

الدلالة الاحصائية في الدراسات التربوية والنفسية: مراجعة منهجية للأدب التربوي في الوطن العربي
Statistical significance in educational and psychological studies: a systematic review of educational literature in the Arab world

د. إبراهيم بن سعيد بن حميد الوهبي

أستاذ القياس والتقويم المساعد بجامعة الشرقية - سلطنة عمان

Dr. Ibrahim Said Humaid Al-Wahaibi

assistant professor of measurement and evaluation at a'sharqiyah university- sultanate of oman

ملخص الدراسة:

هدفت الدراسة إلى التحقق من مدى اهتمام الباحثون العرب بدراسة الدلالة الإحصائية، من مختلف الجوانب النظرية والتطبيقية، واستخدمت المنهج البليومتري، حيث تكونت عينة الدراسة من جميع الدراسات العربية المتعلقة بالدلالة الإحصائية والمنشورة في دار المنظومة خلال ربع قرن (1996-2021)، وتوصلت النتائج إلى أن عدد الدراسات العربية التي بحثت في الدلالة الإحصائية بلغت (32) دراسة فقط، وتعتبر قليلة جداً، أي بمعدل أقل من دراستين في العام الواحد، مما يؤكد على افتقار المكتبة العربية بهذا النوع من الدراسات، كما أن أغلب هذه الدراسات بحثت في الجانب التطبيقي للدلالة الإحصائية، حيث وجد (8) دراسات فقط بحثت في الجوانب النظرية للدلالة الإحصائية من مثل استخداماتها الخاطئة وجوانب قصورها، وأوضح النتائج أيضاً أن الباحث لهذا الدراسة هو أكثر الباحثين العرب الذين كتبوا في هذا المجال حيث نشر حوالي خمس الدراسات في مجالات الدلالة الإحصائية وتطبيقاتها، وخلصت الدراسة إلى جملة من التوصيات أهمها دعوة الباحثين العرب في إعداد دراسات في مؤشرات الدلالات الإحصائية تتبع المنهج النوعي.

الكلمات المفتاحية: الدلالة الإحصائية؛ دار المنظومة، الوطن العربي

Abstract:

The study aimed to verify the extent of Arab researchers' interest in studying statistical significance, from various theoretical and applied aspects, and used the bibliometric approach, as the study sample consisted of all the Arab studies related to statistical significance published in the Dar al-Mizamah during (1996-2021), and the results were reached. Noting that the number of Arab studies that looked at statistical significance amounted to (32) studies only, and they are considered very few, that is, at a rate less than two studies per year, which confirms the lack of this type of studies and research in the Arab library, and most of these studies have been discussed in this aspect. Applied to the statistical significance, where he found (8) studies only that investigated the theoretical aspects of the statistical significance, such as its erroneous uses and its shortcomings, and the results also indicated that the researcher for this study is the most Arab researchers who wrote in this field, where he published about five studies in the fields of statistical significance and their applications. The study concluded with a set of recommendations, the most important of which is inviting Arab researchers to prepare studies on statistical significance indicators that follow the qualitative approach.

Key words: statistical significance; House system, the Arab world

مقدمة:

تمثل الدلالة الإحصائية محور نتائج الدراسات التربوية والنفسية، بل ومعظم الدراسات الإنسانية التي تستخدم منهجية البحث الكمي، فلا تكاد تخلو منها نتائج أي دراسة عربية كانت أو أجنبية، حيث شاع استخدام الاختبارات الدلالة الإحصائية (NHST) لاختبار البيانات الإحصائية منذ أكثر من نصف قرن (الوهبي، 2020)، ورافق هذا الاستخدام الوقوع في تفسيرات خاطئة لنتائج هذه الاختبارات الإحصائية، مما سعت كثير من الدراسات الأجنبية إلى تصحيح المسار لهذه الاستخدامات (Amrhein & Ioannidis, 2019; Greenland, 2019; Greenland & McShane, 2019; Hubbard & Lindsay, 2008; Hurlbert & Levine & Utts, 2019; Johnson, 2019; Kaufman, 1998; Kmetz, 2019; Shaver, 1992; Thompson, 1997; Nix & Barnette, 1998)، وحيث أن استخدامات اختبارات الدلالة الإحصائية منتشرة أيضاً في الدراسات العربية، فإن هذا الانتشار رافقه أيضاً قيام الباحثون العرب بتفسيرات خاطئة لنتائج هذه الاختبارات الإحصائية، وهذا ما

أكدته الدراسة الاستطلاعية التي قام بها الوهبي(2020)، الأمر الذي يتطلب البحث والتقصى لمدى اهتمام الباحثون العرب بدراسة استخدامات الدلالة الإحصائية، ومحاولة تطويرها.

