حجوم الأثر في الدراسات الأولية والتحليل البعدي
طرق حسابها وأساليب تفسيرها

إعداد

أ. د. عبد المنعم أحمد حسن

أستاذ المناهج وطرق التدريس

كلية التربية- جامعة الأزهر بالقاهرة

رؤى تربوية

هذا الباب يصدر عن مجلة التربية جامعة الأزهر ليضع بين يدي الباحثين رؤى تربوية سواء كانت رؤى تؤسس لإجراءات بحثية، أو رؤى نقدية تحليلية لما يقدمه الفكر التربوي من بحوث ودراسات وقضايا، تهدف المجلة من ذلك إلى تطوير الفكر والإجراءات البحثية لدى الباحثين.

ويأتي هذا الباب تحقيقا لرؤية المجلة وأهدافها، والتي من أهمها: المشاركة بفاعلية في الارتقاء بالعلوم التربوية في العالم العربي منطلقة إلى العالمية.ومعاونة الباحثين في تطوير معارفهم وإنتاجهم العلمي من خلال بحوث علمية رصينة في المجالات التربوية اتبعت إجراءات تحكيم متميزة.

وتأتي المقالة الأولى في هذا الباب لتتناول قضية مهمة من القضايا التي تتناول البحوث التجريبية وبحوث التحليل البعدي وهي حجوم الأثر طرق حسابها وكيفية تفسيرها، لتمكن الباحثين من تقديم نتائجهم بشكل إجرائي يوضح لمتخذي القرار والباحثين الآخرين آثار التجارب التربوية في المجالات المختلفة، ومقدار التأثير الناتج عنها.

كاتب هذا المقال استاذ وعلم من أعلام المناهج وطرق التدريس وله باع كبير في مجال دراسات التحليل البعدي، والإحصاء التربوي، والتحليل الناقد للبحوث والدراسات اكتسبه من خبرته الطويلة في هذا المجال، الأستاذ الدكتور عبد المنعم أحمد حسن، الذي عمل استاذ ورئيسا لقسم المناهج وطرق التدريس بكلية التربية جامعة الأزهر الشريف، كما عمل رئيسا لتحرير المجلة الدولية للأبحاث التربوية (مجلة كلية التربية جامعة الامارات سابقا) وعضوا بالمجلس العلمي لترقيات أعضاء هيئة التدريس بجامعة الإمارات ومقررا للجنة العلمية الدائمة لترقية الأساتذة والأساتذة المساعدين تخصص مناهج وطرق تدريس وتكنولوجيا التعليم بجامعة الأزهر، وله العديد من الأبحاث المنشورة في المجلات العربية والأجنبية.

هذا المقال يعرض أساليب وطرق حساب حجوم الأثر من التحليلات الإحصائية المختلفة،وكيفية تفسيرها؛ لتسهم في تقديم توصيات واضحة عن تأثير التجارب البحثية المختلفة.

والمجلة إذ تقدم هذا النتاج الفكري الراقي فإنها تتقدم بخالص الشكر والتقدير لكاتب المقال، وتتمنى أن يكون وجبة دسمة يفيد منها الباحثون في المجالات التربوية المختلفة

رئيس التحرير.

حجوم الأثر في الدراسات الأولية والتحليل البعدي

طرق حسابها وأساليب تفسيرها

عبد المنعم أحمد حسن
قسم المناهج وطرق التدريس،كلية التربية- جامعة الأزهر بالقاهرة
البريد الإلكتروني: ah21348@azhar.edu.eg

مقدمة:

لعل المتتبع للدراسات والبحوث التي تنشر في الدوريات التربوية والنفسية سواء كانت عربية أو أجنبية يلاحظ أن هناك عدداً من أوجه النقد التي وجهت لاختبارات الدلالة الإحصائية للفروض الصفرية، فقد رأى عدد من المتخصصين ومنهم (Krantze, 1999; Nix & Barnnette, 1998) أن هذا النوع من الاختبارات أدى إلى إعاقة التقدم العلمي في الميدان التربوي والنفسي، وذلك لأن الباحثين وقعوا في أخطاء عديدة عند تفسيرهم لنتائج دلالات هذه الفروق والتي منها : الخلط بين الدلالة الإحصائية للنتائج والدلالة العملية لها أو ما يشار إليه بمقدار حجم الأثر "Effect Size"، وهذا الخلط يتمثل في أن بعض الباحثين يرى أنه إذا تغيرت قيمة الدلالة الإحصائية من 0.05 إلى 0.01 مثلاً فهذا يعني أن هذه النتائج ذات فاعلية أكبر أو قيمة عملية أكبر (Nickerson, 2000) ، ومن بين الأخطاء أيضاً إعتقاد بعض الباحثين أنه إذا كانت النتائج دالة احصائياً فهي مهمة وأنها إذا كانت غير دالة احصائياً فهي قليلة أو عديمة الأهمية (Thompson, 1996). هذا بالاضافة الى أن عدداً آخر من الباحثين قد لايدرك أن الدلالة الإحصائية قد تكون نتيجة لكبر حجم عينة البحث بغض النظر عن كون هذه النتيجة لها دلالة عملية أم لا ، أي أن الدلالة الإحصائية دالة لحجم العينة بدلاً من كونها دالة للفروق العملية بين المتوسطات (نصار ، 2006) وأخيراً فقد يعتقد بعض الباحثين أن وجود دلالة احصائية لنتائج بحث ما يعني قابلية هذه النتائج للتكرار على عينات أخرى مسحوبة من نفس المجتمع ، لكن واقع الأمر كما يرى تومسون (Thompson, 1999) أن اختبار الدلالة الإحصائية يحدد درجة احتمال أن يكون الفرض الصفري حقيقي في المجتمع عند التحليل الاحصائي لنتائج عينة عشوائية مسحوبة من هذا المجتمع ولا يعني بالضرورة أن سحب عينات عشوائية أخرى من نفس المجتمع سيؤدي إلى الحصول على النتائج نفسها.

وبسبب هذا النقد لاختبارات الدلالة الإحصائية للفروض فإن عدداً آخر من الباحثين ينصحون باستخدام مؤشرات حجم الأثر باعتبارها تؤدي بالباحثين إلى مزيد من فهم النتائج التي توصلوا إليها (السعيد ، 2003 ، منصور، 1997 ، نصار ، 2006) ، بل إن بعض الباحثين ذهب إلى أنه يجب استخدام المؤشرين معاً، مؤشر الدلالة الإحصائية لاختبار الفروض الصفرية ، ومؤشرات حجم الأثر- باعتبارهما مكملان لبعضهما بعضاً ، أو باعتبارهما وجهان لعملة واحدة (منصور ، 1997 ، Fan, 2001).

ولقد أشار هذان الباحثان (منصور 1997 وFan,2001) وغيرهما (Thompson, 1999; Trusty, Thompson & Petrocelli, 2004) إلى عدد من مميزات استخدام مؤشرات حجم الأثر نذكر منها: أن قيمة حجم الأثر المحسوبة من بيانات دراسة ما توضح مدى قوة العلاقة بين المتغير المستقل والمتغير التابع، أو أنها تحدد أثر المتغير المستقل على المتغير التابع ، وأن قيمة حجم الأثر

لا تتوقف فقط عند مجرد مقارنتها بمعايير محددة مثل معايير كوهين (Cohen) التي حددها ليصف بها متى تكون هذه القيمة مرتفعة أو متوسطة أو منخفضة بل تتعدى ذلك لتسمح للباحث بأن يقارن قيمة حجم الأثر التي حصل عليها بمقادير حجوم الأثر في الدراسات المشابهة لدراسته. كما تسمح قيمة حجم الأثر التي يحصل عليها بعض الباحثين بمقارنة نتائج الدراسات التي أجريت حول ظاهرة ما أو مشكلة تربوية أو نفسية معينة مما يسمح بتعميم نتائج هذه الدراسات كما يحدث في دراسات التحليل البعدي Meta- analysis. هذا بالإضافة إلى المرونة التي تتمتع بها بعض هذه المؤشرات، فعلى سبيل المثال يمكن تحويل مؤشر حجم الأثر "*d*" لكوهين الذي يشير إلى الفرق المعياري بين متوسطين إلى حجم الأثر *r* وهو مؤشر يشير إلى قوة واتجاه العلاقة بين المتغيرات (Cohen, 1988).

وبناء على هذه المميزات فقد لوحظ أن كثيراً من المتخصصين في الإحصاء التربوي والنفسي ورؤساء تحرير الدوريات العلمية المحلية والعالمية ينصحون الباحثين الذين يريدون نشر بحوثهم في هذه الدوريات بأن يبينوا قيمة مؤشرات حجم الأثر لبحوثهم (Meline & Wang, 2004; Thompson, 2002; Trusty, Thompson &, Petrocelli, 2004) . بل إن دليل النشر العلمي للبحوث الصادر عن جمعية علم النفس الأمريكية في طبعاته الخامسة الصادرة في عام (2005) والسادسة الصادرة في عام (2010) والسابعة الصادرة في عام (2020) أوصى الباحثين بضرورة استخدام هذه المؤشرات في تفسير نتائج البحوث التربوية والنفسية وعلى وجه الخصوص بعد فحص قيمة الدلالة الإحصائية الإحتمالية *p* (American Psychological Association (APA), 2005, 2010, 2020).

الدلالة الإحصائية وحجم الأثر

الدلالة الإحصائية

هناك نوعان من الدلالة الإحصائية : النوع الأول هو قيمة إحتمالية يحددها الباحث قبل أن يجمع بياناته ويحللها، وهذه القيمة توضح مقدار الخطأ الذي يقبل الباحث أن يقع فيه نتيجة لرفض الفرض الصفري. بمعنى أنه إذا اتخذ الباحث قراراً على أساس النتائج التي حصل عليها تجريبياً برفض الفرض الصفري، فإن احتمال خطأ هذا القرار يكون أقل من أو مساوي لهذه القيمة التي يرمز لها بالحرف ألفا (α) (علام، 1993أ، ص96) وعادة ما يستخدم الباحثون في البحوث التربوية والنفسية قيمة تساوي 0.05 أو 0.01 في أغلب الأحيان. أما النوع الثاني فهو قيمة الدلالة الإحصائية الحقيقية أو الدلالة الإحتمالية المحسوبة عند استخدام الرزم الإحصائية مثل SPSS، فالنتائج المحسوبة باستخدام مثل هذه الرزمة تكون مصحوبة بقيمة احتمالية يشار إليها بالحرف الإنجليزي *p* ، وهذه القيمة لا يحددها الباحث مثل (α). وبمعنى آخر فهي تقدير لاحتمال حدوث النتائج من عينة ما عندما يكون الفرض الصفري صحيحاً، وإذا كانت قيمة (*p*) أقل من أو تساوي قيمة (α) المحددة من قبل الباحث ، فإن الباحث يرفض الفرض الصفري، ويقال حينئذ أن النتائج دالة احصائياً، أما إذا كانت قيمة الاحتمال (*p*) المحسوبة أكبر من قيمة (α) المحددة من قبل الباحث، فإن الباحث سوف يفشل في رفض الفرض الصفري، ويقال حينئذ أن النتائج غير دالة احصائياً. وإذا كانت النتائج دالة احصائياً فإن هذا لا يعني أن هذه النتائج ذات قيمة عملية أو مهمة – كما سبق القول- بل إنها تشير إلى مدى ثقتنا في أن النتائج التي حصلنا عليها لا ترجع إلى أخطاء المعاينة Sampling Errors .

حجم الأثر

يُعد حجم الأثر من المفاهيم الإحصائية المهمة التي تستخدم في كثير من الاختبارات الإحصائية للفرض الصفري. فإذا اعتبرنا أن الفرض الصفري يُعبر عن "غياب الظاهرة قيد البحث في مجتمع الدراسة، فإن حجم الأثر يُعبر عن درجة وجود هذه الظاهرة في ذلك المجتمع . وكلما كانت قيمة حجم الأثر كبيرة دل ذلك على وجود الظاهرة بدرجة أكبر في ذلك المجتمع" (Cohen, 1988, p. 9-10). ولما كان الباحث نادراً ما يحصل على بياناته من مجتمع الدراسة ، فإن حجم الأثر الذي يحسبه الباحث ليس سوى "تقدير" لحجم العلاقة بين المتغيرات في عينة من هذا المجتمع.

ويجب أن نميز هنا بين حجم الأثر الذي يحدده الباحث قبل إجراء بحثه Hypothesized Effect Size وحجم الأثر الذي يحصل عليه من نتائج البحث Obtained Effect Size فحجم الأثر الذي يحدده الباحث قبل إجراء بحثه هو الذي يستخدمه عند حساب حجم العينة المطلوبة في بحث ما للحصول على مقدار القوة الإحصائية لاختبار إحصائي Power of the Test أما حجم الأثر المحسوب والذي يحصل عليه الباحث بعد تحليل بياناته فهو الذي يعطينا قيمة الدلالة العملية للنتائج التي حصلنا عليها من عينة مسحوبة من مجتمع ما (Huck, 2012; Knapp, 1998).

وعلى هذا يمكن تعريف حجم الأثر بأنه مقدار العلاقة أو الفرق الناتج بين مجموعتين عند التعرض لمعالجات تجريبية. فعندما تكون البيانات مستمرة فيعبر عن حجم الأثر بالفرق المعياري بين المتوسطات، أو بقوة العلاقة الإرتباطية بينهما. أما إذا كانت البيانات تصنيفية فيعبر عن حجم الأثر لها بنسبة مئوية قد تكون نسبة الأرجحية أو نسبة المخاطرة. وغالبا ما يشير حجم الأثر للأهمية العملية للنتائج التي تم التوصل إليها في سياق معين، بالإضافة إلى أنه يستخدم في التحليل البعدي لنتائج البحوث لتقدير متوسط حجوم الأثر لكل الدراسات المتضمنة في هذا النوع من التحليل.

ويستخدم الباحثون مؤشرات مختلفة لتقدير حجم الأثر أو الدلالة العملية ، وهذه المؤشرات تختلف باختلاف هدف البحث وتصميمه وطريقة تحليل بياناته . ونظراً للعدد الكبير من هذه المؤشرات والمرتبطة بالإحصاء البارامتري فقد صُنفت في مجموعتين رئيستين هما : مجموعة مؤشرات الفروق المعيارية بين المتوسطات Measures of Standardized Differences Indices والتي منها على سبيل المثال مؤشر كوهين "d" ومؤشر جلاس "Δ" ومؤشر هدجز "g" . ومجموعة مؤشرات قوة العلاقة أو الارتباط بين المتغيرات والتي تشير إلى حجم التباين المُفسر Variance Accounted for Indices ومن أمثلة هذه المؤشرات r^2 "η^2" ، "ω^2" ، "R^2" (Thompson, 2002). كما صنف تومسون هاتين المجموعتين إلى مؤشرات مُعدّلة وأخرى غير مُعدّلة ، فعلى سبيل المثال عند استخدام تحليل التباين بأنواعه قد يستخدم الباحث المؤشر غير المعدل (η^2) أو المؤشر المعدل (ω^2). هذا بالإضافة إلى وجود مجموعة أخرى من مؤشرات حجم الأثر أو الدلالة العملية والتي تستخدم في الإحصاء اللابارامتري كما سيرد ذكر ذلك.

استخدام حجم الأثر كبديل لقيم الدلالة الاحصائية p

عادة ما يرد في تقارير البحوث والدراسات قيم p التي تعبر عن الدلالة الإحصائية الإحتمالية عند إختبار الفروض. فقد يستخدم بعض الباحثين قيم p كبديل لحجم الأثر خاصة عندما تكون قيم p دالة إحصائياً فتؤخذ هذه القيمة كدليل على كبر تأثير حجم المعالجة، أما إذا كانت قيمة p غير دالة إحصائيا فتؤخذ كدليل لضعف تأثير المعالجة. لكن لابد من ملاحظة أن

قيم p تتأثر بحجم عينة البحث فكلما كان حجم العينة كبيراً كانت p دالة إحصائياً و كلما صغر حجم العينة تكون قيمة p غير دالة إحصائيا. فعلى سبيل المثال إذا كان لدينا دراستين الأولى ذات حجم أثر كبير قيمته 0.6 وحجم العينة 20 فرداً ومن ثم جاءت نتائج الدراسة غير دالة إحصائيا أما الدراسة الثانية فكان لها حجم أثر ضعيف قيمته 0.2 لكن حجم العينة يساوي 500 فرداً ومن ثم جاءت النتائج دالة إحصائياً حينئذ يمكن أن نستنج أنه قد يكون حجم الأثر كبيراً وحجم العينة صغيراً لكن النتائج غير دالة إحصائياً وبالعكس قد يكون حجم الأثر صغيراً وحجم العينة كبيراً لكن النتائج دالة إحصائياً. وفي كلتا الحالتين يجب عدم الإعتماد على قيم الدلالة الإحصائية p كمؤشر لحجم الأثر. وفي البحوث الأولية يمكن للباحث أن يتجنب هذا الخلط من خلال حساب قيمة حجم الأثر وفترة الثقة له (Borenstien & Hedges, 2019)

<u>مصطلحا حجم الأثر و أثر المعالجة</u>

يميز بورنستين وهيدجز (Borenstien & Hedges, 2019) بين مصطلحين هما حجم الأثر وأثر المعالجة ففي المجال الطبي يشار إلى حجم الأثر بمصطلح "أثر المعالجة" (Treatment Effect) ويستخدم في هذا المجال مؤشرات لحجم الأثر مثل نسبة المخاطرة وفرق المخاطرة ونسبة الأرجحية. أما في المجالات النفسية والتربوية والإجتماعية فيستخدم عادة مصطلح "حجم الأثر" (Effect Size). وفي هذه المجالات يستخدم مؤشرات لحجم الأثر مثل الفرق غير المعياري بين متوسطين أو الفرق المعياري بين متوسطين أو معاملات الإرتباط بين متغيرين أو أكثر. ويحدد بورنستين وهدجيز (Borenstien & Hedges, 2019, p. 209) أربعة إعتبارات يمكن النظر إليها عند إختيار أحد مؤشرات حجم الأثر وهي:

1. أن حجوم الأثر المحسوبة من الدراسات الأولية يمكن مقارنتها ببعضها. بمعنى أن حجوم الأثر لا تعتمد على تصميم البحوث والتي يمكن أن تتغير من بحث لآخر (فقد تتغير حجوم العينات أو قد يستخدم في البحث متغيرات وسيطة مثلاً).
2. أن يكون حجم الأثر الذي تم حسابه له معنى وقابل للتفسير. وهذا يعني أن المتخصصين في المجال موضع البحث يفهمون معنى حجم الأثر الذي تم حسابه ويستطيعون تفسيره. كما يمكن لهؤلاء الباحثين تحويل أحد مؤشرات حجم الأثر إلى مؤشر أخر (مثال ذلك تحويل مؤشر حجم الأثر d إلى المؤشر r).
3. أنه يمكن حساب حجم الأثر من الدراسات والبحوث المنشورة أو غير المنشورة التي ستضُمن في التحليل البعدي، وهذا يعني أنه ليس من المرجّح أن يتطلب التحليل البعدي اللجوء إلى الدرجات الخام لكل بحث أولي ومن هذه البيانات الخام يتم حساب حجم الأثر. وإنما يتم في الواقع حساب حجوم الأثر للدراسات والبحوث من ملخصات النتائج التي تحتوي على المتوسطات والإنحرافات المعيارية.
4. أن يحقق حجم الأثر المستخدم في التحليل البعدي الخصائص والشروط الإحصائية اللازمة، فعلى سبيل المثال يجب أن يكون التوزيع العيني لحجم الأثر معلوماً ومن ثم يمكن حساب تباينه وفترة الثقة له.

ومن الناحية العملية فإن قرار إختيار حجم الأثر المناسب في التحليل يعتمد على ما لدى الباحث من بيانات الدراسات الأولية. فإذا كان لدى الباحث المتوسطات والإنحرافات المعيارية فيمكنه حينئذ استخدام الفرق المعياري بين المتوسطات ، وإذا كانت البيانات المتوفرة من البحوث الأولية هي معاملات الإرتباط، فإن حجم الأثر المناسب هنا هو معامل الإرتباط نفسه.

حساب حجوم الأثر

من الناحية العملية لا يتعامل الباحثون عادة مع حجم الأثر في المجتمع وإنما يتعاملون مع حجوم الأثر لعينات مسحوبة من المجتمع وبالتالي يحاولون تعميم نتائجهم على المجتمع. وبطبيعة الحال ستختلف حجوم الأثر في العينات عن حجم الأثر في المجتمع الذي سحبت منه هذه العينات، ونتيجة لهذا الإختلاف نحصل على التوزيع العيني لحجوم الأثر (Sampling Distribution) وبناء على النظرية الإحصائية يمكننا حساب تباين حجوم الأثر في هذه العينات.

ونظراً لتعدد حجوم الأثر تتعدد طرق الحصول عليها من بيانات البحوث، ومن ثم سنعرض فيما يلي عدداً من طرق حساب حجوم الأثر التي يمكن أن تستخدم سواء للدراسات والبحوث الأولية أو تستخدم في التحليل البعدي للبحوث والدراسات. ولقد إعتمدنا في ذلك على ما ورد في عدة كتب ومقالات منها الكتب والمقالات التالية (Borenstien & Hedges, 2019, Borenstien et al, 2009, Card, 2012,, Cumming, 2012, & Ellis, 2010). وسيتم العرض بالترتيب التالي:

- حجوم الأثر القائمة على المتوسطات
  - حجم الأثر للفرق بين متوسطين لمجموعتين مستقلتين أو مترابطتين
  - حجم الأثر للفرق المعياري بين متوسطين لمجموعتين مستقلتين أو مترابطتين
- حجوم الأثر القائمة على معاملات الإرتباط

حجوم الأثر القائمة على المتوسطات

عندما تحتوي الدراسات المنشورة على المتوسطات والانحرافات المعيارية فإنه يمكن حساب حجم الأثر من متوسطات الدرجات الخام مباشرة أو يحسب حجم الأثر من الفرق المعياري بين المتوسطات كما يلي:

أولاً: حساب حجم الأثر من متوسطات الدرجات الخام (حجم الأثر غير المعياري) ويرمز له بالرمز (D):

عندما تكون نتيجة أو نتائج البحوث معروضة على هيئة متوسطات وانحرافات معيارية تم حسابها من الدرجات الخام فإنه يمكن حساب حجم الأثر من هذه المتوسطات بصورة مباشرة. والميزة الأساسية في هذه الحالة هي أن حجم الأثر المحسوب بهذه الطريقة له معنى مباشر. فعلى سبيل المثال إذا أخذنا في الاعتبار مجموعتين إحداهما تلقت معالجة والأخرى ضابطة وإذا افترضنا أن μ_1، μ_2 هما متوسطا هاتين المجموعتين في المجتمع، فإن الفرق بين هذين المتوسطين والذي يعبر عن حجم الأثر في هذه الحالة يعطى من المعادلة:

$$\Delta = \mu_1 - \mu_2 \qquad (1)$$

حيث Δ هي الفرق بين المتوسطين في المجتمع أما عند تقدير حجم الأثر لعينتين فسنستخدم الحرف الإنجليزي D وفيما يلي سنبين كيفية تقدير حجم الأثر D لهذين البارامترين من الدراسات التي تحتوي على مجموعات مستقلة والدراسات التي تحتوي على مجموعات معتمدة أو متكافئة

أ. حساب حجم الأثر D لمجموعتين مستقلتين:

يتم حساب قيمة D لمجموعتين (عينتين) مستقلتين كما يلي:

نفترض أن $\bar{X}_1$، $\bar{X}_2$ هما متوسطا المجموعتين المستقلتين، فإن حجم الأثر والذي يمثله الإحصائيD يعطى من العلاقة الآتية:

$$D = \bar{X}_1 - \bar{X}_2 \qquad (2)$$

لاحظ أننا استخدمنا الحرف D لحساب الفرق بين متوسطي الدرجات الخام مباشرة، وسنستخدم الحرف d عند حساب الفرق المعياري بين المتوسطين.

وإذا افترضنا أن S_1, S_2 هما الانحرافان المعياريان لدرجات المجموعتين وأن n_1, n_2 هما عدد أفراد كل مجموعة.

وإذا افترضنا تجانس الانحرافين المعيارين، فإن تباين حجم الأثر V_D يعطي من المعادلة التالية:

$$V_D = \frac{n_1 + n_2}{n_1 n_2} \times S^2_{pooled} \qquad (3)$$

أما S_{pooled} والتي تشير إلى التباين المشترك فنحصل عليها من المعادلة الآتية:

$$S_{pooled} = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}} \qquad (4)$$

أما إذا كان الانحرافان المعياريان في المجتمع غير متجانسين أي أن $\sigma_1 \neq \sigma_2$ فإن تباين حجم الأثر يعطى من المعادلة التالية:

$$V_D = \frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} \qquad (5)$$

وفي كلتا الحالتين (سواء كان الإنحرافان المعياريان متساويين أم غير متساويين) فإن الخطأ المعياري يمكن إيجاد قيمته من المعادلة الأتية:

$$SE_D = \sqrt{V_d} \qquad (6)$$

مثال:

إذا افترضنا أن لدينا مجموعتين إحداهما تجريبية والأخري ضابطة وكان متوسط المجموعة الأولى $\bar{x}_1 = 103$ ومتوسط المجموعة الثانية $\bar{x}_2 = 100$ وأن $n_1 = n_2 = 50$ وأن $s_1 = 5.5$ و$s_2 = 4.5$ إحسب قيمة حجم الأثر وقيمة الانحراف المعياري المشترك لحجم الأثر وتباين حجم الأثر والخطأ المعياري لحجم الأثر.

الحل

قيمة حجم الأثر هي

$$D = 103 - 100 = 3.00$$

2. وإذا افترضنا أن الانحرافين المعياريين في المجتمع متساويان أي أن $\sigma_1 = \sigma_2$ أي متجانسين فإن قيمة الأنحراف المعياري المشترك في المجموعتين تعطي من المعادلة التالية:

$$S_{pooled} = \sqrt{\frac{(50-1)\times 5.5^2 + (50-1)\times 4.5^2}{50+50-2}}$$
$$= 5.0249$$

3. ومقدار تباين حجم الأثر هو

$$V_D = \frac{50+50}{50x50} \times 5.0240^2 = 1.0100$$

4. الخطأ المعياري لحجم الأثر هو

$$SE_D = \sqrt{1.0100 =} \; 1.0050$$

أما إذا كان $\sigma_1^2 \neq \sigma_2^2$ فإن تباين حجم الأثر يتم حسابه من المعادلة التالية:

$$V_D = \frac{5.5^2}{50} + \frac{4.5^2}{50} = 1.010$$

ويحسب الخطأ المعياري من المعادلة التالية:

$$SE_D\sqrt{1.010} = 1.0050$$

من هذا المثال نلاحظ أن المعادلتين 3، 5 تعطيان نفس القيمة، وهذه الحالة تحدث فقط عندما يكون حجما العينتين متساويين والتباينين متساويين في المجموعتين.

