

بحوث العلوم النفسية: الاستدلال الإحصائي والأزمة في قابلية التكرار

**Psychological Science Research: Statistical Inference and the
Reproducibility Crisis**

أحمد محمد حسين عشر 2.1

ahmed.unshur@igdore.org



<https://orcid.org/0000-0001-9616-9553>

¹ قسم علم النفس، كلية الآداب، جامعة إفريقيا العالمية، الخرطوم، السودان

² Institute for Globally Distributed Open Research and Education (IGDORE)

المستخلص:

تتراكم المعرفة نتيجة لتتابع البحوث وتطورها، ويعتبر البحث العلمي أداة ووسيلة موضوعية للكشف عن الحقيقة العلمية، والعلم يقوم بالتصحيح الذاتي ليصحح مساره. تواجه بحوث العلوم النفسية قضايا منهجية مرتبطة بالاستدلال الإحصائي وبالتحديد سوء فهم واستخدام الدلالة الإحصائية أو القيم الاحتمالية (p-values)، وعدم قابلية النتائج للتكرار. يقوم الباحثون بعملية تجريف البيانات (Data-dredging) لإيجاد نتائج ذات دلالة إحصائية تبرر النشر وذلك نتيجة للتنافس العالي في البيئة الأكاديمية. تقدم هذه الورقة جهود العلماء والباحثين في علم النفس والإحصاء تجاه هذه القضايا، بما فيها بيان أصدرته الجمعية الإحصائية الأمريكية (ASA) American Statistical Association عن الدلالة الإحصائية والقيم الاحتمالية، والإطار العلمي المفتوح Open Science Framework (OSF)، ومشروع تعاوني قام به التعاون العلمي المفتوح بحيث تم تكرار 100 دراسة تجريبية وارتباطية للحصول على تقدير مبدئي في قابلية نتائج البحوث النفسية للتكرار. كما تقدم الورقة عدداً من الحلول المقترحة. يؤمل أن تحفز هذه الورقة النقاش في وسط الباحثين وأن تفتح آفاقاً جديدة للبحوث النفسية.

الكلمات المفتاحية: البحث العلمي، الاستدلال الإحصائي، قابلية التكرار، العلم المفتوح، علم النفس.

Abstract:

Knowledge accumulates due to the succession and development of research, and scientific research is considered an objective tool and means to reveal scientific truth. Science performs self-correction to correct its course. Psychological science research faces methodological issues related to statistical inference, namely misunderstanding and misuse of statistical significance, (i.e., p-values), and irreproducibility of findings. Data-dredging became a process carried out by researchers to find statistically significant results that justify publication, as a result of high competition in the academic environment. This paper presents the efforts of scientists toward these issues, including a statement on statistical significance and p-values issued by the American Statistical Association (ASA), the Open Science Framework (OSF), and the Open Science Collaboration which replicated 100 experimental and correlational studies to obtain an initial estimate of the reproducibility of psychological research findings. The paper also presents a number of suggested solutions. Hopefully, this paper will stimulate discussion among researchers and contribute to opening a new outlook for psychological research.

Keywords: Scientific Research, Statistical Inference, Reproducibility, Open Science, Psychology

مقدمة:

يمثل البحث عن المعرفة إحدى الجوانب التي اهتم بها الإنسان عبر العصور المختلفة؛ وذلك للإجابة عن تساؤلاته حول نفسه والعالم المحيط به، فالإنسان منذ أن خلق محباً للاستطلاع، كما أنه لا يستطيع أن يحيا وسط الظواهر والأشياء دون أن يكون لنفسه عنها بعض الأفكار التي تساعد على تحديد سلوكه تجاهها، وتساعد في تحسين أساليب حياته (البيسوني، 2013). يعتبر العلم عملية جمع وتقييم الأدلة الإمبريقية بصورة منهجية، للإجابة عن أسئلة أو اختبار أفكار معينة، والبحث العلمي أداة ووسيلة موضوعية للكشف عن الحقيقة العلمية، وهو طريق مقبول لتثبيت وترسيخ الحقيقة في المجالات الإنسانية، حيث يتم عرضها ونقدها بموضوعية، كما أنه الطريق الميسر لتوسيع الاتفاق العقلي بين الناس، وجعل أحكامنا أكثر قبولاً ودقة لدى الآخرين (شحاتة، 2008). تتراكم المعرفة نتيجة لتتابع البحوث وتطورها، والتقدم الذي وصل إليه الإنسان في الوقت الحاضر حصل نتيجة لتراكم المعرفة في جميع مجالات العلم (أبوعلام، 2007). المنهج العلمي أسلوب لاكتساب المعرفة، باستخدام أساليب منهجية وموضوعية لفهم الظواهر، وينطوي على المنطق وفحص البيانات (Crano et al., 2015). والباحث أثناء بحثه لحقائق المعرفة والعلوم، يهدف إلى إحداث إضافات أو تعديلات في ميادين العلوم، الأمر الذي يؤدي إلى تطويرها وتقدمها.

