مؤشرات الدلالة العملية في حالة اختبار تحليل التباين

محمد حسين سعيد حسين(•)

الملخص:

إن غياب استخدام وعرض نتائج حجم التأثير أو الدلالة العملية، بجانب عدم الالمام بالقضايا المرتبطة بها، يعد سببا رئيسيا في إعاقة انتشار واستخدام الممارسات الإحصائية المرتبطة بها في البحوث النفسية والتربوية في العالم العربي، على الرغم من أنه وعلى نحو متزايد، يُطلب من الباحثين في العديد من المجلات المتخصصة حساب مؤشرات حجم التأثير وعرضها ومناقشتها كجزء روتيني من أي تقرير للبحوث النفسية والتربوية. كما يعد تحليل التباين من الأساليب الإحصائية التي تستخدم في المقارنة بين أكثر من متوسطين مرة واحدة لتقليل الأخطاء التي يمكن أن تحدث في حالة استخدام اختبار "ت" للمقارنة بين كل متوسطين في كل مرة. وفي تحليل التباين لا تعطي قيمة "ف" أي معلومات حول الدلالة العملية. لذا تناقش هذه الورقة مؤشرات حجم التأثير في حالة اختبار تحليل التباين وذلك من خلال: توضيح العلاقة بين دلالة قيمة "ف" وحجم العينة، ومفهوم وأهمية حجم التأثير، وأهم مؤشرات حجم التأثير المستخدمة في حالة تحليل التباين (مربع ايتا ومربع ايتا الجزئي ومربع اوميجا ومربع ابسيلون)، ثم أداء تلك المؤشرات. وانتهت المقالة إلى أنه يجب استخدام قيم حجم التأثير المصححة مثل مربع ابسيلون ومربع اوميجا جنبًا إلى جنب مع قيم الدلالة الإحصائية P في حالة تحليل التباين بدلاً من أحجام التأثير غير المصححة مثل مربع ايتا ومربع ايتا الجزئي، كما يجب تضمين الباحثين في بحوثهم قيم "ف" ودرجات الحرية للتأثير ودرجات الحرية للخطأ، أو قيم مؤشرات حجم التأثير غير المتحيزة حتى يمكن تسهيل تحليل القوة للباحثين الآخرين، كما يجب الحذر عند استخدام مؤشرات حجم التأثير حتى لا يساء تفسيرها.

الكلمات المفتاحية: الدلالة العملية، حجم التأثير، تحليل التباين، مربع ايتا، مربع اوميجا، مربع ابيسلون.

(•) أستاذ بكلية التربية جامعة بني سويف للمراسلات في شأن هذا البحث ترسل mohamedhussein40@yahoo.com

Indicators of Practical Significance in the Case of the Analysis of Variance Test

Mohamed Hussein Said Hussein(•)

Abstract:

The absence of the use and presentation of the results of effect size or practical significance, along with the lack of knowledge of the issues associated with them, is a major reason for impeding the spread and use of statistical practices associated with them in psychological and educational research in the Arab world, although it is increasingly required of researchers in many It is professional journals to calculate, present, and discuss effect size indices as a routine part of any psychological and educational research report. Analysis of variance is also one of the statistical methods that are used to compare more than two means at once to reduce errors that can occur in the case of using "t" test to compare two means each time. In ANOVA the value of "f" does not give any information about practical significance. Therefore, this paper discusses the effect size indicators in the case of the analysis of variance test, through: clarifying the relationship between the significance of the "f" value and the sample size, the concept and importance of effect size, and the most important effect size indicators used in the case of analysis of variance (eta squared, partial eta squared, omega squared, and epsilon squared), then perform those indicators. The article concluded that corrected effect size values such as epsilon squared and omega squared should be used along with the values of statistical significance P in the case of analysis of variance instead of uncorrected effect sizes such as eta squared and partial eta squared, and researchers should include in their research the values of "f" and degrees of freedom for effect and degrees of freedom for error, or values of unbiased effect size indicators so that power analysis can be facilitated for other researchers, and care must be taken when using effect size indicators so as not to be misinterpreted.

Keywords: Practical significance, effect size, analysis of variance, eta squared, omega squared, epsilon squared.