ب . حساب قيمة حجم الأثر غير المعياري D لمجموعتين (مترابطتين) أو لإختبارين قبلي وبعدي

المعادلات المذكورة سابقا يمكن استخدامها مع مجموعتين مستقلتين، ولكن هناك تصميمات بحثية يستخدم فيها مجموعات متكافئة مثل التوائم المتماثلة أو تلاميذ بنفس نسبة الذكاء يتم مزاوجتهم. ومن ثم فإن وحدة التحليل الأحصائي هى الأزواج. وإذا كان هذا الارتباط مرتفعًا فإن قيمة التباين ستقل ومن ثم تزداد الدقة. وإذا كان لدينا درجات مختلفة لكل زوج من الأزواج فإننا سنحصل على متوسط الفرق بين الدرجات $\overline{X}_{Dif}$ والانحراف المعياري للفرق بين الدرجات S_{Diff} ومن ثم فإن حجم الأثر في هذه الحالة يعطى من المعادلة التالية:

$$D = \overline{x}_{diff} \qquad (7)$$

وتباين حجم الأثر يعطى من المعادلة:

$$V_D = \frac{s_{Diff}^2}{2} \qquad (8)$$

حيث n هى عدد الأزواج

بينما الخطأ المعياري لحجم الأثر يعطى من المعادلة التالية:

$$SE_D = \sqrt{V_D} \qquad (9)$$

مثال:

إذا كان فرق المتوسطين بين مجموعتين في إحدى التجارب يساوي 5 وأن الأنحراف المعياري للفرق يساوي 10 وأن عدد الأزواج n يساوي 50 زوجا فإن:

$$D = \bar{X}_{Diff} = 5$$

$$V_D = \frac{10^2}{50} = 2.00 \qquad (10)$$

$$SE_D = \sqrt{2.00} = 1.4142 \qquad (11)$$

وكبديل لما سبق إذا كان لدينا متوسط كل مجموعة وانحرافها المعياري فإن الفرقD يساوي

$$D = \bar{X}_1 - \bar{X}_2 \qquad (12)$$

$$V_D = \frac{S_{Diff}^2}{n} \qquad (13)$$

$$SE_{Diff} = \sqrt{V_{Diff}} \qquad (14)$$

في هذه الحالة لابد من حساب الانحراف المعياري للفرق بين الدرجات من الانحراف المعياري لكل زوج من الأزواج من المعادلة الآتية:

$$S_{Diff} = \sqrt{S_1^2 + S_2^2 - 2r \times S_1 \times S_2} \qquad (15)$$

حيث r هي معامل الارتباط بين الأزواج المتكافئة، وإذا كانت $S_1=S_2$ فيمكن تبسيط المعادلة(15) لتصبح كما يلي:

$$S_{Diff} = \sqrt{2 \times S_{pooled}^2 (1-r)} \qquad (16)$$

وفي كلتا الحالتين كلما اقتربت قيمة r من الواحد الصحيح فإن قيمة الخطأ المعياري للفرق ستقل، وإذا كانت r تساوي صفر فإن الخطأ المعياري في هذه الحالة يصبح مساويًا للخطأ المعياري لمجموعتين مستقلتين.

مثال:

افترض أن متوسط أزواج من الأقارب A، B هو 105، 100 على الترتيب بانحرافين معياريين هما 10، 10 وأن معامل الإرتباط بين مجموعتي الأزواج هو 0.5 وأن عدد الأزواج هو 50 زوجا، أوجد مقدار حجم الأثر، وتباينه والخطأ المعياري له.

الحل

$$D = 105 - 100 = 5.00$$

$$V_D = \frac{S_{Diff}^2}{n}$$

للحصول على قيمة تباين حجم الأثر لابد من إيجاد قيمة $\boldsymbol{S_{Diff}}$ والتي نحصل عليها من المعادلة التالية:

$$S_{Diff} = \sqrt{S_1^2 + S_2^2 - 2r \times S_1 \times S_2}$$
$$= \sqrt{10^2 + 10^2 - 2 \times 0.50 \times 10 \times 10} = 10$$

وبناء عليه فإن مقدار تباين حجم الأثر هو

$$V_D = \frac{S_{Diff}^2}{n} = \frac{10^2}{50} = 2$$

ومقدار الخطأ المعياري هو

$$SE_{Diff} = \sqrt{V_{Diff}} = \sqrt{2.00} = 1.414$$

وإذا كان الإنحرافان المعياريان متساويين فيمكن استخدام المعادلة المبسطة التالية لحساب الإنحراف المعياري للفرق

$$S_{Diff} = \sqrt{2 \times S_{pooled}^2 (1 - r)} = \sqrt{2 \times 10^2 (1 - 0.50)}$$
$$= 10$$

وهذه المعادلات تطبق أيضا في حالة ما إذا كان لدينا إختبار قبلي وأخر بعدي، فالاختباران القبلي والبعدي يشبهان متوسطا المجموعتين المتكافئتين، وn هى عدد الأفراد، r هو معامل الارتباط بين درجات الاختبارين القبلي والبعدي.

ثانيًا: حجم الأثر المعياري لمجموعتين مستقلتين

كما ذكرنا أعلاه فإن حجم الأثر للفرق بين المتوسطين الخام (D) يمكن استخدامه عندما تستخدم الدراسات مقاييس أو اختبارات متشابهة ومن ثم يمكن جمع حجوم الأثر معا كما يحدث في التحليل البعدي. أما في حالة استخدام دراسات مختلفة ذات مقاييس مختلفة فلا يمكن استخدام الفرق بين المتوسطات الخام لحساب قيم حجم الأثر ومن ثم لا يمكن جمع حجوم الأثر هذه كما يحدث في التحليل البعدي.

وفي حالة استخدام مقاييس أو اختبارات مختلفة فهنا نحسب الفرق بين المتوسطين ثم نقسمه على الإنحراف المعياري لكل دراسة (أي نحسب الفرق المعياري بين متوسطين) ومن ثم نحصل على حجم الأثر وحجم الأثر هذا هو الذي يمكن استخدامه للمقارنة بين الدراسات المختلفة أو استخدامه في التحليل البعدي لعدد من الدراسات. وهذه هي الطريقة التي استخدمها كوهين(1979، 1987) في حساب قوة الاختبار الإحصائي.

وإذا كان لدينا مجموعتان متوسط الأولى μ_1 وانحرافها المعياري ρ_1 ، ومتوسط الثانية μ_2 وانحرافها المعياري ρ_2 وإذا كان الانحرافان المعياريان متساويين فإننا يمكننا حساب الفرق المعياري للمتوسطين في المجتمع الذي سحبت منهما العينتان بالمعادلة التالية:

$$\delta = \frac{\mu_1 - \mu_2}{\sigma} \qquad (17)$$

حيث δ هي حجم الأثر في المجتمع.

μ تمثل المتوسط سواء في المجموعة الأولي أو المجموعة الثانية.

σ تعبر عن الانحراف المعياري المشترك في المجتمع.

وفيما يلي سنبين كيفية حساب الفرق المعياري بين متوسطي عينتين مستقلتين.

أ- **حساب حجم الأثر d، g من الدراسات التي تستخدم مجموعات مستقلة:**

يمكن تقدير d ، g **من المعادلات التالية:**

$$d = \frac{\overline{x}_1 - \overline{x}_2}{s_{pooled}} \qquad (18)$$

حيث $\overline{x}_1 - \overline{x}_2$ عبارة عن الفرق بين متوسط المجموعة الأولى والمجموعة الثانية.

s_{pooled} هي الإنحراف المعياري المشترك بين المجموعتين، والذي يمكن حسابه من المعادلة التالية:

$$s_{pooled} = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}} \qquad (19)$$

حيث s_1, s_2 هما الانحرافان المعياريان للمجموعة الأولى والمجموعة الثانية وكل من n_1 ، n_2 هما عدد الأفراد في كل مجموعة.

والسبب الذي يدعونا إلى حساب الانحراف المعياري المشترك بين المجموعتين هو أنه حتى لو افترضنا أن الانحرافين المعياريين في المجتمعين متساويان، فإنه من غير المحتمل أن يكون الانحرافان المعياريان للعينتين أو المجموعتين متساويين، ولكن من خلال إيجاد الانحراف المعياري المشترك فإننا نحصل على تقدير أكثر دقة للقيم المشتركة بين هذين الانحرافين المعياريين. ويُشار إلى الفرق المعياري بين المتوسطين بالحرف d أو مؤشر كوهين d في حالة استخدام العينات.

لاحظ أننا استخدمنا الحرف اللاتيني δ للإشارة إلى حجم الأثر في المجتمع كما استخدمنا الحرف الإنجليزي d للإشارة إلى حجم الأثر في العينات

أما قيمة تباين حجم الأثر d فيمكن حسابها من المعادلة التالية:

$$V_d = \frac{n_1 + n_2}{n_1 n_2} + \frac{d^2}{2(n_1 + n_2)} \qquad (20)$$

ويمكن حساب قيمة الخطأ المعياري لحجم الأثر d من المعادلة التالية:

$$SE_d = \sqrt{V_d} \qquad (21)$$

ولقد بين هيدجز (Hedges) أن قيمة حجم الأثر d متحيزة قليلًا وأن هذا التحيز يمكن التغلب عليه باستخدام تعديل بسيط يعطينا قيمة غير متحيزة لحجم الأثر، ويرمز لحجم الأثر غير المتحيز بالحرف g أو هدجيز g . ولقد أعطى هدجيز المعادلة التالية لتعديل حجم الأثر:

$$J = 1 - \frac{3}{4df - 1} \qquad (22)$$

حيث df تشير إلى درجات الحرية التي استخدمت لتقدير s_{pooled} ودرجات الحرية لمجموعتين مستقلتين هي $(n_1 + n_2 - 2)$.

وبناء على ما سبق فإن:

- حجم الأثر المعدل يصبح:

$$g = j \times d \qquad (23)$$

- وتباين حجم الأثر يصبح:

$$V_g = j^2 \times V_d \qquad (24)$$

- والخطأ المعياري يصبح:

$$SE_g = \sqrt{V_g} \qquad (25)$$

مثال:

إذا افترضنا أن لدينا مجموعتين تجريبيتين متوسط المجموعة الأولى 103 ومتوسط المجموعة الثانية 100 والانحراف المعياري للمجموعة الأولى 5.5 والانحراف المعياري للمجموعة الثانية 4.5 وأن عدد الأفراد في المجموعتين متساوي وقيمته 50 فإننا يمكن باستخدام المعادلات السابقة أن نحسب كل القيم التي عبرت عنها المعادلات كما يلي :

الحل

- نحسب أولاً قيمة الانحراف المعياري المشترك كما يلي:

$$s_{pooled} = \sqrt{\frac{(50-1)5.5^2 + (50-1)4.5^2}{50+50-2}} = 5.0249$$

- ومن ثم يمكن حساب قيمة حجم الأثر d كما يلي:

$$d = \frac{103-100}{5.0249} = 0.5970$$

- نحسب تباين حجم الأثر كما يلي:

$$V_d = \frac{50+50}{50 \times 50} + \frac{0.5970^2}{2(50+50)} = 0.0418$$

- نحسب الخطأ المعياري لحجم الأثر كما يلي:

$$SE_d = \sqrt{0.0418} = 0.2044$$

- وعند استخدام تعديل هيدجيز نحصل على القيم التالية:

$$J = 1 - \frac{3}{4(98-1)} = 0.9923$$

ومن ثم فإن حجم الأثر بعد التعديل g وتباينه والخطأ المعياري له يعطى كما يلي:

$$g = 0.9923 \times 0.5970 = 0.5942$$
$$V_g = 0.9923^2 \times 0.0418 = 0.0411$$
$$SE_d = \sqrt{0.0411} = 0.2028$$

من الملاحظ أن معامل التصحيح لهيدجيز دائما أقل من الواحد الصحيح وبالتالي فإن قيمة (g) ستكون أقل من قيمة (d) وأن تباين (g) سيكون أقل من تباين (d)، ويجب أن نلاحظ

أيضا أن معامل التصحيح (J) سيكون قريبا من الواحد إلا إذا كانت درجات الحرية أقل من 10 وكما يتضح من المثال المحلول، فإن الفرق بين (g) و (d) غير ذي قيمة تذكر.

ب- **حساب حجم الأثر المعياري لمجموعتين مترابطتين (اختبار قبلي وبعدي)**

يمكننا تقدير الفرق المعياري بين متوسطين من البحوث التي استخدمت تصميم المجموعات المترابطة (أي المجموعات ذات الاختبار القبلي – والبعدي) أو مجموعات متكافئة ومن ثم نستخدم المعادلة التالية:

$$d = \frac{\bar{Y}_{diff}}{S_{pooled}} = \frac{\bar{Y}_1 - \bar{Y}_2}{S_{pooled}} \qquad (26)$$

نلاحظ أن المعادلة السابقة تشبه المعادلة (18) والتي تستخدم لحساب قيمة حجم الأثر المعياري لمجموعتين مستقلتين، وكما ذكرنا من قبل فعند التعامل مع مجموعات مستقلة فإن الوحدة الطبيعية للانحراف هي الانحراف المعياري المشترك بين المجموعتين، وهذه القيمة يمكن حسابها بسهولة. وعلى العكس من ذلك عندما نتعامل مع مجموعات متكافئة أو متوسطات قبلية وبعدية فإن الوحدة الطبيعية للانحراف هي الانحراف المعياري للفرق بين الدرجات وهذه القيمة يجب أن يعرفها الباحث الذي يجري البحث. ولحساب قيمة d من الانحراف المعياري للفرق، فإننا يجب أن نحسب الانحراف المعياري داخل المجموعتين أي الانحراف المعياري المشترك والذي يجب أن يكون موجودًا في مقام المعادلة (26).

إذاً عند التعامل مع مجموعات متكافئة أو متوسطات قبلية وبعدية فإن الانحراف المعياري المشترك يمكن حسابه من الانحراف المعياري للفرق بين المتوسطين باستخدام المعادلة التالية:

$$S_{pooled} = \frac{S_{diff}}{\sqrt{2(1-r)}} \qquad (27)$$

- حيث r هي معامل الارتباط بين درجات الاختبارين القبلي والبعدي أو بين درجات مجموعتي المزواجة. وبعد ذلك يمكن تطبيق المعادلة (26) لحساب قيمة d ، أما تباين d فنحصل عليه من المعادلة التالية:

$$V_d = \left(\frac{1}{n} + \frac{d^2}{2n}\right) \times 2(1-r) \qquad (28)$$

حيث n هى عدد الأزواج، أما الخطأ المعياري لحجم الأثر يحسب من المعادلة التالية:

$$SE_d = \sqrt{V_d} \qquad (29)$$

وحيث إن معامل الارتباط r بين درجات الاختباريين القبلي والبعدي مطلوب لحساب الانحراف المعياري المشترك، والذي يمكن حسابه من الانحراف المعياري للفرق فلابد أن نفترض أن معامل الارتباط معلوم أو يمكن حسابه (كما نجده في جدول إختبار t لعينات معتمدة عند

استخدام برنامج SPSS)، والا إضطررنا إلى حساب معامل الارتباط من الدراسات المشابهة، ومن المحتمل أيضاً أن نجري ما يسمى بتحليل الحساسية(Sensitivity Analysis) مستخدمين في ذلك مدى من معاملات الارتباط المحتملة.

ولحساب قيمة هدجيز g والاحصاءات المرتبطة به فإننا نستخدم المعادلات 22، 23، 24، 25 السابق الإشارة إليها أعلاه مع ملاحظة أن درجات الحرية في هذه الحالة هي ($df = n-1$)

مثال:

افترض أن لدينا مجموعة من الطلاب أجري عليها إختباراً قبليا كانت قيمته 100، ثم إختباراً بعدياً جاءت قيمته 103، وأن الانحراف المعياري للفرق بين المتوسطين كان (s_{diff}= 5.5) وأن عدد أفراد هذه المجموعة من الطلاب 50 فرداً وأن معامل الارتباط بين الاختبارين القبلي والبعدي هو (r = 0.7) . إحسب حجم الأثر، وتباين حجم الأثر ، والخطأ المعياري لحجم الأثر.

الحل

1. نحسب قيمة (s_{pooled}) كما يلي:

$$s_{pooled} = \frac{5.5}{\sqrt{2(1-0.7)}} = 7.1005$$

2. نحسب قيمة حجم الأثر d كما يلي:

$$d = \frac{103-100}{7.1005} = 0.4225$$

3. نحسب تباين حجم الأثر كما يلي:

$$V_d = \left(\frac{1}{50} + \frac{0.4225^2}{2\times 50}\right) x2(1-0.7) = 0.0131$$

4. نحسب الخطأ المعياري لحجم الأثر كما يلي:

$$SE_d = \sqrt{0.0131} = 0.1143$$

5. ولحساب قيمة g نستخدم معامل التصحيح J على النحو التالي:

- قيمة التعديل:

$$J = \left(1 - \frac{3}{4(49-1)}\right) = 0.9864$$

- قيمة g

$$g = 0.9864 \times .4225 = .4160$$

- قيمة التباين:

$$V_g = 0.9864^2 \times 0.0131 = 0.0127$$

- قيمة الخطأ المعياري:

$$SE_v = \sqrt{0.0127} = 0.1126$$

تضمين دراسات ذات تصميمات مختلفة في التحليل نفسه:

من وجهة النظر الاحصائية فإن قيمة حجم الأثر d أو g سيكون لها نفس المعنى بغض النظر عن تصميمات البحوث، ومن ثم نستطيع حساب حجم الأثر والتباين لكل دراسة مستخدمين المعادلات المناسبة، ثم بعد ذلك نُضمّن كل الدراسات في دراسة واحدة للتحليل البعدي. وبينما لا يوجد أي معوقات فنية لاستخدام دراسات ذات تصميمات مختلفة في التحليل البعدي ذاته، إلا أنه هناك احتمالا بأن هذه الدراسات قد تختلف عن بعضها اختلافًا جوهريًا كما سنرى فيما بعد.

وفي أي تصميم بحثي فإن اختبار الفرق بين المتوسطين قد يكون $\overline{x}_1 - \overline{x}_2$ أو قد يكون $\overline{x}_2 - \overline{x}_1$ هو من اختيار الباحث إلا إذا قرر الباحث استخدام أحد المتوسطين قبل الآخر في معادلة حساب حجم الأثر. وإذا اختار الباحث أحد المتوسطين قبل الأخر فلابد له أن يستخدم الترتيب نفسه في كل الدراسات، فعلى سبيل المثال إذا كان الفرق الموجب بين المتوسطين يشير إلى أن المجموعة التجريبية كانت أفضل من المجموعة الضابطة فلابد أن يطبق ذلك على كل الدراسات أي لابد أن يأتي متوسط المجموعة التجريبية قبل متوسط المجموعة الضابطة، أو أن يأتي متوسط الاختبار البعدي قبل متوسط الاختبار القبلي.

وفي بعض الدراسات قد يعبر عن النتائج بعدد الإجابات الصحيحة، وفي دراسات أخرى قد يعبر عن النتائج بعدد الإجابات الخطأ فيكون من الضروري حينئذ أن يحدد الباحث إشارة حجم الأثر ويوحد ذلك في كل الدراسات

<u>حجوم الأثر القائمة على معاملات الارتباط</u>

تعرض بعض البحوث والدراسات نتائجها في صورة معاملات إرتباط بين متغيرين أو أكثر، وفي هذه الحالة يُعد معامل الإرتباط مؤشراً لحجم الأثر، وعادة ما يشار إلى معامل الإرتباط في المجتمع بالحرف اللاتيني (ρ) وينطق رو. ومن المعلوم أن معامل الإرتباط يأخذ القيم ± 1 ومعامل الإرتباط الذي قيمته صفر يعني أنه لا توجد علاقة بين المتغيرات. ومعامل الإرتباط الذي قيمته أقل من الصفر يعني أن القيم المرتفعة لأحد المتغيرات ترتبط بالقيم المنخفضة لمتغير آخر. أما إذا كانت قيمة معامل الإرتباط أكبر من الصفر فهذا يعنى أن القيم المرتفعة لأحد المتغيرات ترتبط بالقيم المرتفعة لمتغير آخر والعلامة الموجبة أو السالبة التي توضع قبل قيمة معامل الإرتباط تشير إلى إتجاه العلاقة بين المتغيرات.

ويستخدم كثير من الباحثين معامل الإرتباط كمؤشر لحجم الأثر وذلك لعدد من الأسباب، أولها أن معامل الإرتباط إحصاءة مألوفة لمعظم الباحثين، وثانيها أن معامل الإرتباط متعدد الإستخدامات ، وثالثها أن معامل الإرتباط يمكن استخدامه لمعرفة الفروق بين المتوسطات فضلا عن معرفة قوة العلاقة بين المتغيرات.

وبطبيعة الحال يوجد أنواع مختلفة من معاملات الإرتباط مثل معامل إرتبط بيرسون ومعامل إرتباط سبيرمان ومعامل إرتباط كندال ومعامل الإرتباط المتسلسل وغيرهم. وفي هذا

الجزء سنتعامل مع معامل إرتباط للمتغيرات المستمرة وهو معامل إرتباط بيرسون ونرجئ بقية معاملات الإرتباط لجزء آخر في هذا العرض.

حساب حجم الأثر وتباينه من معامل الارتباط لبيرسون r

تتم الإشارة لمعامل الإرتباط في المجتمع بالحرف ρ أما تقديره في العينة فيشار إليه بالحرف r ويتم حسابه بالطريقة المعروفة أو باستخدام أي برنامج إحصائي، لكن تباين معامل الإرتباط V_r فيتم حسابه من المعادلة رقم 30 الأتية.

$$V_r = \frac{(1-r)^2}{n-1} \qquad (30)$$

حيث n تعبر عن حجم العينة.

ومعظم بحوث التحليل البعدي لا تتعامل مباشرة مع قيم r عند حساب المتوسط الكلي لحجوم الأثر وبدلاً من ذلك يتم تحويل معامل الإرتباط إلى مقياس أخر يسمى بتحويل فيشر Z (Fisher's Z Transformation) (لاحظ عدم الخلط بين تحويل فيشر Z وبين الدرجة المعيارية Z) ثم بعد الإنتهاء من الحسابات باستخدام هذا التحويل يعود الباحث مرة أخرى لتحويل قيم فيشر Z إلى قيم معامل الإرتباط مرة أخرى.

ومعادلة تحويل معامل الإرتباط إلى معامل فيشر Z يُعبر عنها بالمعادلة التالية:

$$Z = 0.5 \times \ln\left(\frac{1+r}{1-r}\right) \qquad (31)$$

وإذا كان الباحث ليس ملماً بالتعامل مع اللوغاريتمات فيمكنه إيجاد قيم فيشر Z من أحد الجداول الموجودة بكتب الإحصاء ومنها على سبيل المثال كتاب جلاس وهوبكن (Glass & Hopkin, 1996).

كما يمكننا إيجاد قيمة تباين معامل فيشر من المعادلة البسيطة التالية.

$$V_Z = \frac{1}{(n-3)} \qquad (32)$$

أما الخطأ المعياري لقيمة فيشر Z فنحصل عليه من المعادلة التالية.

$$SE_Z = \sqrt{V_Z} \qquad (33)$$

وبعد الحصول على قيم Z نحولها مرة أخرى إلى معامل الإرتباط باستخدام المعادلة التالية أو باستخدام الجدول الموجودة بكتاب جلاس وهوبكن (Glass & Hopkin, 1996) كما ذكرنا.

$$r = \frac{e^{2Z} - 1}{e^{2Z} + 1} \qquad (34)$$

مثال:

إذا حصل باحث في دراسته على معامل إرتباط ($r = 0.50$) وكان حجم العينة لديه يساوي 100 فرد فيستطيع حينئذ حساب قيمة فيشر Z كما يلي:

$$Z = 0.5 \times \ln\left(\frac{1+0.5}{1-0.5}\right) = 0.5493$$

$$V_Z = \frac{1}{100-3} = 0.0103$$

$$SE_Z = \sqrt{0.0103} = 0.1015$$

ثم بعد الحصول على قيم فيشر Z تحول هذه القيم لمعامل الإرتباط مرة أخرى على النحو التالي

$$r = \frac{e^{2(0.5493)} - 1}{e^{2(0.5493)} + 1} = 0.50$$

مما سبق نلاحظ أنه إذا إحتوت الدراسات أو البحوث على معاملات إرتباط فإن هذه المعاملات تعبر عن حجوم الأثر لهذه الدراسات. ويمكننا تحويل قيمة أي معامل إرتباط إلى معامل فيشر Z وتجرى كل التحليلات اللازمة باستخدام هذا المعامل، وبعد الإنتهاء من هذه الخطوات نحول قيم Z مرة أخرى لقيم r المناظرة لها.

حساب حجوم الأثر من المعلومات المتاحة في تقارير البحوث

عندما يتاح للباحث نتائج بحث منشور محتوية على المتوسطات والإنحرافات المعيارية وحجوم العينات أو معاملات الإرتباط يصبح من السهل عليه حساب حجوم الأثر من هذه المعلومات. لكن أحياناً لا تتاح كل هذه المعلومات في تقارير البحوث المنشورة ومن ثم يقع على عاتق الباحث أو القائم بإجراء التحليل البعدي استخدام الإختبارات الإحصائية المتاحة في تقارير البحوث لحساب حجوم الأثر وفيما يلي سنعرض لطرق حساب حجوم الأثر لكل من نتائج الإختبارات الأتية مع إعطاء بعض الأمثلة المناسبة في كل حالة.

- إختبار النسبة التائية t
- تحليل التباين
- معاملات الإرتباط
- معامل الإنحدار الخطي الثنائي والمتعدد
- معامل χ^2
- بعض الإختبارات اللابارامترية

أولاً: حساب حجم الأثر من نتائج اختبار t

يوجد ثلاثة أنواع من إختبار النسبة التائية أولاهما النسبة التائية لعينة واحدة، وثانيهما النسبة التائية لعينتين مستقلتين أو مجموعتين مستقلتين ، وثالثهما النسبة التائية لعينتين متزاوجتين (معتمدتين) أو النسبة التائية لنتائج إختبارين قبلي وبعدي.

1. حساب قيمة مؤشر حجم الأثر (d) من إختبار (t) لعينة واحدة

يتم استخدام إختبار النسبة التائية لعينة واحدة لمقارنة متوسط درجات عينة معينة بمتوسط محدد سلفاً أو بمتوسط للمجتمع الذي أشتقت منه العينة. والمعادلة العامة لإختبار النسبة التائية لعينة واحدة هي:

$$t = \frac{\bar{x} - \mu}{s/n}$$

ويمكن حساب حجم الأثر لعينة واحدة من المعادلة التالية

$$d = \frac{\overline{X}}{SD}$$

حيث $\overline{X}$ هي قيمة فرق المتوسطين (المتوسط المحسوب من قبل الباحث مطروحا من المتوسط الإفتراضي)، SD هي الإنحراف المعياري للدرجات وهذه البيانات يمكن الحصول عليها من الجداول الناتجة عند استخدام برنامج SPSS كما سنبين في المثال التالي:

كما يمكن حساب قيمة حجم الأثر بالتعويض في قيمة t من جدول النتائج المعطى من برنامج SPSS من المعادلة التالية:

$$d = \frac{t}{\sqrt{N}}$$

حيث N تشير إلى العدد الكلي لأفراد العينة.

مثال:

قام باحث بقياس إتجاهات عينة قدرها 30 طالباً من طلاب الصف الأول الثانوي نحو الرياضيات مستخدماً مقياساً من نوع ليكرت تم تقنينه على عينات كثيرة بحيث كان متوسط درجات عينات التقنين يساوي 50 أراد هذا الباحث معرفة ما إذا كان يوجد فرق دال إحصائيا بين درجات إتجاهات عينة الطلاب التي إختارها وبين المتوسط المعياري لعينات التقنين المقدر بخمسين درجة. لذلك قام بتحليل نتائجه باستخدام برنامج SPSS وحصل على النتائج المبينة بالجداول التالية. من هذه الجداول حسب الباحث قيمة t لعينة واحدة وحسب أيضا قيمة حجم الأثر d على النحو التالي.

ملحوظة يمكن حساب قيمة (t) من هذه المعلومات ببساطة بالتعويض في المعادلة السابق ذكرها أعلاه.

N	Mean	Standard Deviation SD	Standard Error of the Mean
30	54.63	10.327	1.886

One Sample t test

t	df	sig	Mean Difference
2.457	29	.02	4.633

$$d=\frac{\overline{X}}{SD}=\frac{4.633}{10.327}=.448$$

وحجم الأثر المعدل g يمكن حسابه من المعادلتين السابقتين أرقام 22، 23 كما يلي

$$J=1-\frac{3}{4df-1}=1-\frac{3}{4\times 29-1}=0.9739$$

$$g=J\times d=0.9739\times 0.448=0.436$$

ويمكن للباحث حساب قيمة d مباشرة باستخدام قيمة t من الجدول المذكور أعلاه كالآتي:

$$d=\frac{t}{\sqrt{N}}=\frac{2.475}{\sqrt{29}}=0.456$$

كما يمكن للباحث أيضاً حساب حجم الأثر باستخدام المعادلة التالية والتي تُعطي نسبة التباين المفسر.

$$\eta^2=\frac{t^2}{t^2+df}=\frac{2.457^2}{2.456^2+29}=.17$$

ملحوظة مهمة يعطي برنامج SPSS النسخة 28 قيمة حجم الأثر d وقيمة حجم الأثر المعدل g دون الحاجة لإجراء الحسابات السابقة لقيم d، g.

