واقع استخدام الدلالة الإحصائية والدلالة العملية للاختبارات الإحصائية في رسائل الماجستير في الجامعات العمانية خلال الفترة (1998-2018)

إبراهيم بن سعيد بن حميد الوهيبي[1,*] داود بن عبد الملك الحدابي[2] حسين بن علي بن طالب الخروصي[3]
جامعة الشرقية، سلطنة عمان الجامعة الإسلامية العالمية (ماليزيا) جامعة السلطان قابوس(سلطنة عمان)

The reality of Using statistical significance and the practical significance of statistical tests in master's theses in Omani universities during the period (1998-2018)

Ibrahim AlWahaibi [1],*
ishalwahaibi@gmail.com
A'SHARQIYAH UNIVERSITY(Sultanate of Oman)

Dawood alHadabi[2]
Dawood@iium.edu.my
International Islamic University (Malaysia)

Hussein AlKharousi[3]
hussein5@squ.edu.om
3. Sultan Qaboos University (Sultanate of Oman)

تاريخ الاستلام:2019/12/19؛ تاريخ القبول:2020/01/28؛ تاريخ النشر:2022/02/28

ملخص: هدفت الدراسة إلى بحث واقع استخدام مؤشرات الدلالة الإحصائية والدلالة العملية في رسائل الماجستير التربوية بالجامعات العمانية، وهي جامعة السلطان قابوس، وجامعة نزوى، وجامعة ظفار، وجامعة صحار، خلال الفترة (1998-2018).وقد اتبعت الدراسة المنهج الوصفي التحليلي، وتم اختيار عينة عشوائية منتظمة بلغت (628) رسالة ماجستير، من جميع الجامعات بنسبة (33.14%) من مجتمع الدراسة والبالغ (1895) رسالة ماجستير، وتوصلت نتائج الدراسة إلى أن عدد الاختبارات الإحصائية المستخدمة في رسائل الماجستير بالجامعات العمانية بلغ (11926) اختبارا، وتم استخدام مؤشرات الدلالة العملية كمقاييس مكملة للدلالة الإحصائية في (7%) فقط من إجمالي الاختبارات الإحصائية، وكما أن أغلب نتائج اختبارات (ت) واختبارات تحليل التباين كانت ليست ذات دلالة إحصائية عند مستوى دلالة أقل من (0.05)، والعكس في نتائج اختبارات العلاقات (معامل ارتباط بيرسون، ومعامل الانحدار)، وتم حساب الدلالة العملية لجميع الاختبارات الإحصائية، وكانت متوسطة في إجمالي اختبارات (ت) واختبارات معامل ارتباط بيرسون بالجامعات العمانية، وصغيرة في إجمالي اختبارات تحليل التباين، وكبيرة في اختبارات معامل الانحدار.

الكلمات المفتاحية . الدلالة الإحصائية؛ الدلالة العملية؛الاختبارات الإحصائية

Abstract. The present study aimed to investigate the reality of the use of indicators of statistical significance and Practical significance in the Master's Thesis at Omani universities, Sultan Qaboos University, Nizwa University, Dhofar University, and Sohar University during 1998-2018. A random sample of (628) master's Theses was selected, which is (33.14%) from the total population (1895).
The results of the study showed that the number of statistical tests used in master's theses reached (11926) tests, and practical significance indicators were used as measures that complement the statistical significance in only (7%) of the total statistical tests, and that most of the results of (t) tests and analysis of variance tests were not statistically significant at a level of significance less than (0.05), and the opposite in the results of Pearson correlation coefficient and regression coefficient. The practical significance was calculated for all tests. It was found medium in (t) tests and Pearson correlation coefficient tests, small in analysis of variance tests, and large in regression coefficient tests.

Keywords: statistical tests; practical significance; statistical significance

***corresponding author**

1. مقدمة

لقد اهتمت سلطنة عمان بالبحث العلمي والتربوي منذ العقد الثاني من عمر النهضة الحديثة، حيث تم افتتاح أول جامعة حكومية في السلطنة وهي جامعة السلطان قابوس عام 1986م، والتي تعد أهم مؤسسة بحثية في السلطنة، وتضم تسعة مراكز بحثية تمثل مختلف الكليات العلمية والإنسانية، وتتبع هذه المراكز نائب رئيس الجامعة للدراسات العليا والبحث العلمي، وتمثل كلية التربية بالجامعة الجهة الرئيسة للبحث التربوي، وبدأت برامج الماجستير بها منذ العام 1992م بثلاثة برامج في المناهج وطرق التدريس، ووصلت حاليا إلى عشرة برامج للدراسات العليا في الكلية، كما تقوم الكلية بإصدار مجلة علمية محكمة منذ عام 1996م (المهدي، الشنفري، الفهدي، ولاشين، ٢٠١٤)، كما أن الجامعات الخاصة بالسلطنة -نزوى وصحار و ظفار والعربية المفتوحة والشرقية-، أولت اهتمامها بالبحث التربوي، فمثلا في جامعة نزوى تم اعتماد برامج الدراسات العليا (الماجستير) عام 2007م، في التخصصات التربوية، كالإدارة التربوية والمناهج وطرق التدريس والإرشاد النفسي والتوجيه، مما أثمر ذلك في رفد البحث التربوي بعدد كبير من الرسائل الجامعية التي أجيزت في هذه الجامعات.

ولقد قام الباحث بإجراء دراسة استكشافية مسحية اتبعت المنهج البيبليومتري (Bibliometric Methodology) لرسائل الماجستير في المجالات التربوية التي تمت إجازتها في الجامعات العمانية، وخلصت الدراسة إلى جملة من النتائج تم عرضها في الفصل الثالث من هذه الدراسة، ومن أهمها:

- الجامعات العمانية التي بها برامج للدراسات العليا هي خمس جامعات فقط: جامعة حكومية واحدة (جامعة السلطان قابوس)، وأربع جامعات خاصة وهي جامعات نزوى وصحار وظفار والعربية المفتوحة، وتم استبعاد الجامعة العربية المفتوحة لتطبيقها نظام الامتحان الشامل في برامج الماجستير ، وعدم الأخذ بنظام الرسالة إلا من قِبل طالبين فقط كما أفادت بذلك عمادة الدراسات العليا بالجامعة.
- عدد رسائل الماجستير في المجالات التربوية التي أُجيزت في الجامعات العمانية الخمس، منذ فتح أول برنامج للماجستير في كلية التربية بجامعة السلطان قابوس عام ١٩٩٢م حتى نهاية الفصل الثاني من العام الجامعي 2018/2017 ، بلغت عدد (1965) دراسة، واحتلت كلية التربية في جامعة السلطان قابوس النسبة الأكبر من هذه الرسائل حيث بلغ عددها (1316) بنسبة مئوية بلغت حوالي (67%).

وعلى الرغم من ضخامة هذا العدد من الأطروحات والرسائل إلا أنها لم تحظ بأي دراسة لتحليل وتقويم توجهاتها وخصائصها المنهجية والأساليب الإحصائية المستخدمة-على حد علم الباحث- سوى الدراستين اللتين تم ذكرهما سابقا واللتان أُجريتا في جامعة السلطان قابوس فقط، وهما دراسة عطاري (2004)، والتي هدفت إلى دراسة واقع اتجاهات البحث التربوي في سلطنة عمان من خلال تحليل عدد (265) أطروحة دكتوراه ورسالة ماجستير، تناولت التعليم في السلطنة في الفترة 1970 – 2002 والمجازة من أي جامعة من الجامعات العربية أو الأجنبية، وقد استخدمت هذه الدراسة المنهج البيبلومتري في تحليل هذه الرسائل، إلا أنها لم تتطرق إلى تحليل المنهجية البحثية، وتقويمها، وخاصة طرق اختيار العينة، وأدوات جمع البيانات، وكذلك لم تبحث في تحليل وتقويم الأساليب الإحصائية المستخدمة في تلك الرسائل، ولم تبحث في تتبع مؤشرات الدلالة العملية والقوة الإحصائية للاختبارات، أما الدراسة الثانية التي حصل عليها الباحث والتي أُجريت بجامعة السلطان قابوس هي دراسة حاردان(2014) فهدفت إلى تحديد خصائص المنهجية المستخدمة في رسائل الماجستير المجازة بكلية التربية بجامعة السلطان قابوس خلال الفترة (2002-2013)، وكذلك التحقق من توافر معايير البحث العلمي في المنهجية المستخدمة في هذه الرسائل.