(•)Professor of Educational Psychology, Faculty of Education, Beni-Suef University mohamedhussein40@yahoo.com

مقدمة:

إن غياب تقديم التوصيات باستخدام وعرض نتائج الدلالة العملية أو حجم التأثير، بجانب عدم الإلمام بالقضايا المرتبطة بها، يعد سببا رئيسا في عدم استخدام الممارسات الإحصائية المرتبطة بها في البحوث النفسية والتربوية في العالم العربي. على الرغم من أنه وعلى نحو متزايد، بدأ يُطلب من الباحثين حساب إحصاءات حجم التأثير وتقديمها ومناقشتها كجزء روتيني من أي تقرير للبحوث النفسية والتربوية (Bakeman, 2005)، حيث يشير دليل جمعية علم النفس الأميريكية[1] إلى أنه من الضروري دائمًا تقرير أحجام التأثير لتقدير حجم أو أهمية النتيجة؛ حيث يعد إعداد التقارير الكاملة لتقديرات الأحجام المناسبة للتأثير أحد الحد الأدنى من التوقعات لجميع مجلات (American Psychological Association, 2009)، وفي الآونة الأخيرة، طلب عدد كبير من المجلات العلمية الإبلاغ عن مقاييس حجم التأثير جنبًا إلى جنب مع قيمة مستوى الدلالة[2] عند عرض نتائج التحليل الإحصائي (Skidmore, & Thompson, 2012).

كما يعد تحليل التباين[3] (اختبار ف) من الأساليب الإحصائية التي تستخدم في المقارنة بين أكثر من متوسطين مرة واحدة لتقليل الأخطاء التي يمكن أن تحدث في حالة استخدام اختبار "ت" للمقارنة بين كل متوسطين في كل مرة، وبالتالي يزيد احتمال اتخاذ قرار صحيح عند اختبار الفروض الإحصائية المرتبطة بالمقارنة بين تلك المتوسطات باستخدام تحليل التباين. ومع ذلك، فإنه لا يوضح مدى اختلاف متوسط المجموعات المقارنة عن بعضها البعض أو مدى الاختلاف الذي حدث في المتغير التابع نتيجة تأثير المتغير المستقل. بعبارة أخرى، أثناء اختبار الدلالة الإحصائية للفروق بين

1- American Psychological Association (APA)
2- P Value
3- Analysis of Variance (ANOVA-F)

مستويات المتغير المستقل، لا يعطي اختبار "ف" أي معلومات حول أهميته العملية (Yigit & Mendes, 2018).

إن الدلالة الإحصائية تعبر عن احتمال وجود علاقة بين المتغيرات، وهي تختلف عن الدلالة العملية؛ فقد تكون بعض النتائج ذات دلالة إحصائية ولكن الآثار المترتبة على هذه النتائج تكون غير ذات أهمية أو دلالة عملية. ويقصد بالدلالة العملية الأساليب التي يتم من خلالها معرفة حجم الفرق أو حجم العلاقة بين متغيرين أو أكثر ويسمى حجم التأثير، وكل من الدلالة الإحصائية والدلالة العملية يكمل كل منهما الآخر ويؤدي استخدامهما معًا إلى إثراء نتائج البحوث النفسية والتربوية (منصور، ١٩٩٧). وفي مقالة سابقة (حسين، ٢٠٢٢) تم تناول مؤشر "كوهين" وطرق استخدامه وتفسيره في حالات اختبار "ت" المختلفة للمقارنة بين المتوسطات. وفي هذه المقالة يتم مناقشة مؤشرات حجم التأثير في حالة اختبار تحليل التباين، وذلك من خلال: توضيح العلاقة بين دلالة قيمة "ف" وحجم العينة، ومفهوم وأهمية حساب مؤشرات حجم التأثير، وأهم مؤشرات حجم التأثير المستخدمة في تحليل التباين مثل: مربع ايتا ومربع ايتا الجزئي ومربع اوميجا ومربع ابسيلون.