2. حساب قيمة حجم الأثر باستخدام مؤشر كوهين (d) من اختبار (t) للعينات المعتمدة

يستخدم إختبار النسبة التائية للعينات المعتمدة لمقارنة متوسطي مجموعتين متزاوجتين أو للفرق بين متوسطي إختبارين أحدهماقبلي والآخر بعدي. وتعطي قيمة النسبة التائية في هذه الحالة من المعادلة التالية:

$$t=\frac{\overline{d}}{s_{\overline{d}}}$$

حيث ($\overline{d}$) هي الفرق بين المتوسطين، ($S_{\overline{d}}$) هي الخطأ المعياري بين المتوسطين ويعطي بالمعادلة الأتية:

$$s_{\overline{d}} = \frac{s_d}{\sqrt{n}}$$

ويمكن حساب حجم الأثر d لكوهين في هذه الحالة من العلاقة

$$d = \frac{\overline{x}}{SD}$$

أو من العلاقة

$$d = \frac{t}{\sqrt{N}}$$

كما يمكن حساب حجم الأثر بإيجاد نسبة التباين الفسر η^2 من العلاقة التالية في حالة وجود عدد أفراد العينة N في تقرير البحث.

$$\eta^2 = \frac{(N)Mean^2}{(N)Mean^2 + (N-1)SD^2}$$

كما يمكن حساب قيمة η^2 من المعادلة التالية بمعلومية قيمة t

$$\eta^2 = \frac{t^2}{t^2 + N - 1}$$

مثال:

أراد باحث تقييم فاعلية برنامج تدريبي على مهارات الحاسوب لثلاثين معلماً من معلمي العلوم بإحدى المدارس. قام الباحث بتطبيق بطاقة ملاحظة على هؤلاء المعلمين مرتين مرة قبل تطبيق البرنامج التدريبي ومرة بعد تطبيقه. حلل الباحث النتائج التي حصل عليها باستخدام برنامج SPSS فحصل على النتائج المبينة بالجدولين الآتيين، من هذين الجدولين أحسب قيمة مؤشر كوهين d من النتائج التالية :

Teast	Mean	N	SD
Pretest	4.5	30	1.843
Posttest	5.76	30	1.493

t paired test

Mean Difference	SD	t	df	sig
1.167	2.260	2.827	29	.008

$$d = \frac{\bar{x}}{SD} = \frac{1.167}{2.260} = .516$$

يمكن أيضا باستخدام المعادلتين 22، 23 حساب قيمة حجم الأثر المعدل g
كما يمكن حساب قيمة d باستخدام قيمة t من المعادلة التالية:

$$d = \frac{t}{\sqrt{N}} = \frac{2.827}{\sqrt{30}} = \frac{2.827}{5.477} = .516$$

وإذا أراد الباحث حساب نسبة التباين المفسر فيمكنه استخدام إحدى المعادلتين التاليتين

$$\eta^2 = \frac{(N)Mean^2}{(N)Mean^2 + (N-1)SD^2} = \frac{30(1.167)^2}{30(1.167)^2 + (30-1)2.26^2} = .216$$

$$\eta^2 = \frac{t^2}{t^2 + N - 1} = \frac{2.827^2}{2.827^2 + (30-1)} = \frac{7.991}{7.991 + 29} = .216$$

ملحوظة مهمة: يعطي برنامج SPSS النسخة 28 قيمة حجم الأثر d وقيمة حجم الأثر المعدل g دون الحاجة لإجراء الحسابات السابقة لقيم d، g.

3. حساب قيمة حجم الأثر باستخدام مؤشر كوهين (d) من إختبار (t) للعينات المستقلة

يستخدم إختبار النسبة التائية في هذه الحالة لمقارنة متوسطي مجموعتين مستقلتين، والمعادلة المستخدمة في حالة تساوي العينتين هي:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

أما إذا كان حجما العينتين غير متساويين فتُستخدم المعادلة التالية:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_{pooled}^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

ويُحسب حجم الأثر للمجموعات المستقلة مباشرة من المتوسطين والإنحراف المعياري المشترك بين هذين المتوسطين باستخدام المعادلة التالية.

$$d = \frac{\bar{X}_t - \bar{X}_c}{S_{pooled}}$$

حيث d هي معامل أو مؤشر كوهين لحجم الأثر

$\overline{X}_t$ هي متوسط درجات المجموعة التجريبية

$\overline{X}_c$ هي متوسط المجموعة الضابطة

S_{pooled} هي الإنحراف المعياري المشترك

ولحساب الإنحراف المعياري المشترك في مقام المعادلة السابقة نستخدم المعادلة التالية:

$$s_{pooled} = \sqrt{\frac{s_t^2(n_t - 1) + s_c^2(n_c - 1)}{n_t + n_c - 2}}$$

كما يمكن حساب قيمة (d) من المعادلة التالية باستخدام قيمة النسبة التائية t

$$d = t\sqrt{\frac{n_t + n_c}{n_t n_c}}$$

وأيضاً يمكن حساب قيمة حجم الأثر من نسبة التباين المفسر η^2 من المعادلة التالية

$$\eta^2 = \frac{t^2}{t^2 + (n_t + n_c - 2)}$$

مثال:

قام أحد الباحثين بالبحث عن فاعلية برنامج تدريبي في تدريس مهارات استخدام الحاسب الآلي في إجراء التجارب العملية في العلوم. إختار الباحث عينة عشوائية مكونة من 30 طالبا من طلاب الصف الثالث الإعدادي قسمهم إلى مجموعتين إحداهما تجريبية درست البرنامج التدريبي والأخرى ضابطة درست نفس المهارات لكن عملياً في المختبر دون استخدام الحاسب الآلي. حلل الباحث نتائجه التي جاءت كما في الجدولين التاليين. احسب حجم الأثر الناتج باستخدام نتيجة إختبار t مرة ونسبة التباين المفسر مرة أخرى.

	n	Mean	SD	Standard Error
Experimental Group	15	45.20	24.969	6.447
Control Group	15	22.07	27.136	7. 006

t test for independent Groups

t	*df*	*Sig*	Mean Difference	Standard Error
2.43	28	.022	23.133	9.521

لحساب قيمة مؤشر حجم الأثر d لابد أولا من حساب قيمة الإنحراف المعياري المشترك S_{pooled} من المعادلة

$$s_{pooled} = \sqrt{\frac{s_1^2(n_t-1)+s_2^2(n_c-1)}{n_t-n_c-2}} = \sqrt{\frac{(15-1)(24.2)^2+(15-1)(27.136)^2}{15+15-2}} = 25.709$$

ثم بعد ذلك يتم التعويض في معادلة حجم الأثر d كما يلي:

$$d = \frac{\overline{x}_t - \overline{x}_c}{s_{pooled}} = \frac{45.20-22.o7}{25.709} = .89$$

كما يمكن حساب قيمة d بمعلومية t وعدد الأفراد n في كل مجموعة كما يلي:

$$d = t\sqrt{\frac{n_t+n_c}{n_t n_C}} = 2.34\sqrt{\frac{30}{225}} = .886$$

أيضاً يمكن حساب قيمة حجم الأثر بإيجاد نسبة التباين المفسّر η^2 على النحو التالي :

$$\eta^2 = \frac{t^2}{t^2+(n_t+n_C-2)} = \frac{2.43^2}{2.43^2+28} = .174$$

ثانيا: حساب قيمة حجم الأثر من تحليل التباين

يُعد تحليل التباين أحد عناصر النموذج الخطي العام (General Linear Model) . وهذا التحليل يتكون من عدد من الأساليب الإحصائية التي تُستخدم فيها متغيرات تصنيفية تتعلق بالعوامل المستقلة ومتغيرات فترية تتعلق بالعوامل التابعة. ونظراً لتعدد هذه الأساليب الإحصائية سنعرض في هذا الجزء لكيفية حساب حجم الأثر من نتائج خمسة أنواع من تحليل التباين وهي:

1. تحليل التباين الأحادي (One Way Analysis of Variance)
2. تحليل التباين الثنائي (One Way Analysis of Variance)
3. تحليل التتباين المتلازم الأحادى (One Way Analysis of Covariance)
4. تحليل التباين الأحادي مع تكرار القياس (Repeated Measure Onw Way Analysis of Variance)
5. تحليل التباين المتعدد في إتجاه واحد (Multivariate Analysis of Variance)

1. حساب حجم الأثر من نتائج تحليل التباين الأحادي

يستخدم تحليل التباين الأحادي لمقارنة متوسطات مجموعتين أو أكثر من مجموعتين عندما يكون لدى الباحث متغير مستقل له مستووين أو أكثر من مستويين ومتغير تابع واحد. وهذا التحليل يعطي نتائج أكثر صدقاً عندما يكون عدد الأفراد في المجموعات متساوي، وأن التباينات بين المجموعات متجانسة، وأن درجات العامل التابع موزعة توزيعاً إعتدالياً في كل مجموعة.

ويعطي برنامج SPSS نسبة التباين المفسر η^2 الجزئية (Partial η^2) ويمكن حساب هذه القيمة من المعادلة التالية:

$$\eta^2{}_{partial}\ \frac{SS_{Treatment}}{SS_{Traetment}+SS_{Error}}$$

حيث $SS_{Treatment}$ هي مجموع مربعات إنحرافات متوسط كل معالجة (طريقة تدريس كما في المثال الأتي) عن المتوسط العام لكل المجموعات مضروبا في عدد أفراد كل مجموعة كما في جدول تحليل التباين أما SS_{Error} فهي مجموع مربعات الفرق بين درجات كل فرد في المجموعة ومتوسط الدرجات في هذه المجموعة.

وتتراوح قيمة (η^2) بين الصفر والواحد الصحيح. ويرى كوهين إلى أن قيمة (η^2) التي تساوي 0.01 تشير إلى حجم أثر ضعيف، وتشير قيمة (η^2) التي تساوي 0.06 إلى حجم أثر متوسط وتشير قيمة (η^2) التي تساوي 0.14 إلى حجم أثر مرتفع (أنظر الجزء الخاص بتفسيرات حجوم الأثر)

ولما كانت قيمة η^2 ليست دقيقة تماماً فقد استخدم الباحثون معادلة ω^2 التي تعطي قيمة أدق لنسبة التباين المفسر (Mayer & Well, 2003, p. 230)، وهذه القيمة يمكن إيجادها من المعادلة التالية:

$$\omega^2=\frac{(a-1)(F_a-1)}{(a-1)(F_a-1)+na}$$

حيث a تشير إلى عدد المعالجات ، *n* هي عدد أفراد العينة في كل معالجة ، أما F_A فهي قيمة النسبة الفائية للمعالجات كما ترد في جدول تحليل التباين.

كما يمكن حساب قيمة حجم الأثر من المعادلة التالية التي تعطي حجم الأثر لكوهين والذي يرمز له في هذه الحالة بالحرف $\hat{f}$ من إحدى المعادلتين التاليتين:

$$\hat{f}=\sqrt{\frac{(a-1)(MS_{treatment}-MS_{Error})}{anMS_{error}}}$$

أو من المعادلة المبسطة الأتية (Mayer & Well, 2003, p. 232)

$$\hat{f}=\sqrt{\frac{(a-1)(F_a-1)}{an}}$$

كما يمكن إجراء التحويل بين $\hat{f}$ و ω^2 من المعادلتين التاليتين:

$$\omega^2 = \frac{\hat{f}^2}{\left(1+\hat{f}^2\right)}$$

$$\hat{f}^2 = \frac{\omega^2}{\left(1-\omega^2\right)}$$

مثال

أراد باحث التعرف على أكثر الطرق فاعلية في تدريس الجمع لتلاميذ الصف الثاني الإبتدائي. لذلك قام الباحث بإختيار عينة عشوائية من هؤلاء التلاميذ عددها 15 تلميذاً، وزّعهم عشوائيا على ثلاث مجموعات. تلقت المجموعة الأولى دروس الجمع بالطريقة العادية، وتلقت المجموعة الثانية نفس الدروس باستخدام المعداد اليدوي، أما المجموعة الثالثة فتلقت نفس الدروس باستخدام برنامج حاسوبي في الجمع. قام الباحث بتحليل النتائج التي توصل إليها باستخدام برنامج الإحصاء SPSS فحصل على الجداول الأتية. من هذه الجداول إحسب قيم حجم الأثر (η^2 ، ω^2) أيضا إحسب قيمة حجم الأثر $\hat{f}$ لكوهين.

جدول يبين درجات التلاميذ في كل مجموعة (ملحوظة تم استخدام 5 أفراد في كل مجموعة ليسهل على من يريد حساب تحليل التباين يدويا أن يقوم بذلك).

مجموعة (1)	مجموعة (2)	مجموعة (3)
0	6	6
4	8	5
0	5	9
1	4	4
0	2	6

جدول يبين الإحصاء الوصفي للبيانات

Descriptive Statistics

Dependent Variable: Acievement

Group	Mean	Std. Deviation	N
Group1	1.0000	1.73205	5
Group2	5.0000	2.23607	5
Group3	6.0000	1.87083	5
Total	4.0000	2.87849	15

جدول يبين نتائج تحليل التباين الأحادي

Tests of Between-Subjects Effects

Dependent Variable: Acievement

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	70.000[a]	2	35.000	9.130	.004	.603
Intercept	240.000	1	240.000	62.609	.000	.839
Group	70.000	2	35.000	9.130	.004	.603
Error	46.000	12	3.833			
Total	356.000	15				
Corrected Total	116.000	14				

a. R Squared = .603 (Adjusted R Squared = .537)

حساب قيمة η^2

$$\eta^2_{\ partial}\ \frac{SS_{Treatment}}{SS_{Traetment}+SS_{Error}}=\frac{70}{70+46}=0.603$$

حساب قيمة ω^2

$$\omega^2=\frac{(a-1)(F_a-1)}{(a-1)(F_a-1)+na}=\frac{(3-1)(9.130-1)}{(3-1)(9.130-1)5\times 3}=0.530$$

حساب قيمة حجم الأثر $\hat{f}$ لكوهين يمكن أن يتم بإحدى المعادلتين التاليتين

$$\hat{f}=\sqrt{\frac{(a-1)(MS_{treatment}-MS_{Error})}{anMS_{error}}}=\sqrt{\frac{(3-1)(35-3.833}{3\times 5\times 3.833}}=1.04$$

أو من المعادلة المبسطة التالية

$$\hat{f}=\sqrt{\frac{(a-1)(F_a-1)}{na}}=\sqrt{\frac{(3-1)(9.130-1)}{5\times 3}}=1.04$$

ويمكن تحويل $\hat{f}^2$ إلى ω^2 كما يلي:

$$\omega^2=\frac{\hat{f}^2}{\left(1+\hat{f}^2\right)}\frac{1.04^2}{\left(1+1.04^2\right)}=0.52$$

لاحظ أن قيم η^2 ، ω^2 تعطينا نسبة التباين المفسر أما $\hat{f}$ فتعطينا قيمة حجم الأثر لكوهين

لاحظ أيضا أن برنامج SPSS يعطي قيمة η^2 مباشرة في جدول تحليل التباين.

حساب حجوم الأثر من التصميمات العاملية في تحليل التباين:

التصميمات العاملية (Factorial Designs) هي تلك التي تحتوي على أكثر من عامل مستقل وعامل تابع واحد. فعلى سبيل المثال فقد يحتوي التصميم العاملي على عاملين مستقلين لكل عامل منهما مستويان فيشار إليه بأنه تصميم عاملي (2 × 2)، أو قد يحتوي أحد العاملين على ثلاث مستويات والعامل الآخر على مستويين فيشار إليه بأنه تصميم عاملي (3 × 2) أوقد يحتوي التصميم على ثلاثة عوامل مستقلة يحتوي كل عامل منها على مستويين فيشار إلي بأنه تصميم عاملي (2 × 2 × 2) وهكذا.

وتتميز التصميمات العاملية عن التصميم الأحادي بأنها توفر جهد ووقت الباحث، كما تمكنه من دراسة أثر التفاعل بين العوامل أو المتغيرات. ونظراً لتعقد التصميمات العاملية سنكتفي هنا بعرض كيفية حساب حجوم الأثر من تحليل التباين ثنائي الإتجاه (Tow Way Analysis of Variance). كما سنعرض أيضا لتحليل التباين مع تكرار القياس، وتحليل التباين المتلازم (أو ما يسمى أحياناً بتحليل التغاير)

2. **حساب حجم الأثر من تحليل التباين الثنائي**

يعطي برنامج SPSS قيمة η^2 مباشرة من جدول تحليل التباين لكل من المتغير المستقل الأول والمتغير المستقل الثاني والتفاعل بينهما. ويمكن للباحث أن يستخدم المعادلات التالية لحساب هذه القيم يدوياً من المعادلات التالية:

بالنسبة للعامل الأول تعطى قيمة η^2 من المعادلة التالية

$$\eta^2_{Factor1} = \frac{SS_{Factor1}}{SS_{Factor1} + SS_{Error}}$$

والمعادلة نفسها تستخدم أيضا مع العامل الثاني على النحو التالي:

$$\eta^2_{Factor2} = \frac{SS_{Factor2}}{SS_{Factor\ 2} + SS_{Error}}$$

وبالنسبة للتفاعل بين العاملين الأول والثاني فيعطي بالمعادلة التالية

$$\eta^2_{Main-Interaction} = \frac{SS_{Main-Interaction}}{SS_{Main-Interaction} + SS_{Error}}$$

كما يمكن حساب حجم الأثر $\hat{f}$ لكوهين من المعادلات التالية:

$$\hat{f}_A = \sqrt{\frac{(a-1)(F_A-1)}{abn}}$$

$$\hat{f}_B = \sqrt{\frac{(b-1)(F_B-1)}{abn}}$$

$$\hat{f}_{AB} = \sqrt{\frac{(a-1)(b-1)(F_{AB}-1)}{abn}}$$

حيث a تمثل العامل الأول، b تمثل العامل الثاني، ab تمثل تفاعل هذين العاملين.

أما قيم ω^2 للعامل الأول والثاني والتفاعل بينهما فيمكن حسابهم من العادلات التالية:

$$\omega_a^2 = \frac{(a-1)(F_A-1)}{(a-1)(F_A-1)+abn}$$

$$\omega_b^2 = \frac{(a-1)(F_B-1)}{(a-1)(F_B-1)+abn}$$

$$\omega_{ab}^2 = \frac{(a-1)(F_{AB}-1)}{(a-1)(F_{AB}-1)+abn}$$

كما يمكن تحويل ω^2 إلى $\hat{f}^2$ والعكس من المعادلات التالية:

$$\omega_A^2 = \frac{\hat{f}_A^2}{\left(1+\hat{f}_A^2\right)}$$

$$\omega_B^2 = \frac{\hat{f}_B^2}{\left(1+\hat{f}_B^2\right)}$$

$$\omega_{AB}^2 = \frac{\hat{f}_{AB}^2}{\left(1+\hat{f}_{AB}^2\right)}$$

مثال

استخدم باحث 3 استراتيجيات لتدريس المفاهيم لعينة من طلاب وطالبات المرحلة الإعدادية بهدف تحديد فاعلية كل من استراتيجيات التدريس والجنس والتفاعل بينهما على تحصيل أفراد العينة. لذا إختار الباحث 24 طالبا و24 طالبة لبحثه بطريقة عشوائية ثم وزع الباحث الطلاب والطالبات عشوائيا في 6 مجموعات. المجموعة الأولى تكونت من 8 طالبات درسن المفاهيم باستراتيجية تقليدية، والمجموعة الثانية تكونت من 8 طلاب درسوا المفاهيم باستراتيجية تقليدية أيضا، المجموعة الثالثة تكونت من 8 طالبات درسن المفاهيم من خلال استراتيجية التعلم بالإستقصاء، أما المجموعة الرابعة فتكونت من 8 طلاب درسوا المفاهيم أيضا باستراتيجية التعلم بالإستقصاء، المجموعة الخامسة تكونت من 8 طالبات درسن المفاهيم من خلال استراتيجية التعلم التعاوني، أما المجموعة السادسة المكونة من 8 طلاب درسوا باستراتيجية التعلم التعاوني أيضا.

جدول يبين درجات الطلاب والطالبات في طرق التدريس الثلاث

استراتيجية التدريس	تقليدية		استقصاء		تعاونية	
جنس الطالب	ذكور	إناث	ذكور	إناث	ذكور	إناث
	30	55	45	67	50	65
	30	65	60	65	55	70
	30	70	85	60	80	60
	55	55	65	70	65	60
	35	55	70	65	70	60
	20	60	70	60	75	55
	45	50	80	60	75	60
	40	50	60	50	65	55

قام الباحث بتحليل الدرجات التي حصل عليها باستخدام برنامج الإحصاء SPSS وحصل على النتائج التالية:

Tests of Between-Subjects Effects

Dependent Variable: Achievement

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	5479.167[a]	5	1095.833	13.197	.000	.611
Intercept	163333.333	1	163333.333	1967.025	.000	.979
Gender	168.750	1	168.750	2.032	.161	.046
Strategy	3332.292	2	1666.146	20.065	.000	.489
Gender * Strategy	1978.125	2	989.063	11.911	.000	.362
Error	3487.500	42	83.036			
Total	172300.000	48				
Corrected Total	8966.667	47				

a. R Squared = .611 (Adjusted R Squared = .565)

من جدول تحليل التباين الثنائي أوجد قيم η^2 لكل من المتغير المستقل الأول (Gender) والمتغير المستقل الثاني (Strategy) والتفاعل بينهما. ثم جد قيمة كل من $\hat{f}$ ، ω^2

حساب قيمة η^2 لعامل الجنس

$$\eta^2_{gender} = \frac{SS_{gender}}{SS_{gender} + SS_{Error}} = \frac{168.750}{168.750 + 3487.500} = 0.0461$$

حساب قيمة η^2 لعامل الإستراتيجية

$$\eta^2_{Strategy} = \frac{SS_{Strategy}}{SS_{Strategy} + SS_{Error}} = \frac{3332.292}{3332.292 + 3487.50} = .397$$

حساب قيمة η^2 لعامل للتفاعل بين الجنس والإستراتيجية

$$\eta^2 = \frac{SS_{Interaction}}{SS_{Interaction} + SS_{Error}} = \frac{1978.125}{1978.125 + 3487.50} = .098$$

قارن بين القيم المحسوبة للمؤشر η^2 والقيم الجدولية لتحليل التباين.

حساب قيمة ω^2 ، $\hat{f}$ من بيانات جدول تحليل التباين

$$\omega^2_{gender} = \frac{(a-1)(F_A-1)}{(a-1)(F_A-1)+abn} = \frac{(2-1)(2.032-1)}{(2-1)(2.032-1)+48} = 0.02$$

$$\hat{f}_{gender} = \sqrt{\frac{(a-1)(F_A-1)}{abn}} = \sqrt{\frac{(2-1)(2,032-1)}{48}} = 0.15$$

$$\omega^2_{strategy} = \frac{(b-1)(F_B-1)}{abn} = \frac{(3-1)(20.065-1)}{(3-1)(20.065-1)+48} = 0.44$$

$$\hat{f}_{strategy} = \sqrt{\frac{(b-1)(F_B-1)}{abn}} = \sqrt{\frac{(3-1)(20.065-1)}{48}} = 0.8912$$

$$\omega^2_{ab} = \frac{(a-1)(b-1)(F_{AB}-1)}{(a-1)(b-1)(F_{AB}-1)+abn} = \frac{1\times 2(11.911-1)}{1\times 2(11.911-1)+48} = 0.313$$

$$\hat{f}_{ab} = \sqrt{\frac{(a-1)(b-1)(F_{AB}-1)}{abn}} = \sqrt{\frac{1\times 2(11.911-1)}{48}} = 0.658$$

3. **حساب قيمة حجم الأثر لتحليل التباين المتلازم (المصاحب) الأحادي**

يستخدم تحليل التباين المتلازم بهدف إجراء تعديلات على البيانات التجريبية في ضوء الفروق بين أفراد العينة قبل بداية التجربة في متغير أو أكثر يسمى بالمتغير المصاحب، وتستخدم درجات هذا المتغير في تعديل الفروق بين مجموعات المعالجة التجريبية وفي تقليل خطأ التباين. وتتوقف هذه التعديلات على مقدار الإرتباط بين المتغيرين المصاحب والتابع (علام، 1993أ، ص. 338)

ويعطي برنامج SPSS قيمة η^2 لكل من المتغير المستقل والمتغير المتلازم (Covariance) ويمكن حسابهما من المعادلة التالية:

$$\eta^2 = \frac{SS_{factor-covariance}}{SS_{factor-covariance} + SS_{error}}$$

حيث $SS_{factor-covariance}$ تشير إلى مجموع المربعات للعامل المستقل أو للعامل المصاحب أما SS_{error} فتشير إلى مجموع مربعات الخطأ في جدول تحليل التباين.

مثال:

إختار أحد المعلمين عينة عشوائية قدرها 12 طالبا ممن يدرسون مقرراً في الإحصاء التربوي. وزع المعلم هؤلاء الطلاب إلى مجموعتين كل مجموعة تتكون من 6 طلاب. درست المجموعة الأولي بطريقة المحاضرة ودرست المجموعة الثانية بطريقة التعلم التعاوني. طبق المعلم على طلابه إختبارين أحدهما إختبار للإستعداد في الرياضيات واستخدمه كمتغير متلازم (مصاحب) أما الإختبار الثاني فكان إختباراً للتحصيل في مقرر الإحصاء استخدمه كمتغير تابع. حلل المعلم النتائج التي حصل عليها باستخدام تحليل التباين المتلازم في إتجاه واحد وحصل على النتائج التالية:

جدول درجات الطلاب في إختباري الإستعداد والتحصيل

الطريقة	المحاضرة		التعلم التعاوني	
الإختبار	الإستعداد	التحصيل	الإستعداد	التحصيل
	4	1	1	1
	3	2	3	2
	5	3	2	4
	6	4	4	5
	7	5	5	6
	9	6	7	6

جدول تحليل التباين المتلازم من برنامج SPSS

Tests of Between-Subjects Effects

Dependent Variable: Quiz score

Source	Type I Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	31.693[a]	2	15.846	16.667	<.001	.787
Intercept	168.750	1	168.750	177.483	<.001	.952
Aptitude	20.881	1	20.881	21.961	.001	.709
Group	10.812	1	10.812	11.372	.008	.558
Error	8.557	9	.951			
Total	209.000	12				
Corrected Total	40.250	11				

a. R Squared = .787 (Adjusted R Squared = .740)

جدول المتوسطات المعدلة من برنامج SPSS

Group

Dependent Variable: Quiz score

Group	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Traditional lecture method of instruction	2.686[a]	.423	1.729	3.642
Small group and self-directed instruction	4.814[a]	.423	3.858	5.771

a. Covariates appearing in the model are evaluated at the following values: Aptitude = 4.6667.

ملحوظة: الدرجات الخام مأخوذة من كتاب (Lomax & Hash-Vaughn, 2012) الفصل الرابع

من جدول تحليل التباين المتلازم أوجد قيمة η^2 للمتغير المستقل (طريقة التدريس) من المعادلة التالية:

$$\eta^2 = \frac{SS_{group}}{SS_{group} + SS_{error}} = \frac{10.812}{10.812 + 8.557} = 0.5582$$

قارن القيمة التي تم حسابها من المعادلة السابقة بقيمة η^2 الموجودة بجدول تحليل التباين المتلازم.