مشكلة الدراسة وأسئلتها:

تتمثل مشكلة الدراسة في الاستخدام غير الدقيق للدلالة الإحصائية في الدراسات الإنسانية، مما أدى إلى ظهور انتقادات أثبتت حوله (Gigerenzer, 2004; Perezgonzalez, 2014, 2015)، ومصدر هذه الانتقادات أن اختبارات الدلالة الإحصائية (NHST) تعاني من بعض القصور في استخداماته، كما أنها فُسرت بغير الطريقة التي أنشأت من أجلها (الوهبي، 2020)، وعلى الرغم من حالات القصور التي تعترى الدلالة الإحصائية إلا أنها لا تزال تستخدم في أوساط الباحثين، ولم يتم التوقف عنها، أو إصلاح من جوانب قصورها، مما أدى إلى توصل مختلف الدراسات والأبحاث إلى مجموعة من القرارات غير الدقيقة والتي لا يمكن الاعتماد عليها، (Alwahaibi, et al., 2020) الأمر الذي دعا بعض الجمعيات الإحصائية مثل الجمعية الأمريكية للإحصاء (ASA)، إلى إصدار بيان عام 2016 تحذر فيه من التفسير الخاطئ لاستخدامات الدلالة الإحصائية، واشتمل البيان على ستة مبادئ (Amrhein & Greenland & McShane, 2019; Gagnier & Morgenstern, 2017; Thomas & Barach & Wilkinson & Farooqi & Wasserstein & Lazar, 2016; Lipshultz, 2017)، ولم يتوقف الأمر عند هذا الحد في الأدب الأجنبي المهتم بهذا الموضوع، بل وأقيمت ندوات وأصدرت أعداد خاصة من مجلات إحصائية عالمية مرموقة، كما هو الحال في مجلة الإحصاء الأمريكية American Statistician حيث أصدرت عدد خاص 2019 نُشر فيه أكثر من (40) بحثاً، حول "الاستدلال الإحصائي في القرن الحادي والعشرين: عالم يتجاوز $P > 0.05$ "، لمناقشة الجوانب المختلفة للدلالة الإحصائية واختباراتها المستخدمة، وإيجابياتها وسلبياتها (Wasserstein & Schirm & Lazar, 2019)، ونظراً لأهمية الدلالة الإحصائية في المجال التربوي، فإنه يتطلب البحث المتواصل عن الآليات والإجراءات التربوية لكيفية استخدامها، ومواكبة الاهتمام العالمي بهذا المصطلح التربوي، لذا فقد سعت هذه الدراسة إلى مراجعة منهجية للأدبيات المتعلقة بالدلالة الإحصائية ونوعية المناهج العلمية المستخدمة في هذه الدراسات، وكيفية تناولها بالوطن العربي على أساس تحليل وصفي خلال ربع قرن من الآن (1996-

(2021)، والكشف عن مستوى اهتمام الباحثين العرب في المكتبة العربية بهذا الموضوع من خلال البحث في إحدى قواعد البيانات والمعلومات الالكترونية التي تختص بالأبحاث والدراسات العلمية، وهي دار المنظومة، وبالتحديد فإن أسئلة الدراسة تتمثل في:

كم عدد الدراسات التي بحثت الدلالة الإحصائية الموجودة في قواعد بيانات دار المنظومة في الفترة من 1996م إلى 2021م ؟

ما نسبة الدراسات العربية التي بحثت في الأدب النظري للدلالة الإحصائية؟

من هم الباحثون الذين لديهم أكبر عدد من الدراسات في الدلالة الإحصائية في الوطن العربي؟

أهداف الدراسة:

تهدف هذه الدراسة إلى التحقق من مدى اهتمام الباحثون العرب بدراسة الدلالة الإحصائية، من مختلف الجوانب النظرية والتطبيقية.

أهمية الدراسة:

نظرا لكون قيمة البحث العلمي في مختلف المجالات تكمن في دقة نتائج الدراسات العلمية التي يتم التوصل إليها، وحيث أن هذه النتائج كثيرا ما يتم استخراجها والحصول عليها باستخدام الاختبارات الإحصائية، فتطور البحث العلمي مرهون بالاستخدام الصحيح لهذه الاختبارات الإحصائية، والتطوير المستمر فيها، والأدب الأجنبي ثري بدراسة الاستخدامات المختلفة لأكثر الاختبارات الإحصائية المستخدمة في تحليل البيانات وهي اختبارات الدلالة الإحصائية، لذا فإن أهمية هذه الدراسة تبرز في التعرف على ثرات الأدب العربي في تطوير البحث العلمي، من حيث البحث في الوسائل المعينة للحصول على نتائج دقيقة ومحددة، كما تنبع أهمية هذه الدراسة من كونها تعطي مؤشر جيد لمدى مواكبة الأدب العربي لمستحدثات الدلالة الإحصائية.