العلاقة بين دلالة قيمة "ف" وحجم العينة:

قبل توضيح المقصود بحجم التأثير في تحليل التباين، رأيت أنه من الأفضل توضيح العلاقة بين دلالة قيمة "ف" وعلاقتها بحجم العينة، حتى يدرك الباحثون أن قيمة p الصغيرة والتي تُستخدم كمعيار للدلالة الإحصائية لا تعني بالضرورة أن قوة العلاقة بين المتغيرات المستقلة والتابعة في دراسة معينة كبيرة. فقد قدم (Maxwell & Delaney, 2004) صيغة رياضية للعلاقة بين قيمة أي اختبار إحصائي وحجم التأثير وحجم العينة على النحو التالي: قيمة الاختبار الإحصائي = حجم التأثير X حجم العينة. ومن هذه المعادلة يتبين أنه مع تثبيت قيمة حجم التأثير وزيادة حجم العينة فإن قيم الاختبار الاحصائي تزيد على الرغم من أن حجم التأثير ثابتا، أي أن التغير في قيمة الاختبار الإحصائي تكون نتيجة للتغير في حجم العينة. فلو افترضنا أن حجم تأثير

متغير في متغير آخر هو ٠,٥ فإن قيمة الاختبار الاحصائي لحجم عينة ١٠ هو ٥، وعندما يكون حجم العينة ١٠٠ ومع نفس حجم التأثير وهو ٠,٥ تكون قيمة الاختبار الاحصائي ٥٠، وهكذا تزيد قيمة الاختبار الاحصائي مع زيادة حجم العينة على الرغم من أن تأثير المتغير في الآخر ثابتة ولم تتغير. لذا يعد حجم العينة العامل الأساسي في تقرير ما إذا كانت نتائج أي دراسة دالة أو غير دالة إحصائيا (Snyder & Lawson, 1993). فحتى الفروق الصغيرة جدًا يمكن العثور عليها لتكون ذات دلالة إحصائية مع عينات ذات حجم كبير جدًا، وقد لا يتم العثور على دلالة إحصائية مع عينات صغيرة الحجم على الرغم من وجود أحجام تأثير كبيرة. ويبين جدول (١) التالي العلاقة بين دلالة قيمة "ف" وحجم العينة مع تثبيت حجم التأثير عند (٣٣ %).

جدول (١) العلاقة بين دلالة قيمة "ف" وحجم العينة مع تثبيت حجم التأثير (٣٣ %)

مصدر التباين	مجموع المربعات SS	درجات الحرية df	متوسط المربعات SM	قيمة F	قيمة الدلالة p value
ن=١٦					
بين المجموعات	٣٣٣,٣٣	٢	١٦٦,٦٦٥٠٠	٣,٢٤٩٩٥	٠,٠٧١٧
داخل المجموعات	٦٦٦,٦٧	١٣	٥١,٢٨٢٣١		
المجموع	١٠٠٠	١٥	٦٦,٦٦٦٦٧		
ن=١٧					
بين المجموعات	٣٣٣,٣٣	٢	١٦٦,٦٦٥٠٠	٣,٤٩٩٩٥	٠,٠٥٨٥
داخل المجموعات	٦٦٦,٦٧	١٤	٤٧,٦١٩٢٩		
المجموع	١٠٠٠	١٦	٦٢,٥٠٠٠٠		
ن=١٨					
بين المجموعات	٣٣٣,٣٣	٢	١٦٦,٦٦٥٠٠	٣,٧٤٩٩٤	٠,٠٤٧٨
داخل المجموعات	٦٦٦,٦٧	١٥	٤٤,٤٤٤٦٧		
المجموع	١٠٠٠	١٧	٥٨,٨٢٣٥٣		
ن=١٩					
بين المجموعات	٣٣٣,٣٣	٢	١٦٦,٦٦٥٠٠	٣,٩٩٩٩٤	٠,٠٣٩٠
داخل المجموعات	٦٦٦,٦٧	١٦	٤١,٦٦٦٨٨		
المجموع	١٠٠٠	١٨	٥٥,٥٥٥٥٦		

قام الباحث بتعديل محتوى جدول (١) السابق من (Snyder &

(Lawson, 1993) حيث عرضا في جدولهما قيم "ف" المحسوبة وقيم "ف" الجدولية وقيمة حجم التأثير r^2 لكل عينة في عمود بالجدول، مع إضافة عمود لبيان قبول أو رفض الفرض الصفري، وقام الباحث باستبدال هذا العمود بقيمة الدلالة الإحصائية p المناظرة لدرجات الحرية، كما تم الاكتفاء بكتابة قيمة حجم التأثير أعلى الجدول. وفي الجدول السابق يتضح أن قيمة الدلالة الإحصائية انخفضت من ٠,٠٧١٧ (وهي قيمة غير دالة إحصائيا) إلى ٠,٠٣٩٠ (وهي قيمة دالة إحصائيا) وذلك مع زيادة حجم العينة من ١٦ إلى ١٩ أي بزيادة ٣ أفراد فقط تغيرت قيمة دلالة "ف" من عدم الدلالة الاحصائية الى الدلالة الاحصائية بهذه الزيادة فقط، على الرغم من ثبات قيمة حجم التأثير عند ٣٣%.