يشير علي (2005) إلى أن أسس التفكير والبحث العلمي تطورت ببطء واستغرقت عدة قرون في التاريخ الإنساني، والتفكير العلمي - كأسلوب منظم في معالجة المشاكل - له أصوله ومناهجه وشروطه الخاصة به، وهو أسلوب يمكن أن يُعتمد عليه في المعالجة الموضوعية للمشاكل التي تواجه الإنسان، ويحتاج إلى قدر كبير من الالتزام بأسس منظمة للتفكير. جوهر العلوم بما فيها العلوم النفسية هو النقد الذاتي (Self-criticism) القاسي والصارم، وتخضع العلوم النفسية نتائجها على التدقيق البحثي بحيث يتم الاحتفاظ بالنتائج

التي تنجو من الفحص الدقيق بصورة مؤقتة، بينما يتم التعديل أو التخلي عن غيرها. وبهذه الصورة تقوم العلوم النفسية مثل بقية العلوم الأخرى بالتصحيح الذاتي والتقدم التدريجي (Lilienfeld & Waldman, 2017). يؤكد المنسي (2011) أن الحقيقة العلمية ليست لها قيمة ما لم يتم التأكد من صدقها وثباتها بطريقة البحث. وقد ظهر مجال يعرف بـ"ما وراء المعرفة" (Metascience)، وهو نهج يحول فيه العلم عدسة التدقيق عليه. ما وراء العلم (علم العلم) يستخدم أساليب صارمة لدراسة أثر الممارسات العلمية على صحة النتائج العلمية (Schooler, 2014).

مشكلة الدراسة:

يعتمد البحث العلمي على تطبيق المنهج العلمي مما يؤدي إلى أبحاث راسخة واكتشافات ذات أهمية عالية. وتنتج البحوث النفسية نتائج حول السلوك والعمليات العقلية، ولكن هناك قلق متزايد في أن البحوث المنشورة غير جديرة بالثقة (Ioannidis, 2005)؛ وذلك بسبب قضايا منهجية متعلقة بالاستدلال الإحصائي (سوء فهم واستخدام الدلالة الإحصائية أو القيم الاحتمالية p-values) وعدم القابلية للتكرار. وتتحدد مشكلة الدراسة في التساولين الآتيين:

1. ما هي الجهود المبذولة للتصدي عن الممارسات الإحصائية غير الجيدة في البحوث النفسية؟
2. ما هي الجهود المبذولة للتغلب على عدم قابلية التكرار في البحوث النفسية؟

أهمية الدراسة:

تبرز أهمية الدراسة من خلال تركيزها على بعض القضايا المنهجية التي تدور في مجال علم النفس، والتي ينبغي التعرف عليها من قبل الباحثين وبالتحديد الجدد منهم. وتقدم الدراسة جهود العلماء والباحثين تجاه هذه القضايا، ويمكن أن تفتح باباً للنقاش حول طبيعة إجراء البحوث النفسية.

أهداف الدراسة:

تتمثل أهداف الدراسة في الآتي:

1. إبراز جهود الباحثين للتصدي للممارسات الإحصائية غير الجيدة في البحوث النفسية.

2. الكشف عن جهود الباحثين للتغلب على عدم قابلية التكرار في البحوث النفسية.

حدود الدراسة:

أجريت الدراسة في الفترة من 2018 – 2019.

المنهج:

قام الباحث بالبحث في APA PsycINFO و Google Scholar وذلك باستخدام الكلمات المفتاحية الآتية (Reproducibility, Statistical Inference, Psychology). تم جمع 6 دراسات ومن ثم تم اختيار دراستين للمراجعة، الأولى ركزت على الاستدلال الإحصائي (Wasserstein & Lazar, 2016) والثانية على قابلية التكرار (Open Science Collaboration, 2015).

الاستدلال الإحصائي:

الإحصاء هو دراسة جمع ووصف البيانات واستخلاص الاستنتاجات منها (Thijssen, 2014). والاستدلال الإحصائي هو أن يتم استخدام البيانات للإجابة عن أسئلة جوهرية موضوعية (Welsh, 1996). في العقود الأخيرة حدثت زيادة في تكميم البحث العلمي وانتشرت البيانات الكبيرة (Big Data) المعقدة مما أدى إلى توسع نطاق تطبيقات الأساليب الإحصائية (Wasserstein & Lazar, 2016). وأصبح اختبار دلالة الفرضية الصفرية (NHST) Null Hypothesis Significance Testing الأسلوب الأكثر استخداماً

على نطاق واسع في تحليل البيانات التي يتم جمعها في البحوث النفسية (Nickerson, 2000). وقد يتصور المرء أن أسلوباً تمّ احتضانه من قبل مجتمع بحثي بأكمله أن يكون مفهوماً بشكل جيد وغير مثير للجدل بعد عقود طويلة من الاستخدام المستمر. ومع ذلك، اختبار دلالة الفرضية الصفرية مثير للجدل للغاية، وقد صدرت نداءات للتخلي أو منع استخدامه (Baguley, 2012).

