حوليات الآداب والعلوم الاجتماعية

فصلية علمية محكمة - مجلس النشر العلمي - جامعة الكويت

حجم الأثر وما وراء التحليل

Effect Size and Meta-Analysis

ممدوح عبد المنعم الكناني
قسم علم النفس التربوي - كلية التربية
جامعة المنصورة
جمهورية مصر العربية

هشام فتحي محمد جاد الرب
رئيس وحدة الاستشارات البحثية والإحصائية
جامعة الكويت
الكويت

جامعة الكويت
KUWAIT UNIVERSITY

مجلس
النشر العلمي

ISSN: 5248 - 1560

الرسالة 640 - الحولية 44

1445 هـ / مـارس 2024 م

• ثمن العدد

الكويت: 500 فلس، البحرين: دينار واحد، قطر: 10 ريالات، الإمارات 10 دراهم، السعودية: 10 ريالات، عمان : ريال واحد.

ثمن النسخة في دول الوطن العربي ما يعادل دولاراً واحداً.

ثمن النسخة في الدول الأجنبية ما يعادل ثلاثة دولارات.

• الاشتراك السنوي (أفراد):

سنة واحدة: الكويت 4 دينار كويتي، الدول العربية: 6 دينار كويتي، الدول الأجنبية 22 دولار.

سنتان: الكويت 7 دينار كويتي، الدول العربية: 10 دينار كويتي، الدول الأجنبية 37 دولار.

ثلاث سنوات: الكويت: 10 دينار كويتي، الدول العربية: 14 دينار كويتي، الدول الأجنبية 52 دولار.

أربع سنوات: الكويت: 13 دينار كويتي، الدول العربية: 18 دينار كويتي، الدول الأجنبية 67 دولار.

• الاشتراك السنوي (مؤسسات):

سنة واحدة: الكويت 22 دينار كويتي، الدول العربية: 22 دينار كويتي، الدول الأجنبية 90 دولار.

سنتان: الكويت 37 دينار كويتي، الدول العربية: 37 دينار كويتي، الدول الأجنبية 150 دولار.

ثلاث سنوات: الكويت: 53 دينار كويتي، الدول العربية: 53 دينار كويتي، الدول الأجنبية 210 دولار.

أربع سنوات: الكويت: 67 دينار كويتي، الدول العربية: 67 دينار كويتي، الدول الأجنبية 270 دولار.

جميع المراسلات الخاصة بشروط النشر أو أية استفسارات أخرى

توجه إلى رئيس تحرير حوليات الآداب والعلوم الاجتماعية

ص.ب: 17370 الخالدية – الكويت: 72454 – تلفون: 24830256 – فاكس 24830256

ISSN 1560 - 5248 Key title: Hawliyyat Kulliyyat al-Adab

E-mail: aass@Ku.edu.Kw

http://apc.kuniv.edu.kw/AASS/

تتوفر نصوص البحوث كاملة لدى:

EBSCO Publishing Products

دار المنظومة: www.mandumah.com

UIrich's Periodicals Directory 263117

إصدارات مجلس النشر العلمي

مجلة العلوم الاجتماعية 1973، مجلة الكويت للعلوم والهندسة 1973، مجلة دراسات الخليج والجزيرة العربية 1975، لجنة التأليف والتعريب والنشر 1976، مجلة الحقوق، 1977، حوليات الآداب والعلوم الاجتماعية 1980، المجلة العربية للعلوم الإنسانية 1981، مجلة الشريعة والدراسات الإسلامية 1983، المجلة التربوية 1983،المجلة العربية للعلوم الإدارية 1991.

حوليات الآداب والعلوم الاجتماعية

تصدر عن مجلس النشر العلمي ـ جامعة الكويت

جميع الحقوق محفوظة للمجلة

فصلية علمية محكمة تتضمن مجموعة من الرسائل
تعنى بنشر الموضوعات التي تدخل في تخصصات
العلوم الإنسانية والاجتماعية

الحولية الرابعة والأربعون
الرسالة الأربعون بعد المئة السادسة
1445 هـ / 2024م

حوليات الآداب والعلوم الاجتماعية – جامعة الكويت

هيئـة التحريـر

- **أ.د. تغريد محمد القدسي - غبرا**
رئيسة هيئة التحرير
قسم دراسات المعلومات - جامعة الكويت
Taghreed.alqdsi@ku.edu.kw

- **أ.د. باقر سليمـان النجار**
جامعة البحرين - البحرين
أستاذ زائر بقسم الاجتماع - جامعة الكويت
baqer.alnajjar@ku.edu.kw

- **أ.د. نعمان محمود أحمد جبران**
جامعة إربد - الأردن
أستاذ زائر بقسم التاريخ - جامعة الكويت
numan.jubran@ku.edu.kw

- **أ.د. عدنان يوسف العتـوم**
قسم علم النفس - جامعة اليرموك - الأردن
atoum@yu.edu.jo

- **أ.د. مشاعل الحملي**
قسم اللغة الإنجليزية وآدابها - جامعة الكويت
mashael.alhamly@ku.edu.kw

- **د. عبد الله محمد الجسمي**
قسم الفلسفة - جامعة الكويت
abdulla.hasan@ku.edu.kw

- **د. إبراهيم ناجي الهدبان**
قسم العلوم السياسية - جامعة الكويت
ibrahim.alhada@ku.edu.kw

- **د. أحمد مبارك الحصم**
قسم الجغرافيا - جامعة الكويت
ahmed.hassam@ku.edu.kw

- **د. أسعد هاشم الصالح**
قسم اللغة العربية - جامعة أنديانا
الولايات المتحدة الأمريكية
alsaleha@indiana.edu.kw

- **أ. مهـا إبـراهيـم المسعد**
مديرة التحرير - جامعة الكويت
maha.almsad@ku.edu.kw

الهيئـة الاستشارية

- **أ.د. إبراهيـم السعافيـن**
قسم اللغة العربية - الجامعة الأردنية

- **أ.د. حمدي حسن أبو العينين**
كلية الإعلام - جامعة مصر الدولية

- **أ.د. سـاري حنفـي**
رئيس الجمعية الدولية لعلم الاجتمـاع
الجامعة الأمريكية - بـيروت
sh41@aub.edu.lb

- **أ.د. عبد الله الوليعي**
قسم الجغرافيا - جامعة الملك سعود

- **أ.د. مـأمـون فنـدي**
مدير معهد لندن للدراسات الإستراتيجية
المملكة المتحدة

- **أ.د. سليمان حوامدة**
جامعة شمال تكساس
Suliman.hawamdeh@unt.edu

- **أ.د. مهى عبد المجيد زراقط**
كلية الاعلام - الجامعة اللبنانية
Maha.zaraket@ul.edu.lb

قواعد النشر في حَوْليّات الآداب والعلوم الاجتماعيَّة

- حوليات الآداب والعلوم الإجتماعية: فصلية تنشرُ البحوثَ والدِّراساتِ الأصيلةَ باللّغتَيْن العربيَّةِ والإنجليزيَّة في مجال العلوم الاجتماعيّة والإنسانيّة التي يتوفر فيها ما يأتي:
- أنْ تُمثِّلَ الدِّراسةُ إضافةً جديدةً في حقل التّخصّص.
- لم يسبق نشر الدراسة بأي صورة كانت، ولم يسبق أيضًا تقديمها للنشر إلى جهة أخرى أثناء ورودها إلى الحوليات. ويلتزم الباحث بكتابة إقرار وتعهد بأن البحث المقدم لم يسبق نشره في أي وعاء نشر، أو أرسل إلى جهة أخرى.
- ألا يقل عدد كلمات الدراسة عن (15000) كلمة، شاملة المراجع والهوامش والجداول (بحدود 50 صفحة) وألا يزيد عدد الكلمات عن (60000) كلمة في حدود 200 صفحة.
- يطبع البحث بواسطة معالج النصوص Microsoft Word وعلى مسافة ونصف، وبنط Simplified Arabic 14 للبحوث باللغة العربية، وعلى مسافتين، وبنط Times New Roman 12 في حالة البحوث باللغة الإنجليزية. على أن يكون عنوان البحث معبراً عن الدراسة وبحدود 15 كلمة.
- يرفق الباحث ملخصاً للبحث، مطبوعًا باللغتين العربية والإنجليزية، في حدود (250) كلمة. على أن يحوي ملخص البحث: هدف الدراسة وأسئلتها، منهج الدراسة المستخدم، أبرز النتائج المستخلصة وأهم الاستنتاجات، إضافة إلى الكلمات الدالة (المفتاحية).
- يرفق الباحث مع البحث سيرة علمية مختصرة، مطبوعة باللغتين العربية والإنجليزية، تشمل أهمّ مؤلفاته وأبحاثه.
- تقدم الخرائط والأشكال والرسوم (إن وجدت) بأصولها الصالحة للطباعة بصيغة JPG، وبمستوى دقة 600 × 800.
- في حالة رغب الباحث بنشر الصور أو الخرائط أو الأشكال البيانية الملونة يلتزم بدفع تكاليفها.
- يراعي الباحث عند كتابة البحث الالتزام بالنسخة الأخيرة (السابعة) من نظام جمعية علم النفس الأمريكية American Psychological Association (APA7) من حيث كتابة المراجع والهوامش في متن البحث، إضافة لقائمة المراجع. على الوجه الموضح أدناه.
- كتابة المرجع في المتن: اسم العائلة للمؤلف متبوعًا بفاصلة، ثم سنة النشر. (يرجى الرجوع إلى دليل التوثيق وفقًا لنظام (APA7).

مثال (Courtois, 2001)

- قائمة المراجع: يرجى الرجوع إلى دليل التوثيق وفقًا لنظام (APA7) للتفاصيل.

Jones, J. (2005). Writing with style. Style Writing Journal, 12 (6),1433

ويمكن زيارة موقع APA لمعرفة القواعد الخاصة بنظام (APA7) على:

http://www.apastyle.org

- يجب أن تشتمل جميع البحوث على قائمة المراجع كاملة في نهاية البحث، على أن يكون بنط الكتابة بالنص الروماني (Times Roman Script).
- لمعرفة قواعد النشر وأخلاقياته الرجاء مراجعة موقع الحوليات الإلكتروني:

http://apc.kuniv.edu.kw/

شروط قبول البحوث في الحوليات:

- يجب أن يقدم البحث عن طريق نظام الـ OJS وليس البريد الإلكتروني؛ ليتسنى للحوليات البدء بالتقييم. - تقبل الحوليات البحوث التي تقدم إلكترونياً فقط من خلال الموقع:

http:// journals.KU.edu.KW/aass

- لا تقبل مجلة الحوليات البحوث التي سبق نشرها بأي شكل أو وسيط.
- لا تقبل مجلة الحوليات نشر أبحاث الماجستير أو الدكتوراه أو أي مستلات منها.
- لا ترد ولا تسترجع أصول البحوث المقدمة للنشر، سواء نشرت أم لم تنشر.
- لا يجوز نشر البحوث مع جهات أخرى إلا بعد موافقة الحوليات على ذلك وإذا ثبت نشرها فستتخذ إدارة الحوليات الإجراءات القانونية المتبعة بهذا الشأن.
- يمكن للباحث نشر بحثه مع جهات أخرى، بعد الحصول على إذن كتابي سابق من رئيس التحرير، وبعد انقضاء ثلاث سنوات على نشره في الحوليات.
- يعرض البحث الذي تتوافر فيه القواعد المذكورة سابقًا وبعد موافقة هيئة التحرير، على محكمين اثنين لتقرير مدى صلاحية البحث للنشر في المجلة، ويكون التحكيم ثنائي الحجب (Double Blind Review).
- ستحاول المجلة نشر البحوث بعد تحكيمها خلال مدة لا تزيد عن ستة شهور.
- تمنح المجلة للباحث ثلاثين نسخة مجانية مطبوعة من بحثه المنشور.
- ترسل جميع المراسلات الخاصة بالحوليات ما عدا البحوث، إلى رئيس التحرير عن طريق البريد الإلكتروني للمجلة.

رئيس تحرير حوليات الآداب والعلوم الاجتماعية

ص.ب : 17370 الخالدية

رمز بريدي : 72454 - الكويت

ISSN 1560 Key title: Hawliyyat Kulliyyat al-Adab

http://apc.kuniv.edu.kw/AASS

E-mail: aass@ku.edu.kw

الإصدار الثالث – الحولية الرابعة والأربعون – مارس 2024م

نبدأ عامًا جديدًا 2024م راجين للجميع الخير كله، ونتطلع إلى عالم يتحرر من الصعوبات التي كبلتنا طوال عقود ليست بالقليلة. هذا ونرجو لأهلنا في فلسطين السلام والتحرر الذي نصبو إليه جميعًا. يجيء هذا العدد ونحن لا نزال نعاني آثار الصراعات والمنازعات، ناهيك عن الحرب المستعرة ضدنا وضد ما نمثله بوصفنا عربًا ومسلمين.

الرسالة الأولى بعنوان: حجم *الأثر وما وراء التحليل*، للأستاذ الدكتور ممدوح عبد المنعم الكناني، من قسم علم النفس التربوي في كلية التربية بجامعة المنصورة؛ والأستاذ الدكتور هشام فتحي جاد الرب، من قسم علم النفس بكلية العلوم الاجتماعية في جامعة الكويت. تعرض الدراسة لتعريف حجم الأثر ولمنهجية ما وراء التحليل، وكيف يمكن القيام به، كما قدم البحث لأهمية الانتقال من الدلالة الإحصائية إلى العملية. ثم أخيرًا يعرض البحث للمقارنة بين المجموعات الفرعية في الخصائص الإحصائية لنتائج ما وراء التحليل ولموضوع ما وراء التحليل. ويتميز البحث بأنه يسوق أمثلة واقعية عديدة، إضافة إلى الأشكال التوضيحية؛ لتقريب فهم ما وراء التحليل.

الرسالة الثانية في هذه الحولية بعنوان: *التاريخ السياسي لإمارة بنو عناز في بلاد الجبال (حلوان) 381-446 هـ/ 991-1054م*، للأستاذ الدكتور سليمان العبد الله الخرابشة، من جامعة اليرموك في إربد بالمملكة الأردنية الهاشمية. تعرف الدراسة أسرة بني عناز الكردية التي تمكن أمراؤها من إقامة إمارة خاصة على يد زعيمهم أبي الفتح محمد بن عناز، ومن تبعه من الأمراء بعده. امتد حكم هؤلاء الأمراء خلال العهد البويهي والسلجوقي في الفترة (381-446هـ/991-1054م). كان الأمراء خلالها مستقلين وأصحاب سيادة ونفوذ. ضمت إمارتهم المدن والقرى العديدة قبل أن يضعف نفوذهم نتيجة النزاعات لحساب السلاجقة الذين قاموا بدورهم من حيث الاستيلاء على إقليمهم، إقليم الجبال.

الرسالة الثالثة بعنوان: *صعوبات تنظيم الانفعال كمتغير وسيط بين التعاطف وإجهاد الشفقة لدى مقدمي الرعاية الصحية (أطباء وممرضين)*، للدكتورة نجوى إبراهيم الشناوي من قسم علم النفس في كلية الآداب بجامعة طنطا في جمهورية مصر العربية. تهدف الدراسة إلى تفحص العلاقة بين إجهاد الشفقة، والتعاطف، وصعوبات تنظيم الانفعال. وقد تفحصت كل ذلك لدى 300 طبيب وممرض من مقدمي الرعاية الصحية. وانتهت الدراسة إلى وجود علاقات إيجابية بين إجهاد الشفقة، والتعاطف وصعوبات تنظيم الانفعال. والمثير للانتباه هو أن بعض هذه الانفعالات وجدت لدى مقدمي الرعاية الصحية من أطباء وممرضين كذلك.

الرسالة الرابعة في هذه الحولية بعنوان: *تَكْميمُ البَحرِ الطَّويلِ في الشِّعرِ العَرَبيّ: عَيْنِيَّةُ مُتَمِّم بن نُوَيْرَة نموذج*، للدكتورة ريم محمود فاخوري من قسم اللغة العربية في كلية الآداب بجامعة حلب؛ والدكتور أحمد كمال جطل من قسم العلوم الأساسية في كلية الهندسة الكهربائية والإلكترونية بالجامعة نفسها. يحاول البحث دراسة نظرية الفيزياء الكميَّة وتجلياتها في الشعر العربي. يرسخ البحث عملية نقد الشعر العربي ويرسيها بوصفها ممارسة ضرورية.

الرسالة الخامسة في هذا الإصدار هي بعنوان: *الرؤية الثنائية في رواية (سيدات القمر) لجوخة الحارثي*، للأستاذة الدكتورة آمنة يوسف محمد عبده من قسم اللغة العربية في كلية اللغات بجامعة صنعاء. تنظر الباحثة إلى التقنية التي استخدمها الراوي لبناء النص؛ بمعنى آخر استخدام نظرية السرد الحديثة.

أما الرسالة السادسة في هذه الحولية فهي بعنوان: *أسلوب التعليل بين التنظير النحوي والواقع النصي*، للدكتور أحمد عطيفي شحاتة من قسم اللغة العربية وآدابها في كلية الآداب بجامعة القاهرة. البحث يهدف إلى النظر العميق لأسلوب التعليل للوصول إلى كيفية تركيب الجملة التعليلية. تطرق البحث إلى أنواع التعليل الصريح والضمني والحقيقي، وفرق بينها مستخدمًا وصف النصوص ومحاولة الاستقصاء من خلال ذلك.

أ.د. تغريد القدسي - غبرا
رئيسة التحرير

Doi 10.34120/0757-044-640-001

الرسالة (640)
قدمت في : 2020/2/21
أجيزت في: 2022/6/13
● **للاستشهاد:**
الكناني، ممدوح عبدالمنعم وجاد الرب، هشام فتحي محمد. (2024). حجم الأثر وما وراء التحليل. *حوليات الآداب والعلوم الاجتماعية (44)*640.

حجم الأثر وما وراء التحليل

Effect Size and Meta-Analysis

ممدوح عبد المنعم الكناني
قسم علم النفس التربوي - كلية التربية
جامعة المنصورة
جمهورية مصر العربية

هشام فتحي محمد جاد الرب
رئيس وحدة الاستشارات البحثية والإحصائية
جامعة الكويت
الكويت

حوليات الآداب والعلوم الاجتماعية ـ الحولية الرابعة والأربعون 1445هـ/2024م

المؤلف:

أ.د. ممدوح عبد المنعم الكناني.

- أستاذ علم النفس التربوي بكلية التربية جامعة المنصورة.
- عميد كلية التربية الأسبق، جامعة المنصورة.

بعض الإنتاج العلمي:

أولاً – الكتب:

1- سيكولوجية التعلم وأنماط التعليم.
2- الأسس النفسية الابتكار.
3- القياس والتقويم النفسي.
4- مناخ الابتكارية في الأسرة والمدرسة والمجتمع.
5- الفروق الفردية في الذكاء والقدرات العقلية.
6- سيكولوجية الإبداع وأساليب تنميته.
7- الرعاية النفسية والتربوية للموهوبين.
8- سيكولوجية الإبداع والتجديد لدى الأفراد والمنظمات.
9- سيكولوجية إبداع الطفل.
10- قراءات ودراسات في إبداع الطفل.
11- تنمية إبداع الطفل.
12- الإحصاء الوصفي والاستدلالي في العلوم السلوكية والاجتماعية.

ثانياً – الأبحاث:

1- التنبؤ بقابلية الفرد للتعلم الذاتي من خلال سماته الشخصية.
2- المناخ الابتكاري في الاسرة، وعلاقته ببعض الخصائص الاجتماعية والشخصية لدى أعضاء الأسرة.
3- الجنس والخبرة غير المباشرة، وأثرهما على تعلم المهارات الإدراكية الحركية.
4- بعض خصائص شخصية المعلمين وعلاقتها برضاهم عن عملهم وبصحتهم النفسية.
5- الاندفاع – التروي كأسلوب معرفي ، وعلاقته بانتقاء الإجابة الصحيحة في الاختبارات العقلية ذات الاختيار من متعدد.
6- مدى تحقق التنظيم الهرمي للحاجات عند (ماسلو) في البيئة المصرية.
7- بحوث الابتكارية في البيئة المصرية ، بين النظرية والتطبيق.
8- مدى التغير في الاتجاهات النفسية والتربوية للمعلمين قبل وبعد تخرجهم من كليات التربية.
9- العلاقات التفاعلية بين التفكير الابتكاري والتفكير الناقد والذكاء في مستوياتهم المختلفة.
10- علاقة مركز التحكم (الداخلي – الخارجي) في التدعيم ببعض المتغيرات الدافعية.
11- مدى التباين بين تصرفات المعلمين في المواقف التربوية وتوقعات الطلاب والوالدين لتلك التصرفات.
12- مشكلات الأطفال المبتكرين المترتبة على خصائص شخصيتهم.
13- الأخطاء الإحصائية الشائعة في البحوث التربوية والنفسية.

المؤلف:

أ.د. هشام فتحي محمد جادالرب.

- الدكتوراه في جامعة واين استيت، الولايات المتحدة الأمريكية.
- رئيس وحدة الاستشارات البحثية والإحصائية، جامعة الكويت.
- أستاذ القياس والإحصاء النفسي المشارك، جامعة الكويت.
- أستاذ القياس والإحصاء النفسي، جامعة المنصورة.
- باحث خبير وإحصائي تطبيقي، جامعة مدريد، 2007.
- مستشار إحصائي جامعة وأين استيت الأمريكية، 2004.

بعض الإنتاج العلمي:

أولاً – الكتب:

1- Ben Zayed, S., Gani, A.B., **Gadelrab, H.F.,** & Bin Othman, M.K. (2020). Decades trend of emergency department system operations. In S. Ben Zayed et al. (Eds.), *Operational Management in Emergency Healthcare.* Springer.

2- *مفاتح الإحصاء: مدخل مبسط لأساسيات الإحصاء*. دار العلم، 2017.

3- *الإحصاء الاستدلالي في العلوم السلوكية*. (مشترك)، دار الكتاب الحديث، 2016.

4- **Gadelrab, H.F.** (2010). *Evaluating the fit of structural equation models: Sensitivity to specification error and descriptive goodness-of-fit Indices.* LAP Lambert Academic Publishing,

5- *جودة التعليم في ضوء تقييم القيمة المضافة*. (مشترك)، دار الفكر العربي، 2011.

ثانياً – الأبحاث:

1- Zine-El-Abiddine, F., Al-Tammemi, A.B., **Gadelrab, H.F.,** Lin, C-Y, Aljaberi, M.A., Alhuwailah, A., Roubi, M.L. (2021). Arabic COVID-19 psychological distress scale: Development and initial validation. *BMJ Open, 11 :e046006. doi:10.1136/ bmjopen-2020-046006.*

2- **Gadelrab, H.F.,** Alkhadher, O., Almoshawah, S., …& Slimene, S. (2020). Organizational justice in Arab Countries: Investigation of the Measurement and Structural Invariance. *Cross-cultural Research, 54*(1), 3 - 27.

3- **Gadelrab, H.F.** (2019). An Investigation of Differential Relationships of Implicit and Explicit Aggression: Validation of an Arabic Version of the Conditional Reasoning Test for Aggression. *Journal of Personality Assessment, 101*(6), 609-620.

4- **Gadelrab, H.F.** & Alkhadher, O. (2017). To Translate or to Develop a Measure? The Case of a New Arabic Measure of Organizational Justice. *International Journal of Selection and Assessment. 25*(1), 83 - 91.

5- Alkhadher, O. & **Gadelrab, H.F.** (2016). Organizational Justice Dimensions: Validation of an Arabic Measure. *International Journal of Selection and Assessment. 24*(4), 337- 351.

المحتوى

الملخص 15

مقدمة 17

المبحث الأول ـ الدلالة الإحصائية وحجم الأثر 19

الدلالة الإحصائية 19

أوجه النقد للدلالة الإحصائية 22

حجم الأثـر 22

تفسير حجم الأثـر 26

العوامل المؤثرة في دقة حجم الأثر 32

ملاحظات ختامية وملخص 33

المبحث الثاني ـ تقدير مؤشرات حجم الأثر وتحويلاتها 35

مؤشر الفرق الخام بين المتوسطات 35

مؤشر الفرق المعياري بين المتوسطات 36

حجم الأثر باستخدام معاملات الارتباط ومربعاتها 46

المبحث الثالث ـ طرق تقدير قيمة متوسط وتباين حجم الأثر في إطار نموذجي الآثار الثابتة والعشوائية 69

نموذجا الآثار الثابتة والعشوائية 77

نموذج الآثار الثابتة Fixed Effect Model (FEM) 77

نموذج الآثار العشوائية Random Fixed Effect Model (REM) 84

المبحث الرابع : المقارنة بين المجموعات الفرعية في خصائصها الإحصائية 105

المقارنة بين المجموعات الفرعية في بعض الخصائص الإحصائية في إطار نموذج الآثار الثابتة 105
تقدير قيمة تقديرات حجم الأثر للمجموعات الفرعية 120
نسبة التباين المفسر بالمجموعات الفرعية 131
المبحث الخامس ـ ما وراء التحليل : مفهومه وأهدافه وأهميته وخطواته 133
مفهوم ما وراء التحليل Meta-Analysis 133
أهداف ما وراء التحليل 134
أهمية ما وراء التحليل 135
خطوات ما وراء التحليل 136
المبحث السادس ـ قوة الاختبارات الإحصائية التي يتضمنها ما وراء التحليل 147
أنواع القرارات عند اختبار القروض إحصائياً 147
اختبار الفرض الصفري والفرض البديل 148
القوة الإحصائية والتخطيط للدراس 150
العوامل المؤثرة في قوة الاختبار الإحصائي 150
دراسات القوة الإحصائية 154
تحليل حساسية الاختبار 155
قوة الاختبارات الإحصائية المتضمنة بما وراء التحليل 155
قوة اختبار التجانس أو التباين بين المجموعات في إطار نماذج الآثار العشوائية 176
المراجع 179

الملخّص

الأهداف: هدف هذا البحث إلى عرض مقدمة نظرية للتعريف بحجم الأثر وعرض منهجية ما وراء التحليل، وكيفية القيام بها وأهميتها من خلال أمثلة متنوعة لطرق تقدير قيمة حجم الأثر المختلفة، ومتوسطها وتباينها عبر الدراسات المختلفة. يعرض البحث أهمية الانتقال إلى ما بعد الدلالة الإحصائية إلى الدلالة العملية وتقدير حجم الأثر، متناولاً مفهوم حجم الأثر، وتفسيره، والعوامل المؤثرة في دقة تقديره؛ وشارحاً أنواع مؤشرات حجم الأثر، وطرق التحويلات الرياضية بين تلك المؤشرات. **المنهج**: من الضروري التمييز بين نموذج الآثار الثابتة والعشوائية عند القيام بدراسة ما وراء التحليل؛ لذلك يعرض البحث طرق تقدير متوسط وتباين حجم الأثر في ضوء نموذج الآثار الثابتة ونموذج الآثار العشوائية. ينتقل البحث لعرض موضوع مهمّ عند القيام بدراسة ما وراء التحليل، وهو المقارنة بين المجموعات الفرعية في الخصائص الإحصائية لنتائج ما وراء التحليل، ومنه ينتقل البحث لعرض موضوع ما وراء التحليل: مفهومه وأهدافه وأهميته وخطواته، وينتهي البحث بعرض قوة الاختبارات الإحصائية التي يتضمنها ما وراء التحليل. **النتائج**: عرض البحث عدداً كبيراً من الأمثلة الواقعية والأشكال التوضيحية، لتسهيل فهم أسس ما وراء التحليل، ولمساعدة الباحث المهتم في تخطيط وتنفيذ دراسة ما وراء التحليل الخاصة به.

الكلمات الدالة (المفتاحية): حجم الأثر، تحليل الميتا، نموذج الآثار الثابتة، نموذج الآثار العشوائية.

مقدمة

يعد هذا البحث أساساً نظرياً لموضوع حجم الأثر وما وراء التحليل (أو التحليل التلوي أو البعدي) Meta-Analysis ، الذي يعتبر من الموضوعات الضرورية للباحثين في جميع ميادين العلوم الإنسانية والطبيعية، الذين يهدفون إلى إجراء تكامل بين النتائج التي توصلت إليها البحوث في أحد موضوعات تخصصهم. ويعرض الشق الأول من هذا البحث الفروق الأساسية بين الدلالة الإحصائية والدلالة العملية لنتائج البحوث، ومنها يتعرض لموضوع حجم الأثر، وتقدير مؤشراته في نماذج الآثار الثابتة Fixed Effects، والآثار العشوائية Random Effects. بينما يهتم الشق الثاني من البحث بما وراء التحليل (أو التحليل البعدي)، الذي يُعرف بأنه تحليل إحصائي لمجموعة كبيرة من النتائج المستخلصة من عدة دراسات فردية مستقلة في موضوع معين، بهدف التركيب والتكامل والتعميم، ويمكن اعتباره منهجاً وصفياً تحليلياً، يهدف إلى استخلاص النتائج الكامنة وراء عدة نتائج مستمدة من دراسات فردية ذات خصائص معينة. وتتحدد أهمية ما وراء التحليل في زيادة موضوعية ودقة نتائج متوسطات حجم الأثر الناتجة عنه، ومن خلال المراجعة المنتظمة للأبحاث الفردية يمكن معرفة الفجوات أو النواقص البحثية في أدبيات موضوع الدراسة، كما أنه يزيد من القوة الإحصائية للاختبارات الإحصائية المستخدمة لاستكشاف نواتج وخصائص ما وراء التحليل.

ويشمل هذا البحث ستة مباحث، يتضمن المبحث الأول الفرق بين الدلالة الإحصائية والدلالة العملية (حجم الأثر)، والعوامل المؤثرة في دقة حجم الأثر. ويتناول المبحث الثاني طرق تقدير حجم الأثر لأهم الاختبارات البارامترية، وبعض الاختبارات اللابارامترية، كما يتضمن هذا المبحث تحويل حجم الأثر الذي لا يناسب الباحث إلى حجم أثر آخر يتفق مع تحليلاته الإحصائية عند القيام بما وراء التحليل. ويشمل المبحث الثالث طرق تقدير متوسط حجم الأثر وتباينها في إطار نموذج الآثار الثابتة والآثار العشوائية، وكذلك خصائص هذين النموذجين والفرق بينهما. ويشمل المبحث الرابع المقارنة بين المجموعات الفرعية من حيث متوسطاتها، وتبايناتها.

والمبحثان الخامس والسادس مخصصان لاستخدام تقديرات حجم الأثر في القيام بما وراء التحليل. يتضمن المبحث الخامس مفهوم ما وراء التحليل وأهدافه وأهميته وخطواته، وطرق المراجعة الممنهجة للدراسات المتضمنة في ما وراء التحليل، كما يشمل خطوات ما وراء التحليل ومتطلباته، وكذلك الصدق الداخلي والخارجي لما وراء التحليل. ويتضمن المبحث السادس قوة الاختبارات الإحصائية المستخدمة بما وراء التحليل والعوامل المؤثرة في القوة وطرق تقدير قيمة تلك القوة لبعض الاختبارات الإحصائية المتضمنة بالتحليل البعدي.

المبحث الأول

الدلالة الإحصائية وحجم الأثر

يعتمد الباحثون بشكل أساسي على اختبارات الدلالة الإحصائية Statistical Significance عند تقرير وتفسير النتائج التي يتوصلون إليها في دراساتهم. ويشير فام (1997) إلى أن الدلالة الإحصائية تهتم بمستوى الثقة في النتائج المستنتجة، فعندما يكون الفرق بين متوسطين دالاً إحصائياً عند مستوى دلالة، فإننا نقرر رفض الفرض الصفري عند مستوى الدلالة 05. ؛ وهذا يعني أننا نثق بهذا القرار بنسبة 95%. إلاّ أن زيادة حجم العينة قد يؤدي إلى وجود دلالة إحصائية عند مستوى الدلالة 05. على الرغم من أن الفرق بين المتوسطين في حقيقة الأمر قد يكون صغيراً. ويرى كيرك (Kirk, 1996) أن الدلالة الإحصائية تشير إلى النتائج إذا ما كانت تعزى للمصادفة، بينما يوضح حجم الأثر Effect Size فائدة النتائج من الناحية العملية أو التطبيقية؛ ولذلك يجب أن تهتم بها الدراسات العملية والتطبيقية.

ويلاحظ أن معظم الباحثين يقتصرون في تقرير نتائج بحوثهم على الدلالة الإحصائية دون الكشف عن مقدار العلاقة القائمة بين المتغيرين، ولذلك توجد مغالاة في تفسير النتائج اعتماداً على مستوى الدلالة على الرغم من أنه قد لا تكون لها قيمة من الناحية التطبيقية أو العملية (Rosenthal, 1994). فإذا وجد الباحث أن القيمة دالة إحصائياً فإن ذلك لا يدل على أن حجم الأثر أو درجة العلاقة بين المتغيرين مرتفعة أو منخفضة. وهذا يعني أنه لا يمكننا التنبؤ بحجم الأثر من مجرد تعرّف مستوى الدلالة الإحصائية لاختبار إحصائي في بحث ما. فحجم الأثر جاء ليكمل المعلومات التي يعطيها مستوى الدلالة؛ ولم يأتِ ليحل محلها.

الدلالة الإحصائية

تشير الدلالة الإحصائية إلى درجة الثقة التي نعطيها لنتائج الفروق بين المتوسطات أو العلاقات بين المتغيرات بصرف النظر عن حجم هذه الفروق أو العلاقات، وأن هذه الفروق أو الارتباطات ليست نتيجة المصادفة أو أخطاء المعاينة والقياس. ومنذ أكثر من مائة عام كانت الدلالة الإحصائية هي الأساس في تحليل البيانات في مجال العلوم التربوية والاجتماعية؛ إذ يمكن للباحثين الحصول على نتائج ذات دلالة إحصائية، بغض النظر عن التأثير الفعلي للمعالجات في الدراسة. فقد تكون هذه النتائج دالة إحصائياً بسبب كبر حجم العينة مثلاً، وعادة ما يكون اهتمام الباحثين مركزاً على معرفة مدى وجود فروق دالة بين متوسطات المجموعات بغض النظر عن قيمة هذه الفروق ومدى فائدتها للممارسة

العملية. ويبين فان وشـــين (Fan & Chen, 2001) أنه كلما زاد حجم العينة قلّ خطأ المعاينة Sampling Error، وكلما قل خطأ المعاينة كان تقدير معلمة التوزيع (البارامتر) أكثر دقة، وهذا يعني أن فرصـــة الباحث تزيد في الحصـــول على النتائج ذات الدلالة الإحصـــائية بزيادة حجم العينة. والعكس أيضـاً صـحيح، فقد تكون الفروق بين المتوسـطات كبيرة نسـبياً، ولكن بالنظر إلى صـغر حجم العينات، فقد لا تصـل هذه الفروق إلى مسـتوى الدلالة الإحصـائية، حتى لو كان مسـتوى الدلالة الذي اختاره الباحث متواضعاً (فام،1997).

لنأخذ المثال التالي ليوضـح تأثير حجم العينة على الدلالة الإحصـائية (Cortina & Nouri, 2000) :في عينة حجمها 1500 قد نصـل إلى معامل ارتباط مقداره 0.2 وهو دال عند مسـتوى05. إلا أن القيمة الفعلية أو العملية له (r^2) لا تتعدى 0.04 والذي يمكن تفســـيره بأن نســـبة التباين في المتغير التابع الذي يرجع للمتغير المســـتقل هي 4% فقط، أما الباقي وقيمته 96% فإنه يرجع لعوامل أخرى. لنأخذ مثالاً آخر من الاختبار الإحصـائي الشـهير (مربع كاي x^2). يوضـح الجدول على اليمين توزيع بيانات للفروق بين النسـاء والرجال في الاسـتماع للموسـيقي البيانات على اليمين (A) تتضـمن عينة حجمها 20 فقط، في حين البيانات على اليسار (B) تتضمن بيانات 200 فرد (أي عشرة أضعاف حجم البيانات A).

(A)	لا يسمع الموسيقى	يسمع الموسيقى
النساء	2	8
الرجال	8	2

(B)	لا يسمع الموسيقى	يسمع الموسيقى
النساء	20	80
الرجال	80	20

في البيانات (B) تبلغ قيمة مربع كاي $x^2 = 5$ عند درجة حرية =1 وهي دالة عند05.، ومعامل الارتباط الرباعي لهذه البيانات $r_t = .809$؛ في حين أن قيمة مربع كاي من البيانات (A) تبلغ $x^2 = 69.62$ عند درجة حرية =1 وهي دالة عند،001. وقيمة معامل الارتباط الرباعي لم تتغير $r_t = .809$. لاحظ أنه على الرغم من أن قيمة معامل الارتباط الرباعي لم تتغير، فإن تغيراً واضحاً قد حدث في قيمة مربع كاي ومستوى الدلالة.

يعرض جدول (1/1) (في Holmes, 1983) قيم مســـتوى الاحتمال p (الذي يحدد مســـتوى الدلالة الإحصـــائية) في أربع دراســـات؛ حيث كانت قيمته 0.182 في الدراســـات الثلاث الأولى، بينما

كانت قيمته 0.0039 في الدراسة الرابعة. قد يظن البعض أنه نظراً لانخفاض قيمة احتمال الحصول على النتيجة بوساطة المصادفة في الدراسة الرابعة، فإن حجم الأثر يكون أكبر في الدراسة الأخيرة من الدراسات الثلاث الأخرى. لكن كما يوضح جدول (1/1) أن حجم الأثر للدراسات الأربع متساوٍ، والسبب في انخفاض قيمة p في الدراسة الأخيرة، هو ارتفاع حجم العينة؛ ومن ثم، انخفاض الخطأ المعياري، على الرغم من أن لها حجم الأثر نفسه للدراسات الثلاث الأخرى.

جدول (1/1)

اختلاف قيمة الاحتمال p في أربع دراسات لها حجم الأثر نفسه

الدراسة	D	N	*P*
A	0.30	80	0.182
B	0.30	80	0.182
C	0.30	80	0.182
D	0.30	400	0.0039
Total	0.30	640	0.001

وقد بين طومسون وشوماخر (Thompson & Schumacher, 1997) أن دراستين يمكن أن تجريا بعدد أفراد عينتين متفاوتين قد تؤديان إلى نفس قيمة حجم الأثر على الرغم من أن إحداهما قد تظهر فروقاً ذات دلالة إحصائية والأخرى قد تشير إلى عدم وجود فروق ذات دلالة إحصائية؛ مما يتطلب إجراء تحليل قبلي لقوة الاختبار قبل البدء بإجراء الدراسة لمعرفة عدد الأفراد الذي يحتاج إليه الباحث لإظهار الفروق الإحصائية في ظروف الدراسة عند علمه بتقدير حجم الأثر، فالباحث الذي استخدم عينة مكونة من 1000 شخص في دراسته يمكن أن يتبين أنه يمكن تكرار نتائج دراسته، والحصول على نتائج ذات دلالة إحصائية مشابهة باستخدام حجم عينة أقل. وفي المقابل فإن باحثاً آخر قد استخدم حجم عينة أقل وليكون 20 فرداً. وبوجود حجم أثر متوسط قد يشير إلى وجود فاعلية أو أثر عملي للمعالجة في دراسته ولكنها غير دالة إحصائياً؛ ومن ثم، يمكنه أن يقدم توصيات لباحثين آخرين بإعادة الدراسة مع زيادة لحجم العينة على الرغم من عدم ظهور دلالة إحصائية للفرق بين معالجات دراسته.

أوجه النقد الموجهة للدلالة الإحصائية

مما سبق يمكن استنتاج أوجه النقد الموجهة للدلالة الإحصائية، هي:

-الاعتقاد الخاطئ بأن قيمة p تقيس احتمال خطأ المعاينة والثقة في النتيجة إذا تكررت.

-المعلومات التي توفرها اختبارات الدلالة الإحصائية خاصة جداً؛ مما يجعلها غير قابلة للتعميم في كثير من المواقف.

- الدلالة الإحصائية لا تذكر شيئاً عن حجم الأثر مباشرة؛ ومن ثم لا نستطيع الخروج منها بمضمون عملي أو تطبيقي.

- أن مستوى الدلالة يقدم الدليل على قيمة احتمالات الخطأ في رفض الفرض الصفري في حال صحته فقط.

- نظراً لأن القيم الحرجة بالجداول الإحصائية هي التي يتم على أساسها رفض أو عدم رفض الفرض الصفري، فإنه إذا كانت هذه الدرجة مناظرة للاحتمال p الذي يساوي 0.051 فإن الباحث لا يستطيع رفض الفرض الصفري. وهذا يعني عدم الدقة في اتخاذ القرار في هذه الحالة، علماً بأن الفرق بين 0.051 و0.05 بسيط، ولكن على الباحث ان يكون متشدداً في اتخاذ القرار.

- يؤثر انتهاك افتراضات الأسلوب الإحصائي، وعدم استيفائها في البيانات كثيراً على قيمة الاحتمال والدلالة الإحصائية.

- حجم الأثر المتوسط قد لا يكون دالاً إحصائياً إذا كان حجم العينة صغيراً. وكذلك فرق المعالجة قد يكون دالاً إحصائياً، ولكنه لاَ يكون مهمّاً إذا كان حجم العينة كبيراً.

حجم الأثر

يسد مفهوم حجم الأثر ثغرة أدركها كثير من الباحثين، الذين اعتمدوا فقط على مستوى الدلالة التي تعبر عن مدى الثقة التي نوليها في نتائج الدراسات الخاصة بالفروق والعلاقات بصرف النظر عن حجم العلاقة. وتقوم فكرة حجم الأثر Effect Size (ES) على صياغة الفروق بين المتوسطات، وذلك باستخدام الانحراف المعياري كوحدة قياس لمقدار الفرق بين تلك المتوسطات، أو التعبير عن العلاقة بين المتغيرات المستقلة من جهة والمتغيرات التابعة من جهة أخرى، عن طريق استخراج حجم تباين المتغير التابع الذي يمكن تفسيره عن طريق المتغير المستقل، كما هو الحال في تحليل التباين (Cohen, 1988).

تم تعرّف مفهوم الدلالة العملية في بداية القرن العشرين بوساطة بيرسون، وفُسّر مربع إيتا η^2 في تحليل التباين بوساطة فيشر Fisher (1925)، كما استخدم مربع إيبسلون (ε^2) بوساطة كيلى

(1935)؛ ومن ثم عُمّم مفهوم حجم الأثر في كتب الإحصـاء. وقد شـجعت رابطة علم النفس الأمريكية منذ أكثر من 20 عاماً الباحثين على إضـافة حجم الأثر إلى جانب نتائج الاختبارات الإحصـائية في التقارير البحثية.

وقد قدمت رابطة علم النفس الأمريكية APA عدة توصـيات فيما يتعلق بحجم الأثر (في أبوعلام، 2006)، منها:

1- توثيق البيانات المتعلقة بحجم الأثر في المجتمع، وأحجام العينات، وخصائص مقاييس البحث التي تعتبر أساساً لاختبار القوة الإحصائية للاختبارات الإحصائية بالدراسة.

2- يجب إعطاء حجم الأثر المناسب لبيانات البحث مع النتائج الأولية وقيم *p*؛ لأن ذلك يثري البحث ويسهل القيام بالتحليل البعدي (ما وراء التحليل).

3-ضرورة استخدام حدود الثقة لقيم حجم الأثر التي يتم تقديرها، ويجب أن نعرف أن فترة الثقة إذا لم تتضمن القيمة صفراً فإننا نرفض الفرض الصفري عند الاحتمال *p*، والتي تقل عن 0.05.

4- لا تعتبر النتيجة التي نحصل عليها مهمة بمجرد أنها دالة إحصائية؛ لأن الحصول علي الدلالة الإحصائية لا يعني الدلالة العملية.

ونتيجة لذلك جاءت دعوة الباحثين إلى كتابة مسـتوى الدلالة الإحصـائية؛ بحيث تكون معززة بقيمة مؤشر حجم الأثر. وقد استجاب كثير من الباحثين لذلك، حتى إن محرري بعض الدوريات العلمية المتخصصة يشترطون تعزيز النتائج بقيم مؤشرات حجم الأثر للبحوث المقدمة للنشر (LeCroy & Krysik, 2007.)، وقد أوصت جمعية الباحثين التربويين الأمريكية AERA بدعم النتائج الإحصائية بقيم حجم الأثر وفترات الثقة المرتبطة بها وعدم اقتصـار البحث على الفحص التقليدي لدلالة الفروق الإحصـائية بين مجموعات الدراسـة. فمن بين 104 بحوث منشـورة في دوريتين شـهيرتين في البحث التربوي والنفسي تم تقييمها قبل عام 2000، توصل يوتكنر وزملاؤه (Yetkiner et al. , 2004) إلى أن 23% منها دعمت البحث بقيمة حجم الأثر، وذلك في دراسات ارتباطية تعتمد على معامل ارتباط بيرسون وتحليل الانحدار؛ حيث يظهر حجم الأثر كجزء من النتائج الإحصـائية الذي تزودنا به برامج التحليل الإحصـائي المختلفة تلقائياً. لكن عند النظر للبحوث المنشـورة بين عامي 2000 و 2007 لوحظ ازدياد في النسبة وصلت إلى 79% في إحدى الدوريات، ولوحظ أن البحوث التي لم تزود بقيم حجم الأثر أغفلت أيضاً التزويد بمعلومات كافية تقود إلى تقدير قيمته (Yetkiner et al., 2004).

وقد راجع كيرك (Kirk , 1996) عدداً من الدوريات وتوصـل إلى أن إجراء كتابة حجم الأثر غير كاف ويختلف من دورية إلى أخرى؛ حيث كانت نسـبة ظهور حجم الأثر في الدوريات التي تمت

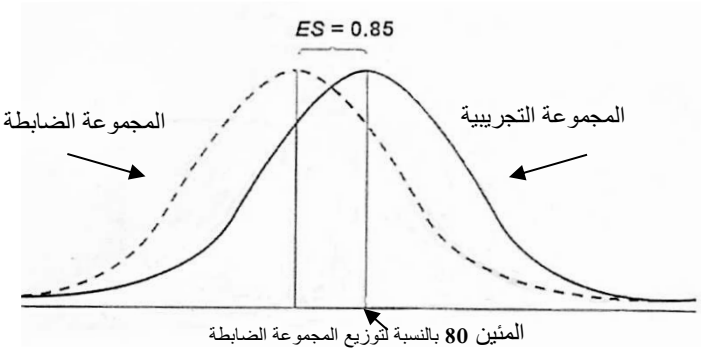
مراجعتها لا تزيد على 77%، وبناء على النتائج التي توصـــل إليها كيرك قامت بعض الدوريات بإضافة تعليمات جديدة للنشر، تطلب من الباحثين الإشارة في النتائج إلى مستوى الدلالة وحجم الأثر أو قوة العلاقة بين المتغيرات.

يوضح شكل (2/1) مفهوم حجم الأثر في إطار الفرق المعياري للمتوسطات بين مجموعتين: الأولى تجريبية (تخضـع للمعالجة)، والأخرى ضـابطة (لا تخضـع للمعالجة)؛ مثل مؤشـرات كوهين Cohen's d وهيدجز Hedges' g ومؤشر Glass's Δ (أو جلاس دلتا).

$$ES=\frac{\mu_1-\mu_2}{\sigma}=\Delta=\delta \qquad ES=\frac{\overline{X_e}-\overline{X_c}}{Sc}=g \qquad ES=\frac{\overline{X_e}-\overline{X_c}}{Sp}=d$$

شكل (1/1)

تفسير حجم الأثر في حالة المجموعتين التجريبية والضابطة (في Cohen, 1988)



من شـكل (1 /1) يتبين أن حجم أثر المعالجة 0.85، وأن متوسـط المشـاركين في المجموعة التجريبية يفوق 80% المشـــاركين في المجموعة الضـــابطة. إن حجوم الأثر هي مجموعة من المؤشرات تقيس مقدار أثر المعالجات أو فعاليتها، التي تختلف عن اختبارات الدلالة الإحصائية في أنها مســتقلة عن حجم العينة. وتعتبر مقاييس حجم الأثر العملة الموحدة لدراســات "ما وراء التحليل" Meta-Analysis، ويختلف حجم الأثر باختلاف نوع الدراســات ما بين دراســات تجريبية للفروق (مجموعة تجريبية ومجموعة ضابطة)، ودراسات ارتباطية وغيرها (سكران، 2006). يوضح جدول (2/1) مثالاً لثلاث قيم لحجم الأثر بطريقة كوهين d، مستخلصة من ثلاث دراسات افتراضية.

جدول(2/1)

مثال لثلاث قيم لحجم الأثر مستخلصة من ثلاث دراسات افتراضية

الدراسة	$\overline{X_1}$	$\overline{X_2}$	$\overline{X_1}$ - $\overline{X_2}$	S	d	T
(1)	220	200	20	48	0.42	دالة عند 0.01
(2)	210	200	10	48	0.21	دالة عند 0.01
(3)	170	200	-30	48	-0.63	دالة عند 0.01

تشير إشارة السالب في جدول (2/1) في الدراسة الثالثة إلى أن الأثر سلبي، ويلاحظ أن كل النتائج في الجدول دالة إحصائياً (قيم t دالة)، ولكن حجم الأثر في الدراسة الأولى ضعف حجم الأثر في الدراسة الثانية. وهذا يعني أن معرفة الدلالة الإحصائية لا تعطينا الكثير من المعلومات عن حجم الأثر. إن الدلالة الإحصائية تخبرنا بأن نتائج التجربة تقنعنا بأن هناك أثراً لا يرجع إلى المصادفة، لكنها لا تخبرنا بحجم هذا الأثر.