مفهوم وأهمية مؤشرات حجم التأثير:

حجم التأثير أو الدلالة العملية هو مؤشر معياري ويقدر معلمة مستقلة عن حجم العينة، ويحدد مدى الفروق في المجتمع أو العلاقة بين المتغيرات التفسيرية والاستجابة أو المتغيرات المستقلة والمتغيرات التابعة، وهناك فئتان من حجم التأثير هما: الفروق المعيارية بين المتوسطات[1] ومقاييس الارتباط أو نسبة التباين المفسر (Richardson, 1996). والدلالة العملية هي نسبة ما يفسره المتغير المستقل من تباين للمتغير التابع (الصياد، ١٩٨٨م). وكلما كانت قيمة الدلالة العملية كبيرة دل ذلك على وجود الظاهرة بدرجة أكبر في المجتمع (Cohen, 1988).

يعتبر حجم التأثير من المخرجات الهامة لنتائج البحوث التربوية والنفسية. يريد الباحثون معرفة ما إذا كان للتدخل أو البرنامج التجريبي تأثير أكبر من الصفر ، أو (عندما يكون من الواضح وجود تأثير) ما مدى حجم هذا التأثير، وغالبًا ما يطلب من الباحثين كتابة تقارير عن أحجام التأثير لبحوثهم

1- Standardized Mean Differences

(Lakens, 2013). فقد تؤدي النتيجة الدالة إلى الاستدلال على وجود ارتباط ما، ولكن هذا لا يعني بأي حال من الأحوال وجود درجة مهمة من الترابط بالضرورة، على العكس من ذلك، يمكن أن يحدث دليل على وجود ارتباط إحصائي قوي في البيانات حتى عندما تكون النتائج غير دالة إحصائيا. فحجم التأثير يزود الباحثين بمعلومات تتجاوز حدود اختبارات الدلالة الإحصائية، ويمنحهم فهما وتفسيرا واضحا للنتائج التي توصلوا إليها.

ويذكر (Lakens, 2013) ثلاثة أسباب لاستخدام حجم التأثير في البحوث النفسية والتربوية وهذه الأسباب هي: أولاً، تسمح للباحثين بعرض حجم التأثير في مقياس موحد يمكن فهمه بغض النظر عن المقياس الذي تم استخدامه لقياس المتغير التابع، فأحجام التأثير المعيارية تسمح للباحثين بإيصال الأهمية العملية لنتائجهم (الفوائد العملية للنتائج في الحياة اليومية)، بدلاً من الإبلاغ فقط عن الأهمية الإحصائية. ثانيًا، تسمح أحجام التأثير للباحثين باستخلاص استنتاجات في دراسات وبحوث التحليل البعدي[1] من خلال مقارنة أحجام التأثير المعيارية عبر الدراسات. ثالثًا، يمكن استخدام أحجام التأثير من الدراسات السابقة عند التخطيط لدراسة جديدة وتحديد حجم عينتها، حيث يمكن أن يوفر تحليل القوة[2] للاختبار الاحصائي مؤشرا على متوسط حجم العينة الذي تحتاجه الدراسة لملاحظة او التوصل الى نتيجة ذات دلالة إحصائية مع الاحتمالية المرغوبة باستخدام برنامج G*Power. نظرًا لأن أحجام العينات الصغيرة جدًا تزيد من فرصة حدوث خطأ من النوع الثاني (عدم ملاحظة تأثير ذي دلالة إحصائية عند وجود تأثير حقيقي في المجتمع)، ولأن جمع عدد كبير جدًا من المشاركين غير فعال ومكلف، فإن أحد الأساليب لتحديد حجم عينة الدراسة هو تحليل القوة. ففي تحليل القوة، يحسب الباحثون حجم العينة اللازم للوصول إلى حجم تأثير معين، بمستوى دلالة محدد مسبقًا،