ونظرا لعدم قيام (حاردان، 2014؛ عطاري، 2004) بتقويم المنهجية البحثية والأساليب الإحصائية المستخدمة في عينة دراستهما، ولم يتطرقا إلى تقويم مؤشرات الدلالات الإحصائية والعملية وقوة الاختبارات الإحصائية، إلا إشارة مختصرة فقط إلى حجم الأثر في دراسة (حاردان، 2014) حيث بينت نتائجها عدم توافر هذا المعيار الإحصائي في عينة دراستها، وكذلك مقرر الإحصاء الوحيد الذي يدرسه الطلبة الملتحقين ببرامج الماجستير في تخصصي المناهج وطرق التدريس والإدارة التربوية بكلية التربية بجامعة السلطان قابوس وهو "مساق الإحصاء النفسي والتربوي(علنف 6205) يقتصر على الأساليب الإحصائية

الأولية والمتوسطة" فقط (حاردان، 2014، ص143)، والأساليب الإحصائية المتوسطة التي أوضحتها حاردان(2014) في دراستها لا تتعدى تحليل التباين الأحادي فقط، كما أن رسائل الماجستير المنجزة في الجامعات الخاصة العمانية (نزوى، صحار، ظفار، العربية المفتوحة) منذ افتتاح برامج الدراسات العليا فيها لم تحظ حتى الآن –على حد علم الباحث- بأي دراسة لتحليلها وتقويمها بشكل عام أو بتقويم جانب معين منها كمؤشرات الدلالة العملية مثلا، وحيث إن الباحثين عادة ما يقتصرون على الدلالة الإحصائية فقط في تفسير نتائج دراساتهم ولا يهتمون بمؤشرات الدلالة العملية أو قوة الاختبارات الإحصائية (البارقي، 2012؛ الحداد، 2006؛ الدرابيع، 2003؛ الشريفين، 2017؛ محمود، 2003)، وطلبة الدراسات العليا في الجامعات العمانية حالهم كحال غيرهم من الباحثين، كلُ ذلك يؤكد على أهمية مراجعة هذا الكم الكبير من هذه الرسائل والدراسات التربوية، للوقوف على تقويم واقع استخدامات الدلالة الإحصائية في تفسير النتائج، ومدى استخدام طلبة الدراسات العليا بالجامعات العمانية لمؤشرات الدلالة العملية في دراساتهم، وهو ما تهدف إليه الدراسة الحالية.

1.1.أسئلة الدراسة

السؤال الأول: هل تقترن الدلالة الإحصائية بالد الة العملية في نتائج رسائل الماجستير بالجامعات العمانية؟

السؤال الثاني: ما واقع استخدام مؤشرات الدلالة العملية في رسائل الماجستير بالجامعات العمانية؟

2.1.أهداف الدراسة : تهدف هذه الدراسة إلى تحديد مستوى استخدم مؤشرات الدلالة العملية في رسائل الماجستير بالجامعات العمانية، ومعرفة مدى اقترانها بالدلالة الإحصائية.

3,1. أهمية الدراسة : تسعى الدراسة الحالية إلى توضيح أهمية استخدام مؤشرات الدلالة العملية في تفسير نتائج رسائل الماجستير بالجامعات العمانية، لذلك فإنها قد تُفيد المؤسسات العليا المسؤولة عن البحث التربوي في سلطنة عمان كمجلس البحث العلمي و عمادات الدراسات العليا في جامعة السلطان قابوس والجامعات الخاصة، من خلال النظرة الشاملة والمتكاملة عن درجة الثقة في نتائج الدراسات التربوية (رسائل الماجستير) في سلطنة عمان، وتحديد جوانب القصور في تفسير هذه النتائج، مما قد يسهم ذلك في تطوير إستراتيجية التخطيط والتطوير للبحث التربوي في السلطنة، وكذلك في رفدها بمعلومات كمية ونوعية عن واقع استخدام مؤشرات الدلالة العملية في الدراسات التربوية في سلطنة عمان، مما قد يساعد في توجيه برامج الدراسات العليا في الجامعات العمانية نحو الأفضل.

4,1. حدود الدراسة : **الحدود المكانية والزمانية:** اقتصرت هذه الدراسة على رسائل الماجستير التي أجيزت من الجامعات العمانية وهي (جامعة السلطان قابوس، جامعة نزوى، جامعة صحار، جامعة ظفار)، في المجالات التربوية وهي (المناهج وطرق التدريس، علم النفس، الإدارة التعليمية، والتربية الرياضية، أصول التربية، الإشراف التربوي، الإرشاد النفسي، الإرشاد والتوجيه)، والمجازة من هذه الجامعات، منذ اعتماد برامج الدراسات العليا بها وحتى نهاية الفصل الدراسي الثاني من العام الأكاديمي 2018/2017، والجدول(1) المشار إليه في الملحق (1) يوضح الفترة الزمنية لرسائل الماجستير في كل جامعة من الجامعات محل الدراسة.

الحدود الموضوعية: اقتصرت هذه الدراسة على رسائل الماجستير في المجالات التربوية التي أجيزت من الجامعات العمانية التي استخدمت تصاميم كمية واستخدمت الإحصاء الاستدلالي في الإجابة عن أسئلتها، وبالتحديد الاختبارات المعلمية الآتية (اختبارات ت T-Test، اختبارات تحليل التباين (الأحادي، الثنائي، المتعدد)، معامل ارتباط بيرسون، معامل الانحدار)

5,1. مصطلحات الدراسة:

رسائل الماجستير: يقصد بها الباحث جميع الرسائل العلمية التربوية في مرحلة الدراسات العليا الأولية المجازة من كليات التربية أو العلوم الاجتماعية في الجامعات العمانية (جامعة السلطان قابوس، جامعة نزوى، جامعة ظفار، جامعة صحار)، منذ انطلاقة برامج الدراسات العليا في هذه الكليات وحتى نهاية الفصل الدراسي الثاني من العام الأكاديمي 2018/2017.

الدلالة الإحصائية Statistical Significance: تمثل الدلالة الإحصائية احتمالية العلاقة بين عدد من المتغيرات بسبب معين غير الصدفة (Kenton,2019)، ويعرفها الباحث إجرائيا بمستوى الدلالة الملحق بنتائج الاختبارات الإحصائية المستخدمة في

رسائل الماجستير المجازة بالجامعات العمانية ، والتي بناء عليه يتم قبول أو رفض الفرضيات الصفرية، وعادة ما يكون هذا المستوى يساوى القيمة (0.05).

الدلالة العملية Practical Significance: تعرف الدلالة العملية بمقدار الخروج عن الفرضية الصفرية (Cohen,1988)، ويتم تعريفها إجرائيا بحجم الفروق الفعلية بين المتوسطات الحسابية لمستويات المتغيرات المستقلة، التي يتم الحصول عليها من خلال تطبيق معادلات رياضية.

6,1. الإطار النظري والدراسات السابقة .لم يتوقف الأمر عند إصدار البيانات من الجمعيات العالمية المختصة في الإحصاء مثل American Psychiatric Association (APA) وAmerican Statistical Association (ASA)التي تبين الاستخدامات الخاطئة للدلالة الإحصائية، بل تطور إلى أبعد من ذلك بكثير حيث قدمت اثنان من المجلات العلمية المرموقة (American Statistician, NATURE) توصياتهما بأن مصطلح "الأهمية الإحصائية" قد تقاعد (Johnson,2019)؛ وبناء على هذه التوصية فإن بيان الجمعية الأمريكية للإحصاء (ASA) الذي صدر عام 2016م قد تم توقيفه، وفي شهر أكتوبر (2017)، عقدت الجمعية (ASA) ندوة حول الاستدلال الإحصائي، استمرت لمدة يومين، دعت فيه الباحثين إلى الكتابة الصريحة حول الأهمية الإحصائية وتفسيراتها المختلفة والجدل حولها، وتم في الندوة وضع الأسس لإصدار عددٍ خاصٍ من مجلة الإحصاء الإمريكية American Statistician حول هذا الموضوع، (Wasserstein&Schirm& Lazar, ,2019) وفي شهر مارس من عام (2019م)، تم إصدار هذا العدد الخاص رقم (73)، وعُرض فيه أكثر من 40 بحثًا حول "الاستدلال الإحصائي في القرن الحادي والعشرين: عالم يتجاوز P 0.05>"، بين فيه الباحثون آراءهم في الأهمية الإحصائية، وكانوا بين مؤيد ومعارض لها، وإن أهم توصية لتلك البحوث والتقارير لُخِّصت في سبع كلمات (قبول عدم اليقين يكون مدروسا ومنفتحا ومتواضعا) (Acceptuncertainty. Be thoughtful, open, andmodest)، واختصرت في الحروف " ATOM" (Wasserstein&Schirm& Lazar, ,2019, p2)، وأوضح المعارضون للأهمية الإحصائية بأن عبارة "ذات دلالة إحصائية" يجب التخلي عنها من قِبل الباحثين ورؤساء تحرير المجلات العلمية (Hurlbert& Levine & Utts,2019)، وقدم هيرلبرت وليفين ويوتس (Hurlbert& Levine & Utts,2019) مقترحا للتخلي عن الأهمية الإحصائية، أيَّده ووافق عليه (48) من المؤسسات الإحصائية، والإحصائيين والعلماء الآخرين، ويشمل هذا المقترح ما يأتي:

أولا: يجب على الباحثين عرض قيم p فقط بدون تصنيفها بـ"ذات دلالة إحصائية" أو "ليست ذات دلالة إحصائية".