كما يمكن أيضا حساب قيمة حجم الأثر d في حالة استخدام تحليل التباين المتلازم الأحادي من المعادلة الأتية:

$$d = \sqrt{\frac{F\left(n_1 + n_2\right)}{n_1\, n_2}} \times \sqrt{\left(1 - R^2\right)}$$

$$d = \sqrt{\frac{11.372 \times 12}{6 \times 6}} \times \sqrt{\left(1 - 0.787\ \right)} = 0.898$$

حيث R^2 هي معامل إرتباط المتغير المتلازم أو معامل الإرتباط المتعدد إذا تعددت المتغيرات المتلازمة (هذه القيمة موجودة أسفل جدول تحليل التباين المتلازم باللون الأحمر).

كما يمكن حساب $\hat{f}$ من المعادلة الأتية

$$\hat{f} = \sqrt{\frac{(a-1)(F_A - 1)}{na}} = \sqrt{\frac{(2-1)(11.372)}{12}} = 0.92$$

4. **حساب قية حجم الأثر من تحليل التباين الأحادي مع تكرار القياس**

في هذا النوع من التحليل الإحصائي يتعرض كل فرد من أفراد العينة للقياس في أحد المتغيرات أكثر من مرة. فعلى سبيل المثال قد يجتاز مجموعة من الطلاب 3 إختبارات (قبلي وبعدي ومرجأ) ومن ثم يتم البحث عن دلالات الفروق بين متوسطات الإختبارات الثلاثة. وبعد الحصول على هذه النتائج قد يرغب الباحث في إيجاد قيمة حجم الأثر نتيجة تكرار القياس. وحجم الأثر قد يعبر عنه في هذه الحالة بمؤشر η^2 .

وإيجاد قيمة حجم الأثر η^2 في هذه الحالة يتم باستخدام نتائج تحليل التباين الأحادي مع تكرار القياس وتظهر هذه القيمة في جدول تحليل التباين عندما يطلب القائم بالتحليل هذه القيمة من البرنامج. ويمكن أيضا حساب هذه القيمة من المعادلة الأتية.

$$\eta^2 = \frac{SS_{factor}}{SS_{factor} + SS_{Error}}$$

مثال

عُرضت أوراق إجابات 8 طلاب في موضوع للتعبير على 4 مصححين. طُلب من كل مصحح إعطاء درجة من 10 لكل ورقة إجابة من الإجابات الثمانية. وجاءت تقديرات المصححين كما يلي. إحسب قيمة η^2 لهذه التقديرات.

	Rater1_raw	Rater2_raw	Rater3_raw	Rater4_raw
1	3.00	4.00	7.00	8.00
2	6.00	5.00	8.00	9.00
3	3.00	4.00	7.00	9.00
4	3.00	4.00	6.00	8.00
5	1.00	2.00	5.00	10.00
6	2.00	3.00	6.00	10.00
7	2.00	4.00	5.00	9.00
8	2.00	3.00	6.00	10.00

باستخدام برنامج SPSS نحصل على متوسطات تقدير كل من المصححين الأربعة وقيمة النسبة الفائية ومستوى الدلالة الإحصائية وقيمة حجم الأثر η^2 .

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
factor1	Sphericity Assumed	198.125	3	66.042	73.477	<.001	.913
	Greenhouse-Geisser	198.125	1.428	138.760	73.477	<.001	.913
	Huynh-Feldt	198.125	1.691	117.163	73.477	<.001	.913
	Lower-bound	198.125	1.000	198.125	73.477	<.001	.913
Error(factor1)	Sphericity Assumed	18.875	21	.899			
	Greenhouse-Geisser	18.875	9.995	1.888			
	Huynh-Feldt	18.875	11.837	1.595			
	Lower-bound	18.875	7.000	2.696			

من الجدول أعلاه يتضح أن قيمة η^2 تساوي 0.913 (موضحة باللون الأحمر في الجدول) كما يمكن حساب هذه القيمة يدويا كما يلي:

$$\eta^2 = \frac{SS_{factor}}{SS_{factor} + SS_{Error}} = \frac{198.125}{198.125 + 18.875} = .913$$

وهذه القيمة تشير إلى أن 91% تقريبا من التباين الناتج يفسر الإختلاف بين هؤلاء المقدرين.

أما ω^2 فيمكن حسابها من العلاقة الأتية

$$\omega^2 = \frac{(a-1)(F_A-1)}{(a-1)(F_A-1)+an}$$

$$\omega^2 = \frac{(4-1)(73.477-1)}{(4-1)(73.477-1)+4\times 8} = 0.872$$

ويمكن حساب $\hat{f}$ من العلاقة الأتية:

$$\hat{f} = \sqrt{\frac{(a-1)(F_A-1)}{an}}$$

$$\hat{f} = \sqrt{\frac{(4-1)(73.477-1)}{4\times 8}} = 2.61$$

وكما سبق القول يمكن تحويل $\hat{f}^2$ إلى ω^2 أو ω^2 إلى $\hat{f}^2$ من المعادلة الأتية

$$\omega^2 = \frac{\hat{f}^2}{\left(1 - \hat{f}^2\right)}$$

5. **<u>حساب قيمة حجم الأثر لتحليل التباين عديد المتغيرات التابعة في متغير مستقل واحد</u>**

في الأمثلة السابقة كنا نتعامل مع أثر متغير مستقل واحد أو أكثر في متغير تابع واحد (مثال ذلك استخدام إختبار النسبة التائية لمجموعتين مستقلتين أو استخدام تحليل التباين الأحادي أو الثنائي الإتجاه). أما في تحليل التباين عديد المتغيرات التابعة فنحن نتعامل مع أثر متغير مستقل أو أكثر في متغيرين تابعين أو أكثر وبهذا الأسلوب يمكن لنا أن نتحكم في خطأ النوع الأول وهو رفض فرض صفري صحيح. وبذلك يمكن إعتبار تحليل التباين عديد المتغيرات التابعة تعميماً لكل من إختبارات النسبة التائية وتحليل التباين.

وكمثال لهذا النوع من التحليل دعنا نفترض أن باحثاً أراد دراسة أثر ثلاث استراتيجيات لتعلم العلوم في تذكر تلاميذ الصف الخامس الإبتدائي للمفاهيم العلمية وتطبيقهم لها في هذا المثال لدينا متغير مستقل واحد وهو نوع الإستراتيجية ومتغيرين تابعين هما التذكر والتطبيق. والبيانات الخام لهذه التجربة مأخوذة من كتاب (Green & Salkind, 2014).

جدول الدرجات الخام للمثال السابق

المجموعة	المجموعة الأولى		المجموعة الثانية		المجموعة الثالثة	
مستوى التحصيل	التذكر	التطبيق	التذكر	التطبيق	التذكر	التطبيق
	3	1	6	7	4	3
	4	4	7	4	6	6
	4	3	5	6	4	4
	4	5	6	3	4	3
	2	3	5	2	5	6
	3	2	5	7	5	5
	4	3	5	4	5	5
	3	3	5	5	2	4
	3	5	8	7	3	3
	3	3	6	5	4	5

يعطي برنامج SPSS أكثر من قيمة لـ η^2 ولكن الأكثر استخداما هي قيمة η^2 المناظرة للإحصائي (Λ Wilk's Lambda) كما في الجدول التالي.

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Intercept	Pillai's Trace	.962	326.035[b]	2.000	26.000	<.001	.962
	Wilks' Lambda	.038	326.035[b]	2.000	26.000	<.001	.962
	Hotelling's Trace	25.080	326.035[b]	2.000	26.000	<.001	.962
	Roy's Largest Root	25.080	326.035[b]	2.000	26.000	<.001	.962
group	Pillai's Trace	.602	5.811	4.000	54.000	<.001	.301
	Wilks' Lambda	.421	7.028[b]	4.000	52.000	<.001	.351
	Hotelling's Trace	1.318	8.240	4.000	50.000	<.001	.397
	Roy's Largest Root	1.275	17.215[c]	2.000	27.000	<.001	.560

a. Design: Intercept + group

b. Exact statistic

c. The statistic is an upper bound on F that yields a lower bound on the significance level.

*لاحظ اننا إقتصرنا على الجزء المطلوب لإيجاد قيمة η^2 أما بقية نتائج التحليل فلم ندرجها هنا.

من الجدول يتضح لنا أن قيمة F (وقيمتها تساوي 7.028 كما هي موضحة باللون الأحمر دالة إحصائيا) مما يعني أن هناك فروقا بين متوسطات التذكر والتطبيق عند التعرض للطرق الثلاثة وأن قيمة η^2 تساوي 0.351

ويمكن حساب قيمة η^2 باستخدام قيمة (Wilk's Lambda Λ) من المعادلة التالية:

$$\eta^2 = 1 - \Lambda^{\frac{1}{s}}$$

حيث s هي عدد مستويات المتغير المستقل مطروحا منها 1 . ومن المثال الذي معنا نجد أن Λ تساوي 0.421 وأن (s= 3-1=2) ومن ثم فإن

$$\eta^2 = 1 - \Lambda^{\frac{1}{s}} = 1 - .421^{\frac{1}{2}} = 1 - .6448 = .351$$

أنظر قيمة Λ في الجدول باللون الأخضر

ثالثا: حساب مؤشرات قوة العلاقة أو الإرتباطات بين المتغيرات

أشرنا في الصفحات السابقة لعدد من مؤشرات قوة العلاقة بين المتغيرات، مثل d، f، η^2 كما أشرنا كذلك لمؤشر أخر وهو ω^2 والمؤشران الأخيران يشيران إلى حجم الأثر من خلال قيمة التباين المفسر وفيما يلي سنعرض لعدد أخر من مؤشرات حساب قوة العلاقة بين المتغيرات مثل:

1. معامل إرتباط بيرسون
2. معامل إرتباط سبيرمان للرتب
3. معامل الإرتباط لكندال (τ)
4. معامل الإرتباط الثنائي والثنائي المتعدد

1. **حساب حجم الأثر من معامل ارتباط بيرسون**

يمكن للباحث الذي يستخدم معامل الإرتباط لبيرسون (Pearson *r*) بين متغيرين درجاتهما مستمرة (Continuous Variables) أن يستخدم هذا المعامل لبيان قوة العلاقة بين هذين المتغيرين. فإذا أعطت النتائج مثلا معامل إرتباط بيرسون (*r*) فإن الباحث يمكنه الحكم على قوة الإرتباط أو حجم الأثر بمقارنة قيمة معامل الإرتباط بالقيم التي اشار إليها كوهين والتي تحدد ما إذا كان حجم الأثر ضعيفاً أو متوسطاً أو كبيراً وهذه القيم جاءت على النحو التالي:

إذا كانت قيمة (r) = 0.10 فهذا يعني أن حجم الأثر ضعيف، وفي هذه الحالة فإنه يُفسر 01. من التباين الكلي. وإذا كانت قيمة (r) = 0.30 فهذا يعني أن حجم الأثر متوسط ، وفي هذه الحالة فإنه يُفسر9% من التباين الكلي. أما إذا كانت قيمة (r) = 0.50 فهذا يعني أن حجم الأثر كبير، وفي هذه الحالة فإنه يُفسر 25% من التباين الكلي.

ومن المهم أن نلاحظ أن قيمة حجم الأثر والتي تساوي 0.4 مثلاً لا تكون مساوية لضعف قيمة حجم الأثر التي تساوي 0.2 لأن معاملات الإرتباط ليست خطية تماماً. (أنظر الجزء الخاص بتفسيرات حجوم الأثر)

وباستخدام برنامج SPSS يمكن الحصول على قيمة لمعامل إرتباط بيرسون بين متغيرين تتراوح بين (± 1).

ومن الملاحظات المهمة هنا أننا إذا حسبنا مربع معامل الإرتباط ، فإن هذا المقدار والذي يسمى بمعامل التحديد يشير إلى مدى مشاركة تباين أحد المتغيرين مع المتغير الأخر. فعلى سبيل المثال إذا كان معامل إرتباط بيرسون ($r = -0.44$) بين درجات درجات قلق الإمتحان وبين درجات أداء مجموعة من الطلاب على إمتحان في التحصيل فإن ($r^2 = 0.194$) وهذه القيمة تدلنا على مقدار التباين المشترك بين درجات قلق الإمتحان ودرجات أداء الطلاب في إختبار التحصيل لكنها لا تشير إلى العلاقة السببية بين هذين المتغيرين. وبتحويل الكسر العشري إلى نسبة مئوية أي بضرب 0.194 في 100 نحصل على النسبة المئوية 19.4% وهنا يمكن القول أن تباين درجات أداء الطلاب في الإختبار التحصيلي يتأثر بمقدار 19.4% بتباين درجات القلق وأن الباقي وهو80.6% من تباين الدرجات يتأثر بأداء الطلاب في إمتحان التحصيل وبعوامل أخرى. (أنظر الجزءالخاص بتفسيرات حجوم الأثر)

2. **حساب حجم الأثر لمعامل ارتباط الرتب لسبيرمان**

يستخدم معامل إرتباط الرتب لسبيرمان (Spearman Rho ρ) وهو معامل إرتباط لابارامتري عندما تكون البيانات رتبية أو كانت البيانات مستمرة لكن لا تنطبق عليها افترضات استخدام معامل ارتباط بيرسون مثل اعتدالية توزيع الدرجات، ومعامل إرتباط سبيرمان يستخدم مثله مثل معامل الإرتباط لبيرسون كمؤشر لحجم الأثر لكن تفسيره يختلف عن تفسير معامل إرتباط بيرسون كما سيأتي في الملاحظات التي في نهاية هذا الجزء.

3. **حساب حجم الأثر لمعامل كندال Kendall's tau τ**

هذا المعامل يستخدم في الإحصاء اللابارامتري كبديل لمعامل سبيرمان عندما تكون البيانات قليلة العدد وتحتوي على رتب متساوية بين المتغيرين (Tied Ranks) وعلى الرغم من أن معامل ارتباط الرتب لسبيرمان هو الأكثر استخداماً في البحوث إلا أن معامل كندال هو الأكثر دقة. وكلا المعاملين يمكن استخدامهما لإيجاد حجم الأثر. أنظر أيضا الملاحظات في نهاية هذا الجزء للفرق بين هذا المعامل ومعامل بيرسون وسبيرمان.

4. **حساب حجم الأثر لمعاملي الإرتباط الثنائي والثنائي المتسلسل النقطي**

يستخدم هذان المعاملان عندما يكون أحد المتغيرين ثنائي (Dichotomous) - أي تصنيفي يحتوي على صنفين- مثل الذكور والإناث لكن الفرق بين معامل الإرتباط الثنائي (Biserial) ومعامل الإرتباط الثنائي المتسلسل النقطي (Point Biserial) يتوقف على ما إذا كان المتغير الثنائي متقطع (Discrete) أو أن يحتوي على متغير يمكن أن يكون مستمراً (Continuous) وكمثال للمتغير المتقطع الثنائي هو أن يكون الفرد إما فقيرًا أو غنياً ، راسباً أو ناجحاً ، ذكراً أو أنثى، يمتلك سيارة أو لايمتلك سيارة. أما المتغير الثنائي المتسلسل فبين المتغيرين نوع من التسلسل فعندما نصنف طالبا بأنه راسب فقد يكون الطالب راسبا بأقل من درجة النجاح بدرجتين أو ثلاث وطالب اخر راسب أيضا ولكن أقل من درجة النجاح بعدد كبير من الدرجات وهكذا بالنسبة للنجاح فهناك طالب ناجح بدرجات أعلى من معدل النجاح وأخر متفوق بدرجات عالية. نخلص من هذا أن الباحث يستخدم معامل الإرتباط النقطي المتسلسل (Point Biserial Correlation r_{pb}) عنما تكون البيانات ثنائية فقط (المرأة حامل أو غير حامل) بينما يستخدم معامل الإرتباط المتسلسل r_b (Biserial Correlation Coefficient) عندما يكون أحد المتغيرين ثنائي خلفيته درجات مستمرة (النجاح والرسوب في الإمتحان) (علام، 1993ب).

ملاحظات مهمة تتعلق باستخدام معاملات الارتباط كمؤشر لحجم الأثر

لقد أورد فيلد (Field, 2009, p. 193) عدداً من الملاحظات تتعلق بالفروق بين معاملات الإرتباط كمؤشر لحجم الأثر هي:

أ. في معامل الإرتباط لبيرسون يمكننا تربيع قيمة معامل الإرتباط (r^2) ومن ثم نحصل على نسبة التباين المشترك بين المتغيرين موضع البحث، وينطبق نفس الكلام على معامل سبيرمان لأننا عندما نحسب معامل سبيرمان فإننا نستخدم نفس المعادلة التي تُستخدم لحساب معامل بيرسون، لكن (r^2) بالنسبة لمعامل سبيرمان تعبر عن نسبة التباين المشترك بين الرتب التي يشترك فيها كل من المتغيرين.

ب. لا يشبه معامل إرتباط كندال (τ) معامل إرتباط بيرسون أو معامل إرتباط سبيرمان من الناحية العددية ومن ثم فإن (τ^2) لا تعبر عن نسبة التباين المشترك بين المتغيرين أو نسبة التباين المشترك بين الرتب. كما أن قيمة كندال (τ) تقل بنسبة تتراوح بين 66% - 75% عن قيمة معامل إرتباط بيرسون أو سبيرمان. لذلك فعند استخدام معامل كندال (τ) لا يجب أن نستخدم قيمة (τ^2) لتفسير حجم الأثر.

جـ. يختلف كل من معامل الإرتباط النقطي المتسلسل ومعامل الإرتباط المتسلسل في المقدار فالقيمة العددية لمعامل الإرتباط المتسلسل أكبر من القيمة العددية لمعامل الإرتباط النقطي، ولذلك يجب أن يكون الباحث حذراً ومتأكدا من أن المتغير موضع الدراسة ثنائيا أم ثنائياً بظهير مستمر.

رابعا: حساب حجم الأثر لمعامل الإنحدار الخطي الثنائي والمتعدد

يُعد موضوع الإنحدار (Regression) من الموضوعات التي تهتم بالتنبؤ؛ فالباحث قد يهتم بالتنبؤ بمتغير باستخدام متغير أخر أو أكثر من متغير. وعادة ما يسمى المتغير المُنبئ (Predictor Variable) بالعامل المستقل والمتغير المتنبأ به (Criterion or Outcome Variable) بالعامل التابع. وبوجه عام يوجد أنواع من الإنحدار منها الإنحدار الخطي الثنائي البسيط، والإنحدار الخطي المتعدد والإنحدار الخطي المتعدد الهرمي والإنحدار اللوجستي وغير ذلك وفي هذا الجزء سنبين كيفية حساب حجم الأثر لنوعين من معاملات الإنحدار هما معامل الإنحدار الثنائي، ومعامل الإنحدار المتعدد.

1. حساب حم الأثر من معامل الإنحدار الخطي الثنائي

يستخدم معامل الإنحدار الخطي الثنائي إما لتفسير تأثير متغير مستقل على متغير تابع أو لتقدير مدى قدرة متغير مستقل بالتنبؤ بالمتغير التابع. وعند استخدام برنامج SPSS فإننا نحصل على أربعة أنواع من معاملات الإرتباط هي: معامل إرتباط بيرسون (r) ومعامل الإرتباط المتعدد (Multiple Correlation Coefficient R) ومربع معامل الإرتباط المتعدد (R^2) ومربع معامل الإرتباط المتعدد المعدل ($R^2_{adjusted}$). وبرغم غزارة هذه المعلومات إلا أننا نلاحظ أن ($r = R$)، ($r^2 = R^2 =$ $R^2_{adjusted}$) وبناء على ذلك فيجب أن يأخذ الباحث معامل الإرتباط r أو معامل التحديد r^2 ويقارن قيمة r بمؤشرات كوهين السابق الإشار إليها (1. تأثير ضعيف، 3. تأثير متوسط، 5. تأثير قوي) وذلك بغض النظر عن إشارة معامل الإرتباط. (أنظر أيضا الجزء الخاص بتفسيرات حجوم الأثر)

مثال:

أراد باحث الإجابة عن السؤال التالي: هل يمكن لإختبار القبول التي تجريه الكلية أن يتنبأ بدرجات طلاب الفرقة الأولى في أحد مقررات اللغة الإنجليزية بنهاية العام الدراسي الأول لهم؟ حصل الباحث على النتائج الأتية لعشرة طلاب ثم قام بتحليلها باستخدام الإنحدار الثنائي البسيط مستخدماً برنامج SPSS .

	Abilitytest	endyeartest
1	37.00	23.00
2	45.00	36.00
3	43.00	27.00
4	50.00	34.00
5	65.00	45.00
6	72.00	49.00
7	61.00	42.00
8	57.00	38.00
9	48.00	30.00
10	77.00	47.00

جدول 1 معامل إرتباط بيرسون

Correlations

		endyeartest	Abilitytest
Pearson Correlation	endyeartest	1.000	.949
	Abilitytest	.949	1.000
Sig. (1-tailed)	endyeartest	.	.000
	Abilitytest	.000	.
N	endyeartest	10	10
	Abilitytest	10	10

جدول 2 قيم R ، R^2 ، R^2 المعدلة

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.949[a]	.901	.888	2.92382	.901	72.585	1	8	.000

Predictors: (Constant), Abilitytest

جدول 3 نتيجة إختبار F

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	620.510	1	620.510	72.585	.000[b]
	Residual	68.390	8	8.549		
	Total	688.900	9			

a. Dependent Variable: endyeartest

b. Predictors: (Constant), Abilitytest

جدول 4 يبين قيمة معامل الإنحدار

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.	Correlations Zero-order	Correlations Partial	Correlations Part
1	(Constant)	2.013	4.221		.477	.646			
	Abilitytest	.632	.074	.949	8.520	.000	.949	.949	.949

a. Dependent Variable: endyeartest

من جدول 1 نلاحظ أن قيمة معامل إرتباط بيرسون r في جدول 1 تساوي قيمة معامل R في جدول 2 وكل منهما تساوي 0.949 .

من جدول 3 نلاحظ أن قيم النسبة الفائية F والتي تساوي 72.585 دالة إحصائيا وهي تعني أن معامل الإرتباط بين المتغيرين لا يساوي صفر.

ومن جدول 4 نلاحظ أن معامل الإنحدار يساوي 0.632 وهو يمثل ميل خط أفضل مطابقة مع ملاحظة أن هذه القيمة لا تساوي معامل الإرتباط بين المتغيرين.

وعلى أية حال ما يهمنا هنا هو حجم الأثر الذي يمكن الحصول عليه من قيمة R^2 أو قيمة r^2 التي تساوي 0.901 وهي في هذه الحالة تشير إلى نسبة التباين المفسر أو حجم الأثر. وبمعنى آخر فإن 90.1% من تباين العامل التابع (درجات إختبار نهاية العام) تم تفسيرها من خلال العامل المستقل (درجات إختبار القبول في اللغة الإنجليزية) . أو أن كلا المتغيرين يتشاركان بنسبة تباين قدرها 90.1% . وهذا يعني أيضا أن 9.9 من التباين يمكن تفسيره بوجود عوامل أخرى مؤثرة على درجات إختبار نهاية العام.

2. **<u>حساب حجم الأثر لمعامل الإنحدار الخطي المتعدد</u>**

Multiple Linear Regression

يستخدم الإنحدار الخطي المتعدد مثله مثل الإنحدار الثنائي إما في التنبؤ أو التفسير؛ ففي التنبؤ يحاول الباحث الكشف عن العوامل المستقلة التي يمكنها التنبؤ بالعامل التابع ، أما في التفسير فيحاول الباحث تفسير العامل التابع في ضوء أثر العوامل المستقلة. ويلاحظ أن الإنحدار الخطي يحتوي على عامل تابع واحد أو اثنين أو أكثر من العوامل المستقلة. ويعطي برنامج SPSS قيم كل من R و R^2 و R^2 المعدلة. و R تشير إلى مقدار حجم الأثر وتقارن بمؤشرات كوهين السابق الإشارة إليها أما R^2 و R^2 المعدلة فيعبران معامل التحديد أو نسبة التباين المفسر. فمثلاً إذا كانت قيمة R = .5. فإن R^2 = .25. وهذا يعني أن 25% من تباين درجات العامل التابع تشترك مع تباينات العوامل المستقلة أو تُعزى إلى العوامل المستقلة. وهذا يعني أيضا أن 75% من التباين يمكن تفسيره بوجود متغيرات أخرى مؤثرة غير العوامل المستقلة.

خامساً: حساب حجم الأثر لمعامل χ^2 وجداول الاقتران

يُعد إختبار مربع كاي (χ^2) من أشهر الإختبارات الإحصائية استخداماً عندما تكون البيانات المتاحة اسمية أو رتبية. وبوجه عام يوجد ثلاثة أنواع من إختبار مربع كاي هي:

1. إختبار مربع كاي لعينة واحدة (One Sample χ^2 Test)، ويسمى أحياناً بإختبار حسن المطابقة.
2. إختبار مربع كاي للعينات المستقلة (Independent Samples χ^2 Test) ويسمى أحياناً بمربع كاي للتجانس بين النسب.
3. إختبار مربع كاي المستقل (Chi-square Test of Independence).

وفيما يلي سنعرض لنوعين من هذه الأنواع الثلاثة

1. **حساب حجم الأثر من مربع كاي لعينة واحدة**

مثال:

نفترض أن باحثاً أراد إختبار الفرض القائل بأنه لا يوجد فرق دال إحصائيا بين تفضيلات طلاب عينة من المدارس الثانوية الأزهرية للكليات المختلفة في الجامعة. فإذا إختار الباحث عينة عشوائية تشتمل على 781 طالباً من هذه المدارس وكان تكرار تفضيلاتهم لخمسة كليات جامعية هي كالآتي:

الكلية	الطب	الصيدلة	الهندسة	الزراعة	التربية
التكرار	131	117	263	147	123

تم حساب حجم الأثر باستخدام إختبار χ^2 لحسن المطابقة من برنامج SPSS وجاءت النتائج كالتالي:

Prefernces

	Observed N	Expected N	Residual
Medicine	131	156.2	-25.2
Pharmacy	117	156.2	-39.2
Engineering	263	156.2	106.8
Agriculture	147	156.2	-9.2
Education	123	156.2	-33.2
Total	781		

Test Statistics

	Prefernces
Chi-Square	94.525[a]
df	4
Asymp. Sig.	<.001

a. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 156.2.

من الجدول الأول نحصل على عدد الطلاب الذين يفضلون كل كلية من الكليات المذكورة ومن الجدول الثاني نحصل على قيمة (χ^2) تساوي 94.525 وهي دالة إحصائيا عند مستوى 0.001 وهذا يعنى أن هناك فرقا دالاً إحصائيا بين تفضيلات الطلاب.

لحساب حجم الأثر لمربع كاي لعينة واحدة يمكننا استخدام المعادلة التالية

$$\text{Effect Size} = \frac{\chi 2}{(\text{Total Sample Size Across All Categories})(\text{Number of Categories}-1)}$$

$$Effect\ Size = \frac{94.525}{781(5-1)} = 0.033$$

وعلى الرغم من وجود فرق دال إحصائيا بدرجة كبيرة إلا أن حجم الأثر جاء ضعيفاً.

2. <u>حساب حجم الأثر باستخدام مربع كاي لعينات مستقلة</u>

يستخدم مربع كاي للعينات المستقلة مع البيانات التي تكون على هيئة تكرارات ، وتوضع هذه التكرارات في جداول تسمى جداول الإقتران (Contingency Tables) وهذه الجداول تكون لها بعدين أحدهما رأسي والأخر أفقي. وقد يتكون البعد الرأسي من عمودين أو أكثر ، كما يمكن أن يتكون البعد الأفقي من صفين أو أكثر. فعلى سبيل المثال قد يشير البعد الأفقي إلى المرحلة الدراسية (إبتدائي، إعدادي، ثانوي) وقد يشير البعد الرأسي إلى الجنس (ذكور وإناث).