1- Meta-Analysis
2- Power Analysis

وقـوة إحصـائية مرغوبـة. إن الحـد الأدنـى المقبـول عمومًـا للقـوة هـو ٠,٨٠ (Cohen, 1988). يعتمد هذا الحد الأدنى على فكرة أنه مع وجود مستوى دلالة يبلـغ ٠,٠٥، يكـون الخطـأ مـن النـوع الثـاني (١ – القوة) هـو ٠,٢٠. فأخطـاء النوع الثاني يمكن أن يكون لها عواقب أكثر خطورة من أخطاء النوع الأول ، (Fiedler et al., 2012). وبالتالي ، على الرغم من أن قوة اختبار احصائي مساوية لـ ٠,٨٠ هي الحد الأدنى الموصى به ، فإن المستويات الأعلى (على سبيل المثال ، ٠,٩٥) تكون مرغوبة أكثر ، طالما أنها ممكنة عمليًا.

على الرغم مما سبق عرضه لتوضيح أهمية حجم التأثير أو الدلالة العملية إلا أنها لا تغني عن الدلالة الإحصائية، لأن مؤشرات الدلالة العملية لا تزود الباحث بأي إثبات أو برهان على احتمالية حدوث النتائج، لذلك من الضروري أن تكون الدلالة الإحصائية والدلالة العملية متلازمتين.

مؤشرات حجم التأثير في حالة تحليل التباين:

يمكن تجميع طرق حساب حجم التأثير في عائلتين هما: عائلة d وتتكون من الفروق المعيارية للمتوسط، وعائلة r كمقاييس لقوة الارتباط، وتستند أحجام تأثير عائلة d على الفرق بين الملاحظات ، مقسومًا على الانحراف المعياري لهذه الملاحظات، وتصف أحجام تأثير عائلة r نسبة التباين للمتغير التابع والتي يتم تفسيرها من خلال المتغير المستقل (Rosenthal, 1994). وفي تحليل التباين يتم تقسيم مجموع المربعات الكلي[1] إلى أجزاء مرتبطة بالتأثيرات المختلفة التي يحددها التصميم (Cohen & Cohen, 1983) فالتباينات في الدراسة تنشأ بسبب مصدرين من التباين: عوامل المعالجة[2] في الدراسة أو مجموع المربعات بين المجموعات[3] والفروق بين الأفراد[4] أو مجموع المربعات

1- Total Sum of squares (SSt)
2- Manipulated Factors
3- Sum of squares Between Groups (SSb)= SSEffect
4- Individual Differences

داخل المجموعات[1] التي ترجع إلى خصائص المشاركين والظروف التي لم يتم ضبطها (Olejnik & Algina, 2003). ويوجد مجموعة من مؤشرات حجم التأثير تستخدم في تحليل التباين، ومن تلك المؤشرات الأكثر شيوعًا واستخداما في تحليل التباين هي: مربع ايتا[2] η^2 ومربع ايتا الجزئي[3] η_p^2 ومربع اوميجا[4] ω^2 ومربع ابسيلون[5] ε^2، وفيما يلي تفصيل طريقة حساب كل مؤشر.

مؤشر مربع ايتا:

يرمز له بالرمز η^2 وقدم هذا المؤشر فيشر ويُطلق عليه مربع إيتا وهي نسبة التأثير إلى إجمالي التباين. ويتم حسابه من المعادلة: η^2 = مجموع المربعات بين المجموعات (التأثير) SS_{Effect} / مجموع المربعات الكلي SS_{Total} (Maxwell, et al., 1981). إن مربع ايتا يساوي ٠,١٩ يعني أن ١٩٪ من إجمالي التباين في المتغير التابع يمكن تفسيره من خلال المتغير المستقل.