ثانيا: تقديم طلبات رسمية ومباشرة إلى محرري ولجان التحرير في المجلات العلمية لتعديل تعليماتهم للمؤلفين لتشمل عدم السماح للدراسات التي لا تلتزم بالمقترح رقم (1) أعلاه، وقد يتم إنجاز هذه المهمة تحت رعاية ASA ، والجمعية الإحصائية الملكية ، وربما غيرها من الجمعيات الإحصائية، ويرى الباحث بأن من أهم مشكلات الدلالة الإحصائية التي أوجدت المعارضة ضدها هو تصنيفها إلى مستويين "ذات دلالة إحصائية" و "ليست ذات دلالة إحصائية"، ويتفق الباحث مع مقترح هيرلبرت وليفين ويوتس Hurlbert& Levine &Utts في ترك ذكر هذا التصنيف، إلا أن ترك هذا التصنيف لا يؤدي إلى ترك استخدام اختبارات الدلالة الإحصائية، فهذه الاختبارات لها أهميتها وفائدتها في التحليل الإحصائي للبيانات، لكن المشكلة التي سببها الانتقادات ضدها تكمن في سوء استعمالها فقط.

أما كمتز (Kmetz,2019) فيرى أن تتم معالجة قضية التخلص من الأهمية الإحصائية و قيم p بهدوء والتحلي بالصبر، ولا يشجع على الهجوم المباشر ضدها، لكونها تمكنت في البحث العلمي وتعززت في أوساط الباحثين لفترة تزيد على الخميسن عاما؛ لذلك يحتاج الأمر إلى مزيد من الوقت، واقترح في بادي الأمر ضرورة الإشارة إلى بيان ASA عند تقديم دراسة أو بحث أو ورقة عمل، أو مراجعة إلى أي مجلة أو مؤسسة تحرير، ويتفق الباحث مع طريقة كمتز Kmetz؛ حيث إن الدعوة إلى التخلص من الأهمية الإحصائية باتت غير مجدية حيث إنها ولدت منذ فترة ليست بالقليلة، ولا يزال صداها يتردد بين فترة وأخرى إلا أنها لم تُؤتي ثمارها حتى الآن؛ لذلك فإن هذا يدل على أنها فكرة مستعجلة وقد يشوبها بعض الخلل، والتأني في معالجة مشكلة الاستخدامات الخاطئة للدلالة الإحصائية قد يفتح الآفاق نحو تصحيحها وتطويرها وتحسينها، بدلا من التخلي عنها مباشرة.

كما قام كل من أمرهين و جرينلاند و مكشين (Amrhein& et al.,2019) بكتابة مسودة رسالة "تقاعد الأهمية الإحصائية" يدعون فيها إلى وقف استخدام قيم P بالطريقة التقليدية ثنائية التفرع لتحديد ما إذا كانت النتيجة تدحض فرضية علمية أو تدعمها، والتخلي عن مفهوم الأهمية الإحصائية بأكمله، وتمت دعوة المؤلفين ومحرري المجلات إلى التعليق على هذه الرسالة والتوقيع عليها إذا وافقوا على دعوة التخلي عن الدلالة أو الأهمية الإحصائية بالطريقة الحالية، ونظرا للانتشار الواسع للاستخدامات الخاطئة للدلالة الإحصائية فقد وافق على هذه الرسالة (250) خلال الـ 24 ساعة الأولى، وبعد حوالي أسبوع بلغ عدد الموقعين أكثر من (800) شخصا كلهم من الإحصائيين والباحثين السريريين والطبيين وعلماء الأحياء وعلماء النفس من أكثر من 50 دولة، ومن جميع القارات باستثناء القارة القطبية الجنوبية، ووصف أحد المدافعين هذا التفاعل الكبير ضد سوء استخدام الدلالة الإحصائية وقيم p بأنه "ضربة جراحية ضد الاختبارات غير المدروسة ذات الأهمية الإحصائية"، وعلى الرغم من أن عدد الموقعين بلغ أكثر من (800) شخصا في ظرف أسبوع واحد، فإن الباحث يرى بأن هذا العدد وإن كان كبيرا في حجمه؛ فإنه لا يزال قليلا على مستوى الباحثين وعلماء الإحصاء في مختلف دول العالم، كما أن هذا العدد لا يدل على إجماع المختصين بالإحصاء على ضرورة التخلي عن الدلالة أو الأهمية الإحصائية، حيث لا يمكن حصر المختصين بالإحصاء، فلو افترضنا جدلا أن عددهم يصل إلى (16000) شخص، فإن المعارضين يمثلون (5%) فقط، كما أن هؤلاء الموقعين ليسوا كلهم يدعون إلى تقاعد الأهمية الإحصائية، والاستغناء عنها بشكل كامل، حيث إن بعضهم يعارض سوء استعمالها فقط، ولا يدعوا إلى تركها.

وقام لوينيدس (Loannidis,2019) بتحليل ملخصات البحوث الطبية الحيوية التي تستخدم قيم p خلال الفترة (1990-2015)، وتوصل إلى أن 96% منها كانت قيم p تساوي (0.05) أو أقل، مما يُشير إلى تحيز البحوث إلى النتائج ذات الدلالة الإحصائية، وخلصت دراسته إلى التشكيك في النتائج "ذات دلالة إحصائية" عند مستوى (0.05)، وأوصى بعدم استخدام قيم P، ويرى الباحث بأن لوينيدس Ioannidis استعجل في توصيته بعدم استعمال قيم p؛ حيث إنه لم يوضح ولم يبحث الأسباب التي جعلت (96%) من نتائج البحوث التي حللها تكون ذات دلالة إحصائية، فقد تكون هذه النتائج صادقة، وقد يكون من ضمن الأسباب الخطأ الذي ارتكبته بعض المجلات المحكمة التي لا تنشر إلا الدراسات التي تشير نتائجها إلى دلالة إحصائية.

وعلى الرغم من كثرة النقد الموجه إلى اختبارات الدلالة الإحصائية، والدعوة إلى ما يسمى بـ"تقاعد الأهمية الإحصائية" والتخلص من قيم p و البحث عن بدائل أخرى، إلا أن هناك فريقا آخر يدعو إلى بقاء هذا النوع من الاختبارات الإحصائية، حيث يعتقد جودمن (Goodman,2019) بصعوبة التخلص من الأهمية الإحصائية وقيم p، وتفسير ذلك ليس فلسفيًا ولا علميًا، بل اجتماعيا؛ حيث إن جميع الباحثين يستخدمون هذا النوع من الاختبارات منذ فترة طويلة، فالتخلص منها بالكلية يشبة الاستغناء عن المال، ويطرح العديد من الأسئلة مثل: هل سنحصل على ممارسة عملية جيدة إذا تم التخلص من القيم P؟ ، هل سيؤدي التخلص من قيم P إلى تحسين التصميم التجريبي؟، هل سيتحسن القياس؟، هل سيساعد على مواءمة السؤال العلمي مع تلك التحليلات؟ ويُشير إلى أن الادعاءات التي ينادي بها المعارضون للأهمية الإحصائية لم تتحقق منذ أكثر من (90) سنة، لذلك تعزى مشكلات القيم P إلى حد كبير إلى التطبيق غير المدروس لاختبار أهمية الفرضية الصفرية (NHST) وسوء استخدامها بشكل غير مناسب، وهذا ما أكده الباحث سابقا مع طريقة كمتز Kmetz.

وأشار جودمن (Goodman,2019) إلى أن الأكاديمية الوطنية للعلوم National Academy of Sciences في أمريكا أصدرت عام 2018 تقريراً بعنوان "تعزيز النزاهة في مجال البحوث" تم فيه تعريف النزاهة البحثية إلى ما وراء الطرق التقليدية للتزوير والاحتيال والانتحال ليشمل مجموعة متنوعة من الممارسات الاستنتاجية المضللة التي تسمى "ممارسات البحث الضارة"، حيث يتضح أن العديد من هذه الممارسات -إن لم يكن جميعها- تنطوي على انحراف عن التصميم الإحصائي والسلوك والتحليل والتفسير الإحصائي، ودعا جودمان Goodman إلى الحاجة إلى التغيير الاجتماعي الكبير في المؤسسات الأكاديمية، وفي المجلات، وفيما بين وكالات التمويل والهيئات التنظيمية، ويقترح إقامة شراكة "مع حركات الإصلاح العلمي والإصلاحيين في التخصصات والمجلات ووكالات التمويل والهيئات التنظيمية لتعزيز ومكافأة العلوم" القابلة للتكرار "وتقليل تأثير الأهمية الإحصائية على النشر والتمويل والترويج.