أشار العديد من الباحثين في مجال علم النفس وفي مجالات أخرى مثل العلوم الطبية والإحصاء عن المشاكل المرتبطة إلى الاستدلال الإحصائي والنتائج التي تتوصل إليها البحوث العلمية. قام إيونيدس (Ioannidis, 2005) بنشر مقال علمي يقترح فيه أن معظم نتائج البحوث المنشورة زائفة، كما نشرت نوزو (Nuzzo, 2014) مقالاً حول الأخطاء الإحصائية في المنهج العلمي. عليه قام الإحصائيون بالبحث عن طرق أفضل للتفكير بشأن تقييم البيانات لمساعدة الباحثين على تجنب فقدان معلومات مهمة، أو الاهتمام بالدلالات الكاذبة.

تقوم العديد من النتائج العلمية المنشورة بمفهوم "الدلالة الإحصائية" "Statistical Significance"، وعادة ما يتم تقييمها بالقيمة الاحتمالية (P-value). فبينما يمكن أن تكون القيمة الاحتمالية مقياس إحصائي مفيد، عادة ما يساء استخدامها وتفسيرها (Wasserstein & Lazar, 2016). يرى نيكرسون

(Nickerson, 2000) أن العديد من الصعوبات المفاهيمية لدى الباحثين حول اختبار دلالة الفرضية الصفرية ترجع إلى عدم التمييز بين الاحتمالات المطلقة والشرطية (Absolute and Conditional Probabilities)، وعلى وجه الخصوص الفشل في فهم أن القيمة الاحتمالية التي تنتجها اختبارات الدلالة الإحصائية هي احتمال شرطي؛ أي احتمال الحصول على النتيجة الإحصائية التي تم الحصول عليها على أساس افتراض

في أن الفرضية الصفرية صحيحة. ترى نوزو (2014) Nuzzo أن القيم الاحتمالية غير موثوقة أو موضوعية كما يفترض معظم العلماء، وتشجع التفكير المشوش، ومثال ذلك هو ميلها إلى صرف الانتباه عن الحجم الحقيقي للأثر.

يمثل النشر (Publication) "عملة" العلوم الأكاديمية (Academic Science) ويزيد احتمال التوظيف والتمويل والترقية (Nosek et al., 2012)، والبيئية الأكاديمية اليوم أكثر تنافساً من أي وقت مضى، وهناك تركيز عال على تكديس منشورات ذات نتائج دالة إحصائياً (Nuzzo, 2015) بحيث يستخدم الباحثون عملية تجريف البيانات (Data-dredging) لإيجاد نتائج ذات دلالة إحصائية تبرر النشر (Begley, 2015).

الممارسات الإحصائية غير الجيدة المرتبطة بسوء فهم واستخدام الدلالة الإحصائية أدت إلى أن تأخذ بعض الجهات مواقف بشأنها. وقامت الجمعية الإحصائية الأمريكية American Statistical Association (ASA) بإصدار بيان عن الدلالة الإحصائية والقيم الاحتمالية نشرته في مجلة The American Statistician (TAS)، حيث قدمت فيه ستة مبادئ بمصطلحات غير فنية والتي من شأنها أن تحسن إجراء وتفسير العلم الكمي (Quantitative Science) وفقاً لإجماع واسع النطاق من قبل الإحصائيين (Wasserstein & Lazar, 2016)، وتتمثل في الآتي:

المبدأ الأول: يمكن للقيم الاحتمالية أن تشير كيف تتعارض البيانات مع نموذج إحصائي محدد.

توفر القيمة الاحتمالية نهج لتلخيص عدم التوافق بين مجموعة معينة من البيانات ونموذج مقترح لتلك للبيانات. فالسياق الأكثر شيوعاً هو نموذج بُني تحت مجموعة من الافتراضات، جنباً إلى جنب مع ما يسمى "بالفرضية الصفرية". تفترض الفرضية الصفرية عدم وجود أثر، مثل عدم وجود اختلاف بين مجموعتين، أو

عدم وجود علاقة بين العامل والنتيجة. فكلما صغرت القيمة الاحتمالية كلما زاد عدم التوافق الإحصائي للبيانات مع الفرضية الصفرية، إذا كانت الافتراضات الأساسية المستخدمة لحساب القيمة الاحتمالية داعمة. ويمكن تفسير هذا التناقض بأنه يلقي الشك أو يقدم أدلة ضد الفرضية الصفرية أو الافتراضات التي تقوم عليها.