حدود الدراسة:

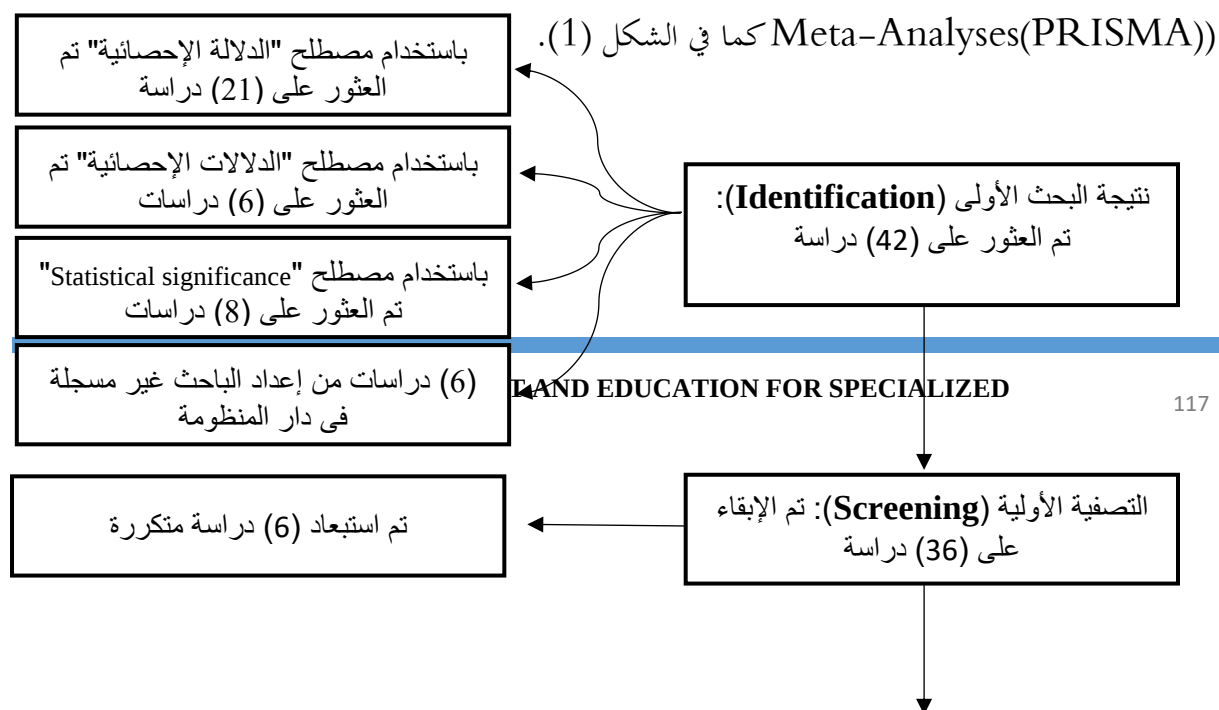
اقتصرت الدراسة على الدراسات المنشورة في قواعد بيانات "دار المنظومة" في ربع قرن (1996-2021) فقط.

مصطلحات الدراسة:

الدلالة الإحصائية: يعرفها الوهبي (2020) بأنها نتيجة القيمة الاحتمالية (p-value) المستخدمة للحكم على الاختبارات الإحصائية المستخدمة في الدراسات، وتصنف إلى مستويين: (دالة) إذا كانت قيمة (p-value) أقل من (0.05)، و(غير دالة) إذا كانت أكبر من (0.05).

منهجية الدراسة:

