء على النتائج أعلاه وفيما يتعلق بالاختبار مان ويتني Mann Whitney فقد اقترن 82% من الفرضيات الدالة إحصائيا الخاضعة لهذا الاختبار بقوة اختبار صغيرة و18% منها بقوة اختبار متوسطة، الأمر الذي يشير إلى عدم الدقة في القرار الإحصائي، ما يستوجب طرح عديد من التساؤلات حول دقة النتائج الخاصة بهذا الاختبار ومصداقيتها إذ أن الاختبار الإحصائي لم يكن ذا قوة كافية للكشف عن الفروق أما فيما يتعلق باختبار كروسكال واليسKruskal Wallis فقد بلغت نسبة الفرضيات الدالة إحصائيا الخاضعة لهذا الاختبار والتي اقترنت بقوة اختبار صغيرة 10% و20% من الفرضيات اقترنت بقوة اختبار متوسطة و70% منها اقترنت بقوة اختبار كبيرة، وأخيرا كان عدد الفرضيات الدالة إحصائيا والتي تم فحصها باستخدام اختبار ويلككسون Wilcoxon (27) فرضية، اقترن 7% منها بقوة اختبار صغيرة و11% منها بقوة اختبار متوسطة و82% منها بقوة اختبار كبيرة، الأمر الذي يشير إلى دقة النتائج الخاصة بهذا

الإحصائي حيث إن النسبة الأكبر من الفرضيات الدالة إحصائيا تم اتخاذ القرار فيها برفض الفرضية الصفرية باستخدام اختبارات ذات قوة كبيرة

رابعا: الإجابة عن السؤال الرابع ونصه: "هل توجد علاقة ارتباطية دالة إحصائيا عند مستوى الدلالة (0.05) بين قيم ثبات الأدوات المستخدمة لجمع البيانات لاختبار الفرضيات الإحصائية اللامعلمية وقيم الدلالة العملية لها في رسائل الماجستير وأطروحات الدكتوراة التربوية في الجامعة الأردنية في الفترة الزمنية (2017–2021)؟" تم حساب معامل ارتباط بيرسون بين قيم الدلالة العملية ومعاملات الثبات الخاصة بأدوات جمع البيانات المستخدمة في رسائل الماجستير وأطروحات الدكتوراة التربوية في الجامعة الأردنية ضمن الفترة الزمنية (2017–2021)، وقد بلغت قيمة معامل ارتباط بيرسون (0.42) وكانت هذه القيمة دالة إحصائيا عند مستوى الدلالة (0.05)، الأمر الذي يشير إلى وجود علاقة ارتباطية طردية بين معامل الثبات ومؤشرات الدلالة العملية إذ إنه بارتفاع قيم معاملات الثبات المسجلة لأدوات جمع البيانات التي تم استخدامها من قبل الباحثين كانت البيانات التي تم الحصول عليها باستخدام هذه الأدوات أكثر دقة ومصداقية، الأمر الذي يترافق معه ارتفاع قيم مؤشرات الدلالة العملية. ونجد من خلال قيمة معامل التحديد(0.176) أن ما نسبته 17.4% من التباين في قيم مؤشرات الدلالة العملية يعزى للزيادة في قيم الثبات لأدوات جمع البيانات.

التوصيات:

نظرا لأهمية الدلالة العملية وقوة الاختبارات الإحصائية المستخدمة من قبل الباحثين في عديد من الدراسات والأبحاث، ونتيجة لما تبين من ضعف اهتمام الباحثين بحساب الدلالة العملية إلى جانب الدلالة الإحصائية فقد توصلت هذه الدراسة إلى تقديم التوصيات الاتية للباحثين ومتخذي القرار في المجلات والدوريات البحثية:

- أن يقدم المشرفون توجيها للباحثين بضرورة حساب حجم الأثر كمؤشر للدلالة العملية، الأمر الذي يساعد على اتخاذ قرارات أكثر مصداقية وموثوقية.
- أن يعتمد المحكمون للدراسات والأبحاث المنشورة في المجلات التربوية، والمشرفين على رسائل الماجستير وأطروحات الدكتوراة تضمين الباحثين لحجم الأثر في أبحاثهم شرطا لقبولها، أومعيارا من معايير تقييمها.
- إجراء مزيد من دراسات ما وراء التحليل للأبحاث المنشورة في عديد من المجلات.

References:

Al_Batsh, M. & Abu Zainah, F (2007). **Research methodology** (1st ed). Amman: Dar Al_Maseera.

Al-Haj Mahmoud, F (2003). **The reality of the statistical significance, the practical significance, and the power of the test for the statistical tests used in master's theses of the Faculty of Educational Sciences at Mutah University,** Unpublished Master Thesis, Mutah University, Al_Karak, Jordan.

Al_Kilani, A. & Alshareefen, N (2011). **Introduction to research in educational and social sciences** (3st ed). Amman: Dar Al_Maseera.

Allam, S. (2005). **Parametric & nonparametric inferential statistical methods** (1st ed). Cairo: Dar AL-Fikr Al-Arabi.

Al_Maliki, F. (2020). The statistical power of educational research in the Arab Gulf Journals between the past and the present. **The Arab Gulf Message**,157(41),37-54.

Al_Najjar, A. (2005). The statistical and practical significance of the chi-square test in the published administrative research, Evaluation Study. **The Arab Journal of Administrative Sciences, 12**(5),169-193.

Al_Qudah, M. (2016). **The degree of agreement between the result of statistical significance and practical significance in the results of statistical analyzes in master's theses in the college of Educational Sciences at Mutah University**. Unpublished Master Thesis, Mutah University, Al_karak, Jordan.

Al_Shareefen, N. (2017). Meta analysis of the research published in the Jordanian Journal of Educational Sciences: Practical Significance and Test power. **Journal of Association of Arab Universities for Education and Psychology**, muj15(3),130-170.

Aron, A. & Aron, E. (1997). **Statistics for the behavioral and social sciences: Abrief course.** New Jersey: Prentice Hall.

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

Cohen, J. (1988). **Statistical power analysis for the behavioral sciences** (1st ed). New York: Academic Press.

Daniel, T. D. (1993). **Statistical Power analysis of the quantitive techniques used in the journal of research in music education,** 1987 through 1991. Doctoral dissertation, Auburn University.

Hasse, T. (2001). Statistical Significance should not be considered one of lifes guarantees: effect size is needed. **Educational and Psychological Measurement**,61,219-224.

Huston, L. (1993). Meaningfulness**,** 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, New Orleans, LA** (ERIC Document, No. 364608).

Kirk, R. E. (1996). Practical significance: a concept whose time has com. **Educational and Psychological Measurement,56,**746-759**.**

Konstantopoulos, S. (2008). Computing power of tests on the variance treatment effects in designs with two levels of nesting. **Multivariate Behavioral Research,43,**327-352.

Park, H. M (2008). Hypothesis **testing and statistical power of a test**. Working Paper. The diversity information technology services (UITS) Center for statistical and mathematical computing, Indiana University. Rrtrieved,2008, from http://www.indiana.edu/ststmath.

Rossi, L (1990). Statistical power Analysis: **American Education Research Journal**, 39(9), 391-401.

Wilkinson , L. (1992). Statistical significance testing : alternatives and considerations**.paper presented at the educational research association , Houston,TX,January31,1992.**