المبدأ الثاني: القيم الاحتمالية لا تقيس احتمال صحة الفرضية تحت الدراسة، أو احتمال أن البيانات تم إنتاجها عن طريق الصدفة العشوائية وحدها.

كثيراً ما يرغب الباحثون في تحويل القيمة الاحتمالية إلى بيان عن صحة الفرضية الصفرية، أو حول احتمال أن الصدفة العشوائية أنتجت البيانات المرصودة. القيمة الاحتمالية ليست كذلك، القيمة الاحتمالية بيان حول البيانات بالنسبة لتفسير افتراضي، وليست بيان حول التفسير نفسه.

المبدأ الثالث: الاستنتاجات العلمية والقرارات التجارية أو السياسية لا ينبغي أن تستند فقط على ما إذا تجاوزت القيمة الاحتمالية على عتبة معينة.

الممارسات التي تحد من تحليل البيانات أو الاستدلال الإحصائي لقواعد مثل ($p < 0.05$) لتبرير الاستنتاجات العلمية يمكن أن تؤدي إلى معتقدات خاطئة وضعف في اتخاذ القرارات. لا تصبح النتيجة فوراً "صحيحة" على جانب واحد و "خاطئة" من جهة أخرى. يجب على الباحثين أن يجلبوا العديد من العوامل السياقية لاستخلاص الاستنتاجات العلمية، بما في ذلك تصميم الدراسة، وجودة القياسات، والأدلة الخارجية للظاهرة تحت الدراسة، وصحة الافتراضات التي تقوم عليها تحليل البيانات. الاعتبارات البراغماتية غالباً ما تتطلب قرارات ثنائية "نعم - لا"، وهذا لا يعني أن القيم الاحتمالية يمكن أن تضمن أن القرار صحيح أو خاطئ. يؤدي الاستخدام الشائع للدلالة الإحصائية (والذي يفسر بصفة عامة $p \leq 0.05$) كرخصة للقيام بادعاء استنتاج علمي (أو الحقيقة الضمنية) إلى تشويه كبير في العملية العلمية.

المبدأ الرابع: الاستدلال الصحيح يتطلب الشفافية وتقديم تقارير كاملة.

لا يجب تقديم تقارير عن القيم الاحتمالية والتحليلات ذات الصلة بها بصورة انتقائية. فإجراء تحليلات متعددة على البيانات وتقديم التقرير عن تحليلات ذات قيم احتمالية معينة فقط (عادة تلك التي تجاوزت عتبة الدلالة Significance Threshold) يجعل القيم الاحتمالية المذكورة في التقرير غير قابلة للتفسير. اختيار النتائج العاهدة، والمعروف أيضا بمصطلحات مثل تجريف البيانات، مطاردة الدلالة، تنقيب الدلالة، الاستدلال الانتقائي، وقرصنة القيمة الاحتمالية (p-hacking) يؤدي إلى زيادة زائفة على النتائج ذات الدلالة الإحصائية في الأدبيات المنشورة، وينبغي التجنب عن ذلك.

لا يحتاج المرء إلى أن يقوم باختبارات إحصائية متعددة لتظهر هذه المشكلة. كلما يختار الباحث ما سيتم عرضه بناء على النتائج الإحصائية، يصبح التفسير الصحيح لتلك النتائج تحت خطر شديد إذا لم يتم إعلام القارئ عن الاختيار وأساسه. يجب على الباحثين كشف عدد الفرضيات التي تم فحصها أثناء الدراسة، وجميع قرارات جمع البيانات، وجميع التحليلات الإحصائية التي أجريت، وجميع القيم الاحتمالية التي تم حسابها. الاستنتاجات العلمية الصحيحة القائمة على القيم الاحتمالية والإحصاءات ذات الصلة بها لا يمكن استخلاصها على الأقل دون معرفة عدد ونوع التحليلات التي أجريت، وكيف تم اختيار تلك التحليلات (بما فيها القيم الاحتمالية) في تقديم التقرير.

المبدأ الخامس: القيم الاحتمالية أو الدلالة الإحصائية لا تقيس حجم الأثر أو أهمية النتيجة.

الدلالة الإحصائية لا تعادل الأهمية العلمية، البشرية أو الاقتصادية. القيم الاحتمالية الصغيرة لا تعني بالضرورة وجود تأثيرات ذات أهمية كبيرة، والقيم الاحتمالية الكبيرة لا تعني عدم وجود الأهمية أو عدم وجود الأثر. أي أثر، مهما صغر، يمكن أن ينتج قيمة احتمالية صغيرة إذا كان حجم العينة أو دقة القياس مرتفعاً بما

فيه الكفاية، والتأثيرات الكبيرة يمكن أن تنتج قيماً احتمالية غير مثيرة للإعجاب إذا كان حجم العينة صغيراً أو إذا كانت القياسات غير دقيقة. وبالمثل، فإن الآثار التقديرية المتطابقة سيكون لديها قيم احتمالية مختلفة إذا اختلفت دقة التقديرات.