وقد أوضح روزينثال (Rosenthal,1991; 1994) أن هناك طرقاً ومؤشرات أكثر شيوعاً لتقدير قيمة حجم الأثر هي:

- طرق ومؤشرات مباشرة أو أولية تحسب حجم الأثر في دراسات الفروق؛ مثل مقارنة أثر مجموعتين في المتغير التابع، كما هو الحال في المؤشرات: كوهين d وهيجنز g وجلاس دلتا.
- طرق ومؤشرات مباشرة تحسب حجم الأثر في الدراسات الارتباطية؛ مثل قياس قوة العلاقة الخطية بين متغيرين، كما هو الحال في البيانات التي يمكن الحصول عليها من معاملات الارتباط ومربعاتها؛ مثل: r، r^2، R^2،(مربع أيتا) η^2 ،(مربع أوميجا) ω^2.
- طرق ومؤشرات مباشرة تحسب حجم الأثر في دراسات العلاقة بين المتغيرات الثنائية ذات الحدين، كما هو الحال في البيانات المستمدة من جداول (2x2) ومن إحصائي مربع كاي بدرجة حرية (1)؛ مثل نسبة المخاطرة Risk Ratio (RR)، ونسبة الأرجحية Odd Ratio (OR)، وعرض حجم الأثر ذي الحدين Binomial Effect Size Display (BESD).
- طرق ومؤشرات غير مباشرة أو محولة، وفيها تحوّل قيم بعض إحصاءات الاختبارات الإحصائية؛ مثل t،F،χ2، إلى أحد المؤشرات التي يحددها الباحث كعامل مشترك يستخدم في المقارنة أو في ما وراء التحليل.

وقد حدد عودة والخليلي (1998) مزايا الدلالة العملية أو حجم الأثر في الآتي:

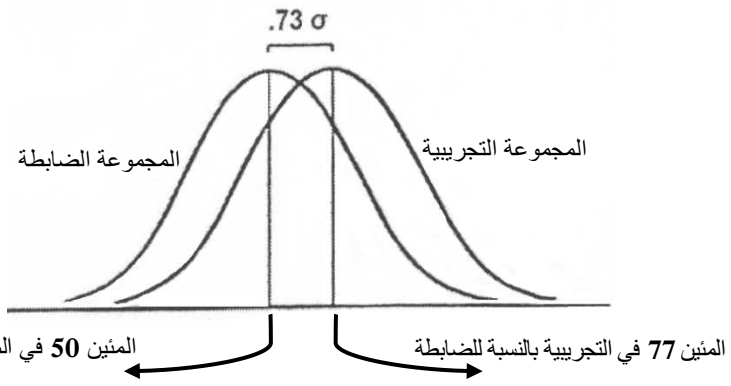
- إنها تعطي تقديراً لمســـتوى وجود الظاهرة في المجتمع، وذلك من خلال قيمة على مقياس متصـل يبدأ من الصـفر، عندما يكون الفرض الصـفري صـحيحاً، وقيمة أكبر من الصـفر إذا كان الفرض الصفري خاطئاً؛ ومن ثم فهو مؤشر لمدى الانحراف عن الفرض الصفري.
- إنها تعطي قيمة محددة تعتبر محكاً لمستوى حجم الأثر بأنه صغير أو متوسط أو كبير.
- يمكن استخدام نتائجها كمقياس كمي للمقارنة بين دراستين.
- يمكن استخدامها في تحليل القوة الإحصائية لتقدير حجم العينة اللازمة لدراسة ما.

تفسير حجم الأثر

يمكن تفسير حجم الأثر في إطار مدى التداخل بين توزيع المجموعتين التجريبية والضـابطة، من خلال المقارنة بين المئينيات أو تحديد مسـتويات لها. إن تقديرات حجم الأثر؛ مثل d وg، مشـابهة للدرجـة المعياريـة في التوزيع الاعتدالي المعيـاري. فحجم أثر مقداره d=.8 يعني أن درجـة الطـالب المتوسط في المجموعة التجريبية تزيد على درجات 79% من أفراد المجموعة الضابطة، ويتضح ذلك من جدول (3/1) وشكل (2/1).

شكل (2/1)

تفسير حجوم الأثر في تمثيله بالمئينيات على المنحنى الاعتدالي(Nouri & Greenberg, 1995)



يوضح شكل (2/1) أن متوسط المجموعة التجريبية يعادل المئين 77 في المجموعة الضابطة التي متوسطها يعادل المئين 50 (Redfield, Rousseau, 1981)، علماً بأن σ هو الانحراف المعياري المجمع للمجموعتين. ويبين جدول (3/1) أنه إذا كانت قيمة d لكوهين مساوية للصفر فإن ذلك يعني انطباق منحنيي توزيع المجموعتين التجريبية والضابطة، وهذا يعني عدم التداخل بينهما. ومن هذا الجدول يتبين أيضاً أنه إذا كانت d=0.8 فإن ذلك يعني وجود 47.7 % من عدم التطابق بين توزيع الدرجات في كلتا المجموعتين. وإذا كانت قيمة d= 1.7 فإن ذلك يعني أن عدم التداخل أو التطابق بين التوزيعين يساوي 75.4%.

وبلغة المئينيات، فإن d=0 يعني أن متوسط المجموعة التجريبية يمثل المئين 50 في المجموعة الضابطة، وإذا كانت d=0.8 فإن ذلك يعني أن متوسط المجموعة التجريبية يناظر المئين 79 للمجموعة الضابطة، وإذا كانت قيمة d= 1.6 فإن ذلك يعني أن متوسط المجموعة التجريبية يناظر المئين 95 للمجموعة الضابطة، كما يتضح من جدول (3/1). يعرض جدول (4/1) تمثيل حجم الأثر ذي الحدين لفرق المتوسط المعياري (BESD)، ومعامل الارتباط (r) المناظر لها.

جدول (3/1)

العلاقة بين d و P_s، P، BESD، CLES (Lenhard, & Lenhard, 2016)

حجم الأثر d	نسبة عدد الأفراد في المجموعة الضابطة الأقل من متوسط المجموعة التجريبية (P_S)	رتبة الشخص في مجموعة ضابطة تشمل 25 فرداً	احتمال وجود شخص في أي من المجموعتين بناءً على درجته (P)	الارتباط المكافئ (الفرق بين نسبتي النجاح بالمجموعتين) في حجم أثر ذي حدين BESD	احتمال أن شخصاً من المجموعة التجريبية سيكون أعلى من شخص بالمجموعة الضابطة إذا تم اختيارهما عشوائياً CLES
0.0	% 50	13th	0.50	0.00	0.50
0.1	% 54	12th	0.52	0.05	0.53
0.2	% 58	11th	0.54	0.10	0.56
0.3	% 62	10th	0.56	0.15	0.58
0.4	% 66	9th	0.58	0.20	0.61
0.5	% 69	8th	0.60	0.24	0.64
0.6	% 73	7th	0.62	0.29	0.66
0.7	% 76	6th	0.64	0.33	0.69
0.8	% 79	6th	0.66	0.37	0.71
0.9	% 82	5th	0.67	0.41	0.74
1.0	% 84	4th	0.69	0.45	0.76

حجم الأثر d	نسبة عدد الأفراد في المجموعة الضابطة الأقل من متوسط المجموعة التجريبية (Ps)	رتبة الشخص في مجموعة ضابطة تشمل 25 فرداً	احتمال وجود شخص في أي من المجموعتين بناءً على درجته (P)	الارتباط المكافئ (الفرق بين نسبتي النجاح بالمجموعتين) في حجم أثر ذي حدين BESD	احتمال أن شخصاً من المجموعة التجريبية سيكون أعلى من شخص بالمجموعة الضابطة إذا تم اختيارهما عشوائياً CLES
1.2	% 88	3rd	0.73	0.51	0.80
1.4	% 92	2nd	0.76	0.57	0.84
1.6	% 95	1st	0.79	0.62	0.87
1.8	% 96	1st	0.82	0.67	0.90
2.0	% 98	1st (or1st out of 44)	0.84	0.71	0.92
2.5	% 99	1st (out of 160)	0.89	0.78	0.96
3.0	% 99.9	1st (out of 740)	0.93	0.83	0.98

جدول(4/1)

تمثيل حجم الأثر لذى الحدين لفرق المتوسط المعياري (BESD) ومعامل الارتباط (r) المناظر لها (Lenhard, & Lenhard, 2016)

حجم الأثر(d)	(r)	نسبة الحالات بالمجموعتين الضابطة والتجريبية فوق الوسيط	الفرق بين النسبتين BESD
.10	.05	.47/.52	.05
.20	.10	.45/.55	.10
.30	.15	.42/.57	.15
.40	.20	.40/.60	.20
.50	.24	.38/.62	.24
.60	.29	.35/.64	.29
.70	.33	.33/.66	.33
.80	.37	.31/.68	.37
.90	.41	.29/.70	.41
1.00	.45	.27/.72	.45
1.10	.48	.26/.74	.48
1.20	.51	.24/.75	.51

وقد قدم كوهين (Cohen , 1977) ثلاثة قياسات إضافية تساعد في تفسير حجم الأثر، هي:

1- U_1، وهي نسبة المساحة الكلية في التوزيع، التي يغطيها توزيعا المجموعتين التجريبية والضابطة، ويكون فيها عدم تداخل، ويمكن حسابها على النحو الآتي:

أ- البحث عن d/2 في جدول التوزيع الاعتدالي، ثم تسجيل المناطق المحيطة من اليمين ولتكن A ومن اليسار ولتكن B.

ب- حساب نسبة عدم التداخل U_1، وهي الفرق بين A وB.

2- U_2، وهي النسبة في المجموعة التجريبية التي تزيد على النسبة في المجموعة الضابطة، ونبحث عن d/2 في جدول التوزيع الاعتدالي وستكون U_2 هي النسبة على اليسار.

3- U_3، وهي النسبة في المجموعة التجريبية، التي تتجاوز معدل الشخص في المجموعة الضابطة، ولحسابها نحدد قيمة d من جدول التوزيع الاعتدالي، وستكون هي النسبة إلى اليسار، ولذلك تستخدم U_3 بصورة أكثر، وذلك لسهولة تفسيرها.

ويعرض جدول (5/1) العلاقة بين قيمة كوهين d، واحتمالية التفوق (PS) Probability of Superiority، وتحدد احتمال أن تكون درجة فرد اختير عشوائياً من المجموعة التجريبية أكبر من درجة شخص اختير عشوائياً من المجموعة الضابطة، وU_1 (النسبة المئوية لعدم التداخل بين التوزيعين).

جدول (5/1)

العلاقة بين d، r، η^2، PS، U_1 (Lenhard, & Lenhard, 2016)

d	R	r^2 or η^2	PS	U_1
0.0	00	000	50	0
0.1	0.05	0.002	53	8
0.2	0.10	0.10	56	15
0.3	0.15	0.22	58	21
0.4	0.20	0.038	61	27
0.5	0.24	0.59	64	33
0.6	0.33	0.11	66	38
0.7	0.37	0.14	69	43
0.8	0.41	0.17	71	47
0.9	0.45	0.20	74	52
1.0	0.45	0.20	74	55
1.1	0.48	0.23	78	59
1.4	0.57	0.30	80	62
1.4	0.57	0.33	82	65
1.5	0.60	0.36	84	68
1.6	0.65	0.39	86	71

d	R	r^2 or η^2	PS	U_1
1.7	0.63	0.42	87	73
1.8	0.65	0.42	89	75
1.9	0.69	0.45	90	77
2.0	0.71	0.74	91	79
2.0	0.71	0.50	92	81
2.2	0.77	0.89	94	84
2.4	0.79	0.59	96	87
2.6	0.79	0.63	97	89
2.8	0.81	0.66	98	91
3.0	0.85	0.69	98	93
3.2	0.86	0.74	99	94
3.4	0.86	0.74	99	95
3.6	0.87	0.76	99	96
3.8	0.89	0.78	100	97
4.0	0.89	0.80	100	98

إن مؤشر كوهين U_3 (Cohen's U_3) يقيس النسبة المئوية لمجموعة المعالجة (التجريبية) التي تعلو متوسط المجموعة الضابطة، وهذا ما يبينه جدول (5/1). أما جدول (1-6)؛ فيعرض نسبة أفراد المجموعة التجريبية (المعالجة) فوق متوسط المجموعة الضابطة.

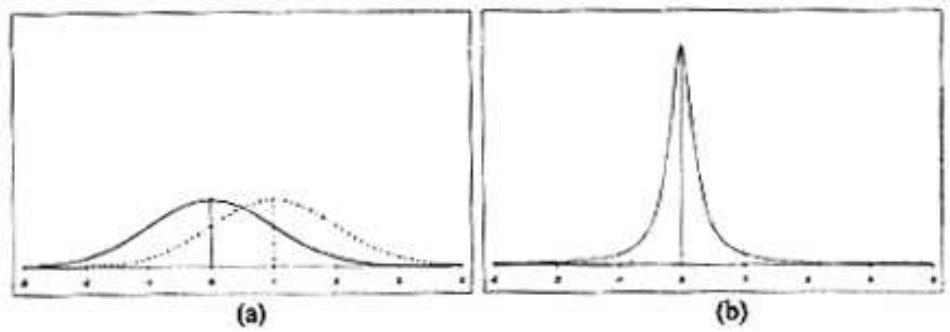
جدول (6/1)

نسبة أفراد المجموعة التجريبية (المعالجة) فوق متوسط المجموعة الضابطة (Lenhard, & Lenhard, 2016)

حجم الأثر d	النسبة فوق متوسط المجموعة الضابطة	النسبة المضافة فوق الاحتمال الأصلي (0.50)
.10	.54	.04
.20	.58	.08
.30	.62	.12
.40	.66	.16
.50	.69	.19
.60	.73	.23
.70	.76	.26
.80	.79	.29
.90	.82	.32
1.00	.84	.34
1.10	.86	.36
1.20	.88	.38

شكل (3/1)

حجما الأثر المتساويان يختلفان في المعنى من حيث التداخل بين المنحنيين A وB الاعتدالي وغير الاعتدالي (في Cohen, 1988)



لذلك توجد طريقة أخرى لتفسير حجم الأثر، وهي مقارنتها عن طريق تقدير حجم الأثر المعياري الذي حدده العديد من الباحثين. مثال ذلك تحديد كوهين (Cohen, 1977) لثلاثة مستويات لحجم الأثر: منخفض(0.2) ومتوسط (0.5)، ومرتفع (0.8)، ويبين جدول (7/1) ثلاثة مستويات (منخفض- متوسط - مرتفع) لبعض مؤشرات تقدير قيمة حجم الأثر. وأكد كوهين أن حجم الأثر المتوسط هو القيمة التي يرغب أن يحصل عليها الباحث الدقيق في دراسته، وفي حال عدم وجود أبحاث سابقة في مجال الدراسة فإن الباحثين -على الأغلب- يستخدمون هذا الحجم (حجم الأثر المتوسط) في الدراسة الأولى من نوعها في المجال؛ كونها تمثل متوسط قيم حجم الأثر في حقول مختلفة من البحث.

جدول (7/1)

مستويات حجم الأثر لعدد من المؤشرات المستخدمة في تقدير قيمته (Ellis, 2010)

ما الذي نقيسه؟	مستويات حجم الأثر			مؤشرات حجم الأثر
	مرتفع	متوسط	منخفض	
العلاقة بين متغيرين متصلين	0.60	0.40	0.20	ϕ_ε, ϕ, ρ, R, r
تفسير تباين المتغير التابع بالمتغير المستقل	0.86	0.16	0.04	ϕ^2، R^2، r^2
الفرق المعياري بين مجموعتين	0.80	0.50	0.20	d, g, Δ
الفرق بين معاملي ارتباط مستقلين	0.50	0.30	0.10	Cohen's a = r_A – r_B
الفرق بين نسبتين	0.80	0.50	0.20	Cohen's h = P_A – P_B
اختبار الإشارة حيث P=.5	0.25	0.15	0.05	Cohen's h = P – 0.50
χ^2 لحسن المطابقة أو الاستقلال	0.50	0.30	0.10	Cohen's w
مربع معامل الارتباط المتعدد وتفسيره	0.15	0.06	0.01	η^2, ω^2, ε^2
تحليل التباين في اتجاه واحد	0.40	0.25	0.10	Cohen's f
نسبة الأرجحية OR ونسبة المخاطرة RR	4.0	3.0	2.0	RR, OR

العوامل المؤثرة في دقة حجم الأثر

إن دقة حجم الأثر يمكن أن يعبر عنها بالخطأ المعياري لحجم الأثر أو فترة الثقة له، فكلما قل الخطـأ المعياري لحجم الأثر وضاقت فترة الثقة له، كان ذلك مؤشراً لدقة حجم الأثر، الذي يؤثر بدوره في دقة التحليل البعدي أو ما وراء التحليل، وفيما يلي بعض العوامل المؤثرة في دقة حجم الأثر (Cohen, 1988):

1- **محدودية المدى**: إن اختلاف العينات من حيث تباينها أو تجانسها الداخلي تؤثر في حجم أثرها. فكلما زاد تجانس درجات المجموعة أدى ذلك إلى حجم أثر أكبر من الذي نحصل عليه من عينة أخرى تباينها أكبر. ولذلك يُفضل استخدام الانحراف المعياري للمجتمع الذي أخذت منه هذه العينات.

2- **التوزيعات الاعتدالية وغير الاعتدالية**: يتوقف تفسير حجم الأثر على نوع التوزيع الذي أخذت منه، أهو اعتدالي أم غير اعتدالي، ويُفضل عدم المقارنة بين حجوم الأثر التي أُخذت من توزيعات غير اعتدالية.

3 -**الخطأ المعياري**: كلما قل الخطأ المعياري كان ذلك مؤشراً لدقة اتخاذ القرار فيما يتعلق بحجم الأثر. ويلاحظ أن الخطأ المعياري مهم في تقدير قيمة فترة الثقة لحجم الأثر.

4- **ثبات الأداة وصدقها**: كلما زاد ثبات الأداة وصدقها كان حجم الأثر المحسوب أكثر دقة.

5- **نوع التصميم البحثي**: إن التصميم التجريبي للمجموعات المرتبطة أو المتناظرة تعطي تقديراً لحجم الأثر أكثر دقة من التصميم التجريبي لمقارنة المجموعات المستقلة.

6-**حجم العينة**: كلما زاد حجم العينة زاد دقة وتقدير احصاءاتها. إن الدراسات ذات الحجم الكبير تحقق تقديرات لحجم الأثر ذات دقة أكبر.

ملاحظات ختامية وملخص

يستخلص من العرض السابق ملاحظات، أهمها (أبو هاشم، 2004؛ الكناني، 2020):

1- الدلالة الإحصائية لا تذكر شيئاً يتعلق بحجم الأثر مباشرةً؛ ولذلك لا نستطيع الخروج منها بما يشير إلى مضمون تطبيقي أو عملي.

2- حجم الأثر هو مفتاح ما وراء التحليل؛ إذ يجعله ممكناً؛ فهو يمثل قيمة واتجاه العلاقة التي تهتم بها.

3- إن معظم اختبارات الدلالة الإحصائية يمكن تحويلها إلى مقاييس لحجم الأثر.

4- لا يمكن التنبؤ بحجم الأثر بدقة من تعرّف مستوى الدلالة الإحصائية فقط؛ فقد يكون مستوى الدلالة الإحصائية مرتفعاً وحجم الأثر متوسطاً او ضعيفاً، وقد يكون السبب الرئيسي في عدم الوصول إلى الدلالة الإحصائية هو صغر حجم العينة مع أن حجم الأثر كبير، وقد يكون مستوى الدلالة ضعيفاً في حين أن حجم الأثر كبير.

5- يوجد أنواع متعددة لمقاييس حجوم الأثر، كل واحد منها يناسب مواقف وتصميمات بحثية معينة.

6- تتوقف الدلالة العملية على نوع البحث وعلى خبرة الباحث وحكمه علي النتائج، ومستوى حجم الأثر لا يناظر دائماً مستوى الدلالة العملية؛ فقد يكون حجم الأثر صغيراً ومع ذلك يرى الباحث أن النتائج لها دلالة عملية؛ لأن ذلك يتوقف علي عوامل أخري تدخل في تقدير الباحث.

7- إذا لم توجد دلالة إحصائية للنتيجة، وكان حجم العينة كبيراً بما يكفي لاكتشاف حجم أثر متوسط، فإننا نتوقع حجم أثر ضعيف.

8- لا يتأثر حجم الأثر بحجم العينة، في حين تتأثر مستويات الدلالة الإحصائية لنتائج الاختبارات بحجم العينات.

9- الدراسات ذات أحجام العينات الكبيرة، والأحداث الكثيرة لها وزن أكبر عند تقدير قيمة متوسط حجم الأثر.

10- تحديد حجم الأثر مهم للباحثين في تصميم دراسة ذات قوة إحصائية أعلى بدلاً من إضاعة الوقت في أشياء تافهة؛ إذ تزداد قوة الاختبار الإحصائي بزيادة حجم الأثر في تلك الدراسة وهذا ما تبينه الجداول الإحصائية.

11- من المنطقي إذا ما تم حساب حجم الأثر باستخدام عينة كبيرة الحجم، فإن قيمته سوف تكون أكثر دقة مما لو حُسب من عينة صغيرة الحجم.

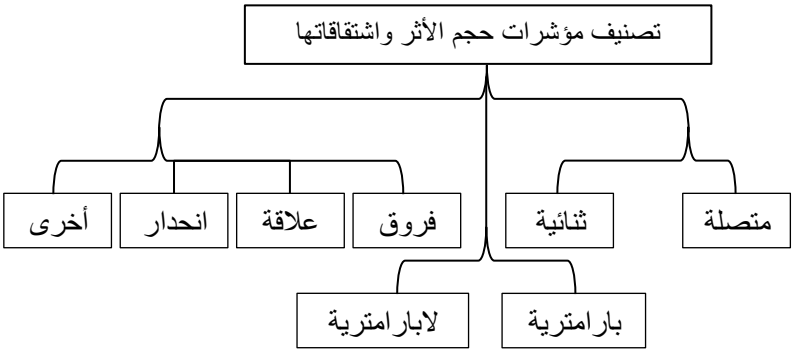
المبحث الثاني

تقدير مؤشرات حجم الأثر وتحويلاتها

هناك عدد كثير من المؤشـرات والمعادلات لتقدير قيمة حجم الأثر، ويحتاج بعض الباحثين إلى توحيد طريقة تقدير قيمة حجم الأثر من الدراسـات المتنوعة للقيام بدراسـة ما وراء التحليل (أو التحليل البعدي)، ولكن جميع الصـيغ ليسـت متكافئة في صـدقها أو اسـتدلالها. سـنعرض في هذا المبحث أنواع مؤشـرات حجم الأثر الأكثر شـهرة وتحويلاتها من شـكل إلى آخر. يعرض شـكل (1/2) تصـنيفاً عاماً لمؤشرات حجم الأثر.

شكل (1/2)

تصنيف مؤشرات حجم الأثر واشتقاقاتها



مؤشر الفرق الخام بين المتوسطات

$$RD = \bar{x}_1 - \bar{x}_2$$

يسـمى الفرق بين المتوسـطين بحجم الأثر الخام Raw Difference (RD)، أما إذا نسـب هذا الفرق إلى الانحراف المعيـاري لإحـدى المجموعتين المتجـانســتين أو الانحراف المعيـاري المجمع Pooled Standard Deviation؛ فإن حجم الأثر في هذه الحالة يكون أكثر عمومية ويشـبه الدرجة المعيارية (SMD)؛ مثل قيمتي g و d.

ويفضل استخدام متوسط الفرق غير المعياري (مثل الفرق الخام بين مجموعتين، وفترة الثقة) عندما تكون:

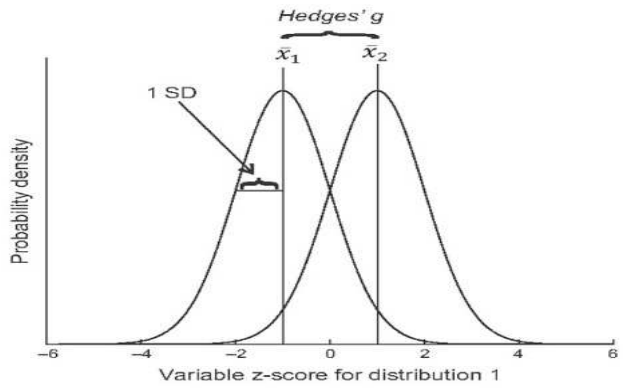
1- العينة ذات مدى محدود.
2- المجتمع الأصلي غير اعتدالي.
3- المجموعتين التجريبية والضابطة لهما انحرافيان معياريان مختلفان.
4- المقياس المستخدم منخفض الثبات أو ثباته غير معروف.

مؤشر الفرق المعياري بين المتوسطات

يمكن تحديد حجم الأثر من خلال تحديد الفرق المعياري بين المتوسطات Standardized Mean Difference (SMD)، وهو يساوي الفرق بين متوسطي مجموعتين مقسوماً على الانحراف المعياري المجمع لهاتين المجموعتين، و يشار إلى حجم الأثر ذلك بمؤشر كوهين Cohen's d، وهو يساوي أيضاً هيدجز Hedges's g. ويلاحظ من شكل (1/2) أنه عندما يكون التباين في المجموعتين صغيراً، فإن الفرق بين متوسطي المجموعتين يكون ظاهراً أو كبيراً. أما إذا كان التباين في المجموعتين كبيراً، فإن الفرق بين متوسطي المجموعتين يكون منخفضاً وأقل دلالة.

شكل (1/2)

الفرق المتوسط المعياري



وقد اقترح هنتر وآخرون (Hunter, et al., 1986) استخدام الانحراف المعياري المجمع من المجموعتين؛ لأنه أقل تأثراً بأخطاء المعاينة مقارنة بالانحراف المعياري للمجموعة الضابطة.

ويُقدر الانحراف المعياري المجمع بالمعادلة $S_\rho = \sqrt{\frac{(n_e-1)S_e^2+(n_c-1)S_c^2}{n_e+n_c-2}}$

والخطأ المعياري للفرق بين المتوسطين = $SE_{SMD} = \sqrt{S_\rho^2\left(\frac{1}{n_e}+\frac{1}{n_c}\right)}$

وللحصول على الانحراف المعياري المجمع بالمجتمع فإننا نقدره من S_p باستخدام المعادلة:

$$\sigma_p = S_p\sqrt{df/N}$$

حيث df درجة الحرية، وN العدد الكلي للمشاركين.

حيث S_ρ هو الانحراف المعياري المجمع للمجموعتين التجريبية والضابطة.

ويفترض لاستخدام هذه المعادلة تجانس الانحراف المعياري بالمجموعتين أو مجتمعهما

كذلك فإن تباين SMD يساوي:

$$V_{SMD} = \left[\frac{n_e+n_c}{n_e \cdot n_c}\right]\left[\frac{(SMD)^2}{2\,(n_e+n_c)}\right]$$

والخطأ المعياري لـ (SMD) $SE_{.SMD} = \sqrt{V_d}$

مثال: من البيانات التالية $\bar{x}_1$، $\bar{x}_2$ تساويان 55 و75، وn_1، n_2 تساويان 155 ,156 وS_1، S_2 تساويان 47 و 64، احسب SMD و S_ρ

RD = 55 – 75 = -20

$$S_\rho = \sqrt{\frac{(155-1)(47)^2+(156-1)(64)^2}{155+156-2}} = \sqrt{3156}=56.18$$

الخطأ المعياري لـ SMD

$$SE_{SMD} = \sqrt{3156\left(\frac{1}{155}+\frac{1}{156}\right)} =6.37$$

$$SMD = \frac{-20}{6.37} = -3$$

فترة الثقة لـ MD بمستوى 95%

SMD = [-20 ± 1.96(6.3)]=[-32.49, – 7.51]

مؤشر كوهين d

عرف كوهين (Cohen, 1988) المؤشر d بأنه الفرق بين المتوسطين $\bar{x}_1$ و $\bar{x}_2$ مقسوماً على الانحراف المعياري لإحدى المجموعتين بشرط التجانس بينهما، ويكون الفرق بين المتوسطين في اتجاه التحسن أو الاتجاه المتوقع إذا كانت قيمته موجبة والعكس صحيح. ويفضل استخدام الانحراف المعياري المجمع في بعض الصيغ الخاصة بالمؤشر.

$$d= (\bar{x}_e-\bar{x}_c) /S_p$$

حيث $\bar{x}_e$ متوسط المجموعة التجريبية، و$\bar{x}_c$ متوسط المجموعة الضابطة، وS_p الانحراف المعياري المجمع، ويتحدد باستخدام المعادلة:

$$S\rho = \sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}$$

ويحسب الخطأ المعياري لـ d من المعادلة التالية لمجموعتين مستقلتين:

$$SE_d= \sqrt{\frac{n_1+n_2}{n_1\, n_2}+\frac{d^2}{2(n_1+n_2)}}$$

وفي حالة العينتين المرتبطتين يحسب الخطأ المعياري لـ d باستخدام المعادلة:

$$SE_d = \sqrt{\frac{2(1-r_{12})}{n}+\frac{d^2}{2(n-1)}}$$

مثال: إذا كان متوسطا مجموعتين $\bar{x}_1 = 103$ و$\bar{x}_2 = 100$ وانحرافهما المعياري 5.5 و4.5، وكان حجم العينتين $n_1=n_2=50$. فما قيمة S_p, d, V_d, SE_d.

$$S_p= \sqrt{\frac{(50-1)\times 5.5^2+(50-1)\times 4.5^2}{50+50-2}} =5.025$$

$$d= \frac{103-100}{5.025} =0.597$$

$$V_d= \frac{50+50}{50\times 50}+\frac{5.025^2}{2(50+50)} = 0.0418$$

$$SE_d= \sqrt{0.0418} =0.204$$

أما في حالة مجموعتين مرتبطتين فإن: $d = \overline{D}/S_{\overline{D}}$

حيث $\overline{D}$ هو متوسط الفروق المتناظرة بين القياسين القبلي والبعدي مثلاً، ويتحدد $S_{\overline{D}}$ باستخدام المعادلة

$$S_{\overline{D}} = \sqrt{S_1^2 + S_2^2 - 2r.S_1.S_2}$$

وهذه الطريقة مناسبة لتحديد درجة تأثير المعالجة في منهج المجموعة الواحدة التي تستخدم الفرق بين القياسين القبلي والبعدي؛ حيث يحسب S_p باستخدام المعادلة:

$$SE_{\overline{D}} = S_{\overline{D}}/\sqrt{2(1-r)}$$

حيث: r هو معامل الارتباط بين درجات القياسين القبلي والبعدي.

$S_{\overline{D}}$ هو الانحراف المعياري للفرق بين درجات القياسين القبلي والبعدي.

$SE_{\overline{D}}$ هو الخطأ المعياري للفرق بين درجات القياسين القبلي والبعدي.

مثال: إذا كان متوسطا درجات مجموعتين مرتبطتين $\bar{x}_1$ و $\bar{x}_2$ هما 100,105، وأعداد أفراد المجموعة 50، ومعامل الارتباط يساوي 0.50، فإنه يمكن تقدير حجم الأثر

$S_{\overline{D}} = \sqrt{(10)^2 + (10)^2 - 2(0.50)(10)(10)} = 10$

$\overline{D} = 105-100 = 5$

حجم الأثر d $d = 5/10 = 0.5$

مؤشر هيدجز g (Hedges's, 1983) ومؤشر جلاس دلتا Glass's Delta

تحسب g مباشرة باستخدام المعادلة $g = \frac{\overline{x}_e - \overline{x}_c}{S_\rho}$

حيث $\overline{x}_e$ ، $\overline{x}_c$ هما متوسطا المجموعتين و S_ρ الانحراف المعياري المجمع من المجموعتين المتجانستين. ويمكن تحديد الانحراف المعياري المجمع باستخدام المعادلة $S_\rho = S_e\sqrt{n-1}$

حيث S_e هو الخطأ المعياري للمتوسط و n عدد أفراد المجموعة، وكذلك فإن:

$$S_\rho = \sqrt{\frac{(n_e-1)S_e^2 + (n_c-1)S_c^2}{n_e + n_c - 2}}$$

لاحظ أنه عند تقدير قيمة حجم الأثر لا نستخدم الانحراف المعياري لتوزيع المتوسطات، ولكننا نستخدم الانحراف المعياري لأفراد العينة الضابطة، على افتراض أن العينتين لهما الانحراف المعياري نفسه، وتكون g مساوية $g = (\overline{x}_e - \overline{x}_c)/S_\rho$

وبالنسبة للمجتمع: $\delta = \Delta = (\mu_e - \mu_c)/\sigma$

وعندما يساوي Se، Sc الانحرافين المعياريين للمجموعتين، فإن:

$$g = \frac{\bar{x}_e - \bar{x}_c}{\sqrt{(n_e-1)S_e^2 + (n_c-1)S_c^2}} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{MS_w}}$$

$$g = \frac{t\,(n_e + n_c)}{\sqrt{n_e\, n_c\, (n_e + n_c - 2)}}$$

$$g = \frac{2\sqrt{F}}{\sqrt{df}} = 2\sqrt{F/df}$$

في حالة استخدام t للمقارنة بين المجموعتين التجريبية والضابطة، فإن:

$$g = t\sqrt{\frac{1}{n_e} + \frac{1}{n_c}} = t\sqrt{\frac{n_e + n_c}{n_e\, n_c}}$$

وفي حالة تساوي المجموعتين في العدد n، فإن $g = \frac{2t}{\sqrt{2n}}$

وفي حالة استخدام Z للمقارنة بين المجموعتين، فإن $g = Z\sqrt{\frac{n_e + n_c}{n_e\, n_c}}$

وفي حالة تساوي عدد المجموعتين، فإن $g = 2\,Z/\sqrt{2\,n} = \sqrt{2f/n}$

تقدير g من قيم ودلالة *p* للمجموعتين التجريبية والضابطة

يمكن حســـاب قيمة g من معرفة دلالة وقيمة *p* ووجهة اختبارها (ذات طرف واحد أم ذات طرفين)؛ وذلك بتحديد قيمة Z المناظرة لقيمة *p*. وتوضـــع g مســـاوية للصـــفر في حالة عدم الدلالة، ولتقدير g في هذه الحالة نتبع الآتي:

1- حساب قيمتي p_e ، p_c

2- تحويل قيمتي p_e، p_c إلى مناظرهما Z_e ، Z_c باستخدام جدول المساحات بالمنحنى الاعتدالي المعياري.

3- نســـتخدم المعادلة $g = (Z_e - Z_c)\sqrt{\frac{n_e + n_c}{n_e\, n_c}}$، وهي معادلة جلاس وزملائه (Glass, et al., 1981)

و في حالة $n_e = n_c$ تصبح المعادلة

$$g = (Ze - Zc)\sqrt{\frac{2}{n}}$$

$$g = \frac{2\,(Z_e - Z_c)}{\sqrt{2n}}$$

وفي حال الفرق بين نسبتين نحسب نسبة النجاح في المجموعتين P_c وP_e ، ونسبة الفشل في المجموعتين q_e و q_c.

ويحسب الخطأ المعياري للفرق بين النسبتين: $S_p = \sqrt{\frac{(n_e-1)p_eq_e+(n_c-1)p_cq_c}{n_e+n_c-2}}$

كما نحسب دلالة الفرق بين النسبتين من المعادلة: $Z = (Pe - Pc)/S_p$

كما يمكن استخدام t

$$t = \frac{P_{exp}-P_{con.}}{\sqrt{[P(1-P)]\left[\frac{1}{n_{exp.}}+\frac{1}{n_{con.}}\right]}}$$

حيث:

P_{exp} : النسبة المئوية للنجاح بالمجموعة التجريبية

$P_{con.}$: النسبة المئوية للنجاح بالمجموعة الضابطة

$$P = \frac{P_{exp} - P_{con.}}{2}$$

وباستخدام χ^2 من الجدول 2 × 2 $g = \sqrt{\frac{\chi^2\,(n_e+n_c)}{n_e\,n_c}}$

و في حالة $n_e = n_c$ فإن المعادلة تصبح $g = \sqrt{\frac{2\,\chi^2}{n}}$

وفي حالة تحليل التباين بين مجموعتين مستقلتين $g = \sqrt{\frac{F\,(n_e+n_c)}{n_e\,n_c}}$

و في حالة $n_e = n_c$ فإن المعادلة تصبح $g = \sqrt{\frac{2\,F}{n}}$

وتحســـب قيمة g من تحليل التباين للمقارنة بين المجموعتين التجريبية والضـــابطة في حالة df= 1 للبسط أما إذا كانت df أكبر من ذلك؛ فإننا لا نستخدم المعادلة السابقة.

وفي حالة المقارنة بين مجموعتين مرتبطتين أو متناظرتين؛ مثل الفرق بين القياسين القبلي B والبعدي A نستخدم المعادلة

$$g = \frac{(\overline{X}_A - \overline{X}_B)}{S_{A-B}} = \frac{\overline{D}}{S_D} = \frac{\overline{D}}{\sqrt{\frac{SS_D}{n-1}}}$$

حيث $\overline{X}_A$, $\overline{X}_B$ متوسطا الدرجات قبلي وبعدي، S_{A-B} الانحراف المعياري للفرق بين الدرجات المتناظرة قبلي وبعدي.

كما يمكن تحديد S_{A-B} من خلال تحليل التباين لمجموعتين مرتبطتين من المعادلة

$$S_{A-B} = \sqrt{2MS_E} = \sqrt{2MS_w}$$

وتحسب قيمة g أو d غير المتحيزة من المعادلة

$$d_{unb.} = d(1 - \frac{3}{4\,df - 1})$$

حيث درجة الحرية df = (n-1) لإحدى المجموعتين المرتبطتين.

أو درجة الحرية = $(n_1 + n_2 - 2)$ للمجموعتين المستقلتين.

مؤشر دلتا Δ لجلاس

يمكن تقدير قيمة حجم الأثر دلتا لجلاس (Glass, et al., 1981) من المعادلة:

$$\Delta = \frac{\overline{X}_1 - \overline{X}_2}{S_{con.}}$$

حجم الأثر في المجتمع:

$$\theta = \frac{\mu_1 - \mu_2}{\sigma_{con.}}$$

التحويل بين تقديرات حجم الأثر وتقدير بعض قيم حجم الأثر الشـــهيرة (في Cohen, 1988; Borenstein, et al., 2009; Rosenthal, 1994; Sánchez-Meca, et al., 2003)

تحويل d إلى g

في حالة مجموعتين مرتبطتين، فإن: $g = d\sqrt{\frac{df}{n}} = d\sqrt{\frac{N-1}{n}}$

و في حالة مجموعتين مستقلتين، فإن: $g= d\sqrt{\frac{df}{N}} =d\sqrt{\frac{N-2}{N}}$

حيث: $N= n_1 +n_2$

ويلاحظ أن g تعطي تقديرات كبيرة مع العينات الصغيرة.

تحويل r *إلى* g

في حالة مجموعتين مرتبطتين:

$$g = \frac{r}{\sqrt{1-r^2}}\sqrt{\frac{df}{n}}$$

$$g = \frac{r}{\sqrt{1-r^2}}\sqrt{\frac{n-1}{n}}$$

وفي حالة مجموعتين مستقلتين $g = 2r/\sqrt{1-r^2}$

تقدير g باستخدام قيمة t للمقارنة بين مجموعتين متناظرتين أو مرتبطتين

$$g = t/\sqrt{n}$$

في حالة مجموعتين مستقلتين $g = t/\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$

في حالة مجموعتين مرتبطتين $g = (\bar{x}_e - \bar{x}_c)/S_{e-c}$

حيث S_{e-c} هي الانحراف المعياري للفرق بين الدرجات المتناظرة

ومن تحليل التباين $S_{e-c} = \sqrt{2MS_w}$

تقدير d من خلال قيمة Z، t، F

$$d = \frac{t(n_1+n_2)}{\sqrt{n_1 n_2 df}} = \frac{2t}{\sqrt{df_w}} = \frac{2z}{\sqrt{df_w}}$$

حيث إن df=n-1 في حالة الفرق بين مجموعتين مرتبطتين $d = t\sqrt{2(1-r_{xy})/n}$

و df= n_1+n_2-2 في حالة الفرق بين مجموعتين مستقلتين $d = t\sqrt{(n_1-n_2)/n_1 n_2}$

وفي حالة تحليل التباين في اتجاه واحد $d= 2\sqrt{\frac{F}{df_w}}$

حيث df_w هي درجة حرية الخطأ أو درجة الحرية داخل المجموعات.

وحالة الفرق بين القياسين القبلي والبعدي، نستخدم المعادلة: $d= \overline{D} / SE_{\overline{D}}$

حيث: $SE_{\overline{D}} = S_{\overline{D}} / \sqrt{2(1-r)}$

حيث r هو معامل الارتباط بين درجات الأفراد في القياسين القبلي والبعدي.

و $S_{\overline{D}}$ هو الانحراف المعياري للفروق بين القياسين.

مثال: من القياسين القبلي والبعدي كان متوسطهما $\bar{x}_1 = 103$ و الأخرى $\bar{x}_2 = 100$ والانحراف المعياري لفروق الدرجات 5.5، وكان حجم العينة 50 والارتباط بين القياسين القبلي والبعدي 0.7 فإنه يمكن حساب d على النحو الآتي:

$$SE_{\overline{D}} = \frac{5.5}{\sqrt{2(1-0.7)}} = 7.1005$$

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

تقدير قيمة d من g

من مجموعة أو مجوعتين مرتبطتين $d= g\sqrt{(n_1 + n_2)/df}$

$$= g \sqrt{\frac{n_1+ n_2}{n_1+ n_2-1}}$$

من مجموعتين مستقلتين $d= g\sqrt{(n_1 + n_2)/df}$

$$= g \sqrt{\frac{n_1+ n_2}{n_1+ n_2-2}}$$

تقدير قيمة d في حالة معاملات الارتباط R^2, η^2,r:

في حالة ارتباط بيرسون $d= 2r /\sqrt{(1 - r^2)}$

في حالة مربع ايتا $d= 2\eta /\sqrt{(1 - \eta^2)}$

في حالة الارتباط المتعدد $d= 2R /\sqrt{(1 - R^2)}$

تحويل الارتباط الثنائي r_{pb} **إلى** d $d= \frac{r_{pb}}{(1-r^2)p_1p_2}$

حيث إن r_{pb} معامل الارتباط الثنائي و P_1، P_2 نسبتي تحقيق الحدث في المجموعتين.

تقدير قيمة d إذا توفرت بيانات عن نسب من حققوا تقدير معين (P_c, P_e)

$$d = Z_C - Z_E$$

حيث Z_e هي الدرجة المعيارية المقابلة للنسبة المئوية العليا (P_e) من المساحة تحت المنحنى الاعتدالي Z_c معرفة بالمثل، يتم الحصول على قيم Z_c، Z_e من جداول التوزيع الاعتدالي عن طريق قيم P_c، P_e.

تقدير قيمة d عند توفر بيانات عن مستويات الدلالة وحجم العينة

في بعض التقارير البحثية، رغم استخدام اختبارات الدلالة إلا أنه في التقرير البحثي يسجل فقط مستوى الدلالة الإحصائية (p) وحجم العينة. في مثل هذه الحالات يتم تحويل قيمة مستوى الدلالة الإحصائي إلى قيمة t المناظرة (Z أو r) عن طريق جداول t الإحصائية وباستخدام حجم العينة لتحديد درجات الحرية، وبعد ذلك يُحسب حجم الأثر باستخدام المعادلة $r = Z/\sqrt{n}$.

تحويل χ^2 إلى d:

عندما df= 1 $\quad d= \sqrt{4\,\chi^2/(n - \chi^2)}$

عندما df >1 $\quad d= \sqrt{4\,\chi^2/n}$

تحويل OR إلى d:

$$d= \ln(OR)(1.81)/\pi$$

حيث OR هي نسبة الأرجحية و ln هو اللوغاريتم الطبيعي.

ويمكن تقدير قيمة d بفترة ثقة يحددها الباحث 99% أو 95%، وذلك باستخدام المعادلة التالية: $CI_d= d \pm Z_{\alpha/2}.\ S_d$

مثال: إذا كان $n_a=n_b=20$ و $d=0.7$ $\quad SE= 0.326$

$$CI_d= 0.7 \pm 1.96(0.326)= [0.06,\ 1.34]$$

ويلاحظ أن المؤشر d أكثر تحيزاً لحجم أثر المجتمع؛ إذ يعطي قيمة مرتفعة، وخاصة مع العينات الصغيرة، ولذلك يفضل استخدام المؤشر d مع العينات الكبيرة.

حجم الأثر باستخدام معاملات الارتباط ومربعاتها

توجد أنواع مختلفة من معاملات الارتباط نتيجة اختلاف طبيعة البيانات، إلا أن معامل ارتباط بيرسون يعتبر أكثر معاملات الارتباط استخداماً، ويستخدم لمعرفة العلاقة الخطية بين متغيرين متصلين. ويعد مربع معامل الارتباط، لبيرسون أحد مؤشرات حجم الأثر؛ إذ يدل على مقدار تأثر تباين أحد المتغيرين بتباين المتغير الآخر، وهذه هي الفكرة نفسها لحجم الأثر، ومن هنا يمكن القول: إن مربع معامل ارتباط بيرسون هو أحد مؤشرات حجم الأثر ويفسر (مثل مربع إيتا) على أنه النسبة المئوية للتباين المُفَسر (فام، 1997). أما بالنسبة للدراسات التنبئية، التي تعتمد على الانحدار المتعدد؛ فيمكن الحصول على قيمة (R^2)، وهو مربع معامل الارتباط المتعدد بين مجموعة من المنبئات من جهة، والمتغير المتنبأ به من جهة أخرى (نصار، 2006). ومع ذلك ينصح بعض الباحثين بعدم استخدام مربع معامل الارتباط كحجم أثر في بعض الحالات. ويذكر جريسوم وكيم (Grissom, & Kim, 2005) الأسباب التالية للتخلي عن القيمة (r^2) كمؤثر لحجم الأثر:

1- أن تربيع قيمة r الصغيرة الدالة ينتج عنها قيم صغيرة. مثال ذلك معامل الارتباط 0.3 يكون مربعه 0.09، وهذا تقدير أقل من قيمته العملية.
2- أن القيمة المربعة لـــ r ليس لها اتجاه (موجب أو سالب)؛ ولذلك يجب على الباحث أن يذكر قيمة r واتجاهه إلى جانب r^2.
3- ظهور أنواع أخرى لحجوم الأثر أكثر قوة إحصائية من مربع معامل الارتباط.

وعلى الرغم من النقد السابق لمعاملات الارتباط ومربعاتها فإنها تستخدم على نطاق واسع في بحوث ما وراء التحليل (التحليل البعدي) كمقياس للعلاقة الخطية بين متغيرين متصلين خطيين (X, Y). وهو يحسب بالمعادلة:

تقدير قيمة r من الانحدار الخطي البسيط

ان معادلة الانحدار الخطي البسيط هي $y= b_0+ b_1x_1$ ، ومنها نحسب معامل الارتباط r_{xy} من المعادلة $r_{xy}= b_1 (\frac{S_x}{S_y})$ (Tabachnick, & Fidell, 2018)

حيث S_x هو الانحراف المعياري للمتغير المنبئ x و S_y الانحراف المعياري للمتغير المتنبأ به y و b_1 هو ميل المتغير x على y، أو معامل الانحدار غير المعياري.

معامل الارتباط ϕ

يحسب معامل الارتباط ϕ (Frey, 2018) باستخدام المعادلة $\phi = \sqrt{\chi^2/n}$ أو r_ϕ عند درجة حرية تساوي 1 لقيمة χ^2 وهو يستخدم في حالة الجداول (2 x2) مثل اختبار الاستقلال بين متغيرين ثنائيين، عند درجات حرية تساوي واحداً، فإذا كانت ϕ=0 فإن ذلك يعني الاستقلال بين المتغيرين. وهو يفسر مثل معامل الارتباط r. ويمكن تحويل قيمة d إلى r إذا توافرت قيمة d من المعادلة:

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

$$r = d/\sqrt{d^2(n_1 + n_2)^2/n_1 n_2}$$

معامل فاي لكرامر ϕ_c

يستخدم في حالة الجدول المزدوج بأي تقسيم (R ≥2, C ≥2) (Frey, 2018)

$$\phi_c = \sqrt{\frac{\chi^2}{n\,(K-1)}}$$

حيث n عدد الأفراد، K عدد الصفوف (R) أو عدد الأعمدة (C) (أيهما أصغر)

والجدول التالي يعرض حجم الأثر لكوهين (1988)، المتعلق بمعامل ϕ لكرامر في حالات وجود تقسيمات مختلفة خاصة بالجانب الأصغر لجدول التوافق، وهي 2, 3, 4. لاحظ أنه عندما يكون الجانب الأصغر 2 فإن درجات الحرية تساوي الواحد، وهذا يعطي حالة خاصة، هي قيمة ϕ (Rosenthal, & Rosnow, 2003).