مثال:

إفترض أن باحثاً أراد إختبار الفرض القائل بأنه لا يوجد فرق دال إحصائيا في إتجاهات طلاب المرحلة الثانوية نحو العمل المهني، لذلك إختار عينتين عشوائيتين إحداهما للبنبن مكونة من 80 طالباً والأخرى للبنات ومكونة من 50 طالبة. وطلب من كل فرد في العينتين تحديد إتجاهه على مقياس ليكرت ذو تدريج ثلاثي (موافق، محايد، غير موافق). يوضح الجدول التالي النتائج التي حصل عليها الباحث (مأخوذ من علام، 1993أ، ص. 265).

	طلاب	طالبات	المجموع
موافق	60	20	80
محايد	12	16	28
غير موافق	13	14	27
المجموع	80	50	135

تم حساب حجم الأثر باستخدام إختبار χ^2 لحسن المطابقة من برنامج SPSS وجاءت النتائج كالتالي:

Attitude * gender Crosstabulati

			gender		
			Male	Female	Total
Attitude	Disagree	Count	13	14	27
		Expected Count	17.0	10.0	27.0
	Neutral	Count	12	16	28
		Expected Count	17.6	10.4	28.0
	Agree	Count	60	20	80
		Expected Count	50.4	29.6	80.0
Total		Count	85	50	135
		Expected Count	85.0	50.0	135.0

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	12.366[a]	2	.002
Likelihood Ratio	12.362	2	.002
Linear-by-Linear Association	9.174	1	.002
N of Valid Cases	135		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 10.00.

من هذا الجدول يتضح أن قيمة مربع كاي تساوي 12.366 وهي قيمة دالة إحصائيا عند مستوى 0.002

وعند حساب حجم الأثر من جداول الإقتران فإن برنامج SPSS يزودنا بعدد من المعاملات (المؤشرات) التي يمكن أن تشير إلى حجم الأثر، ومن بين المعاملات الأكثر استخداماً معامل فاي φ ومعامل كريمر V مع ملاحظة أن φ لجداول الإقتران 2 x2 هي حالة خاصة من معامل إرتباط بيرسون. ويبين الجدول الأتي قيمتي معامل فاي ومعامل كرامر الناتجة من تحليل البيانات باستخدام برنامج SPSS .

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	.303	.002
	Cramer's V	.303	.002
	Contingency Coefficient	.290	.002
N of Valid Cases		135	

ويمكن حساب قيمة φ يدويا من المعادلة التالية:

$$\phi = \sqrt{\frac{\chi^2}{N}}$$

وباستخدام هذه المعادلة نجد أن

$$\varphi = \sqrt{\frac{12.366}{135}} = 0.3026$$

وهي نفس القيمة التي حصلنا عليها من برنامج SPSS وتُفسر قيمة فاي بنفس المؤشرات التي يُفسر بها معامل الإرتباط (0.1 حجم أثر صغير، 0.3 حجم أثر متوسط، 0.5 حجم أثر كبير) وبطبيعة الحال فإن تفسير هذه القيم يتوقف على طبيعة وسياق البحث كما سنبين ذلك لاحقاً.

وهنا يجدر الإشارة إلى ملاحظتين الملاحظة الأولى هي إذا كانت جداول الإقتران تحتوي عددا من الصفوف والأعمدة أكثر من 2 X 2 فإن قيمة فاي ستكون أكبر من 1 ومن ثم يصعب تفسيرها. ولذا يجب استخدام معامل كرامر والذي يُرمز له بالرمز V والذي يعمل على تعديل قيمة فاي لتكون بين صفر ، 1. أما الملاحظة الثانية فهي أنه في حالة جداول الإقتران 2 × 2، أو 2 × 3، أو 3 × 2 فإن قيمة معامل فاي وقيمة معامل كرامر تكون متساوية (Morgan et al, p. 143) ويمكن حساب قيمة معامل كرامر V من المعادلة التالية:

$$V = \sqrt{\frac{\varphi^2}{(number\ of\ rows\ or\ number\ of\ column, which\ is\ smaller - 1)}}$$

وبالتعويض بقيمة معامل φ الذي حصلنا عليها في معادلة كريمر نحصل على قيمة هذا المعامل وهي أيضا تساوي القيمة الناتجة من استخدام برنامج SPSS .

$$V=\sqrt{\frac{0.303^{2}}{(2-1)}}=0.303$$

سادساً: حساب حجم الأثر لبعض الاختبارات اللابارامترية

1. إختبار رتب إشارات المجموعات المتزاوجة أو المرتبطة (إختبار ويلكوكسون)
2. إختبار مان- ويتني لمجموعتين مستقلتين
3. إختبار فريدمان لأكثر من مجموعتين مترابطتين
4. إختبار كروسكال- واليس لأكثر من مجموعتين مستقلتين

1. حساب حجم الأثر باستخدام إختبار ويلكوكسون (Wilcoxon Signed Ranks Test)

إختبار ويلكوكسون هو إختبار لابارامتري يستخدم عندما يريد الباحث مقارنة بيانات مرتبطة أو بيانات معتمدة وهو يشبه إختبار النسبة التائية (t Test) للعينات المترابطة أو العينات المعتمدة ويختلف إختبار ويلكوكسون عن إختبار النسبة التائية في أنه يتعامل مع بيانات رتبية في حين أن إختبار النسبة التائية يتعامل مع بيانات مستمرة. وأيضا يُستخدم إختبار ويلكوكسون حينما تكون البيانات تخالف الإفتراضات (الشروط) الواجب تحققها عند استخدام الإحصاء البارامتري.

ويمكن حساب قيمة حجم الأثر باستخدام قيمة Z من جدول النتائج المعطى عند استخدام برنامج SPSS وتحويل هذه القيمة إلى r كما في المثال التالي ثم مقارنة r بقيم مؤشرات حجم الأثر.

مثال

إفترض أن باحثاً قام بتدريس مقرر في تكنولوجيا التعليم لأثني عشر طالباً من طلاب الدبلوم الخاص في التربية بإحدى كليات التربية وقام بتطبيق إختبار قبلي على الطلاب وبعد إنتهاء المقرر قام بتطبيق إختبار بعدي على نفس الطلاب وجاءت نتائجه كما يلي:

إختبار قبلي	31	14	53	18	21	44	12	36	22	29	17	40
إختبار بعدي	31	14	50	30	28	48	35	32	23	34	27	42

حلل الباحث نتائجه باستخدام إختبار ويلكوكسون من برنامج SPSS وحصل على النتائج الأتية:

Ranks

		N	Mean Rank	Sum of Ranks
Posttest - Pretest	Negative Ranks	2[a]	3.75	7.50
	Positive Ranks	8[b]	5.94	47.50
	Ties	2[c]		
	Total	12		

من الجدول السابق نحصل على قيمة ويلكوكسون (T= 7.5) وهي أصغر القيمتين في عمود مجموع الرتب (Sum of Ranks) .

كما يزودنا برنامج SPSS أيضا بقيمة Z والتي تساوي 2.040- ومستوى الدلالة الإحصائية وهو يساوي 0.041 كما بالجدول التالي

Test Statistics[a]

	Posttest - Pretest
Z	-2.040[b]
Asymp. Sig. (2-tailed)	.041

a. Wilcoxon Signed Ranks Test

b. Based on negative ranks.

ولحساب قيمة حجم الأثر نستخدم المعادلة الأتية:

$$Effect\ Size\ r = \frac{|Z|}{\sqrt{N}}$$

$$r = \frac{|-2.040|}{\sqrt{12}} = \frac{2.040}{3.46} = 0.589$$

حيث $|Z|$ هي القيمة المطلقة (أي القيمة العددية بغض النظر عن الإشارة موجبة أو سالبة)، و N هي عدد أفراد العينة. وقيمة حجم الأثر هنا عبارة عن معامل إرتباط وتفسر بنفس الطريقة التي يتم تفسير معامل الإرتباط بها.

2. <u>حساب حجم الأثر باستخدام اختبار مان-ويتني</u>

إختبار مان- ويتني هو أيضا إختبار لابارامتري يستخدم عندما يريد الباحث مقارنة بيانات مجموعتين مستقلتين وهو يشبه إختبار النسبة التائية (t Test) للمجموعات المستقلة ويختلف إختبار مان- ويتني عن إختبار النسبة التائية في أنه يتعامل مع بيانات رتبية في حين أن إختبار النسبة التائية يتعامل مع بيانات مستمرة. وأيضا يُستخدم إختبار مان- ويتني حينما تكون البيانات تخالف الإفتراضات أي الشروط الواجب تحققها عند استخدم الإحصاء البارمتري.

ويمكن حساب قيمة حجم الأثر باستخدام قيمة Z من جدول النتائج المعطي عند استخدام برنامج SPSS وتحويل هذه القيمة إلى r كما في المثال التالي ، ثم مقارنة r بقيم مؤشرات حجم الأثر.

مثال

أراد باحث إجراء تجربة طبق فيها طريقتين لتدريس الكلمات باللغة الإنجليزية لسبعة تلاميذ في الصف الخامس الإبتدائي. وبعد مرور 4 أسابيع تدريسية أجرى الباحث إختباراً في إستدعاء الكلمات وجاءت النتائج كما في الجدول الأتي:

طريقة التدريس 1	48	40	39	50	41	39	53
طريقة التدريس 2	14	18	20	10	12	102	17

قام الباحث بتحليل هذه النتائج باستخدام إختبار مان – ويتني من برنامج SPSS وحصل على النتائج الأتية:

Ranks

	Methods	N	Mean Rank	Sum of Ranks
Score	Method1	7	10.00	70.00
	Method2	7	5.00	35.00
	Total	14		

هذا الجدول يبن لنا أن عدد الأفراد في كل مجموعة هو 7 وأن متوسط رتب المجموعة الأولى هو 10 وأن متوسط رتب المجموعة الثانية هو 5 وأن مجموع رتب المجموعة الأولى هو 70 بينما مجموع رتب المجموعة الثانية هو 35

Test Statistics[a]

	Score
Mann-Whitney U	7.000
Wilcoxon W	35.000
Z	-2.236
Asymp. Sig. (2-tailed)	.025
Exact Sig. [2*(1-tailed Sig.)]	.026[b]

ويبين الجدول الثاني أن النتائج دالة إحصائيا حيث $(U=7, n_1=n_2=7, p=0.026)$
حيث U تعبر عن قيمة إختبار مان- ويتني. ويمكن حساب حجم الأثر من المعادلة الأتية :

$$Effect\ Size=\frac{|Z|}{\sqrt{N}}$$

$$r=\frac{|-2.236|}{\sqrt{14}}=\frac{2.236}{3.73}=0.63$$

3. **حساب حجم الأثر باستخدام اختبار فريدمان للمجموعات المعتمدة (المترابطة)**

يُعد إختبار فريدمان من الإختبارات اللابارامترية التي تستخدم لمقارنة أكثر من عينتين مترابطتين أو عينة واحدة تكرر القياس عليها، وهو يقابل تحليل التباين الأحادي مع تكرار القياس في الإختبارات البارامترية. وعندما تكون نتائج إختبار فريدمان دالة إحصائياً وأردنا معرفة الفروق بين العينات أو بين القياس المتكرر فحينئذ نستخدم إختبار ويلكوكسون الذي سبق الإشارة إليه لإجراء المقارنات البعدية.

ولحساب حجم الأثر يوصي (Green & Salkind, 2014, p.67) باستخدام معامل الإتفاق لكندال والذي يشار إليه بالحرف W (Kendall's Coefficient of Concordance) ونحصل على هذا المعامل مباشرة من برنامج SPSS كما سيبن في المثال التالي. وتتراوح قيمة معامل إتفاق كندال بين الصفر والواحد الصحيح. وكلما إقتربت القيمة من الواحد الصحيح دل ذلك على حجم أثر مرتفع.

ويذكر عدد أخر من الباحثين أنه لا توجد طريقة مباشرة لقياس حجم الأثر عند إستخدام إختبار فريدمان ولذا يقترحون إجراء المقارنات البعدية باستخدام إختبار ويلكوكسون والذي نحصل منه على قيمة Z بين كل مقارنتين ثم نحول قيمة Z إلى معامل الإرتباط r من المعادلة $r=\frac{|Z|}{\sqrt{N}}$ ويدل معامل الإرتباط في هذه الحالة على حجم الأثر.

مثال

قام أحد الباحثين بإجراء بحث للتعرف على ما إذا كان هناك فرق دال إحصائياً في المهارت العملية للطلاب في مقرر للكيمياء كلما تقدموا في دراسة المقرر، لذلك لاحظ الباحث مهارات الطلاب على ثلاث فترات زمنية، مرة قبل المقرر، ومرة ثانية بعد مرور شهرين من تدريس المقرر ، والمرة الثالثة بعد مرور أربعة أشهر على دراستهم للمقرر. سجل الباحث نتائجه على برنامج SPSS كالتالي.

	Baseline	TowMonth	fourMonth
1	66.00	67.00	69.00
2	49.00	50.00	56.00
3	51.00	52.00	49.00
4	65.00	65.00	69.00
5	42.00	43.00	46.00
6	38.00	39.00	40.00
7	33.00	31.00	39.00
8	41.00	41.00	44.00
9	46.00	47.00	48.00
10	45.00	46.00	46.00
11	36.00	33.00	34.00
12	51.00	55.00	67.00

وبإجراء التحليل الإحصائي باستخدام برنامج SPSS نحصل على الجدول الأتي ومنه نحصل على قيمة مربع كاي وهي تشير إلى إختبار فريدمان والدلالة الإحصائية. ويتضح من الجدول أن قيمة فريدمان وهي 10.978 دالة إحصائيا عند $p = 004$ ولإجراء المقارنات البعدية نستخدم إختبار ويلكوكسون مع تعديل فيمة ألفا لتصبح 0.016 بدلا من 0.05 (هذه القيمة تم حسابها باستخدام تعديل بونفروني أي بقسمة ألفا على عدد المقارنات)

Test Statistics[a]

N	12
Chi-Square	10.978
df	2
Asymp. Sig.	.004

a. Friedman Test

كما يمكن أيضا حساب قوة العلاقة لإختبار فريدمان كما سبق الإشارة إلى ذلك باستخدام معامل الإتفاق لكندال كما يبينه الجدول التالي والذي منه يتضح أن حجم الأثر أو قوة العلاقة هي:

$$Kendall's\ W = 0.457$$

Test Statidtics

N	12
Kendall's W[a]	.457
Chi-Square	10.978
df	2
Asymp. Sig.	.004

a. Kendall's Coefficient of Concordance

4. حساب حجم الأثر من إختبار كروسكال- واليس للعينات المستقلة

إختبار كروسكال- واليس هو أحد الإختبارات اللابارامترية الذي يستخدم لمقارنة ثلاث مجموعات مستقلة أو أكثر، وهو بذلك يشبه الإختبار البارمتري تحليل التباين في إتجاه واحد. وعادة ما يتم استخدام إختبار كروسكال- واليس لتحليل رتب الدرجات وليس الدرجات الأصلية. وعندما تُظهر النتائج فرقاً دالاً إحصائياً يجب إجراء المقارنات البعدية بين نتائج المجموعات باستخدام إختبار مان- ويتني.

ومن الملاحظ أن برنامج SPSS لا يعطي قيمة حجم الأثر بصورة مباشرة من نتيجة إختبار كروسكال- واليس، ولكن كما يقترح (Green & Salkind, 2014, p. 346) يمكن إيجاد حجم الأثر η^2 التي تمثل نسبة التباين المفسر من المعادلة الأتية:

$$\eta^2 = \frac{\chi^2}{N - 1}$$

حيث (χ^2) تعبر عن قيمة كروسكال واليس من نتيجة جدول التحليل باستخدام برنامج SPSS وأن، N هي العدد الكلي لأفراد العينة.

أما في حالة إجراء المقارنات البعدية باستخدام إختبار مان- ويتني فيحسب حجم الأثر بين كل مجموعتين باستخدام المعادلة الأتية:

$$r = \frac{|Z|}{\sqrt{N}}$$

مثال:

أراد أحد الباحثين التعرف على ما إذا كان معلمو اللغة العربية أو العلوم أو الرياضيات أكثر قبولاً لتطبيق المستحدثات التكنولوجية في تدريسهم، لذلك تخير الباحث 3 عينات عشوائية من هذه التخصصات تكونت من 8 من معلمي اللغة العربية، و6 من معلمي العلوم ، و 10 من معلمي الرياضيات. طبق الباحث استبانة تطبيق المستحدثات التكنولوجية من تصميمه على هؤلاء المعلمين وجاءت نتائجه كما في الجدول التالي: (الجدول مأخوذ من علام، 1993أ، ص ص. 427-430)

معلمو الإبتدائي	معلمو الإعدادي	معلمو الثانوي
105	101	128
130	150	260
100	130	247
108	142	280
103	214	266
109	247	273
121		291
130		256
		247
		214

تم تحليل البيانات السابقة باستخدام برنامج SPSS وجاءت نتائج التحليل كما في الجداول التالية:

Ranks

	Group	N	Mean Rank
Scores	Elementary	8	5.56
	Preparatory	6	11.42
	Secondary	10	18.70
	Total	24	

Test Statistics[a,b]

	Scores
Chi-Square	15.570
df	2
Asymp. Sig.	.000

a. Kruskal Wallis Test

b. Grouping Variable: Group

يبين الجدول الأول عدد أفراد كل مجموعة من المعلمين والعدد الكلي لهؤلاء المعلمين. كما يبين الجدول متوسط الرتب لكل مجموعة. أما الجدول الثاني فيعطينا قيمة كروسكال- واليس والذي يشار إليها في كتب الإحصاء بالحرف (H) وتعبر عنه في جدول SPSS قيمة (χ^2) التي تساوي 15.570 وهي قيمة دالة إحصائياً.

باستخدام χ^2 من الجدول أعلاه في المعادلة السابقة نحصل على قيمة حجم الأثر للمعالجات الثلاث.

$$\eta^2 = \frac{\chi^2}{N-1} = \frac{15.570}{24-1} = \frac{15.570}{23} = 0.14$$

أما إذا أردنا حساب حجم الأثر بين كل معالجتين على حدة فنستخدم مان – ويتنى كإجراء للمقارنات البعدية ثم نحسب حجم الأثر باستخدام المعادلة المذكورة سابقاً وهي

$$r = \frac{Z}{\sqrt{N}}$$

طرق تفسير حجوم لأثر

إفترض أنك تقرأ تقريراً لأحد البحوث التجريبية الذي أجري لدراسة حجم الأثر لطريقة تدريس حديثة مقابل طريقة تدريس تقليدية، وأشارت الدراسة إلى أن حجم الأثر $d = 0.55$ أو اشارت إلى أن حجم الأثر $r = 0.10$ فماذا تعني هذه المعلومات لصانع القرار أو للمعلم داخل الصف أو حتى لباحثين آخرين ليسوا على دراية بمعنى حجم الأثر. أو قد يسأل كل هؤلاء عن أهمية هذه النتائج أو ما الدلالة العملية لها؟ إن الإجابة عن هذه الأسئلة وغيرها هي موضوع هذا الجزء. ,سوف نبدأ أولاً بعرض التفسير غير الدقيق لحجوم الأثر ثم يلي ذلك عرض لبعض الطرق لتفسير حجوم الأثر.

التفسير غير الدقيق لحجوم الأثر

إن بعض طرق عرض نتائج حجم الأثر سواء من قبل الباحثين أو الممارسين قد تكون غير دقيقة وأحياناً قد تكون مضللة للقارئ. ولبيان ذلك سوف نعرض لطريقتين شائعتين يتم من خلالهما عرض بيان أهمية نتائج البحث. هاتان الطريقتان هما استخدام قيمة الدلالة الإحصائية الإحتمالية p واستخدام مؤشرات كوهين التي تصنف حجوم الأثر إلى ثلاثة مستويات صغيرة ومتوسطة وكبيرة.

أ. استخدام إختبار الدلالة الإحصائية p

قد يستخدم بعض الباحثين قيمة الدلالة الإحصائية p للتعبير عن أهمية النتائج أو للتعبير عن قوة المعالجة التجريبية، فعندما يحصل الباحث على قيمة دالة إحصائياً عند مستوى 0.05 فقد يقرر أن نتائج البحث ذات أهمية معينة أما إذا حصل على قيمة دالة إحصائياً عند مستوى 0.01 فقد يعتبر هذه النتائج أكثر أهمية من تلك التي جاءت عند مستوى 0.05 والأكثر من ذلك إذا جاءت النتائج دالة عند مستوى 0.001 فقد يعتبرها أكثر أهمية من تلك التي جاءت عند مستوى 0.05 أو 0.01 كل هذه التفسيرات خاطئة فهي تعنى أن الباحث لا يعرف معنى الدلالة الإحصائية

التي هي مجرد إحتمال وهذا الإحتمال يتوقف على حجم العينة فكلما كبر حجم عينة الدراسة جاءت النتائج دالة إحصائياً بغض النظر عن قوة المعالجة التجريبية. هذا من ناحية ومن ناحية أخرى تتوقف قيمة الدلالة الإحصائية على مقدار تباين درجات العامل التابع وعلى وجود أو عدم وجود متغير مصاحب بل وعلى نوع الإختبار الإحصائي المستخدم.

ب. استخدام مؤشرات كوهين التي تصنف حجم الأثر إلى صغير ومتوسط وكبير

إن استخدام مؤشرات كوهين لتصنيف حجم الأثر *d* أو *r* يؤدي عادة إلى تفسيرات خاطئة إذا لم يؤخذ السياق الذي يتم فيه البحث في الإعتبار فما قد يكون حجم أثر صغير في سياق معين قد يكون ذا أهمية كبيرة وقد يكون حجم أثر كبير في سياق آخر غير ذي أهمية.

لقد إقترح كوهين المستويات الثلاثة لحجم الأثر للفرق المعياري بين متوسطين (*d*) والتي حددها بأن يكون المستوى صغيراً عندما يكون حجم الأثر مساوياً 0.20 ويكون المستوى متوسطاًعندما يكون حجم الأثر مساوياً 0.50، ويكون المستوى كبيراً عندما يكون حجم الأثر مساوياً 0.80 هذه المستويات توصل إليها من دراسته لعدد كبير من البحوث المنشورة في إحدى الدوريات العلمية المتخصصة وهي دورية Journal of Abnormal and Social Psychology . ولقد أشار كوهين أنه باستخدام أي من هذه المستويات الثلاثة لحجم الأثر في وجود مستوى الدلالة الإحصائية وحجم العينة يمكن تحديد القوة الإحصائية للبحث قبل إجرائه.

ولتوضيح هذه الفكرة دعنا نفترض أن باحثاً أراد استخدام إختبار النسبة التائية لمقارنة متوسطي عينتين تجريبيتين وأراد الحصول على قوة إحصائية قدرها 0.70 ففي هذه الحالة يمكن للباحث أن يختار عينة قدرها 78 فردا يتم توزيعهم على مجموعتين بكل مجموعة 39 فرداً وأن يحدد مستوى الدلالة عند 0.05 وأن يحدد حجم أثر متوسط قدره 0.5 . كما يمكن للباحث أيضا إذا حدد حجم أثر كبير يساوي 0.80 ومستوى دلالة قدره 0.05 وقوة إحصائية قدرها 0.80 وباستخدام برنامج G Power فيمكنه تحديد حجم العينة المطلوب والذي في هذه الحالة يساوي 21 فرداً في كل مجموعة أو 42 فرداً كعينة كلية.

إذن الفكرة الأساسية تتلخص في أنه لابد عند تفسير حجم الأثر من أن يضع الباحث في إعتباره السياق الذي أُجري فيه البحث فعندما حدد كوهين هذه المستويات الثلاثة المشار إليها حددها في سياق بحوث علم النفس وليس في سياق بحوث تربوية (Cohen, 1988). ولبيان فكرة أهمية السياق في تفسير حجوم الأثر أعطى (Lipsey et al, 2012, p.4) مثالاً بقولهم " لا يمكن تحديد أطوال الأطفال في ثلاث مستويات قصير، متوسط ، وطويل القامة منسوبة إلى البشر ولكن لابد من إرجاع هذه الأطوال إلى أعمارهم فقد نحدد عمر 5 سنوات مثلاً ثم نحدد مستويات الأطوال عند هذه السن أو يمكن أن نحدد جنس الأطفال أولا ثم نحدد أطوال البنين أو البنات في سن معينة.

بعد عرض هاتين النقطتين السلبيتين عند تفسير حجوم الأثر نعرض لعدد من الطرق الصحيحة لعرض وتفسير حجوم الأثر.

التعبير عن حجوم الأثر Describing Effect Sizes

عند التعبير عن حجوم أثر البحوث والدراسات يُفضل دائما التعبير عنها بنفس مستوى القياس الذي استخدم في تحليل البيانات ليسهل فهمها. فعلى سبيل المثال إذا كانت الدراسة تقارن بين

عدد الناجحين وعدد الراسبين فى مقرر معين (بيانات ثنائية) فيمكن ببساطة عرض النتائج على هيئة نسب مئوية لكل مجموعة على حدة وبهذه الطريقة في العرض يسهل على غير المتخصصين في مجال الإحصاء فهم المقصود بتأثير المقرر على الطلاب الذين درسوه.

أما إذا كانت البيانات مستمرة فيمكن التعبير عنها بحجم الأثر للفرق المعياري بين متوسطي مجموعتين (*d*) مع الإشارة إلى النسب المئوية لأفراد المجموعتين. فعلى سبيل المثال إذا أجري بحث لتحديد فاعلية استخدام برنامج معين كبرنامج الرحلات التعليمية الإستكشافية على الإنترنت لمجموعة من الطلاب مقابل مجموعة أخرى لم تدرس بهذا البرنامج وبتطبيق إختبار بعدي على المجموعتين وحساب حجم الأثر (*d*) ووجد أنه يساوي 0.20 فإنه من الأفضل في هذه الحالة أن نبين للقارئ النسبة المئوية لتحصيل الذين استخدموا برنامج الرحلات التعليمية مقابل النسبة المئوية لتحصيل الذين لم يستخدموا هذا البرنامج.

وعلى الرغم من سهولة هذا الأمر في البحوث الأولية إلا أن ذلك ليس سهلاً في بحوث التحليل البعدي، وذلك لأنه من الممكن أن يستخدم الباحثون الذين يدرسون نفس المشكلة مقاييس مختلفة في بحوثهم لقياس المتغير التابع نفسه. فعلى سبيل المثال دعنا نذكر المثال الذي عرضه (Valentine et al, 2019) ففي الولايات المتحدة الأمريكية لابد أن يجتاز الطلاب المتقدمين للإلتحاق بالجامعات أحد إختبارين هما إختبار تقييم القدرات الدراسية (Scholastic Assessment test "SAT") أو الإختبار الأمريكي للكليات ("American College Testing "ACT) واختبار (SAT) له إنحراف معياري قدره 117 نقطة بينما إختبار (ACT) له إنحراف معياري قدره 5 نقاط تقريباً. فإذا إهتم باحث بإجراء تحليل بعدي للبحوث التي أجريت لدراسة فاعلية برنامج تدريبي معين لإعداد الطلاب لإجتياز هذه الإختبارات فسيجد الباحث أن بعض الدراسات استخدمت إختبار (SAT) ودراسات أخرى استخدمت إختبار (ACT) وهو ما يعني أن نتائج الدراسات التي ستستخدم في التحليل البعدي لم يتم قياسها بنفس المقياس، لذلك يستخدم الباحث حجم الأثر (*d*) للفرق المعياري بين المتوسطات ليتحرر من إختلاف المقياسين. ولكن قد لا يستطيع القارئ العادي أو المسؤول التربوي أن يفهم المقصود بحجم الأثر. ولتوضيح ذلك دعنا نفترض أن حجم الأثر الذي حسبه الباحث جاء مساوياً 0.08 وهنا يجب على الباحث أن يحول قيمة حجم الأثر إلى عدد من النقاط التي يحصل عليها من إجتاز إختبار (SAT) أو إختبار (ACT) وذلك بضرب قيمة حجم الأثر في الإنحراف المعياري لدرجات كل إختبار على حدة للحصول على قيمة المعالجة في كل مقياس. وفي المثال الذي معنا سنلاحظ أن أثر المعالجة في إختبار (SAT) سوف يرفع درجات الطلاب بمقدار 9.4 نقطة (0.08×117)، بينما أثر المعالجة على إختبار (ACT) سوف ترفع درجات الطلاب بمقدار 0.40 نقطة (0.08×5). وإذا إفترضنا أن إختبار (SAT) له متوسط قدره 515 درجة وأن إختبار (ACT) له متوسط قدره 21 درجة. فحينئذ يمكن عرض النتائج للقارئ العادي أو المسؤول الجامعي أو حتى الطلاب كما يلي:

1. أن درجات الطالب المتوسط على إختبار (SAT) سوف تزداد من 515 درجة إلى 524.4 درجة.
2. أن درجات الطالب المتوسط على إختبار (ACT) سوف تزداد من 21 درجة إلى 21.4 درجة.