كما أن كولقوهون (Colquhoun,2019) يؤكد على ضرورة تقليل الحوافز التي تعطى للباحثين لإنتاج عمل غير موثوق به، وأوضح بأن نقد الجمعة الأمريكية للإحصاء SAS لقيم P كان نقدا مقنعا، حيث إنها فشلت في التوصل إلى أي استنتاجات حول ما ينبغي القيام به حيال الاستخدامات الخاطئة للدلالة الإحصائية؛ لذلك توقع بأنه ذلك البيان لا يكون له تأثير كبير على مستخدمي الإحصاءات، ويذكر كريستين لينوكس Kristin Lennox بأن الإحصائيين لم يتفقوا حتى الآن على حل مناسب تجاه هذه المشكلة؛ لذلك لم توجد توصيات جذرية وواضحة وصريحة نحو الاستمرارية في استخدام الأهمية الإحصائية أم العدول عنها البتة.

ويعارض كمبيل و جوستابسون (Campbell & Gustafson,2019) فكرة أن عتبة الدلالة الإحصائية الأكثر صرامة ستؤدي إلى أن تكون العلوم المنشورة أقل موثوقية، بل يؤكدون إلى أن الانتقال من $p < 0.05$ إلى $p < 0.005$ قد يكون مفيدًا للعلوم المنشورة من حيث الموثوقية، مما يشير إلى أنهم يؤيدون الدلالة والأهمية الإحصائية بل ويدعون إلى ضرورة أن تكون قيم p أكثر انخفاضا.

وشكك هوبارد (Hubbard,2019) في نجاح الجهود التي بذلتها مؤخراً الجمعية الإمريكية للإحصاء ASA لتحسين الممارسة الإحصائية، والمتمثلة في مواجهة سوء استخدام وتفسير أهمية الفرضية الصفرية (NHST) والقيم p، وقدم هوبارد Hubbard أدلة دامغة على أن هذه ستكون مهمة صعبة للغاية، حيث إن هذه المشكلة معروفة، ولم تفعل ASA شيئًا لمنع استخدامها خلال الفترة 1960-2007، على العكس من ذلك فقد زاد توظيف NHST خلال تلك الفترة، وتوقع هوبارد بأن ASA قد تنظر لاحقا في توصيتها بالتخلي عن استخدام القيم p.

وقام ريكر و آخرون (Fricker& et al., 2019) بتقييم التحليلات الإحصائية المستخدمة في 31 بحثا في مجلة علم النفس الاجتماعي الأساسي والتطبيقي (BASP) المنشورة عام 2016م، بعد أن حظرت المجلة على الباحثين استخدام قيم P في بحوثهم، وتوصلوا في بحثهم هذا إلى أن مؤلفي علم النفس الاجتماعي يميلون إلى المبالغة في استنتاجاتهم حيث تتجاوز ما يمكن أن تدعمه بيانات دراساتهم؛ مما يؤكد على عدم الاستعداد لترك الدلالة الإحصائية، وعدم وجود البدائل الجذرية التي تغني عنها، وهو ما أكده ماثيوز Matthews في دراسته (2019).

وأشار ماثيوز (Matthews,2019) إلى عدم وجود توافق في الآراء بشأن البدائل المقبولة للاستخدامات الخاطئة الروتينية لاختبارات أهمية الفرضية الصفرية (NHST)، الأمر الذي يترك للباحثين المجال في الاستمرار في استخدام اختبارات NHST، بغض النظر عن عيوبها؛ لذا درس ماثيوز Matthews إمكانية تحليل المصداقية (AnCred) لحل هذا المشكلة، باستخدام أمثلة من واقع الحياة، حيث أوضح قدرة ANCred على تزويد الباحثين بإطار عمل بسيط ولكنه قوي لتقييم نتائج الدراسة التي تتجاوز ثنائية النتيجة الإحصائية (دالة إحصائيا/ أو ليست دالة إحصائيا)، وأكد على أن أي اقتراح يشجع على إجراء تغييرات في الممارسات الاستدلالية يجب أن يقبل وجود كل شيء في اختبارات الفرضية الصفرية NHST.

ويؤكد كينيدي شافير (Kennedy-Shaffer,2019) على الدور الذي لعبته اختبارات الأهمية الإحصائية وقيم p في تطور العلوم الإحصائية، خلال القرن العشرين، منذ أن طورها فيشر عام (1925م)، ويدعو إلى فهم كيف نشأت هذه الطرق وانتشرت، وكيف أصبحت $p<0.05$ معيارًا للاستدلال العلمي، والميزة التي قدمتها في ذلك الوقت، وكيف تم تفسيرها، حيث إنه يمكن لهذا المنظور التاريخي أن يسترشد بعمل الإحصائيين اليوم من خلال تشجيع التفكير الناقد المدروس في كيفية إدراك العلماء لعملهم، بما في ذلك البدائل المقترحة لقيم p.

أما جرينلاند (Greenland,2019) فيعزو المشكلة في الأهمية الإحصائية إلى مصادر الانتقادات الخاطئة لقيم P، والتي من أهمها تفسير قيم p الأقل من (0.05) بعبارة "ذات دلالة إحصائية" والأكبر من (0.05) بعبارة " ليست ذات دلالة إحصائية"؛ لذلك يقترح كتابة قيم p كما هي فقط بدون التفسير السابق.

ويتضح مما سبق بأن كل باحث ينظر إلى المشكلة من منظوره الخاص، فمثلا مشكلة تصنيف قيم p إلى عبارة "ذات دلالة إحصائية" و "ليست ذات دلالة إحصائية"، شدد فيها كل من هيرلبرت وليفين ويوتس Hurlbert& Levine &Utts و جعلها سببا للتخلي عن الدلالة الإحصائية البتة، بينما كانت نظر لوينيدس Ioannidis إليها أقل حدة حيث أوصى بعدم استعمال قيم P؛ لأن نتائج الدراسات التي حللها تتحيز نحو النتائج ذات الدلالة الإحصائية، أما جرينلاند Greenland فكانت نظرته

إيجابية حيث إنه عزا مشكلة الأهمية الإحصائية إلى تصنيف قيم P إلى مستويين، واقترح كتابة قيم p كما هي بدون التصنيف الثنائي، ويتفق الباحث مع الرأي الثالث وهو رأي جرينلاند Greenland، حيث إن هذا الرأي يعالج الأمور بنظرة إيجابية ويدعو إلى إيجاد الحلول المناسبة للمشكلات.

2. الطريقة والأدوات: نظرا لكون الدراسة تهدف إلى تحليل واقع مؤشرات الدلالة العملية في رسائل الماجستير بالجامعات العمانية؛ فإن المنهج المناسب لها هو المنهج الوصفي التحليلي، والذي يعتمد على دراسة الظواهر والوقائع كما هي موجودة في المجتمع.

1.2. مجتمع الدراسة: يتكون مجتمع الدراسة من جميع رسائل الماجستير في المجالات التربوية، والمجازة من مؤسسات التعليم العالي بسلطنة عمان، منذ بداية إنشائها وحتى نهاية الفصل الدراسي الثاني من العام الأكاديمي 2018/2017، وهذه المؤسسات هي جامعة السلطان قابوس وهي الجامعة الحكومية الوحيدة بسلطنة عمان، وأربع جامعات خاصة وهي: جامعة نزوى، جامعة صحار، جامعة ظفار، والجامعة العربية المفتوحة، وهذه الجامعات هي التي لديها برامج الماجستير في التربية في سلطنة عمان، وللحصول على البيانات الدقيقة لهذه البرامج قام الباحث بالتواصل المباشر والكتابي بالوحدات المسؤولة عن برامج الدراسات العليا بهذه الجامعات، وذلك من أجل التعرف على أنواع برامج الدراسات العليا التي تنفذها هذه الجامعات، وبداية افتتاحها، ونظام الدراسة فيها، وكذلك من أجل الحصول على قوائم مطبوعة أو إلكترونية بجميع رسائل الماجستير المجازة منها، وتبين من تلك الإجراءات أن بعض هذه الجامعات يطبق نظام كتابة الرسالة البحثية فقط، وهي جامعة السلطان قابوس، وبعضها لديها النظامان: نظام كتابة الرسالة البحثية، أو نظام الامتحان الشامل، وهي جميع الجامعات الخاصة الأربع، إلا أن الجامعة العربية المفتوحة تبين أن أغلب طلابها يأخذون نظام الامتحان الشامل، ولم يجد الباحث إلا طالبين فقط اختارا نظام كتابة الرسالة البحثية، لذلك تم استبعاد هذه الجامعة من مجتمع الدراسة، واقتصر المجتمع على ثلاث جامعات خاصة بالإضافة إلى جامعة السلطان قابوس، وبلغ مجتمع الدراسة (1895) رسالة ماجستير.