قيم كوهين الخاصة بمعامل ϕ لكرامر أو ϕ_c

عدد أقسام الجانب الأصغر بجدول التوافق	حجم الأثر		
	الضعيف	المتوسط	المرتفع
في حالة وجود قسمين (df_{small}=1)	0.10	0.30	0.50
في حالة وجود ثلاثة أقسام (df_{small}=2)	0.07	0.21	0.35
في حالة وجود أربعة أقسام (df_{small}=3)	0.06	0.17	0.29

معامل الارتباط الثنائي الأصيل

يمكن حساب قيمة معامل الارتباط الثنائي الأصيل من المعادلة (Frey, 2018)

$$r_{pb} = \frac{(y_1 - y_2)\sqrt{p_1 p_2}}{S_y}$$

حيث y_1 وy_2 متوسطا المجموعتين، وP_1 وP_2 نسبة المشاركين بكل مجموعة، وS_y الانحراف المعياري للمتغير y. ويلاحظ أن معامل الارتباط الثنائي الحقيقي يميل إلى إعطاء قيم أقل لمعامل الارتباط في المجتمع، وخاصة عندما يكون حجم العينة صغيراً (أقل من 20)، وكذلك عدم التقارب بين حجمي عينتي المتغير الثنائي. ولذلك يصحح هذا التحيز باستخدام المعادلة التالية:

$$r'_{pb} = r_{pb} + \frac{r_{pb}\,(1-r_{pb}^2)}{2(n-3)}$$

معامل الارتباط الرباعي

يمكن تقدير قيمة معامل الارتباط الرباعي، عندما يكون هناك متغيران اسميان ثنائيان (0/1)، حيث يمكن تحديد قيم التكرارات في الخلايا التي تنتج من العلاقة بين المتغيرين (Frey, 2018).

	0	1
0	a	B
1	c	D

$$r = \cos\left[\frac{180°}{1+\sqrt{\frac{ad}{cb}}}\right]$$

تحويلات بعض قيم الاختبارات الإحصائية من وإلى r_{pb}

للقيام بدراسة منهجية تعتمد على حجم الأثر، لابد من توحيد قيم حجم الأثر عبر الدراسات؛ لذلك من المهم عرض التحويلات الرياضية بين أشهر قيم الاختبارات الإحصائية ومعامل الارتباط الثنائي (في Cohen, 1988; Borenstein, et al., 2009; Rosenthal, 1994; Sánchez-Meca, et al., 2003)

تحويل t إلى r_{pb}

في حالة المقارنة بين مجموعتين مستقلتين $r_{pb} = \sqrt{\frac{t^2}{t^2+(n_1+n_2-2)}}$

في حالة المقارنة بي مجموعتين مرتبطتين $r_{pb} = \sqrt{\frac{t^2}{t^2+(n-1)}}$

في حالة تحليل التباين في اتجاه واحد واستخدام F $r_{pb} = \sqrt{\frac{F}{F+(n_1+n_2-2)}}$

تحليل التباين في اتجاه واحد باستخدام t $r_{pb} = t/\sqrt{(t^2 + df_w)}$

تحويل d إلى r_{pb}

من مجموعتين مرتبطتين $r_{pb}= \sqrt{\frac{d^2}{(d^2+1)}} = \frac{d}{\sqrt{(d^2+1)}}$

من مجموعتين مستقلتين $r_{pb}= d / \sqrt{d^2 + (\frac{1}{p_1 p_2})}$

حيث $P_1= (\frac{n_1}{n_1+n_2})$ هي نسبة المشاركين في المجموعة الأولى، $P_2=(\frac{n_2}{n_1+n_2})$ هي نسبة المشاركين في المجموعة الثانية. وعندما $P_1=P_2=0.5$ و $n_1 = n_2$ فإن المعادلة هي:

$$r_{pb}=\frac{d}{\sqrt{(d^2+4)}}$$

تحويل قيمة مان-ويتني U إلى r_{pb}

$$r_{pb} = 1 - \frac{2u}{n_e n_c}$$

حيث U هي قيمة مان ويتني، و n_e, n_c هي أعداد المجموعتين التجريبية والضابطة.

تحويل r_{pb} إلى r_b

$$r_b= \frac{r_{pb}\sqrt{n_e n_c}}{|z|\ (n_e+n_c)}$$

حيث r_{pb} هو معامل الارتباط الثنائي الأصيل، وr_b هو معامل الارتباط الثنائي غير الأصيل، وZ هي الدرجة المعيارية عند نقطة قطع النسبتين p و q.

ويمكن تقدير قيمة r باستخدام قيمة z و n باختبار للفرق بين مجموعتين سواء كانتا مرتبطتين أم مستقلتين، من خلال المعادلة التالية $r = \frac{z}{\sqrt{n}}$

مثال: إذا كانت النتيجة دالة عند $p = 0.007$ ، فإن قيمة Z المقابلة تساوي 2.44 لاختبار ذي ذيل واحد. فإذا كان عدد الأفراد المشاركين 40 فإن حجم الأثر يساوي:

$$r = \frac{z}{\sqrt{n}} = \frac{2.44}{\sqrt{40}} = 0.39$$

تقدير قيمة r من تحليل التباين الأحادي

$$r= \sqrt{SS_B/SS_T}$$

حيث SS_B هو متوسط مجموع المربعات بين المجموعات، و SS_T هو متوسط مجموع المربعات الكلي.

ملاحظات (في Ellie, 2010):

1- يمكن تسجيل r=0 عندما تشير الدراسة إلى عدم وجود دلالة إحصائية ولا يمكن تقدير قيمة ارتباط معين.

2- يجب اســـتخدام حجم الأثر المصـــحح كلما أمكن، وذلك بتقدير قيمة معامل الارتباط غير المتحيز من المعادلات التالية:

في حالة r $r_{adj}= \sqrt{1-\frac{(1-r^2)(n-1)}{n-2}}$

في حالة R $R_{adj}= \sqrt{1-\frac{(1-R^2)(n-1)}{n-k-1}}$

حيث k عدد متغيرات التنبؤ التي يشملها معامل الارتباط.

3- في حالة التعامل إحصائياً مع r يتم تحويله إلى Z_r، لأن توزيع r ملتو، بينما توزيع Z_r يميل إلى الاعتدالية، وهي تحسب بالمعادلة:

$$Z_r= \frac{1}{2}\ln(\frac{1+r}{1-r})$$

ويحسب تباين Z_r باستخدام المعادلة:

$S^2_{z_r} = \frac{1}{n-3}$ (أو مربع خطأ Z_r)

ويحسب تقدير معامل الارتباط بالمجتمع بفترة ثقة معنية باستخدام المعادلة:

$$IC_{Zr}= Z_r \pm Z_{\alpha/2}.\ S_{Zr}$$

4- يمكن تقدير قيمة r من تصميم تجريبي لمجموعتين بتحويل t إلى r.

5- قيمة r_{pb} تكون مؤشراً جيداً لحجم الأثر من جدول (2×2) الذي يشمل مستويين للمتغيرين.

تقدير قيمة حجم الأثر باستخدام R^2, ε^2, ω^2, η^2 (انظر Tabachnick, & Fidell, 2018)

- **تقدير مربع معامل الارتباط المتعدد R^2 (باستخدام تحليل التباين)**

أ) في حالة تحليل التباين البسيط $R^2= (SS_B)^2/(SS_T)^2$

ب) في حالة تحليل التباين المتعدد $R^2= (SS_{eff})^2/(SS_T)^2$

ج) في حالة تحليل الانحدار $R^2= (SS_{reg})^2/(SS_T)^2$

حيث SS_B التباين بين المجموعات، SS_{eff} التباين نتيجة المؤثر، SS_{reg} تباين الانحدار.

والخطأ المعياري لمربع الارتباط المتعدد، ويتحدد بالمعادلة

$$SE_{R^2}=\sqrt{\frac{4R^2(1-R^2)(n-k-1)}{(n^2-1)(n+3)}}$$

حيث n عدد أفراد المجموعة، k عدد المتغيرات المتنبأ بها.

وتكون فترة الثقة لمعامل الارتباط المتعدد هي: $IC_{R^2} = R^2 \pm Z_{\alpha/2}.\ SE_{R^2}$

ويمكن تقدير قيمة معامل الارتباط المتعدد R غير المتحيز باستخدام المعادلة التالية، وخاصة عندما يكون حجم العينة صغيراً

$$R_{adj}=\sqrt{1-\frac{(n-1)(1-R^2)}{n-k-1}}$$

حيث k عدد المنبئآت، n عدد أفراد العينة.

- **تقدير قيمة مربع إيتا η^2:**

يستخدم في حالة تحليل التباين بأنواعه، وهو مقياس لدرجة الارتباط في العينة وتفسر η^2_{xy} مضروبة في 100 بأنها نسبة التباين في Y المُفَسَر بالتباين في X والشيء نفسه بالنسبة لـ R^2. أما η^2 و R^2 الجزئية؛ فهي نسبة التباين الكلي الذي يرجع إلى المتغير المعطي مع عزل أثر العوامل الأخرى المصاحبة له. وقد تَبَين أن η^2 متحيزة؛ حيث تعطى تقدير أكبر للتباين المفسر، ولكنه يقل عندما يكون حجم العينة كبيراً.

ويمكن تقدير قيمته من تحليل التباين البسيط في اتجاه واحد من المعادلة التالية:

$$\eta^2 = \frac{SS_{B.}}{SS_T} = \frac{SS_{ef}}{SS_T}$$

وفي تحليل التباين لقياسات متتابعة (المتكررة) نستخدم المعادلة:

$$\eta^2 = \frac{SS_B}{SS_B + SS_P}$$

وفي تحليل التباين العاملي (الثنائي):

$$\eta^2_A = \frac{SS_A}{SS_T} = \frac{SS_A}{SS_A+SS_w}$$
$$\eta^2_B = \frac{SS_B}{SS_T} = \frac{SS_B}{SS_B+SS_w}$$

$$\eta^2_{AB} = \frac{SS_{AB}}{SS_T} = \frac{SS_{AB}}{SS_{AB} + SS_w}$$

مربع إيتا الجزئي

$$\eta^2_P = \frac{SS_{ef.}}{SS_{ef.} + SS_{e.}}$$

$$\eta^2_P = \frac{(df_{ef.})(F_{ef.})}{(df_{ef.})(F_{ef.}-1)+df_{e.}} = \frac{F_{ef.}}{F_{ef.}+df_{w.}}$$

- **تقدير مربع أوميجا ω^2 ومربع أبيسلون ε^2**

أ- في تحليل التباين أحادي: $\omega^2 = \frac{SS_{ef.}-(df_{ef.})(MS_{w.})}{SS_{T.}+MS_{w.}}$

$$\omega^2 = \frac{(F_{eff}-1)(df_{eff})}{(F_{eff}-1)(df_{eff})+nk} = \frac{(k-1)(F-1)}{(k-1)(F-1)+N}$$

ب- في تحليل التباين العاملي (الثنائي): $\omega^2 = \frac{(df_A)(F_A-1)}{(df_A)(F_A-1)+N_a}$

ج- مربع إيبسلون $\varepsilon^2 = \frac{SS_B-(K-1)MS_W}{MS_t}$

ملاحظات عن η^2 و ω^2 و ε^2 (في Cohen, 1988; Ellie, 2010):

1- مربع إيتا η^2 مثل R^2 يعرفان بأنهما نسبة الارتباط، وهي نسبة التباين في المتغير التابع التي ترجع إلى المتغير المســتقل، فإذا كانت $\eta^2=0.45$ فإن ذلك يعني أن 45% من التباين في المتغير التابع (DV) يرجع إلى المتغير المستقل(IV).

2- مربع أوميجا ω^2 ومربع أبسيلون ε^2 يعتبران قيمتين غير متحيزتين.

3- عندما يكون لدينا تحليل تباين لمجموعتين مستقلتين فإن $\eta=r$.

4- مربع أوميجا ω^2 أوسع استخداماً من مربع إيبسلون ε^2 في تقدير حجم أثر المعالجات التجريبية.

5- η^2 لا تكون أقل من الصفر، بينما قد تكون نظرياً قيمة ω^2 و ε^2 أقل من الصفر.

6- حجم الأثر المعياري هو التباين المفسر مقارنة بالتباين غير المفسر.

7- في حالة استخدام تحليل التباين يفضل قيمة مربع إيتا (η^2) أكثر من d، وt.

8- يستخدم R^2 في حالة الانحدار الخطي، بينما يستخدم η^2 في حالة التنبؤ غير الخطي. وفي حالة الانحدار الخطي نجد أن $\eta^2=R^2$.

تحويل t^2 إلى η^2 (أنظر Tabachnick, & Fidell, 2018)

في حالة مجموعتين مرتبطتين: $\eta^2= t^2/ (t^2+df) = t^2/ [t^2+ (n-1)]$

في حالة مجموعتين مستقلتين: $\eta^2= t^2/ [t^2+ (n_1+n_2-2)]$

والشيء نفسه بالنسبة لتحويل Z^2 *إلى* η^2: $\eta^2= Z^2/ (Z^2+df)$

تحويل d^2 *إلى* η^2: $\eta^2= d^2/ (d^2+4)$

تحويل ω^2 *إلى* η^2:

$$\eta^2= \frac{\left[df_{ef.}+\frac{nk\omega^2}{(1-\omega^2)}\right]\left[\frac{1}{df_{w.}}\right]}{\left[df_{ef.}+\frac{nk\omega^2}{(1-\omega^2)}\right]\left[\frac{1}{df_{ef.}}\right]+1}$$

حيث k عدد المجموعات، df_w درجة حرية الخطأ، df_{ef} درجة حرية المتغير الأساسي، n عدد أفراد العينة.

تحويلات بين η^2، Cohen's f^2، ε^2:

$$f^2 = \eta^2/(1-\eta^2)$$

$$\eta^2 = f^2/(1+f^2)$$

في حالة تحليل التباين:

$$\eta_p^2 \text{ (الجزئية)} = \frac{F(df_{num.})}{[F(df_{num.})+df_{den}]}$$

$$\varepsilon^2 = \frac{(F-1)df_{num}}{F(df_{num})+df_{den}}$$

$$f_p^2 = \frac{\eta_p^2}{(1-\eta_p^2)}$$

مثال: من الجدول التالي لملخص تحليل التباين الثنائي احسب ω_A^2، ω_B^2، ω_{AB}^2 وكذلك تأثير المتغيرات A وB والتفاعل بينهما على المتغير التابع

مصدر التباين	SS	df	MS	F
العامل A	120	1	120	* 24
العامل B	20	2	10	2
التفاعل A × B	80	2	40	* 8
داخل المجموعات	120	24	5	
المجموع	340	29		

التأثير الأساسي للمتغير المستقل A:

$$\omega_{\mathrm{A}}^2 = \frac{\mathrm{SS_A} - (\mathrm{a}-1)\mathrm{MS_w}}{\mathrm{SS_t} + \mathrm{MS_w}} = \frac{120 - (2-1)5}{340 + 5} = .333$$

التأثير الأساسي للمتغير المستقل B:

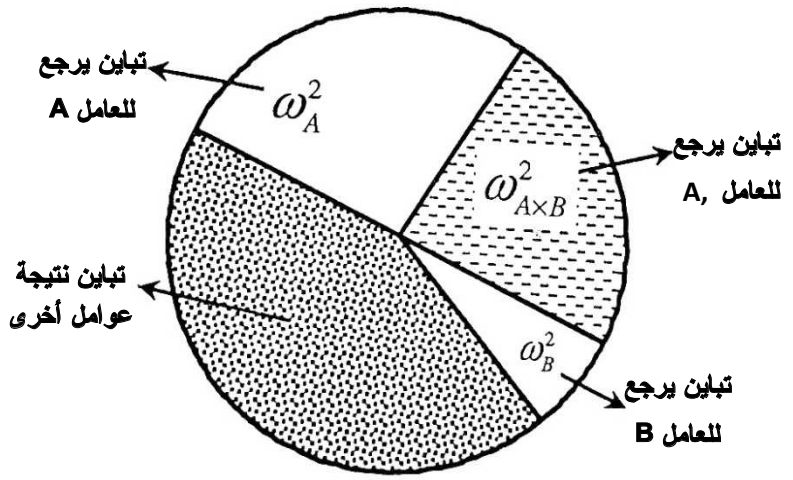
$$\omega_{\mathrm{B}}^2 = \frac{\mathrm{SS_B} - (\mathrm{b}-1)\mathrm{MS_w}}{\mathrm{SS_t} + \mathrm{MS_w}} = \frac{20 - (3-1)5}{340 + 5} = .043$$

تأثير التفاعل بين المتغيرين المستقلين (A × B):

$$\omega_{\mathrm{A\times B}}^2 = \frac{\mathrm{SS_{A\times B}} - (\mathrm{a}-1)(\mathrm{b}-1)\mathrm{MS_w}}{\mathrm{SS_t} + \mathrm{MS_w}} = \frac{80 - (2-1)(3-1)5}{340 + 5} = .203$$

شكل (2/2)

حجم *الأثر للعاملين المستقلين A, B والتفاعل بينهما (المؤلفان)*



مما سـبق يتبين أن 33.3 % من التباين الكلي يقترن بتأثير العامل A، وأن 4.3 % من التباين الكلي يقترن بتـأثير العـامل B، وأن 20.3 % يقترن بتـأثير التفـاعل بين المتغيرين؛ وهـذا يعني أن (33.3+4.3+20.3) أو 57.9 % من التباين الكلي في المتغير موضـوع الدراسـة يقترن بالعاملين A والتفاعل بينهما.

تحويل η^2 *إلى* ω^2

$$\omega^2 = \frac{[\mathrm{df_w}(\eta^2)/(1-\eta^2] - \mathrm{df_{ef}}}{[\mathrm{df_w}(\eta^2)/(1-\eta^2] + \mathrm{nk}}$$

$$\omega^2=1-\frac{MS_w}{MS_T}$$

مربع إيبسلون

$$\varepsilon^2=\frac{SS_{B.}-df_{B.}(MS_{w.})}{SS_T}$$

قيمة إيبسلون

$$\varepsilon=\sqrt{\frac{df_{B.}(F-1)}{df_{B.}(F)+df_{w.}}}$$

$$\text{Cohen's } f^2=R^2/(1-R^2)$$

$$\text{Cohen's } f^2=\eta^2/(1-\eta^2)$$

مثال: إذا علمت القيم التالية من تحليل التباين، احسب قيم **η^2 و ω^2 و ε^2**

$SS_{B.}=937.82$ $SS_{w.}=714.38$ $SS_{T.}=1652.20$

$MS_{B.}=234.46$ $MS_{w.}=9.53$ $F(4.75)=24.6$ ($p<0.001$)

$$F=\sqrt{\frac{(K-1)(F-1)}{N}}=\sqrt{\frac{4(23.6)}{86}}=1.086$$

$$\eta^2=\frac{SS_{B.}}{SS_{T.}}=\frac{937.820}{1652.20}=0.5676$$

$$\varepsilon^2=\frac{SS_{B.}-(K-1)MS_{w.}}{SS_T}=\frac{937.82-(4)(9.53)}{1652.20}$$

$$=\frac{937.82-38.12}{1652.20}=\frac{899.7}{1652.2}=0.541$$

$$\omega^2=\frac{SS_{B.}-(K-1)MS_{w.}}{SS_T+MS_{w.}}=\frac{937.82-(4)(9.53)}{1652.20+9.53}=\frac{937.82}{1661.73}=56.44$$

مثال: في حالة ثلاث معالجات مرتبطة كل مجموعة (4)، وكانت نتائج تحليل التباين على النحو التالي:

$SS_{ef.}=1095.5$ مجموع مربعات التأثير

$SS_{t.s}=596.5$ مجموع مربعات تفاعل المشاركين والمعالجة

$SS_{T.}=2290.25$ مجموع المربعات الكلي

$SS_{sub.}=598.25$ مجموع المربعات بين المشاركين

ملحوظة للحصول على MS نقسم SS على درجات الحرية df لكل قيمة مناظرة.

$$\omega^2 = \frac{(\mathrm{K}-1)(\mathrm{MS_{ef.}} - \mathrm{MS_{t.s}})}{\mathrm{SS_T} + \mathrm{MS_{sub.}}} = \frac{(3)\left[\frac{1095.5}{3-1}\right] - \left[\frac{596.5}{(3-1)(4-1)}\right]}{(2290.25) + \left[\frac{598.25}{4-1}\right]} = 0.36$$

هذا يعني أن 36% من التباين الكلي يرجع إلى تأثير المتغير المستقل على المتغير التابع.

تقدير قيمة حجم الأثر من مستوى الدلالة α أو الاحتمال p

في بعض التقارير البحثية يسجل الباحث فقط مستوى الدلالة الإحصائية α أو قيمة الاحتمال p وحجم العينة. وفي مثل هذه الحالة يمكن تحويل قيمة مستوى الدلالة إلى القيمة المناظرة لها (Z أو r) باستخدام جدول المساحات أسفل المنحنى الاعتدالي المعياري أو Z، وباستخدام حجم العينة لتحديد درجة الحرية المناظرة. وبعد ذلك يحسب حجم الأثر باستخدام المعادلة: حجم الأثر= $r = \frac{Z}{\sqrt{n}}$. ويوضح بورنستين وزملاءه (Borenstein, 2000; Borenstein et al., 2009; Rothstein & Borenstein, 2014) أنه يمكن تقدير قيمة حجم الأثر من خلال قيمة مستوى الدلالة إذا كان اتجاه النتيجة معلوماً عن طريق تحديد الإحصائي المناظر لقيمة مستوى الدلالة. ولذلك يجب تحديد اتجاه الفرض أو اتجاه الأثر (موجباً أم سالباً) لمعرفة قيمة p أتتبع ذيلاً واحداً أم ذيلين. وبالنسبة لاختبار من طرف واحد، تحسب قيمة p باستخدام المعادلة التالية:

$$p = 1 - \phi(\pm|z|)$$

حيث نختار الإشـــارة (+) عندما يكون الفرق في الاتجاه المتوقع، وإشـــارة (-) في الحالة الأخرى، وبالنسبة لاختبار من طرفين، تحسب قيمة P باستخدام المعادلة التالية:

$$P = 2 - \phi(\pm|z|)$$

حيث $\phi|z|$ هو التوزيع المتجمع، وهذا موجود كجدول في معظم كتب الإحصاء ومن جدول (z).

تقدير قيمة حجم الأثر للمتغيرات الثنائية

إن المتغيرات الثنائية تحدد وقوع الحدث من عدمه مثل الدراسات الطبية التي تتعامل مع متغير مثل حياة المريض أو وفاته، الإصابة أو عدم الإصابة. ويمكن تقدير قيمة حجم الأثر لهذه الدراسات من خلال الخطوات التالية (Borenstein, et al., 2008):

1- اختبار إحدى النتائج كحدث ثنائي (يحدث أو لا يحدث).

2- حساب الاحتمالية للحدث الذي نهتم به في المجموعة التجريبية P_e والمجموعة الضابطة P_c.

3- تحويل قيم الاحتمالية إلى قيمة Z من خلال إيجاد الدرجات Z_e, Z_c المناظرة إليها في جدول التوزيع الطبيعي.

4- تقدير قيمة g باستخدام المعادلة التالية:

$$g= (Z_e - Z_c) \sqrt{\frac{n_e + n_c}{n_e\, n_c}}$$

حيث n_e حجم العينة للمجموعة التجريبية، n_c حجم العينة للمجموعة الضابطة.

Z_e قيمة Z المقابلة للاحتمال، P_e في المجموعة التجريبية، Z_c قيمة Z المقابلة للاحتمال P_c في المجموعة الضابطة.

تقدير حجم الأثر للفرق بين معامل الارتباط بمجموعتين مستقلتين (Cohen, 1988):

وذلك باستخدام المعادلة التالية:

$$\text{Conhen's } q= Zr_1 - Zr_2$$

حيث Zr_1 هو تحويل فيشر لمعامل الارتباط بالمجموعة الأولى r_1، و Zr_2 هو تحويل فيشر لمعامل الارتباط بالمجموعة الثانية r_2. كما يمكن استخدام كوهين q في عينة واحدة بالمقارنة بين معامل الارتباط الملاحظ أو المحسوب ومعامل الارتباط النظري أو المتوقع بالمجتمع، ثم تحويلهما من خلال جدول فيشر، ثم العودة مرة أخرى إلى r.

تقدير قيمة حجم الأثر باستخدام نسبتين مستقلتين (Cohen, 1988):

يمكن استخدام المعادلة التالية:

$$Z_{(P_1 - P_2)} = \frac{P_1 - P_2}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}} = \frac{P_e - P_c}{\sqrt{pq(\frac{1}{n_e} + \frac{1}{n_c})}}$$

حيث $q= 1-p$ و $p= p_e - p_c$.

مثال: نفترض أن عدد الذين تحسنوا باستخدام العلاج النفسي في المجموعتين (1 و 2) هما: $n_1= 31$، $P_1= 0.61$، والذين تحسنوا باستخدام العلاج بالعقاقير $n_2= 36$، $P_2= 0.31$.

$$Z_{(P_1 - P_2)} = \frac{0.61 - 0.31}{\sqrt{\frac{(0.47)(0.53)}{31} + \frac{(0.47)(0.53)}{36}}} = 2.47$$

ويمكن تحويل هذه القيمة إلى P مرة أخرى باستخدام جدول المساحات أو إلى r كحجم أثر.

تقدير قيمة حجم الأثر باستخدام التكرارات بجدول (2×2)

نسبة الأرجحية Odds Ratio (OR) (Borenstein, et al., 2008)

نسبة الأرجحية هي احتمالية النجاح بالمجموعتين التجريبية (التي تخضع للمعالجة أو التدخل) بالنسبة إلى احتمالية النجاح في المجموعة الضابطة.

مع الالتزام بأماكن الحروف بالجدول (2×2) نحدد نسبة الأرجحية بالمعادلة:

	نجاح	فشل
المجموعة التجريبية	A	b
المجموعة الضابطة	C	d

$$OR= (a/b) /(c/d)$$

أو نسبة الأرجحية = $OR=\frac{ad}{cb}$

مثال: من الجدول التالي، احسب OR

المجموعة	فشل	نجاح
التجريبية ←	10	10
الضابطة ←	19	1

$$OR= (10/10) /(1/19)=19$$

كما يمكن تعريف الأرجحية بأنها: عدد الأفراد ذوي الحدث الذي نهتم به مقارنة بالأفراد الذين لا يمتلكون ذلك الحدث.

مثال: إذا شرب القهوة 24 فرداً، أصيب منهم 6 بصداع، لذلك فإن أرجحية حدوث الصداع لمن يشرب القهوة يساوي 6/18 أو 33%.

مما سبق يتبين أن OR تستخدم عندما يكون الناتج ثنائياً، كما يتبين أن مدى OR يراوح بين صفر و ∞، وعندما تكون OR >1، فإن هذا يشير إلى زيادة احتمالية الأرجحية بالنسبة للمجموعة المرجعية أو الضابطة. ونستخدم التحويل ln OR بدلاً من OR؛ لأن توزيع OR يتمركز حول 1 وبمدى يراوح بين صفر و∞، أما ln OR؛ فإنه يتمركز حول الصفر ويراوح بين ∞- و ∞+. والخطأ المعياري لإحصاء ln OR هو

$$SE (\ln OR) = \sqrt{\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}}$$

وفي حالة وجود صفر في إحدى الخلايا نحدد ln OR بالصيغة:

$$\ln OR= \ln \left(\frac{(a+0.5)(d+0.5)}{(c+0.5)(b+0.5)}\right)$$

بخطأ معياري:

$$S_e (\ln OR) = \sqrt{\frac{1}{(a+0.5)}+\frac{1}{(b+0.5)}+\frac{1}{(c+0.5)}+\frac{1}{(d+0.5)}}$$

مثال: احسب OR من الجدول التالي:

	فشل العلاج	نجاح العلاج
المجموعة التجريبية	66/a	34/b
المجموعة الضابطة	34/c	66/d

$OR = (66/34) / (34/66) = 3.73$

حجم الأثر لـ OR $ES_{OR} = \ln OR = 1.316$

الخطأ المعياري $SE_{(\ln OR)} = \left(\frac{1}{66} + \frac{1}{34} + \frac{1}{66} + \frac{1}{34}\right) = 0.299$

فترة الثقة 95% $95\%\ CI_{\ln OR} = 1.316 \pm 1.96(0.299) = [0.37, 1.90]$

بعد التحويل فترة الثقة 95% ← [1,478, 6.686]

ملاحظات (المصدر Borenstein, et al., 2009; Rosenthal, 1994; Sánchez-Meca, et al., 2003)

1- العلاقات السالبة تشير إلى قيم ln OR تكون محصورة بين 0 و 1.

2- العلاقات الموجبة تشير إلى قيم ln OR تكون محصورة بين 1 و∞.

3- عندما OR=1 فإن ذلك يعني عدم وجود فروق بين المجموعتين في نسبة الأرجحية، وعندها ln OR = 0.

4- تحول النتائج إلى Ln OR باستخدام اللوغاريتم الطبيعي $ES_{OR} = \ln(OR)$، ثم نعود عكسياً إلى النتيجة النهائية باستخدام اللوغاريتم الطبيعي $OR = e^{ES_{OR}}$

نسبة المخاطرة أو المجازفة Risk Ratio (RR) (في Rosenthal, 1994)

هو مؤشر يشير لنسبة المخاطرة في مجموعتين مختلفتين؛ حيث يكون البسط هو النسبة في المجموعة الضابطة والمقام هو النسبة في المجموعة التجريبية. وهي نسبة الخسارة أو المجازفة في المجموعتين. ولذلك تكون معادلة تقدير قيمة RR من الجدول (2×2)، وهي بذلك تقدم مقارنة بين مجموعتين من حيث الخسارة أو الضرر أو المرض، أما في نسبة الأرجحية فإنها تهتم بالمكسب أو الخلو من المرض. ويمكن تعريف المخاطرة بأنها عدد الأفراد ذوي الحدث موضع الاهتمام مقارنة بالعدد الكلي للأفراد.

	تحسن المرض	بقاء المرض	
المجموعة التجريبية	a	b	n_1
المجموعة الضابطة	c	d	n_2

$$RR= \frac{a/n_1}{c/n_2} = \frac{a/(a+b)}{c/(c+d)}$$

مثال: 24% فرد.

وعندما تكون نسبة المخاطرة أقل من 1 تكون المخاطرة منخفضة في المجموعة التجريبية مقارنة بالمجموعة الضابطة. أما إذا كانت نسبة المخاطرة أكبر من 2؛ فإن ذلك يشير إلى أن المخاطرة كبيرة. وبذلك تكون المخاطرة المتوسطة محصورة بين 1 و2. إن نسبة المخاطرة تقيس مخاطر حدث معين بالنسبة للمتغير المستقل وإن احتمالية المخاطرة المنخفضة تعادل تقريباً نسبة الأرجحية. ويلاحظ أنه عندما يكون الحدث شائعاً أو نادراً فإن الفرق بين المخاطرة والأرجحية يكون كبيراً. ويتحدد الخطأ المعياري للنسبة RR باستخدام المعادلة:

$$SE\ (\ln RR) = \sqrt{\frac{1}{a} + \frac{1}{b} - \frac{1}{a+b} - \frac{1}{c+d}}$$

وفي حالة عدم وجود تكرارات في إحدى الخلايا فإننا نضيف 0.5 بكل تكرار بالخلايا، مثلما فعلنا مع OR.

مثال: من الجدول (2 x2) التالي احسب نسبة المخاطرة:

	تحسن المرض	بقاء المرض	
المجموعة التجريبية	22	14	36
المجموعة الضابطة	10	22	32

$$RR= P_1/P_2 = \frac{14/36}{22/32} =0.57$$

عندما تساوي RR القيمة 1 فإن هذا يعني عدم وجود فروق بين المجموعتين.

عندما تكون RR أكبر من 1 فإن ذلك يعني أن المخاطرة في المجموعة التجريبية أكبر من المجموعة الضابطة.

فرق المخاطرة Risk Difference (RD) (في Rosenthal, 1994)

هو مؤشر للفرق بين نسبة المخاطرة في المجموعة التجريبية ونسبة المخاطرة في المجموعة الضابطة، وهو يتحدد بالمعادلة:

$$RD= \frac{a}{n_1} - \frac{c}{n_2}$$

ويتحدد الخطأ المعياري لفرق المخاطرة بالمعادلة:

$$SE_{RD}=\sqrt{\frac{ab}{n_1^3}+\frac{cd}{n_2^3}}$$

مثال: من الجدول التالي احسب فرق المخاطرة

	الموت	الحياة	
المجموعة الضابطة	11	143	154
المجموعة التجريبية	4	142	146

نسبة الموتى في المجموعة الضابطة = $\frac{11}{154}$

نسبة الموتى في المجموعة التجريبية = $\frac{4}{146}$

فرق المخاطرة $RD=\frac{a}{a+b}-\frac{c}{c+d}=.071-.027=.044=\%4.4$

وهذا يعني أن هناك فرقاً في المخاطرة بين المجموعة التجريبية والمجموعة الضابطة يساوي 4.4%.

الخطأ المعياري لـ RD

$$SE_{RD}=\sqrt{\frac{11(143)}{154^3}+\frac{4(142)}{146^3}}=0.025$$

$$=\sqrt{\frac{7.1(1-7.1)}{154}+\frac{2.7(1-2.7)}{146}}=0.025$$

فترة الثقة

$$CI_{RD}=0.044\pm 1.96(0.025)=[-.005,.093]$$

قيمة حجم الأثر في حالة ذي الحدين (BESD)

أسلوب تقدير قيمة حجم الأثر في حالة ذي الحدين The Binomial Effect Size Display تم تقديمه في عام 1982 على يد كل من روزنثال و روبين (Rosenthal & Rubin, 1982; 1986)؛ وذلك بهدف عرض قيمة حجم الأثر التجريبي؛ مثل نسبة الأرجحية (النجاح، الشفاء، التحسن) ونسبة المخاطرة (المرض، الوفاة،...) وهو يعطي الفرق بين تأثير المجموعة التجريبية والمجموعة الضابطة. ويستخدم BESD عندما يكون هناك توقع لحدوث واحدة من نتيجتين؛ مثل النجاح مقابل الفشل. وفيما يلي أسهل طريقة لتقدير قيمة BESD (Rosenthal & Rubin, 1986)

1) تحويل حجم الأثر إلى r
2) حساب نسبة نجاح المجموعة التجريبية باستخدام المعادلة:

$$Success\ (E)=0.5+(r/2)$$

3) حساب نسبة نجاح المجموعة الضابطة باستخدام المعادلة:

$$Success\ (C)=0.5 - (r/2)$$

مثال: من الجدول التالي يتبين لنا طرق تقدير قيمة RD وRR و OR، ومنه يمكن تقدير قيمة r المناظرة، ثم نحسب BESD المعياري على النحو التالي:

المجموعة	يموت	يعيش	عرض BESD اعتماداً على OR	عرض BESD اعتماداً على RR	عرض BESD اعتماداً على RD
Control	A	B			
Treat	C	D			
الدراسة (1)	53	47	1.27	1.13	0.06
	47	53			
الدراسة (2)	75	25	9.00	3.00	0.50
	25	75			
الدراسة (3)	95	5	361.00	9.00	0.90
	5	95			

توجد طريقة بسيطة للتوفيق وتقنين العلاقة بين المؤشرات الثلاثة (OR, RR,RD) حيث يحسب r المناظر ثم نحسب BESD المعياري كما جددنا سابقاً.

تقدير قيمة حجم الأثر لبعض الاختبارات اللابارامترية (الشربيني، 2011)

فيما يلي عدد من مؤشرات حجم الأثر المشتقة في بعض الاختبارات اللابارامترية:

اختبار مان ويتني Mann-Whitney

في اختبار مان ويتني لكي نحسب حجم الأثر فإننا نحسب معامل الارتباط الثنائي الأصيل باستخدام المعادلة:

$$r_{pb} = 1-(2U / n_1n_2)$$ (الشربيني، 2011)

حيث U القيمة الصغرى باختبار مان ويتني، و n_1, n_2 عدد أفراد المجموعتين.

وعندما يكون حجم العينة مناسباً وافتراض الاعتدالية متوافراً، فإنه يمكن تحويل U إلى z على النحو التالي:

$$|Z| = [U - (n_1n_2/2)]/[\sqrt{n_1n_2(n_1 + n_2 + 1)/12}$$

تكون Z دالة إذا كانت أكبر من أو تساوي القيمة الحرجة لـ Z لاختبار ذي ذيلين ودلالة $\alpha/2$، بالاعتماد على المنحنى الاعتدالي.

مثال: إذا كانت قيمة حجمي العينتين $n_1= n_2= 20$

$U=103$

$$r_{pb} = \left[1 - \frac{2(103)}{20(20)}\right]$$

$$|Z| = [103 - (20 \times 20)/2]/\sqrt{20 \times 20(41)/12}$$

$$= (103 - 200)/\sqrt{4100/3} = 97/37 = 2.624$$

$|Z| = 2.624$ وهي دالة عند 0.01

$$r = \frac{Z}{\sqrt{N}} = \frac{2.624}{\sqrt{40}} = 0.415$$

حيث $N=n_1+n_2$

كما يستخدم مع اختبار مان ويتني وويلكوكسن للعينات المستقلة المعادلة التالية: (Kerby, 2014)

$$r = \frac{2(\bar{R}_1 - \bar{R}_2)}{(n_1 + n_2)}$$

حيث $\bar{R}_1$ متوسط رتب المجموعة الأولى، وحجمها يساوي n_1.

$\bar{R}_2$ متوسط رتب المجموعة الثانية، وحجمها يساوي n_2.

مثال :

المجموعة (1)	المجموعة (2)		
R_1	**R_2**	**n**	**$\bar{R}$**
10	8	6	4.17
9	6	4	7.50
7	5	**Σ 10**	
4	3		
ΣR_1 30	2		
	1		
	ΣR_2 30		

$$r = \frac{2(7.50 - 4.17)}{10}$$

$$r = \frac{2(3.33)}{10} = 0.666$$

مثال:

النوع	N	متوسط الرتب	مجموع الرتب
إناث	539	424.97	229056.50
ذكور	180	166.46	29783.50
المجموع	719		
مان ويتني U	= 13493.5		
ويلكوكسن W	= 29783.5		
قيمة Z	= -14.638		

$$\therefore r = \frac{2(\bar{R}_1 - \bar{R}_2)}{(n_1 + n_2)}$$

$$= \frac{2(424.97 - 165.46)}{719} = 0.722$$

$$r = 1 - (2U/n_1 n_2)$$

$$= 1 - \frac{(13493.5)}{539(180)} = 0.722$$

ويمكن تحويل قيمة ويلكوكسن W إلى قيمة U باستخدام المعادلة التالية:

$$U = W - \frac{n(n+1)}{2}$$

كما يمكن استخدام المعادلة باختبار مان ويتني:

$$r = 2(\bar{R}_1 - \bar{R}_2)/\,(n_1 + n_2)$$

حيث $\bar{R}_1$ متوسط رتب المجموعة (1) و $\bar{R}_2$ متوسط رتب المجموعة (2).

معامل الارتباط الثنائي لرتب الأزواج المرتبطة (لاختبار ويلكوكسن)

إذا وجد الباحث أن الفرق بين الأزواج المرتبطة من الدرجات أو الفرق بين القياس القبلي والقياس البعدي دالاً فإنه يمكن تحديد قوة العلاقة بين المتغيرين المستقل والتابع باستخدام معامل الارتباط الثنائي لرتب الأزواج المرتبطة Matched – Pairs Rank Biserial Correlation، وهذا المعامل يحدد درجة العلاقة بين إشارة فروق الرتب والرتبة المناظرة لها، وهو يستخدم الصيغة الآتية (الشربيني، 2011):

$$r_{pb} = \frac{4/T - (\Sigma R +)_2(\Sigma R-)}{n(n+1)}$$

معامل الارتباط الثنائي

حيث: $\Sigma R +$ هو مجموع الرتب ذات الإشارة الموجبة.

$\Sigma R -$ مجموع رتب الإشارة السالبة.

N هو عدد أزواج الدرجات المتناظرة.

T القيمة الأصغر من $\Sigma R +$ و $\Sigma R -$

وقيمة معامل الارتباط هذه تراوح بين +1 و-1 وإن إشـــارته تعتمد على أي من المجموعتين نعتبرها المجموعة الأولى. كما يستخدم في تقدير قيمة معامل الارتباط الثنائي الحقيقي للمعادلة التالية:

$$r_{pb} = 1 - \frac{4\,T_+}{n(n+1)}$$

حيث T_+ مجموع الرتب ذات الإشارة الموجبة، n عدد أفراد المجموعة (عدد أزواج الدرجات).

مثال

الأفراد	الدرجات قبل	الدرجات بعد	التغيير	رتب التغيير	
A	20	38	+18	8	
B	22	37	+15	7	
C	19	33	+14	6	
D	20	29	+9	5	
E	22	14	-8		4
F	18	12	-6		3
G	24	20	-4		2
H	20	22	+2	1	
			Σ	27	9

$$r = \left(\frac{27}{30}\right) - \left(\frac{9}{30}\right) = 0.75 - 0.25 = 0.50$$

$$r = 4\left|9 - \frac{36}{2}\right| / [8(9)]$$

$$= 4|-9|/72 = 36/72 = 0.50$$

تقدير قيمة حجم الأثر عند استخدام اختبار كروسكال ووواليس لعدة مجموعات مستقلة (في Steyn, 2020):

يمكن تقدير قيمتي مربع إيبسلون ومربع إيتا، من خلال المعادلتين:

$$\eta^2 = \frac{H-K+1}{N-K} \qquad \varepsilon^2 = \frac{H(n+1)}{(N^2-1)}$$

حيث K عدد المجموعات، H القيمة المحسوبة من اختبار كروسكال ووواليس، N عدد الأفراد الكلي بالمجموعات.

تقدير قيمة حجم الأثر عند استخدام اختبار فريدمان لعدة مجموعات مرتبطة (في Steyn, 2020):

وفيه نستخدم معامل كندال تاو W، وهو يتحدد في هذه الحالة بالمعادلة:

$$W = \frac{F_r}{N(k-1)} = \frac{\chi^2}{N(k-1)}$$

حيث n عدد أفراد المجموعة، وk عدد المجموعات، F_r هي القيمة المحسوبة في اختبار فريدمان، التي يعبر عنها بقيمة χ^2 بدرجة حرية (K-1).

تحويل قيم الإحصاءات في الاختبارات اللابارامترية إلى قيم Z و r (في Steyn, 2020):

يمكن استخدام قيمة Z لتقدير قيمة حجم الأثر r في اختبار مان-ويتني باستخدام المعادلات التالية:

$$\mu_u = \frac{n_1 n_2}{2}$$

$$\sigma_u = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}$$

$$Z = \frac{U - \mu_u}{\sigma_u} =$$

$$\therefore r = \frac{Z}{\sqrt{n}}$$

ولحساب Z و r باختبار ويلكوكسن لاختبار إشارة الرتب، للأعداد الكبيرة نستخدم المعادلات:

$$\eta_T = N(N+1)/4$$

$$\sigma_u = \sqrt{N(N+1)(2N+1)/4}$$

$$Z_T = \frac{T - \eta_T}{\sigma_u}$$

ولحساب Z لاختبار الإشارة أو اختبار ذي الحدين للأعداد الكبيرة نستخدم المعادلة:

$$Z = \frac{X - Np}{\sqrt{Npq}}$$

متوسط ذي الحدين = Np وتباين ذي الحدين = Npq

X= العدد الملاحظ من الأحداث p

وعند درجات الحرية df=1 فإن قيمة Z^2 تتبع توزيع χ^2 وفقاً للمعادلة:

$$\chi^2 = Z^2 = \frac{(X - Np)^2}{Npq}$$

المبحث الثالث

طرق تقدير قيمة متوسط وتباين حجم الأثر
في إطار نموذجي الآثار الثابتة والعشوائية

لتقدير قيمة حجم الأثر المركب أو متوسط تقديرات حجم الأثر، يجب أن نحدد أولاً درجة تباين أو تجانس تلك الحجوم، حتى يمكن تحديد النموذج المستخدم في دمج أو تجميع حجوم الأثر؛ أهو نموذج حجم الأثر الثابت أم نموذج حجوم الأثر العشوائية. ويرى هايدو-مدينا وآخرون (Huedo-Medina, et al., 2006: 7-10) أن هناك مصدرين للتباين بين الدراسات:

1- التباين الناتج عن خطأ المعاينة (Q_W)، ويسمى التباين داخل الدراسة، الذي ينتج بسبب اختلاف العينات داخل الدراسة الأولية.
2- التباين بين الدراسات (Q_B)، وهو يشير إلى التباين الحقيقي بين الدراسات المتضمنة في التحليل.

تحديد تباين أو تجانس تقديرات حجم الأثر

عندما نتحدث عن التشتت أو التباين في تقديرات حجم الأثر بالدراسات المتعددة في موضوع معين، فإننا يجب أن نركز على تباين تقديرات حجم الأثر الحقيقية، علماً بأن التباين الملاحظ يتضمن هذا التباين الحقيقي، بالإضافة إلى الخطأ العشوائي أو خطأ المعاينة. يرى هنتر وآخرون (Hunter et al., 1986) أن تقدير قيمة حجم الأثر العام أو المتوسط يتطلب أن تكون عينة حجوم الأثر متجانسة باعتبارها من مجتمع واحد. والاختبار الإحصائي لمدى تجانس تقديرات حجم الأثر هو اختبار الفرض الصفري الذي ينص على أنه لا توجد فروق ذات دلالة إحصائية بين تقديرات حجم الأثر ($d_1=d_2=d_3=\ldots\ldots d_k=d$) والفرض البديل هو أن أحد تقديرات حجم الأثر يختلف عن باقي تقديرات حجم الأثر. ويعتبر تقييم التجانس بين الدراسات المتضمنة في دراسات ما وراء التحليل قضية أساسية تؤثر على إمكانية التعميم لنتائج ما وراء التحليل، ويستخدم مقياس درجة عدم الاتساق أو التباين بين نتائج الدراسات المتضمنة فيما وراء التحليل لتقييم التجانس (T^2, I^2, Q).

إحصاء كوكران Cochran's Q (Aert, et al., 2019)

الإحصاء المستخدم في تحديد تباين حجوم الأثر هو إحصاء كوكران Cochran`s Q، وهو يقيس انحراف تقديرات حجم الأثر الملاحظ عن متوسط حجم الأثر المشترك، وهو يتحدد بالمعادلة التالية:

مؤشر كوكران للتجانس الكلي:

$$Q = \sum_{i=1}^{K} w_i(\bar{y}_i - \bar{y})^2$$

$$Q = \sum w_i.\bar{y}_i^2 - [(\sum w_i.\bar{y}_i)^2 / \sum w_i]$$

حيث $\bar{y}_i$ = حجم الأثر بالدراسة (i)، $\bar{y}$= متوسط تقديرات حجم أثر الدراسات التي عددها K، w_i= وزن الدراسة (i)

ويتحدد وزن الدراسة w_i باستخدام إحدى المعادلات التالية:

أ- وزن الدراسة $w_i = \frac{1}{n_i}$ حيثn_iعدد أفراد الدراسة.

ب- وزن الدراسة $w_i = \frac{1}{V_i}$ حيثV_iهي التباين داخل الدراسة (i).

ويستخدم ما سبق في حالة نموذج الأثر الثابت.

$S^2_{e\bar{y}}$ =مربع الخطأ المعياري لمتوسط حجم الأثر، $w_{yi} = \frac{1}{S^2_{ey_i}}$

وفي حالة نموذج الآثار العشوائية فإن وزن الدراسة يساوي معكوس التباين داخل الدراسةV_i+التباين بين الدراسات T_i

$w_i = {}^{1}/_{(T_i + V_i)}$ حيثT_iيمثل التباين بين الدراسات.

خصائص إحصاء كوكران Q (Aert, et al., 2019)

1- أن توزيع Q يتبع توزيع x^2 عند درجات حرية (K-1)، حيث K هي عدد الدراسات أو عدد المتغيرات موضـــوع التحليل أو الانحدار، وعندما تكون x^2 غير دالة فإن ذلك يعد دليلاً على تجانس نتائج الدراسات.

2- إن Q_T مقياس معياري لا يتأثر بوحدة قياس حجم الأثر.

3- التباين الكلي لـــ Q_T يمكن تقسيمه إلى جزأين:

أ- التباين داخل الدراسات أو Qw أو تباين خطأ العينة.

ب- التباين بين الدراسات Q_B.

4- التباين الملاحظ يعكس أخطاء المعاينة في الدراسات. عندما تكون Q غير دالة، فإن ذلك يعني قبول الفرض الصفري الذي ينص على أن تقديرات حجم الأثر متجانسة، أما إذا كانت Q دالة؛ فإن ذلك يعني أن تقديرات حجم الأثر متباينة.

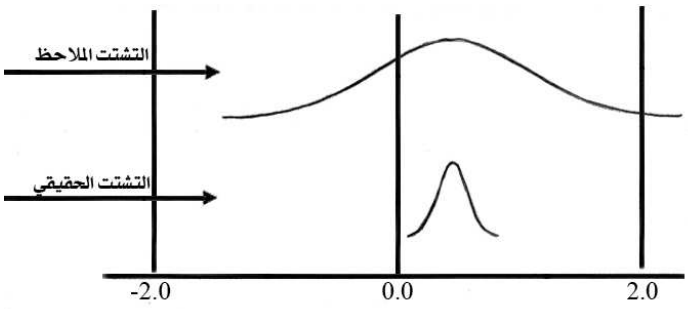
5- وفي حالة عدم تحقق التجانس، يجب إعادة تجميع الدراسات عن طريق تصنيف Cluster Method أو استخدام المتغيرات المعدلة moderator variables لتقسيم الدراسات؛ مثل: العمر، المرحلة، الحالة الاقتصادية، وبعد ذلك يُعاد التحليل لكل مجموعة من المجموعات الجديدة (Schwarzer, 1989).

6- Q هو مقياس إحصائي يعزل التباين الحقيقي لحجوم الأثر عن تباينها الملاحظ.

ويبين الشكل (1/3) العلاقة بين التباين الملاحظ والتباين الحقيقي.

شكل (1/3)

العلاقة بين التباين الملاحظ والتباين الحقيقي



7- لا يصلح استخدام إحصاء Q لتقدير قيمة التباين الحقيقي.