استخدمت هذه الدراسة المنهج المسحي الببليومتري (Bibliometric Methodology) للدراسات العربية المنشورة في قواعد بيانات دار المنظومة والتي قامت بتحليل الدلالة الإحصائية في متغيراتها الرئيسة، وقد تم اعتماد دار المنظومة كأساس للبحث عن دراسات الدلالة الإحصائية العربية المنشورة كونها تمثل أحد قواعد البيانات العربية المتخصصة في نشر الدراسات والأبحاث الأكاديمية، وارتباطها مع عدد كبير من المجالات والجمعيات العلمية المحكمة في الدول العربية، كما أنها متخصصة في نشر رسائل الماجستير واطروحات الدكتوراه العربية في مختلف العلوم والتخصصات الأكاديمية، والإنسانية، كما تتضمن دار المنظومة قاعدة بيانات الكشاف، وقاعدة المعلومات التربوية (EduSearch)، وقاعدة معلومات الاقتصاد والإدارة: (EcoLink)، وقاعدة معلومات العلوم الإسلامية والقانونية (Islamic info)، وقاعدة معلومات اللغة والأدب (AraBase)، وقاعدة معلومات العلوم الإنسانية (Humanindex)، وكلها قواعد بيانات تشمل جميع الدوريات والمجلات العلمية العربية المختلفة، وأبحاث ودراسات المؤتمرات والندوات في مختلف المجالات العلمية، التربوية والاجتماعية، والاقتصادية، والدينية، والإنسانية وغيرها (الوهبي وآخرون، 2021)، ولقد تم الاعتماد على مكونات التقارير المستخدمة في المراجعات المنهجية والتحليلات الوصفية Reviews and Preferred Reporting Items for Systematic



مجتمع الدراسة وعينتها:

شملت الدراسة جميع الدراسات العربية التربوية التي والتي بحثت الدلالة الإحصائية في متغيراتها الرئيسة، المنشورة خلال ربع قرن من الآن، في الفترة من (1996-2021) والبالغ عددها (32) دراسة فقط.

نتائج الدراسة:

السؤال الأول: كم عدد الدراسات التي بحثت الدلالة الإحصائية الموجودة في قواعد بيانات دار المنظومة في الفترة من 1996م إلى 2021م ؟

يوضح الجدول (1) أعداد الدراسات والنسب المئوية حسب الفترات الزمنية بالسنوات

الجدول(1): توزيع الدراسات حسب الفترات الزمنية (1996-2021)

الفترة الزمنية	عدد البحوث	النسبة المئوية
2000-1996	4	12.5%
2005 – 2001	8	25%
2010 – 2006	4	12.5%
2015 – 2011	2	6%
2021 – 2016	14	44%
المجموع	32	100%

يتبين من الجدول(1) أن الدراسات التي بحثت في الدلالة الإحصائية في الوطن العربي خلال ربع قرن في الفترة (1996-2021) تعتبر قليلة جدا جدا، حيث بلغت (32) دراسة فقط، أي بمعدل أقل من دراستين في العام الواحد، مما يؤكد على افتقار المكتبة العربية بهذا النوع من الدراسات والأبحاث، ولا توجد مقارنة أبدا بما أنتجته المجلات الأجنبية في هذا المجال، حيث أن ما قام به الباحثون العرب خلال ربع قرن يعتبر أقل من عدد الدراسات التي صدرت في عدد واحد فقط من إحدى المجلات الإحصائية بالولايات المتحدة الأمريكية، وهي مجلة الإحصاء

الإمريكية American Statistician حيث أصدرت عدد خاص 2019 نُشر فيه أكثر من (40) بحثاً، حول "الاستدلال الإحصائي في القرن الحادي والعشرين: عالم يتجاوز $P>0.05$ "، لمناقشة الجوانب المختلفة للدلالة الإحصائية واختباراتها المستخدمة، وإيجابياتها وسلبياتها، كما تشير النتائج إلى أن الخمس سنوات الماضية (2016-2021) تعد أكثر فترة تم نشر فيها الدراسات التي بحثت في الدلالة الإحصائية، حيث تم نشر عدد (14) دراسة، بمعدل (44%) من إجمالي الدراسات المنشورة، وهذه النتيجة تعطي مؤشراً يبشر ببداية الوعي لدى الباحثين العرب بأهمية البحث في هذا المجال، ويأمل الباحث أن تكثر الدراسات التي تبحث في الدلالة الإحصائية، والاختبارات المكمل لها، خلال قادم السنوات.

السؤال الثاني: ما نسبة الدراسات العربية التي بحثت في الأدب النظري للدلالة الإحصائية؟

أظهرت نتائج الدراسة إلى أن أغلب الدراسات العربية التي بحثت في الدلالة الإحصائية هي تلك التي تقصت درجة استخدامها في الدراسات والبحوث التي أجيّزت في الجامعات أو نُشرت في المجلات المحكمة، أما الدراسات التي بحثت في الأدب النظري للدلالة الإحصائية من الاستخدامات الخاطئة أو جوانب قصورها، أو طرائق تطويرها، فهي قليلة جداً، بلغت (8) دراسات فقط بما نسبته (20%) فقط، مما يشير إلى الحاجة الملحة لمزيد من هكذا دراسات، لما لهذا الموضوع من أهمية كبيرة في تبصير الباحثين العرب به، وإطلاعهم على المستجدات الحديثة التي يتم مناقشتها عالمياً في هذا الموضوع، والمساهمة العربية في تطوير الدلالة الإحصائية واختباراتها المختلفة.