2.2. عينة الدراسة: من خلال جدول تحديد حجم العينة المقترح من قِبل Krejcie& Morgan(1970)، فإن العدد الفعلي لعينة الدراسة بلغ (628) رسالة ماجستير، وبنسبة (33.14%) من إجمالي المجتمع.

3,2. أداة الدراسة: لتحقيق أهداف الدراسة؛ قام الباحث بالاطلاع على مجموعة من الدراسات والبحوث (المالكي، 2018؛ Hartgerink , van Aert, Nuijten, Wicherts, & van Assen ,2016؛ البارقي، 2012؛ الحداد، 2006؛ الجودة، 2004؛ محمود، 2003؛ Hess,2003؛ WARD,2002؛ بابطين، 2002؛ إبراهيم، 2000) وذلك للاستفادة منها في تصميم أداة الدراسة، والتي تمثلت في استمارة تحليل المحتوى لجمع البيانات الإحصائية للاختبارات الإحصائية المراد تحليلها.

4,2. المعالجة الإحصائية: للإجابة على أسئلة الدراسة؛ تم استخدام الأساليب الإحصائية الآتية من خلال برنامجي EXCEL و SPSS، حيث تم استخدام الإحصاءات الوصفية: التكرارات والنسب المئوية، المتوسطات الحسابية، الانحرافات المعيارية، كما تم استخدام مؤشرات (d, μ^2، f، F^2)؛ لحساب مؤشرات الدلالة العملية للاختبارات الاحصائية المحددة في الدراسة.

3. النتائج ومناقشتها.

1,3. نتائج السؤال الأول:هل تقترن الدلالة الإحصائية بالدلالة العملية في نتائج رسائل الماجستير بالجامعات العمانية؟

للإجابة على هذا السؤال فقد تم عد جميع الاختبارات الإحصائية اختبارات (ت)، اختبارات تحليل التباين، اختبارات معامل ارتباط بيرسون، اختبارات معامل الانحدار التي تضمنتها هذه الدراسة، برسائل الماجستير بالجامعات العمانية، وتحديد أي من هذه الاختبارات تم حساب الدلالة العملية لها، ويوضح الجدول(2) المشار إليه في الملحق(1) أن إجمالي الاختبارات الإحصائية المستخدمة في رسائل الماجستير بجميع الجامعات العمانية الأربع بلغت (11926) اختبارا، ويُعدُّ هذا الرقم كبيرا إذا ما قورن بأعداد الاختبارات الإحصائية التي فحصتها الدراسات السابقة التي اطلع عليها الباحث، حيث إن أعلى عدد لها كان في دراسة (البارقي،2012) والذي بلغ (2223) اختبارا، وبلغت نسبة الاختبارات الإحصائية بهذه الدراسة أكثر من خمسة أضعاف الاختبارات الإحصائية التي حصرها البارقي(2012)، وبذلك تميزت هذه الدراسة بكثرة الاختبارات الإحصائية المشمولة فيها، وقد

يعود سبب ذلك إلى أن الفترات الزمنية التي حددتها عينات الدراسات السابقة كانت محدودة، في حين اعتمدت هذه الدراسة كل الفترات الزمنية التي بدأت فيها برامج الماجستير بالجامعات العمانية، فوصلت الفترة الزمنية إلى (20) سنة دراسية كما هو الحال في جامعة السلطان قابوس حيث شملت عينة الدراسة رسائل الماجستير المجازة منذ 1998م حتى 2018م، كما أن كبر حجم العينة بهذه الدراسة له دور في زيادة عدد الاختبارات الإحصائية، حيث بلغت عينة الدراسة (628) والتي تمثل نسبة (33.14%) من إجمالي المجتمع والذي بلغ (1895) رسالة ماجستير، وبهذا فإن نتائج هذه الدراسة تكون أكثر دقة وأكثر مصداقية، وتعدُّ هذه النتائج المرجع الأول والأساسي لمعرفة واقع الدلالة العملية في رسائل الماجستير في سلطنة عمان.

وكان لاختبارات (ت) بأنواعها الثلاثة (ت لعينة واحدة، ت لعينتين مستقلتين، ت لعينتين مرتبطتين) نصيبٌ أكبر من هذه الاختبارات، من حيث الاستخدام؛ حيث بلغت نسبة استخدامها (43.5%)، وجاءت اختبارات تحليل التباين بأنواعه الثلاثة (الأحادي والثنائي والمتعدد)، في المرتبة الثانية من حيث الاستخدام، حيث بلغت نسبتها (35.7%)، كما أوضحت النتائج بأن طلبة الماجستير في الجامعات العمانية لا يميلون إلى استخدام اختبارات معامل الانحدار في دراساتهم، لذلك كانت نسبة استخدام هذا النوع من الاختبارات ضئيلة جدا لم تتجاوز (2.5%)، وكان أشهر الاختبارات المستخدمة في رسائل الماجستير بالجامعات العمانية هو اختبار ت لعينتين مستقلتين، يليه اختبار تحليل التباين الأحادي ثم معامل ارتباط بيرسون بنسب استخدام بلغت (36.84%، 30.45%، 18.55%) على التوالي.

أما فيما يخص الدلالة العملية في الرغم من كثرة الاختبارات الإحصائية التي تضمنها رسائل الماجستير بالجامعات العمانية إلا ان الغالبية العظمى منها اقتصرت على مؤشرات الدلالة الإحصائية فقط، ولم تستخدم مؤشرات الدلالة العملية إلا نسبة (7.24%) فقط منها، كما أن اختبارات تحليل التباين الثنائي والثلاثي هي أكثر الاختبارات التي تم حساب الدلالة العملية لها مقارنة ببقية الاختبارات حيث بلغت نسبة استخدام الدلالة العملية (27.5%) يليها اختبارات (ت) لعينتين مرتبطتين بنسبة مئوية بلغت (17%)، وبالنسبة لأشهر ثلاثة اختبارات إحصائية وأكثرها استخداما هي: اختبار ت لعينتين مستقلتين، وتحليل التباين الأحادي، ومعامل ارتباط بيرسون، فهي أقل الاختبارات اقترانا بالدلالة العملية في الجامعات العمانية.

واختلفت نتائج هذه الدراسة مع نتائج مجموعة من الدراسات (kirk,1996; Sun, Pan, & Wang, 2010) التي أوضحت نتائجها إلى أن حوالي ما لا يقل عن (50%) من الاختبارات الإحصائية تم حساب الدلالة العملية لها، بينما تتفق نتائج الدراسة مع نتائج الكثير من الدراسات، حيث تتكرر هذه النتائج في العقود الثلاثة الأخيرة، فأقدم دراسة اطلع عليها الباحث وأشارت إلى ضعف استخدام مؤشرات الدلالة العملية هي دراسة الصياد(1988)، وتوالت بعدها الدراسات المتفقة معها في هذه النتائج مثل (الوهيبي، الحدابي، والخروصي، 2019؛ المالكي،2018؛ القضاة، 2016؛ البارقي، 2012؛ الشمراني، 2012؛ Gentry & Peters,2009؛ الحداد، 2006؛ الجودة، 2004؛ الدرابيع، 2003؛ محمود، 2003؛ بابطين، 2002؛ إبراهيم، 2000؛Plucker,1997)، وهذا يشير إلى أن طلبة الماجستير في الجامعات العمانية يكتفون بنتائج الدلالة الإحصائية فقط للحكم على دقة وأهمية النتائج التي يتوصلون إليها، وهذا مؤشر سلبي، والأخذ بنتائج وتوصيات هذه الدراسات فيه مخاطرة، حيث أن الدلالة الإحصائية ما هي إلا خطوة أولى للحكم على صحة النتائج وأهميتها ودقتها، ولا بد من استخدام مؤشرات أخرى مكملة لمؤشرات الدلالة الإحصائية، كمؤشرات الدلالة العملية، حيث توجد أسباب مختلفة تؤدي إلى الدلالة الإحصائية للنتائج كارتفاع حجم العينة مثلا. وقد يعود سبب عدم اهتمام طلبة الماجستير في الجامعات العمانية باستخدام مؤشرات الدلالة العملية كأساليب مكملة للدلالة الإحصائية إلى الآتي:

- قلة المراجع العربية التي تتحدث عن مؤشرات الدلالة العملية وأهميتها وطرق استخدامها، وهذا ما لاحظه الباحث –سابقا- أثناء اطلاعه على الأدب التربوي.

-عدم تدريس هذه المؤشرات في مناهج ومقررات الإحصاء لطلبة الماجستير، كما أشارت إلى ذلك (حاردان، 2014)، وتقتصر مقررات الإحصاء ومناهج البحث على الإحصاء الوصفي والإحصاء الاستدلالي البسيط والمتوسط فقط.