المبدأ السادس: القيمة الاحتمالية وحدها لا توفر مقياساً جيداً للأدلة المتعلقة بالنموذج أو الفرضية.

ينبغي أن يدرك الباحثون أن قيمة احتمالية دون سياق أو أدلة أخرى توفر معلومات محددة. على سبيل المثال، أخذ قيمة احتمالية قريبة بـ (0.05) وحدها فقط يقدم أدلة ضعيفة ضد الفرضية الصفرية. وبالمثل، قيمة احتمالية كبيرة نسبياً لا توحى أدلة لصالح الفرضية الصفرية؛ قد تكون العديد من الفرضيات الأخرى بنفس القدر أو أكثر اتساقاً مع البيانات المرصودة. لهذه الأسباب، تحليل البيانات لا ينبغي أن ينتهي بحساب القيمة الاحتمالية عندما تكون هناك أساليب أخرى مناسبة وممكنة.

أكد فاسرشتاين و لازار (2016) Wasserstein & Lazar أن الممارسة الإحصائية الجيدة، باعتبارها عنصراً أساسياً في الممارسة العلمية الجيدة، تركز على مبادئ تصميم وإجراء الدراسة الجيدة، ومجموعة متنوعة من الملخصات العددية والبيانية للبيانات، وفهم الظاهرة تحت الدراسة، وتفسير النتائج في سياق، وتقديم تقارير كاملة، وفهم منطقي وكمي صحيح لما تعنيه ملخصات البيانات، بحيث لا يجب أن يكون مؤشراً واحداً بديل للمنطق العلمي.

الأزمة في قابلية التكرار:

يمثل التكرار إحدى أهم المبادئ الرئيسة للمنهج العلمي (Replication) (Leek & Peng, 2015). فمن الصعب التأكد مما إذا كانت نتيجة ما ذات دلالة إحصائية حقيقية وصادقة، وتكرار البحوث هو المعيار الأكثر أهمية لتقييم متانة نتائجها، ولاسيما عندما يتم القيام بذلك من قبل باحثين مستقلين (Lindsay, 2015).

وعلى مدى السنوات الأخيرة، أثير شك في علم النفس باعتباره مجالاً يتميز بأبحاث قابلة للتكرار (Yong, 2012)، وأشار بعض العلماء إلى حالة الضجة الراهنة بـ"أزمة في قابلية التكرار" "Reproducibility Crisis" (Bartlett, 2014). وتشير الأزمة في قابلية النتائج للتكرار إلى أزمة منهجية في العلوم حيث وجد العلماء أن نتائج العديد من الدراسات العلمية يصعب أو يستحيل تكرارها على تحقيق لاحق، إما عن طريق باحثين مستقلين أو من قبل الباحثين الأصليين أنفسهم (Schooler, 2014).

يقول روسل (Russell 2013): "يجب على العلم - كنظام - أن يولي أهمية أكبر لقابلية التكرار؛ لأنها الفاصل بين العلم والحكايات". يرى أسيندريف وآخرون (Asendorpf et al., 2013) أن مصطلح "أزمة" "Crisis" قد يكون وصفاً مبالغاً فيه، حيث إن حجم المشكلة في قابلية التكرار في مجال علم النفس غير معروفة إلى حد كبير. ولكن في نفس الوقت، أصبح من الواضح بشكل متزايد بأن علماء النفس ليس بمقدورهم الافتراض في قابلية نتائجهم للتكرار. ويعتقد بينغ (Peng 2015) أن سوء فهم واستخدام الاستدلال الإحصائي له دور كبير في أزمة قابلية التكرار. وتؤكد بيكر (Baker 2016) أن عملية غربة الاكتشافات لاستبعاد النتائج الزائفة تكون محرجة، كما تشير إلى وجود ندرة في الاتصال بين الباحثين؛ وإذا تواصل الباحثون الذين يكررون الدراسة مع الباحثين الأصليين، طلباً للمساعدة، فإنهم يخاطرون بإظهار أنفسهم بمظهر المتشكك، أو غير الكفاء. ونادراً ما يتم إجراء دراسات التكرار في العلوم النفسية، كما يتم استخفافها من قبل المحررين والمراجعين على حدّ سواء (Makel et al., 2012).