8- نظراً لأن Q تتبع في توزيعها χ^2 بدرجات حرية (K-1) لذلك يمكن تحديد قيمة p المناظرة لـχ^2 ؛ ومن ثم قيمة α، فإذا كانت p أقل من 0.05 فإن ذلك يعني أن الدراسات لا تشترك في حجم أثر عام أو مشترك Common Effect Size.

9- يعتمد اختبار الدلالة لـــ Q على الدقة التي نقدر بها الأثر وخاصة عدد الدراسات التي يشملها تحليل مجموعة الدراسات الفردية.

10- إن دلالة قيمة p تعكس التباين في حجم الأثر الحقيقي، إلا أن عدم دلالة Q لا يؤخذ كدليل على أن حجم الأثر يكون ثابتاً؛ لأن نقص الدلالة يمكن أن يرجع إلى انخفاض القوة الإحصـــائية للاختبار، كذلك فإن صـغر عدد الدراسـات وارتفاع التباين داخل الدراسـة قد يؤدي إلى نقص الدلالة p.

11- إن إحصاء Q و p تعالج فقط اختبار الدلالة ولا تستخدم كبديل لكمية التباين الحقيقي.

12- إحصاء Q لا تعطي تقديراً كمياً للتباين الحقيقي بين تقديرات حجم الأثر الحقيقية.

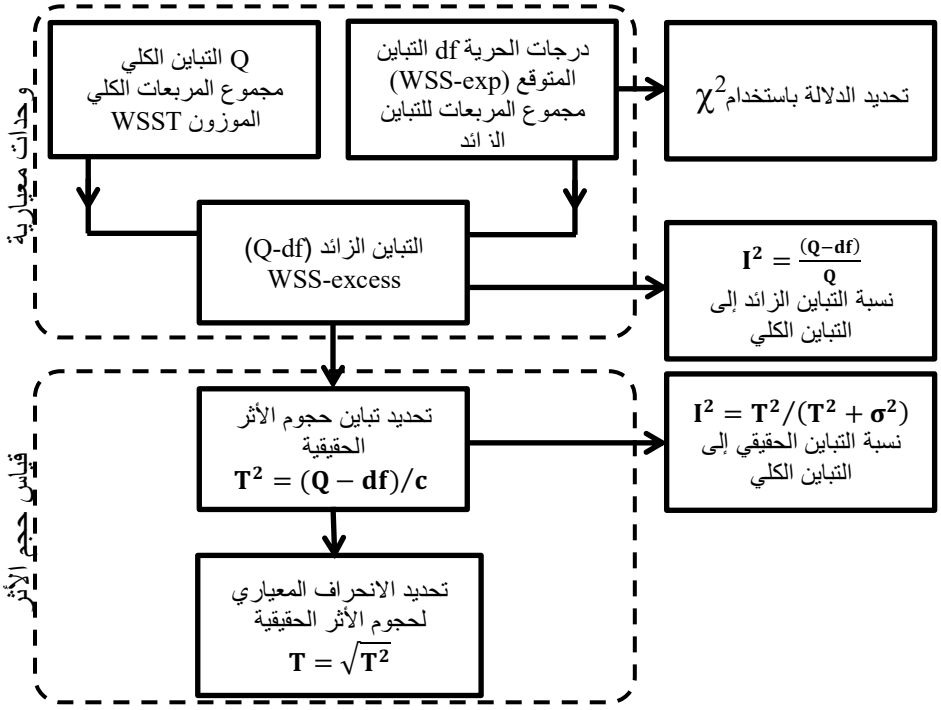
13- الاختبار Q ليس قوياً في حالة صـغر عدد الدراسـات التي يشـملها التحليل وكبر التباين داخل الدراسات.

14- يفضل استخدام نموذج الآثار العشوائية في حالة دلالة Q.

15- لمعرفة ما إذا كانت الدراسـات الموجودة تتبع لنفس المجتمع، يجب حسـاب إحصـائي التجانس Q.

شكل (2/3)

كيفية اشتقاق T، Q و I^2 و T^2 من بعضها البعض (المصدر Sutton, 2009 بتصرف)



إحصاء I^2

يمكن استخدام إحصاء Q في بناء إحصاء I^2 و T^2 لتحديد نسبة التباين الملاحظة في تقديرات حجم الأثر الذي يفسره التباين في تقديرات حجم الأثر الحقيقية. ويوضح شكل (1/3) العلاقة بين Q وكل من I^2 و T^2. وقد استخدم هيجنس وآخرون (Higgins et al., 2003) إحصاء I^2 لتحديد نسبة التباين في حجوم الأثر المفسرة وليس أخطاء المعاينة إلى التباين في حجوم الأثر الكلية، وهو يتحدد بالمعادلة التالية:

$$I^2 = \frac{(Q - df)}{Q} \times 100$$

وتشير المعادلة السابقة إلى أن النسبة المئوية للتباين الزائد $(Q - df)$ منسوبة إلى التباين الكلي، كما تشير أيضاً إلى النسبة المئوية في التباين الملاحظ في تقديرات حجم الآثار التي تعكس الفروق في الدرجات الحقيقية. ويمكن أيضاً تفسيرها على أنها النسبة في التباين الملاحظ التي ترجع التباين الحقيقي إلى تقديرات حجم الأثر.

$$I^2 = \frac{V_{B.}}{V_{T.}} \times 100$$

وتمثل المعادلة التالية التباين بين الدراسات مقسوماً على التباين الكلي:

$$I^2 = \frac{T^2}{T^2 + V_i} \times 100$$

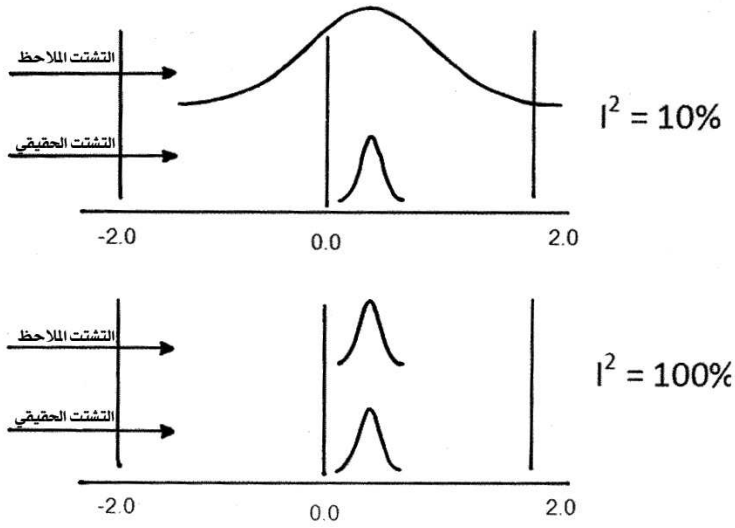
والآلية المستخدمة لعزل التباين الحقيقي تتم من خلال الفرق بين التباين الملاحظ Q والتباين المتوقع، وهو df (ويسمى هذا الفرق بالتباين الزائد excess) منسوباً إلى التباين الكلي الموزون Q ومضروباً في 100.

خصائص I^2 لتقدير قيمة تباين تقديرات حجم الأثر (Sutton, 2009)

يمتاز مؤشر I^2 بأنه يسلط الضوء على أي تأثير لعدم التجانس Heterogeneity، وسهولة تقدير قيمته وتفسيره بغض النظر عن نوع البيانات المستخدمة في التحليل الأولي، وهو لا يعتمد على عدد الدراسات المتضمنة فيما وراء التحليل (Higgins et al, 2003). وشكل (3/3) يبين العلاقة بين التشتت الملاحظ والتشتت الحقيقي عندما I^2=10% وعند I^2=100%.

شكل (3/3)

العلاقة بين التشتت الملاحظ والتشتت الحقيقي عندما $I^2=10\%$ وعند $I^2=100\%$ (المصدر: المؤلفان)



ويمكن تحديد الخصائص التالية لمؤشر I^2 (Sutton, 2009):

1- تحدد القيمة (Q- df) مدى انخفاض أو ارتفاع التباين.

2- تجيب قيمةI^2على السؤال: ما نسبة التباين الملاحظ التي تعكس الفروق الحقيقية؟

3- القيمة المتوقعة لـ Q تساوي df، وذلك في حالة افتراض أن كل الدراسات تشارك في حجم الأثر المشترك أو العام.

4- (Q-df) هو جزء من التباين ينشأ نتيجة الفروق في الآثار الحقيقية بين الدراسات وليس المصادفة.

5- قيمة I^2 تراوح بين 0%و 100%، كما يتضح من شكل (3/3)، وهي مقسمة إلى المستويات التالية:

أ) تشير القيمة 0% إلى أنه لا يوجد تباين ملاحظ وأن التباين المحسوب يرجع إلى خطأ المعاينة.

ب) مستوى التباين يكون منخفضاً عندما تكون I^2 أقل من 25% ويناسب النموذجين الثابت والعشوائي، وهذه النسبة تقود إلى خطأ المعاينة.

ج) مستوى التباين يكون متوسطاً عندما تراوح قيمة I^2 بين 25% و 50%.

د) مستوى التباين يكون مرتفعاً عندما تراوح قيمة I^2 بين 50% و 75%.

هـ) مستوى التباين يكون مرتفعاً جداً عندما تكون قيمة I^2 أكبر من 75%، وعندما يصل إلى 100% فإن جميع التباين المحسوب هو تباين حقيقي بين الدراسات ولا وجود لتباين بسبب الخطأ (Huedo-Medina, et al., 2006)

6- I^2 تشبه مفهوم معامل الثبات ومربع معامل الارتباط المتعدد باعتبارها النسبة من التباين الكلي التي تفسر بمتغيرات نسبية معينة.

7- I^2 ليس مقياساً مطلقاً للتباين أو عدم التجانس.

8- عندما تكون I^2 أكبر من 5%2 يفضل استخدام نموذج الآثار العشوائية أما إذا كانت أقل من 25% فيمكن استخدام نموذج الآثار الثابتة.

9- تهدف قيمة Q إلى تحديد مدى دلالة التباين بين مجموعة الدراسات المتضمنة في التحليل بينما تعني قيمة I^2 النسبة المئوية من هذا التباين المفسرة بالمعالجة أو المتغير المستقل (لا يفسر نتيجة المصادفة).

استخدام τ^2 أو T^2 (تنطق مربع تاو) لتقدير قيمة التباين بين تقديرات حجم الأثر

تباين حجوم الأثر الحقيقية τ^2 وتقدير T^2 إلى بارامتر التباين الكلي لحجوم الأثر الحقيقية بالمجتمع (بارامتر المجتمع)، وحيث إننا نعمل دائماً مع عينات من المجتمع فإننا نقدره باستخدام طريقة العزوم لدير سيمونين ولايرد DerSimonian & Laird (1986)، وهي تتحدد بالمعادلة التالية:

في حالة تقديرات حجم الآثار الثابتة: $T^2 = (Q - df)/C$

حيث: $C = \Sigma w_i - \left[\frac{\Sigma(w_i^2)}{\Sigma w_i}\right]$، و$df=(k-1)$، وk هي عدد الدراسات أو عدد المتغيرات موضوع التحليل.

خصائص T^2 و T (DerSimonian & Laird, 1986)

1- T^2 له نفس وحدة قياس التباين داخل أو بين الدراسات.

2- T^2 هو مؤشر للتباين الحقيقي الذي لا يرجع إلى المصادفة.

3- إذا كان عدد الدراسات قليلاً فإن تقدير T^2 سوف يكون منخفض الدقة والقوة.

4- قيمة T^2 و T أكبر من أو تساوي الصفر.

5- تعتمد قيم T^2 و T على درجة التشتت أو التباين الزائد (Q-df)، وهذا الأخير يعكس الفروق الحقيقية بين الدراسات.

6- تعطينا قيمة T^2 التباين بين حجوم الأثر الحقيقية بالدراسات المتضمنة بالتحليل، وهذا يعني أن T^2 تعكس درجة التباين الحقيقي، أما T^2؛ فهي تباين حجوم الأثر الحقيقية بالمجتمع.

7- اتساع فترة الثقة لـ T^2 و I^2 يدل على تباين أكبر.

8- يتضمن التباين الملاحظ كلاً من التباين الحقيقي والخطأ العشوائي (خطأ المعاينة).

مثال: جدول (1/3) يبين حجم الأثر والتباين والوزن لـ(6) دراسات (المصدر Hunter, & Schmidt, 2004)

جدول (1/3)
حجم الأثر والتباين والوزن لـ(6) دراسات

الدراسات	حجم الأثر y_i	التباين V_i	الوزن W_i	W_i^2	$W_i y_i$	$W_i y_i^2$	متوسط حجم الأثر	$(y_i - \bar{y})^2$	$W_i(y_i - \bar{y})^2$
1	0.095	0.033	30.352	921.21	2.869	0.271	0.4134	0.102	3.103
2	0.277	0.031	32.568	1060.68	9.033	2.505	0.4134	0.019	0.610
3	0.367	0.050	20.048	401.93	7.349	2.694	0.4134	0.002	0.046
4	0.664	0.011	95.111	9046.01	63.190	41.983	0.4134	0.063	5.950
5	0.462	0.043	23.439	549.37	10.824	4.999	0.4134	0.002	0.053
6	0.185	0.023	42.698	1823.12	7.906	1.464	0.4134	0.052	2.241
المجموع			244.215	13802.33	101.171	53.915			12.003

من الجدول يمكن حساب متوسط حجم الأثر $\bar{y}$ وإحصاء Q كما يلي:

$$\bar{y} = (101.171)/(244.215) = 0.4134$$

$$Q = 53.915 - \frac{(101.171)^2}{244.215} = 12.003$$
$$df = 6 - 1 = 5$$
$$Q - K = 12.003 - 5 = 7.003$$

وحيث إن p لقيمة Q=12.003 ودرجة حرية 5 تساوي 0.035، فإن ذلك يعني أن Q دالة ومن ثم نرفض الفرض الصفري القائل إن كل حجوم الأثر تشترك في حجم أثر مشترك؛ ومن ثم نقبل الفرض البديل الذي يعني أن حجوم الأثر الحقيقية تختلف بين الدراسات وبعضها لذلك نستخدم الصيغ أو المعادلات التالية:

$$C = 244.215 - \left(\frac{13802.33}{244.215}\right) = 187.6978,$$

تباين حجوم الأثر الحقيقية

$$T^2 = \frac{12.003 - 5}{187.698} = 0.0373,$$

$$T = \sqrt{0.0373} = 0.1932$$

الانحراف المعياري لحجوم الأثر

$$I^2 = \left(\frac{12.003 - 5}{12.003}\right) \times 100\% = 58.34\%$$

النسبة المئوي لتفسير التباين

$$V_{T^2} = 2 \times \left(\frac{31.020}{187.698^2}\right) = 0.0018$$

تباين T^2

$$SE_{T^2} = \sqrt{0.0018} = 0.0420$$

الخطأ المعياري لـ T^2

نموذجا الآثار الثابتة والآثار العشوائية

لتقدير قيمة حجم الأثر المركب (المتوسط، العام، الملخص، الموزون) فإن ذلك يتطلب أن تحدد أولاً تقديرات حجم الأثر أهي متساوية أم عشوائية متباينة، وبمعنى آخر تحديد نموذج تقديرات حجم الأثر المستخدم في هذه الحالة أهو نموذج حجم الأثر الثابت أم هو نموذج تقديرات حجم الأثر العشوائية.

نموذج الآثار الثابتة Fixed Effect Model (FEM)

نستخدم نموذج الأثر الثابت إذا توافر الشرطان التاليان: الأول أن نعتقد أن كل الدراسات المتضمنة في التحليل تكون وظيفياً متماثلة، والثاني أن يكون هدفنا تقدير قيمة حجم الأثر المشترك لمجتمع مماثل، ولا نعمم على مجتمعات أخرى. نموذج الأثر الثابت يفترض أن كل الدراسات المتضمنة بما وراء التحليل، سحبت من مجتمع مشترك، وبطريقة أخرى، كل العوامل التي يمكن أن تؤثر في حجم الأثر تكون هي نفسها بكل دراسات المجتمع، ولذلك فإن حجم الأثر يكون هو نفسه بكل دراسات المجتمع. أما

حجم الأثر الملاحظ فإنه يختلف من دراسة إلى أخرى، ويرجع ذلك إلى الخطأ العشوائي المتأصل بكل دراسة (Hunter, & Schmidt, 2004).

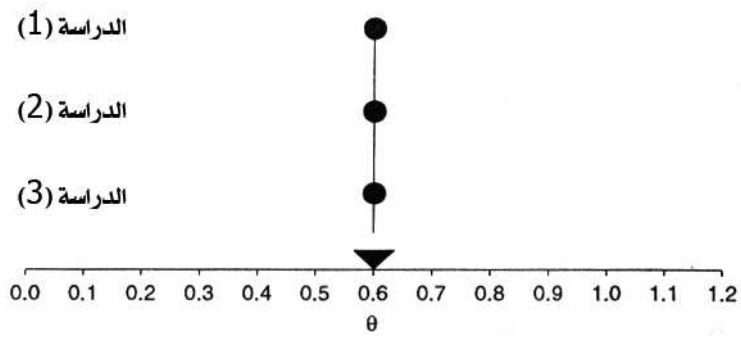
مثال ذلك افترض أن شركة دواء سوف تستخدم ألف مريض للمقارنة بين مخدر ومهدئ، وأن العينة التي يمكن التعامل معها كتجربة هي 100 مريض في الوقت نفسه. ولذلك فإن الشركة في حاجة إلى تكرار ذلك الإجراء عشر مرات. وكانت تلك الدراسات العشر متماثلة في تطبيق المتغيرات ومخرجاتها. وبصفة خاصة في سحب المرضى عينة الدراسة من مجتمع واحد، وكذلك استخدام الباحثين أنفسهم. كل الدراسات تتوقع المشاركة في أثر مشترك، وبذلك يتحقق الشرط الأول، والهدف من التحليل هو معرفة إذا ما كان المخدر يعمل في المجتمع الذي سحبت منه عينات المرضى (لا يقدر استدلالياً على مجتمع آخر) وهذا يوفر الشرط الثاني.

افتراضات نموذج الآثار الثابتة (Borenstein, et al., 2009)

1- أن الدراسات المتضمنة بالتحليل سحبت من مجتمع مشترك أو واحد يقيس المعلم نفسه (μ). ويسمى حجم الأثر الحقيقي أو الأثر المشترك العام Common.
2- أن العوامل المؤثرة في حجم الأثر هي نفسها المؤثرة في كل دراسات المجتمع الواحد،
3- الآثار الملاحظة يمكن أن تكون كلها متماثلة وإذا اختلفت من دراسة إلى أخرى فإن ذلك يكون بسبب الخطأ العشوائي بكل دراسة (خطأ المعاينة).
4- الأثر المركب يمكن أن يكون هو نفسه في كل الدراسات الفرعية.
5- اتساع فترة الثقة للأثر المركب أو المشترك سوف يقترب من الصفر، ويظهر ذلك في شكل (3/3) حيث تكون فترة الثقة مساوية للصفر.

شكل (3/3)

نموذج الأثر الثابت والآثار الحقيقية (بتصرف من Borenstein, et al., 2009)



6- أوزان الدراسات تحدد بنسبة مئوية (حصة) هي معكوس التباين داخل الدراسة $(1/V_i)$ أو مربع الخطأ العشوائي داخل الدراسة $(SE^2{}_{y_i})$.

7- تشارك كل الدراسات المتضمنة بالتحليل في حجم الأثر المشترك.

مما سبق نستنتج ما يلي:

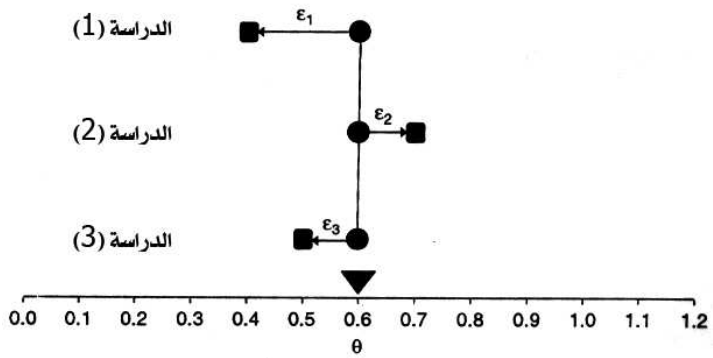
الأثر الملاحظ بالدراسة i = الأثر الحقيقي بالدراسة i + الخطأ العشوائي (خطأ المعاينة) بالدراسة i

$$y_i = \theta_i + e_i$$

يلاحظ من شكل (3/3) أن حجم الأثر المشترك العام Overall يساوي 0.6 ويمثل بمثلث، والأثر الحقيقي (θ) بكل دراسة يمثل بدائرة. حجم الأثر الحقيقي بكل دراسة يساوي 0.6 ويمثل بدائرة، وهذه الدوائر تقع فوق المثلث مباشرة. في الشكل (4/3) يمثل حجم الأثر الملاحظ $(\bar{y})$ بمربع. ونظراً لأن معظم الدراسات تستخدم عينات ليست كبيرة جداً، فإن هناك خطأ معاينة (ε_i)؛ ولذلك يظل الأثر الحقيقي بكل دراسة كما هو، ولكن الأثر الملاحظ (المربع) يختلف من دراسة إلى أخرى.

شكل (4/3)

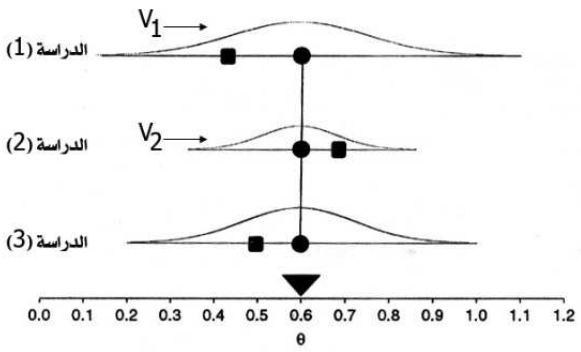
نموذج الأثر الثابت والآثار الملاحظة (بتصرف من Borenstein, et al., 2009)



فكما يتضح من شكل (4/3) في الدراسة (1) إذا كان خطأ المعاينة (ε_1) يساوي (0.20-) فإن الأثر الملاحظ (y_1) يساوي: 0.4=0.2-0.6 ، وفي الدراسة الثانية إذا كان خطأ المعاينة (ε_2) مساوياً (0.10) فإن (y_2) يساوي: 0.70=0.10+0.60 ، وفي الدراسة الثالثة إذا كان خطأ المعاينة (ε_3) مساوياً (0.10-) فإن الأثر الملاحظ (y_1) يساوي: 0.5=0.1-0.6. وبصفة عامة فإن المعادلة العامة لما سبق هي: $y_i = \theta + \varepsilon_i$

شكل (5/3)

الخطأ المعياري لتقديرات حجم الأثر العشوائية (بتصرف من Borenstein, et al., 2009)



وحيث إن الخطأ في أي دراسة يكون عشوائياً، فإنه يمكن تقدير خطأ توزيع المعاينة (خطأ المعاينة) كما في شكل (5/3)، الذي يوجد فيه منحنى اعتدالي لكل حجم أثر حقيقي بكل دراسة، ويعتمد اتساع هذا المنحنى على التباين في تلك الدراسة، الذي يعتمد بدوره على حجم عينة الدراسة [في الدراسة (1) العينة صغيرة وفي الدراسة (2) العينة كبيرة] ويُحسب الخطأ المعياري في هذه الحالة بالجذر التربيعي للتباين.

وفي ما وراء التحليل، عادة ما يقيم الوزن بكل دراسة بمعكوس تباين تلك الدراسة $W_i = \frac{1}{V_i}$ حيث V_i هو التباين داخل الدراسة i، وحجم الأثر المشترك بنموذج الأثر الثابت يساوي $\bar{y} = (\sum W_i Y_i)/\sum W_i$. ولاختبار الفرض الصفري فأن الأثر الحقيقي المشترك θ يساوي صفراً نستخدم المعادلة $z = \frac{\bar{y}}{\sqrt{V_i}}$. إذا كانت دقة الدراسات المتضمنة فيما وراء التحليل متساوية، كان بإمكاننا ببساطة تقدير قيمة متوسط حجوم الأثر، ولكن إذا كان بعض الدراسات أكثر دقة من البعض الآخر فإنه يجب أن نعطي وزناً أكبر للدراسات المتوافرة فيها مزيد من المعلومات، وكذلك تقدير قيمة المتوسط الوزني وذلك بإعطاء وزن أكبر لبعض الدراسات ووزن أقل للبعض الآخر.

يستخدم نموذج الأثر الثابت إذا كان هناك سبب للاعتقاد بأن جميع الدراسات متطابقة وظيفياً، وإذا كان هدفنا تقدير قيمة حجم الأثر المشترك، الذي يمكن تعميمه على أمثلة أخرى من المجتمع نفسه. مثال ذلك إجراء شركة للأدوية خمس دراسات لتقييم فاعلية دواء ما، وقد تعاملت الدراسات الخمس مع الباحثين أنفسهم وبالطريقة نفسها مع المرضى، لذلك كان المتوقع أن يحدث أثر متساو أو متطابق في الدراسات الخمس (كما لو كانت دراسة واحدة كبيرة). وينشأ نتيجة ما سبق أن التحليل باستخدام هذا النموذج يكون على مدى ضيق بالمجتمع (Borenstein, et al., 2009).

مثال لتقدير قيمة $\bar{y}_i$، W_i، V_i، $\bar{y}$، Q باستخدام جدول (2/3)، الذي يعرض بيانات حجم الأثر باستخدام نسبة الأرجحية في (6) دراسات في إطار نموذج الآثار الثابتة (في Cumming, 2012 بتصرف):

جدول (2/3)

بيانات حجم الأثر الأولية لست دراسات وتحليلها

الدراسات	حجم الأثر y_i	التباين V_i	الوزن W_i	$W_i y_i$	$W_i y_i^2$	W_i^2	W_i^3
1	-0.366	0.185	5.402	-1.978	0.724	29.18	157.66
2	-0.288	0.290	3.453	-0.993	0.286	11.92	41.18
3	-0.384	0.156	6.427	-2.469	0.948	41.30	265.42
4	-1.322	0.058	17.155	-22.67	29.971	294.30	5048.71
5	-0.417	0.282	3.551	-1.480	0.617	12.61	44.76
6	-0.159	0.160	6.260	-0.998	0.159	39.19	245.33
المجموع			42.248	-30.59	32.705	428.50	5803.06

يعرض جدول (3/3) قيم الأوزان وحجم الأثر المتوسط وتباين كل دراسة في إطار نموذج الآثار الثابتة.

جدول (3/3)

قيم w_i، $\bar{y}$، V للدراسات الست

الدراسات	حجم الأثر y_i	التباين V_i	الوزن W_i	المتوسط $\bar{y}$	$(y_i - \bar{y})^2$	$W(y_i - \bar{y})^2$
1	-0.366	0.185	5.402	-0.724	0.128	0.692
2	-0.288	0.290	3.453	-0.724	0.191	0.658
3	-0.384	0.156	6.427	-0.724	0.116	0.743
4	-1.322	0.058	17.155	-0.724	0.357	6.127
5	-0.417	0.282	3.551	-0.724	0.094	0.335
6	-0.159	0.160	6.260	-0.724	0.319	1.996
المجموع						10.551

متوسط تقديرات حجم الأثر OR

$$\bar{y}_{OR} = \frac{-30.594}{42.248} = -0.724$$

تباين تقديرات حجم الأثر OR

$$Q = \sum_{i=1}^{K=1} W_i (y_i - \bar{y})^2 = 10.551$$

أو تباين تقديرات حجم الأثر OR

$$Q = 32.705 - \frac{(-30.594)^2}{42.248} = 10.551$$

وفي إطار افتراض أن كل الدراسات تشارك في الأثر المشترك العام، فإن القيمة المتوقعة لـ Q تتحدد بالمعادلة: $df = 6 - 1 = 5$، ويكون التباين الزائد 5.5512=5-10.5512= (Q – df). وهذه الزيادة ترجع إلى الفروق في تقديرات حجم الأثر وأن قيمة p المقترنة بـ Q أو $\chi 2$ تساوي 0.061 عند درجات حرية 5 وهذا يعني قبول الفرض الصفري، مما يعني وجود تجانس بين حجوم الأثر.

قيمة C
$$C = 42.248 - \left(\frac{428.50}{42.248}\right) = 32.1052$$

تباين حجوم الأثر
$$T^2 = \frac{10.5512 - 5}{32.1052} = 0.1729$$

الانحراف المعياري لحجوم الأثر
$$T = \sqrt{0.1729} = 0.4158$$

النسبة المئوية لـ I^2
$$I^2 = \left(\frac{10.5512 - 5}{10.5512}\right) \times 100 = 52.61\%$$

ولتقدير قيمة الخطأ المعياري لـ T^2 فإننا في حاجة من الجدولين السابقين إلى SW_1 و SW_2 و SW_3 واستخدام الصيغ التالية:

$$A = \left[df + 2\left(\sum W_i - \frac{\sum W_i^2}{\sum W_i}\right)T^2 + \left(\sum W_i^2 - 2\left(\frac{\sum W_i^3}{\sum W_i}\right) + \left(\frac{(\sum W_i^2)^2}{(\sum W_i)^2}\right)\right)T^2\right]$$

$$A = \left[5 + 2\left(42.25 - \frac{428.5}{42.25}\right)0.1729 + \left(428.5 - 2\left(\frac{5803.1}{42.25}\right) + \left(\frac{(428.5)^2}{(42.25)^2}\right)\right)0.1729^2\right]$$

$$= 23.7754$$

تباين T^2
$$V_{T^2} = 2 \times \left(\frac{23.7754}{32.1052^2}\right) = 0.0461$$

الخطأ المعياري لـ T^2
$$SE_{T^2} = \sqrt{0.0461} = 0.2148$$

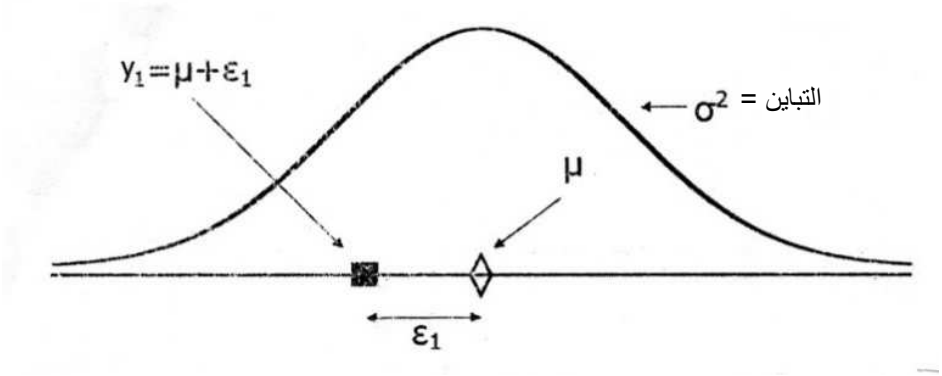
دلالة الخطأ المعياري
$$Z = 0.4158/0.2148 = 1.936$$

نموذج الآثار العشوائية Random Effect Model (REM)

على العكس من نموذج الآثار الثابتة، عندما يجمع الباحثون بيانات من عدة دراسات أجريت بوساطة عدد متنوع من الباحثين في مجتمعات مختلفة، في هذه الحالة لا تكون هذه الدراسات متعادلة وظيفياً؛ لأنها تختلف في طرق من شأنها أن يكون لها تأثير على النتائج، ولذلك لا يجب أن نفترض حجم أثر مشترك عام. بالإضافة إلى أن التحليل باستخدام هذا النموذج يهدف عادة إلى التعميم على مدى أوسع بالمجتمع. إن ما سبق يتطلب نموذج الآثار العشوائية.

شكل (6/3)

الأثر الحقيقي والملاحظ (في Borenstein, et al., 2009)



يتضح من شكل (6/3) أن حجم الأثر المشترك μ وتقديرات حجم الأثر الملاحظة تتوزع حول μ بتباين قدره σ^2، الذي يعتمد على حجم العينة. وتقديرات حجم الأثر الملاحظة هي عينة من توزيع حجم الأثر الحقيقي ذي المتوسط μ. وتتحدد قيمة الأثر الملاحظ Y وهو يساوي $(\mu+\varepsilon_i)$ حيث ε_i هو الخطأ داخل الدراسة.

بفرض أن الأثر الملاحظ Y_1 هو عينة من توزيع ذو حجم أثر حقيقي مقداره θ_1 وتباين σ^2 وحجم الأثر الحقيقي θ_i هو عينة من توزيع له متوسط μ وتباين T^2 (أنظر شكل 7/3) يمكن استنتاج أنه لأي دراسة (i) تنتمي لهذا المجتمع فإن:

$$\theta_i = \mu + \xi_i \qquad Y_i = \theta_i + \varepsilon_i \qquad \therefore\ Y_i = \mu + \xi_i + \varepsilon_i$$

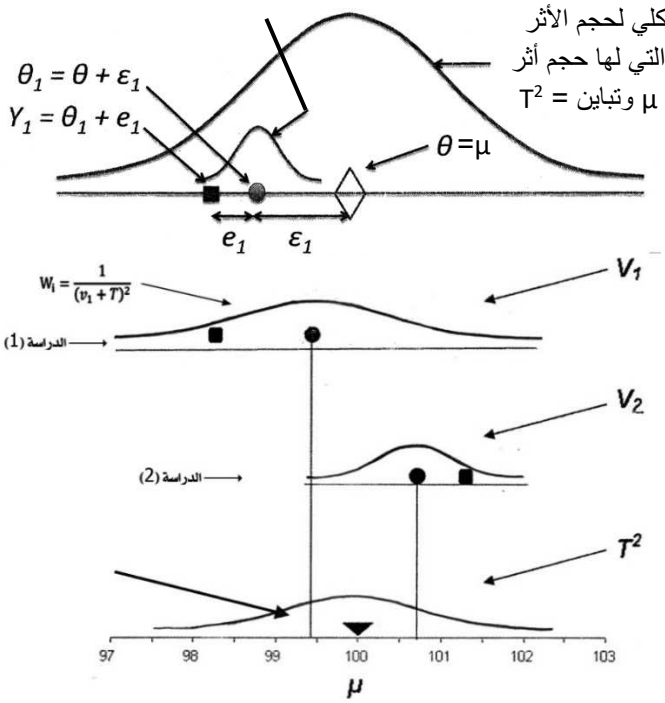
في نموذج الآثار العشوائية يفترض أن الدراسات محسوبة من مجتمعات يختلف بعضها عن بعض في الطرق التي تؤثر في أثر المعالجة، ويرجع هذا الاختلاف لسببين: الأول هو الخطأ العشوائي

خلال الدراسات كما في نموذج الأثر الثابت، والثاني هو التباين الحقيقي في حجم الأثر، من دراسة إلى أخرى (Cumming, 2012).

شكل (7/3) *التباين الداخلي وبين الدراسات (في Cumming, 2012 بتصرف)*

توزيع المجتمع الفرعي لحجم الأثر الحقيقي للدراسات التي تنتمي لها دراسة (1) بحجم أثر حقيقي بمتوسط Θ_1 وتباين = σ^2

توزيع المجتمع الكلي لحجم الأثر الحقيقي للدراسات التي لها حجم أثر حقيقي بمتوسط = μ وتباين = T^2



الافتراضات التي يتضمنها نموذج الآثار العشوائية (Borenstein, et al., 2009)

يفترض نموذج الآثار العشوائية عدداً من الافتراضات، من أهمها:

1- الدراسات المستخدمة في عملية التحليل هي عينة عشوائية من دراسات المجتمع الافتراضي وهي متكافئة وظيفيًا.

2- أن التباين بين تقديرات حجم الأثر يرجع إلى الخطأ العشوائي داخل كل دراسة (خطأ المعاينة) بالإضافة إلى التباين بين الدراسات أو بين تقديرات حجم الأثر وهو بؤرة التركيز في هذا النموذج.

3- في هذا النموذج يتم حساب وزن الدراسة باستخدام معكوس التباين المكون من التباين داخل الدراسات v_i (مثل نموذج الأثر الثابت) والتباين بين الدراسات T_i^2

$$W_i^* = 1/(V_i + T_i^2)$$

$$V_i^* = (V_i + T_i^2)$$

$$T^2 = [Q - (K-1)]/C$$

$$C = \sum w_i^* - [\sum (w_i^*)^2 / \sum w_i^*]$$

4- حجم الأثر الملاحظ (Y_i) =حجم الأثر الحقيقي (θ) + الخطأ (داخل الدراسة e_i وبين الدراسات ϵ_i).

5- أن تقديرات حجم الأثر الملاحظة والحقيقية تتوزع وفقًا للمنحنى الاعتدالي؛ فالدراسات المتضمنة بالتحليل البعدي حول موضوع معين تختلف في أحجام عيناتها وأعمار المشاركين بكل دراسة واختلاف تصميم تلك الدراسات وشدة وكثافة وسائل وعوامل التدخل أو المعالجة، وكذلك أدوات الدراسة.

6- متوسط تقديرات حجم الأثر في نموذج الآثار العشوائية يتحدد بالمعادلة

$$\overline{Y} = \left(\sum_{i=1}^{k} W^*.Y_i\right) / \sum_{i=1}^{k} W_i^*$$

الخطأ المعياري في إطار نموذج الآثار العشوائية (Cumming, 2012)

في إطار نموذج الأثر الثابت يحسب الخطأ المعياري لمتوسط حجم للأثر باستخدام المعادلة:

$$SE_{\bar{y}} = \sqrt{\frac{\sigma^2}{n(k)}}$$

حيث σ الانحراف المعياري بالدراسة i، وn هو المتوسط التوافقي لتقديرات حجم العينات في الدراسات المختلفة، وK هو عدد الدراسات. ويعني ذلك أنه مع حجم عينة كبير وكاف داخل الدراسات، فإن الخطأ المعياري يؤول للصفر. أما في إطار نموذج الأثر العشوائي؛ فيحسب الخطأ المعياري لمتوسط حجم الأثر، باستخدام المعادلة التالية:

$$SE_{\bar{y}} = \sqrt{\frac{\sigma^2}{n(k)} + \frac{T^2}{k}}$$

ونلاحظ أن الجزء الأول أسفل الجذر هو نفسه الموجود في المعادلة الأولى وهذا جزء يؤول إلى الصفر عندما يكون حجم الدراسة كبيراً وكافياً. أما الجزء الثاني؛ فإنه يؤول إلى الصفر نظرياً فقط عندما يكون عدد الدراسات كبيراً جداً أو غير محدود، وهو يعكس التباين بين الدراسات. ويلاحظ أن زيادة حجم العينة داخل الدراسات غير كاف لاختزال الخطأ المعياري؛ إذا وجد عدد قليل من الدراسات. لذلك فإن الخطأ المعياري في نموذج الأثر العشوائي قد يظل كبيراً، حتى إذا كان حجم العينات في الدراسات (n) كبيراً جداً.

مما سبق يتبين أنه لمزيد من الدقة في حساب الإحصاءات باستخدام نموذج الآثار العشوائية، فإننا في حاجة إلى عينات ضخمة في عددها حتى يؤول الخطأ داخل الدراسة إلى الصفر، بالإضافة إلى زيادة عدد الدراسات المتضمنة في التحليل حتى يؤول الخطأ المعياري في تقدير μ إلى الصفر، أما في حالة صغر عدد الدراسات؛ فإن هذين المصدرين للخطأ سوف ينزعان للارتفاع. وإذا كان تباين الآثار الحقيقية للدراسات تافهاً فإن نموذج الآثار العشوائية سوف يختزل إلى نموذج الآثار الثابت، ولذلك لا يوجد مانع من استخدام نموذج الآثار العشوائية في هذه الحالة، وإذا كانت قيمة Q دالة فإن استخدام نموذج الآثار العشوائية يكون له مسوغ إحصائي واضح.

ويضيف ديكوستر (Decoster, 2009: 37-39) أنه في أغلب الحالات يكون استخدام نموذج الأثر العشوائي ملائماً بشكل أكبر من استخدام نموذج الأثر الثابت. يستخدم بعض الباحثين نموذج الأثر الثابت لسهولة العمليات الحسابية المستخدمة فيه مقارنة بنموذج الأثر العشوائي؛ وبعد توافر البرمجيات الإحصائية لإجراء العمليات الحسابية بسهولة فقد توجه ما وراء التحليل لاستخدام نموذج الأثر العشوائي بشكل أكبر. يعرض جدول (4/3) مقارنة بين نموذج الأثر الثابت والعشوائي.

جدول(4/3)

مقارنة بين نموذج الأثر الثابت والعشوائي (المؤلفان)

نموذج الأثر الثابت	نموذج الآثار العشوائية
1- يفترض أن تقديرات حجم الأثر الحقيقية متساوية في كل الدراسات التي يشملها التحليل ولذلك يفترض التجانس بين تلك الآثار.	1- يفترض أن تقديرات حجم الآثار الحقيقية بالمجتمع هي عينة عشوائية تختلف من دراسة إلى أخرى ولذلك يفترض التباين أو عدم التجانس بين تلك الآثار.
2- عند تقدير قيمة الخطأ المعياري لحجم الأثر نستخدم التباين داخل الدراسة V_i أو خطأ المعاينة العشوائية بكل دراسة.	2- عند تقدير قيمة الخطأ المعياري لحجم الأثر نستخدم التباين بين الدراسات T^2 وداخل الدراسة V_i.
3- الخطأ المعياري لحجم الأثر المشترك يكون أقل واقعية ومعقولية من نموذج الآثار العشوائية.	3- الخطأ المعياري لمتوسط تقديرات حجم الأثر يكون أكثر واقعية ومعقولية من نموذج الأثر الثابت.
4- قرار عدم التباين هو نتيجة معرفة وخبرة وليس نتيجة اختبار إحصائي.	4- قرار التباين هو نتيجة اختبارات إحصائية.
5- اختبار دلالة حجم الأثر المشترك أقل تحفظاً (زيادة إيجابية زائفة).	5- اختبار دلالة الأثر المركب أكثر تحفظاً.
6- فترة الثقة لحجم الأثر المشترك تكون أقل اتساعاً.	6- فترة الثقة لمتوسط حجم الأثر تكون أكثر اتساعاً.
7- قدرة أقل على تعميم النتائج.	7- قدرة أكبر على تعميم النتائج.
8- وزن الدراسة $= \frac{1}{V}$	8- وزن الدراسة $= \frac{1}{(V+T^2)}$
9- الأثر الملاحظ= الأثر المشترك الحقيقي + الخطأ (داخل الدراسات).	9- الأثر الملاحظ= متوسط الأثر المركب + الخطأ (داخل الدراسات وبين الدراسات)
10- يوجد متوسط أثر حقيقي واحد.	10- يوجد متوسط أثر مركب لتوزيع تقديرات حجم الأثر العشوائية.
11- الفروق في حجم الأثر الملاحظ ترجع إلى خطأ المعاينة أو خليط المشاركين بالدراسة.	11- الفروق في حجم الأثر الملاحظ ترجع إلى اختلاف الدراسات وخليط المشاركين وفي وسائل تنفيذ التدخل.
12- أوزان الدراسات أقل توازناً حيث تكون الدراسات الكبيرة أكبر وزناً بينما تكون الدراسة الصغيرة ذات وزن أقل نسبياً.	12- أوزان الدراسات أكثر توازناً في عملية التحليل.
13- تباين توزيع تقديرات حجم الأثر الثابتة أقل من نظيره بنموذج الآثار العشوائية.	13- تباين توزيع تقديرات حجم الأثر العشوائية أكبر من نظيره بنموذج الآثار الثابتة.
14- العوامل المؤثرة في تقديرات حجم الأثر تكون نفسها في كل الدراسات ولذلك يكون حجم الأثر الحقيقي هو نفسه في كل الدراسات.	14- العوامل المؤثرة في تقديرات حجم الأثر تختلف داخل الدراسات كما تختلف بين الدراسات مثل العمر الزمني مستوى التعليم والصحة لدى المشاركين بالدراسات.
15- الخطأ المعياري لحجم الأثر المشترك يشمل مصدراً واحداً وهو يساوي $Se_{\bar{y}} = \sqrt{\frac{\sigma^2}{nk}}$ σ = الانحراف المعياري بالدراسة. الخطأ المعياري يمكن أن يصل إلى الصفر حتى في وجود دراسة أو أكثر كبيرة ودقيقة.	15- الخطأ المعياري لحجم الأثر المركب يشمل مصدرين، وهو يساوي $Se_{\bar{y}} = \sqrt{\frac{\sigma^2}{nk} + \frac{T^2}{k}}$ n = عدد المشاركين بالدراسة، k عدد الدراسات.
16- الدراسة الصغيرة ذات تأثير أكبر على حجم الأثر المشترك أما في الدراسة الكبيرة؛ فالعكس صحيح.	16- الدراسة الكبيرة ذات تأثير أقل على حجم الأثر، أما الدراسة الصغيرة؛ فالعكس صحيح.

مثال لتقدير قيمة بعض مكونات الآثار الثابتة (في Cumming, 2012 بتصرف)

الدراسات	حجم الأثر y_i	التباين الداخلي V_{Y_i}	الوزن W_i	تقدير قيمة بعض القيم		
				$w_i y_i^2$	$w_i y_i^2$	w_i^2
1	0.095	0.033	30.352	2.869	0.271	921.21
2	0.277	0.031	32.568	9.033	2.505	1060.68
3	0.367	0.050	20.048	7.349	2.694	401.93
4	0.664	0.011	95.111	63.190	41.983	9046.01
5	0.462	0.043	23.439	10.824	4.999	549.37
6	0.185	0.023	42.698	7.906	1.464	1823.12
المجموع			244.215	101.171	53.915	13802.33

تحديد قيم w_i، y_i، $\bar{y}$، $V_{\bar{y}}$، $S_{e\bar{y}}$، $LL_{\bar{y}}$، $UL_{\bar{y}}$، Z، Q، T^2 في حالة نموذج الأثر الثابت.

وزن الدراسة $$W_i = \frac{1}{0.0329} = 30.3515$$

أثر الدراسة في وزنها $$W_i y_i = 30.3515 \times 0.0945 = 2.8690$$

الأثر المشترك $$\bar{y} = \frac{101.171}{244.215} = 0.4143$$

تباين $\bar{y}$ $$V_{\bar{y}} = \frac{1}{244.215} = 0.0041$$

الخطأ المعياري لـ $\bar{y}$ $$s_{e\bar{y}} = \sqrt{0.0041} = 0.0640$$

الحد الأدنى لفترة الثقة $$LL_{\bar{y}} = 0.4143 - 1.96 \times 0.0640 = 0.2889$$

الحد الأعلى لفترة الثقة $$UL_{\bar{y}} = 0.4143 + 1.96 \times 0.0640 = 0.5394$$

قيمة Z لتحديد دلالة $$Z = \frac{0.4143}{00640} = 6.4739$$

p لاختبار ذو ذيل واحد $$p = 1 - \emptyset(6.4739) < 0.0001$$

p لاختبار ذو ذيلين $$p = 2[1 - \emptyset(|6.4739|)] < 0.0001$$

التباين الكلي $$Q = 53.915 - \left(\frac{101.171^2}{244.215}\right) = 12.0033$$

$$df = (6 - 1) = 5,$$

قيمة C $$C = 244.215 - \left(\frac{13802.325}{244.215}\right) = 187.698$$

التباين بين الدراسات $$T^2 = \frac{12.0033 - 5}{187.698} = 0.0373$$

مثال لتقدير قيمة بعض مكونات الآثار العشوائية (في Cumming, 2012 بتصرف)

الدراسات	حجم الأثر y_i	التباين داخل الدراسات V_{y_i}	التباين بين الدراسات T_i^2	التباين الكلي $V_{Y_i} + T^2$	الوزن W^*	W^*Y_i
1	0.095	0.033	0.037	0.070	14.233	1.345
2	0.277	0.031	0.037	0.068	14.702	4.078
3	0.367	0.050	0.037	0.087	11.469	4.204
4	0.664	0.011	0.037	0.048	20.909	13.892
5	0.462	0.043	0.037	0.080	12.504	5.774
6	0.185	0.023	0.037	0.061	16.466	3.049
المجموع					90.284	32.342

تحديد قيم W_i، y_i، $\bar{y}$، $V_{\bar{y}}$، $S_{e\bar{y}}$، $LL_{\bar{y}}$، $UL_{\bar{y}}$، Z، Q، T^2 في حالة نموذج الآثار العشوائية لنفس المثال.