السؤال الثالث: من هم الباحثون الذين لديهم أكبر عدد من الدراسات في الدلالة الإحصائية في الوطن العربي؟ بينت النتائج أيضاً أن الباحث لهذه الدراسة هو أكثر الباحثين العرب الذين قاموا بنشر دراسات في مجال الدلالة الإحصائية، حيث نشر (5) بحوث بالإضافة إلى أطروحته للدكتوراه، مما يمثل حوالي خمس الدراسات -19%- التي بحثت في الدلالة الإحصائية في الوطن العربي.

توصيات الدراسة، ومقترحاتها:

توصلت الدراسة إلى مجموعة من التوصيات والمقترحات من أهمها:

تضمنين مقررات الإحصاء التربوي بالجامعات العربية لجوانب قصور الدلالة الإحصائية واستخداماتها الخاطئة في تفسير النتائج.

عدم الاكتفاء باختبارات الدلالة الإحصائية في تحليل نتائج الدراسات، وإنما تعزز بالبدائل المكملة والداعمة كمؤشرات الدلالة العملية أو حجم الأثر.

عدم التحيز للنتائج الدالة إحصائية، وتقبل النتائج الأخرى غير الدالة إحصائية.

ضرورة قيام الباحثون العرب في إعداد دراسات في مؤشرات الدلالات الإحصائية تتبع المنهج النوعي.

المراجع:

AlWahaibi, I.S.; AlHadabi, D.A.; & AlKharusi, H.A. (2020). The degree of agreement between the statistical significance and the practical indications in the statistical tests for master's theses in Omani universities. *Journal of Human Development and Education for Specialized Research*, 6 (1), 100-130.

DOI/URL: <http://jhdesr.siats.co.uk/060104>

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

Gagnier, J., & Morgenstern, H. (2017). Misconceptions, Misuses, and Misinterpretations of P Values and Significance Testing . *The Journal of Bone and Joint Surgery*, 99 (18), 1598 -1603, <http://dx.doi.org/10.2106/JBJS.16.01314>.

Gigerenzer, G. (2004). Mindless statistics. *J. Soc. Econ*, 33, 587-606. doi:10.1016/j.socec.2004.09.033.

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>.

Hubbard, R. , & Lindsay, R. M. (2008). Why P Values Are Not a Useful Measure of Evidence in Statistical Significance Testing. *THEORY & PSYCHOLOGY*, 18 (1), 69–88, DOI: 10.1177/0959354307086923, <http://tap.sagepub.com>

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>.

[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>.

Kaufman, A. S. (1998). Introduction to the Special Issue on Statistical Significance Testing. *Research in the Schools*, 5 (2), P1, Fall 1998.

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>.

Moher D., Liberati A., Tetzlaff J., Altman G. & Prisma Group. (2019). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS medicine* 6, 7, e1000097.

Nix, T. W., & Barnette, J. J. (1998). The Data Analysis Dilemma: Ban or Abandon. A Review of Null Hypothesis Significance Testing. *Research in the Schools*, 5 (2), 3-14.

Perezgonzalez, J. D. (2014). A reconceptualization of significance testing. *Theor. Psychol.* 24 (6), 852- 859. doi:10.1177/0959354314546157.

Perezgonzalez, J. D. (2015). Fisher, Neyman-Pearson or NHST? A tutorial for teaching data testing. *Frontiers in Psychology*, 6, 1-11.

Shaver, James P. (1992). What Statistical Significance Testing Is, and What It Is Not. Paper Presented at Annual Meeting of the American Educational Research Association (San Francisco, CA, April 20-24, 1992).

Thomas R.L., Barach P.R., Wilkinson J.D., Farooqi A.A., & Lipshultz S.E. (2017). The danger of relying on the interpretation of p-values in single studies: Irreproducibility of results from clinical studies. *Progress in Pediatric Cardiology*, 44, 57-61.

Thompson, B. (1997). Rejoinder: Editorial Policies Regarding Statistical Significance Tests: Future Comments. *Educational Research*, 26, (5), 29- 32, Jun-Jul 1997.

Wasserstein, R.L., & Lazar, N.A. (2016). The ASA's statement on p-values: context, process, and purpose. *Am Stat.* 70 (2), 129-33.

Wasserstein, R.L., Schirm, A. L., & Lazar, N. A. (2019). Moving to a World Beyond "p < 0.05". *The American Statistician*, 73 (1), 1-19.