التطور العلمي يتطلب الالتزام بالشفافية (Transparency) (Lewandowsky & Bishop, 2016)، وظهرت حركة جديدة في البحث العلمي تعرف بالعلم المفتوح (Open Science). حيث يشير مصطلح العلم المفتوح إلى عملية جعل محتوى وطريقة إنتاج الأدلة والنتائج العلمية شفافة وسهلة الوصول للآخرين (Munafò

(et al., 2017). ويقوم الباحثون والعلماء في مجال علم النفس بمبادرات لتطوير البحث العلمي وإيجاد حلول للمشاكل التي تواجه العلم. فمثلاً، في عام 2013، قام بريان نوسيك وجيفري سبيس (Brian Nosek & Jeffrey Spies) بتأسيس مركز العلوم المفتوحة (COS) Center for Open Science في شارلوتسفيل، الولايات المتحدة، وهو مؤسسة علمية تكنولوجية غير ربحية تهدف إلى زيادة انفتاح، ونزاهة البحوث العلمية، وقابلية نتائجها للتكرار. وقام المركز بإنشاء الإطار العلمي المفتوح Open Science Framework (OSF) ، وهي خدمة حرة ومفتوحة المصدر تسهل التعاون المفتوح في البحوث العلمية.

قام التعاون العلمي المفتوح⁽¹⁾ (2015) Open Science Collaboration من خلال استخدام الإطار العلمي المفتوح بجهد تعاوني واسع النطاق للحصول على تقدير مبدئي في قابلية نتائج العلوم النفسية للتكرار، بحيث تم تكرار 100 دراسة تجريبية وارتباطية نُشرت في ثلاث مجلات علمية محكمة⁽²⁾ في عام 2008 . واستشار التعاون العلمي المفتوح مع الباحثين الأصليين الذين قاموا بإجراء الدراسات، كما استخدم نفس أدوات الدراسات الأصلية مع القيام بمراجعات داخلية؛ لكي تحتفظ الدراسات المكررة بدقة الدراسات الأصلية. واستطاع التعاون العلمي المفتوح (2015) تكرار 39% فقط من النتائج الأصلية، وأكد أن الفشل في التكرار لا يعني بالضرورة أن النتائج الأصلية كانت خاطئة أو مزيفة. ومع ذلك، فإن هذه النتائج تشير إلى أن علماء النفس ومستهلكي البحوث النفسية، بما فيها وسائل الإعلام، عليهم أن يضعوا وزناً أقل بكثير مما يضعونه حالياً على النتائج غير المكررة (Waldman & Lilienfeld, 2015). وقدّم مشروع التعاون العلمي المفتوح أدلة تراكمية

(1) مشروع تكون من فرق عمل تكونت من 270 باحث.

(2) Psychological Science (PSCI), Journal of Personality and Social Psychology (JPSP), and Journal of Experimental Psychology: Learning, Memory, and Cognition (JEP:LMC).

من العديد من النتائج في البحوث النفسية، واقترح أنه لا يزال هناك المزيد من العمل الذي يجب أن نقوم به للتحقق من ما إذا نعرف ما نعتقد أننا نعرفه (Open Science Collaboration, 2015).

الإطار العلمي المفتوح (OSF) Open Science Framework:

تم إنشاء الإطار العلمي المفتوح (OSF)⁽³⁾ من قبل مركز العلوم المفتوحة. حيث يوفر الإطار للباحثين دعماً مجانياً ومفتوح المصدر في إدارة المشاريع عبر دورة حياة البحث بأكملها، كما يسهل عملية التعاون العلمي. يعمل الإطار كنظام سير العمل (Workflow System) بحيث يمكن للباحثين ربط الخدمات والتطبيقات العديدة التي يستخدمونها بالإطار لزيادة الكفاءة وتبسيط عملية إجراء البحوث، كما يعمل كأداة تعاون (Collaboration Tool) بحيث يمكن للباحثين العمل على المشاريع البحثية بصورة خاصة مع عدد محدود من المتعاونين وجعل أجزاء منها عامة، أو جعل المشروع بأكمله متاحاً للجمهور لنشره على نطاق أوسع. ويعمل الإطار أيضاً كمستودع مرن (Flexible Repository) بحيث يمكن للباحثين تخزين وأرشفة بياناتهم وبروتوكولاتهم وأدواتهم البحثية فيه.

من يمكن أن يستخدم الإطار العلمي المفتوح؟

1. يمكن للباحثين استخدام الإطار العلمي المفتوح لإدارة مشاريعهم أو تسجيل دراساتهم.
2. يمكن أن تستخدم المؤتمرات الأكاديمية الإطار العلمي المفتوح للاجتماعات لتبادل الملصقات والشرائح المقدمة في المؤتمر مع جمهور أوسع لزيادة تأثير العمل المقدم.

⁽³⁾ Open Science Framework (OSF): <https://osf.io/>

3. يمكن أن تستخدم الدوريات العلمية الإطار العلمي المفتوح لدعم مبادرات تبادل البيانات والأدوات، التقارير المسجلة (Registered Reports) والتسجيل المسبق (Preregistration) للبحوث.
4. يمكن أن تستخدم المؤسسات الإطار العلمي المفتوح لتوفير الباحثين مشاعات (Commons) علمية حرة ومفتوحة المصدر.
5. يمكن للممولين استخدام الإطار العلمي المفتوح لدعم مبادرات تبادل البيانات والأدوات، وذلك لتوسيع أثر العمل الذي يمولونه.