وزن الدراسة (1)
$$W_1^* = \frac{1}{(0.0329 + 0.0373)} = \frac{1}{(0.070)} = 14.2331$$

متوسط تقديرات حجم الآثار العشوائية
$$\bar{y}^* = \frac{32.342}{90.284} = 0.3582$$

تباين الآثار العشوائية
$$V_{\bar{y}}^* = \frac{1}{90.284} = 0.0111$$

الخطأ المعياري لمتوسط حجم الأثر
$$Se_{\bar{y}}^* = \sqrt{0.0111} = 0.1052$$

الحد الأدنى للفترة
$$LL_{\bar{y}} = 0.3582 - 1.96 \times 0.1052 = 0.1520$$

الحد الأعلى للفترة
$$UL_{\bar{y}} = 0.3582 + 1.96 \times 0.1052 = 0.5645$$

مستوى دلالة $\bar{y}$
$$Z = \frac{0.3582}{0.1052} = 3.4038$$

p للاتجاه الواحد
$$p = 1 - \emptyset(3.4038) = 0.0003$$

p للاتجاهين
$$p = 2[1 - \emptyset(|3.4038|)] = 0.0007$$

أساليب تقدير قيمة حجم الأثر المركب

متوسط حجم الأثر الموزون / الملخص / المشترك العام

لكي نحسب حجم الأثر الموزون أو المشترك من عدة دراسات مجتمعة، فإن هذه الدراسات يجب أن تستخدم المؤشر نفسه لتقديرات حجم أثر تلك الدراسات، كما يجب أن نحدد مدى تجانس أو تباين تلك التقديرات لحجم الأثر، لتحديد نموذج الأثر المناسب، أهو نموذج الأثر الثابت أم نموذج الأثر العشوائي؟ وقد تناولنا ذلك في المبحث الثاني وفي بداية هذا المبحث. وفيما يلي بعض طرق تقدير قيمة حجم الأثر المركب (Cohen, 1988)

تقدير قيمة حجم الأثر المركب $\bar{d}$

$$\bar{d} = \sum_{i=1}^{k} (W_i d_i) / \sum_{i=1}^{k} W_i$$

$$W_i = 1/V_i$$

$$V_i = \frac{n_e + n_c}{n_e . n_c} + \frac{(d_i)^2}{2(n_e + n_c)}$$

الخطأ المعياري لـ $\bar{d}$

$$SE_{\bar{d}} = \sqrt{V_{\bar{d}}} = \sqrt{\frac{1}{\sum w_i}}$$

دلالة $\bar{d}$

$$Z = \frac{\bar{d}}{Se_{\bar{d}}}$$

فترة الثقة لـ $\bar{d}$

$$CI_{\bar{d}} = \bar{d} \pm Z_{\alpha/2}(Se_{\bar{d}})$$

إذا وقع الصفر داخل هذه الفترة فإننا نقبل الفرض الصفري، وهذا يعني أن المتغير المستقل موضع التحليل في مجموعة الدراسات موضع البحث ليس له أثر دال على المتغير التابع. ويلاحظ أنه كلما كانت فترة الثقة متسعة كان تقدير حجم الأثر أقل دقة، ولكن إذا كانت فترة الثقة ضيقة زادت دقة التقدير. من المعادلات السابقة يشير $\bar{d}$ هو حجم الأثر المركب، و w_i هو وزن الدراسة i، V_i تباين حجم الأثر بالدراسة i، d_i هو حجم الأثر بتلك الدراسة. والأوزان المقترحة يمكن أن تكون أحجام العينات في كل دراسة n_i أو مقلوب قيمة تباين حجم الأثر بكل دراسة V_i.

طريقة بيتو Peto **لتقدير قيمة متوسط الأثر الموزون** $\overline{OR}$ (في Brockhaus, et al., 2015)

وتسمى هذه الطريقة طريقة الخطوة الواحدة One step لتقدير قيمة نسبة الأرجحية المركب $\overline{OR}$ ، الذي يعتمد على مقياس لوغاريتم أرجحية النجاح، وتعتمد هذه الطريقة على معادلات مختلفة وفيما يلي جدول يعقبه عدة معادلات:

	يحدث	لا يحدث
المجموعة التجريبية	A_i	B_i
المجموعة الضابطة	C_i	D_i

تفترض هذه الطريقة أن $y_i = \frac{O_i - E_i}{I_i}$

حيث تساوي الكمية الملاحظة أو المقدرة القيمة E_i والتي تتحدد بـ $O_i = A_i$، و E_i هي الكمية أو التقدير المتوقع، وهي تساوي

$$E_i = \frac{(A_i + B_i) \times (A_i + C_i)}{n_i}$$

$$n_i = A_i + B_i + C_i + D_i \quad \text{و}$$

و I_i تساوي $I_i = \frac{(A_i+B_i)\times(C_i+D_i)\times(A_i+C_i)\times(B_i+D_i)}{n^2{}_i \times (n_i-1)}$

وتباين لوغاريتم تقدير نسبة الأرجحية لدراسة واحدة يساوي:

$$VlnOR_i = \frac{1}{I_i}$$

الوزن المعطى للدراسة يساوي معكوس تباينها، وهو يتحدد بالمعادلة:

$$W_i = \frac{1}{V_Y}$$

الوزن يساوي تماماً I_i، والمتوسط الموزون لتقديرات لوغاريتم نسبة الأرجحية يساوي:

$$InOR = \frac{\sum_{i=1}^{k} W_i Y_i}{\sum_{i=1}^{k} W_i}$$

وتباينها يساوي $v_{InOR} = \frac{1}{\sum_{i=1}^{k} W_i}$

والطريقة البديلة لكتابة هذه النتائج هي $InOR = \frac{\sum_{i=1}^{k}(O_i - E_i)}{\sum_{i=1}^{k} I_i}$

وتباينها يعطى بالمعادلة $v_{InOR} = \frac{1}{\sum_{i=1}^{k} I_i}$

والخطأ المعياري باستخدام وحدات لوغاريتمية يساوي

$$SE_{InOR} = \sqrt{V_{InOR}}$$

فترة الثقة للأثر لملخص باستخدام وحدات لوغاريتمية تحسب على النحو التالي:

$$LL_{InOR} = InOR - (1.96)SE_{InOR}$$

$$UL_{InOR} = InOR - (1.96)SE_{InOR}$$

وتتحدد قيمة Z بالمعادلة $Z = \frac{InOR}{Se_{InOR}}$

ثم تحول Z إلى ما يقابلها من قيمة P في حالتي اختبار ذي طرف واحد واختبار ذي طرفين

وتتحدد قيمة P لاختبار ذي طرف واحد من خلال المعادلة $\rho = 2\left[1 - \left(\emptyset(|Z|)\right)\right]$

مثال لطريقة بيتو (في Brockhaus, et al., 2015)

يعرض جدول (5/3) نتائج (6) دراسات تتضمن نسبة الأرجحية OR بجدول (2×2)

جدول (5/3)

نتائج (6) دراسات تتضمن حساب نسبة الأرجحية OR

الدراسات	A	B	C	D	N	O	E	O-E	I	Y	W=I	WY= O-E
1	12	53	16	49	130	12	14	-2	5.5349	-0.3613	5.5349	-2.000
2	8	32	10	30	80	8	9	-1	3.5316	-0.2832	3.5316	-1.000
3	14	66	19	61	160	14	16.5	-2.5	6.5896	-0.3794	6.5896	-2.500
4	25	375	80	320	800	25	52.5	-27.5	22.8332	-1.2044	22.8332	-27.500
5	8	32	11	29	80	8	9.5	-1.5	3.6677	-0.4090	3.6677	-1.500
6	16	49	18	47	130	16	17	-1	6.3256	-0.1581	6.3256	-1.000
المجموع											48.4827	-35.500

لوغاريتم OR $\quad \mathrm{InOR} = \frac{-35.5000}{48.4827} = -0.7322$

تباين لو OR $\quad \mathrm{V_{InOR}} = \frac{1}{48.4827} = 0.0206$

الخطأ المعياري لـ لوغاريتم OR أو lnOR $\quad \mathrm{SE_{InOR}} = \sqrt{0.0206} = 0.1436$

$$\mathrm{LL_{InOR}} = -0.7322 - 1.96 \times 0.1436 = -0.0137$$

$$\mathrm{UL_{InOR}} = -0.7322 - 1.96 + 0.1436 = -0.4507$$

$$\mathrm{Z} = \frac{-0.7322}{0.1436} = -5.098$$

قيمة p لطرف واحد $\quad \rho = 1 - \emptyset(-5.098) = 0.00000017$

قيمة p لطرفين $\quad \rho = 2[1 - (\emptyset(|-5.098|))] = 0.00000034$

في الخطوة الأخيرة، نحول القيم عكسيًا إلى وحدات خام كما هو معروض فيما يلي

$$\mathrm{OR} = \exp(-0.7322) = 0.841$$

$$\mathrm{LL_{OR}} = \exp(-1.0137) = 0.363$$

$$\mathrm{UL_{OR}} = \exp(-0.4507) = 0.637$$

مما سبق يتبين ان طريقة الخطوة الواحدة One step تتبع الخطة الأساسية نفسها لتقدير قيمةOR، وفيها نستخدم أوزان معكوس التباين وتتابع نقاط التحليل ولكنه يختلف في الطريقة قليلاً فقط، لحساب لوغاريتم نسبة الأرجحية وتباينها، وطريقة بيتو هي امتداد لنموذج الآثار العشوائية. وعندما يوجد خلية أو أكثر فارغة في جدول(2X2) فإن صيغة معكوس التباين لا تستطيع أن تتعامل مع الصفر؛ ولذلك فإن هذه الطريقة تضيف 0.5 لكل تكرار بالخلايا الأربع للجدول.

طريقة مانتيل-هانزيل Mantel-Haenszel (في Robins, et al., 1986)

وهذه الطريقة تخص نسبة الأرجحية OR بجدول (2×2) وفيها تستخدم المعادلات الآتية:

حجم الأثر بالدراسة $y_i = \frac{A_i.D_i}{B_i.C_i}$

	يحدث	لا يحدث
المجموعة التجريبية n_1	A_i	B_i
المجموعة الضابطة n_2	C_i	D_i

ويتحدد وزن الدراسة (i) في هذه الطريقة من خلال المعادلة التالية:

$$W_i = \frac{B_i C_i}{n_i}$$

حيث $n_i = A_i + B_i + C_i + D_i$

ويحسب المتوسط الوزني للأثار باستخدام المعادلة

$$\overline{OR}_{MH} = \left(\sum_{i=1}^{K} W_i y_i\right) / \sum_{i=1}^{k} W_i$$

وتتحدد المعادلة السابقة لغوياً بأن متوسط حجم الآثار الوزني يساوي مجموع حاصل ضرب حجم أثر الدراسة في وزنها، مقسوماً على مجموع أوزان الدراسات. وفي هذه الطريقة نلاحظ أن حجم الأثر يقاس بوحدات خام بدلاً من لوغاريتم الوحدات، ولكننا نحتاج إلى تحويل نسبة الأرجحية إلى وحدات لوغاريتمية لحساب الدرجة Z وفترات الثقة. وفيما يلي معادلات هذه القيم منفصلة:

$$R_i = \frac{A_i D_i}{n_i} \qquad S_i = \frac{B_i C_i}{n_i}$$

$$E_i = \frac{(A_i + D_i) A_i D_i}{n_i^2} \qquad F_i = \frac{(A_i + D_i) B_i C_i}{n_i^2}$$

$$G_i = \frac{(B_i+C_i)A_iD_i}{n_i^2} \qquad H_i = \frac{(B_i+C_i)B_iC_i}{n_i^2}$$

وتستخدم المعادلة التالية لتقدير قيمة تباين الأثر المجمع بوحدات لوغاريتمية وهو يساوي:

$$V_{InOR_{MH}} = 0.5\left(\frac{\sum_{i=1}^{K} E_i}{\left(\sum_{i=1}^{k} R_i\right)^2} + \frac{\sum_{i=1}^{k} F_i + \sum_{i=1}^{k} G_i}{\sum_{i=1}^{k} R_i \times \sum_{i=1}^{k} S_i} + \frac{\sum_{i=1}^{k} H_i}{\left(\sum_{i=1}^{k} S_i\right)^2}\right)$$

والخطأ المعياري له $SE_{InOR_{MH}} = \sqrt{V_{InOR_{MH}}}$

فترة الثقة 95% للأثر الملخص باستخدام وحدات لوغاريتم، وسوف يتم حسابها على النحو التالي:

$$LL_{InOR_{MH}} = InOR_{MH} - (1.96 \times SE_{InOR_{MH}})$$

$$UL_{InOR_{MH}} = InOR_{MH} - (1.96 \times SE_{InOR_{MH}})$$

وتحسب قيمة Z بالمعادلة $Z = \frac{\ln_{MH}}{SE_{InOR_{MH}}}$

قيمة *p* لاختبار ذي طرف واحد $\rho = 1 - \emptyset(\pm|Z|)$

حيث نختار Z+ إذا كان الفرق المتوقع في الاتجاه الموجب، و Z– إذا كان في الاتجاه السالب قيمة *p* لاختبار ذو طرفين $\rho = 2[1 - (\emptyset(\pm|Z|))]$

مثال لطريقة مانتيل-هانزيل: يعرض جدولا (6/3)، و(7/3) نتائج (6) دراسات تتضمن نسبة الأرجحية OR بجدول (2×2)؛ حيث يتم تقدير حجم الأثر المركب بطريقة مانتل-هانزيل (في Robins, et al., 1986 بتصرف).

جدول (6/3)

نتائج (6) دراسات تتضمن حساب نسبة الأرجحية OR

الدراسات	تكرار الخلايا الأربع				n_i	OR_i	Wi	OR(WT)
	A	B	C	D				
1	12	53	16	49	130	0.6934	6.5231	4.5231
2	8	32	10	30	80	0.7500	4.0000	3.0000
3	14	66	19	61	160	0.6810	7.8375	5.3375
4	25	375	80	320	800	0.2667	37.5000	10.0000
5	8	32	11	29	80	0.6591	4.4000	2.9000
6	16	49	18	47	130	0.8526	6.7846	5.7846
المجموع							67.0452	31.5452

جدول (7/3)

قيم R و S و E و F و G و H

الدراسة i	R	S	E	F	G	H
1	4.5231	6.5231	2.1224	3.0608	2.4007	3.4622
2	3.0000	4.0000	1.4250	1.9000	1.5750	2.1000
3	5.3375	7.8375	2.5020	3.6738	2.8355	4.1637
4	10.0000	37.5000	4.3125	16.1719	5.6875	21.3218
5	2.9000	4.4000	1.3413	2.0350	1.5588	2.3650
6	5.7846	6.7846	2.8033	3.2879	2.9813	3.4967
المجموع	31.5425	67.0452	14.5064	30.1295	17.0388	36.9157

من الجدولين السابقين نجد أن نسبة الأرجحية $\mathrm{OR_{MH}}$ تساوي:

$$\mathrm{OR_{MH}} = \frac{\mathrm{OR_i WT_i}}{\mathrm{WT_i}} = \frac{31.5452}{67.0452} = 0.4705$$

$\therefore$ لوغاريتم نسبة الاحتمالية تساوي $\mathrm{InOR_{MH}} = \ln(0.4705) = -0.7539$

وتحسب فترة الثقة للأثر الملخص بوحدات لوغاريتمية كما يلي. انظر جدول (8/3)

في الدراسة الأولى على سبيل المثال نجد أن:

$$\mathrm{R} = \frac{(12)(49)}{130} = 4.5231 \qquad \mathrm{S} = \frac{(53)(16)}{130} = 6.5231$$

$$\mathrm{F} = \frac{(12+49)(53)(16)}{130^2} = 3.068 \qquad \mathrm{E} = \frac{(12+49)(12)(49)}{130^2} = 2.1224$$

$$\mathrm{H} = \frac{(53+16)(53)(16)}{130^2} = 3.4622 \qquad \mathrm{G} = \frac{(53+16)(12)(49)}{130^2} = 2.4007$$

ويتحدد تباين لوغاريتم نسبة الاحتمالية كما يلي:

$$\mathrm{V_{InOR_{MH}}} = 0.5\left(\frac{14.5046}{31.5452^2} + \frac{30.1295 + 17.0388}{31.5452 \times 67.0452} + \frac{36.9157}{67.0452^2}\right) = 0.0225$$

ويساوي الخطأ المعياري له $SE_{InOR_{MH}} = \sqrt{0.0225} = 0.1502$

وفترة الثقة $LL_{InOR_{MH}} = -0.7539 - 1.96 \times 0.1502 = -1.0482$

$UL_{InOR_{MH}} = -0.7539 - 1.96 + 0.1502 = -0.45951$

مستوى دلالة حجم الأثر المركب $Z = \frac{-0.7539}{0.1502} = -5.0211$

قيمة p لاختبار ذي طرف واحد $\rho = 1 - \emptyset(+5.0211) = 0.00000026$

قيمة p لاختبار ذي طرفين $\rho = 2\left[1 - \left(\emptyset(-5.0211)\right)\right] = 0.00000051$

وفي النهاية تحول هذه القيم إلى وحدات خام.

$$OR_{MH} = 0.4705$$

$$LL_{OR_{MH}} = \exp(-1.0482) = 0.3506$$

$$UL_{OR_{MH}} = \exp(0.4596) = 0.6315$$

طريقة هيدجز وزملائه (Hedges et al., 1999)

تركز هذه الطريقة على تقدير قيمة حجم الأثر الموزون لمعاملات الارتباط، ولعمل ذلك نحول r إلى Z_r ثم نحسب متوسط حجم الأثر الموزون لتلك الدرجات المحولة. وفي حالة نموذج الأثر الثابت نستخدم المعادلات التالية:

$$W_i = n_i - 3 \qquad V_i = 1/W_i = 1/(n_i - 3)$$

$$\therefore V_{\bar{z}r} = \frac{1}{\Sigma w_i} = \frac{1}{\Sigma w_{(n_i-3)}}$$

$$\therefore SE_{(\bar{z}r)} = \sqrt{\frac{1}{\Sigma w_i}} = \sqrt{\frac{1}{\Sigma w_{(n_i-3)}}}$$

$$\bar{Z}_{r_i} = \frac{\sum_{i=1}^{k} W_i Z_{r_i}}{\sum_{i=1}^{k} W_i} \qquad W_i = n_i - 3$$

ولذلك يمكن تقدير قيمة متوسط حجم الأثر باستخدام المعادلة التالية:

$$\bar{Z}_{r_i} = \frac{\sum_{i=1}^{k}(n_i - 3)Z_{r_i}}{\sum_{i=1}^{k}(n_i - 3)}$$

حيث K هي عدد الدراسات، و n_i يمثل عدد أفراد العينة بالدراسة i. ولتقدير قيمة درجة تباين تقديرات حجم الأثر بمجموعة الدراسات نستخدم المعادلات التالية:

$$Q = \sum_{i=1}^{K} W_i\left(Z_{r_i} - \bar{Z}_r\right)^2 = \sum_{i=1}^{K}(n_i - 3)(Z_{r_i} - \bar{Z}_r)$$

علماً بأن Q تتبع توزيع χ^2 بدرجات حرية K-1 ولذلك يمكن تحديد مستوى دلالة Q لكي نقبل الفرض الصفري أو نرفضه. وللوصول لدلالة متوسط تقديرات حجم الأثر $\bar{Z}_r$ نستخدم المعادلة $Z = \frac{\bar{z}_r}{SE_{\bar{z}r}}$

وفي حالة استخدام نموذج الآثار العشوائية فإننا نستخدم المعادلات التالية:

$$\overline{Z^*}_r = \frac{\sum_{i=1}^{k} W^* \bar{Z}_{r_i}}{\sum_{i=1}^{k} W^*_i}$$

حيث تحديد الأوزان W^*_i باستخدام المعادلة $W^*_i = 1/\left(\frac{1}{W_i} + T^2\right)$

التباين بين الدراسات يمكن أن يقدر بعدة طرق (Hedges & Vivea, 1998) ورغم ذلك استخدم هيدجز Q السابقة و K والثابت C

$$T^2 = \frac{Q - (K-1)}{C}$$

حيث K عدد الدراسات أو المتغيرات موضوع الدراسة، ويتم تحديد المقدار الثابت C باستخدام المعادلة:

$$C = \sum_{i=1}^{K} W_i - \frac{\sum_{i=1}^{K} (W_i)^2}{\sum_{i=1}^{K} W_i}$$

إذا كانت قيمة T^2 سالبة فإننا نستبدل بها القيمة صفراً (لأن التباين بين المجموعات لا يمكن أن يأخذ قيمة سالبة) وتستخدم W^*_i لتقدير قيمة متوسط حجم الأثر $\bar{Z}_r$، وفي النهاية نحدد فترة الثقة لمتوسط حجم الأثر ولحسابها فإننا في حاجة لمعرفة الخطأ المعياري لحجم الأثر، الذي يحسب باستخدام المعادلة

$$SE_{(\overline{Z^*}_r)} = \sqrt{1/\sum_{i=1}^{K} W_1^*}$$

تدور فترة الثقة حول متوسط حجم الأثر، ويتم حسابها باستخدام الخطأ المعياري والقيمة الحرجة بالمنحنى الاعتدالي لاختبار ذي طرفين، التي تساوي 1.96 عند مستوى ثقة 95%، ويمكن تقدير قيمة الحدين الأعلى والأدنى باستخدام معادلة فترة الثقة (95%): $CI = \overline{Z^*}_r \pm 1.96(SE_{(\overline{Z^*}_r)})$

وهذه القيم تحول ثانية للرجوع إلى r

مثال لطريقة هيدجز وزملائه (Hedges et al., 1999)

من المعلومات عن الدراسات التالية يمكن تقدير Q، T^2، $\bar{Z}_r$ وفترة الثقة 95%. لاحظ أن $\bar{Z}_r = \frac{-1933.83}{2167} = -0.892$.

الدراسات	N	W	W^2	r	Z_r	WZ_r	$Z_r - (-0.892)$	$[Z_r-(-0.892)]^2$	$W[Z_r - (-0.892)]^2$
1	135	132	17424	-0.68	-0.83	-109.44	0.06	0.004	0.53
2	1235	1232	1517824	-0.79	-1.07	-1320	-0.18	0.032	39.49
3	570	567	321489	-0.74	-0.95	-538.92	-0.06	0.003	1.91
4	190	187	34969	0.12	0.12	22.55	1.01	1.026	191.89
5	52	49	2401	0.24	0.24	11.99	1.14	1.293	63.36
المجموع		2167	1894107			-1933.83			297.18

$$Q = \sum_{i=1}^{K} W_i \left(Z_{r_i} - \bar{Z}_r\right)^2 = 297.18$$

$$C = \sum_{i=1}^{K} W_i - \frac{\sum_{i=1}^{K} (W_i)^2}{\sum_{i=1}^{K} W_i} = 2167 - \frac{1894107}{2167} = 1292.93$$

$$T^2 = \frac{Q - (K-1)}{C} = \frac{297.18 - (5-1)}{1292.93} = 0.227$$

ولتقدير قيمة $\overline{Z_r^*}$، $SE_{(\overline{Z_r^*})}$، CI نصمم جدول وندون به البيانات كما يلي:

الدراسات	n_i	W_i	$1/W_i$	$(1/W_i) + T^2$	$1/\left[\left(\frac{1}{W_i}\right) + T^2\right]$	Z_r	$W^*.Z_r$
1	135	132	0.008	0.23	4.30	-0.83	-3.562
2	1235	1232	0.001	0.23	4.42	-1.07	-4.740
3	570	567	0.002	0.23	4.41	-0.95	-4.188
4	190	187	0.005	0.23	4.34	0.12	0.523
5	52	49	0.020	0.25	4.07	0.24	0.997
المجموع		2167			21.53		-10.97

$$\overline{Z^*}_r = \frac{\sum_{i=1}^{k} W^* \bar{Z}_{r_i}}{\sum_{i=1}^{k} W^*_i} = \frac{-10.97}{21.53} = -0.509$$

ثم يتم تحويل $\bar{r}$ إلى $\overline{Z_r^*}$ باستخدام جداول فيشر أو استخدام المعادلة 2، وهذا يعطي الأثر الحقيقي الذي يساوي 0.469- (لاحظ أن هذه القيمة أقل من حجم الأثر في نموذج الأثر الثابت (0.666-)).

وللحصول على فترة الثقة يجب تقدير قيمة $\overline{Z^*}_r$ وذلك باستخدام المعادلة:

$$SE_{(\overline{Z^*}_r)} = \sqrt{1/\sum_{i=1}^{k} W^*_i} = \sqrt{1/21.23} = 0.215$$

$$CI_{UL} = \bar{Z}_r \pm Z_{\alpha/2}.SE(\overline{Z^*}_r)\ (95\%)$$

$$= 0.509 \pm 1.96(0.215)]$$

$$= \bar{Z}_r \pm 0.422$$

$$= [-0.9319\ \&\ -0.087]$$

وبتحويل قيمتي r_z إلى r فإن فترة الثقة تكون

$$CI_{UL} = [-0.731\ \&\ -0.087]$$

طريقة هنتر وشميدت (Hunter & Schemdt, 2006)

هي طريقة فعالة تؤكد عزل مصادر الخطأ؛ مثل خطأ المعاينة وأخطاء القياس، والاختلاف الرئيسي بين هذه الطريقة وطريقة هيدجز وزملائه تبدو في عدم تحويل تقديرات حجم الأثر r عند حساب متوسط حجم الأثر الموزون، ووزن الدراسة باستخدام حجم العينة. ويحسب متوسط حجم الأثر الموزون من خلال المعادلة:

$$\bar{r} = \left(\sum_{i=1}^{k} n_i r_i\right) / \sum_{i=1}^{k} n_i$$

ويتحدد تباين تقديرات حجم الأثر من خلال متوسط مربع الخطأ الموزون، باستخدام المعادلة التالية:

تباين تقديرات حجم الأثر $\sigma_r^2 = \sum_{i=1}^{k} n_i(r_i - \bar{r})^2 / \sum_{i=1}^{k} n_i$

تباين الخطأ $\sigma_e^2 = (1 - r)^2 / \sum_{i=1}^{k}(n_i - 1)$

تباين حجم الأثر بالمجتمع $\sigma_i^2 = \sigma_r^2 - \sigma_{e\bar{r}}^2$

وتحسب فترة الثقة باستخدام المعادلة $CI_{95\%} = \bar{r} \pm 1.96\sqrt{\sigma_p^2}$

ومثلما فعلت طريقة هيدجز وزملائه، فإن دلالة متوسط حجوم الأثر يتم الحصول عليه بحساب الدرجة المعيارية Z، وذلك بقسمة متوسط تقديرات حجم الأثر على الخطأ المعياري. إلا أن أسلوب تقدير قيمة الخطأ المعياري يختلف في طريقة هنتر وشميدت.

الخطأ المعياري $Se_{\bar{r}} = \sqrt{\frac{\sigma_r^2}{k}} = \frac{S_r}{\sqrt{k}}$

كذلك يحسب مربع خطأ المعاينة باستخدام المعادلة

$$SE_{\bar{r}} = \Sigma n_i (r_i - \bar{r})^2 / \Sigma n_i$$

$$\therefore Z = \bar{r}/SE_{\bar{r}}$$ دلالة Z تساوي

مثال لطريقة هنتر وشميدت (Hunter & Schemdt, 2006)

من المعلومات عن الدراسات التالية يمكن تقدير $\bar{r}$ و σ_r^2 و σ_e و σ_p^2، و $IC_{\bar{r}}$

الدراسات	n_i حجم العينة	r حجم الأثر بالدراسة	$n_i r_i$	$[r-(-0.666)]^2$	$n_i[r-(-0.666)]^2$
1	135	-0.68	-91.80	0.000196	0.26
2	1235	-0.79	-975.65	0.015376	18.99
3	570	-0.74	-421.80	0.005476	3.12
4	190	0.12	22.80	0.617796	117.38
5	52	0.24	12.48	0.820836	42.68
المجموع	2182		-1453.97		182.43

$$N = \sum_{i=1}^{K} n_i = 2182\ , \quad \sum n_i r_i = -1453.97$$

$$\therefore \bar{r} = \frac{-1453.97}{2182} = -0.666,\ \sigma_{\bar{r}}^2 = \frac{182.43}{2182} = 0.0836$$

$$\sigma^2{}_e = [(1-(-0.666)^2)]^2 / [\frac{2182}{5} - 1] = 0.000711$$

$$\sigma^2{}_P = 0.0836 - 0.000711 = 0.0829$$

$$CI_{95\%} = \left[-0.666 \pm 1.96\sqrt{0.0829}\right] = [-0.10\& -1.23]$$

ويمكن تقدير قيمة مدى تجانس تقديرات حجم الأثر (لمعاملات الارتباط) باستخدام اختبار مربع كاي:

$$\chi^2 = \sum_{i=1}^{k} \frac{(n_i-1)(r_i-\bar{r})^2}{(1-r^2)^2}$$ بدرجات حرية k-1

وقد قارن جونسين وآخرون (Johnson et al., 1995) من خلال قاعدة بيانات موحدة، وتوصلوا إلى أن طرق كل من هيدجز وأولكن (الأثر الثابت)، و روزينتال وروبن، وهنتر وشميدت، متقاربة فيما يتعلق بمتوسط حجم الأثر ومدى تجانسها أو تباينها، ولكنها اختلفت في دلالة متوسط حجم الأثر خلال عملية ما وراء التحليل. حيث تبين أن طريقة هنتر وشميدت تصل إلى تقديرات محافظة (ومن ثم اتساع فترة الثقة) أكثر من الطريقتين الأخريين، ولذلك فإن طريقة هنتر وشميدت يمكن أن تستخدم عند الاهتمام بالحصول على نتائج متحفظة.

طريقة ستوفر Stouffer (Stouffer, et al., 1949)

وفيها نحسب z_i لكل قيمة p لطرف واحد، وهو يكون محسوب مباشرة من حجم الأثر وخطئه المعياري بدلاً من الطريقة الأولى؛ حيث يتم الحساب على النحو الآتي:

القيمة المتوسطة للأثر $\bar{Z}_{stof.} = \left(\sum_{i=1}^{k} Z_i\right)/\sqrt{k}$

حيث k هي عدد الدراسات، و $\bar{Z}_{stof.}$ هي الدرجات المعيارية من توزيع المنحنى الاعتدالي.

مثال لطريقة ستوفر(Stouffer, et al., 1949)

يشمل الجدول الآتي 33 دراسة تتضمن حجم الأثر لاستخدام أحد العقاقير الطبية باستخدام نسبة المخاطرة RR؛ والمطلوب حساب كل من $\bar{Z}_{sto.}$ ودلالتها.

الدراسات	نسبة المخاطرة (RR)	الحد الادنى	الحد الاعلى	Z	قيمة p لطرفين	قيمة p لطرف واحد	لو p لطرف واحد
1	0.229	0.030	1.750	-1.420	0.155	0.078	-2.554
2	0.571	0.196	1.665	-1.026	0.305	0.152	-1.881
3	1.349	0.743	2.451	0.984	0.325	0.837	-0.177
4	0.703	0.534	0.925	-2.519	0.012	0.006	-5.134
5	1.223	0.669	2.237	0.654	0.513	0.743	-0.296
6	1.011	0.551	1.853	0.034	0.973	0.513	-0.667
7	0.779	0.478	1. 268	-1.005	0.315	0.157	-1.849
8	0.457	0.252	0.828	-2.581	0.010	0.005	-5.315
9	2.377	0.649	8.709	1.307	0.191	0.904	-0.100
10	0.964	0.332	2.801	-0.068	0.946	0.473	-0.749
11	1.048	0.481	2.282	0.117	0.907	0.547	-0.604
12	2.571	0.339	19.481	0.914	0.361	0.820	-0.199
13	0.922	0.609	1.394	-0.386	0.699	0.350	-1.051
14	0.608	0.417	0.886	-2.590	0.010	0.005	-5.338
15	0.702	0.443	1.110	-1.514	0.130	0.065	-2.734
16	0.282	0.034	2.340	-1.172	0.241	0.121	-2.116
17	1.161	0.840	1.604	0.905	0.366	0.817	-0.202
18	0.813	0.263	2.506	-0.361	0.718	0.359	-1.025
19	0.612	0.356	1.050	-1. 782	0.075	0.037	-3.286
20	0.880	0.619	1.250	-0.713	0.476	0.238	-1.436
21	0.827	0.749	0.914	-3.738	0.000	0.000	-9.286
22	0.429	0.041	4.439	-0.710	0.477	0.239	-1.432
23	0.079	0.005	1.350	-1.752	0.080	0.040	-1.828
24	0.333	0.038	2.925	-0.991	0.322	0.161	-3.222
25	1.095	0.073	16.427	0.066	0.948	0.526	-0.642
26	0.500	0.132	1.887	-1.023	0.306	0.053	-1.876
27	0.621	0.151	2.555	-0.660	0.510	0.255	-1.367
28	0.174	0.040	0.761	-2.323	0.020	0.010	-4.596
29	0.604	0.188	1.944	-0.845	0.398	0.199	-1.614
30	0.462	0.048	4.461	-.668	0.504	0.252	-1.378
31	0.654	0.322	1.331	-1.171	0.241	0.121	-2.114
32	0.769	0.704	0.839	-5.862	0.000	0.000	-19.940
33	0.244	0.051	1.164	-1.770	0.077	0.038	-3.260
المجموع				-33.677			-89.268

من الجدول السابق يمكن تقدير قيمة نسب المخاطرة لدواء ما بفترة ثقة 95% في 33 دراسة حيث نحسب Z من المعادلات التالية:

$$\bar{Z} = \frac{\sum \ln(RR_i)}{[[\ln(UL_i) - \ln(LL_i)]/[2(1.96)]]}$$

حيث UL_i هو الحد الأعلى للفئة، وLL_i هو الحد الأدنى للفئة i

ويمكن الوصول لدلالة الإحصاء Z من المعادلة $Z = \frac{\Sigma \ln(RR_i)}{SE_{\ln(RR_i)}}$

حيث البسط هو لوغاريتم نسبة المخاطرة والمقام هو الخطأ المعياري لنسبة المخاطرة. ونحصل على قيم p لاتجاه واحد ويتم الحصول منها على قيم Z بالتوزيع الاعتدالي المعياري. ومن الجدول السابق نجد أنه يتضمن مجموع قيم (Z) تساوي 33.677-؛ ومجموع لوغاريتم Ln لقيم p لطرف واحد تساوي 89.268-

$$\chi^2 = (-2)\Sigma \text{Ln Pi}$$

$$\therefore \chi^2 = (-2)(-89.268) = 178.536$$

ومن خلال قيمة χ^2 عند درجات حرية 66 (ضعف عدد الدراسات أو 2K) تحول إلى p وهي قيمة صغيرة جدًا ($p = 2.6\times10^{-12}$). وبطريقة ستوفر نجد أن قيمة Z تساوي $Z = -33.677/\sqrt{33}$. وبتحويل قيمة Z الكلية السابقة إلى قيمة p ($p = 2.3\times10^{-9}$). وتؤكد قيمة p باستخدام طريقتي فيشر وستوفر فائدة الدواء (حيث كانت قيمة الإحصاء دالة إحصائياً). هذا الاختبار لا يخبرنا بأي شيء حول قيمة حجم الأثر المشترك أو دلالته الإحصائية. ويمكن استخدام هذه الطريقة إذا كانت لكل النتائج دلالة إحصائية (p)، حيث يمكنك تجميع قيم p من دراسات مختلفة استخدمت أي اختبار إحصائي (F، t، r،.. ....) والقيام بتحليل بعدي لنتائج تلك الدراسات.

المبحث الرابع

المقارنة بين المجموعات الفرعية

في خصائصها الإحصائية

للمقارنة بين المجموعات الفرعية في حجم أثرها يجب أن نركز على حساب حجم الأثر بكل مجموعة، ثم المقارنة بين حجوم الأثر بمجموعات الدراسة، لكي نحدد في أي المجموعات يكون حجم الأثر أكبر. مثل المقارنة بين أثر أسلوبين من التدريس في مجموعتين مختلفتين، وفي هذه الحالة يجب أن يكون مؤشر حجم الأثر بالدراسات كلها بالنمط نفسه، وليكن g أو d (الفرق المتوسط المعياري)، أو نسبة الأرجحية أو نسب المخاطرة بعد تحويلها لوغاريتمياً، وكذلك معاملات الارتباط بعد تحويلها إلى Z فيشر. وتوجد طرق متعددة لمقارنة المجموعات الفرعية، وهي اختبارات t أو Z واختبار Q المعتمد على تحليل التباين، واختبار Q المعتمد على قياس التباين Heterogeneity. وفيما يلي سوف نقارن العديد من خصائص الإحصائية بين المجموعات الفرعية وخاصة المقارنة بينهم في تقديرات حجم الأثر والتباين. وسوف يتم عرض ذلك في حالة نموذج الأثر الثابت وفي نموذج الآثار العشوائية، وأهمها:

1) حساب بعض القيم الإحصائية للمجموعات الفرعية.
2) تحديد دلالة الفرق بين تقديرات حجم الأثر بالمجموعات الفرعية (قيمة z ودلالتها).
3) تحديد الفروق بينها باستخدام Q_w , Q_B, Q.
4) تحديد الفروق بينها باستخدام T^2 المجمعة.

المقارنة بين المجموعات الفرعية في بعض الخصائص الإحصائية في إطار نموذج الآثار الثابتة

لنفرض أن لدينا عشر دراسات تمت خمس منها على مجموعة (A)، والخمس الأخرى على مجموع (B) كما في جدول (1/4). يريد الباحث مقارنة حجم الأثر في المجموعتين (A) و (B). يمكن أن نرى من الجدول (1/4) أن حجم الأثر المشترك بالدراسات الخمس (A) يساوي 0.324 وتباينه 0.003، وأن حجم الأثر المشترك بالدراسات الخمس (B) 0.611 وتباينه يساوي 0.002 ولمقارنة المجموعتين نحتاج إلى إجراء عمليتين، وهما:

أ- مقارنة متوسط الأثر والتباين الداخلي بكلتا المجموعتين الفرعيتين.

ب- مقارنة متوسط الأثر والتباين بين المجموعتين الفرعيتين.

يبين جدول (1/4) قيم تقديرات حجم الأثر بكل مجموعة فرعية وكذلك الأثر الملخص (أو متوسط الأثر) بكل منهما والتباين الداخلي والتباين بين الدراسات ووزن كل منها (في (Boernstein, et al., 2009

جدول (1/4)

بيانات حجم الأثر بالدراسات العشر، والمجموعتين الفرعيتين A و B

	الدراسات	حجم الأثر Y_i	التباين داخل $V_{\bar{Y}}$	التباين بين الدراسات T^2	التباين الكلي V	الوزن W_i	قيم مشتقة W_{iYi}	$(W_{iYi})^2$	W^2
	1	0.110	0.0100	0.0000	0.0100	100.000	11.000	1.210	10000.000
	2	0.224	0.0300	0.0000	0.0300	33.333	7.467	1.673	1111.111
A	3	0.338	0.0200	0.0000	0.0200	50.000	16.900	5.712	2500.000
	4	0.451	0.0150	0.0000	0.0150	66.667	30.067	13.560	4444.444
	5	0.480	0.0100	0.0000	0.0100	100.000	48.000	23.040	10000.000
	A	0.324	.003			350.000	113.433	45.195	28055.556
	1	0.440	0.015	0.0000	0.0150	66.667	29.333	12.907	4444.444
	2	0.492	0.0200	0.0000	0.0200	50.000	24.600	12.103	2500.000
B	3	0.651	0.0150	0.0000	0.0150	66.667	43.400	28.253	4444.444
	4	0.710	0.0250	0.0000	0.0250	40.000	28.400	20.164	1600.000
	5	0.740	0.0120	0.0000	0.0120	83.333	61.667	45.633	6944.444
	B	0.611	0.003			306.667	187.400	119.061	19933.333
	المجموع الكلي	0.458	0.002			656.667	300.833	164.255	47988.889

وفيما يلي تقدير قيمة بعض الإحصاءات المشتقة من بيانات الجدول (1/4)

1- تقدير قيمة القيم الإحصائية لمجموعة الدراسات الفرعية (A) باستخدام نموذج الأثر الثابت:

- متوسط حجم أثر مجموعة الدراسات A

$$\bar{y}_A = \frac{113.433}{350.000} = 0.3241$$

- تباين المجموعة A

$$V_{\bar{y}A} = \frac{1}{350.000} = 0.0029$$

- الخطأ المعياري لـ $\bar{y}_A$

$$SE_{\bar{y}A} = \sqrt{0.0029} = 0.0535$$

- الحد الأدنى لفترة الثقة $\bar{y}_A$

$$LL_{\bar{y}A} = 0.3241 - 1.96 \times 0.0535 = 0.2193$$

- الحد الأعلى لفترة الثقة $\bar{y}_A$

$$UL_{yA} = 0.3241 + 1.96 \times 0.0535 = 0.4289$$

- قيمة اختبار Z_A ودلالة $\bar{y}_A$

$$Z_A = \frac{0.3241}{0.0535} = 6.0633$$

- قيمة p المناظرة لـ Z_A (دالة $\bar{y}_A$)

$$P(Z_A) =< 0.0001$$

- قيمة Q_A

$$Q_A = 45.195 - \frac{133.433^2}{353.000} = 8.4316$$

- دلالة C_A

$$P(Q = 8.4316, df = 4)\ 0.0770$$

- قيمة C_A

$$C_A = 350.000 - \frac{28055.006}{350.000} = 269.8413$$

- التباين بين دراسات المجموعات A

$$T_A^2 = \frac{8.4316 - 4}{269.8413} = 0.0164$$

- قيمة I_A^2

$$I_A^2 = \frac{(8.4316 - 4)}{8.4316} \times 100 = 52.5594$$

2- تقدير قيمة القيم الإحصـــائية لمجموعة الدراســـات الفرعية (B) وذلك باســـتخدام نموذج الأثر الثابت:

- متوسط حجم أثر مجموعة الدراسات B

$$\bar{y}_B = \frac{187.400}{306.667} = 0.3241$$

- تباين المجموعة B

$$V_{\bar{y}B} = \frac{1}{306.667} = 0.0029$$

- الخطأ المعياري لـ $\bar{y}_B$

$$SE_{\bar{y}B} = \sqrt{0.0033} = 0.0571$$

- الحد الأدنى لـ $\bar{y}_B$

$$LL_{\bar{y}B} = 0.06111 - 1.96 \times 0.0571 = 0.4992$$

- الحد الأعلى لـ $\bar{y}_B$

$$UL_{\bar{y}B} = 0.6111 + 1.96 \times 0.0571 = 0.7230$$

- قيمة اختبار Z_B ودلالة Y_B

$$Z_B = \frac{0.6111}{0.0571} = 10.7013$$

- قيمة p المناظرة لـ Z_B

$$P(Z_B) =< 0.0001$$

- قيمة Q_B

$$Q_B = 45.195 - \frac{187.400^2}{306.667} = 4.5429$$

- دلالة C_B

$$P(Q = 4.5429, df = 4) = 0.337$$

- قيمة C_B

$$C_B = 306.667 - \frac{19933.333}{306.667} = 241.667$$

- التباين بين دراسات المجموعات B

$$T_B^2 = \frac{4.5429 - 4}{241.667} = 0.0022$$

- قيمة I_B^2

$$I_B^2 = \frac{(4.5429 - 4)}{4.5429} \times 100 = 11.9506$$

3- تقدير قيمة القيم الإحصائية للدراسات العشر بالمجموعتين الفرعيتين معاً

- متوسط حجم الأثر الكلي

$$\overline{Y}_{AB} = \frac{300.833}{656.667} = 0.4581$$

- التباين الكلي للدراسات ككل

$$V_{AB} = \frac{1}{656.667} = 0.0015$$

- الخطأ المعياري لمتوسط الأثر الكلي

$$SE_{AB} = \sqrt{0.0015} = 0.0390$$

- الحد الأدنى لفترة الثقة للأثر الكلي

$$LL_{\overline{Y}AB} = 0.4581 - 1.96 - \times 0.0390 = 0.3816$$

- الحد الأعلى لفترة الثقة للأثر الكلي

$$UL_{\overline{Y}AB} = 0.4581 - 1.96 \times 0.0390 = 0.5346$$

- قيمة اختبار Z ودلالة $\overline{Y}_{AB}$

$$Z_B = \frac{0.4581}{0.0390} = 11.7396$$

- قيمة P المناظرة لـ Z (مستوى دلالة $\bar{y}_{AB}$)

$$P(Z) < 0.0001$$

- قيمة Q_{AB}

$$Q_{AB} = 164.255 - \frac{300.833^2}{656.667} = 26.4371$$

- دلالة Q

$$P(Q = 26.4371, df = 9) = 0.0017$$

- قيمة C_{AB}

$$C_{AB} = 656.667 - \frac{47988.889}{656.667} = 583.5871$$

- التباين بين المجموعات

$$T_{AB}^2 = \frac{226.4371 - 9}{538.5871} = 0.0299$$

- النسبة المئوية للتباين

$$I_{AB}^2 = \frac{(26.4371 - 9)}{26.4371} \times 100 = 65.96$$

ويلخص جدول (2/4) نتائج التحليل للمجموعتين الفرعيتين في إطار نموذج الآثار الثابتة.

جدول (2/4)

ملخص تحليل المجموعتين الفرعيتين والمجموعة الكلية

الرموز الإحصائية	A	B	AB (المشترك)
$\overline{Y}$	0.3241	0.6111	0.4581
V	0.0029	0.0033	0.0015
$SE_{\overline{y}}$	0.0535	0.0571	0.0390
$LL_{\overline{y}}$	0.2193	0.4992	0.3816
$UL_{\overline{y}}$	0.4289	0.7230	0.5346
Z	6.0633	10.7013	11.7396
P2	0.0000	0.0000	0.0000
Q	8.4316	4.5429	26.4371
Df	4.0000	4.0000	9.0000
P-value	0.0770	0.3375	0.0017
Numerator	4.4316	0.5429	17.4371
C	269.8413	241.6667	583.5871
T^2	0.0164	0.0022	0.0299
I^2	52.5594	11.9500	65.9569

وتتم المقارنة بين متوسطي حجم الأثر بثلاث طرق متكافئة، تصل إلى نفس القيمة *p*، وفيما يلي العمليات الحسابية لتلك الطرق.

1- مقارنة متوسطي تقديري حجم أثر مجموعتين فرعيتين باستخدام اختبار Z

هذه الطريقة تعتمد على متوسط تقديرات حجم الأثر وتباينها بالمجموعات الفرعية للدراسات أكثر من الأفراد المشاركين، وبافتراض أن Θ_A وΘ_B هما متوسطا حجم الأثر الحقيقي للمجموعتين الفرعيتين A و B، وأن $\overline{Y}_A$ و$\overline{Y}_B$ هما متوسطا حجم الأثر المقدر وتباينهما بالمجموعتين V_A و V_B، وأن الفرق بين الأثرين هو D_{YAB}.

الفرق بين حجما الأثر بالمجموعتين A وB $\quad D_{\overline{Y}AB} = \overline{Y}_A - \overline{Y}_B$

قيمة اختبار Z (دلالة الفرق بين المجموعتين) $\quad Z = \dfrac{D_{\overline{y}AB}}{S_{eD\overline{Y}}}$

الخطأ المعياري للفرق بين حجمي أثر المجموعتين $\quad SE_{DY} = \sqrt{V_{yA} + V_{yB}}$

وبالنسبة لاختبار ذو طرفين فإن قيمة p تحسب بالمعادلة التالية: $p = 2[C1 - \phi|Z|]$

حيث $\phi|Z|$ هو التوزيع المتجمع الاعتدالي المعياري، وفي المثال الحالي:

- الفرق بين حجمي الأثر $D_{\bar{Y}AB} = 0.6111 - 0.3241 = 0.2870$
- الخطأ المعياري للفرق بين حجمي أثر المجموعتين $SE_{D\bar{Y}} = \sqrt{0.0029 + 0.00033} = 0.0782$
- قيمة اختبار Z لتقدير قيمة $D_{\bar{Y}}$ $Z = \frac{0.2870}{0.0782} = 3.6691$
- مستوى دلالة Z (دلالة الفرق بين متوسطي حجم الأثر $p = 2[1 - (\phi(|306691|))] = 0.0002$

وهذا يعني وجود فرق دال بين حجمي الأثر في المجموعتين A و B

2- مقارنة عدة مجموعات فرعية باستخدام تحليل التباين (في Cumming, 2012)

تُســـتخدم الطريقة الأولى عند وجود مجموعتين فرعيتين فقط، أما إذا وجدت ثلاث مجموعات فرعية أو أكثر؛ فإننا نســـتخدم تحليل التباين والمقارنات المتعددة، وفيه نجزئ التباين الكلى (تباين كل المشـاركين حول المتوسـط العام للمجموعات) إلى تباينين: التباين داخل المجموعات (تباين المشـاركين بكل مجموعة حول متوسـطها) والتباين بين المجموعات (تباين متوسـطات المجموعات حول المتوسـط العام للمجموعات) ثم نختبر الدلالة الإحصائية للتباينات التالية:

1- مجموع مربعات الانحرافات بكل الدراسات الخمس A (Q_A).
2- مجموع مربعات الانحرافات بكل الدراسات الخمس B (Q_B)
3- Q_W وهي تساوي ($Q_A + Q_B$) أو مجموع مربعات الانحرافات بالدراسات العشر مجتمعة.
4- Q_B وهي تســـاوي مجموع مربعات الانحرافات الموزونة بين متوســـطات آثار الجماعات الفرعية حول المتوسط العام.
5- Q_T، وهي تساوي مجموع مربعات الانحرافات الموزونة حول المتوسط العام للآثار.

وفيما يلي يعرض جدول (3/4) نتائج تحليل التباين للمجموعات الفرعية (A وB) وفقاً لنموذج الأثر الثابت.

جدول (3/4)

نتائج تحليل التباين للمجموعات الفرعية

مصدر التباين	Q	Df	P
A	8.4316	4	0.0770
B	4.5429	4	0.3375
Within	12.9745	8	0.1127
Between	13.4626	1	0.0002
Total	26.4371	9	0.0017

ويتحدد مجموع مربعات الانحرافات الموزونة داخل المجموعات باستخدام المعادلة التالية:

$$Q_W = \sum_{J=1}^{K} Q_j$$

حيث K عدد المجموعات، و j رقم المجموعة

$$Q_w = 8.4316 + 4.5429 = 12.9745$$

حيث $Q = Q_w + Q_B$

$$\therefore Q_B = Q - Q_w$$

$$Q_B = 26.4371 - 12.9745 = 13.4626$$

ويتم تقييم إحصاء Q وفقاً لتوزيع χ^2 عند درجات الحرية المناظرة، كما يلي:

أ- Q الكلية تساوي 26.437 ودرجة حرية تساوي 9 ولذلك فإن P = .0017، وهذا يشير إلى وجود فروق دالة بين المجموعات العشر.