مؤشر مربع ايتا الجزئي:

يرمز له بالرمز η_p^2 وقدم هذا المؤشر كوهين. على الرغم من أن مربع ايتا طريقة فعالة لمقارنة أحجام التأثير داخل الدراسة الواحدة، إلا أنه لا يمكن مقارنة مربع ايتا بسهولة بين الدراسات، لأن التباين الكلي في الدراسة (SS_{total}) يعتمد على تصميم الدراسة، ويزيد عندما يتم بحث تأثير العديد من المتغيرات (Lakens, 2013). كما أوصى(Tabachnick & Fidell, 2001) باستخدام مربع ايتا الجزئي حتى يمكن مقارنة أحجام التأثير بين الدراسات ، والتي تعبر عن مجموع مربعات التأثير ومجموع مربعات الخطأ المرتبطة بالتأثير. حيث يستخدم η_p^2 كحل لمشكلة المقارنة في سياق التصميمات العاملية ، لحذف تأثير

1- Sum of squares Within Groups (SSw)=(SSError)
2- Eta Squared
3- Partial Eta Squared
4- Omega Squared
5- Epsilon Squared

العوامل الأخرى من المقام (Bakeman, 2005)؛ حيث يساهم مجموع المربعات الناتجة عن كل عامل وعن التفاعلات متعددة العوامل في مجموع المربعات الكلي لمجموعة البيانات. وبالتالي ، قد لا تكون قيمة مربع ايتا لهذا العامل قابلة للمقارنة عبر الدراسات (Olejnik & Algina, 2003). ويتم حسابه من المعادلة: η_p^2 = مجموع المربعات للتأثير SS_{Effect} / (مجموع المربعات الكلي SS_{Total} + مجموع مربعات الخطأ SS_{Error}) (Maxwell, Camp, & Arvey, 1981). كما يمكن حساب مربع ايتا الجزئي من قيمة "ف" ودرجات حريتها من المعادلة: η_p^2 = قيمة "ف" × درجات الحرية للتأثير / ((قيمة "ف" × درجات الحرية للتأثير) + (درجات الحرية للخطأ)) (Lakens, 2013). فلو كانت نتائج تحليل التباين لاحدى الدراسات هي: ف (١، ٣٠) = ١٠ فإن قيمة مربع ايتا الجزئي = (١٠×١) / ((١٠×١)+(٣٠)) = ٠,٢٥. وفي تحليل التباين أحادي الاتجاه فإن قيمة مربع ايتا تكون مساوية لقيمة مربع ايتا الجزئية Levine et al., (2002). وتوفر الإصدارات الحالية من برنامج SPSS η_p^2 وليس η^2 عند طلب إحصائيات حجم التأثير.

مؤشر مربع ابسيلون:

قدم هذا المؤشر كيلي ويرمز له بالرمز ε^2 ويمكن حساب مربع ابسيلون من المعادلة ε^2 = ((مجموع المربعات للتأثير SS_{Effect}) − (درجات الحرية للتأثير df_{Effect} × متوسط مربعات الخطأ MS_{Error})) / (مجموع المربعات الكلي SS_{Total}) (Maxwell, Camp, & Arvey, 1981).

مؤشر مربع أوميجا:

في عائلة r لأحجام التأثير، يُعرف تصحيح مربع ايتا باسم مربع أوميغا. تم اقتراح مربع اوميجا لتصحيح هذا التحيز (Hayes, 1963). ويمكن حساب مربع اوميجا من المعادلة ω^2 = ((مجموع المربعات للتأثير SS_{Effect}) − (درجات الحرية للتأثير df_{Effect} × متوسط مربعات الخطأ MS_{Error})) / (مجموع

المربعات الكلي SS_{Total} + متوسط مربعات الخطأ MS_{Error}) (Maxwell, Camp, & Arvey, 1981).

أداء مؤشرات حجم التأثير في حالة تحليل التباين:

وجد (Keselman, 1975) أن القيم المتوسطة لمربع أوميغا تكون دائمًا أقرب إلى قيم المجتمع، بينما تكون القيم المتوسطة لمربع إبسيلون دائمًا أكبر قليلاً من مربع أوميغا ولكنها أصغر من القيم المتوسطة لمربع ايتا. وأشار (Grissom & Kim, 2004) إلى أن المؤشر البديل الأقل تحيزًا إلى حد ما لكل من η^2 وε^2 هو ε^2، والمؤشر غير المتحيز تقريبًا هو ω^2. وكذلك توصل (Matsumoto et al., 2011) إلى نتائج مشابه. ويرى Olejnik & Algina, (2000) أن مربع إبسيلون يصحح بسط مربع إيتا بطرح مربع متوسط الخطأ من مجموع تباين المربعات، وأن مربع أوميغا يضبط مربع إبسيلون عن طريق إضافة خطأ متوسط مجموع المربعات إلى المجموع الكلي للمربعات في مقام مربع إبسيلون.