من هذا يتبين لنا أن عرض النتائج بهذه الصورة يُسهّل على المهتمين بهذه النتائج فهم المقصود بأن حجم الأثر كان يساوي 0.08 .

بعد هذا العرض سوف نبين كيف يمكن تحويل بيانات الدراسات التي يتم التعبير عنها بحجم الأثر للفرق المعياري بين المتوسطات، أو بمعاملات الإرتباط ، أو بنسبة التباين المفسر، وكذلك حجم الأثر الثنائي، ومعاملات كوهين المعبر عنها بالحرف U وحجم الأثر الذي يعبر عنه باللغة الشائعة.

التعبير عن حجم الأثر بمعاملات الارتباط

لعل كثيرا ممن يتعاملون مع البحوث العلمية قد لا يستوعبون المقصود بالفرق المعياري بين متوسطين بالقدر الكافي، لكن قد يكون من السهل عليهم التعامل مع معاملات الإرتباط لسهولة فهمها. هذا بالإضافة إلى أن كثيراً من البحوث يمكن تحليل نتائجها باستخدام معاملات الإرتباط. ولقد أشرنا في جزء سابق إلى أنه يمكن تحويل الفرق المعياري بين متوسطين إلى معامل الإرتباط كما يمكن تحويل معامل الإرتباط إلى الفرق المعياري بين متوسطين (أنظر على سبيل المثال Field, 2004, pp. 56-58).

وعلى الرغم من سهولة استخدام معاملات الإرتباط كمؤشرات لحجم الأثر إلا أن هناك تحذيرين مهمين ذكرهما (Valentin et al, 2019) التحذير الأول يتعلق باعتقاد بعض الباحثين أن معامل الإرتباط لا يبين العلاقة السببية المحتملة بين المتغيرات، لكن حقيقة الأمر أن العلاقة السببية ترجع بالدرجة الأولى إلى تصميم البحث وليس إلى التحليل الإحصائي للبيانات. كما أن كثيرا من الباحثين قد لا يعلمون أن معاملات الإرتباط يمكن استخدامها لتحليل نتائج البحوث التجريبية.

أما التحذير الثاني فيتعلق بأن بعض الباحثين قد يلجأون إلى إيجاد مربع معامل الإرتباط للحصول على نسبة التباين المفسر، لكن مربع معامل الإرتباط قد يؤدي بقارئ البحث إلى الإعتقاد بصغر حجم الأثر. فعلى سبيل المثال إذا كان حجم الأثر معبراً عنه بمعامل إرتباط يساوي 0.3 فإن مربع هذا المعامل الذي يعبر عن نسبة التباين المفسر تساوي 0.09 وهي نسبة تبدو صغيرة للقارئ غير المتخصص دون ربطها بالسياق الذي تم فيه البحث.

التعبير عن حجم الأثر باستخدام U لكوهين

قدم كوهين في كتابه تحليل القوة الإحصائية في العلوم السلوكية Statistical Power Analysis for Behavioral Sciences الصادر عام 1977 في طبعته الأولى وعام 1988 في طبعته الثانية قدم ثلاثة مقاييس لتحويل حجوم الأثر إلى نسب مئوية تحت المنحنيين الإعتداليين لدرجات مجموعتين إحداهما تجريبية والأخرى ضابطة وأشار إلى هذه المقاييس الثلاثة بالأحرف U_1 ، U_2 ، U_3 (Cohen, 1988, pp. 20-22)

وتشير هذه المقاييس الثلاثة إلى التداخل أو عدم التداخل لمنحنيات توزيع درجات المجموعة التجريبية والمجموعة الضابطة. ومن ثم يمكن تحويل حجوم الأثر إلى نسب مئوية باستخدام جداول وضعها كوهين لهذا الغرض وهنا سنعرض للجداول المتعلقة بالمقاييس الثلاثة بالإضافة إلى قيم معامل الإرتباط (r) ومربع معامل الإرتباط (r^2) (أنظر جدول في الصفحة التالية)

ويعتبر المقياس U_3 هو أكثر المقاييس الثلاثة شيوعاً بين الباحثين لسهولته لأنه يشير إلى نسبة زيادة متوسط درجات المجموعة التجريبية عن متوسط درجات المجموعة الضابطة والتي اعتبر كوهين متوسطها 50%. فعلى سبيل المثال إذا كان حجم الأثر للفرق المعياري بين متوسطين (d) يساوي صفر فإن U_3 تعبر عن أن 50% من درجات توزيع المجموعة التجريبية يتجاوز متوسط توزيع درجات المجموعة الضابطة، وإذا كان حجم الأثر (d) يساوي 1.00 (أي إنحراف معياري واحد)

فإن U_3 تعبر عن أن 84.1% من متوسط درجات توزيع المجموعة التجريبية تتجاوز متوسط توزيع درجات المجوعة الضابطة. وبمعنى أخر فإن نسبة التحسن في المجموعة التجريبية تساوي 34.1 % أي84.1% مطروحا منها 50%. (أنظر أيضا الجدول في ملحق 2).

ولتسهيل الأمر على الباحثين دون اللجوء للمعادلات التي ذكرها كوهين يمكن إيجاد قيمة U_3 من خلال برنامج (Google- Sheets) أو برنامج (Excel) باستخدام الدالة الأتية NORM.S.DIST(d)*100 فعلى سبيل المثال إذا كان حجم الأثر (*d*) يساوي 0.80 فعند استخدام هذه الدالة من برنامج (Google- Sheets) نحصل على قيمة U_3 تساوي 78.8%.

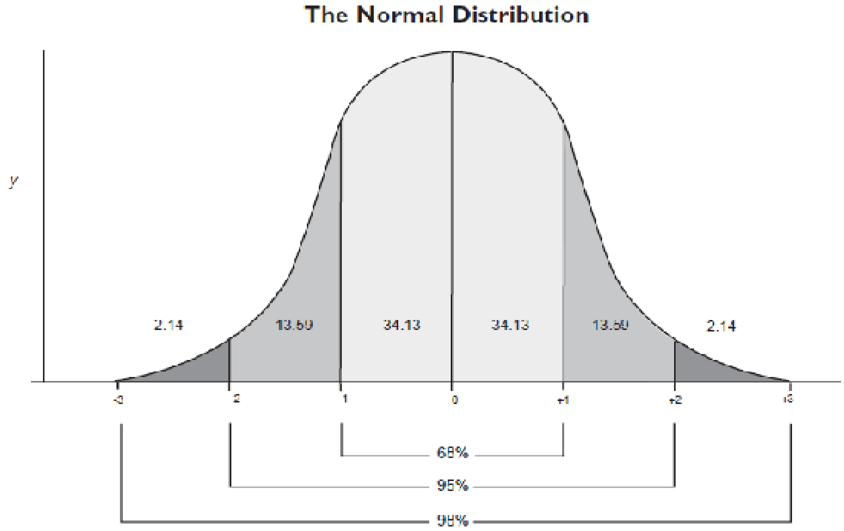
جدول يبين قيم U ، r ، r^2

d	r	r^2	U_1	U_2	U_3	WWC-II	CLES	BESD
+4.00	+.89	80.0%	97.7%	97.7%	99.9%	50	99.8%	94.7%
+3.00	+.83	69.2	92.8	93.3	99.9	50	98.3	91.6
+2.00	+.71	50.0	81.1	84.1	97.7	48	92.1	85.4
+1.50	+.60	36.0	70.7	77.3	93.3	43	85.6	80.0
+1.00	+.45	20.0	55.4	69.1	84.1	34	76.0	72.4
+0.50	+.24	5.9	33.0	59.9	69.1	19	63.8	62.1
+0.40	+.20	3.5	27.4	58.0	65.5	16	61.1	59.8
+0.30	+.15	2.2	21.3	56.0	61.8	12	58.4	57.4
+0.20	+.10	1.0	14.8	54.0	57.9	8	55.6	55.0
+0.10	+.05	0.25	7.7	52.0	54	4	52.8	52.5
0.00	.00	0	0	50	50	0	50	50
−0.10	−.05	0.25	7.7	48.0	46	−4	47.2	47.5
−0.20	−.10	1.0	14.8	46.0	42.1	−8	44.4	45.0
−0.30	−.15	2.2	21.3	44.0	38.2	−12	41.6	42.6
−0.40	−.20	3.5	27.4	42.0	34.5	−16	38.9	40.2
0.50	.24	5.9	33.0	40.1	30.9	19	36.2	37.9
1.0	.45	20.0	55.4	30.9	15.9	34	24.0	27.6
−1.50	−.60	36.0	70.7	22.7	6.7	−43	14.4	20.0
−2.00	−.71	50.0	81.1	15.9	2.3	−48	7.9	14.6
−3.00	−.83	69.2	92.8	6.7	0.1	−50	1.7	8.4
−4.00	−.89	80.0	97.7	2.3	0.1	−50	0.2	5.3

التعبير عن حجم الأثر باستخدام جدول مارزانو وزميليه

لقد قام كل من مارزانو وزميليه (Marzano et al, 2001) بعمل جدول لتحويل حجوم الأثر للفرق المعياري بين متوسطين إلى نسبة كسب مئوية خاصة عند استخدام الباحث للتحليل البعدي للبحوث والدراسات. فعلى سبيل المثال إذا إفترضنا أن حجم الأثر (*d*) يساوي 1.00 في دراسة للتحليل البعدي فهذا يعني أن متوسط درجات أفراد المجموعة التجريبية أعلى من متوسط درجات المجموعة الضابطة بمقدار 1.00 إنحراف معياري. كما بين مارزانو وزميلاه أن هذه النتيجة تعني أن فرداً ما عند المئين 50 في المجموعة التجريبية سيكون على بعد مقداره إنحراف معياري واحد من فرد آخر عند المئين 50 في المجموعة الضابطة. وهذا يعني سهولة تحويل حجم الأثر إلى

نسبة كسب مئوية. ولإجراء ذلك نفترض أن توزيع درجات تحصيل مجموعة من الطلاب مثلاً موزعة توزيعاً إعتدالياً (أنظر الشكل التالي)

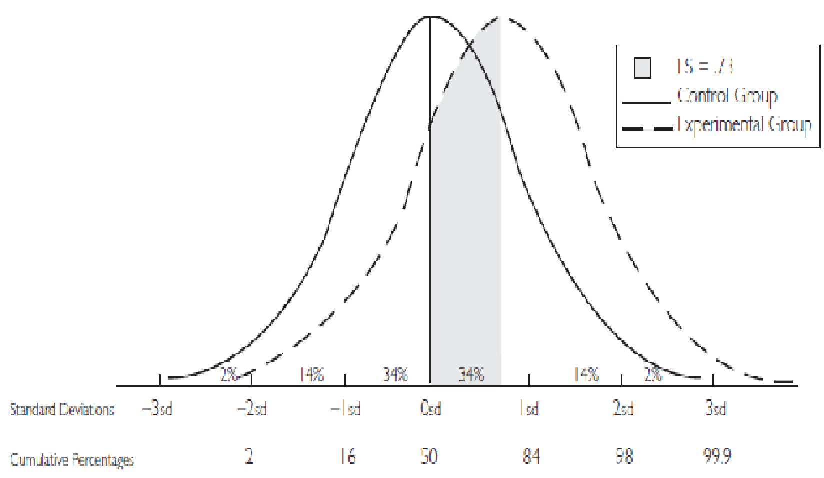


من هذا الشكل يتضح أن هناك مدى للإنحراف المعياري يمتد من 3 إنحراف معياري فوق المتوسط و 3 إنحراف معياري تحت المتوسط. كما يبين الشكل أن 34% من الدرجات تحت المنحنى الإعتدالي تقع بين المتوسط وإنحراف معياري واحد، وأن حوالي 14% من الدرجات تقع بين بين الإنحراف المعياري الأول والإنحراف المعياري الثاني وهكذا.

وبالعودة للمثال الذي معنا والذي يشير إلى أن حجم الأثر يساوي 1.00 (أي إنحراف معياري واحد) فإننا نستطيع تفسير هذه القيمة في ضوء النسبة المئوية للكسب أي أن حجم الأثر الذي يساوي 1.00 يشير إلى نسبة كسب قدرها 34% نقطة. وهكذا نرى أن تحويل حجم الأثر إلى نسبة كسب مئوية يساعد الباحثين على عرض وتفسير النتائج بلغة يسهل فهمها.

وكمثال آخر ذكره مارزانو وزميلاه هو أن إحدى الدراسات التي أجريت في التحليل البعدي لعدد 14 دراسة أولية عن أثر استخدام استراتيجية طرح الأسئلة ذات المستويات العليا من التفكير على تحصيل الطلاب توصلت إلى أن متوسط حجم الأثر لهذه الدراسات يساوي 0.73 وهذا يعني أن الطالب المتوسط الذي درس باستراتيجية طرح الأسئلة ذات المستويات العليا من التفكير جاءت درجاته أعلى من درجات زميله الذي لم يدرس بهذه الإستراتيجية بمقدار 0.73 إنحراف معياري (أنظر الجزء المظلل بالشكل التالي)

Average Effect Size Using Higher-Level Questions



وبالرجوع إلى الجدول الذي وضعه مارزانو وزميلاه (ملحق 3) يمكن تحويل حجم الأثر هذا إلى نسبة كسب مئوية قدرها 27 نقطة أي أن درجة الطالب المتوسط في مجموعة استراتيجية طرح الأسئلة ذات المستويات العليا يتفوق على زميله الذي لم يدرس بهذه الإستراتيجية بمقدار 27 نقطة.

ويمكن أيضاً استخدام الدالة NORM.S.DIST(*d*) من برنامج (Google- Sheets) للحصول على قيمة U_3 كما سبق ثم طرح قيمة U_3 من 50% نحصل على نسبة الكسب المئوية التي إقترحها مارزانو وزميلاه. فعلى سبيل المثال لو إفترضنا أن حجم الأثر بين متوسطين (*d*) يساوي 1.61 فيمكن استخدام الدالة المذكورة لحساب قيمة U_3 وتساوي 94.6% وبطرح هذه القيمة من القيمة عند المتوسط والتي تساوي 50% نحصل على النسبة المئوية للكسب والتي تساوي 44.6 % (أي 45% تقريبا) (قارن هذه النتيجة بما جاء في جدول مارزانو في ملحق 3).

<u>التعبير عن حجم الأثر باستخدام اللغة الشائعة</u> Common Language Effect Size (CLES)

قدم كل من (McGraw & Wong, 1992) طريقة لتحويل حجوم الأثر إلى لغة سهلة يمكن أن يفهمها غير المتخصصين في الإحصاء أطلقا عليها اسم اللغة الشائعة للتعبير عن حجم الأثر وإختصاراً يرمز لها بالأحرف (CLES). وهذه الطريقة تعبر عن إحتمال تفوق مجموعة مختارة عشوائياًعن إحتمال تفوق مجموعة أخرى مختارة عشوائياً أيضا. كما تشبه هذه الطريقة تلك التي أشار إليها كوهين بالحرف *U* . فعلى سبيل المثال إذا كان حجم الأثر للفرق المعياري بين متوسطين يساوي صفر فإن قيمة هذا التحويل تساوي 50% وتقترب هذه القيمة من الواحد الصحيح كلما زادت قيمة حجم الأثر. ويمكن الحصول على قيمة CLES من برنامج (Google-Sheets) من الدالة

$$NORM.S.DIST\left(\frac{d}{\sqrt{2}}\right)$$

فعلى سبيل المثال إذا كان لدينا نتائج تجربة جاء حجم الأثر لها (*d*) يساوي 0.30 فيمكن استخدام برنامج (Google- Sheets) لحساب قيمة CLES المناظرة وفي هذه الحالة تساوي 0.58 تقريباً وهذا يعني أن الطالب في المجموعة التجريبية ستكون درجته أعلى بمقدار 58% من درجة الطالب في المجموعة الضابطة.

التعبير عن حجم الأثر للبيانات الثنائية Binomial Effect Size Display (BESD)

إقترح هذا المؤشر (Rosenthal & Rubin, 1982) بعد أن لاحظا أن عدداً لابأس به من المتخصصين في علم التفس والإحصاء لا يفسرون حجم الأثر المعبر عنه بمعامل الإرتباط r تفسيراً صحيحاً خاصة عن تربيع قيمة هذا المعامل للحصول على نسبة التباين المفسر، فعلى سبيل المثال إذا كان معامل الإرتباط يساوي 0.30 فإن نسبة التباين المفسر في هذه الحالة تساوي 0.09 وهي نسبة ضعيفة ولا تعني أن للمعالجة أثراً جيداً في رأي هؤلاء المتخصصين. لهذا قدم هذان الباحثان مؤشرهما الذي أطلقا عليه مؤشر حجم الأثر للبيانات الثنائية أو للبيانات ذات الحدين (Binomial Effect Size Display "BESD") وأشارا إلى أهمية هذا المؤشر لأنه يسهل فهمه على الباحثين، والممارسين الميدانيين، والطلاب، وعامة الناس، وأن هذا المؤشر قابل للتطبيق في سياقات مختلفة، وأنه من السهل الحصول عليه إحصائيا بمعادلات بسيطة. هذه المعادلات عرضها (Valentine et al, p. 440) وهي كالتالي:

نسبة درجات المجموعة التجريبية التي ستكون أعلى من الوسيط = $0.5+\dfrac{r}{2}$

نسبة درجات المجموعة الضابطة التي ستكون أقل من الوسيط = $0.5-\dfrac{r}{2}$

حيث r هي معامل الإرتباط الثنائي المتسلسل بين درجات المجموعتين.

هاتان المعادلتان تعملان بصورة أكثر دقة عندما يتساوى أفراد المجموعتين ويكون تباين المجموعتين متجانساً. فعلى سبيل المثال إذا كان لدينا مجموعتان تلقت إحداهما معالجة تجريبية والمجموعة الأخرى ضابطة، وكان معامل الإرتباط بين درجاتهما يساوي 0.32 فإن معدل النجاح في المجموعة الأولى يساوي $0.5+\dfrac{.32}{2}=.66$ أو كما بين رونتال وروبين (Rosenthal & Rubin, 1982, p. 167) أن هذه المعالجة رفعت نسبة النجاح في المجموعة التجريبية من 34% إلى 66% (أنظر الجدول الأتي)

أما معدل النجاح في المجموعة الضابطة فيساوي $0.5-\dfrac{.32}{2}=.34$ ويمكن عرض هذه النتائج كما في الجدول التالي:

	النسبة المئوية للدرجات فوق الوسيط	النسبة المئوية للدرجات تحت الوسيط
المجموعة التجريبية	66	34
المجموعة الضابطة	34	66

كما يمكن حساب هذا المؤشر (BESD) باستخدام حجم الأثر d من خلال تحويله إلى معامل الإرتباط بالمعادلة التالية

$$r = \frac{d}{\sqrt{d^2 + 4}}$$

فعلى سبيل المثال إذا كان حجم الأثر d يساوي 0.50 فيمكن باستخدام المعادلة السابقة إيجاد قيمة r والتي تساوي 0.24 ومن هذه القيمة يمكن الحصول على مؤشر BESD كما سبق. ويشير الجدول التالي المأخوذ من مقالة روزنتال وروبين إلى قيم r ، r^2 والزيادة في معدلات النجاح ومقدر الفرق بين معدلات النجاح

r^2	r	Success rate increased From	Success rate increased To	Difference in success rates
.01	.10	.45	.55	.10
.04	.20	.40	.60	.20
.09	.30	.35	.65	.30
.16	.40	.30	.70	.40
.25	.50	.25	.75	.50
.36	.60	.20	.80	.60
.49	.70	.15	.85	.70
.64	.80	.10	.90	.80
.81	.90	.05	.95	.90
1.00	1.00	.00	1.00	1.00

مقارنة حجوم الأثر المحسوبة بحجوم أثر معيارية

في الأجزاء السابقة عرضنا لكيفية تفسير حجوم الأثر التي يقوم الباحث بحسابها سواء لبحث أولي أو لبحث في التحليل البعدي إلى مؤشرات إحصائية بسيطة يمكن فهمها بواسطة الباحثىن والممارسين التربويين. وفي الجزء التالي سنعرض لطرق أخرى لتفسير حجوم الأثر المحسوبة بالإعتماد على معايير ترتبط أساساً بالبحوث التربوية. هذه الطرق عرضها باحثون مثل (Hill et al, 2008; Lipsey et al, 2012; Valentine, et al, 2019) يمكن تلخيصها كما يلي:

1. مقارنة حجوم الأثر المحسوبة بمعايير متوقعة للنمو الأكاديمي.
2. مقارنة حجوم الأثر المحسوبة بأهداف السياسة التعليمية أو الفجوات في الأداء الأكاديمي الطلاب.
3. مقارنة حجوم الأثر المحسوبة بحجوم أثر من دراسات مشابهة أو من دراسات التحليل البعدي.

أولا- مقارنة حجوم الأثر المحسوبة بمعايير متوقعة للنمو الأكاديمي

في هذه الطريقة يتم مقارنة حجم الأثر الذي يحصل عليه الباحث بمعايير تم تحديدها لمستويات التحصيل المتوقعة للطلاب من مستوى رياض الأطفال إلى مستوى الصف الثاني عشر (Hill et al, 2008). لقد وضعت هيل وزملاؤها مستويات لحجوم الأثر للفرق المعياري بين المتوسطات (d) وفق

درجات التحصيل في الإختبارات المقننة لأربع مواد دراسية هي القراءة والعلوم والرياضيات والدراسات الإجتماعية. هذه المستويات تبين مستوى تحصيل الطالب المتوسط من عام دراسي إلى العام الدراسي التالي له (أنظر الجدول التالي).

وإذا تتبعنا متوسط حجم الأثر في أي مادة دراسية من المواد المعروضة في الجدول سنلاحظ إنخفاضاً في متوسطات حجوم الأثر بصورة تدريجية من عام دراسي إلى الذي يليه حتى تصبح متوسطات حجوم الأثر صغيرة جدا عند الصف الثاني عشر. ففي مادة القراءة مثلاً نجد أن متوسط حجم الأثر للدراسات التي تمت في التحليل البعدي بين الصفين الأول والثاني يساوي 0.97 أما بين الصفين الخامس والسادس نلاحظ أن متوسط حجم الأثر هبطت قيمته وأصبحت 0.32 ، أما بين الصفين الثامن والتاسع هبطت قيمة متوسط حجم الأثر لتصبح 0.24 مع وجود بعض الإستثناءات.

ويمكن باستخدام الجدول أن يقوم باحث بإجراء بحث أو بإجراء تحليل بعدي في القراءة مثلا على تلاميذ الصف الأول الإبتدائي وجاء حجم الأثر يساوي 0.15 فيمكن للباحث ان يقارن هذه القيمة بالقيمة الموجودة بالجدول والتي تساوي 0.97 ويصل إلى استنتاج مؤداه أن حجم الأثر الذي توصل إليه من بحثه يعتبر ضعيفا مقارنة بحجم الأثر المناظر في الجدول. أما إذا كان الباحث قد أجرى بحثه على طلاب الصف العاشر وحصل على حجم أثر يساوي 0.15 أيضا وبمقارنة هذه القيمة بما يناظرها في الجدول والتي تساوي 0.19 يمكن للباحث أن يقول أنه توصل إلى حجم أثر مرتفع.

جدول # يبين حجوم الأثر لأربعة مواد دراسية كما جاءت في دراسة (Lipsey et al, 2012)

Grade Transition	Reading	Math	Science	Social Studies
Grade K - 1	1.52	1.14	--	--
Grade 1 - 2	0.97	1.03	0.58	0.63
Grade 2 - 3	0.60	0.89	0.48	0.51
Grade 3 - 4	0.36	0.52	0.37	0.33
Grade 4 - 5	0.40	0.56	0.40	0.35
Grade 5 - 6	0.32	0.41	0.27	0.32
Grade 6 - 7	0.23	0.30	0.28	0.27
Grade 7 - 8	0.26	0.32	0.26	0.25
Grade 8 - 9	0.24	0.22	0.22	0.18
Grade 9 - 10	0.19	0.25	0.19	0.19
Grade 10 - 11	0.19	0.14	0.15	0.15
Grade 11 - 12	0.06	0.01	0.04	0.04

الجدول منقول عن (Lipsey et al, 2012, p. 28)

ملحوظة: حجوم الأثر في الجدول تنطبق على الطلاب الذين أجريت عليهم دراسات (Hill et al, 2008) وهم طلاب في مدارس الولايات المتحدة الأمريكية ، وحتى يتسنى لنا في مصر والعالم العربي إجراء مثل هذه الدراسات يجب أخذ هذه المعايير بحذر شديد أو إيجاد معايير خاصة بنا.

ثانيا- مقارنة حجوم الأثر المحسوبة بأهداف السياسة التعليمية أو الفجوات في أداء الطلاب

تركز بعض البحوث على دراسة فاعلية برامج دراسية معينة على أداء الطلاب ومدى تقدمهم نحو تحقيق الأهداف المرجوة منها. فإذا وجدت فروق واضحة (أي هناك فجوة أو فجوات) بين تحصيل طلاب المدارس الحكومية والمدارس الخاصة في اللغة الإنجليزية مثلا فقد يكون هناك برامج لتقليل هذه الفروق. فعلى سبيل المثال إذا قام باحث أو مجموعة من الباحثين بحساب حجم الأثر للبرامج القائمة بالمدارس الحكومية، ووجد هذا الباحث أو هؤلاء الباحثين أن حجم الأثر في هذه البرامج يساوي 0.14 بينما كان حجم الأثر في برامج المدارس الخاصة يساوي 0.23 ففي هذه الحالة توجد فجوة مقدارها 0.9 إنحراف معياري (أو d = 0.9). فإذا قام باحث أو مجموعة من الباحثين بتطبيق برنامج جديد في اللغة الإنجليزية على طلاب المدارس الحكومية وحُسب حجم الأثر لهذا البرنامج الجديد ووجد أنه يساوى 0.7 فهذا يعنى أن الفجوة بين برنامج اللغة الإنجليزية الجديد وبرنامج اللغة الإنجليزية بالمدارس الخاصة تقلصت إلى 0.2 بدلا من أن كانت 0.9 . أي أن البرنامج الجديد أكثر فاعلية من برنامج المدارس الحكومية.