ب- Q_w. تساوي 12.975 ودرجة حرية تساوي 8 ولذلك فان P= 0.1127 وهذا يشير إلى أنها غير دالة.

ج- Q_B. تساوي 13.4626، ودرجة حرية تساوي 1، لذلك فان P تساوي 0.0002 وهذا يعني أنها دالة، وهذا يعني وجود فروق بين المجموعتين الفرعيتين لصالح المجموعة (B).

د- Q_A= 8.432، df_A= 4، P= 0.077 ؛ وهذا يعني أن التباين داخل المجموعة الفرعية A غير دالة (يوجد تجانس بين هذه الدراسات).

ه- Q_B= 4.543، df_B= 4، P=. 3375 وهذا يعني وجود تجانس داخل دراسات المجموعات الفرعية B؛ وهذا يعني أن تلك المجموعات تشترك في حجم أثر مشترك عام.

3- **مقارنة حجمي الأثر بالمجموعتين الفرعيتين باستخدام اختبار Q** (Brockhaus, et al., 2015)

في مثالنا الحالي نجد أن حجمي الأثر بالمجموعتين الفرعيتين A، B يساويان 0.324 و0.611 على الترتيب، وتباينهما 0.003 و 0.003 على الترتيب. ذلك نطبق معادلات تقدير قيمة Q كما يعرض جدول (4/4).

جدول (4/4)

نتائج مقارنة حجمي الأثر بالمجموعتين الفرعيتين باستخدام اختبار Q

الدراسات الفرعية	حجم الأثر $\overline{Y}$	التباين داخل V_y	التباين بين T^2	التباين الكلى V	الوزن W	قيم مشتقة		
						Wy	Wy^2	W^2
A	0.3241	0.0029	0.0000	0.0029	350.000	113.433	36.763	122500.000
B	0.6111	0.0033	0.0000	0.0033	306.667	187.400	114.518	94044.444
					656.667	300.833	151.281	216544.444

- حجم الأثر الكلي $\overline{Y} = \frac{300.833}{656.667} = 0.4581$
- تباين y داخل الدراسات $V_y = \frac{1}{656.667} = 0.0015$
- التباين بين الدراسات Q_B $Q_B = 151.281 - (\frac{300.833^2}{656.667}) = 13.4626$
- درجة الحرية لـ Q_B $df = 2 - 1 = 1$
- قيمة P دالة $P\ (Q_B = 13.4626, df = 1) = 0.002$

وحيث إن Q_B دالة الفرق لصالح حجم الأثر الأكبر ونلاحظ أن الطريقتين المستخدمتين لـــQ_B تعطى القيمة نفسها (13.4626) عندما توجد درجة حرية مساوية للواحد؛ إذ إن قيمة Z تساوي الجذر التربيعي لـ Q. وقد بلغت قيمة Z= 3.6691، كما كانت قيمة *p* متساوية (*p* = 0.0002).

فترة الثقة للفرق بين حجمي أثر المجموعتين A وB

- الفرق بين حجمي الأثر $D_y = y_A - y_B$
  $D_y = 0.6111 - 0.3241 = 0.2870$
- الخطأ المعياري للفرق $S_{eDy} = \sqrt{0.0029 + 0.0033} = 0.0782$
- اختبار Z $\therefore Z = \frac{0.2870}{0.0782} = 3.67$
  $Z = \sqrt{Q}$
  $Q = 13.463$

- فترة الثقة للفرق بين حجماياالأثر $Ic = D_{YAB} \pm 1.96\ (S_{eDy})$
- فترة الثقة $Ic_{Dy} = 0.2870 \pm 1.96\ (0.0782)$

$= [\ 0.4403, 0.1337\ ]$

تقدير قيمة T^2 في إطار نموذج الأثر الثابت

الأثر الثابت الحقيقي في نموذج الآثار الثابتة يكون هو نفسه في كل الدراسات، بينما في نموذج الآثار العشوائية يمكن أن تختلف الآثار الحقيقية من دراسة إلى أخرى، وهذا الفرق بين النموذجين له أثره في طرق وزن الدراسات الأثر المتوسط وخطئه المعياري. وفي إطار نموذج الأثر الثابت نفترض أن كل الدراسات في المجموعة الفرعية A تشترك في حجم مشترك عام، وكذلك كل الدراسات في المجموعة الفرعية B تشترك في حجم أثر مشترك عام، بينما يسمح نموذج الآثار العشوائية بوجود تشتت حقيقي للأثار داخل دراسات المجموعة الفرعية A، وكذلك داخل دراسات المجموعة الفرعية B.

ومن الأمثلة الجيدة على نموذج الأثر الثابت، نفترض أن شركة للأدوية أشركت 1000 مريض لمعالجتهم من مرض معين، وتم تقسيمهم إلى عشر جماعات، بكل جماعة 100 مريض، وهذه الجماعات متماثلة في كل أوجه المهمة. ولذلك يمكن أن نفترض أن الأثر الحقيقي سوف يكون هو نفسه في الدراسات العشر (لأن الدراسات يجريها الباحثون أنفسهم، الذين يستخدمون المجتمع نفسه والطرق نفسها). وإذا استخدمنا الدواء D_a مع المجموعات الخمس A والدواء D_b مع المجموعات الخمس B فإننا يمكن أن نستخدم نموذج الأثر الثابت مع المجموعتين الفرعيتين كل على حدة؛ حيث تعمل كل الدراسات بالباحثين أنفسهم وتستخدم المجتمع نفسه والطرق نفسها.

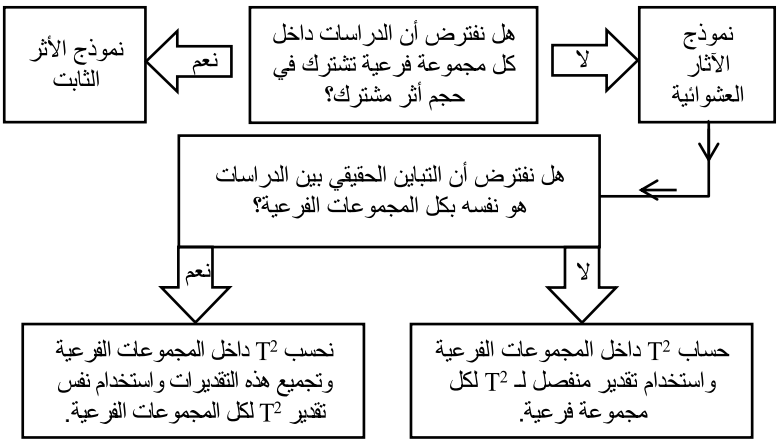
ولتطبيق نموذج الآثار الثابتة، نحتاج إلى تقدير T^2، حيث نحسب التباين داخل مجموعة الدراسات الفرعية بالمجموعة A (التباين بين أثار الدراسات الخمس ومتوسط حجم أثرها $\bar{Y}_A$) ، وكذلك التباين داخل الدراسات الفرعية الخمس بالمجموعة B (التباين بين أثار الدراسات الخمس ومتوسط حجم أثرها $\bar{Y}_B$). وكذلك تحديد التباين داخل الدراسات العشر معاً، وذلك من خلال الفرق بين أثارها المنفردة ومتوسط حجم الأثر العام.

وفي مثالنا الحالي نجد أن T^2 داخل المجموعة الفرعية A يساوي 0.016، وفي المجموعة الفرعية B يساوي 0.002، وفي المجموعة الكلية للدراسات يساوي 0.023 . ويجب ملاحظة أنه إذا كان عدد الدراسات في المجموعة الفرعية 5 أو أقل فإن قيمة T^2 داخل المجموعة الفرعية يحتمل أن تكون غير دقيقة، ولذلك يجب أن نتأكد من أن عدد الدراسات كاف لنصل إلى تقدير دقيق مقبول لـ T^2، كذلك فإن اختلاف T^2

مجموعة فرعية إلى أخرى يتطلب استخدام تقديرات منفصلة لـــ T^2، وكذلك فإن زيادة الدقة التي نحصل عليها بتجميع دراسات أكثر يحتمل أن تقترب من الفروق الحقيقية بين الدراسات في قيم T^2 ($T^2= (Q-df/c)$).

والمخطط المعروض في شكل (1/4) يبين أن الدراسات داخل كل مجموعة فرعية إذا اشتركت في حجم الأثر العام، فإننا نستخدم نموذج الأثر الثابت لتحديد قيمة أوزان كل دراسة، وتكون قيمة T^2 مساوية للصفر. وفي إطار نموذج الآثار العشوائية نقدر قيمة T^2 من داخل المجموعات الفرعية كل على حدة واستخدام التقدير المنفصل بكل مجموعة فرعية subgroup. وفي حالة وجود عدد قليل من الدراسات داخل المجموعات الفرعية فإن التقديرات يمكن أن تكون غير دقيقة؛ ولذلك يفضل تجميع تلك التقديرات.

شكل (1/4)
مخطط يبين أسلوب انتقاء نموذج الأثر لتقدير قيمة حجم الأثر ودلالته لمجموعتين فرعيتين باستخدام نموذج الآثار العشوائية (المؤلفان)



في نموذج الآثار العشوائية سوف نحسب تقديرات منفصلة تقديرات حجم الأثر ودلالتها وكذلك T^2 بكل دراسة ومجموعة فرعية، ويبين جدول (5/4) حجم أثر وتباين كل دراسة من الدراسات العشر، وأثر وتباين المجموعتين الفرعيتين وكذلك الأثر المركب العام وتباينه (انظر Cumming, 2012).

جدول (5/4)

تقدير قيمة التقديرات المنفصلة لـ $\bar{y}$ و W و T^2 باستخدام نموذج الآثار العشوائية

الدراسات		حجم الأثر Y	التباين داخل V_Y	التباين بين T^2	التباين الكلى V	الوزن W	قيم مشتقة		
							W.y	$W.y^2$	W^2
	1	0.110	0.0100	0.0164	0.0264	37.846	4.163	0.458	1432.308
	2	0.224	0.0300	0.0164	0.0464	21.541	4.825	1.081	464.017
A	3	0.338	0.0200	0.0164	0.0364	27.455	9.280	3.137	753.788
	4	0.451	0.0150	0.0164	0.0314	31.824	14.353	6.473	1012.757
	5	0.480	0.0100	0.0164	0.0264	37.846	18.166	8.720	1432.308
مجموع A		0.325	0.0060			156.512	50.787	19.868	5095.179
	1	0.4440	0.0150	0.0022	0.0172	57.983	25.512	11.225	3362.002
	2	0.492	0.0200	0.0022	0.0222	44.951	22.116	10.881	2020.582
B	3	0.651	0.0150	0.0022	0.0172	57.983	37.747	24.573	3362.002
	4	0.710	0.0250	0.0022	0.0272	36.702	26.058	18.501	1347.034
	5	0.740	0.0120	0.0022	0.0142	70.193	51.943	38.438	4927.012
مجموع B		0.610	0.0040			267.811	163.376	103.619	15018.633
المجموع الكلي		0.505	0.0020			424.323	214.163	123.487	20113.812

1- تقدير متوسط حجم الأثر بالدراسات A في إطار نموذج الآثار العشوائية:

- حجم أثر المجموعة A $\quad \bar{Y}_A = \frac{50.787}{156.512} = 0.3245$
- تباين حجم الأثر $\bar{y}_A \quad V^*_{YA} = \frac{1}{156.512} = 0.0064$
- الخطأ المعياري لـ $\bar{y}_A \quad SE^*_{YA} = \sqrt{0.0064} = 0.0799$
- الحد الأدنى لفترة الثقة لـ $\bar{y}_A \quad LL^*_{YA} = 0.3245 - 1.96 \times 0.0799 = 0.1678$
- الحد الأعلى لفترة الثقة لـ $\bar{y}_A \quad UL^*_{YA} = 0.3245 + 1.96 \times 0.0799 = 0.4812$
- اختبار Z لدلالة حجم الأثر $\bar{Y}_A \quad Z_{YA} = \frac{0.3245}{0.0799} = 4.0595$
- قيمة P لتحديد دلالة $\bar{y}_A \quad P(Z^*_{YA}) < 0.0001$
- التباين بين مجموعة الدراسات A عند درجات حرية 4

$$Q^*_A = 19.868 = (\frac{50.787}{156.512})^2 = 3.3882$$

ويلاحظ أننا لتقدير قيمة Q_A^*استخدمنا أوزان الآثار العشوائية، وهو يعكس التشتت داخل وبين الدراسات، ويشتق منه اختبار مدى تجانس الدراسات داخل المجموعة الفرعية A وتقدير T^2.

2- تقدير حجوم الأثر والتباين بين الدراسات B في إطار الآثار العشوائية:

- حجم أثر المجموعة B $\overline{Y}_B^* = \frac{163.376}{156.512} = 0.6100$
- تباين حجم الأثر $\bar{y}_B$ $Y_{YB}^* = \frac{1}{267.811} = 0.0037$
- الخطأ المعياري لـ $\bar{y}_B$ $SE_{YB}^* = \sqrt{0.0037} = 0.0611$
- الحد الأدنى لفترة الثقة لـ $\bar{y}_B$ $LL_{YB}^* = 0.6100 - 1.96 \times 0.0611 = 0.4903$
- الحد الأعلى لفترة الثقة لـ $\bar{y}_B$ $UL_{YB}^* = 0.6100 + 1.96 \times 0.0611 = 0.7298$
- اختبار Z لدلالة $\overline{Y}_B$ $Z_{YB}^* = \frac{0.6100}{0.0611} = 9.9833$
- قيمة P لتحديد دلالة $\overline{Y}_B^*$ $P(Z_{YB}^*) < 0.0001$
- التباين بين مجموعة الدراسات B عند درجات حرية 4

$$Q_B^* = 103.619 = (\frac{163.376}{267.811})^2 = 3.9523$$

3- تقدير متوسط حجم الأثر والتباين بن الدراسات العشر معاً في إطار نموذج الآثار العشوائية:

- حجم الأثر للدراسات العشر معاً $Y_{AB}^* = \frac{214.163}{424.323} = 0.6100$
- تباين حجم الأثر للدراسات العشر معاً $V_{YAB}^* = \frac{1}{424.323} = 0.0037$
- الخطأ المعياري لحجم الأثر $SE_{YAB}^* = \sqrt{0.0024} = 0.0485$
- الحد الأدنى لفترة الثقة لـ $\bar{y}_{AB}$ $LL_{YAB}^* = 0.5047 - 1.96 \times 0.0485 = 0.496$
- الحد الأعلى لفترة الثقة لـ $\bar{y}_{AB}$ $UL_{YAB}^* = 0.5047 + 1.96 \times 0.0485 = 0.5999$
- اختبار Z لدلالة حجم الأثر الكلي $Z^* = \frac{0.5047}{0.0485} = 10.3967$
- قيمة P المناظرة لـ Z $P(Z^*) < 0.0001$
- التباين بين الدراسات العشر معاً $Q^* = 123.487 = (\frac{214.163}{424.323})^2 = 15.3952$

ويبين جدول (6/4) ملخصاً للقيم التي حسبناها سابقاً $\overline{Y}$، V ، $S_{e\bar{y}}$، Z، Q ، P(z)

جدول (6/4)

ملخص نتائج تحليل مقارنة حجمي الأثر في المجموعتين

الإحصائيات	Aالمجموعة	Bالمجموعة	ABالمجموعة الكلية
$\overline{Y}$	0.3245	0.6100	0.5047
V	0.0064	0.0037	0.0024
SE_y	0.0799	0.0611	0.0485
$LL_{\overline{y}}$	0.1678	0.4903	0.4096
$UL_{\overline{y}}$	0.4812	0.7298	0.5999
Z	4.0595	9.9833	10.3967
P^2	0.0000	0.0000	0.0000
Q	3.3882	3.9523	15.3952

1- المقارنة بين مجموعتين فرعيتين في متوسطي حجمي أثرهما باستخدام اختبار Z في إطار نموذج الآثار العشوائية (Riley, et al., 2011)

حيث إننا نريد مقارنة متوسطي حجمي الأثر في المجموعتين الفرعيتين A و B، علماً بأن متوسطي حجمي الأثر يساويان 0.325 و0.610 بانحرافين معياريين 0.006 و0.004 على الترتيب. وسوف نستخدم معادلات مماثلة لتلك التي استخدمناها سابقاً والتغيير هنا هو استخدامنا للرمز (*) للإشارة إلى أن الاحصاءات تعتمد على أوزان الآثار العشوائية وليس أوزان الآثار الثابتة. وكذلك الاعتماد على الفرض الصفري $\bar{y}_A = \bar{y}_B$ وليس $\Theta_A = \Theta_B$ ، التي نستخدمها للإشارة إلى القيم المشتركة عندما نتعامل مع نموذج الأثر الثابت.

وبافتراض أن Y_A^* وY_B^* هما الأثران المقدران للمجموعتين وتباينهما V_{YA}^* وV_{YB}^* فإذا استخدمنا D_{AB}^* للإشارة إلى الفرق بين متوسطي حجم الأثرين.

الفرق بين حجمي الأثر بالمجموعتين A وB $D_{YAB}^* = \overline{Y}_B - Y_A^* = 0.6100 - 0.3254 = 0.2856$

الخطأ المعياري لـ D_{AB} $\sqrt{V_{yA}^* + V_{yB}^*} Z_{eDAB}^* = = S_{eDAB}^* = \sqrt{0.0064 + 0.0037} = 0.1006$

وتكون معادلة الاختبار الإحصائي للمقارنة بين الأثرين هي:

- الفرق بين الأثرين Z اختبار $Z_{DAB}^* = D_{DAB}^* / S_{eDAB}^*$

$$Z_{DAB}^* = \frac{0.2856}{0.1006} = 2.8381$$

- مستوى دلالة Z $p^* = 2\ [\ 1 - \ (\phi(|Z_D^*|)]$

حيث $\phi|Z|$هي التوزيع المتجمع الاعتدالي المعياري والقيمة P المناظرة لـــــ Z=2.8381 هي0.0045، وهذا يبين وجود فرق دال إحصائياً بين الأثرين بالمجموعتين الفرعيتين A وB.

2- المقارنة بين مجموعتين فرعيتين A وB في اختبار Q المعتمد على تحليل التباين، في إطار نموذج الآثار العشوائية (Riley, et al., 2011)

إن التغيير الأساسي في هذه الحالة هو تطبيق الأوزان كما في نموذج الآثار العشوائية، وفي هذه الطريقة نحسب القيم التالية:

Q_A^*، وهي مجموع مربع الانحرافات الموزونة لكل دراسة من الدراسات الخمس بالمجموعة الفرعية A حول متوسطها المشترك.

Q_B^*، وهي مجموع مربع الانحرافات الموزونة لكل دراسة من الدراسات الخمس بالمجموعة الفرعية B حول متوسطها المشترك.

Q_W، وهي مجموع Q_A^* وQ_B^* ولعدة مجموعات يكون $Q_W^* = \sum_{J=1}^{K} Q_K^*$

Q_{Be}، وهي مجموع مربعات الانحرافات الموزونة لمتوسطي المجموعتين الفرعيتين حول المتوسط العام.

Q_T، وهي مجموع مربعات الانحرافات الموزونة لكل الآثار حول المتوسط العام grand.

$$Q_W^* = Q_A^* + Q_B^*$$

ودرجة حرية (df_B+df_A)

التباين الكلى الداخلي = التباين داخل المجموعة A + التباين داخل المجموعة B

لأي عدد من المجموعات الفرعية $Q^*_W = \sum_{J=1}^{K} Q_j^*$

ويكون التباين بين المجموعات $Q^*_B = Q^* - Q_W$

جدول (7/4)

قيم Q_A^*، Q_B^*، Q_W^*. Q_T^* في إطار نموذج الآثار العشوائية

مصادر التباين	Q*	Df	P
Q_A	3.3882	4	
Q_B	3.9523	4	
Within (Qw.)	7.3406	8	
Between ($Q_{Be.}$)	8.0547	1	0.0045
Total (Q)	15.3952	9	

نلاحظ أن Q_B^* تساوي 8.0547 عند درجات حرية 1 وهي دالة عند P = 0.0045، وهذا يعني أن حجم الأثر لصـــالح المجموعة الفرعية B؛ إذ إن حجم أثرها 0.61 في مقابل حجم أثر المجموعة A الذي يســـاوي 0.3245. وســـوف نطبق المعادلات التالية من خلال بيانات جدول (8/4):

جدول (8/4)

بيانات الفرق بين المجموعتين A، B في حجم الأثر

المجموعات الفرعية	حجم الأثر Y	التباين داخل V_y	التباين بين T^2	التباين الكلي V	الوزن W	قيم مشتقة		
						WY	WY^2	W^2
A	0.3245	0.0064	0.0000	0.0064	156.512	50.787	16.480	24495.944
B	0.6100	0.0037	0.0000	0.0037	267.811	163.376	99.666	71722.774
					424.323	214.163	116.146	9621.718

قيمة حجم الأثر $= 0.5047\overline{Y}^* \frac{214.163}{424.323}$

تباين حجم الأثر $\frac{1}{424.323} = 0.0024V_Y^* =$

التباين بين الدراسات $Q_{B.} = 116.146 - \frac{(214.163)^2}{424.323} = 8.0547$

$df = 2 - 1 = 1$

مستوى دلالة QB $P\ (Q = 8.0547, df = 1) = 0.0045$

تقدير قيمة الفروق وفترة الثقة كما يلي:

الفرق بين متوسطي حجمي الأثر $D_y^* = 0.6100 - 0.3245 = 0.2856$

الخطأ المعياري لـ $D\bar{y}$ $SE^*_{Dy} = \sqrt{0.0064 + 0.0037} = 0.1006$

الحد الأدنى لفترة الثقة $LL^*_D = 0.2856 - 1.96\ (0.1006) = 0.0883$

الحد الأعلى لفترة الثقة $UL^*_D = 0.2856 + 1.96\ (0.1006) = 0.4828$

$Z = .2856/31006 = 2.8389$

وهذا يعني أن الفرق دال أو حقيقي بين متوسطي حجمي الأثر في الدراستين A وB وأن هذا الفرق يحتمل أن يقع في المدى من 0.09إلى 0.48

تقدير التباين المجمع لـــ T^2 ($T_A^2 + T_B^2$) في إطار نموذج الآثار العشوائية (Riley, et al., 2011)

سوف نحسب متوسطي الآثار للمجموعات الفرعية، وذلك باستخدام نموذج الآثار العشوائية مع التقدير المجمع لـــ T^2، الذي يشار إليه بأنه التباين داخل المجموعات وهو يساوي $T_W^2 = T_A^2 + T_B^2$

التباين داخل المجموعة الفرعية $T^2 = \frac{Q-df}{C}$ حيث $C = \Sigma_{wi} - \frac{\Sigma wi^2}{\Sigma wi}$ و $df = k - 1$

حيث k هي عدد الدراسات.

ولأي عدد من المجموعات تتحدد قيمة T^2_W الكلية من خلال المعادلة التالية:

$$T^2_W = \left[\sum_{i=1}^{k} Q_i - \sum_{i=1}^{k} df_i\right] / \sum_{i=1}^{k} Ci$$

لاحظ أن هذه الطريقة لتقدير T^2_W يمكن أن تصل إلى قيمة سالبة (عندما يكون التشتت الملاحظ أقل مما نتوقع بالمصادفة) وفي هذه الحالة نضع تقدير T^2_W مساوياً للصفر.

تقدير قيمة تقديرات حجم الأثر للمجموعات الفرعية

يعرض جدول (9/4) تقدير حجم الأثر لكل دراسة للوصول إلى قيمة T^2_W

جدول (9/4)

تقدير حجم الأثر لكل دراسة والإحصاءات المشتقة منها

الدراسات		حجم الأثر y_i	التباين داخل V_y	التباين بين T^2	التباين الكلي V	الوزن W	قيم مشتقة WY	WY^2	W^2
	1	0.110	0.0100	0.0097	0.0197	50.697	5.577	0.163	2570.150
	2	0.224	0.0300	0.0097	0.0397	25.173	5.639	1.263	633.678
T^2_A	3	0.338	0.0200	0.0097	0.0297	33.642	11.371	3.843	1131.752
	4	0.451	0.0150	0.0097	0.0247	40.445	18.241	8.226	1635.767
	5	0.480	0.0100	0.0097	0.0197	50.697	24.334	11.681	2570.150
ΣA						200.652	65.161	25.627	8541.498
	1	0.440	0.0150	0.0097	0.0247	40.445	17.796	7.830	1635.767
	2	0.492	0.0200	0.0097	0.0297	33.642	16.552	8.143	1131.752
T^2_B	3	0.651	0.0150	0.0097	0.0247	40.445	26.329	17.140	1635.767
	4	0.710	0.0250	0.0097	0.0347	28.798	20.446	14.517	829.299
	5	0.740	0.0120	0.0097	0.0127	46.030	34.062	25.206	2118.721
ΣB						189.358	115.185	72.837	7351.306
المجموع الكلي						390.010	180.346	98.463	15892.804

وفيما يلي الإحصائيات اللازمة لتقدير قيمة التقدير المجمع لـ T^2(Q ,df)

المجموعة	Q	Df	C
A	8.4316	4	269.8413
B	4.5429	4	241.6667
المجموع	12.9745	8	511.5079

$$T^2_w = \frac{12.9745 - 8}{511.508} = 0.00974$$

حســـابات T^2 المجمع لدراســـات المجموعة الفرعية A (في حالة نموذج الآثار العشـــوائية)

(Riley, et al., 2011)

صيغ T^2 المجمع:

$T^2=(Q-df)/C$ $\quad\quad$ $df=(k-1)$

$C_i = \Sigma w_i - [\Sigma w_i^2 / \Sigma w_i]$

$T^2_W = \left[\sum_{i=1}^{k} Q_i - \sum_{i=1}^{k} df_i\right] / \sum_{i=1}^{k} Ci$

حجم أثر المجموعة الفرعية A $\quad\quad$ $Y^*_A = \frac{65.161}{200.652} = 0.3247$

تباين حجم أثر المجموعة الفرعية A $\quad\quad$ $V^*_{YA} = \frac{1}{200.652} = 0.0050$

الخطأ المعياري لحجم أثر المجموعة الفرعية A $\mathrm{Se}^*_{\mathrm{YA}} = \sqrt{0.0050} = 0.0706$

الحد الأدنى لفترة الثقة $\bar{y}_A$ $\mathrm{LL}^*_{\mathrm{YA}} = 0.3247 - (1.96)(0.0706) = 0.1864$

الحد الأعلى لفترة الثقة $\bar{y}_A$ $\mathrm{UL}^*_{\mathrm{YA}} = 0.3247 + (1.96)(0.0706) = 0.4631$

قيمة اختبار Z $\mathrm{Z}^*_{\mathrm{A}} = \frac{0.3247}{0.0706} = 4.6601$

قيمة P المناظرة (مستوى دلالة Z) $p(\mathrm{Z}^*_{\mathrm{A}}) < 0.0001$

قيمة Q_A $\mathrm{Q}^*_{\mathrm{A}} = 25.627 - (\frac{(65.161)^2}{200.652}) = 4.466$

لاحظ أن Q^* في إطار نموذج الآثار العشوائية يمكن أن تجزئ إلى مكونين، كذلك فإننا لم تعرض P الخاصة بـــ Q^*، بينما في إطار نموذج الأثر الثابت تكون Q مجزئة ونعرض لها P، وهي تعكس التباين بين الدراسات ويشتق منها اختبار مدى التجانس داخل المجموعة الفرعية A وتستخدم لتقدير قيمة $\mathrm{T}^2_{\mathrm{W}}$ المعادلة التالية:

$$\mathrm{T}^2_{\mathrm{W}} = \left[\sum_{i=1}^{k} \mathrm{Q}_i - \sum_{i=1}^{k} \mathrm{df}_i\right] / \sum_{i=1}^{k} \mathrm{Ci}$$

التقدير المجمع لـــ T^2 للمجموعة الفرعية B في إطار نموذج الآثار العشوائية (Riley, et al., 2011)

حجم أثر المجموعة الفرعية B $\mathrm{Y}^*_{\mathrm{B}} = \frac{115.185}{189.358} = 0.6083$

تباين حجم أثر المجموعة B $\mathrm{V}^*_{\mathrm{YB}} = \frac{1}{189.358} = 0.0053$

الخطأ المعياري لحجم أثر المجموعة B $\mathrm{SE}^*_{\mathrm{YB}} = \sqrt{0.0053} = 0.0727$

الحد الأدنى لفترة الثقة $\mathrm{LL}^*_{\mathrm{YB}} = 0.6083 - (1.96)(0.0706) = 0.4659$

الحد الأعلى لفترة الثقة $\mathrm{UL}^*_{\mathrm{YB}} = 0.6083 + (1.96)(0.727) = 0.7507$

اختبار Z لدلالة حجم أثر المجموعة B $\mathrm{Z}^*_{\mathrm{B}} = \frac{0.6083}{0.0727} = 8.3705$

قيمة P (دلالة $\bar{y}^*_B$) $p(\mathrm{Z}^*_{\mathrm{B}}) < 0.0001$

$$Q_B^* = 72.837 - \left(\frac{(115.185)^2}{189.358}\right) = 4.466$$ تباين مجموعة الدراسات B

تقدير قيمة التقدير المتجمع لـــــ $T^2_{ABW.}$ للدراســات العشــر في إطار نموذج الآثار العشــوائية (Riley, et al., 2011)

تحسب T^2 كما تم في حالتي المجموعتين الفرعيتين:

$$Y_{AB}^* = \frac{180.346}{390.010} = 0.4624$$

$$V_{YAB}^* = \frac{1}{390.010} = 0.0026$$

$$SE_{YAB}^* = \sqrt{0.0026} = 0.0506$$

$$LL_Y^* = 0.4624 - (1.96 \times 0.0506) = 0.3632$$

$$UL_Y^* = 0.4624 + (1.96 \times 0.0506) = 0.5617$$

$$Z_{yAB}^* = \frac{0.4624}{0.0506} = 9.1321$$

$$p(Z^*) < 0.0001$$

$$Q_{AB}^* = 98.463 - \left(\frac{180.346^2}{390.010}\right) = 15.0690$$

ويتم تلخيص القيم الإحصائية المحسوبة سابقاً للتقدير المتجمع لـ T^2 كما يلي:

القيم الإحصائية	المجموعة A	المجموعة B	المجموعة Combined
Y	0.3247	0.6083	0.4624
V	0.0050	0.0053	0.0026
SE_y	0.0706	0.0727	0.0506
LL_y	0.1864	0.4659	0.3632
UL_y	0.4631	0.7507	0.5617
Z	4.6001	8.3705	9.1321
P^2	0.0000	0.0000	0.0000
Q	4.4660	2.7706	15.0690

مقارنة حجمي الأثر بالمجموعتين الفرعيتين في إطار نموذج الآثار العشوائية (Riley, et al., 2011)

حجمـا الأثر بـالمجموعتين الفرعيتين A و B همـا 0.325 و0.608 على الترتيـب، وللمقارنـة بينهمـا توجد ثلاث طرق لعمل ذلك وهي متكافئة؛ حيث تحقق نفس القيمـة P. وهذه الطرق هي:

الطريقة الأولى: مقارنة المجموعتين الفرعيتين A و B باستخدام اختبار Z

يمكن استخدام اختبار Z لمقارنة متوسط أثر المجموعة الفرعية A بمتوسط أثر المجموعة الفرعية B في إطار نموذج الآثار العشوائية.

$$D^*_{YAB} = y^*_A - y^*_B$$

حيث y^*_A متوسط أثر المجموعة A و y^*_B متوسط أثر المجموعة B و D_{YAB} هو الفرق بين حجمي الأثر بالمجموعتين الفرعيتين A و B

معادلة اختبار دلالة الفرق بين الأثرين:

$$Z^*_{DyAB} = D^*_{yAB}/Se^*_{yAB}$$

والخطأ المعياري للفرق بين أثر المجموعتين الفرعيتين

$$S_{eDyAB} = \sqrt{V^*_{yA} + V^*_{yB}}$$

علماً بأن Z^* تتبع التوزيع الاعتدالي، وتعطى قيمة P لاختبار ذي طرفين بالمعادلة

$$p = 2\,|1 - \phi|\,Z*_D|$$

حيث $\phi\,|Z|$ هو التوزيع الاعتدالي المعياري، وفي مثالنا الحالي:

$$= 0.6083 - 0.3247 = 0.2835 D^*_{yAB}$$

علمًا بأن Z^* تتبع التوزيع الاعتدالي، وتعطى قيمة P لاختبار ذي طرفين بالمعادلة

$$p = 2\,|1 - \phi|Z^*_D|$$

حيث $\phi\,|Z|$ هو التوزيع المتجمع الاعتدالي المعياري. وفي مثالنا الحالي

$$= 0.6083 - 0.3247 = 0.2835 D^*_{yAB}$$

$$= 0.1013 SE^*_{DAB} = \sqrt{0.0050 + 0.0053}$$

$$Z^*_B = \frac{0.2835}{0.1013} = 2.7986$$

وp لاختبار ذي طرفين، التي تناظر القيمة 2.7986 تكون 0.0051 وهذا يشير إلى أن الأثر المتوسط يحتمل ألا يكون هو نفسه في مجموعتي الدراسات A و والدراسات B.

المقارنة بين المجموعتين الفرعيتين A وB في اختبار Q المعتمد على تحليل التباين

في هذه الحالة سوف نقوم بتقدير قيمة الكميات التالية:

1- Q_A^* = مجموع مربعات الانحرافات الموزونة لكل الدراسات A حول متوسطها.

2- Q_B^* = مجموع مربعات الانحرافات الموزونة لكل الدراسات B حول متوسطها.

3- $Q_{W.}^* = Q_B^* + Q_A^*$ مجموع الانحرافات الموزونة داخل المجموعتين الفرعيتين

4- $Q_{B.}^*$ = مجموع مربعات الانحرافات الموزونة لمتوسطات المجموعات الفرعية عن المتوسط الكلى.

5- Q= مجموع مربعات الانحرافات الموزونة لكل آثار الدراسات حول المتوسط الكلى.

وحيث إن $Q_{W.}^* = Q_A^* + Q_B^*$ لتمثيل مجموع مربعات الانحرافات الموزونة داخل المجموعتين الفرعيتين. وللتعميم فإنه لعدد K من المجموعات الفرعية نستخدم المعادلة:

$$Q_{W.}^* = \sum_{i=1}^{K} Q_i^*$$

وفي مثالنا الحالي $Q_{W.}^* = 4.4660 + 2.7706 = 7.2366$

ويمكن تقدير قيمة Q^* الكلية باستخدام المعادلة التالية:

$$Q^* = Q_{W.}^* + Q_B^*$$

$$\therefore\ Q_B^* = Q^* - Q_W$$

وفي مثالنا الحالي $Q_B^* = 15.0690 - 7.2366 = 7.8324$

علمًا بأن Q_B ودرجات الحرية (K) تتوزع وفق توزيع χ^2 ويلخص الجدول التالي قيم Q^* وتقدير T^2المجمع وذلك في إطار نموذج الآثار العشوائية:

المجموعات	Q^*	df	P
(QA)	4.4660	4	
(QB)	2.7706	4	
Within(Qw.)	7.2366	8	0.0050
Between(QB.)	7.8324	1	0.0051
Total(Q)	15.0690	9	

ومن الجدول السابق تبين أن قيمة Q بين المجموعات تساوي 7.8324 عند درجة حرية (1)، وهي دالة عند 0.0051، وهذا يعني وجود فروق دالة بين المجموعتين الفرعيتين A وB في حجم الأثر لصالح المجموعة الثانية.

مقارنة المجموعتين الفرعيتين A وB باســـتخدام اختبار Q للتباين (في إطار نموذج الآثار العشـــوائية)

(Riley, et al., 2011)

وفي هذه الطريقة نقوم باستخدام المعادلات السابقة لحساب القيم المتضمنة بالجدول التالي:

الدراسات	حجم الأثر Y	التباين داخل V_Y	التباين بين T^2	التباين الكلى V	الوزن W	قيم مشتقة Wy	Wy^2	W^2
A	0.3247	0.0050	0.0000	0.0050	200.652	65.161	21.161	40261.386
B	0.6083	0.0053	0.0000	0.0053	189.358	115.185	70.066	35856.405
					390.010	180.346	91.227	76117.791

- حجم الأثر الكلي $\overline{Y}^* = \frac{180.346}{390.010} = 0.4624$
- تباين حجم الأثر $V_{\overline{Y}}^* = \frac{1}{390.010} = 0.0026$
- التباين بين الدراسات $Q_{B.} = 91.227 - \left(\frac{180.346^2}{390.010}\right) = 7.8324$
- قيمة P المناظرة $P = (Q = 7.8324, df = 1) = 0.0051$

حيث Q تمثل مجموع المربعات الموزونة للدراسات A و B حول المتوسط الكلى. وحيث إن Q = 7.8324 و df =1 فإن P =. 0051 وهذا يشير إلى الدلالة الإحصائية لـ Q.

حساب قيمة ودلالات الفروق بين المجموعتين الفرعيتين في حجم أثرها:

- الفرق بين أثر المجموعتين $D_{yAB}^* = 0.6083 - 0.3247 = 0.2835$
- الخطأ المعياري للفرق بين حجمي أثر المجموعتين $SE_D = \sqrt{V_A + V_B}$

$$SE_{DyAB}^* = \sqrt{0.0050 + 0.0053} = 0.1013$$

- الحد الادنى لفترة الثقة $LL_D = 0.2835 - (1.96 \times 0.1013) = 0.0850$
- الحد الأعلى لفترة الثقة $UL_D = 0.2835 + (1.96 \times 0.1013) = 0.4821$

وهذا يعني أن الفرق بين أثرى المجموعتين الفرعيتين يراوح بين 0.09 و 0.48 بدرجة ثقة

متوسط حجم الأثر: في حالة وجود أكثر من ناتج للمعالجة (Borenstein, et al. 2009, Cumming, 2012)

بفرض أنه توجد خمس دراسات تتناول أثر طريقة تدريس معينة على تحصيل التلاميذ في الرياضيات والقراءة علماً بأن التلاميذ يتبعون التصميم نفسه وتم توزيعهم على المجموعات الخمس بطريقة عشوائية وأن هذين الناتجين يكونان على التلاميذ أنفسهم. يمكن أيضاً بكل دراسة من الدراسات الخمس أن ننظر إلى الناتج في البداية قبل عملية التدخل أو المعالجة والناتج الثاني بعد عملية التدخل أو المعالجة على التلاميذ أنفسهم، وتسمى هذه الحالة بمنهج المجموعة الواحدة أو المجموعتين المرتبطتين أو القياسات المتكررة (الفرق بين القياسين القبلي والبعدي).

الهدف من هذا التحليل هو تحديد متوسط أثر المعالجة (طريقة التدريس) على القراءة والرياضيات بشكل منفصل باستخدام ما وراء التحليل، أو دمج الناتجين معاً في التحليل نفسه. وقد نهدف إلى تحديد الفرق بين حجمي الأثر في الرياضيات والقراءة علماً بأنهما مرتبطان؛ إذ إنهما أجريا على النلاميذ أنفسهم، ويسمى في الحالة الأولى والثانية الأثر المركب بكل دراسة.

ويتضمن الجدول (10/4) نتائج ما وراء التحليل لخمس دراسات ومتوسط الأثر بها ووزن تلك الدراسات وكذلك التباين والخطأ المعياري لمتوسط تقديرات حجم الأثر.

جدول (10/4)

نتائج ما وراء التحليل لخمس دراسات تمت على التلاميذ أنفسهم في الرياضيات والقراءة

الدراسة	الناتج	أثر الدراسة y_i	تباين أثر الدراسة	متوسط أثر الدراسة y	الارتباط (r)	تباين متوسط الأثر	وزن الدراسة w_i	ES*WT
(1)	القراءة	0.300	0.050	0.200	0.500	0.038	26.667	5.333
	الرياضيات	0.100	0.050					
(2)	القراءة	0.200	0.020	0.150	0.600	0.016	62.500	90375
	الرياضيات	0.100	0.020					
(3)	القراءة	0.400	0.050	0.300	0.600	0.040	25.000	7.500
	الرياضيات	0.200	0.050					
(4)	القراءة	0.200	0.010	0.150	0.400	0.007	142.857	21.429
	الرياضيات	0.100	0.010					
(5)	القراءة	0.400	0.060	0.350	0.800	0.054	18.519	6.481
	الرياضيات	0.300	0.06					
المجموع							275.542	50.118

يحسب متوسط حجم الأثر داخل الدراسة للناتجين: القراءة ($\bar{y}_1$) والرياضيات ($\bar{y}_2$) كما يلي: $\bar{y} = (\bar{y}_1 + \bar{y}_2)/2$

ويتحدد تباين متوسط هذا الأثر باستخدام المعادلة:

$$V_{\bar{y}} = \frac{1}{4}\left[V_{\bar{y}1} + V_{\bar{y}2} + 2r\sqrt{(V_{\bar{y}1})(V_{\bar{y}2})}\right]$$

$$= \frac{1}{4}[2v + 2r(v)] = \frac{1}{2}(v)(r+1)$$

حيث:

$\bar{y}_1$ = متوسط أثر القراءة بالدراسة الأولى.
$\bar{y}_2$ = متوسط أثر الرياضيات بالدراسة الأولى.
$V_{\bar{y}1}$ = تباين متوسط أثر القراءة بالدراسة الأولى.
$V_{\bar{y}2}$ = تباين متوسط أثر الرياضيات بالدراسة الأولى.
r = معامل الارتباط بين درجات y_1 وy_2

ويلاحظ أنه إذا كان الارتباط تام بينها (r=1) فإن تباين مجموعهما يعطي ضعف مجموع التباين، وعندما لا يوجد ارتباط بينهما (r=0) فإن تباين مجموعهما يساوي مجموع التباينين منفردين، ويتحدد تباين الفرق بين ناتجين مرتبطين باستخدام المعادلة:

$$V(\bar{y}_1 - \bar{y}_2) = \frac{1}{4}\left[V_1 + V_2 - 2r\sqrt{V_1(V_2)}\right]$$

وعندما $V_1 = V_2$ $\quad V(\bar{y}_1 - \bar{y}_2) = 2(V)(1-r)$

وفي حالة r=0 $\quad V(\bar{y}_1 - \bar{y}_2) = (V_1 + V_2)/4$

وفي حالة r=1 $\quad V(\bar{y}_1 - \bar{y}_2) = \frac{1}{4}\left[V_1 + V_2 - 2r\sqrt{V_1(V_2)}\right]$

من جدول (8/4) يتبين ما يلي:

متوسط حجم الأثر بالدراسة الأولى $\bar{y}_1 = (y_1 + y_2)/2$

$=(0.3+0.1)/2=0.2$

وتباين حجم الأثر بالدراسة الأولى $V\bar{y}_1 = \frac{1}{4}\left[0.05 + 0.05 + 2(0.50)\sqrt{0.05(0.05)}\right]$

$=0.0375$

وحيث إن $V_1 = V_2$ $\quad V\bar{y}_1 = \frac{1}{2}(0.05)(1+0.50)=0.0375$

وقد تم تقدير القيم الإحصـائية بجدول (9/4) من خلال بعض المعادلات السـابقة والحالية، مثال ذلك أن الوزن هو معكوس التباين، وعمل ما وراء التحليل من خلال القيم المركبة الخمس، وقد استخدم نموذج الأثر الثابت في تقدير قيمة متوسط تقديرات حجم الأثر وغيرها من الإحصاءات.

في حالة جمع الناتجين:

$$\bar{y} = 50.118/275.542 = 0.1819$$

$$V_{\bar{y}} = 1/275.542 = 0.0036$$

$$S_{e\bar{y}} = \sqrt{0.0036} = 0.060$$

وفترة الثقة بنسـبة 95 % لمتوسـط حجم الأثر المركب هي [0.300 , 0.064]، ولتحديد دلالة متوسط حجم الأثر المرحب فإن:

$$Z = 0.1819/\sqrt{0.0036} = 3.03$$

وهذا يعني أنه دال عند مستوى 0.01

تناول أكثر من ناتجين بالدراسة الواحدة

إذا كان عدد نواتج الدراسة الواحدة m، فإننا نحسب حجم الأثر المركب composite من خلال المعادلة: $\bar{y} = \frac{1}{m}\sum_{i=1}^{m} y_i$

وتباين حجم الأثر المركب: $V_{\bar{y}} = \left(\frac{1}{m}\right)^2 \left[\sum_{j=1}^{m} V_1 + \sum r_{ij} \sqrt{V_j}\sqrt{V_k}\right]$

وحيث المتغير j، الدراسة i، r_{ij} هو الارتباط بين كل زوج من النواتج.

وفي حالة تساوي التباينات بدراسة ما وراء التحليل:

$$V_{\bar{y}} = \left(\frac{1}{m}\right)(V)[1 + (m-1)r]$$

ويلاحظ أن الزيادة في m أو r أو الإثنين معاً يؤدي إلى ارتفاع التباين مقارنة بمعالجة الفرق بين النواتج واستقلالهم.

في حالة دراسة الفرق بين نواتج المعالجة

في المثال السابق يمكن تقدير قيمة الفرق بين القراءة والحساب باستخدام المعادلة التالية:

في الدراسة الواحدة:

$$\bar{y}_D = y_1 - y_2$$

$$V_{D\bar{y}} = V_{y_1} + V_{y_2} - 2r\sqrt{V_{y_1}.V_{y_2}}$$

وإذا كان التباينان متساويين (V)، فإنه يمكن تبسيط المعادلة السابقة.

$$V_{D\bar{y}} = 2V(1 - r)$$

وفي حالة r=0.5 فإن تباين الفرق يساوي $V_{\bar{y}} = V$

وفي حالة r=1 فإن تباين الفرق يساوي صفراً.

وفي حالة عدم وجود ارتباط $V_{\bar{y}D} = 2V$

فمثلاً في الدراسة الأولى الفرق بين الرياضيات والقراءة هو

$$\bar{y}_{D_1} = 0.300 - 0.100$$

تباين الفرق بين أثر الناتجين بالدراسة الأولى

$$V_{\bar{y}_1} = 0.05 + 0.05 - 2(0.50)\sqrt{0.05(0.05)}$$

$$= 2(0.05)(1 - 0.50) = 0.05$$

وقد استخدمت المعادلات السابقة للوصول إلى الإحصاءات الأتية في إطار نموذج الأثر الثابت:

الدراسة	الناتج	ES	V	ES	r	$V_{\bar{y}_1}$	W	ES*W
(1)	القراءة الرياضيات	0.300 0.100	0.050 0.050	0.200	0.500	0.050	20.000	4.000
(2)	القراءة الرياضيات	0.200 0.100	0.020 0.020	0.100	0.600	0.016	62.500	6.250
(3)	القراءة الرياضيات	0.400 0.200	0.050 0.050	0.200	0.600	0.040	25.000	5.000
(4)	القراءة الرياضيات	0.200 0.100	0.010 0.010	0.100	0.400	0.012	83.333	8.333
(5)	القراءة الرياضيات	0.400 0.300	0.060 0.060	0.100	0.800	0.024	41.667	4.167
Sum							232.50	27750

من الجدول السابق (10/4) يتبين لنا حساب الإحصاءات التالية:

متوسط الفرق بين حجم أثر الفروق بين الناتجين

$$\bar{y} = \frac{27.75}{232.5} = 0.1194$$

متوسط الفرق بيم تباين حجمي أثر الرياضيات والقراءة

$$V_{\bar{y}} = \frac{1}{232.5} = 0.0043$$

والخطأ المعياري له

$$SE_{\bar{y}} = \sqrt{0.0043} = 0.066$$

$$\therefore Z = 0.1194/0.0660 = 1.809$$

وهذا يعني أنه لا يوجد فرق دال بين الناتجين.

نسبة التباين المفسر بالمجموعات الفرعية

في الدراسات الارتباطية نستخدم R^2 لتحديد نسبة التباين المفسر بالنسبة للتباين الكلي من خلال المتغير المستقل باستخدام المعادلة التالية:

$$R^2 = \frac{\text{التباين المفسر}}{\text{التباين الكلى}} = \frac{\sigma^2 exp}{\sigma^2 tot}$$

$$R^2 = 1 - \frac{\sigma^2 unexp}{\sigma^2 tot} = 1 - \frac{\text{التباين غير المفسر}}{\text{التباين الكلي}}$$

والبسط في المعادلة السابقة هو التباين بين الدراسات بكل مجموعة فرعية والمقام هو التباين الكلي بين الدراسات وهو يساوى التباين داخل المجموعات الفرعية مضافاً إليه التباين بين المجموعات الفرعية. والتباين بين الدراسات داخل المجموعات الفرعية يساوي T_W^2 أو التباين غير المفسر في المعادلة، أما التباين الكلي؛ فإنه يساوي التباين داخل المجموعات الفرعية بالإضافة إلى التباين بين المجموعات الفرعية. وفيما يلي بيانات المثال في إطار نموذج الأثر الثابت.

الإحصائيات	A المجموعة	B المجموعة	المجموعة الكلية
Y	0.3241	0.6111	0.4581
V	0.0029	0.0033	0.0015
SE_Y	0.0535	0.0571	0.0390
LL_Y	0.2193	0.4992	0.3816
UL_Y	0.4289	0.7230	0.5346
Z	6.0633	10.7013	11.7396
P^2	0.0000	0.0000	0.0000
Q	8.4316	4.5429	26.4371
df	4.0000	4.0000	9.0000
p value	0.0770	0.3375	0.0017
Numerator	4.4316	0.5429	17.4371
C	269.8413	241.6667	583.5871
T^2	0.0164	0.0022	0.0299
I^2	52.5594	11.9506	65.9569

من النتائج السابقة نجد أن T^2 لكل فئات الدراسة يساوي 0.0299 و T^2 داخل المجموعات يساوي 0.0097 ، ومنها يمكن تقدير قيمة R^2

$$R^2 = 1 - \frac{0.0097}{0.0299} = 0.6745$$

هذا عند تقدير قيمة R^2 في إطار نموذج الآثار العشوائية؛ أما في حالة نموذج الأثر الثابت؛ فإن التباين بين الدراسات يعطى القيمة صفراً؛ لأنه يفترض أن قيمة T^2 متساوية بكل المجموعات الفرعية.