أشار (Okada, 2013) إلى أنه وفي ضوء طريقة حساب كل مؤشر من مؤشرات حجم التأثير، فإن العلاقة بين المؤشرات الثلاثة يمكن توضيحها في المتباينة: $\omega^2 \leq \varepsilon^2 \leq \eta^2$ ، وللتحقق من ذلك وباستخدام دراسة مونت كارلو توصل إلى أنه خلافًا للاعتقاد الشائع بأن ω^2 هو أقل تحيزًا من ε^2 ، إلا أن ε^2 هو أفضل مؤشر لحجم التأثير من حيث عدم التحيز، وعلى الرغم من ذلك لم يحظ باهتمام مثل المؤشرين الآخرين، كما أن قيم η^2 تنحرف باستمرار عن القيمة الحقيقية للمجتمع حتى عندما يكون حجم العينة ١٠٠، لذا يجب الحذر عند تفسير η^2 في العلوم النفسية حيث يتم استخدام تحليل التباين غالبًا مع البيانات دون حجم عينة كبير. كما توصل (بخاري، ٢٠١٦) إلى أن قيم مربع ايتا وقيم مربع أوميجا في حالة الالتزام بشروط تحليل التباين تكون أكبر من قيمها في حالة التحرر من تلك الاشتراطات. وتوصل (Yigit & Mendes, 2018) إلى أن تقديرات مربع ابسيلون ومربع اوميجا كانت غير متحيزة تمامًا

عند مقارنتها بمربع ايتا ومربع ايتا الجزئي التي يتم حسابها مباشرة بواسطة الحزم الإحصائية الشائعة الاستخدام أثناء تحليل التباين.

يذكر (أبو حطب وصادق، ١٩٩٦) أنه في حالة استخدام مربع ايتا يكون التأثير صغيرا عندما يساوي ٠,٠١ ومتوسطا عندما يساوي ٠,٠٦ وكبيرا عندما يساوي ٠,١٤ فأكثر. ويرى (Snyder & Lawson, 1993) أنه يمكن استخدام نفس تلك المعايير لكوهين مع المؤشرين الآخرين نظرا لتقارب قيمهم وطريقة تفسيرهم وذلك عند استخدام عينات كبيرة يزيد حجمها عن ١٠٠.

التوصيات والتطبيقات التربوية:

يتضح مما سبق أنه يجب استخدام حجم التأثير جنبا الى جنب مع الدلالة الاحصائية للاختبارات الاحصائية بصورة عامة وتحليل التباين بصورة خاصة، كما يجب استخدام قيم حجم التأثير المصححة مثل: مربع ابسيلون ومربع اوميجا جنبًا إلى جنب مع قيم الدلالة الإحصائية P في حالة تحليل التباين بدلاً من أحجام التأثير غير المصححة مثل: مربع ايتا ومربع ايتا الجزئي، على الرغم من كونهما الأكثر شيوعًا واستخداما في البحوث، إلا أنه غير موصي باستخدامه، خاصةً عندما يكون حجم العينة صغيرًا. بالإضافة إلى أنه يجب عند استخدام أحجام التأثير تفسيرها من خلال مقارنتها بأحجام تأثير أخرى في الدراسات السابقة، كما يجب تضمين الباحثين في بحوثهم قيم "ف" ودرجات الحرية للتأثير ودرجات الحرية للخطأ أو قيم مؤشرات حجم التأثير غير المتحيزة حتى يمكن تسهيل تحليل القوة للباحثين الآخرين. وعلى الرغم من كل ما سبق إلا أنه يجب الحذر عند استخدام مؤشرات حجم التأثير حتى لا يساء تفسيرها.

المراجع:

أبو حطب، فؤاد، وصادق، آمال (١٩٩٦). **مناهج البحث وطرق التحليل الإحصائي في العلوم النفسية والتربوية والاجتماعية**. القاهرة:

الأنجلو المصرية.

بخاري، ماجد بن عبدالفتاح (٢٠١٦). أثر انتهاك افتراض تجانس التباين على قيم مربع ايتا ومربع أوميجا كمؤشرات لفحص الدلالة العملية في تحليل التباين الأحادي. **دراسات عربية في التربية وعلم النفس، ١٦**، ٢٠٥-١٤٣.