- الخلط بين الدلالة الإحصائية للنتائج ودلالتها العملية، والاعتقاد بأن الدلالة الإحصائية كافية للحكم على أهمية النتائج، وهذا ما أكدته الدراسة الاستطلاعية التي قام بها الباحث إبان دراسته هذه، ويتفق الباحث مع (سلامة، 2004؛ Little,2001; Hubbard & Armstrong,2006) الذين يؤكدون على أهمية إدراك الباحثين لطبيعة الدلالة الإحصائية، و فهمها الفهم الصحيح.

2,3. نتائج السؤال الثاني:ما واقع استخدام مؤشرات الدلالة العملية في رسائل الماجستير بالجامعات العمانية؟

للإجابة على هذا السؤال تم حساب الدلالة العملية باستخدام مؤشرات الدلالة العملية التي تم التطرق إليها في الإطار النظري لهذه الدراسة، ومن ثم تم إيجاد المتوسطات الحسابية والانحرافات المعيارية لها لكل الاختبارات الإحصائية، برسائل الماجستير بالجامعات العمانية؛ وذلك لمعرفة الواقع الفعلي لمستويات الدلالة العملية، وتم تصنيفها حسب المعيار الذي اقترحه (الوهيبي، 2020)، ويشير الجدول (3) المشار إليه في الملحق(1)إلى أن الدلالة العملية المحسوبة لإجمالي اختبارات (ت) كانت متوسطة بناء على المعيار الذي اقترحه (الوهيبي، 2020)، في رسائل الماجستير بالجامعات العمانية، وتتفق هذه النتائج مع نتائج مجموعة من الدراسات التي أشارت إلى أن مستوى الدلالة العملية لاختبارات (ت) يتراوح بين المتوسط والصغير (الصياد، 1988؛ إبراهيم، 2000؛ محمود، 2003؛ الدرابيع، 2003؛ الجودة،2004؛ الحداد، 2006)، بينما تختلف مع دراسات (المالكي، 2018؛ البارقي، 2012؛ الصائغ، 1996) التي توصلت إلى أن الدلالة العملية لإجمالي اختبارات (ت) كانت كبيرة، كما يشير الجدول (4) إلى تفاوت مستويات الدلالة العملية للاختبارات الإحصائية المستخدمة بالجامعات العمانية، فكانت كبيرة في بعض أنواع الاختبارات (اختبار ت لعينة واحدة ولعينتين مرتبطتين، تحليل التباين المتعدد، واختبارات معامل الانحدار)،وتتفق هذه النتيجة مع نتائج دراسة المالكي(2018)، ويؤكد الباحث على ما ذكره المالكي في دراسته (ص129) من أن سبب ارتفاع الدلالة العملية لاختبار ت لعينتين مرتبطتين " قد يعود إلى دور المعالجة المستخدمة، لا سيما أن غالبية اختبارات ت لمجموعتين مرتبطتين تستخدم في البحوث التجريبية وشبه التجريبية؛ لإيجاد دلالة الفروق بين الاختبار القبلي والبعدي، بهدف إيضاح فعالية المعالجة، أو البرنامج المستخدم، وكلما كانت هذه المعالجة أو البرنامج ذات فعالية أكبر؛ زاد متوسط الفروق بين الاختبار القبلي والبعدي للمجموعة التجريبية، أو بين الاختبارين البعدي للمجموعتين التجريبية والضابطة، مما يجعل قيمة الدلالة العملية ترتفع،... "، كما أن الدلالة العملية تكون أحيانا كما في أكثر الاختبارات الإحصائية استخداما في رسائل الماجستير بالجامعات العمانية وهي اختبار ت لعينتين مستقلتين، واختبار تحليل التباين الأحادي، وتتفق هذه النتائج مع نتائج دراسات (المالكي، 2018؛ القضاة، 2016؛ البارقي، 2012؛ الحداد، 2006؛ الصائغ، 1996) التي توصلت إلى مستوى صغير للدلالة العملية لهذا النوع من الاختبارات، كما اتفقت مع دراسة (النجار، 2005) التي توصلت إلى انخفاض الدلالة العملية لاختبار كاي تربيع، بينما اختلفت مع نتائج (الجودة، 2004؛ الدرابيع، 2003؛ إبراهيم، 2000)، وحيث أن الدلالة العملية لاختبار ت لعينتين مستقلتين ولتحليل التباين الأحادي مرتبطة بمقدار الفروق بين المتوسطات كالفروق بين متوسطات المجموعتين التجريبية والضابطة أو الفروق بين متوسطات النوع الاجتماعي أو مستويات المسميات الوظيفية في متغير ما، لذلك فإن انخفاض أو صغر الدلالة العملية يدل على أن الفروق بين هذه المتوسطات ليست كبيرة وإنما هي فروق صغيرة أو لا تمثل فروقاً حقيقية أو جوهرية، مما يدل على أن المعالجات التي استخدمتها رسائل الماجستير في الجامعات العمانية لم يكن لها أثر كبير، وقد يكون سبب ذلك ما يأتي:

- قد تكون هذه المعالجات المستخدمة لم يتم تطبيقها بالطريقة الصحيحة.
- يُحتمل أن تكون بعض عينات تلك الدراسات لم تستوعب الهدف من الدراسة، ولم يتم توضيح أهداف الدراسة لها بصورة جيدة، مما جعلها غير مهتمة بدقة البيانات والمعلومات التي تقدمها من خلال أدوات الدراسة المعطاة، وقد لمس الباحث بعضا من هذه الملاحظات من خلال عمله في مركز الخوارزمي للدراسات والاستشارات التربوية، حيث لاحظ أن بعض أفراد العينة في بعض الدراسات غير مكترث بالاستبانة المقدمة إليه، فتارة نجد أنهم يختارون خياراً واحداً لجميع فقرات الاستبانة، وتارة يتركون فقرات عديدة فيها، وتارة يُحجمون عن تقديم بعض البيانات الديموغرافية لها،...إلخ، كما أن اللامبالاة باستبانات الدراسات هي سمة ظاهرة بشكل واضح في المجتمعات العربية، فبعض الباحثين يعانون من وجود فاقد في الأدوات الموزعة إلى العينة.

4. التوصيات والمقترحات:

بناء على النتائج السابقة يوصي الباحث بالآتي:

1 - إعادة صياغة مقررات الإحصاء ومناهج البحث في الجامعات العمانية؛ بحيث تحتوي على مؤشرات الدلالة العملية بأنواعها المختلفة.

2 - توجيه طلبة الماجستير بأهمية تضمين نتائجهم لبعض مؤشرات الدلالة العملية.

3 - إعادة تطبيق هذه الدراسة لكل جامعة من الجامعات العمانية على حدة.

4 - إجراء دراسة للتعرف على أثر حجم العينة في الدلالة الإحصائية والدلالة العملية بالجامعات العمانية.

المراجع

إبراهيم، إبراهيم رشاد إبراهيم الحاج(2000). *واقع الدلالة الإحصائية والدلالة العملية وقوة الاختبارات الإحصائية المستخدمة في بحوث مجلة دراسات العلوم التربوية في الأعوام (1997-1998) في الجامعة الأردنية* (رسالة ماجستير غير منشورة). كلية الدراسات العليا، الجامعة الأردنية، الأردن.

بابطين، عادل أحمد(2002). *مشكلات الدلالة الإحصائية في البحث التربوي وحلول بديلة* (رسالة ماجستير غير منشورة). جامعة أم القرى، مكة المكرمة، المملكة العربية السعودية.

البارقي، طلال هيازع (2012). *واقع الدلالة الإحصائية والدلالة العملية للبحوث المنشورة في مجلة جامعة أم القرى للعلوم التربوية والاجتماعية والإنسانية، في المدة 1425-1430هـ* (رسالة ماجستير غير منشورة). جامعة أم القرى، مكة المكرمة، المملكة العربية السعودية.

الجودة، ماجد محمود شريف(2004). *الدلالة الإحصائية والدلالة العملية وقوة الاختبار للأبحاث المنشورة في مجلة أبحاث اليرموك سلسلة العلوم الإنسانية والاجتماعية (1985-2001)* (رسالة ماجستير غير منشورة). كلية الدراسات العليا، جامعة اليرموك، الأردن.

حاردان، زينب بنت سهيل بن سعيد(2014). *تقويم منهجية رسائل الماجستير المجازة بكلية التربية في جامعة السلطان قابوس خلال الفترة (2002-2013)* (رسالة ماجستير غير منشورة). جامعة السلطان قابوس، سلطنة عمان.