أهم النتائج:

توصلت الدراسة إلى أن العلماء يبذلون جهوداً لإيجاد حلول للقضايا التي تواجه بحوث العلوم النفسية والمرتبطة بسوء فهم واستخدام القيم الاحتمالية والاستدلال الإحصائي، وضعف أو عدم قابلية النتائج للتكرار.

1. تم اقتراح عدد من الأساليب الإحصائية التي تركز على التقدير Estimation أكثر من الاختبار (Cumming, 2014; Wasserstein & Lazar, 2016; Nuzzo, 2014; Open Testing Science Collaboration, 2015) وتشمل: فترات الثقة Confidence Intervals، فترات التنبؤ Prediction Intervals، حجم الأثر Effect Size، والتحليل التلوي Meta-analysis.
2. الأساليب البايزية Bayesian Methods هي أيضاً من الأساليب المقترحة استخدامها في البحوث النفسية (Wagenmakers et al., 2017).
3. استخدام طرق بديلة مثل منهجية النظرية المجردة Grounded Theory Methodology (GTM) وهو أسلوب منهجي استقرائي مقارنة لإجراء البحوث؛ وذلك لغرض بناء النظريات (Charmez & Henwood, 2007).

4. رفع مصداقية النتائج المنشورة من خلال إجراء دراسات التكرار Replication Studies واستخدام

التقارير المسجلة والتسجيل المسبق (Open Science Collaboration, 2015; Munafò et

al., 2017).

5. تدريب الطلاب والباحثين بمناهج البحث العلمي والتحليل الإحصائي، وحث الباحثين على التعاون،

والاستفادة من الخدمات الحديثة مثل الإطار العلمي المفتوح (Open Science Collaboration,

2015; Munafò et al., 2017).

الخاتمة:

تشير الجهود والمبادرات التي يقوم بها الباحثون في علم النفس حول القضايا المنهجية في المجال إلى وجود الرغبة في تحسين وتقوية البحوث النفسية. وتمثل القدرة على ملاحظة أنماط جديدة وغير متوقعة في البيانات إحدى السمات التي تميز الإبداع العلمي. ومع ذلك، فإن التحدي الكبير الذي يواجه العلماء هو الانفتاح للأفكار الجديدة والمهمة، مع تجنب في نفس الوقت الانسياق وراء الميول في ملاحظة التركيب المنظم في العشوائية (Munafò et al., 2017). وصون المبادئ العلمية ليس شيئاً ينبغي القيام به مرة واحدة وإلى الأبد؛ بل هو تحدّي يجب مواجهته على أساس يومي من قبل الباحثين والمؤسسات العلمية (Ioannidis, 2012). وظل العلم وسيظل يواجه بدايات خاطئة كجزء من عملية الاكتساب الجمعي للمعرفة الجديدة. والممارسات الإحصائية الجيدة، والتماشي مع التطورات الحديثة من خلال التعلم المنهجي المستمر، والاهتمام بتكرار نتائج البحوث العلمية وإيجاد طرق لتحسينها من شأنها أن تساعد المجتمع العلمي وتوجه بحوث العلوم النفسية نحو الطريق الصحيح.

شكر وتقدير:

يتقدم الباحث بالشكر والتقدير للدكتور عادل إسحق والدكتور محمد صلاح عوض على ما قدما من تعليقات واقتراحات مفيدة ساعدت بشكل كبير في إثراء الورقة.

تضارب المصالح:

يقر الباحث بعدم وجود تضارب في المصالح.

المراجع:

أبوعلام، رجاى محمود. (2007). *مناهج البحث في العلوم النفسية والتربوية*. ط 6. القاهرة: دار النشر للجامعات.

البيسوني، محمد سويلم. (2013). *أساسيات البحث العلمي في العلوم التربوية والاجتماعية والإنسانية*. القاهرة: دار الفكر العربي.

شحاتة، حسن. (2008). *المرجع في مناهج البحوث التربوية والنفسية*. القاهرة: مكتبة الدار العربية للكتاب.

علي، السيد محمد السيد. (2005). *مناهج البحث في العلوم الطبيعية وعلاقتها بالحضارة الإسلامية*. ط 1. الأهرام: الدار العالمية للنشر والتوزيع.

المنسي، محمود عبدالحليم. (2011). *مناهج البحث العلمي في المجالات التربوية والنفسية*. الإسكندرية: دار المعرفة الجامعية.

Asendorpf, J. B., Conner, M., De Fruyt, F., De Houwer, J., Denissen, J. J. A.,

Fiedler, K., Fiedler, S., Funder, D. C., Kliegl, R., Nosek, B. A., Perugini, M.,

Roberts, B. W., Schmitt, M., Vanaken, M. A. G., Weber, H., & Wicherts, J. M.