حسين، محمد حسين سعيد (٢٠٢٢). الدلالة العملية ضرورة حتمية في البحوث النفسية والتربوية مؤشر "كوهين" لحالات اختبار "ت". **مجلة دراسات نفسية، ٣٢(٢)**، ٢٤٦-٢٣٥. doi: 10.21608/PSJ.2022.137138.1047

الصياد، عبد العاطي أحمد (١٩٨٨) الدلالة العملية و وحجم العينة المصاحبتين للدلالة الإحصائية لاختبار "ت" في البحث التربوي والنفسي العربي **بحوث مؤتمر البحث التربوي بين الواقع والمستقبل**، القاهرة.

منصور، رشدي فام (١٩٩٧). حجم التأثير الوجه المكمل للدلالة الإحصائية. **المجلة المصرية للدراسات النفسية، ٧(١٦)**، ٧٥-٥٧.

American Psychological Association. (2009). **Publication manual of the American Psychological Association (6thed.)**. Washington DC: American Psychological Association.

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. **Behavior Research Methods, 37** (3), 379-384.

Cohen, J. (1988). **Statistical power analysis for the behavioral sciences**. Routledge Academic.

Cohen, J., & Cohen, P. (1983). **Applied multiple regression/correlation analysis for the behavioral sciences (2nd ed.).** Hillsdale, NJ: Erlbaum.

Fiedler, K., Kutzner, F., & Krueger, J. I. (2012). The long way from α-error control to validity proper problems with a short-sighted false-positive debate. **Perspectives on Psychological Science, 7**, 661-669.

Grissom, R. J., & Kim, J. J. (2004). **Eff ect sizes for research: A broad practical approach**. New York: Psychology Press.

Hayes, W. L. (1963). **Statistics for psychologists**. New York: Holt, Rinehart & Winston.

Keselman, H. J. (1975). A Monte Carlo investigation of three estimates of treatment magni-tude: Epsilon squared, eta squared, and omega squared. **Canadian Psychological Review, 16**, 44–48. doi: 10.1037/h0081789

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-test and ANOVAs. **Frontier in Psychology, 4**, Article 863. doi: 10.3389/fpsyg.2013.00863

Levine, T. R. And Hullett, C. R. (2002). Eta Squared, Partial Eta Squared, and Misreporting of Effect Size in Communication Research, **Human Communication Research, 28**, 4, 612{615.

Matsumoto, D., Kim, J. J., & Grissom, R. J. (2011). Eff ect sizes in cross-cultural research. In D. Matsumoto & F. J. R. Van de Vijver (Eds.), **Cross-cultural research methods in psychology**. New York: Cambridge University Press.

Maxwell, S. E., & Delaney, H. D. (2004). **Designing experiments and analyzing data: A model comparison perspective**. Routledge.

Maxwell, S. E., Camp, J. C., & Arvey, R. D. (1981). Measures of strength of association: A comparative examination. **Journal of Applied Psychology, 66,** 525–534. doi: 10.1037/0021-9010.66.5.525

Okada, K. (2013). Is Omega Squared Less Biased? a Comparison of Three Major Effect Size Indices in One-Way Anova. **Behaviormetrika 40**(2), 2013, 129–147.

Olejnik, S. & Algina, J. (2003). Eralized Eta and Omega Squared Statistics: Measures of Effect Size for Some Common Research Designs. **Psychological Methods, 8**(4), 434–447. doi: 10.1037/1082-989X.8.4.434.

Richardson, J. T. E. (1996). Measures of effect size. Behav-ioral Research Methods, **Instruments, & Computers, 28**, 12–22.

Rosenthal, R. (1994). Parametric measures of effect size. **The handbook of research synthesis,** 231-244.

Skidmore, S. T. and Thompson, B. (2012). Bias and precision of some classical ANOVA effect sizes when assumptions are violated, **Behavior Research Methods, 45**, 2, 536{546

Snyder, P., & Lawson, S. (1993). Evaluating results using corrected and uncorrected eff ect size estimates. **Journal of Experimental Education, 61**, 334–349

Tabachnick, B. G., & Fidell, L. S. (2001). **Using multivariate statistics (4th ed.).** Boston: Allyn and Bacon.

Yigit, S. & Mendes, M. (2018). Which Effect Size Measure is Appropriate for One-Way and Two-Way ANOVA Models? A Monte Carlo Simulation Study. **Statistical Journal, 16(**3), 295–313.