الحداد، سعدة أحمد محمد (2006). *قوة الاختبارات الإحصائية وواقع الدلالة الإحصائية والدلالة العملية في بحوث مجلة الهيئة القومية للبحث العلمي في الجماهيرية العظمى* (رسالة ماجستير غير منشورة). كلية الدراسات العليا، جامعة أم درمان، السودان.

الدرابيع، ماهر يونس(2003). واقع الدلالات الإحصائية والعملية وقوة الاختبارات الإحصائية المستخدمة في بحوث مجلة مؤتة للبحوث والدراسات. مؤتة للبحوث والدراسات – العلوم الانسانية والاجتماعية- الأردن، 18(1)، 151-176.

سلامة، حسن علي حسن (2004). الدلالة الإحصائية والدلالة العلمية في البحوث التربوية. المجلة التربوية- جامعة سوهاج- مصر، *20*، 3-14.

الشريفين، نضال كمال محمد(2017). ما وراء التحليل للأبحاث المنشورة في المجلة الأردنية في العلوم التربوية : الدلالة العملية وقوة الاختبار. مجلة اتحاد الجامعات العربية للتربية وعلم النفس-سوريا، 15(3)، 130-170.

الشمراني، محمد موسى (2012). استخدام مقاييس الدلالة العملية لحجم التأثير في الحكم على قياس أهمية نتائج البحوث العلمية. مجلة كلية التربية بالمنصورة-مصر، 78(3)، 1-30.

الصائغ، ابتسام حسن مدني (1996). *الدلالة الإحصائية والدلالة العملية لاختبارت وف: دراسة تحليلية تقويمية من خلال رسائل الماجستير التي قدمت في كلية التربية بجامعة أم القرى حتى عام 1415هـ* (رسالة ماجستير غير منشورة). جامعة أم القرى، مكة المكرمة، المملكة العربية السعودية.

الصياد، عبد العاطي (1988). الدلالة العملية وحجم العينة المصاحبتين للدلالة الإحصائية لاختبار tفي البحث النفسي والتربوي (دراسة تقويمية). *جامعة الزقازيق*، جمهورية مصر العربية.

عطاري، توفيق(2004). اتجاهات البحث التربوي في سلطنة عمان من خلال تحليل رسائل الماجستير والدكتوراه التي تتناول التقويم في السلطنة في الفترة 2002- 1970 م.مجلة اتحاد الجامعات العربية-الأمانة العامة لاتحاد الجامعات العربية-الأردن، (44)، 161-196.

القضاة، محمد عواد سالم (2016). *درجة التوافق بين نتيجة الدلالة الإحصائية والدلالة العملية في نتائج التحليلات الإحصائية في رسائل الماجستير في كلية العلوم التربوية في جامعة مؤتة* (رسالة ماجستير غير منشورة). جامعة مؤتة، الأردن.

المالكي، فهد بن عبدالله بن عمر (2018). *واقع الدلالة الإحصائية والدلالة العملية وقوة الاختبار الإحصائي، دراسة تطبيقية على بحوث المجلات التربوية بالجامعات السعودية* (رسالة دكتوراه غير منشورة). جامعة أم القرى، مكة المكرمة، المملكة العربية السعودية.

محمود، فريال محمود محمد الحاج (2003). *واقع الدلالة الإحصائية والدلالة العملية وقوة الاختبار للاختبارات الإحصائية المستخدمة في رسائل الماجستير الصادرة عن كلية العلوم التربوية في جامعة مؤتة* (رسالة ماجستير غير منشورة). كلية الدراسات العليا، جامعة مؤتة، الأردن.

المهدي، ياسر فتحي؛ والشنفري، عبدالله مبارك؛ والفهدي، راشد سليمان؛ ولاشين، محمد عبد الحميد(٢٠١٤). آليات مقترحة لتفعيل دور البحث التربوي في صنع السياسة التعليمية بسلطنة عمان من وجهة نظر الباحثين والممارسين. المجلة التربوية الدولية المتخصصة – المجموعة الدولية للاستشارات والتدريب- *الأردن* ، *3*(11)، 153-174.

الوهيبي، إبراهيم بن سعيد(2020).مؤشرات الدلالات الإحصائية و العملية في الدراسات التربوية بالجامعات العمانية: دراسة تحليلية لرسائل الماجستير (رسالة دكتوراه غير منشورة). الجامعة الإسلامية العالمية بماليزيا.

الوهيبي، إبراهيم بن سعيد؛ والحدابي، داود بن عبدالملك؛ والخروصي، حسين بن علي(2019). واقع مؤشرات الدلالة الإحصائية وحجم الأثر في رسائل الماجستير المجازة بكلية التربية بالجامعة الإسلامية العالمية الماليزية خلال الفترة (1993-2017). مجلة كلية التربية- جامعة أسيوط- مصر، *35*(11).

References

Al-Barqi, Talal Hayazah (2012). The reality of the statistical significance and the practical significance of the research published in the Journal of Umm Al-Qura University for Educational, Social and Human Sciences, in the period 1425-1430 AH (unpublished master'sthesis). Umm Al-Qura University, Makkah Al-Mukarramah, Kingdom of Saudi Arabia. (In Arabic).

Al-Haddad, Saada Ahmed Mohamed (2006). The power of statistical tests and the reality of statistical significance and practical significance in the research of the Journal of the National Authority for Scientific Research in the Great Jamahiriya (unpublished master's thesis). College of Graduate Studies, Omdurman University, Sudan. (In Arabic).

Al-Mahdi, Yasser Fathy; Al Shanfari, Abdullah Mubarak; Al-Fahdi, Rashid Suleiman; Walsheen, Mohamed Abdel Hamid (2014). Suggested mechanisms to activate the role of educational research in making educational policy in the Sultanate of Oman from the point of view of researchers and practitioners. Specialized International Educational Journal - International Group for Consultation and Training - Jordan, 3(11), 153-174. (In Arabic).

Al-Maliki, Fahd bin Abdullah bin Omar (2018). The reality of statistical significance, practical significance and the strength of the statistical test, an applied study on the research of educational journals in Saudi universities (unpublished doctoral thesis). Umm Al-Qura University, Makkah Al-Mukarramah, Kingdom of Saudi Arabia. (In Arabic).

Al-Sayegh, Ibtisam Hassan Madani (1996). Statistical significance and practical significance of the T and F test: An analytical and evaluative study through master's theses presented at the College of Education at Umm Al-Qura University until 1415 AH (unpublished master's thesis). Umm Al-Qura University, Makkah Al-Mukarramah, Kingdom of Saudi Arabia. (In Arabic).

Al-Sayyad, Abdel-Aty (1988). The practical significance and sample size associated with the statistical significance of the t-test in psychological and educational research (an evaluation study). Zagazig University, Arab Republic of Egypt. (In Arabic).

Al-Shamrani, Muhammad Musa (2012). Using measures of practical significance for the size of the impact in judging the measurement of the importance of scientific research results. Journal of the Faculty of Education in Mansoura - Egypt, 78(3), 1-30. (In Arabic).

Al-Sharifin, Nidal Kamal Muhammad (2017). Beyond analysis of research published in the Jordanian Journal of Educational Sciences: practical significance and test power. Journal of the Association of Arab Universities for Education and Psychology - Syria, 15 (3), 130-170. (In Arabic).

Al-Wahaibi, Ibrahim bin Saeed (2020). Indicators of statistical and practical indications in educational studies in Omani universities: an analytical study of master's theses (unpublished doctoral thesis). International Islamic University Malaysia. (In Arabic).

Al-Wahaibi, Ibrahim bin Said; Al-Hadabi, Daoud bin Abdul-Malik; and Kharousi, Hussein bin Ali (2019). The reality of the indicators of statistical significance and the size of the impact in the approved master's theses at the Faculty of Education at the International Islamic University of Malaysia during the period (1993-2017). Journal of the Faculty of Education – Assiut University - Egypt, 35 (11). (In Arabic).

Amrhein, V. , Greenland, S. , &McShane, B. (2019). Retire statistical significance. *NATURE*, 567 , 305-307.

Attari, Tawfiq (2004). Trends in educational research in the Sultanate of Oman through the analysis of master's and doctoral theses dealing with evaluation in the Sultanate in the period 1970-2002. Journal of the Federation of Arab Universities - General Secretariat of the Federation of Arab Universities - Jordan, (44), 161-196. (In Arabic).

Babtain, Adel Ahmed (2002). Statistical significance problems in educational research and alternative solutions (unpublished master's thesis). Umm Al-Qura University, Makkah Al-Mukarramah, Kingdom of Saudi Arabia. (In Arabic).

Campbell, H., Gustafson, P. (2019). The World of Research Has Gone Berserk: Modeling the Consequences of Requiring 'Greater Statistical Stringency' for Scientific Publication. *The American Statistician,73*(1), 358-373. https://doi.org/10.1080/00031305.2018.1555101, accessed 12 May. 2019.