المبحث الخامس

ما وراء التحليل: مفهومه وأهدافه وأهميته وخطواته

إن البحوث التي تُجرى حول موضوع واحد قد تتناقض نتائجها بعضها مع بعض، ولعل أكثر من يعاني من هذه المشكلة هم المسؤولون عن وضع السياسات واتخاذ القرارات، حين يريدون الاستناد إلى نتائج هذه البحوث؛ حيث يجدون أنفسهم حائرين أمام الكثير من النتائج المختلفة. من هنا ظهرت منذ وقت مبكر الحاجة إلى البحوث التكاملية، وهي جهود يبذلها فريق من الباحثين بهدف إحداث التكامل بين نتائج الدراسات المنفصلة والوصول من ذلك إلى استنتاجات تستوعبها جميعاً بشكل كلى (أبو حطب وصادق، 1991). وأحد مداخل تلك البحوث التكاملية هو ما وراء التحليل (التحليل البعدي) –Meta Analysis، وهو أسلوب إحصائي يطبق على نتائج البحوث والدراسات الكمية بهدف إحداث تكامل بين نتائجها. ولا يختلف ما وراء التحليل عن غيره من طرق ومناهج البحث من حيث تحديد المشكلة وصياغة الفروض وتحديد وقياس المتغيرات واختيار عينة البحوث وتحليل نتائج بيانات هذه البحوث؛ وصولاً إلى نتائج تحتاج إلى المناقشة والتفسير، وهو بهذه المواصفات منهج إمبريقي كامل قابل للاستعادة والتكرار (أبو حطب وصادق، 1991).

مفهوم ما وراء التحليل Meta-Analysis

ما وراء التحليل هو التحليل الإحصائي لمجموعة كبيرة من نتائج بحوث منفصلة حول موضوع ما والغرض منه هو التكامل بين النتائج المستمدة من مجتمع كبير، عن طريق عدة بحوث أجريت على عينات مختلفة مشتقة من نفس المجتمع، ولذلك فإن ما وراء التحليل يعمل على تكامل نتائج دراسات كثيرة أجريت حول موضوع ما (Glass et al., 1981). ويرى جلاس أنه أسلوب تحليل إحصائي لنتائج مجموعة من البحوث في مجال ما، بهدف استقراء تعميمات مفيدة من النتائج التي توصلت إليها هذه البحوث (Glass , 1986). كما يعرف ما وراء التحليل بأنه منهج بحث إحصائي يكامل النتائج المستخلصة من دراسات تجريبية أو علاجية متعددة ومستقلة، يتم تحليلها للوصول إلى أثر إجمالي أو مُرَكب للمعالجة أو التدخل (Huque, 1988). كما أن العملية التي يتبعها ما وراء التحليل هي تحويل حجوم الأثر من عدة دراسات قابلة للمقارنة فيما بينها على مقياس مشترك، للوصول لنتيجة شاملة بشأن نتائج تلك الدراسات (Hattie, 2009; Hunter & Schmidt, 2004; Sánchez-Meca & Marín-Martines , 2011).

ويشار إلى ما وراء التحليل على أنه مجموعة من الإجراءات الإحصائية تم تصميمها لتجميع النتائج التجريبية والارتباطية خلال الدراسات الفردية المستقلة، التي لها التساؤلات نفسها تقريباً،

ويستخدم " ما وراء التحليل " الإحصائيات الملائمة من الدراسات الفردية كمنطلق للبيانات. فالافتراض الأساسي لهذا التحليل هو أن كل دراسة تمدنا بتقدير للعلاقة داخل العينة، وبتجميع النتائج عبر الدراسات، يمكن تمثيل العلاقات بصورة أكثر وضوحاً من التي نحصل عليها من تقديرات الدراسة الفردية (Lyons, 1997). ويضيف سكران (2006) أن ما وراء التحليل لم يعد مجرد تطبيق لمجموعة من الإجراءات الإحصائية، وإنما يتخطى ذلك بكثير؛، حيث يتم عمل مسح للدراسات المعنية بالظاهرة محل الدراسة، وتفحص الإطار النظري لتلك الدراسات، وكذلك مشكلة البحث والفروض وإجراءات الدراسة والنتائج ثم وضع المعايير والضوابط لاستخلاص الدراسات التي تخضع نتائجها لإعادة التحليل واتخاذ القرارات المناسبة.

ويشير كلارك وكارتر (Clark & Carter, 1997) إلى أن ما وراء التحليل تعتبر الموازي الكمي لمراجعة البحوث التقليدية؛ إلا أن ما وراء التحليل يسمح للباحث بتكميم الاتجاهات الموجودة في البحوث السابقة عن طريق الجمع بين تقديرات حجم الأثر، حيث أن ما يقدمه ما وراء التحليل بالجمع بين نتائج الدراسات السابقة يعمل على زيادة قوة الاختبار الإحصائي مما قد يترتب عليه أن جمع الدراسات غير الدالة إحصائيا يمكن أن يُظهر دلالتها الإحصائية. كما أن عملية تجهيز نتائج البحوث السابقة؛ لإجراء ما وراء التحليل تجبر الباحث على قراءة البحوث بتعمق وعناية شديدة، وبذلك فإن ما وراء التحليل يتفوق على المراجعة التقليدية للبحوث (أبوعلام، 2007).

ويتطلب ما وراء التحليل نوعاً ما من التكويد الكمي (Quantitative Coding) لنتائج الدراسات السابقة والعينات التي أجريت عليها، هذا بالإضافة إلى خصائص تلك الدراسات؛ من حيث تاريخ النشر ومصدره وخصائص المفحوصين وطبيعة المعالجة وتصميم الدراسات وطرق القياس، وغير ذلك من الخصائص التي يراها الباحث ضرورية للبحث المراد القيام به. وفي بعض الحالات يكون من الضروري تحويل نتائج الدراسات السابقة إلى نظام قياسي ومعياري مشترك حتى يمكن التعامل معها إحصائياً والتوصل إلى نتائج دقيقة غير مضللة. وتمتاز طريقة التحليل الإحصائي في ما وراء التحليل بأنها محاولة علمية لتكميم الاتجاهات التي أسفرت عنها البحوث السابقة من خلال النظرة الإجمالية لتقديرات حجم الأثر التي يمكن التوصل إليها من معالجة نتائج هذه الدراسات السابقة، حيث تزاد قوة الاختبار الإحصائي بالجمع بين نتائج الدراسات السابقة (Cumming, 2012).

أهداف ما وراء التحليل

يرى كل من أبو علام (2004) والشربيني (2011) أن الهدف مما وراء التحليل هو التوصل إلى وصف كمي دقيق غير متحيز لنتائج مجموعة من الدراسات والبحوث حول موضوع معين، وتحديد

مدى اتساق تلك النتائج؛ حيث يهتم هذا النوع من التحليل بتقويم دقيق للدراسات التي نشرت بالفعل من خلال تناول منظم متكامل لنتائج البحوث السابق نشرها في ضوء تحديد المشكلة، وتلخيص بعد تمحيص نتائج تلك البحوث السابقة. ويشير هايدو-ميدينا وآخرون (Huedo-Medina, et al., 2006, 193) إلى ثلاثة أهداف لما وراء التحليل هي: تحديد المتغيرات الوسيطة (المؤثرة) من النتائج المتوفرة في الدراسات المختلفة، واختبار مدى تجانس تقديرات حجم آثار الدراسات المتضمنة في التحليل، والحصول على مؤشر عام حول حجم أثر المتغيرات الوسيطة في دراسات في ما وراء التحليل، ضمن فترة ثقة محددة ومعرفة دلالتها الإحصائية.

أهمية ما وراء التحليل

غالباً ما توصف الدراسات المفردة بكونها "متحيزة"، ويأتي التحيز من مصادر عدة، منها انتقائية الباحث للبيانات والدراسات السابقة التي تدعم توجهه المعرفي، أو ذوقه الشخصي، أو من الاستنتاجات التي تدعم آراءه المسبقة. كما قد تتعارض نتائج الدراسات؛ نظراً لتنوع المناهج البحثية التي استخدمتها، والأدوات البحثية التي طبقتها، وطرق المعاينة، واختلاف البيئات الثقافية التي تمت خلالها الدراسة؛ وتباين خلفيات وأمزجة الباحثين الذين قاموا بها؛ هذا فضلاً عن التباين في مداخل فهم المشكلات البحثية (الخضر، 2022)، ومن هنا تبرز أهمية منهج ما وراء التحليل؛ حيث يسعى إلى التقليل من كمية البيانات عن طريق تلخيصها من مصادر متعددة (شريف، 1993)، والمساعدة على تخطيط البحوث، ووضع إطار للإرشادات، كما يساعد في الاستخدام الفعال للبيانات الموجودة، وضمان التعميم، والمساعدة على التحقق من اتساق العلاقات، وتوضيح عدم تناسق البيانات، باستخدام أساليب واضحة (Gopalakrishnan & Ganeshkumar, 2013).

إن ما وراء التحليل يوفر وسيلة لتعميم النتائج، ويصف مجموعة من الإجراءات من خلال المراجعة المنهجية للبحث الذي يدرس أثراً معيناً، ويجمع بين نتائج الدراسات المستقلة لتقدير حجم الأثر، ودمجها في دراسة واحدة (Ellis , 2010). كما يستخدم للكشف عن المعنى الكامن وراء البحوث والأدبيات؛ إذ يمكن الباحثين من تحديد الفجوات في هذه الأدبيات، التي لا تنعكس المعرفة التراكمية فيها. ويوجه الباحثين إلى الاحتياجات البحثية في مجال التخصص، والموضوعات التي يجب التركيز عليها في البحوث المستقبلية (Hunter & Schmidt, 2004). كما أنه يمكن أن يُستخدم لفحص مصداقية النظريات المطروحة في ضوء ما يتم التوصل إليه من نتائج باستخدام عينات مختلفة (Carlson et al., 1990). إن أسلوب ما وراء التحليل يدمج نتائج الدراسات للكشف عن أبسط أشكال العلاقات التي تكمن وراء الأدبيات البحثية؛ ومن ثم يوفر أساساً لتطوير النظريات، وهو بذلك يحقق أهم أهداف العلم

وهو التراكمية، ويعني ذلك تطوير النظريات التي تشرح وتفسر العمليات التي تكمن وراء الظاهرة، واكتشاف العلاقات بين المتغيرات، التي تعد بناء النظرية وأساسها.

ويرى ديكوستر (Decoster, 2009) أن ما وراء التحليل ليس مجرد ملخص للأدب لكن يقدم التفسير والتكامل النظري بشكل عام، كما يقدم المساهمة العملية ذات القيمة الحقيقية أكثر من كونه يقدم الحسابات الإحصائية للوصول إلى حجم الأثر الفعلي في الدراسات والبحوث في موضوع معين. ويرى ميلبى-ليرفاج هولم (Melby-Lervag & Hulme, 2012) أن أهمية ما وراء التحليل تكمن في معالجة الاختلافات بين الدراسات بشكل علمي ممنهج؛ مما يسهم في تقديم نتائج أكثر دقة من نتائج الدراسات الفردية.

مما سبق يمكن للمؤلفين تحديد أهمية ما وراء التحليل في النقاط التالية:

1- يزيد ما وراء التحليل من الموضوعية في استعراض الأدبيات في موضوع معين؛ إذ يوفر معلومات كمية عن نتائج مختلفة لدراسات مختلفة، وهذا يساعد على التغلب على الطريقة الأكثر ذاتية أحادية النظر في استعراض الأدبيات.
2- يمكننا ما وراء التحليل من التعرف بدقة على جوانب تمثل فجوات بحثية في الأدبيات الخاصة بموضوع معين وتحتاج إلى إجراء أبحاث فيها، وهذا يقودنا إلى أهداف بحثية واضحة جديدة.
3- يمدنا ما وراء التحليل بتقدير أكثر دقة لبارامترات المجتمع، وهذا يعني أنه ييسر تعميم النتائج التي حصلنا عليها من البحوث الفردية في موضوع معين؛ حيث يمكن أن تعمم النتائج على مجتمع أكبر.
4- يمكن من خلال ما وراء التحليل التوصل إلى تحديد المتغيرات الوسيطة والمعدلة التي تشكل وتفسر العلاقة بين المتغيرات المستقلة والتابعة.
5- يزيد استخدام ما وراء التحليل من القوة الإحصائية والدقة في استكشاف الأثر المجمع أو المركب؛ حيث تزداد قوة الاختبار الإحصائي بالجمع بين نتائج الدراسات السابقة. وقد يترتب على ذلك اكتشاف أن الجمع بين عدد من الدراسات ذات النتائج غير الدالة التي ظهرت في الاتجاه نفسه، يمكن أن تظهر دلالتها عند الجمع بينها.
6- يمكن من خلال ما وراء التحليل قياس التباين بين تقديرات حجم أثر الدراسات المجمعة.

خطوات ما وراء التحليل

حدد جلاس وزملاؤه (Glass et al., 1981) عدة خطوات لإجراء ما وراء التحليل تتضمن ما يلي:

1- تحديد موضوع الاهتمام: تحديد الموضوع المراد إجراء تكامل لنتائج الدراسات المتعلقة به بشكل دقيق.

2- تجميع الدراسات ذات العلاقة: تجميع الدراسات السابقة ذات العلاقة ببؤرة الاهتمام التي تم تحديدها في الخطوة الأولى.

3- وضع معايير إقصاء وتضمين الدراسات: وضع معايير لتضمين الدراسات أو إقصائها؛ مثل سنة النشر، حجم العينة، التصميم البحثي، نوع المعالجة المستخدمة، ونوع المتغيرات، وجودة الدراسة وغيرها من المعايير التي تحدد تضمين الدراسة في التحليل من عدمه.

4- فحص الدراسات السابقة: يقوم الباحث بفحص محتوى الدراسات السابقة التي جمعت للتأكد من ارتباطها ببؤرة الاهتمام التي تم تحديدها، وتوافقها مع معايير التضمين المحددة مسبقاً.

5- توصيف وتصنيف خصائص الدراسات وتشفيرها: القيام بتكميم خصائص الدراسات الفردية وترميزها.

6- تقدير قيمة تقديرات حجم الأثر: ويتم ذلك من خلال اختيار الطريقة المناسبة لإدخال البيانات في البرنامج الإحصائي الخاص وتقدير قيمة حجم الأثر اعتماداً على نوع المعالجة في الدراسات الأولية.

وحدد ديورلاك ولايبزى (Durlak & Lipsey, 1991) في كتابهما " دليل استخدام ما وراء التحليل " ست خطوات أساسية ضرورية لما وراء التحليل، وهي:

1 – صياغة أسئلة البحث

أ- هل صيغت أسئلة البحث والفروض بشكل واضح؟

ب- هل تم تحديد المتغيرات الرئيسية للدراسة بشكل منطقي؟

ج- هل تم تحديد الأدبيات بشكل كامل؟ وهل يضمن هذا التحديد أهم الأدبيات المتعلقة بالموضوع؟

د- هل تم تحديد المعايير الضمنية وغير الضمنية؟ وهل هذه المعايير معقولة؟

2 – تحليل الدراسات السابقة

أ- هل تم اختيار عينة ممثلة غير متحيزة من الدراسات السابقة؟

ب- هل تم اختيار عينة من الدراسات غير المنشورة، لتجنب التحيز للبحوث المنشورة؟

ج- هل تم الإشارة للدراسات ذات الصلة التي لم يتم تضمينها؟

د- هل تم اختيار الدراسات ذات الصلة بالدراسة بشكل موضوعي؟

3 – إجراءات الترميز (التكويد)

أ- هل تم وصف المشكلات خلال إجراءات الترميز؟

ب- هل نظام الترميز متاح عند طلبه؟ وهل يتضمن هذا النظام بعض المعايير لترميز الجوانب الغامضة؟

4 – مؤشر حجم الأثر

أ- هل تم وصف كل طرق تقدير قيمة حجم الأثر؟

ب- هل إجراءات التعامل مع النتائج غير الدالة إحصائياً واضحة؟ وهل تم الإشارة إلى تلك النتائج؟

ج- هل مدلولات النتائج غير الدالة إحصائياً مرتبطة بالاستنتاجات وتعميمات ما وراء التحليل؟

5 – التحليلات الإحصائية

أ- هل تم ضبط تأثير صغر أحجام بعض العينات؟

ب- هل تم ضبط تأثير النتائج المتطرفة؟

ج- هل تم استخدام وحدة التحليل المناسبة؟

د- هل تم استخدام الإجراءات المناسبة في أثناء التحليل؟

ه- هل تم استخدام متغيرات معدلة لتفسير الفروق في تقديرات حجم الأثر بين الدراسات؟

و- هل كان هناك انتباه كاف للخصائص المنهجية؟

ز- هل خطة القيام بما وراء التحليل حددت النموذج الإحصائي الصحيح المناسب للبيانات؟

6 – الاستنتاجات والتفسير

أ- هل يربط ما وراء التحليل قوة الأثر بالنتائج الإحصائية؟

ب- هل النتائج مرتبطة بالأدبيات التي تمت مراجعتها؟

وقد حدد ليونز-رث (Lyons-Ruth, 1997) الخطوات التي يجب أن تراعى عند القيام بما وراء التحليل، وهي:

1- تعريف مجال البحث: بوساطة المتغيرات المستقلة أو المتغيرات البحثية أو بوساطة أسباب المتغيرات المهمة ونتائجها.

2- وضع محكات لاختيار الدراسات التي يتم مراجعتها: مثل الدراسات المنشورة في مقابل غير المنشورة، الوقت المستغرق للدراسة، تعريف المتغيرات، جودة الدراسة.

3- تحديد نوع حجم الأثر؛ مثل: حجم الأثر " d " لكوهين.

4- البحث عن الدراسات ذات العلاقة من خلال قواعد البيانات المناسبة.

5- اختيار المجموعة النهائية من الدراسات.

6- تحديد البيانات المستخرجة من المتغيرات ذات الاهتمام، حجوم العينات، تقديرات حجم الأثر، ثبات المقاييس، والخصائص الأخرى لكل دراسة.

7- إعطاء رمز لكل خاصية من الخصائص المرتبطة بحجم الأثر لكل دراسة؛ مثل: خصائص العينة، نوع المتغيرات التابعة.

8- عمل الثبات لإجراءات الترميز باستخدام (من 1 إلى 3) أشخاص آخرين يقومون بعملية الترميز لتحديد ثبات الترميز؛ ومن ثم الموضوعية.

9- عندما توجد مقاييس متعددة لتقديرات حجم الأثر، يحدد المؤشر الذي سوف تستخدمه، ويرجع ذلك إلى التنوع النظري والتنوع القياسي بين المتغيرات.

10- تحديد المتوسطات وتباين التقديرات لحجم الأثر داخل الدراسات والمجموعات الفرعية.

11- تحديد متوسط حجم الأثر الموزون، وكذلك Q لاختبار التجانس أو التباين بين الدراسات في تقديرات حجم الأثر لتقدير مصادر التباين بين الدراسات نتيجة (خطأ المعاينة، الضعف الناتج عن خطأ القياس، مدى القيود وتقدير التباين الحقيقي بين الدراسات، تقدير حجم الأثر الحقيقي المصحح من خطأ العينة، ومن خطأ القياس ومدى القيود.

1- مناقشة نتائج الدراسة في ضوء الإطار النظري وما توصلت إليه معظم الدراسات السابقة.

مما سبق يمكن استخلاص الخطوات الآتية في منهج ما وراء التحليل:

1 – تحديد موضوع الاهتمام ومشكلة الدراسة، وتتطلب الإجراءات التالية:

أ- أن يكون موضوع الدراسة جذاباً وممتعاً للباحث وللآخرين.

ب- شرح أساسي للحاجات التي دعت إلى إجراء ما وراء التحليل في موضوعك (مبررات البحث).

ج- وصف للأدب الذي تريد تحليله بصفة عامة.

د- توافر عدد كافٍ من الدراسات في الموضوع.

ه- تجميع البحوث والدراسات السابقة ذات العلاقة بموضوع ما وراء التحليل؛ بحيث يكون عدد هذه الدراسات مناسباً وممثلاً لمجتمع تلك الدراسات.

و- تحديد المتغيرات المستقلة أو المنبئة، وكذلك المتغيرات التابعة أو المتنبأ بها.

ز- الحصول على نسخ من تقارير البحوث الفردية التي تحتوي على التحليل الإحصائي للبحث.

ح- تحديد السنوات التي تغطيها بوساطة قاعدة البيانات وغيرها.

ط- تحديد المتغيرات التفسيرية والمعدلة المناسبة.

ومن المصادر التي يمكن الرجوع إليها للوصول إلى الدراسات ذات الصلة بموضوع ما وراء التحليل:

- قواعد البيانات المتاحة ضمن مجالات البحوث المتخصصة.
- السجلات البحثية، وهي قوائم لحفظ الدراسات التي تتمحور حول موضوع مشترك.
- قوائم المراجع والمجلات لبحوث ومقالات يمكن أن تفيد في الرجوع إلى الدراسات الأولية.
- البحث اليدوي في فهرس المجلات الجيدة للوصول إلى بحث نشط عن موضوع.
- مقابلات مع باحثين وعلماء في المجال للاسترشاد منهم إلى بحوث منشورة أو غير منشورة.
- خطابات إلى باحثين نشطاء في المجال للاسترشاد منهم إلى بحوث منشورة أو غير منشورة.

2 - حصر وتحديد الدراسات السابقة ذات الثقة التي يشملها ما وراء التحليل

يجب على الباحث وصف الطريقة التي اتبعها في تحديد واختيار الدراسات السابقة التي سيقوم بمراجعتها، وهذا يسمح للآخرين بالحكم على مدى شمولية الدراسات التي تمت مراجعتها، ويتيح للباحثين اللاحقين أيضاً التعرف بوضوح على النقاط التي تحتاج إلى مزيد من المراجعة، ويجب أن يسجل الباحث محكات استبعاد واستبقاء أي دراسة من الدراسات السابقة. ويرى البعض أنه نظراً للتحيز الذي يحدث خلال عملية جمع الدراسات؛ نتيجة لصعوبة الحصول على جميع الدراسات حول موضوع معين لاستخدامها فيما وراء التحليل، أو لاختيار عينة ممثلة لها، أو نتيجة لاستخدام الدراسات المنشورة فقط واستبعاد غير المنشورة؛ فإنه من الأفضل استخدام أسلوب تحديد العدد اللازم للدراسات لتعويض النقص فيها؛ لكي يكون من الممكن الاعتماد على نتائجها (Sutton, 2009). وقد يرجع هذا التحيز إلى الأسباب التالية:

أ- أن الدراسات المسجلة لحجوم أثر مرتفعة مناسبة تكون أكثر احتمالاً للنشر، مقارنة بالدراسات المسجلة لحجوم أثر منخفضة، وهذا يقود إلى تحيز في الأدبيات المنشورة.

ب- أن الدراسات ذات النتائج الدالة إحصائياً تكون أكثر احتمالاً في أن تجد طريقها في الأدبيات المنشورة مقارنة بالدراسات التي تسجل نتائج غير دالة إحصائياً، وقد تم تقدير نسبة الأدبيات المنشورة ذات الدلالة الإحصائية في التخصصات المختلفة بين 61% و86% (Dickersin & Min, 1993).

ج- أن الدراسات المنشورة تكون أكثر احتمالاً في وجودها بما وراء التحليل.

د- توجد بحوث غير منشورة جيدة ولكنها لا توجد بدرجة احتمال مرتفعة في عينة دراسات ما وراء التحليل؛ مثل بحوث الماجستير والدكتوراه المودعة بمكتبات الكليات والجامعات المختلفة، وتقارير المؤتمرات.

ه- لغة البحث أو النشر؛ حيث يظهر التميز في الدراسات المكتوبة باللغة التي يجيدها الباحث؛ مثل اللغة الإنجليزية؛ ومن ثم تُهمل البحوث الأخرى المكتوبة باللغة العربية أو الفرنسية أو الألمانية، وكذلك تناول البحوث التي يمكن الوصول إليها بسهولة فقط (المتاحة للباحث).

ولتجنب هذا التحيز يجب أن نتجنب العوامل السابقة التي تؤدي إلى التحيز؛ مثال ذلك أن نتناول دراسات ذات أحجام عينات متنوعة، وبلغات متعددة، ومنشورة وغير منشورة، وبمستويات دلالة وتقديرات حجم أثر متعددة. عادة ما يعرض الباحث آلية الوصول للدراسات المعتمدة في مراجعته الممنهجة من خلال شكل يسمى مخطط بريزما PRISMA، لتوضيح المراحل التي من خلالها تنتقى الدراسات.

3 – تحليل الخصائص المتعلقة بمضمون البحث وتصنيفها وجدولتها وتكويدها

إن تحديد العوامل والمتغيرات (الخصائص) المتعلقة بمضمون البحث والأكثر أهمية لإدراجها في التفريغ يعتمد على اعتبارات نظرية، وبعد تحديد الخصائص المتعلقة بمضمون البحث، التي يعتقد أنها تؤثر في تباين نتائج الدراسات يمكن تطبيق نظام التفريغ أو الترميز. وقد طرح عبد الحميد (1987) أمثلة لأبرز الخصائص التي يجب الاهتمام بها وتتمثل في: النوع الاجتماعي، سنة النشر، المرحلة العمرية، مجال الدراسة، مكان إجراء الدراسة، مدة الدراسة، نوع المعالجة الإحصائية، مصدر الدراسة، حجم العينة، حجم الأثر. كما حدد ديكوستر (Decoster, 2009) أبرز الخصائص التي يحتاج الباحث إلى ترميزها وتتمثل في: تعريف الدراسة ورقم مميز لكل دراسة، المراجع الطويلة والقصيرة، ترميز المتغيرات المعدلة، ترميز خصائص جودة الدراسة، ترميز نوع التصميم والمنهج البحثي المستخدم في

الدراسة. ومن الخصائص والمتغيرات التي يمكن التصنيف على أساسها أو تدوينها أيضاً: أدوات الدراسة وخصائصها السيكومترية، المرحلة العمرية والدراسية للأفراد المشاركين بالدراسة، تحديد مكان إجراء الدراسة.

4 –تقدير قيمة حجم الأثر بالدراسات

وفي هذه الخطوة يجب مراعاة الأمور التالية:

أ- تسجيل تقدير حجم الأثر الموجودة بالدراسات الأولية التي يحتاج إليها الباحث.

ب- تسجيل البيانات اللازمة لتقدير قيمة حجم الأثر التي نحتاج إليها من الدراسة الفردية.

ج- اختيار مؤشر حجم الأثر المستخدم كعامل مشترك أو القياس الموحد بين الدراسات الفردية المتضمنة في ما وراء التحليل.

د- عمل التحويلات اللازمة لتحويل مؤشرات حجم الأثر المختلفة في نوعها من دراسة فردية أولية إلى أخرى إلى المؤشر أو المقياس الموحد الذي اختارته كعامل مشترك.

ه- وصف توزيع تقديرات حجم الأثر بفئة التصنيف.

و- تقدير قيمة مدى تباين (أو تجانس) تقديرات حجم الأثر المجمعة بما وراء التحليل، وذلك باستخدام القيم (Q , I^2 , T^2)؛ من أجل استبعاد القيم المتطرفة وتحديد نموذج الآثار المستخدمة (الثابت أو العشوائي).

ز- تقدير قيمة متوسط حجم الأثر الموزون.

5– تلخيص وتفسير نتائج ما وراء التحليل

في هذه الخطوة يتم الربط بين نتائج الدراسة وتفسير هذه النتائج، ويمكن أن تمهد إلى نماذج أو نظريات مستقبلية. إن ما وراء التحصيل لا ينبغي ان يكون مجرد ملخص للأدبيات البحثية، ولكن ينبغي أن يوفر التفسير والتكامل النظري. إن ما وراء التحليل ليس تقريراً بسيطاً لمتوسط تقديرات حجم الأثر، ولكنه تفسير للنتائج وعرض لكيفية اتفاقها أو تعارضها مع النظريات الرئيسية في الأدبيات البحثية. إن ما وراء التحليل يقدم مساعدة كبيرة للأدبيات البحثية من خلال توفير ملخص لما يمكن العثور عليه في الأدبيات الموجودة، وهذا ينبغي أن يتبعه اقتراحات للمجالات التي لا تزال في حاجة إلى تطوير، ويقدم توصيات للاتجاه للبحث في المستقبل. يجب أن يميز الباحث أيضاً بين الدراسات التي اعتمد عليها البحث كمصادر وتلك التي مثلت عينة الدراسات البحثية في ما وراء التحليل؛ عادة ما تميز الأخيرة بوضع نجمة في قائمة المراجع.

وقد راجع (الخضر، 2022) عدداً كبيراً من المصادر التي ذكرت التفاصيل الإجرائية لكيفية القيام بالمراجعات الممنهجة وما وراء التحليل، وقام بترجمة وتلخيص للخطوات التي عرضها دانيال (Daniels, 2019) بعد مراجعة توصيات الرابطة الأمريكية لعلم النفس في دليلها المرجعي السابع، وصنفها لخطوات أساسية وأخرى تكميلية، وعلى الباحث المهتم الرجوع إلى مقالة الخضر المهمّة في هذا الصدد.

الصدق الداخلي والخارجي لما وراء التحليل

يمكن تصنيف صدق دراسات ما وراء التحليل إلى صنفين رئيسين: الصدق الداخلي والصدق الخارجي. ويتطلب الصدق الداخلي أن تكون الدراسات الأولية ذات جودة مرتفعة؛ لأن ما وراء التحليل لا يمكن أن يكون صادقاً، إذا وجدت مشكلات تتعلق بجودة الدراسات الأولية. كما يتعلق الصدق الداخلي بوجود عدد كاف من الدراسات المتضمنة في ما وراء التحليل، وذلك لكي يكون حجم الأثر المتوسط أكثر دقة. وعلى الرغم من أنه لا يوجد اتفاق على الحد الأدنى لعدد الدراسات المطلوبة لإجراء ما وراء التحليل، فإنه يجب ألا يقل عن عشر دراسات (Hunter & Schmidt, 2004). يتعلق الصدق الداخلي أيضاً بعدد الباحثين الذين يقومون بعملية التصنيف والترميز. إن أسلوب ما وراء التحليل يجب أن يشمل اثنين من الباحثين على الأقل في عملية التصنيف والترميز؛ لضمان ثبات الترميز الذي يجب أن يكون 0.80 فأكثر.

من ناحية أخرى، فإن العوامل التي تؤدي إلى رفع الصدق الخارجي لما وراء التحليل؛ أن تشمل كل دراسة أولية عينة ممثلة من المشاركين بتلك الدراسة، وأن يكون هناك تمثيل عادل للدراسات المتضمنة في التحليل، وطبيعة وأهمية موضوع ما وراء التحليل في حد ذاته. بالإضافة إلى ذلك فإن خلو الدراسات المتضمنة في التحليل من التحيز، يرفع من الصدق الخارجي لدراسة ما وراء التحليل. لذلك من المهم الحرص على أن يتضمن ما وراء التحليل البحوث المنشورة، وغير المنشورة، واستخدام مصادر المعلومات الإلكترونية، للحصول على الدراسات الأولية ذات الصلة بالموضوع، وكذلك الحصول على الدراسات بلغات مختلفة وكذلك الدراسات البعيدة والقريبة مكانياً.

الانتقادات الموجهة إلى ما وراء التحليل والرد عليها

يعرض بورينستين وزملاؤه (Borenstien et al., 2009) أهم الانتقادات الموجهة لما وراء التحليل والردود على هذه الانتقادات، وهي:

1- يخلط ما وراء التحليل بين نتائج الدراسات التي أجريت في سياقات مختلفة، وتحت ظروف متباينة. والرد على ذلك هو أن الهدف من مراجعة الأدبيات السابقة الخروج بتعميم من اختلاف نتائج الدراسات السابقة، وهذا يتم الوصول إليه باستخدام ما وراء التحليل بشكل أسهل من الطرق التقليدية.

2- يتجاهل ما وراء التحليل الاختلافات الكيفية بين الدراسات. والرد على ذلك هو أن ما وراء التحليل لا يتجاهل هذه الاختلافات، بل يقوم بترميزها كمتغيرات وسيطة بحيث يتم اختبار أثرها تجريبيا.

3- يتجاهل ما وراء التحليل جودة الدراسة. والرد على ذلك هو أن ما وراء التحليل لا يتجاهل جودة الدراسات؛ فإما أن يتم تضمين الدراسات ذات الجودة العالية فقط، وإما أن يتم ترميز جودة الدراسة كمتغير وسيط، ويختبر وجود اختلاف بين الدراسات الجيدة والسيئة.

4- لا يقودنا ما وراء التحليل إلى استنتاجات صحيحة تماماً؛ لأن الدراسات الدالة إحصائياً هي التي يتم نشرها. والرد على ذلك هو أن ما وراء التحليل الجيد يعتبر أقل عرضة للتحيز من المراجعات التقليدية؛ حيث يهتم بالدراسات الدالة وغير الدالة إحصائياً، كما يمكن من خلال ما وراء التحليل تقدير تأثير عدم وجود الدراسات غير الدالة إحصائياً فيما يسمى تقدير تأثير تحيز النشر Publication Bias.

ويضيف الكناني (2020) عدداً آخر من أوجه النقد الموجهة لما وراء التحليل هي:

1- صعوبة قيام باحث واحد بعملية تحليل وتلخيص النتائج الكثيرة للدراسات السابقة في موضوع بحثي واحد خلال الفترة الزمنية للدراسة؛ إذ إن ذلك يتطلب الكثير من الوقت والجهد وخبرة جوهرية بميدان الدراسة.

2- الأساليب الإحصائية المستخدمة في عمليات ما وراء التحليل قد تكون صعبة أو غامضة على الباحث القائم به، وخاصة في الحساب اليدوي، كما توجد صعوبة في الوصول إلى البرامج الإحصائية الخاصة بما وراء التحليل والقيام بها باستخدام الكمبيوتر.

3- بعض الدراسات المستخدمة في ما وراء التحليل تحتوي على مشكلات منهجية تمثل ضعفاً وخللاً فيها؛ لأن المعطيات الرديئة تؤدي إلى مخرجات رديئة.

4- وجود نقص أو قصور في تقرير نتائج الدراسة الفردية؛ إذ يوجد نقص في أجزاء مهمة من المعلومات (مثل حجم العينة، مستوى الدلالة، ..) أما إذا شمل تقرير القوة الإحصائية وحجوم الأثر في كل دراسة، فلن تكون هناك مشكلة أو صعوبة.

5- صعوبة الحصول على الكثير من الدراسات ونتائجها في موضوعات نوعية معينة، لكي يتم تحليلها باستخدام ما وراء التحليل، ولذلك يجب أن يبذل الباحث الجهد للحصول على دراسات عن طريق المكتبات، أو الباحثين أنفسهم، أصحاب الدراسات الفردية، رسائل غير منشورة للماجستير والدكتوراه، ودراسات مقبولة للنشر ولم تنشر، المجلات المتخصصة، قواعد البيانات.

6- يتسم ما وراء التحليل بذاتية القائم به، وهذه تبدو في قراراتهم التي تعتمد على حكمهم وتصنيفهم للمتغيرات الوسيطة أو المستقلة.

7- غالباً ما يتم تجاهل التفاعلات الإحصائية بين المتغيرات موضوع البحث.

8- الدراسات السابقة التي يتضمنها ما وراء التحليل تكون أكثر احتمالاً لأن تكون منشورة باللغة الإنجليزية أكثر من اللغات الأخرى.

9- اهتمام القائم بما وراء التحليل بمجلات نشر معينة مرتبط بها دون غيرها، أو يسهل الوصول إليها.

10- صعوبة التوصل إلى تعريف محدد لمستوى جودة الدراسات التي يشملها ما وراء التحليل.

المبحث السادس

قوة الاختبارات الإحصائية التي يتضمنها ما وراء التحليل

أنواع القرارات عند اختبار الفروض إحصائياً

لسنا في حاجة هنا إلى مراجعة فكرة الفرض الصفري والفرض البديل، التي تعرضها كتب أساسيات الإحصاء. ولكننا يمكننا البدء هنا بأنواع القرارات التي يتم اتخاذها بشأن الفرض الصفري. يوضح جدول (1/6) التفاعل بين الواقع الحقيقي للفرض الصفري (صحيح، خاطئ) ونتائج اختبار ذلك الفرض الصفري؛ ومن ثم اتخاذ القرار (الرفض/ عدم الرفض)، وينتج أربعة أنواع من النتائج يظهر فيها خطأ النوع الأول باحتمال (α) وخطأ النوع الثاني باحتمال (β) وكذلك القوة الإحصائية للاختبار، ودرجة الثقة في اتخاذ القرار.

جدول (1/6)

التفاعل بين الواقع الحقيقي للفرض الصفري (صحيح، خاطئ) ونتائج اختبار هذا الفرض الصفري (الرفض/ عدم الرفض)

الواقع الحقيقي للفرض H_0	قرار اختبار الفرض الصفري H_0	
الفرض الصفري غير صحيح	رفض H_0 رفض فرض صفر غير صحيح (قرار صائب) قوة الاختبار باحتمال ($1-\beta$)	عدم رفض H_0 خطأ عدم رفض فرض صفري غير صحيح خطأ النوع الثاني باحتمال (β)
الفرض الصفري صحيح	رفض H_0 خطأ رفض فرض صفري صحيح خطأ النوع الأول باحتمال (α)	عدم رفض H_0 عدم رفض فرض صفري غير صحيح (قرار صائب) باحتمال ثقة ($1-\alpha$)

يتبين من جدول (1/6) ما يلي:

1. أن خطأ النوع الأول α هو احتمال الخطأ في رفض الفرض الصفري H_0.
2. خطأ النوع الثاني β هو احتمال الخطأ في عدم رفض الفرض الصفري غير الصحيح H_0.
3. القوة الإحصائية Statistical Power هي احتمال الرفض الصحيح للفرض الصفري (ايجاد نتيجة دالة حقيقية)، وهي تساوي ($1-\beta$)؛ وهذا يعني صحة رفض الفرض الصفري.
4. مستوى الثقة في النتيجة يساوي ($1-\alpha$)*100

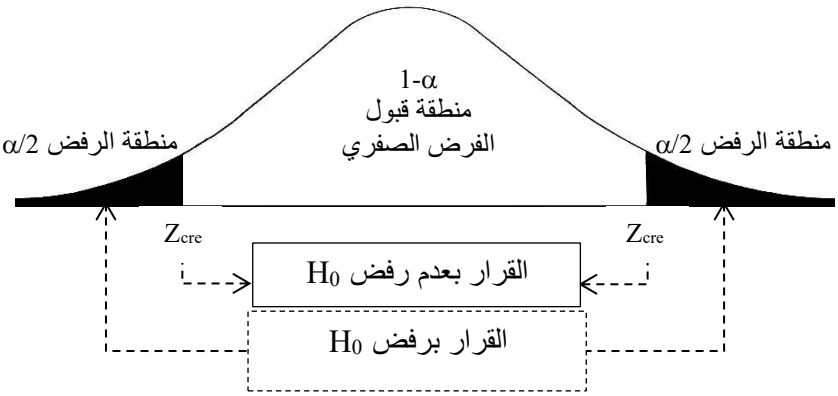
5. العلاقة بين α وβ علاقة عكسية.
6. زيادة القوة ($1-\beta$) تعني نقص β كما تعني زيادة α (بشـرط ثبات العوامل الأخرى المؤثرة على القوة الإحصائية).
7. يمكن حسـاب إحصـاء الخطورة النسـبية، التي تسـاوي احتمالات الوقوع في الخطأ الثاني مقسومة على احتمالات الوقوع في الخطأ من النوع الأول = β/α؛ فإذا كانت $\alpha=0.05$ و $\beta=0.20$ فان الخطورة النسبية تساوي 4=(0.05)/(0.20).

اختبار الفرض الصفري والفرض البديل

الفرض البديل هو الفرض الذي يمكن أن يُقبل كبديل للفرض الصفري، عندما يتم رفض الفرض الصـفري، وقد يكون الفرض البديل موجهاً أو غير موجه؛ فإذا كان موجهاً يكون الاختبار بطرف واحد؛ أي أن منطقة الرفض في جهة واحدة؛ ففي بعض المواقف ربما يفضـــل الباحث الإشـــارة الي وجود فرق لصـــالح جهة دون أخرى، أو تحديد اتجاه العلاقة أو اتجاه الفرق بين متوســـطين. أما إذا كان الفرض غير موجه، فعندها يكون الاختبار بذيلين؛ أي أن منطقة الرفض في الجهتين؛ إذ يشـير الباحث إلى وجود فرق بين مجموعتين أو أكثر، ولكنه لا يحدد لصـالح من هذا الفرق (عودة والخليلي، 1998). يعرض شكل (1/6) منطقتي قبول ورفض الفرض الصفري H_0 في حالة الفرض غير الموجه (طرفان).

شكل (1/6)

منطقتـي رفض أو عدم رفض الفرض الصفري H_0



بعد إخضـــاع الفرض الصـــفري للاختبار، واتخاذ قرار بحقه إما بالرفض وإما بعدم الرفض، وفي الواقع لا يســـتطيع الباحث التأكد مما إذا كان القرار الذي اتخذه صـــائباً أم غير صـائب، وكل ما يعرفه هو احتمالات الخطأ في هذا القرار؛ فإذا كان الفرض الصـفري صـحيحاً؛ ولم يتم رفضـه (مستوى الثقة) أو كان خاطئاً بالفعل ورفضـه الباحث (قوة الاختبار)، فإنه في كلتا الحالتين يكون قد اتخذ قراراً صائباً. أما إذا تم رفض الفرض الصفري وهو في الواقع صحيح فإن الباحث يكون قد وقع في الخطأ من النوع الأول، وإذا لم يتمكن من رفضـه، في حين أنه في الواقع خاطئ فإن الباحث يكون قد وقع في الخطأ من النوع الثاني، ولذلك فإنه أياً كان القرار، فهناك مستوى ما من الخطأ.

إن احتمال رفض الفروض الصـفرية وهي صـحيحة، يسـمي بالخطأ من النوع الأول، وأقصي احتمال لوقوع الباحث في هذا النوع من الخطأ، هو ما يسمي بمستوى الدلالة الإحصائية، ويرمز له بالرمز (α)، ومعنى الدلالة الإحصـــائية أن الفرق بين القيمة النظرية لمعلمة المجتمع والقيمة الناتجة من العينة، هو فرق حقيقي، ولا يعزى إلى المصـــادفة. أما احتمال عدم رفض الفروض الصـفرية وهي صـحيحة، فإن ذلك يسـمى بمسـتوى الثقة ($1-\alpha$). ولكن احتمال رفض الفروض الصفرية وهي في الواقع خاطئة، يسمي بقوة الاختبار الإحصائي ويرمز له بالرمز ($1-\beta$)، أما احتمال عدم رفض الفروض الصفرية وهي في الواقع خاطئة فإنه يسمى بالخطأ من النوع الثاني، واحتمال وقوع الباحث في هذا النوع من الخطأ يرمز له بالرمز β (الشريفين، 2017).

ويرغب الباحث عادة بتصـــميم دراســـة تتســـم بدرجة عالية من القوة؛ بمعنى أنه عندما يختبر الفروض الصفرية، فإنه يحاول دائماً اتخاذ القرار الصائب؛ فهو يرغب في رفض الفروض الصـفرية الخاطئة، وعدم رفض الفروض الصـفرية الصـحيحة؛ بمعني أنه يرغب في التقليل من نوعي الخطأ، وهذا أمر صـــعب نتيجة العلاقة العكســـية بين نوعي الخطأ. لذلك لابد أن يركز الباحثون على نوع من الخطأ أكثر من الآخر، وعادة ما يهتم الباحث باحتمالات الوقوع في الخطأ من النوع الأول بشـكل أكبر، ويترتب على ذلك الاهتمام زيادة احتمال الوقوع في الخطأ من النوع الثاني في بعض الأحيان. لذلك ينصـــح الإحصـــائيون بأن يراعي الباحث عند اختياره لمســـتوى الدلالة طبيعة الدراسـة وفيما تسـتخدم نتائجها. فإذا كانت الدراسـة من النوع الذي يؤدي فيه رفض الفروض الصـــفرية إلى مخاطر أو إنفاق أموال طائلة أو إهدار الوقت والجهد، فإنه ينبغي الحذر من هذا النوع من الخطأ؛ إذ يضـــحي الباحث إلى حد ما بقوة الاختبار من أجل تقليل هذا الخطأ،

كأن يجعل (α=0.01) أو حتى (α=0.001). أما إذا كانت الدراسة استطلاعية ويحاول الباحث التأكد مما إن كانت هناك حاجة إلى مزيد من الدراسات حول الموضوع أو الظاهرة؛ عندئذ فإن زيادة قوة الاختبار تعتبر أمراً مهماً، بينما الخطأ من النوع الأول يصبح أقل أهمية (أبو علام، 1993).

القوة الإحصائية والتخطيط للدراسة

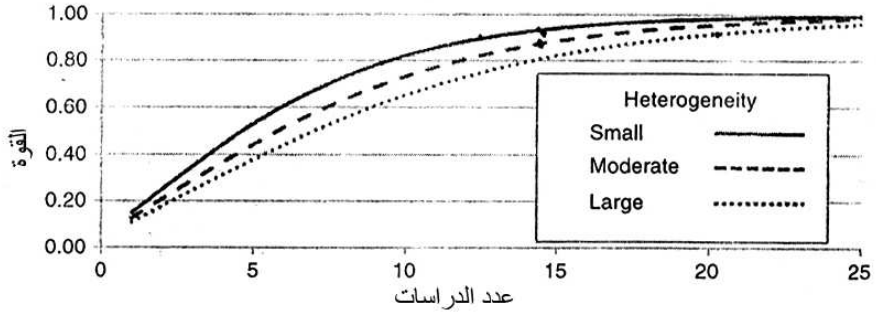
إن القوة الإحصائية مهمة عند التخطيط لدراسة ما؛ فإذا قمت بدراسة فيها قوة منخفضة فإن الدراسة من المحتمل ألا تعطي نتائج دالة حتى إذا كان الفرض البحثي (البديل) صحيحاً؛ ومن ثم، فإن هناك علامات استفهام تتعلق بتخطيط هذه الدراسة. لذلك يتحرى الباحثون طرقاً عملية لزيادة مستوى القوة الإحصائية إلى مستوى مقبول أو حد أدنى لمستوى القوة الإحصائية. يساعد وجود مثل هذا المستوى عند التخطيط للدراسة، خاصة في تحديد عدد المشاركين الذين تتطلبهم تلك الدراسة، وكذلك فإن فهم القوة الإحصائية مهم عند قراءة تقرير البحث للحصول على معنى للنتائج الإحصائية غير الدالة. وقد اقترح كثير من الباحثين، من أبرزهم كوهين (Cohen, 1988) أن لا تقل قوة الاختبار الإحصائي المستخدم لاختبار الفروض الصفرية عن. 80؛ فإذا كانت نتائج الاختبار ذات دلالة إحصائية، وكان حجم الأثر كبيراً فهناك احتمال كبير في الحصول على نتائج ذات دلالة عملية، أما إذا كان حجم الأثر صغيراً وكانت النتيجة ذات دلالة إحصائية فهذا قد يكون بسبب عوامل أخرى. وإذا كانت نتائج الاختبار غير دالة إحصائياً والقوة الإحصائية مرتفعة، فهذا يزيد ثقتنا في صحة الفرض الصفري، أو يعود بنا إلى قيمة الفرض والبناء النظري الذي أُسس عليه الفرض البحثي (Thompson, 1999).

العوامل المؤثرة في قوة الاختبار الإحصائي

أن النتيجة غير الدالة لدراسة منخفضة القوة، تعتبر نتيجة غير حاسمة، وعلى العكس فإن النتيجة غير الدالة لدراسة ذات قوة مرتفعة، قد ترجع إلي وجود فرض بحثي خاطئ، أو إلى حجم أثر أقل من المتوقع. تعتمد القوة على عوامل متعددة؛ من أهمها مستوى الدلالة الإحصائية α؛ أي احتمالية الوقوع في الخطأ من النوع الأول، وعدد المشاركين (أو حجم العينة)، وحجم الأثر، وكذلك تزداد قوة الاختبار الإحصائي في حالة استخدام اختبار ذي اتجاه واحد أو اختبار فرض موجه، بالمقارنة باختبار ذوي طرفين أو اختبار فرض غير موجه وفيما يلي تفصيل لذلك (Ellis, 2010)

- **مستوى الدلالة الإحصائية:** توجد علاقة عكسية بين β، α (مع تثبيت العوامل الأخرى) فإذا زادت قيمة α (زيادة احتمال رفض فرض صفري رغم أنه صحيح) من 0.01 إلى 0.05 مثلا فان قيمة β تنخفض (نقص احتمال قبول فرض صفري رغم أنه خطأ) وبالتالي تزداد القوة (β-1) وعلى العكس، فعندما تقل قيمة α من 0.05 إلي 0.01 تزداد قيمة β وتقل القوة.
- **اتجاه الفرض البديل:** بتثبيت جميع العوامل والمتغيرات الأخرى، فان الاختبارات ذات الطرف الواحد تكون أقوي من الاختبارات ذات الطرفين.
- **حجم العينة**: يؤثر حجم العينة على قوة الاختبار، فعند استخدام حجم كبير للعينة يؤدي الي الحصول على استدلالات دقيقة لمعالم المجتمع، وذلك لان زيادة حجم العينة يقلل الخطأ المعياري، كما يؤدي الي زيادة قوة الاختبار الإحصائي، وذلك عند تثبيت جميع المتغيرات الأخرى. لذلك إن أكثر الطرق شهرة لزيادة القوة، هو زيادة عدد المفحوصين؛ ومن ثم فان اية فروق بين المتغيرات حتى ولو كانت صغيرة فإنها تكون دالة إحصائياً، وبذلك يصبح الاختبار الإحصائي أكثر حساسية، عندما يصل إلى مستويات الدلالة الإحصائية بعينة أصغر من غيره.
- **حجم الأثر:** مقدار أو حجم الأثر هو الفرق بين الدرجة المعيارية للفرض الصفري والدرجة المعيارية للفرض البديل، أو الفرق المرغوب به الذي يمكن استنتاجه. مع ثبات العوامل الأخرى، كلما كان حجم الأثر أكبر ازدادت فرصة رفض الفرض الصفري الخاطئ؛ مما يرفع من قيمة القوة الإحصائية.
- **عدد الدراسات**: يبين شكل (2/6) (في Cooper& Findley, 1982) العلاقة بين عدد الدراسات وقوة التحليل؛ إذ تزداد العلاقة بين عدد الدراسات المتضمنة بالتحليل وقوته، وتزداد العلاقة بينهما كلما زاد مستوى حجم الأثر.