ثالثاً- مقارنة حجوم الأثر المحسوبة بحجوم الأثر من دراسات مشابهة لها

إن إحدى الطرق المستخدمة لمقارنة حجوم الأثر التي يقوم بحسابها الباحثون سواء في دراسات أولية أو في دراسات التحليل البعدي هي مقارنتها ببحوث أو دراسات مشابهة لها من كافة الوجوه (من حيث العينات، التصميم البحثي، المعالجات ...الخ). أي بمقارنة بحوثهم ببحوث عالية الجودة. ومثل هذه الدراسات تتطلب من الباحثين بذل جهود كبيرة في البحث عن الدراسات وتحليلها ومقارنة نتائجها. كما يمكن للباحثين أيضاً مقارنة نتائجهم بنتائج البحوث المشابهة التي أجريت في مجال التحليل البعدي التي أجريت على عدد كبير من البحوث والدراسات الأولية. فعلى سبيل المثال قام (Hattie, 2009) بإجراء دراسات على أكثر من 800 بحث في التحليل البعدي اشتملت على أكثر من 50000 بحث أولي في التحصيل الدراسي في مجالات تربوية مختلفة. وكمثال نذكر ما قام به (Hattie) عند تحليل 6 بحوث في التحليل البعدي هدفت إلى دراسة أثر خرائط المفاهيم في التحصيل الدراسي. إحتوت الدراسات الست في التحليل البعدي على 287 دراسة أولية اشتملت على 332 حجم أثر نتج عنها متوسط لحجوم الأثر يساوي 0.57 (Hattie, 2009, pp. 168-169). فإذا أجرى باحث آخر تحليلاً أوليا أو تحليلاً بعدياً لدراسات مشابهة لتلك التي أجراها هاتي (Hattie) وجاء متوسط حجوم الأثر في هذه الدراسات أكبر من 0.57 فيمكنه القول أن خرائط المفاهيم في بحثه ذات فاعلية أعلى من تلك التي قام بها هاتي (Hattie) وبالعكس إذا جاء متوسط حجم الأثر لنتائج الباحث أقل من 0.57 فهذا يعني أن فاعلية خرائط المفاهيم أقل من فاعلية خرائط المفاهيم في دراسات هاتي (Hattie).

النقد الموجه لاستخدام حجوم الأثر

عرض عدد من الباحثين لأوجه القصور في استخدام حجوم الأثر في البحوث والدراسات والتي قد تتسبب في عدم تفسير النتائج تفسيراً صحيحاً وفيما يلي سنعرض لخمسة منها كما جاءت في (حسن، 2008)

1. تصميم البحث وعلاقته بحجوم الأثر .
2. غياب موجهات محددة تساعد الباحثين على الاختيار بين حجوم الأثر .
3. الاعتماد على مؤشرات حجوم الأثر وحدها في تفسير النتائج .
4. انتهاكات فروض النموذج الاحصائي .
5. ثبات درجات المقاييس المستخدمة .

أولا: تصميمات البحوث وعلاقتها بحجوم الأثر:

يستخدم الباحثون في تحليلاتهم الإحصائية أساليب مختلفة تختلف باختلاف تصميم البحث . ففي التصميمات التجريبية مثلاً قد يستخدم الباحث تصميم العامل الواحد Single Factor Design أو التصميمات العاملية Factorial Designs وتصميم العامل الواحد مثلاً قد يشتمل على عينات مستقلة ويسمى حينئذ بالتصميم العشوائي الكامل Completely Randomized Design ، كما قد يشتمل على عينة واحدة يكرر القياس عليها ويسمى في هذه الحالة بتصميم القياس المتكرر Repeated Measure Design . والأمر كذلك أيضاً عندما يستخدم الباحث التصميمات العاملية والتي يدرس فيها الباحث أثر عاملين مستقلين أو أكثر على عامل تابع واحد وهي على أنواع منها : تصميمات البحوث البعدية أو الارتباطية EX- post – facto Designs وتصميمات التجميع العشوائي Randomized Block Designs والتصميمات الهرمية Hierarchical Designs وغير ذلك من التصميمات (علام ، 1993 أ؛ فهمي ، 2005 ؛ Field, 2005) . وبناءً عليه فإن مقدار حجم الأثر الناتج من تحليل التباين في كل تصميم تجريبي من هذه التصميمات سوف يختلف الأمر الذي يترتب عليه عدم دقة المقارنة بين مقادير حجوم الأثر الناتجة من استخدام تصميمات مختلفة . ولتوضيح ذلك دعنا نعرض للأمثلة الآتية :

1- إن مقدار حجم الأثر الناتج من تحليل التباين باستخدام التصميمات التجريبية تامة العشوائية أقل من مقدار الدلالة العملية الناتج من تحليل التباين باستخدام التصميمات التجريبية ذات القياس المتكرر ، وذلك لأن تباين الخطأ بين المجموعات Between Group Error Variance في التصميمات تامة العشوائية أكبر من تباين الخطأ داخل المجموعات Within Group Error Variance في التصميمات ذات القياس المتكرر (Olejnic & Algina, 2000; p. 280).

2- أن مقدار حجم الأثر الناتج من استخدام الباحث لتصميم التجميع العشوائي Randomized Block Design والذي يحتوي على متغيرين أحدهما يمثل معالجة تجريبية (طرق التدريس مثلاً) والآخر يمثل فروقاً فردية (الاعتمادية في مقابل الاستقلالية مثلا) يكون أكبر من مقدار حجم الأثر الناتج عن استخدام الباحث التصميم تام العشوائية Completely Randomized Design والذي يحتوي على متغير مستقل واحد (طرق التدريس مثلاً) (Olejnic & Algina, 2000, P. 243-253). ولقد فسر هذان الباحثان هذه الزيادة في مقدار حجم الأثر عند استخدام تصميم التجميع العشوائي ذى المتغيرين بوجود عامل الفروق الفردية الذي يؤدي إلى تقليل قيمة التباين المشترك Pooled Variance المستخدم في حساب قيمة حجم الأثر في هذه الحالة مقارنة بقيمة التباين المشترك عند حساب حجم الأثر إذا لم يدخل عامل الفروق الفردية في

التصميم . وبناء على ذلك فإن هذين الباحثين يقترحان استبعاد عامل الفروق الفردية من معادلة حساب حجم الأثر إذا كان لدى الباحث عاملان أحدهما يمثل معالجة والآخر يمثل فروقاً فردية.

3- بالإضافة إلى ما سبق فإن باحثين آخرين مثل "او جرادي" O'Grady وكيرك Kirk لاحظا أن مقدار حجم الأثر قد يزيد أو ينقص بزيادة عدد المقارنات في التصميم الواحد . ولتوضيح حالة نقص مقدار حجم الأثر بزيادة عدد المعالجات (المقارنات) فقد أعطى "أو جرادي" (O'Grady, 1982, p. 773) مثالاً لتجربة تحتوي على مجموعتين تجريبيتين ، عدد أفراد كل منهما "10" ومتوسط المجموعة الأولى "10" ومتوسط المجموعة الثانية "18" والانحراف المعياري المشترك لهما "2" ، حسب الباحث قيمة "η^2" في هذه الحالة فوجدها 0.82 لكن عندما أصبحت هاتان المجموعتان جزءاً من عامل مستقل يحتوي على خمس مجموعات (أي بإضافة ثلاث مجموعات جديدة عدد أفراد كل مجموعة منها "10" ومتوسطاتها على الترتيب: 12 ، 14 ، 16 مع ثبات الانحراف المعياري المشترك) ، وجد أن قيمة (η^2) والتي تمثل مقدار حجم الأثر للمجموعات الخمس Omnibus Effect Size انخفضت وأصبحت 0.69 وهذا يعني أن قيمة حجم الأثر انخفضت بزيادة عدد المقارنات.

وعلى العكس من ذلك فقد أوضح كيرك (Kirk, 1995, p.261) أن قيمة الدلالة العملية الكلية Omnibus Effect Size يمكن أن تزيد بزيادة عدد المعالجات (المقارنات) ، ففي تجربة افتراضية كان لديه مجموعتان عدد أفراد كل منهما 8 ومتوسط الأولى 16 ومتوسط الثانية 17 والانحراف المعياري المشترك للمجموعتين 2.5 وبحساب قيمة (η^2) وجد أنها تساوي 0.04 وهي قيمة صغيرة جداً ولكن عندما أضاف إلى هاتين المجموعتين مجموعة ضابطة عدد أفرادها أيضاً 8 لكن متوسطها يساوي 6 مع ثبات قيمة الانحراف المعياري المشترك ، وبحساب قيمة (η^2) للمجموعات الثلاث أصبحت هذه القيمة مساوية 0.81 وهي قيمة كبيرة جداً . وهذا يعني أن مقدار حجم الأثر يتغير بزيادة أو نقصان عدد المقارنات (المعالجات) الأمر الذي يتطلب أن يركز الباحثون على المقارنات المطلوبة فقط عند حساب مقدار حجم الأثر وليس على مقدار حجم الأثر الكلي Omnibus Effect Size ، وذلك حتى لاتكون تفسيراتهم مضللة للقارئ .

مما سبق يمكن أن نلاحظ أن مقدار حجم الأثر يتغير من تصميم تجريبي إلى تصميم تجريبي آخر ، وأنه يتغير بالزيادة في حالة وجود عامل يمثل الفروق الفردية ، كما أنه يزيد أو ينقص بزيادة عدد المقارنات الأمر الذي قد يؤثر بصورة أو بأخرى في تفسير الباحثين لنتائج بحوثهم أو عند مراجعة الدراسات السابقة ، أو حتى في دراسات التحليل البعدي والذي يقارن فيه الباحثون مقادير حجوم الأثر دون النظر إلى أنواع التصميمات التجريبية المستخدمة الأمر الذي يؤدي إلى خلط الأوراق أو كما يسميها ناقدو دراسات التحليل البعدي بمشكلة البرتقال والتفاح .

ثانيا: غياب موجهات محددة تساعد الباحثين على اختيار مؤشر حجم الأثر المناسب :

أشار كيرك (Kirk, 1996) إلى أنه يوجد الآن ما يقرب من واحد وستين مؤشراً لحساب قيم حجم الأثر للنتائج ، لكن لا توجد ارشادات أو موجهات تبين للباحثين أفضلية اختيار مؤشر دون غيره من المؤشرات مما قد يربك الباحثين خاصة المبتدئين ، وقد يجعلهم يميلون إلى اختيار المؤشر

الذي يجعل نتائجهم تبدو أكثر أهمية من الناحية العملية ، وإهمال المؤشر الذي لا يظهر أهمية هذه النتائج ، بل إن بعض الباحثين قد يستخدمون المؤشر نفسه بطريقة تظهر نتائجهم على النحو الذي يفضلون ، وللدلالة على الحالة الأخيرة نورد المثال الآتي والمأخوذ عن جلاس ، وماكجرو ، وسميث (Glass et al, 1981, as cited in Olejnik & Algina, 2000, p. 246) والذي افترضوا فيه مجموعتين احداهما تجريبية متوسطها 52 وانحرافها المعياري 2 ، وأخرى ضابطة متوسطها 50 وانحرافها المعياري 10 وبحساب مقدار الدلالة العملية "Δ " باستخدام الانحراف المعياري للمجموعة الضابطة (كما أوصى جلاس وزملاؤه) نحصل على دلالة عملية قدرها 0.2 وهي قيمة صغيرة أما إذا استخدم الباحث الانحراف المعياري للمجموعة التجريبية عند حساب قيمة " Δ " فسيحصل على دلالة عملية كبيرة مقدارها 1 وهذه القيمة قد تخدع القارئ ، كما قد تخدع الباحث المبتدئ.

ومن الأمور التي قد تخدع الباحث المبتدئ أيضاً تصنيف مؤشرات الدلالة العملية إلى مؤشرات للفروق المعيارية بين المتوسطات (g , Δ , d) ومؤشرات للتباين المفسر (r^2, ω^2, η^2) مما يوحي بأن استخدام مؤشر ما من أحد التصنيفين هو أفضل عند تفسير النتائج من مؤشر آخر في التصنيف الثاني وهو أمر غير صحيح إذ إنه يمكن تحويل أي مؤشر إلى مؤشر آخر فمثلاً يمكن تحويل مؤشر d إلى r باستخدام العلاقة $r = \frac{d}{\sqrt{d^2 + 4}}$ كما يمكن تحويل مؤشر r إلى d باستخدام العلاقة $d = \frac{2r}{\sqrt{(1 - r^2)}}$ (Thompson, 2002).

ومن المهم ملاحظة أن القيم المعيارية التي وضعت من قبل المتخصصين للحكم على مقادير حجوم الأثر المحسوبة من البيانات تختلف من متخصص إلى آخر دون وجود ما يرشد الباحث إلى أفضلية معيار عن معيار آخر ، فعلى سبيل المثال فإن قيم الفرق المعياري بين متوسطين يمكن الحكم عليها وفق معايير كوهين (Cohen, 1988) أو معايير ماكلين وزملائه (Mclean et al, 2000 كما هو موضح بالجدول الآتي.

القيم المعيارية للحكم على مقدار الدلالة العملية للفرق بين متوسطين

المؤشر	ضعيف	متوسط	كبير
كوهين d	0.2	0.5	0.8
ماكلين وزملاؤه(2000)	0.5	1 – 0.5	أكبر من 1

وعند استخدام المؤشرات التي تعتمد على معاملات الارتباط ، فإن الأمر يبدو أكثر إرباكاً، وكما يظهر من الجدول التالي (أنظر الصفحة التالية) والمأخوذ من الجدول الذي عرضه كوترليك ووليامز (Kotrlik & Williams, 2003) نلاحظ في هذا الجدول اختلاف قيم معاملات الارتباط التي تمثل معايير للحكم على مقدار الدلالة العملية والتي يفضلها بعض الباحثين عن استخدام مؤشرات الدلالة العملية لكوهين d .

ثالثا: الاعتماد على مؤشر حجم الأثر وحده في تفسير النتائج :

لعل من أهم أوجه النقد التي وجهت عند تفسير النتائج باستخدام اختبار الدلالة الإحصائية هي أن الباحث يكون أمام خيارين لا ثالث لهما ، فهو إما أن يرفض الفرض الصفري ، وإما أن يفشل في رفضه وذلك بناء على مقارنته للقيمة الاحتمالية المحسوبة للدلالة الإحصائية (القيمة الإحتمالية P) بقيمة الفا (α) المحددة سلفاً والتي قد تكون 0.05 أو 0.01 وهذا يعني –كما يقول كيرك-(kirk, 1996, p. 748) تحويل متصل من الاحتمالات إلى قرار ثنائي يجعل الباحثين يختلفون عند تفسير النتائج ، فقد يحصل باحثان على النتائج نفسها (الدلالات الإحصائية نفسها) لكنهما قد يتوصلان إلى استنتاجات مختلفة وذلك بسبب الاختلاف الطفيف في حجم العينات ، أو اختلاف تصميم البحث كما يقرر تومسون وزملاؤه. (Thompson et al, 2005, p. 186)

قيم معاملات الارتباط التي تستخدم كمعايير للحكم على الدلالة العملية

اسم الباحث	المؤشر	القيم المعيارية	التفسير
كوهين	معامل الارتباط	0.01	ضعيف
(Cohen, 1988)	r	0.03	متوسط
		0.05	كبير
دافيز	معامل الارتباط	0.09-0.01	ضعيف جداً
(Davis, 1971)	r	0.29-0.10	ضعيف
		0.49-0.30	متوسط
		0.69-0.50	كبير
		0.70 فأكثر	كبير جداً
هنكل ويرزما، وجورز	معامل الارتباط	0.30-0.00	ضعيف جداً
(Hinkle, Wiersma &	r	0.50-0.30	ضعيف
Jurs, 1979)		0.70-0.50	متوسط
		0.90-0.70	كبير
		1.00 -0.90	كبير جداً

إن هذه الصرامة في اتخاذ هذا القرار الثنائي (رفض أو الفشل في رفض الفرض الصفري) هي نفسها التي يستخدمها بعض الباحثين عند استخدام معايير الحكم على قيم حجوم الأثر بكونها كبيرة أو متوسطة أو ضعيفة. أي تحولنا من قرار ثنائي إلى قرار ثلاثي (Thompson et al, 2005 p. 186) إذ يكتفي كثير من الباحثين بحساب مقدار حجم الأثر والإشارة إليه بكونه كبيراً أو متوسطاً أو صغيراً دون تفسير نتائجهم في ضوء سياق البحث، أو الدراسات المشابهة ، أو أهداف البحث وتصميمه أو كل هذه الجوانب مجتمعة مما يجعل تفسيراتهم لنتائج بحوثهم قاصرة.

ولتوضيح قصور تفسير النتائج بالاعتماد على مؤشر حجم الأثر وحده فقد عرض ليفين (Levin, 1998, p. 45-46) لتجربة أجريت على مجموعتين من كبار السن الأولى ضابطة (أ) والثانية تجريبية (ب1) عدد أفراد كل منهما 3 أفراد ، وكان العامل المستقل هو معالجة تساعد أفراد المجموعة

التجريبية (ب1) على تذكر 10 مهام مختلفة مثل: تناول الدواء في موعده ، الاتصال بالطبيب ... الخ بينما لم تتعرض المجموعة الضابطة إلى أية معالجة وجاءت النتائج كما يلي:

	المجموعة الضابطة	المجموعة التجريبية
	1	3
	1	3
	2	4
المتوسط	1.33	3.33
الإنحراف المعياري	0.577	0.577

وقد حسب الباحث مقدار حجم الأثر بطريقتين : مرة باستخدام مؤشر كوهين "*d*" وكانت قيمته 3.5 ومرة أخرى باستخدام (η^2) وكانت قيمتها 0.82 والمؤشران يعبران عن حجم أثر كبير . لكن عندما قارن الباحث المجموعة الضابطة (أ) بمجموعة تجريبية أخرى ودعنا نشير إليها بالرمز (ب2) مكونة من 3 أفراد أيضاً وهذه المجموعة تعرضت لمعالجة تجريبية مخالفة لتلك التي تعرضت لها المجموعة (ب1) جاءت النتائج كما يلي :

	المجموعة الضابطة	المجموعة التجريبية
	1	5
	1	8
	2	10
المتوسط	1.33	7.667
الإنحراف المعياري	0.577	2.517

وبحساب مقدار حجم الأثر في هذه الحالة وجد أن قيمة "*d*" 3.5 وأن قيمة (η^2) 0.82 وهي نفس القيم في التجربة السابقة وبناء على ذلك فإن الباحث غير المتمرس سوف يكتفي بالقول بأن مقدار حجما الأثر كبيران وأن المعالجة فعالة بنفس المقدار في كلتا التجربتين. وهذا التفسير قاصر بطبيعة الحال ، ذلك أن الفاحص المتأني لكل من التجربتين يلاحظ أن المجموعة التجريبية (ب1) يتذكر أفرادها في المتوسط 3 مهام من 10 ، بينما يتذكر أفراد المجموعة التجريبية (ب2) 8 مهام من 10 تقريباً . وهذا يعني أن المعالجة التجريبية في (ب2) أكثر فاعلية من المعالجة التجريبية في (ب1) رغم تساوي مقدار حجوم الأثر لكل منهما ، وهو ما يدعو إلى القول بأن الاعتماد على مقدار حجم الأثر وحده في تفسير النتائج لا يغني بحال من الأحوال عن النظر والتفكر في البيانات الأصلية كي تكون تفسيراتنا أكثر دقة.

رابعا: انتهاك الفروض التي يستند إليها النموذج الاحصائي :

إن استخدام الاحصاء البارامتري (الاستنتاجي) كما هو معلوم يتطلب التأكد من عدم انتهاك الشروط التي يستند إليها النموذج الاحصائي المستخدم لتحليل البيانات ، ذلك أن انتهاك واحد أو أكثر من هذه الشروط يؤدي إلى أن يقع الباحثون في تفسيرات خطأ لنتائج التحليل، ولعل من بين أهم هذه الانتهاكات التي تؤثر بصورة أو بأخرى على قيم حجوم الأثر المحسوبة ، هي تلك المتعلقة

بأخطاء المعاينة ، واعتدالية توزيع الدرجات وتجانس التباينات في مجتمعات الدراسة . وفيما يلي عرض لأثر هذه الانتهاكات على قيم حجوم الأثر للنتائج .

1- خطأ المعاينة

عند اختبار الفروض الإحصائية لا يجري الباحث دراسته على المجتمع ، وإنما يلجأ إلى اختيار عينة تمثل هذا المجتمع تمثيلاً صادقاً ، ثم يجمع بياناته من هذه العينة ، ويحللها باستخدام أسلوب إحصائي مناسب للاستنتاج منها والتعميم على المجتمع ، وهنا يختلف التقدير الذي يحصل عليه الباحث من العينة عن القيمة الموجودة في المجتمع تحت نفس الظروف . والاختلاف بين هاتين القيمتين يرجع إلى عدد من العوامل لعل من أهمها خطأ المعاينة Sampling Error وكلما زاد حجم العينة كلما قل خطأ المعاينة وزادت ثقتنا بالنتائج ، أي زادت فرصة الباحث في الحصول على نتائج دالة احصائياً عند اختبار فروض بحثه باستخدام أي من الطرق المعروفة لاختبار دلالات الفروق ، أو قوة العلاقة بين المتغيرات. ومن هنا وجه عدد من الباحثين النقد لاختبارات الدلالة الإحصائية ، باعتبار أنه يمكن الحصول على الدلالة الإحصائية للنتائج لا بسبب وجود فروق مهمة ، أو علاقة قوية بين المتغيرات وإنما بسبب كبر حجم العينة المستخدمة في البحث (Cohen, 1994; Fan, 2001; Kirk, 1996; Thomson, 1999)

وإذا نظرنا إلى مؤشرات حجوم الأثر، فإننا نلاحظ أن قيم هذه الحجوم التي يقوم الباحث بحسابها إنما هي تقدير لقيمتها في المجتمع الذي سحبت منه عينة بحثه ، وبهذا المعنى فإن هذا التقدير يتأثر هو الآخر بأخطاء المعاينة ، فكما أوضح (Fan, 2001, p. 278) في دراسته الافتراضية "Mont Carlo Study" أنه كلما زاد حجم العينة المستخدمة في حساب مقدار الأثر كلما قل خطأ المعاينة (قل تباين الدرجات) ومن ثم زادت قيمة حجم الأثر المحسوبة. وأنه كلما صغر حجم العينة زاد خطأ المعاينة (زاد تشتت الدرجات) ومن ثم قلت قيمة حجم الأثر المحسوبة من هذه العينة ، وهو ما أكده تومسون من قبل في عام 1999 ، حيث رأى أن التحليلات الإحصائية التقليدية تتأثر بتباين درجات العينات ، ومن ثم تكون مؤشرات الدلالة العملية والتي تعتمد على هذه التباينات متحيزة (Thompson, 1999, P.171) ، لذا فقد اقترح تومسون أن يتم حساب مقدار الدلالة العملية في هذه الحالات من المعادلات التي تُعدل قيم هذا المقدار.

والنتيجة التي يمكن الخروج بها هنا تناقض إعتقاد بعض الباحثين بأن مقدار حجم الأثر لا يتأثر بحجم عينة البحث (انظر مثلاً: نصار ، 2006 ، ص 44-48 ومنصور ، 1997 ، ص59) ولا يفوتنا هنا أن نؤكد أيضاً أن مقدار حجم الأثر يتأثر إذا كانت العينة المستخدمة في البحث غير عشوائية – وهو ما يحدث في كثير من البحوث التربوية والنفسية- الأمر الذي يتطلب أن يلجأ الباحثون إلى استخدام نطاق الثقة (فترات الثقة) Confidence Intervals بالإضافة إلى حساب قيم حجم الأثر والدلالة الإحصائية لإعطاء بحوثهم قدراً أكبر من الموثوقية كما ينصح بذلك تومسون (1999).

2- اعتدالية توزيع الدرجات وتجانس التباينات في المجتمعات

عند استخدام مؤشرات حجوم الأثر لا بد للباحث أن يدرك أن هذه المؤشرات تتأثر بفرضيتي اعتدالية توزيع درجات المتغير التابع ، وتجانس تباينات درجات هذا المتغير في

المجتمع أو المجتمعات التي يريد الباحث أن يعمم نتائجه عليها. وإذا كان توزيع الدرجات ملتوياً بشدة ، وكانت التباينات غير متجانسة فإن مقدار حجم الأثر المحسوب لهذه العينات لن يكون دقيقاً ، فلقد أجرى "هوجرتى وكرومرى" (Hogarty & Kromry, 2001) دراسة افتراضية Mont Carlo Study أظهرت نتائجها أن أكثر مؤشرات حجم الأثر شيوعاً في البحوث التربوية هما مؤشرا كوهين "*d*" وهدجز "*g*" ، وأن هذين المؤشرين هما أكثر المؤشرات حساسية عندما لا يتوفر شرطا الاعتدالية وتجانس التباينات خاصة إذا كان حجم العينات المستخدمة في الدراسة صغيراً.

ومن المعروف أن مؤشرا كوهين "*d*" وهدجز "*g*" يستخدمان متوسطى المجموعتين التجريبية والضابطة في بسط معادلة تحديد مقدار حجم الأثر ، والمتوسطات بطبيعتها تتأثر بالدرجات المتطرفة ، خاصة عند صغر حجم العينة مما يحدث إنتهاكاً واضحاً لشرطي الاعتدالية وتجانس التباينات الأمر الذي ينتج عنه تشوه قيمة الفرق بين المتوسطين ، ومن ثم عدم دقة قيمة حجم الأثر الناتجة عن هذين المتوسطين ، وبالمثل أيضاً يجب أن يُلاحظ أن كل أنواع تحليل التباين MANOVA , ANCOVA , ANOVA ... الخ تتطلب توفر هذين الشرطين بالإضافة إلى شروط أخرى بطبيعة الحال ، وإذا لم تتوفر هذه الشروط (الفروض) فإن مقدار حجوم الأثر المحسوبة سواء باستخدام (η^2) أو "ايبسلون^2" "ε^2" أو "أوميجا^2" "ω^2" سوف يتأثر وذلك كنتيجة مباشرة لتأثر خطأ التباين الذين يستخدم في مقام النسبة الفائية (MS error) ، ولذا ينصح (Onwuebegbuzie & Daniel, 2003) أن يستخدم الباحثون في هذه الحالة إحصاءً لابارامترياً ثم بعد ذلك يمكن استخدام مؤشرات الدلالة العملية اللابارامترية مثل مؤشر كريمر وأندروز (Kraemer & Andrews, 1982) الذي يشار إليه بالحرف "Y" أو مؤشر اللغة الدارجة Common Language والذي يشار إليه بالحرفين (CL) والذي اقترحه "ماجرو ووانج" (McGraw & Wang, 1992)

<u>خامسا: ثبات درجات أدوات القياس :</u>

إن من الأمور المهمة التي عادة ما يراعيها الباحثون عند نشر دراساتهم هي التأكد من ثبات درجات الأدوات التي استخدموها لجمع البيانات . وكما هو معروف فإن الثبات يُعبر عنه احصائياً بنسبة التباين الحقيقي للدرجات إلى التباين الكلي ، والتباين الكلي ليس سوى مجموع التباين الحقيقي وتباين الخطأ ، وبناءً على ذلك ، فكلما زادت قيمة تباين الخطأ "خطأ القياس" كلما نقصت قيمة معامل الثبات ، وهذا بدوره يؤثر على مقدار حجوم الأثر التي ينشدها الباحث من إجراء بحثه (Baugh, 2002) .

والمتتبع للبحوث المنشورة في الدوريات الأجنبية والعربية على السواء يلاحظ أن كثيراً من الباحثين يحددون قيم معاملات ثبات درجات الأدوات التي يستخدمونها في جمع بياناتهم وعادة ما تكون معاملات الثبات هذه مرتفعة ومقبولة ، إلا أن الخطأ الشائع الذي لا يفطن إليه كثير من هؤلاء هو أن معاملات الثبات التي يعرضونها في بحوثهم تم حسابها في الأغلب الأعم على عينات استطلاعية . وليس على عينة البحث نفسه ، وهذا الخطأ ناشئ عن عدم إدراك الباحثين أن الثبات هو بالدرجة الأولى صفة من صفات الدرجات التي نحصل عليها من هذه الأدوات (Thompson, 1998) . ونظراً لأن كل عينة من الأفراد لها خصائصها الفريدة ، فهذه العينات يمكن

أن تعطي درجات مختلفة عند تطبيق الأدوات نفسها عليها ، ويترتب على اختلاف هذه الدرجات اختلاف معاملات الثبات. وهذا يعني أنه إذا كان معامل ثبات درجات أداة ما على عينة استطلاعية مرتفعاً ، فقد لا يكون هذا المعامل بالضرورة مرتفعاً على عينة البحث الأصلية ، وإذا حدث وكان معامل الثبات على عينة البحث منخفضاً – مقارنة بالعينة الاستطلاعية - فإن ذلك سيؤثر سلباً على قوة الاختبار الاحصائي Statistical Power وفي الوقت نفسه سيؤثر سلباً على مقداري الدلالة الإحصائية وحجوم الأثر.