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

Colquhoun, D. (2019). The False Positive Risk: A Proposal Concerning What to Do About p-Value. *The American Statistician, 73*(1), 192–201. https://doi.org/10.1080/00031305.2018.1529622, accessed 11 May. 2019.

Darabi', Maher Younes (2003). The reality of statistical and practical indications and the strength of the statistical tests used in the research of Mutah Journal for Research and Studies. Mutah for Research and Studies - Humanities and Social Sciences - Jordan, 18(1), 151-176. (In Arabic).

Fricker, R., Burke, K., Han, X., &Woodall, W. (2019). Assessing the Statistical Analyses Used in Basic and Applied Social Psychology After Their p-Value Ban. *The American Statistician, 73*(1), 374-384, https://doi.org/10.1080/00031305.2018.1537892, accessed 12 May. 2019.

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

Goodman, S.N.(2019). Whyis GettingRid of p-Values So Hard? Musings on Science and Statistics. *The American Statistician, 73*(1), 26-30. https://doi.org/10.1080/00031305.2018.1558111, accessed 11 May. 2019.

Greenland, S. (2019). Valid p-Values Behave Exactly as They Should: Some Misleading Criticisms of p-Values and Their Resolution With s-Values. *The American Statistician, 73*(1), 106-114, https://doi.org/10.1080/00031305.2018.1529625, accessed 13 May. 2019.

Hardan, Zainabbint Suhail bin Saeed (2014). Evaluating the methodology of the Master's theses in the College of Education at Sultan Qaboos University during the period (2002-2013) (unpublished master's thesis). Sultan Qaboos University, Sultanate of Oman. (In Arabic).

Hartgerink, C. H. J., van Aert, R. C. M., Nuijten, M. B., Wicherts, J. M., & van Assen, M. A. L. M. (2016). Distributions of p-values smaller than .05 in psychology: What is going on? *PEERJ, 4*, [e1935]. https://doi.org/10.7717/peerj.1935.

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

Hubbard, R. (2019). Will the ASA's Efforts to Improve Statistical Practice Be Successful? Some Evidence to the Contrary. *The American Statistician, 73*(1), 31-35. https://doi.org/10.1080/00031305.2018.1497540, accessed 12 May. 2019.

Hubbard, R. , & Armstrong, J. S. (2006). Why We Don't Really Know What Statistical Significance Means: Implications for Educators. *Journal of Marketing Education,28*(2),114-120. DOI: 10.1177/0273475306288399

Hurlbert, S., Levine, R., &Utts, J. (2019). Coup de Grâce for a Tough Old Bull: 'Statistically Significant' Expires. *The American Statistician, 73*(1), 352-357. https://doi.org/10.1080/00031305.2018.1543616, accessed 12 May. 2019.

Ibrahim, Ibrahim Rashad Ibrahim Al-Haj (2000). The reality of statistical significance, practical significance, and the strength of statistical tests used in the research of the Journal of Educational Sciences Studies in the years (1997-1998) at the University of Jordan (unpublished master's thesis). College of Graduate Studies, University of Jordan, Jordan.(In Arabic).

Ioannidis, J. (2019). What Have We (Not) Learnt From Millions of Scientific Papers With p-Values?. *The American Statistician, 73*(1), 20-25. https://doi.org/10.1080/00031305.2018.1447512, accessed 13 May. 2019.

Johnson .V. E.(2019). *Is it the end of 'statistical significance'? The battle to make science more uncertain.* https://theconversation.com/is-it-the-end-of-statistical-significance-the-battle-to-make-science-more-uncertain-114161. accessed 8 May. 2019

Kennedy-Shaffer, L. (2019).Before $p < 0.05$ to Beyond $p < 0.05$: Using History to Contextualize p-Values and Significance Testing. *The American Statistician, 73*(1), 82-90. https://doi.org/10.1080/00031305.2018.1537891, accessed 13 May. 2019.

Kenton,W.(2019). *Statistical Significance*. Investo pedia, https://www.investopedia.com/terms/s/statistically_significant.asp, accessed 1 June.2019.

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

Kmetz, J. (2019). Correcting Corrupt Research: Recommendations for the Profession to Stop Misuse of p-Values. *The American Statistician, 73*(1), 36-45. https://doi.org/10.1080/00031305.2018.1518271, accessed 12 May. 2019.

Little, J.(2001). UNDERSTANDING STATISTICAL SIGNIFICANCE: A CONCEPTUAL HISTORY, J. *TECHNICAL WRITING AND COMMUNICATION,31*(4), 363-372.

Mahmoud, Feryal Mahmoud Mohamed El-Hajj (2003). The reality of statistical significance, practical significance and test power of statistical tests used in master's theses issued by the College of Educational Sciences at Mutah University (unpublished master's thesis). College of Graduate Studies, Mutah University, Jordan. (In Arabic).

Matthews, R. (2019). Moving Toward the Post $p < 0.05$ Era Via the Analysis of Credibility. *The American Statistician, 73*(1), 202-212. https://doi.org/10.1080/00031305.2018.1543136, accessed 12 May. 2019.

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

Quality, Maged Mahmoud Sharif (2004). Statistical significance, practical significance and test power of research published in the Yarmouk Research Journal, Human and Social Sciences Series (1985-2001) (unpublished master's thesis). College of Graduate Studies, Yarmouk University, Jordan. (In Arabic).

Qudah, Muhammad Awad Salem (2016). The degree of concordance between the result of statistical significance and practical significance in the results of statistical analyzes in master's theses at the College of Educational Sciences at Mutah University (unpublished master's thesis). Mutah University, Jordan. (In Arabic).

Salama, Hassan Ali Hassan (2004). Statistical significance and scientific significance in educational research. Educational Journal - Sohag University - Egypt, 20, 3-14. (In Arabic).

Sun, S., Pan, W., & Wang, L. L. (2010). A Comprehensive Review of Effect Size Reporting and Interpreting Practices in Academic Journals in Education and Psychology. *Journal of Educational Psychology. Advance online publication*, (September 20), doi: 10.1037/a0019507. accessed 25 Mrs. 2019.

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

الملاحق:

ملحق(1): جداول الدراسة.

الجدول(1): الفترة الزمنية لرسائل الماجستير في كل جامعة من الجامعات العمانية

الجامعة	الفترة الزمنية لرسائل الماجستير
جامعة السلطان قابوس	**2018-1992**
جامعة نزوى	**2018-2007**
جامعة صحار	**2018-2011**
جامعة ظفار	**2018-2012**

الجدول(2): أعداد الاختبارات الإحصائية التي تم حساب الدلالة العملية لها برسائل الماجستير بالجامعات العمانية ككل

الاختبار الإحصائي	العدد	اقتصرت على الدلالة الإحصائية فقط		استخدمت الدلالة العملية	
		العدد	النسبة	العدد	النسبة
اختبار ت لعينة واحدة	291	251	%86.3	40	%13.7
اختبار ت لعينتين مستقلتين	4393	4028	%91.7	365	%8.3
اختبار ت لعينتين مرتبطتين	499	414	%83.0	85	%17.0
إجمالي اختبارات ت	5183	4693	%90.5	490	%9.5
تحليل التباين الأحادي	3631	3513	96.8%	118	3.2%
تحليل التباين الثنائي والثلاثي	535	388	%72.5	147	%27.5
تحليل التباين المتعدد	93	73	%78.5	20	%21.5
إجمالي اختبارات تحليل التباين	4259	3974	%93.3	285	%6.7
معامل ارتباط بيرسون	2212	2150	%97.2	62	%2.8
معامل الانحدار	272	245	%90.1	27	%9.9
إجمالي الاختبارات	11926	11062	%92.76	864	%7.24

الجدول(3): المتوسطات الحسابية والانحرافات المعيارية لقيم الدلالة العملية لكل الاختبارات الإحصائية، برسائل الماجستير بالجامعات العمانية ، ومستوياتها

الاختبار الإحصائي	العدد	الدلالة العملية		
		المتوسط الحسابي	الانحراف المعياري	المستوى
اختبار ت لعينة واحدة	291	1.662	8.050	كبيرة
اختبار ت لعينتين مستقلتين	4393	0.405	0.856	صغيرة

اختبار ت لعينتين مرتبطتين	499	3.437	6.891	كبيرة
إجمالي اختبارات ت	5183	0.768	3.107	متوسطة
تحليل التباين الأحادي	3631	0.191	0.288	صغيرة
تحليل التباين الثنائي والثلاثي	535	0.530	2.126	متوسطة
تحليل التباين المتعدد	93	0.944	1.474	كبيرة
إجمالي اختبارات تحليل التباين	4259	0.250	0.841	صغيرة
معامل ارتباط بيرسون	2212	0.148	0.172	متوسطة
معامل الانحدار	272	0.813	1.589	كبيرة