(2013). Recommendations for increasing reproducibility in psychology. *European Journal of Personality*, 27, 108 – 119.

Baguley, T. (2012). *Serious stats: A guide to advanced statistics for the behavioral sciences*. New York, NY: Palgrave Macmillan.

Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. *NATURE*, 533, 452 – 454.

Bartlett, T. (2014). Replication crisis in psychology research turns ugly and odd. *Chronicle of Higher Education*. Retrieved from:

<http://chronicle.com/article/Replication-Crisis-in/147301>

Begley, C. G., Buchan, A. M., & Dirnagl, U. (2015). Robust research: Institutions must do their part for reproducibility. *NATURE*, 525, 25 – 27.

Charmaz, K., & Henwood, K. (2007). Grounded theory in psychology. In C. Willig & W. Stainton-Rogers. (Eds.), *Handbook of qualitative research in psychology*. London: SAGE Publishers.

Crano, W. D., Brewer, M. B., & Lac, A. (2015). *Principles and methods of social research*. 3rd Ed. New York & Hove: Routledge.

Cumming, G. (2014). The new statistics: What and how. *Psychological Science*, 25(1), 7 – 29.

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS*

Med 2(8): e124. doi:10.1371/journal.pmed.0020124

Ioannidis, J. P. A. (2012). Why science is not necessarily self-correcting.

Perspectives on Psychological Science, 7(6), 645 – 654.

Leek, J. T., & Peng, R. D. (2015). Reproducible research can still be wrong:

Adopting a prevention approach. *PNAS, 112(6), 1645 – 1646.*

Lewandowsky, S., & Bishop, D. (2016). Don't let transparency damage science,

NATURE, 529, 459 – 461.

Lilienfeld, S. O. & Waldman, I. D. (2017). Psychological science in perspective. In

S. O. Lilienfeld & I. D. Waldman. (Eds.), *Psychological science under scrutiny:*

Recent challenges and proposed solutions. Hoboken, NJ: John Wiley & Sons,

Inc.

Lindsay, D. S. (2015). Replication in psychological science. *Psychological Science,*

26, 1027 – 1032.

Makel, M., Plucker, J., & Hegarty, B. (2012). Replications in psychology research:

How often do they occur?. *Perspectives on Psychological Science, 7, 537 –*

542.

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D.,

du Sert, N. P., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J., & Ioannidis,

J. P. A. (2017). A manifesto for reproducible science. *NATURE HUMAN*

BEHAVIOR, 1, 0021, DOI:10.1038/s41562-016-0021

Nickerson, R. S. (2000). Null hypothesis significance testing: A review of an old

and continuing controversy. *Psychological Methods*, 5(2), 241 – 301.

Nosek, B. A., Spies, J. R., & Motyl, M. (2012). Scientific utopia: II. Restructuring

incentives and practices to promote truth over publishability. *Perspectives on*

Psychological Science, 7, 615 – 631.

Nuzzo, R. (2014). Statistical errors. *NATURE*, 506, 150 – 152.

Nuzzo, R. (2015). How scientists fool themselves – and how they can stop.

NATURE, 526, 182 – 185.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological

science. *SCIENCE*, 349(6251), aac4716.

Peng, R. D. (2015). The reproducibility crisis in science. A statistical counterattack.

SIGNIFICANCE, 12(3), 30 –32.

Russell, J. F. (2013). If a job is worth doing, it is worth doing twice. *NATURE*, 496,

7.

Schooler, J. W. (2014). Metascience could rescue the 'replication crisis'. *NATURE*,

515, 9.

Thijssen, J. (2014). *A concise introduction to statistical inference*. Boca Raton, FL:

CRC Press, Taylor & Francis Group, LLC.

- Wagenmakers, E.-J., Verhagen, J., Ly, A., Matzke, D., Steingroever, H., Rouder, J. N., & Morey, R. D. (2017). The need for bayesian hypothesis testing in psychological science. In S. O. Lilienfeld & I. D. Waldman. (Eds.), *Psychological science under scrutiny: Recent challenges and proposed solutions*. Hoboken, NJ: John Wiley & Sons, Inc.
- Waldman, I. D., & Lilienfeld, S. O. (2015). Thinking about data, research methods, and statistical analyses: Commentary on Sijtsma's (2014) "Playing with data." *Psychometrika*, 81, 16 – 26.
- Wasserstein, R. L., & Lazar, N. A. (2016). The ASA's statement on p-values: context, process, and purpose. *THE AMERICAN STATISTICIAN*, 70(2), 129–133.
- Welsh, A. H. (1996). *Aspects of statistical inference*. Hoboken, NJ: John Wiley & Sons, Inc.
- Yong, E. (2012). Replication studies: bad copy. *NATURE*, 485, 298 – 300.