هذا وقد أشار بعض الباحثين ومنهم (Onwuegbuzie, & Daniel, 2000) إلى أن درجات أداة ما قد تكون ثابتة عندما تطبق الأداة على عينة البحث الكلية أما إذا وزعت هذه العينة إلى عينات فرعية (كما يحدث عندما توزع عينة من الأفراد وفق متغير الجنس ، أو وفق مستويات مختلفة لطرق التدريس) فقد يحدث أن يكون ثبات درجات الأداة على إحدى العينات الفرعية عال نسبياً ، وعلى العينة أو العينات الأخرى في البحث منخفض نسبياً ، وفي هذه الحالة سيحصل الباحث على نتائج مختلفة للدلالة الإحصائية ولحجم الأثر ، وبمعنى آخر فإن مقارنة درجات مجموعات فرعية ذات قيم مختلفة لمعاملات الثبات يمكن أن يؤثر في معدل خطأ النوع الأول أو خطأ النوع الثاني ، ومن ثم في مقدار حجم الأثر ، لذا يُنصح الباحثين دائماً بأن يحسب الثبات على عينة البحث ، وعلى العينات الفرعية للبحث إن وجدت. ووفق رأي هذين الباحثين فإنه من غير المرغوب فيه مقارنة حجوم الأثر لدراسات مختلفة ببعضها (كما يحدث عادة في استخدام أساليب التحليل البعدي "Meta Analysis") دون الأخذ بالاعتبار معاملات ثبات درجات الأدوات المستخدمة في هذه الدراسات .

ومن المهم أيضاً في هذا الصدد الإشارة إلى ما أشار إليه كل من (Nunnally & Brenestin, 1994) من أن معامل ثبات الدرجات يمثل الحد الأقصى الذي يجب ألا يتعداه مقدار حجم الأثر المحسوب من هذه الدرجات. ولبيان هذه النقطة افترض أنه في تجربة ما كان مقدار حجم الأثر d كما حسب من معادلة كوهين يساوي 2 وأن معامل ثبات درجات الاختبار الذي طبق على عينة البحث في هذه التجربة كان مساوياً 0.83 وبتحويل قيمة حجم الأثر من قيمة الفرق المعياري "d" إلى قيمة حجم الأثر معبرا عنها بمعامل الارتباط "r" فإن هذه القيمة ستكون

$$r = \frac{d}{\sqrt{d^2+4}} = \frac{2}{\sqrt{8}} = 0.707$$

ونلاحظ هنا أن قيمة معامل الارتباط المحسوبة أعلاه والتي تمثل الدلالة العملية لم تتخط قيمة معامل الثبات وهو 0.83

ملخص

ناقشنا عدداً من أوجه القصور التي تصاحب استخدام حجوم الأثر في البحوث التربوية والنفسية ، فعلى الرغم من أن دراسات عديدة بينت مزايا استخدام مؤشرات حجوم الأثر إلا أن عدداً قليلاً من تلك الدراسات تعرض لأوجه القصور والتي منها أولاً: علاقة تصميمات البحوث بمقدار حجم الأثر الناتج من تصميم بحثي مقابل تصميم بحثي آخر ، ومنها ثانياً: غياب موجهات محددة تساعد الباحثين على اختيار مؤشر حجم الأثر دون غيره من المؤشرات ، ومنها ثالثاً: أن الاعتماد على مؤشر حجم الأثر وحده لا يؤدي إلى تفسير جيد لنتائج البحث ، ومنها رابعاً: أن انتهاكات الفروض التي تستند إليها النماذج الإحصائية المستخدمة لتحليل البيانات يؤدي بالضرورة إلى حدوث

أخطاء في قيمة حجوم الأثر ومن ثم تفسير هذه القيمة ، وأخيراً فإن عدم ثبات أدوات القياس (أخطأ القياس) تؤدي أيضاً إلى الحصول على نتائج مضللة وحجم أثر غير دقيق.

وفي ضوء ما سبق يمكن التوصل إلى ما يلي :

1. أن معظم من أيدوا استخدام حجوم الأثر في البحوث ركزوا على الجوانب الإيجابية لهذا الاستخدام ولم يتعرضوا للجوانب السلبية فيه ، وهذا قد يعني بالنسبة لهم أن مقدار حجوم الأثر سوف يعالج جوانب النقص التي ظهرت في حساب مقدار الدلالة الإحصائية وهو رأي غير دقيق .
2. إن اختلاف التصميمات التجريبية يتطلب استخدام معادلات مختلفة لحساب قيم حجوم الأثر، الأمر الذي يجعل مقارنة قيم حجوم الأثر الناتجة عن تصميمات مختلفة أمراً يحتاج إلى إعادة النظر فيه.
3. أن تعدد المقارنات يجب أن يؤخذ في الاعتبار، فكما سبق القول يجب أن يركز الباحث على المقارنات المطلوبة فقط عند حساب قيمة حجوم الأثر وليس على المقارنات ككل.
4. أن و جود عامل أو أكثر للفروق الفردية في تصميم تجريبي يؤثر على قيم حجوم الأثر المحسوبة ومن ثم يجب أن يؤخذ ذلك أيضاً في الاعتبار .
5. يجب أيضاً أن يحدد الباحث المعايير التي استخدمها للحكم على قيمة حجوم الأثر سواء استخدم مؤشرات للفرق المعياري بين المتوسطات أو استخدم مؤشرات التباين المفسر أو معاملات الارتباط .
6. يجب أيضاً أن تفسر نتائج البحوث في ضوء كل من الدلالة الإحصائية للنتائج والدلالة العملية لها ، كما يجب أن يفسر الباحثون هاتين الدلالتين في ضوء سياق البحث ، والدراسات السابقة وأهداف البحث وتصميمه . بل إن بعض الباحثين ذهب – كما أشرنا - إلى استخدام نطاق الثقة (فترات الثقة) أيضاً عند تفسير النتائج .
7. عند استخدام الاحصاء البارامتري يجب أن يتأكد الباحث من عدم انتهاكه لفروض النموذج الاحصائي المستخدم فإذا لم ينتهك الباحث أيا من هذه الفروض يمكنه حساب مقدار حجم الأثر وهو مطمئن لنتائجه ، أما إذا حدث إنتهاك لهذه الفروض فعليه أن يستخدم الإحصاء اللابارامتري ومن ثم مؤشرات حجوم الأثر اللابارامترية .
8. إن مشكلة أخطاء القياس والمتعلقة باستخدام أدوات القياس يجب أن تؤخذ أيضاً في الاعتبار، فحساب ثبات درجات أداة البحث على عينة استطلاعية لا يكفي وإنما يجب حساب ثبات درجات الأدوات على عينة أو عينات البحث نفسه إذ قد تتغير قيم معامل الثبات بتغير العينات كما أشرنا إلى ذلك من قبل.

توصيات

وفق ماتم استعراضه يمكن أن نقترح فيما يلي عدداً من التوصيات الموجهة إلى كل من رؤساء تحرير الدوريات العربية ومؤلفي كتب الإحصاء التربوي والنفسي وكتب مناهج البحث والباحثين:

رؤساء تحرير الدوريات العلمية :

اننا ندعوا رؤساء تحرير الدوريات العربية إلى القيام بمسئولياتهم تجاه توضيح القضايا الإحصائية وخاصة تلك المتعلقة بأوجه القصور التي يمكن أن يقع فيها الباحثون عند استخدام التحليلات الإحصائية بوجه عام وحساب وتفسير قيم حجوم الأثر بوجه خاص . وهذا ما يفعله كثير من رؤساء تحرير الدوريات العلمية العالمية . إن توضيح أوجه القصور هذه سوف يفيد الباحثين الذين يودون نشر بحوثهم في هذه الدوريات وغيرها .

مؤلفو كتب الإحصاء التربوي والنفسي كتب ومناهج البحث :

إنه من المهم أن يتحمل مؤلفو كتب الاحصاء ومناهج البحث جزءاً من مسؤولية تثقيف الباحثين وتبصيرهم بكيفية حساب مقادير حجوم الأثر بصورة صحيحة عند استخدام التصميمات الإحصائية بل وتحذيرهم من أوجه القصور التي تواكب استخدام هذه المقادير وكيفية التغلب عليها أو تلافيها .

الباحثون :

لعل المسئولية الأكبر في تجنب أوجه القصور عند استخدام حجوم الأثر تقع على عاتق الباحثين أنفسهم فلا يجب عليهم اتباع طريقة واحدة لحساب قيمة حجم الأثر - كما يلاحظ في كثير من البحوث العربية - أو اتباع معايير محددة للحكم على مقدار حجوم الأثر بالكبر أو الصغر دون الأخذ في الاعتبار أوجه القصور التي أشير إليها هنا وفي دراسات أخرى.

أساتذة الإحصاء التربوي والنفسي ومناهج البحث

لعل من أهم مسئوليات القائمين بتدريس الإحصاء التربوي والنفسي ومناهج البحث أن يبينوا لطلبة الدراسات العليا ليس فقط الأنواع المختلفة لمؤشرات حجوم الأثر وكيفية حسابها وإنما لابد أيضا من بيان مميزات وعيوب استخدام كل مؤشر منها ومن ثم يظهر لهؤلاء الطلبة أن مناهج البحث والإحصاء ليست مجرد خطوات روتينية يجب اتباعها وإنما هي رؤية كلية وعملية تفكر مستمرة .

المراجع

أولاً- المراجع العربية

حسن، عبد المنعم أحمد (2008). أوجه القصور في استخدام مؤشرات الدلالة العملية في البحوث التربوية والنفسية. *دراسات في المناهج وطرق التدريس*، 134، 14 – 39.

السعيد، رضا (2003). حجم الأثر: أساليب احصائية لقياس الأهمية العملية لنتائج البحوث التربوية، *المؤتمر العلمي الخامس عشر للجمعية المصرية للمناهج وطرق التدريس، القاهرة*: 21-22 يوليو، م2، 643-674.

علام ، صلاح الدين محمود (1993 أ). *الأساليب الإحصائية الاستدلالية البارامترية واللابارامترية في تحليل بيانات البحوث النفسية والتربوية*. (ط 1). دار الفكر العربي .

علام، صلاح الدين محمود ،(1993ب). *تحليل البيانات في البحوث النفسية والتربوية*. دار الفكر العربي.

منصور، رشدي (1997). حجم التأثير: الوجه المكمل للدلالة الإحصائية، *المجلة المصرية للدراسات النفسية*، *7*(16) ، 57-75.

نصار، يحيى (2006). استخدام حجم الأثر لفحص الدلالة العملية للنتائج في الدراسات الكمية، *مجلة العلوم التربية والنفسية الصادرة عن كلية التربية، جامعة البحرين*، *7*(2)، 38-59.

ثانيا- المراجع الأجنبية

American Psychological Association (2005). *Publication manual of the American Psychological Association* (5th ed.). American Psychological Association

American Psychological Association (2010). *Publication manual of the American Psychological Association* (6th ed.). American Psychological Association

American Psychological Association (2020). *Publication manual of the American Psychological Association* (7th ed.) https://doi.org./10.1037/0000165-000

Baugh, F. (2002). Correcting effect sizes for Score reliability: A reminder that measurement and substantive issues are linked inextricably. *Educational and Psychological Measurement,62(2),254-263. https://doi.org/10.1177/0013164402062002004*

Borenstein, M., Hedges, L.V. (2019). Effect size for meta- analysis. In H. Cooper, L. V. Hedges, J. C. Valentine (Eds.), *The handbook of research synthesis and meta- analysis* (3rd ed., pp. 208-243). Russell Sage Foundation.

Borenstein, M., Hedges, L.V., Higgins, J.P., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Wiley

Card, N. A. (2012). *Applied mate-analysis for social science research*. The Guilford Press.

Cohen, J. (1988). *Statistical power analysis for behavioral sciences*. Erlbaum.

Cohen, J. (1994). The earth is round ($p < 0.05$). *American Psychologist*, 49, 997-1003. https://psycnet.apa.org/doi/10.1037/0003-066X.49.12.997

Cumming, G. (2012). *Understanding the new statistics: Effect sizes, confidence intervals, and meta- analysis*. Routledge.

Durlak, J. A. (2009). How to select, calculate, and interpret effect sizes. *Journal of Pediatric Psychology, 34*(9). 917- 928.

https://doi.org/10.1093/jpepsy/jsp004

Ellis, P. D. (2010). *The essential guide to effect sizes: Statistical power, meta-analysis, and interpretation of results*. Cambridge University Press

Fan, X. (2001) Statistical significance and effect size in education research: Two sides of a Coin. *The Journal of Educational Research, 94*(5), 275-282. https: doi.org/ 10.1080/00220670109598763

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). Sage.

Glass, G. V., & Hopkin, K. D. (1996). *Statistical methods in education and psychology* (3rd ed.). Allyn and Bacon.

Green, S. B., & Salkind, N. J. (2014). *Using SPSS for Windows and Macintosh: Analyzing and understanding data* (5th ed.). Pearson.

Hattie, J. A. C. (2009*). Visible learning: A synthesis of over 800 meta-analysis relating to achievement.* Routledge.

Hill, C. G., Bloom, H. S., Black, A. R., & Lipsey, M. W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development perspectives, 2*(3), pp. 172-177.

https://doi.org/10.1111/j.1750-8606.2008.00061.x

Hogarty, K., Y. & Kromrey, J., D. (2000, April). *We have been reporting some effect sizes: Can you guess what they mean*? [Paper presentation} American Research Association, Annual Meeting , Seattle, WA.

Huck, W. S. (2012). *Reading statistics and research.* Pearson.

Kirk, R. E. (1995). *Experimental design: Procedures for behavioral sciences* (3rd ed.). Brooks/cole.

Kirk, R. E. (1996). Practical significance. A concept whose time has come. *Education and Psychological Measurement*, 56,746-759.

https://doi.org/10.1177%2F0013164496056005002

Knapp, T. R. (1998). Comments on the statistical significance testing articles. *Research in the Schools*, *5*(2), 39-41

Kotrlik, J. W. & Williams, H. (2003). The incorporation of effect size in information technology, learning, and performance research. *Information Technology, Learning, and Performance Journal, 21*(1), 1-7.

Krantz, D. H. (1999). The null hypothesis testing controversy in psychology. *Journal of the American Statistical Association*, 94, 1372-1381.

Kraemer, H. C., & Andrews, G. (1982). A nonparametric technique for meta-analysis effect size calculation. *Psychological Bulletin*, *91*(2), 404–412. https://doi.org/10.1037/0033-2909.91.2.404

Levin, J. R. (1998). What if there were no more bickering about statistical significance test? *Research in the Schools*, 5,43-53.

Lipsey, M. W., Puzio, K., Yun, C., Hebert, M. A., Steinka- Fry, k., Cole, M. W., Roberto, M., Anthony, K. S., & Busick, M. d., (2012). *Translating the statistical representation of effects of education interventions into more readily interpretable forms*. (NCSER 2013- 3000) National Center for Special Education Research, Institute of Education Sciences, U.S. Department of Education. Retrived June 3, 2021 from http://ies.ed.gov/ncser/.

Lomax, R. G., & Vaughn, D. L. (2012). *Statistical concepts: A second Course* (4th ed.). Routledge.

Marzano, R. S., Pickering, D. G., & Pollock, J. E. (2001). *Classroom instruction that works*. Association for Supervision and Curriculum Development.

Mayer, J. L., Well, A. D. (2003). *Research design and statistical analysis* (2ed ed.). Lawrence Erlbaum Associate.

McGraw, K. N., & Wong, S. P. (1992). A common language effect size statistic. *Psychological Bulletin,111*(2), 361- 365. https://doi.org/10.1037/0033-2909.111.2.361

McLean, J. E.; O'Neal, M. R. & Barnett, J. J. (2000, November). *Are all effect sizes created equal*? [Paper Presentation]. Mid-South Educational Research Association Annual Meeting, Bowling Green, K Y (ERIC Document Reproduction Service No. ED 448188).

Meline, T. & Wang, B. (2004). Effect-size reporting practices in AJSLP and other ASHA Journals, 1999-2003. *American Journal of Speech-Language Pathology*, 13, 202-207. https: doi. org/ 10.1044/1058-0360(2004/021)

Morgan, G.A., Leech, N. L., Gloeckner, G. W., & Barrett, K. C. (2013). *IBM SPSS for introductory statistics: Use and interpretation (5th ed.)*. Roulege.

Nickerson, R. S. (2000). Null hypothesis significance testing. A review of an old and continuing controversy. *Psychological Methods*, 5, 241-301. https: doi. org/ 10.1037/1082-989x.5.2.241

Nix, T. W & Barnett, J. J. (1998). The data analysis dilemma: Ban or abandon. A review of null hypothesis significance testing. *Research in the School*, 5, 3-14.

Nunnally, J. & Berntein, I (1994). *Psychometric theory* (3rd ed.). Mc Grew-Hill.

O'Grady, K. E. (1982). Measures of explained variance: Cautions and limitations. *Psychological Bulletin*, 92, 766-777.

https://psycnet.apa.org/doi/10.1037/0033-2909.92.3.766

Olejnik, S. & Alegina, J. (2000). Measurement of effect size for comparative studies: Applications, interpretations, and limitations. *Contemporary Educational Psychology*, 25, 241-286. Retrieved September. 4, 2007 from http://www.personal.psu.edu/jx614/m554/articles/olejnic&gina2000.pdf

Onwuegbuzie, A, J. & Daniel, L. G. (2000). Reliability generalization: The importance of considering sample specificity, confidence intervals, and subgroup differences [Paper presentation]. Mid-south Educational Research Association Annual Meeting, Bowling Green, KY (ERIC Document Reproduction Service No. ED 448.204).

Rosenthal, R., & Rubin, D. B. (1982). A simple general purpose display of magnitude of experimental effect. *Journal of Educational Psychology*, 73(2), pp. 166-169.https://doi.org/10.1037/0022-0663.74.2.166

Thompson, B. (1996). AERA editorial policies regarding statistical significance testing: Three suggested reform. *Educational Researcher*, *25*(2), 26-30. https://doi.org/10.2307/1176337

Thompson, B. (1998). Five methodological errors in educational research: The pantheon of statistical significance and other faux pas. Invited address presented at the annual meeting of the American Research Association, San Diego: (ERIC Document Reproduction Service No. ED 419-023).

Thompson, B. (1999). If statistical significance tests are broken misused, what practices should supplement or replace them? *Theory & Psychology*, *9*(2), 165-181. https://doi.org/10.1177%2F095935439992006

Thompson, B. (2002). "Statistical", "Practical", and "Clinical": How many kinds of significance do counselors need to consider? *Journal of counseling & Development*,80,64-71. https://doi.org/10.1002/j.1556678.2002.tb00167.x

Thompson, B; Diamond, K; Mc William, R; Snyder, P & Snyder, S. (2005). Evaluating the quality of evidence from correlational research for evidence –based practice. *Exceptional Children, 71*(2), 181-194. https://doi.org/10.1177%2F001440290507100204

Trusty, J. ; Thompson, B.; & Petrocelli, V. (2004). Practical guide for reporting effect size in quantitative research In the Journal of counseling & Development. *Journal of Counseling & Development,* 82, 107-110. *https://psycnet.apa.org/doi/10.1002/j.1556-6678.2004.tb00291.x*

Valentine, J. C., Aloe, A. M., Wilson, S. J. (2019). Interpreting effect sizes. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta- analysis* (3rd ed., pp. 433-452). Russell Sage Foundations

ملحق 1

بعض تحويلات حجم الأثر

عادة ما تستخدم كثير من البحوث الدراسات المنشورة في الدوريات العلمية أو رسائل الماجستير والدكتوراة أو المقدمة في مؤتمرات علمية أنواعاً مختلفة من حجوم الأثر للإجابة عن سؤال محدد يحاول الباحث الإجابة عنه. فعلى سبيل المثال قد تستخدم بعض الدراسات والبحوث مؤشر (d) الذي يشير إلى الفرق المعياري بين المتوسطات بينما تستخدم دراسات وبحوث أخرى مؤشر (r) ويريد الباحث التعامل مع المؤشرين. فعلى سبيل المثال إذا كانت بعض الدراسات تختبر أثر طريقة تدريس على إكتساب تلاميذ الصف الأول الإبتدائي لمهارات الجمع واستخدم باحثو هذه الدراسات مؤشر حجم الأثر (d) أما في دراسات أخرى فقد استخدم الباحثون حجم الأثر (r) للإجابة عن السؤال نفسه ، من هنا يحتاج الباحث لمعرفة كيف يمكن تحويل أي مؤشر من هذه المؤشرات إلى مؤشر أخر. وفيما يلي سنعرض لكيفية تحويل هذه المؤشرات لبعضها بعضا كما إقترح كل من (Bornstein, Hedges, Higgins, & Rothstein, 2009; Durlak, 2009; Mayer & Well, 2003)).

<u>**التحويل من معامل الارتباط (r) إلى الفرق المعياري بين متوسطين**</u>

يمكن إجراء هذا التحويل باستخدام المعادلة (1) التالية

$$d = \frac{2r}{\sqrt{1-r^2}} \qquad (1)$$

كما يمكن حساب قيمة تباين حجم الأثر من المعادلة (2) التالية:

$$V_d = \frac{4V_r}{(1-r^2)^3} \qquad (2)$$

مثال:

إذا كانت قيمة معامل الإرتباط (r=0.50) وقيمة تباين معامل الإرتباط (V_r=0.0058) فحول هذه القيم إلى ما يناظرها من قيم (d)

الحل:

$$d = \frac{2(0.50)}{\sqrt{1-0.50^2}} = 1.1547$$

$$V_d = \frac{4(0.0058)}{(1-0.50^2)^3} = 0.0550$$

<u>**التحويل من الفرق المعياري بين متوسطين (d) إلى معامل الإرتباط (r)**</u>

يتم هذا التحويل باستخدام المعادلة (3) التالية:

$$r = \frac{d}{\sqrt{d^2 + a}} \qquad (3)$$

حيث (a) هي معامل للتصحيح في حالة ما إذا كانت ($n_1 \neq n_2$) ويعطي هذا المعامل من المعادلة التالية:

$$a = \frac{(n_1 + n_2)^2}{n_1 n_2} \qquad (4)$$

وهذا المعامل تتوقف قيمته على النسبة بين (n_1) و (n_2) وليس على القيمة المطلقة لكل منهما، وإذا كانت هذه القيم غير معروفة فنعتبر ($n_1 = n_2$) ونستخدم في هذه الحالة ($a = 4$) .

أما تباين (r) فيمكن الحصول عليه من المعادلة (5) التالية

$$V_r = \frac{a^2 V_d}{(d^2 + a)^3} \qquad (5)$$

مثال:

إذا كانت ($n_1 = n_2$) و (d=1.1547) و (V_d=0.0550) فإن

$$r = \frac{1.1547}{\sqrt{1.547 + 4}} = 0.50$$

$$V_r = \frac{4^2\,(0.055)}{\left(1.574^2 + 4\right)^3} = 0.0058$$

<u>التحويل من $\hat{f}^2$ إلى ω^2 باستخدام المعادلة التالية:</u>

$$\omega^2 = \frac{\hat{f}^2}{\left(1 + \hat{f}^2\right)} \qquad (6)$$

مثال

إفترض أن قيمة $\hat{f}$ تساوي 0.8 إحسب قيمة ω^2

$$\omega^2 = \frac{\hat{f}^2}{\left(1 + \hat{f}^2\right)} = \frac{0.8^2}{\left(1 - 0,8^2\right)} = 0.39$$

وبالمثل يمكن تحويل قيمة $\hat{f}$ إلى قيمة ω^2 كما يلي

$$\hat{f}^2 = \frac{\omega^2}{\left(1-\omega^2\right)} = \frac{0.39}{\left(1-0.39\right)} = 0.639$$

$$\hat{f} = \sqrt{0.639} = 0.799 \approx 0.80$$

التحويل من η^2 الى d من خلال المعادلة الأتية:

$$d = \frac{2\sqrt{\eta^2}}{\sqrt{1-\eta 2}} \qquad (7)$$

التحويل من d الى η^2 من خلال المعادلة الأتية:

$$\eta^2 = \frac{d^2}{d^2+4} \qquad (8)$$

ملحق 2

جدول تحويلات حجوم الأثرلكوهين (Cohen, 1988, p.22)

Equivalents of d

d	U_1	U_2	U_3	r	r^2
0	0.0%	50.0%	50.0%	.000	.000
.1	7.7	52.0	54.0	.050	.002
.2	14.7	54.0	57.9	.100	.010
.3	21.3	56.0	61.8	.148	.022
.4	27.4	57.9	65.5	.196	.038
.5	33.0	59.9	69.1	.243	.059
.6	38.2	61.8	72.6	.287	.083
.7	43.0	63.7	75.8	.330	.109
.8	47.4	65.5	78.8	.371	.138
.9	51.6	67.4	81.6	.410	.168
1.0	55.4	69.1	84.1	.447	.200
1.1	58.9	70.9	86.4	.482	.232
1.2	62.2	72.6	88.5	.514	.265
1.3	65.3	74.2	90.3	.545	.297
1.4	68.1	75.8	91.9	.573	.329
1.5	70.7	77.3	93.3	.600	.360
1.6	73.1	78.8	94.5	.625	.390
1.7	75.4	80.2	95.5	.648	.419
1.8	77.4	81.6	96.4	.669	.448
1.9	79.4	82.9	97.1	.689	.474
2.0	81.1	84.1	97.7	.707	.500
2.2	84.3	86.4	98.6	.740	.548
2.4	87.0	88.5	99.2	.768	.590
2.6	89.3	90.3	99.5	.793	.628
2.8	91.2	91.9	99.7	.814	.662
3.0	92.8	93.3	99.9	.832	.692
3.2	94.2	94.5	99.9	.848	.719
3.4	95.3	95.5	*	.862	.743
3.6	96.3	96.4	*	.874	.764
3.8	97.0	97.1	*	.885	.783
4.0	97.7	97.7	*	.894	.800

ملحق 3

جدول تحويلات حجم الأثر إلى نسبة كسب مئينية (Marzano et al, 2001, p.160)

Conversion Table for Effect Size/Percentile Gain

Effect Size	Percentile Gain	Effect Size	Percentile Loss
0.00	0	0.00	0
0.02	1	−0.02	−1
0.05	2	−0.05	−2
0.08	3	−0.08	−3
0.10	4	−0.10	−4
0.13	5	−0.13	−5
0.15	6	−0.15	−6
0.18	7	−0.18	−7
0.20	8	−0.20	−8
0.23	9	−0.23	−9
0.25	10	−0.25	−10
0.28	11	−0.28	−11
0.31	12	−0.31	−12
0.33	13	−0.33	−13
0.36	14	−0.36	−14
0.39	15	−0.39	−15
0.41	16	−0.41	−16
0.44	17	−0.44	−17
0.47	18	−0.47	−18
0.50	19	−0.50	−19
0.52	20	−0.52	−20
0.55	21	−0.55	−21
0.58	22	−0.58	−22
0.61	23	−0.61	−23
0.64	24	−0.64	−24
0.67	25	−0.67	−25
0.71	26	−0.71	−26
0.74	27	−0.74	−27
0.77	28	−0.77	−28
0.81	29	−0.81	−29
0.84	30	−0.84	−30
0.88	31	−0.88	−31
0.92	32	−0.92	−32
0.95	33	−0.95	−33
1.00	34	−1.00	−34
1.04	35	−1.04	−35
1.08	36	−1.08	−36
1.13	37	−1.13	−37
1.18	38	−1.18	−38
1.23	39	−1.23	−39
1.28	40	−1.28	−40
1.34	41	−1.34	−41
1.41	42	−1.41	−42
1.48	43	−1.48	−43
1.56	44	−1.56	−44
1.65	45	−1.65	−45
1.75	46	−1.75	−46
1.88	47	−1.88	−47
2.05	48	−2.05	−48
2.33	49	−2.33	−49