شكل (2/6)
قوة ما وراء التحليل كدالة لعدد الدراسات ومدى التجانس Cooper& Findley, 1982



من شكل (2/6) نستنتج أنه إذا كان التشتت بين الدراسات منخفضاً فإننا نحتاج إلى نحو 12 دراسة لكي نحصل على قوة 90%. أما إذا كان التشتت بين الدراسات متوسطاً فإننا نحتاج إلى 15 دراسة للوصول إلى المستوى نفسه للقوة، ونحتاج إلى 20 دراسة للوصول إلى المستوى نفسه للقوة إذا كان التشتت بين الدراسات مرتفعاً. وإذا افترضنا أن لدينا 15 دراسة تمثل الدراسات المتضمنة في دراسة ما وراء التحليل؛ فإذا كان حجم الأثر المتوقع منخفضاً، فإننا نتوقع قوة 90%، أما إذا كان حجم الأثر متوسطاً فإننا نتوقع قوة في حدود 92%، أما إذا كان حجم الأثر مرتفعاً فإن القوة الإحصائية ترتفع إلى 95%.

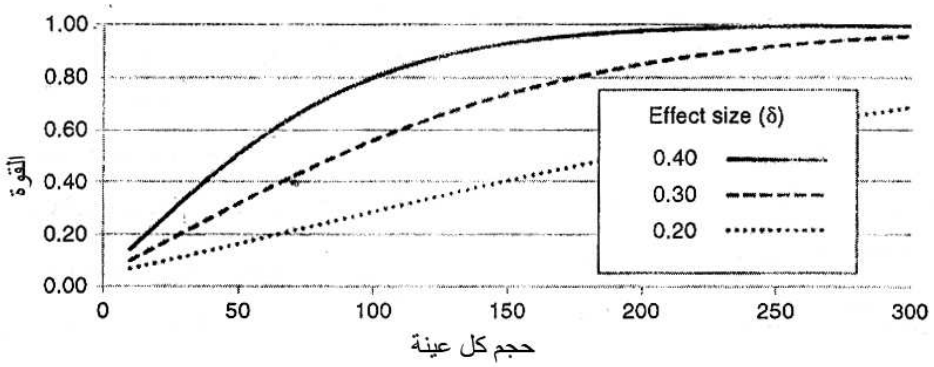
ومن الملاحظات الخاصة بنموذج الأثر العشوائية أنه لتحقيق قوة جيدة في دراسة ما وراء التحليل، يجب أن يشمل التحليل عدداً أكبر من الدراسات، أما إذا كان عدد الدراسات صغيراً، فإن القوة تظل منخفضة حتى ولو كان حجم العينة الكلية كبيراً جداً. وإن القوة تكون أقل في إطار نموذج الأثر العشوائية مقارنة بقوة التحليل في إطار نموذج الأثر الثابت؛ إذ إن الخطأ المعياري يكون أقل في الحالة الأخيرة، عند ثبوت العوامل الأخرى (Riley, et al., 2011).

يعرض شكل (3/6) (في Cooper & Findley, 1982) العلاقة بين حجم العينة والقوة الإحصائية عند مستويات مختلفة من حجم الأثر. يتضح من الشكل أن حجم العينة يرفع من القوة عند تثبيت حجم الأثر، وأن تأثير حجم العينة يكون أقل أهمية عند تقديرات حجم الأثر المرتفعة بعد الوصول إلى عدد كافٍ من الدراسات. مثال ذلك ووفقاً للشكل (3/6) فإننا نكون في حاجة إلى عينة 300 للكشف عن حجم أثر (0.2) للوصول إلى مستوى قوة نحو 70%؛ في حين

نكون في حاجة إلى حجم عينة 120 فقط للوصول إلى قوة إحصائية 90% للكشف عن حجم أثر (0.4).

شكل (3/6)

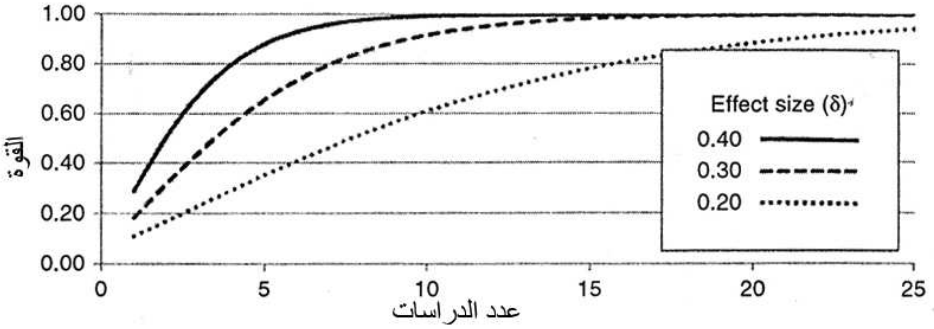
القوة كدالة لحجم العينة ومستويات تقديرات حجم الأثر (في Cooper & Findley, 1982)



يعرض شـــكل (4/6) (في Cooper& Findley, 1982) العلاقة بين القوة وعدد دراسات ما وراء التحليل عند ثلاثة مستويات لحجم الأثر (0.20، 0.30، 0.40) في إطار نموذج الأثر الثابت، واختبار ذي طرفين (0.05). يتضـــح من الشـــكل (4/6) أنه بزيادة عدد الدراســـات تزداد قوة ما وراء التحليل وأن هذه القوة تزداد بزيادة حجم الأثر.

شكل (4/6)

قوة ما وراء التحليل كدالة لعدد الدراسات ومستويات تقديرات حجم الأثر (في Cooper & Findley, 1982)



دراسات القوة الإحصائية

يفترض أن يجري الباحث تحليلاً لقوة الاختبار قبل أو بعد إجراء الدراسة لتحديد عدد أفراد العينة الذين يجب دراستهم (قبل الدراسة)، أو التأكد من أن لديه العدد الكافي من الأفراد لكشف أي فروق ظاهرة بين مجموعات الدراسة (بعد إجراء الدراسة) إذا وجدت. يسمى تحليل قوة الاختبار قبل إجراء الدراسة Priori Power Analysis، وهذا هو الأصل في تحليل القوة؛ إذ يستطيع الباحث، من خلال الرجوع إلى جداول خاصة ومعدة سابقاً من قبل كوهين (Cohen,1988)، أو من خلال برامج حاسوبية خاصة (من أشهرها برنامج GPower)، تحديد عدد أفراد العينة الذين سيحتاج إليهم، قبل إجراء الدراسة، حتي يتوصل لمستوى مقبول من القوة الإحصائية آخذاً بالاعتبار مستوى الخطأ من النوع الأول، وحجم الأثر في المجتمع الذي يتم تقديره من خلال الدراسات السابقة. ويفيد هذا التحليل كوسيلة لضبط قوة الاختبار قبل إجراء الدراسة بشكل فعلي، ويوصى بذلك في الحالات التي لا يمثل فيها الوقت أو التكاليف عاملاً حاسماً عند إجراء الدراسة.

قد يقوم الباحث بتحليل القوة بعد إجراء الدراسة Post-hoc Power Analysis؛ حيث تقدر قيمة قوة الاختبار كدالة لقيمة الخطأ من النوع الأول (α) وحجم الأثر المتوقع في المجتمع، وحجم العينة المستخدمة في الدراسة. وهذا الأسلوب يمكن الباحثين من تقييم مدى فاعلية الاختبار الإحصائي المستخدم في رفض الفروض الصفرية غير الصحيحة بعد الانتهاء من الدراسة. هناك نوع ثالث من تحليلات القوة يسمى تحليل قوة الاختبار المتوازن Compromise Power Analysis، وبه تقدر قيمتا الخطأ من النوعين الأول والثاني وقوة الاختبار من خلال حجم الأثر وحجم العينة ومقدار نسبة الخطأ الاحتمالية التي تساوي الخطأ من النوع الثاني مقسوماً على الخطأ من النوع الأول (β/α)، وتدل قيمة الناتج (واحد صحيح) على تفضيل الباحث للمساواة بين الخطأين في حين أن القيمة أربعة تدل على أن قيمة الخطأ من النوع الثاني مساوية لأربعة اضعاف الخطأ من النوع الأول، ويفيد هذا الأسلوب قبل جمع البيانات أو بعده، وينتج عن التحليل القبلي لقوة الاختبار حجم عينة أكبر مما هو متوافر، ومن خلال التحليل المتوازن يستطيع أن يحسب قيمة الخطأ من النوع الأول وقوة الاختبار الإحصائي. وفي حالة أن الدراسة قد أجريت ولم تُحلل نتائجها بعد، فإن الباحث قد يلجأ إلى تحديد قيمة تضمن له توازن الخطأين معاً ولتكن

مثلاً قيمة متساوية لكليهما آخذين في الاعتبار حجم العينة وحجم الأثر، علماً بأن هذا الأسلوب قد ينتج قيماً غير مرغوب فيها للخطأ من النوع الأول قد تتجاوز 0.05 في حالة حجم أثر وعينة صغيرين أو قيمة تقل عن 0.001 في حالات حجم أثر وعينة كبيرين.

تحليل حساسية الاختبار

في تحليل حساسية الاختبار الإحصائي Sensitivity Analysis يتم تقدير قيمة حجم الأثر كدالة لقيمة الخطأ من النوع الأول وقوة الاختبار وحجم العينة. ويفيد هذا التحليل بشكل خاص في تقييم البحوث المنشورة؛ إذ إنه يزودنا بإجابات عن أسئلة من مثل: ما حجم الأثر الذي يمكن أن تعكسه نتائج الدراسة عند تحديد مستوى قوة الاختبار الإحصائي بـ 0.80 مع الأخذ بالحسبان حجم العينة واحتمال الوقوع في الخطأ من النوع الأول؟ أو ما الحد الأدنى لحجم الأثر الذي يظهر أن الاختبار حساس لوجوده؟

قوة الاختبارات الإحصائية المتضمنة بما وراء التحليل

فيما يلي نعرض لقوة بعض الاختبارات الإحصائية المتضمنة بما وراء التحليل، مثل قوة اختبار Z لتحديد مدى دلالة متوسط تقديرات حجم الأثر ومدى تجانس (أو تباين) تقديرات حجم الأثر، وقوة نموذج المتغير التصنيفي، في إطار نموذجي الآثار الثابتة والعشوائية.

خطوات تقدير قيمة قوة الاختبارات المتضمنة بما وراء التحليل (Hedges, & Pigott, 2011)

1) اتخاذ قرار القيمة الحرجة C_α المناظرة لمستوى الدلالة α.
2) تحديد متوسط حجم الأثر الذي تريده أو تهتم به $\bar{y}$ أو $\overline{E_s}$ أو θ.
3) تحديد عدد الدراسات المتضمنة بما وراء التحليل (k) وعدد المشاركين بهذه الدراساتN.
4) اختيار نموذج الأثر الثابت أو نموذج الآثار العشوائية.

أ) تحديد التباين النموذجي لحجم الأثر $V_{\bar{y}}$ وكذلك داخل الدراسات في إطار نموذج الأثر الثابت.

ب) تحديد التباين النموذجي بين الدراسات τ_x^2 في إطار نموذج الآثار العشوائية.

5) تقدير قيمة قوة الاختبارات النوعية المتضمنة بما وراء التحليل.

تقدير قيمة قوة اختبار Z المحدد لدلالة متوسط تقديرات حجم الأثر

قوة الاختبار الإحصائي هي الرفض الصحيح للفرض الصفري، لصالح الفرض البديل، وهو مهم للتخطيط المستقبلي عندما يكون الاهتمام بالأثر المشترك للتدخل ومعرفة أقل عدد من الدراسات التي نحتاج إليها لاستكشاف حجم الأثر الذي نهتم به ونتوقعه.

- **قوة اختبار Z المستخدم لتحديد مدى دلالة متوسط حجم الأثر في إطار نموذج الأثر الثابت:**

إذا كان توقعنا هو فاعلية المعالجة فإننا نرى أن متوسط حجم الأثر يختلف عن الصفر أو أن هناك فرقاً بين مجموعة المعالجة والمجموعة الضابطة. وهذا يعني:

$$H_o: \bar{y} = 0 \ , \ H_a: \bar{y} \geq 0$$

في هذه الحالة يكون اختبار الفرض البديل ذا اتجاه واحد (طرف واحد)، وتكون القيمة الحرجة هي C_α أما إذا كان $H_o: \bar{y} = 0 \ , \ H_a: \bar{y} \neq 0$؛ فإن الاختبار يكون ذا اتجاهين (طرفين)، وتكون القيمة الحرجة هي $C_{\alpha/2}$

وسوف نرفض الفرض الصفري عندما تكون القيمة المطلقة لـ $Z \leq C_{\alpha/2}$

قيمة Z لاختبار دلالة حجم الأثر تساوي $Z = \frac{(\bar{y}-0)}{\sqrt{V_{\bar{y}}}} = \bar{y} / \sqrt{V_{\bar{y}}}$

حيث $\bar{y}$ هي حجم الأثر المقدر من البيانات، $V_{\bar{y}}$ هو تباين حجم الأثر، Z هي القيمة المحسوبة التي يتم مقارنتها بالقيمة الحرجة (C_α) المستخرجة من التوزيع الاعتدالي المعياري، ولاختبار ذي طرفين تكون القوة مساوية: $p = 2\ (1 - \phi\ |Z|)$

حيث $\phi\ (x)$ هي دالة التوزيع المتجمع للمنحنى الاعتدالي المعياري، وهذا يعني المساحة أسفل المنحنى الاعتدالي المعياري من $(-\infty)$ *إلى* (x). وباستخدام اختبار ذي طرفين تكون معادلة تقدير قيمة القوة هي (Hedges, & Pigott, 2011)

$$\text{Power} = [1 - \phi(C_\alpha - \lambda)] + \phi(-C_\alpha - \lambda)$$

حيث C_α هي القيمة الحرجة للقيمة Z عند مستوى دلالة α و $\lambda = \bar{y}/\sqrt{V_{\bar{y}}}$

مثال ذلك عندما تكون $\alpha=0.05$ فإن $C_\alpha=1.96$.

وحيث إن التباين داخل الدراسات متساوٍ ويساوي $V_{\bar{y}}$؛ لذلك فإن التباين الكلي داخل الدراسات يساوي:

$$V_{\bar{y}} = 1/\sum w_i = 1/(\sum 1/V_y)$$

$$V_{\bar{y}} = 1/(k / V_y) = V_y / k$$

تشير زيادة التباين حول المتوسط في الدراسة الواحدة V_i إلى زيادة خطأ المعاينة بتلك الدراسة، وكل دراسة تتضمن مشاركين مختلفين؛ ولذلك يوجد خطأ المعاينة فقط بنموذج الآثار الثابتة، والبارامتر λ يتحدد بالمعادلة

$$\lambda = \delta/\sqrt{V_\delta} = \bar{y}/\sqrt{V_{\bar{y}}}$$

حيث إن δ هو حجم الأثر الحقيقي أو المفترض، و V_δ هو تباينه.

ويتحدد الخطأ المعياري لمتوسط تقديرات حجم الأثر (في نموذج الأثر الثابت) باستخدام المعادلة:

$$SE_{\bar{y}} = \sqrt{\frac{\sigma^2}{k \cdot n}} = \sqrt{\frac{V_i}{K.n}}$$

حيث $V_i = \sigma^2$ = التباين في الدراسة، و k عدد الدراسات، n عدد المشاركين بكل دراسة، σ الانحراف المعياري. وتبين المعادلة أن الخطأ المعياري لمتوسط حجم الأثر يقل بزيادة حجم العينة الكلية للدراسات (حاصل ضرب k و n)، وإذا زادت قيمة (حاصل ضرب k و n) بدرجة كبيرة جداً فإن الخطأ المعياري يقترب من الصفر، والقوة سوف تقترب من 100%، ويلاحظ أن القوة بأي دراسة من دراسات ما وراء التحليل (بنموذج الأثر الثابت) تكون متماثلة (Hedges, & Pigott, 2011)

مثال 1: يرغب باحث في دراسة فعالية برنامج تدريبي تم دراسة أثره في عدة دراسات منفصلة على المجتمع نفسه؛ لزيادة درجات اختبار SAT في حدود 50 درجة، علماً بأن الانحراف المعياري للاختبار يساوي 100 (انظر Valentine, et al., 2010)

حجم الأثر المرغوب فيه $\bar{y} = \frac{50}{100} = 0.5$

ولتقدير قيمة قوة اختبار حجم الأثر هنا إذا كان عدد أفراد المجموعة الضابطة 20 والمجموعة التجريبية 20، وعدد الدراسات المتضمنة فيما وراء التحليل 10 دراسات؛ فإنه يمكن تقدير قيمة

التباين المشترك داخل الدراسة باستخدام المعادلة:

$$V_i = \frac{n_c+n_e}{(n_c)(n_e)} + \frac{(\bar{y})^2}{2\,(n_c+n_e)}$$

$$V_i = \frac{20+20}{20\,(20)} + \frac{(0.5)^2}{2\,(20+20)} = 0.10$$

$$V^* = \frac{V_i}{k} = 0.10\ /10 = 0.01$$

$$\lambda = \frac{\bar{y}}{V_*} = \frac{0.5}{\sqrt{0.1}} = 5$$

$$C_\alpha = 1.645$$

قوة اختبار تقدير قيمة متوسط حجم الأثر لاختبار ذي طرف واحد:

$$\text{Power} = [1 - \phi(C_\alpha - \lambda)] + \phi(-C_\alpha - \lambda) = [1 - \phi(1.645 - 5)] + 0$$

$$\text{Power} = [1 - \phi(-3.36)] = 1-0 = 1$$

وهذا يعني أنه توجد قوة كافية لإيجاد متوسط حجم أثر مقداره 0.5 في عشر دراسات فقط و α 0.05= لاختبار ذي اتجاه واحد، وحجم عينة (40). وللعدد نفسه من الدراسات (10)، ولكن عندما ينقص حجم الأثر، فإننا نتوقع تغييراً في القوة. فإذا كان حجم الأثر يساوي 0.20 وهذا يناظر فرقاً مقداره (20) بين درجتي المجموعتين، وكان الانحراف المعياري لاختبار SAT يساوي 100 وعند استخدام اختبار ذي ذيلين (0.05) وكان التباين (V_i) مساوياً 0.201 لذلك؛ فإن قوة اختبار حجم الأثر لاختبار ذي طرفين:

$$V^* = 1/\ (\sum 1/V_i) = 1\ /\ (10\ /0.201) = 0.02$$

$$\lambda = \frac{0.20-0}{\sqrt{0.02}} = \frac{0.20}{0.14} = 1.41$$

$$C_{\alpha/2} = 1.96$$

$$\text{Power} = 1\text{- } [\,\phi(1.96 - 1.41) - \phi(-1.96 - 1.41)]$$

$$= 1\text{- } \phi(0.55) + \phi(-3.37)$$

$$= 1\text{- } 0.71 + 0 = 0.29$$

في هذا المثال يتبين لنا نقص القوة بسبب نقص حجم الأثر.

مثال 2: في مجموعة مكونة من 25 مريضاً تم تحديد دواء لمعرفة أثره في علاج الصداع النصفي ومجموعة ضابطة مقارنة عددها 25 مريضاً، وكان حجم الأثر d=0.30، ما القوة

الإحصائية لنتائج (10) دراسات؟ (أنظر Valentine, et al., 2010)

تباين حجم الأثر

$$V_{id} = \frac{n_1+n_2}{(n_1)(n_2)} + \frac{d^2}{2\,(n_1+n_2)}$$

$$V_{id} = \frac{25+25}{25\,(25)} + \frac{(0.30)^2}{2\,(25+25)} = 0.0809$$

$$\lambda = \frac{0.30}{\sqrt{0.0809}} = 1.0574$$

القيمة الحرجة C_α لاختبار ذي طرفين عندما $\alpha = 0.05$ تناظر $Z = 1.96$

$$\text{Power} = 1 - \phi(1.96 - 1.0547) + \phi(-1.96 - 1.0547) = 0.184$$

تباين متوسط حجم الأثر: $V_{\bar{y}} = 0.0809/10 = 0.00809$

$$\lambda = \frac{0.30}{\sqrt{0.00809}} = 3.3354$$

$$C_\alpha = \phi\left(1 - \frac{0.05}{2}\right) = 1.96$$

قوة اختبار حجم الأثر $\text{Power} = 1 - \phi(C_\alpha - \lambda) + \phi(-C_\alpha - \lambda)$

$$= 1 - \phi(1.96 - 3.3354) + \phi(-1.96 - 3.3354) = 0.9155$$

- **قوة اختبار Z الخاص بمدى مدى دلالة متوسط تقديرات حجم الأثر العشوائية** (Riley, et al., 2011)

في نموذج الآثار العشوائية يوجد مصدران للخطأ: الخطأ نتيجة التباين σ^2 والخطأ نتيجة التباين بين الدراسات τ^2، ولذلك يتم تحديد الخطأ المعياري لمتوسط تقديرات حجم الأثر باستخدام المعادلات:

الخطأ المعياري لـ $\bar{y}$ $\qquad SE^*_{\bar{y}} = \sqrt{\frac{\sigma^2}{k.n} + \frac{\tau^2}{k}}$

تباين متوسط الأثر $\bar{y}$ $\qquad V^*_{\bar{y}} = (V_y + \tau^2)/k$

وهذه القيمة (الخطأ المعياري) تقترب من الصفر عندما تكون القيمة k.n كبيرة جداً، وكذلك عندما يكون عدد الدراسات k كبيراً جداً أو أن التباين بين حجوم الأثر الحقيقية يساوي صفراً، أو حجم العينات (n) كبيراً وإذا كان حجم الأثر ثابتاً من دراسة إلى أخرى فإن τ تكون

صـــغيرة. ولذلك فإن قوة التحليل البعدي (ما وراء التحليل) يعتمد على العوامل الداخلية للعينة المؤثرة في الخطأ المعياري لمتوســط تقديرات حجم الأثر، وخاصـــة حجم العينة التراكمية (التي تنتج من مجموع العينات في الدراسات الخاضعة للتحليل)؛ إذ ينزع ما وراء التحليل إلى أن يكون مرتفع القوة مقارنة بأي دراســـة في هذا التحليل، أما إذا اختلف حجم الأثر اختلافاً جوهرياً من دراسة إلى أخرى، وتضمن التحليل عدداً قليلاً من الدراسات؛ فإن ذلك سوف يحد من قوة ما وراء التحليل، حتى لو تضمن التحليل عدداً كبيراً جداً من المشاركين.

وفيما يلي نقارن بين متوسط الأثر المقدر والصفر لتحديد مستوى دلالة هذا المتوسط أو اختلافه عن الصفر، وذلك باستخدام اختبار Z^* كما يلي (أنظر Valentine, et al., 2010)

$$Z^* = (\bar{y}^* - 0) / \sqrt{V_{\bar{y}}^*}$$

حيث $\bar{y}^*$ متوسط حجم الأثر الذي نحصل عليه و $V_{\bar{y}}^*$ هو تباين متوسط حجم الأثر، Z^* هي قيمة الاختبار الإحصائي Z لتحديد مستوى دلالة الفرق.

$$\lambda^* = \frac{(\bar{y}^* - 0)}{\sqrt{V_{\bar{y}}^*}} = \frac{\bar{y}^*}{\sqrt{V_{\bar{y}}^*}}$$

حيث $\bar{y}^*$ متوسط حجم الأثر في إطار نموذج الآثار العشوائية و $V_{\bar{y}}^*$ هو تباين متوسط حجم الأثر $\bar{y}^*$ في إطار نموذج الآثار العشوائية أيضاً.

ولتقدير قيمة قوة اختبار دلالة متوسط تقديرات حجم الأثر نحسب النسبة التي تجتاز $C_{\alpha/2}$ في التوزيع الاعتدالي الخاص بالفرض البديل بمتوســـط $\bar{y}^*$ وتباين مقداره (1)، ويمكن الحصول على هذه النسبة وقوة اختبارنا بتقدير قيمة P.

قوة اختبار دلالة $\bar{y}^*$ (Hedges, & Pigott, 2011)

$$P= 1- \phi (C_\alpha - \lambda^*)$$

حيث $\phi(x)$ هو دالة التوزيع المتجمع للتوزيع الاعتدالي المعياري، التي تساوي المساحة أســـفل المنحنى الاعتدالي المعياري من $-\infty$ إلى x، وفي حالة اختبار ذي طرفين أو ذي طرف واحد فإننا نتبع منطقاً مماثلاً وهو أننا نهتم بنسبة الاختبار الإحصـــائي التي تكون أكبر أو أقل من $C_{\alpha/2}$ أو $|C_{\alpha/2}| \leq Z^*$؛ ولذلك تتحدد قوة الاختبار باستخدام المعادلة:

قوة اختبار Z لتحديد مستوى الدلالة: $P^*=2[1-\phi(|Z^*|)]$

وقوة اختبار Z ذيو الطرفين لدلالة متوسط حجم الأثر، تحسب باستخدام المعادلة

$$Power = 1-\phi(C_\alpha-\lambda^*)+\phi(-C_\alpha-\lambda^*)$$

وتتحدد القيمة الحرجة Z المناظرة لـ C_α كما سبق من خلال المعادلة

$$C_\alpha = \phi[1-(\alpha/2)]$$

وقد وضع هيدجز و بيجو (Hedges & Pigott, 2001) قاعدة أو عرفاً، هي استخدام ثلاثة مستويات للتباين هي (صغير ومتوسط وكبير) حيث نضع τ^2 مساوية لـ 0.33 و 0.67 و 1.00 وهي أضعاف مضروبة في التباين داخل الدراسة، ولذلك فإن التباين الكلي يساوي

$$V_{\bar{y}}^* = 1.33\ V_{\bar{y}}/k$$

$$V_{\bar{y}}^* = 1.67\ V_{\bar{y}}/k \quad \text{أو} = 2.00\ V_{\bar{y}}/k$$

ويحدد هيجنز وطومسون (Higgins & Thompson, 2002) العلاقة بين I^2 و τ^2 من خلال المعادلة التالية: $I^2 = \frac{\tau^2}{\tau^2+V}$

ويحددان ثلاثة مستويات لـ I^2 هي 25% (منخفض)، 50% (متوسط)، 75% (مرتفع). وبالتعويض في المعادلة السابقة نجد ما يلي:

في حالة المستوى المرتفع $I^2 \frac{\tau^2}{\tau^2+V} = 0.75$، وعندما $\tau^2= 3V$: $\therefore I^2 = \frac{3V}{3V+V}$

وعندما تكون I^2 تكون متوسطة فإن $\tau^2= V$ $\therefore I^2 = \frac{V}{V+V} = 50\%$

وعندما تكون I^2 منخفضة فإن $\tau^2= \frac{1}{3}V$ $\therefore I^2 = \frac{\frac{1}{3}V}{\frac{1}{3}V+V} = \frac{V}{4V} = 25\%$

ولتقدير قيمة $V_{\bar{y}}^*$ فإننا نستخدم المعادلة

$$V_{\bar{y}}^* = \frac{1}{\sum_1^k w_i^* 1} = \sum \frac{1}{\tau^2+V} = \frac{\tau^2+V}{k}$$

مثال (3): لتقدير قيمة القوة في إطار نموذج الأثر الثابت لمتوسط الأثر لفئة من الدراسات الخاصة بمجموعتين (تجريبية وضابطة) باستخدام اختبار SAT. فإذا كان حجم الأثر المعياري d=0.10، علماً بأن كل مجموعة تشمل 20 مشاركاً وعدد الدراسات 10 ومستوى الدلالة

α=0.05 لاختبار ذي ذيل واحد. وفي ذلك نفرض أن متوسط درجات المجموعة التجريبية أو (المدربة) أكبر من متوسط درجات المجموعة الضابطة (انظر Hedges, & Pigott, 2011)

إن التباين المشترك داخل الدراسة لحجم الأثر يكون V=0.10 بنموذج الأثر الثابت. **أما بالنسبة لنموذج الآثار العشوائية؛** فسوف نحسب القوة في حالتين: عندما يكون التباين مرتفعاً وعندما يكون منخفضاً. في الحالة الأولى τ^2= 3V=0.30 و في الحالة الثانية τ^2= $\frac{1}{3}$(0.10)=0.033، k=10.

في حالة التباين المرتفع $V_{\bar{y}}^{*}(0.10 + 0.30)/10 = 0.0133$

$\lambda = (0.5\text{-}0)/\sqrt{0.04} = 2.5$

لاختبار ذي اتجاه واحد Power = 1- ϕ(1.645-2.5)

=1- ϕ(-0.855) =1 - 0.2= 0.80

وفي حالة التباين المنخفض $\lambda = (0.5\text{-}0)/\sqrt{0.0133} = 4.34$

Power = 1- ϕ(1.645-4.34)

=1- ϕ(-2.69) =1-0.003= 0.997

مما سبق يتبين لنا أن القوة تقل عندما يزداد التباين، وتزداد عندما يقل التباين.

في المثال السابق كان متوسط حجم الأثر أو الفرق المتوسط المعياري 0.10 ولكن عندما يكون متوسط حجم الأثر المعياري أكبر أو يساوي 0.2 وكان عدد المشاركين بكل مجموعة يساوي 10 وباستخدام اختبار ذي طرفين بمستوى دلالة 0.05 فإن V=0.20 وبالنسبة للتباين المرتفع يكون τ^2= 3V=3(0.20)=0.60 وبالنسبة للتباين المنخفض $\tau^2 = \frac{1}{3}V$ = $\frac{1}{3}$(0.20)=0.07، وعندما تكون k=10 فإن الدرجة المرتفعة من التباين تساوي:

$V_{\bar{y}}^{*}$= (0.2+0.6)/10 =0.08

$\lambda = (0.2\text{-}0)/\sqrt{0.08} = 0.71$

وتكون قوة الاختبار Power= 1- ϕ(1.96-0.71) + ϕ(-1.96-0.71)

=1- ϕ(1.25)+ ϕ(-2.67)

=1- 0.89 +0= 0.11

وفي حالة الدرجة المنخفضة من التباين:

$$V_{\bar{y}}^{*} = (0.2+0.07)/10 = 0.027$$

$$\lambda = (0.2-0)/\sqrt{0.0027} = 1.22$$

$$\text{Power} = 1- \phi(1.96-1.22) + \phi(-1.96-1.22)$$

$$=1- \phi(0.74)+ \phi(-3.18)$$

$$=1- 0.77+0= 0.23$$

وبصفة عامة تكون القوة منخفضة في اختبار متوسط حجم الأثر في إطار نموذج الآثار العشوائية.

مثال (4): إذا افترضـــنا أن حجم الأثر الحقيقي يختلف من دراســـة إلى أخرى، فإن **نموذج حجم الأثر العشـوائي** يكون مناسـباً، فإذا كان حجم الأثر d=0.30 وتوجد 10 دراسـات، وبكل دراسـة 25 مريضاً، فإن التباين داخل الدراسة الواحدة يساوي

$$V_{yi} = \frac{25+25}{25\ (25)} + \frac{(0.30)^2}{2\ (25+25)} = 0.0809$$

$$\lambda^{*} = \frac{0.30}{\sqrt{0.0135}} = 2.5836$$

القيمة الحرجة C_α تساوي:

$$C_\alpha = \phi[\ 1-(\alpha/2)\ ] = 1.96$$

$$\text{Power} = 1- \phi(1.96 - 2.5836) + \phi(-1.96 - 2.5836)$$

$$=1- \phi(-0.6236)+ \phi(-4.5436) = 0.7336$$

مثال(5): لتقدير قيمة عدد الدراسات التي نحتاج إليها لاستكشاف متوسط حجم أثر معين. نفرض أننا نريد حجم أثر OR=1.3 لمجموعتي المدخنين وغير المدخنين، وكانت V=0.09 وكان التباين الكبير المتوقع $\tau^2 = 3V = 3(0.09) = 0.27$؛ فإنه يمكن تقدير قيمة k من المعادلة التالية في حالة التباين المرتفع (أنظر Hedges, & Pigott, 2011)

$$0.01 = \frac{(0.09+0.27)}{K} \qquad \therefore\ k = 0.36/0.01 = 36$$

أما في حالة التباين المنخفض فإن $\tau^2 = \frac{1}{3} V = \frac{1}{3}(0.09) = 0.03$

$$0.01 = \frac{(0.09+0.3)}{K} \qquad \therefore\ k = 0.12/0.01 = 12$$

لاحظ أننا نحتاج إلى عدد أقل من الدراسات لاستكشاف حجم أثر يساوي OR=1.3 في إطار استخدام نموذج الآثار العشوائية، ويكثر هنا العدد المطلوب من الدراسات عندما يكون هناك زيادة في تباين متوسط حجم الأثر.

قوة اختبار درجة تجانس (تباين) تقديرات حجم الأثر (أنظر Ellis, 2010)

فيما يلي نتناول قوة اختبار تجانس (تباين) تقديرات حجم الأثر في إطار نموذج الآثار الثابتة، ونموذج الآثار العشوائية. وسوف يظهر عدم التجانس فيما وراء التحليل عندما يكون التباين بين الدراسات كبيراً ويتجاوز المصادفة.

- **قوة اختبار تجانس تقديرات حجم الأثر الثابتة:**

الفرض الصفري لاختبار التجانس في هذه الحالة هو: $H_0: \bar{y}_1 = \bar{y}_2 = \cdots \bar{y}_k = \bar{y}$

$$Q = \sum_{i=1}^{K} \frac{(y_i - \bar{y})^2}{V_i} = \sum_{i=1}^{K} w_i (y_i - \bar{y})^2 \qquad \text{حيث } w_i = \frac{1}{V_i}$$

وتعتمد قوة هذا الاختبار (Q) على ثلاثة عوامل، هي:

1- ارتفاع التباين داخل الدراسات والتباين بينها.
2- ارتفاع عدد الدراسات.
3- محك الدلالة أكثر تسامحاً أو أقل تحفظاً (بدلاً من 0.01 مثلاً يكون 0.05).

ويتبع إحصاء Q توزيع مربع كاي χ^2 بدرجات حرية (K-1) حيث K عدد الدراسات المتضمنة بالتحليل، ويكون الاختبار ذا اتجاه واحد عندما نرفض الفرض الصفري هو $H_a: Q \geq C_\alpha$، لاحظ أن عدم دلالة χ^2 يشير إلى أن حجوم الأثر متجانسة ولذلك يكون التقدير لحجم أثر مشترك متساوياً. ويتحدد البارامتر λ اللامركزي باستخدام المعادلة: $\lambda = \sum w_i (y_i - \bar{y})$

حيث $\bar{y}$ المتوسط الوزني أو المتوسط المشترك، و y_i هي حجم أثر الدراسة i، و w_i هو وزن الدراسة i : $\bar{y} = (\sum w_i - y_i)/\sum w_i$

ومثلما تحدثنا عن قوة اختبار دلالة متوسط حجوم الآثار، نود أن نعرف نسبة إحصاء Q التي تكون أكبر من C_α. وعندما يكون لدينا توزيع χ^2 وقيمة λ ودرجة حرية (K-1) فإننا سوف نتبع الخطوات السابقة نفسها في تقدير قيمة قوة اختبار التجانس وفقاً للدالة التالية:

$$\text{Power} = 1 - F(c_\alpha/(K-1)\,\lambda)$$

حيث إن $F(c_\alpha/(K-1))\,\lambda$ هي المساحة الأكبر من نقطة القطع c_α بتوزيع χ^2 اللامركزية

وبدرجة حرية K-1 وكذلك القيمة اللامركزية λ، ونبحث إذا ما كان هناك تجانس أم لا، وإذا عبرنا عن الانحراف بالقيمة $0.5\sqrt{V_{\bar{y}}}$ و إذا افترضنا أن حجم العينة N وعدد الدراسات K، ويمكننا تقدير قيمة التباين المشترك $V_{\bar{y}}$ لكل حجم أثر والتباين المقدر لمتوسط حجوم الأثر هذه، فإنه يمكن تقدير قيمة تباين حجم الأثر $V_{\bar{y}}$ والبارامتر اللامركزي λ كما يلي:

$$\lambda = \sum W\left(0.5\sqrt{V_{\bar{y}}}\right)^2 = K\,W\,V_{\bar{y}}(0.5)^2$$

$$\lambda = K\,\frac{1}{V_i}\,\frac{V_i}{K}\,(0.5)^2 = 0.25$$

وإذا أخذنا في الاعتبار قيماً أخرى للفرق بين متوسط الأثر وتقديرات حجم الأثر بالتوزيع مثل $V_{\bar{y}}$ (تباين مرتفع)، $V_{\bar{y}}$ (تباين منخفض) فإن قيم λ كما عرضنا سابقاً تساوي1 و4 على الترتيب.

مثال 5: تم تطبيق اختبار الـــ SAT على مجموعتين، إحداهما ضابطة والأخرى تجريبية، ويرغب الباحث في تحديد قيمة قوة اختبار تجانس تقديرات حجم الأثر لمتوسط معياري قيمته 0.5، علماً بأن العينة الإجمالية 40 مشاركاً (20 بكل مجموعة) وحصلنا على V=0.10 وكان عدد الدراسات 10 يكون $V_{\bar{y}} = \frac{0.10}{10} = 0.01$، وإذا افترضنا أن متوسط الفرق بين حجوم الأثر يساوي $0.5\sqrt{V_{\bar{y}}}$، تكون قيمة λ تساوي:$\lambda = (0.5)^2 = 0.25$ (انظر Ellis, 2010).

والقيمة الحرجة لـــ χ^2 عند c_α تساوي 16.92 بدرجات حرية $(K-1) = 9$ ومستوى دلالة 0.05، وهذا يعني أن 95% من توزيع χ^2 المركزي تقع أسفل القيمة الحرجة 16.92 ولذلك فإن قوة اختبار التجانس تساوي:

$$\text{Power} = 1 - F(c_\alpha/(K-1).\lambda) = 1 - F(16.92/9.0.25) = 0.058$$

وهذا يعني أنه لا توجد لدينا قوة لاستكشاف التجانس عندما يكون الفرق بين حجوم الأثر وحجم الأثر المتوسط مقداره. ($0.5\sqrt{V_{\bar{y}}}$=0.158) وإذا افترضنا أن الفرق بين حجوم الأثر وحجم الأثر المتوسط كبير 2V، فسوف نحصل على البارامتر اللامركزي λ الذي يساوي 4 وتكون قوة اختبار التجانس:

$$\text{Power} = 1 - F(c_\alpha/(K-1).\lambda) = 1 - F(16.92/9.4) = 0.22$$

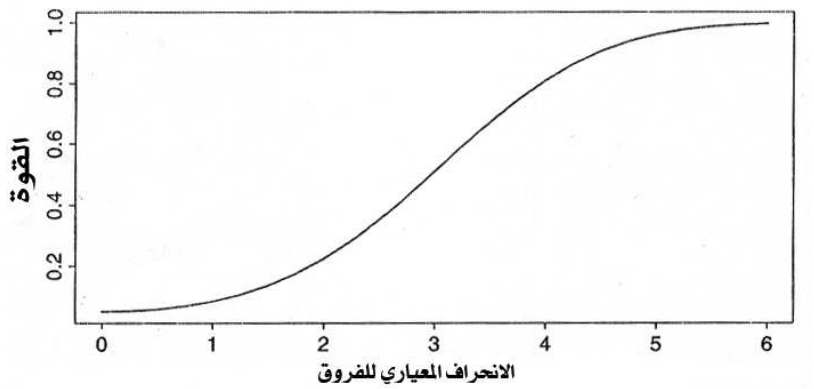
وعندما يكون الفرق بين حجوم الأثر والمتوسط (مقيسة بوحدات الانحراف المعياري) مساوياً

$3\sqrt{V_{\bar{y}}}$ فإن القوة تساوي: $\text{Power} = 1 - F(c_\alpha/(K-1).\lambda) = 0.51$

ويعرض شكل (5/6) العلاقة الطردية بين القوة والفروق بين تقديرات حجم الأثر مقيسة بوحدات الانحراف المعياري.

شكل (5/6)

العلاقة بين القوة والفروق بين تقديرات حجم الأثر مقيسة بوحدات الانحراف المعياري (انظر Ellis, 2010)



- **قوة اختبار دلالة التباين τ^2 في إطار نموذج الآثار العشوائية** (Riely, et al. 2011)

التباين بين حجوم الأثر في نموذج الآثار العشوائية يتضمن مكونين، أحدهما يرجع إلى خطأ في المعاينة والآخر هو التباين بين حجوم الأثر. ولذلك عند القيام باختبار التجانس في إطار نموذج الآثار العشوائية فإننا نهتم بمدى اختلاف توزيع تقديرات حجم الأثر عن الصفر، ولذلك يكون الفرض الصفري في هذه الحالة $H_0: \tau^2 = 0$ والفرض البديل هو $\tau^2 \neq 0$. وهذا الاختبار في نموذج الآثار العشوائية يأخذ الصيغة المشابهة لاختبار التجانس Q في حالة نموذج الآثار الثابتة، الذي يأخذ الصيغة:

$$Q = \sum_{i=1}^{K} w_i^*(y_i - \bar{y})^2 \qquad w^* = \frac{1}{(V+\tau^2)}$$

ثم مقارنة قيمة Q بقيمة χ^2 عند درجات حرية $(K-1)$ ومستوى الدلالة المحدد. وعندما تزيد قيمة Q عن القيمة الحرجة بتوزيع χ^2 المركزي ودرجة حرية $(K-1)$، فإننا نرفض الفرض الصفري ونقرر أن التباين مختلف عن الصفر.

وعندما يكون الفرض الصفري صحيحاً فإن هذا يعني أن $H_0: \tau^2 = 0$

ولتقدير قيمة قوة اختبار τ^2 فإننا نحتاج إلى تقدير قيمة القيم التالية:

$$\mu_Q = C\tau^2 + (K-1)$$

$$C = \sum w_i - \left[\frac{\sum {w_i}^2}{\sum w_i}\right]$$

حيث τ^2 قيمة التباين التي نختبر دلالتها أو اختلافها عن الصفر ويحسب تباين Q أو σ_Q^2 باستخدام أوزان الآثار الثابتة التي تعطى بالمعادلة:

$${\sigma_Q}^2 = 2(K-1) + 4\left[\sum w_i - \frac{\sum {w_i}^2}{\sum w_i}\right]\tau^2 + 2\left[\sum {w_i}^2 - 2\frac{\sum {w_i}^3}{\sum w_i} + \frac{(\sum {w_i}^2)^2}{(\sum w_i)^2}\right]\tau^4$$

ويمكننا بعد ذلك استخدام توزيع χ^2 بدرجة حرية غير صحيحة، وهو تقريب لتوزيع جاما بمتوسط وتباين يعطيان بالمعادلتين السابقتين، ولتقدير قيمة التقريب المصحح نحتاج إلى تقدير قيمة كميات تعتمد على المتوسط والانحراف المعياري السابقين، وهذه الكميات r والتي تتحدد بالمعادلتين التاليتين:

$$r = \sigma_Q^2/2\mu_Q$$

$$s = 2(\mu_Q)^2/\sigma_Q^2$$

وتحسب القوة باستخدام المعادلة التالية:

$$Power = F(c_\alpha/(r/s),0)$$

حيث $F(c_\alpha/(r/s),0)$ هي دالة توزيع متجمع لـ χ^2 بدرجة حرية S (قيمة كسرية أو غير صحيحة) و c_α هي القيمة الحرجة لتوزيع χ^2 ودرجة حرية $(K-1)$ وتحسب المساحة أسفل المنحنى χ^2 وأكبر من القيمة الحرجة c_α وهذه تعطي قوة الاختبار.

اختبار قيم البارامترات التي نحتاج إليها لتقدير قيمة مكون التباين بنموذج الأثر العشوائية:

توجد علاقة بين I^2 و τ^2 هي: $I^2 = \frac{(\tau^2)}{(\tau^2+V)}$

ويحدد هيجنز وطمبسون (Higgins & Thompson, 2002) ثلاثة مستويات لـ I^2

هي: صغيرة $I^2 = 25\%$، متوسطة 50%، مرتفع 75%، كما توجد علاقة بين V، وهي القيمة المشتركة لتباين حجم الأثر داخل الدراسة، ومستويات τ^2، هي: $\tau^2 = V, \quad \tau^2 = V, \quad \tau^2 = \frac{1}{3}V$ 3V منخفضة، متوسطة، مرتفعة للتباين على الترتيب.

مثال 6: تقدير قيمة قوة اختبار التباين τ^2 في المثال 5

يمكن الحصول على الانحراف المعياري المشترك V من خلال المعادلة:

$$V = \frac{20+20}{20(20)} + \frac{(0.2)^2}{2(20+20)} = 0.10$$

ومن خلال قيمة V يمكن تقدير قيمة قوة اختبار τ^2 في المستويات التالية:

عندما يكون التباين صغيراً فإن $\tau^2 = \frac{1}{3}V = 0.033$

وعندما يكون التباين كبيراً فإن $\tau^2 = 3V = 3(0.10) = 0.30$

وحيث إن عدد الدراسات K=10 فإن:

$$C = \sum w_i - \left[\frac{\sum w_i^2}{\sum w_i}\right] = 10w - \frac{10w^2}{10w} = \left(\frac{10}{0.10}\right) - \left[\left(\frac{1}{0.1}\right)^2 / \left(\frac{1}{0.1}\right)\right] = 100 - 10 = 90$$

وبافتراض أن التباين منخفض وعدد الدراسات 10 و C=90 فإن μ_Q تساوي:

$$\mu_Q = 90(0.033) + (10-1) = 2.97 + 9 = 11.9$$

ويمكن تقدير قيمة ${\sigma_Q}^2$ كما يلي:

$${\sigma_Q}^2 = 2(10-1) + 4\left[\sum\left(\frac{1}{0.10}\right) - \frac{\sum\left(\frac{1}{0.10}\right)^2}{\sum\left(\frac{1}{0.10}\right)}\right](0.033)^2$$

$$+2\left[\left(\sum\left(\frac{1}{0.10}\right)^2 - 2\left[\frac{\sum\left(\frac{1}{0.10}\right)^3}{\sum\left(\frac{1}{0.10}\right)}\right]\right) + \left[\frac{\left(\sum\left(\frac{1}{0.10}\right)^2\right)^2}{\left(\sum\left(\frac{1}{0.10}\right)\right)^2}\right](0.033)^4\right]$$

$$= 18 + 0.39 + 2(1000 - 200 + 100)(0.001)^2$$

$$= 18.39 + 0.002 = 18.392$$

ثم نحدد قيمة r وs:

$$r = 18.392 / (2 \times 11.97) = 0.77$$

$$s = 2(11.97)2 / 18.392 = 15.58$$

ومن خلال قيمة r وs فإن قوة اختبار مكون التباين يساوي $\tau^2 = 0.033$ أو درجة منخفضة من التباين في هذه الحالة، وعندما تكون القيمة الحرجة لـــ χ^2 بدرجة حرية $(K-1)$و$\alpha = 0.05$ فان $C_\alpha = 16.92$؛ ومن ثم تكون قوة اختبار مكون التباين تساوي:

$$\text{Power} = F(c_\alpha/(rs),0)$$

$$\text{Power} = F(16.92/(0.77 \times 15.58),0) = 0.13$$

وهذا يعني أن القوة التي لدينا ضـــعيفة عند اتخاذ قرار بشـــأن دلالة $\tau^2 = 0.033$ و أنها تختلف عن الصـــفر. وفي حالة التباين الكبير فإن $\tau^2 = 0.30$ و $\mu_Q = 36$ و$\sigma_Q{}^2 = 64.98$ و r =0.90 و s =20.81 فان الدرجة الحرجة تساوي $C_\alpha = 16.92$ وتكون القوة:

$$\text{Power} = F(c_\alpha/(r\,s),0)$$

$$\text{Power} = F(16.92/(0.90 \times 20.81),0) = 0.59$$

وهذا يعني أن لدينا قوة أكبر لكي تكون $\tau^2 = 0.30$ دالة وتختلف عن الصـــفر، وذلك بالمقارنة بالحالة السابقة عن التباين المنخفض.

تحليل القوة لنموذج المتغير التصنيفي لحجم الأثر (انظر Jackson, & Turner, 2017)

فيما يلي طرق تقدير قيمة تباين المتغير التصنيفي لحجم الأثر، وهذه النماذج تشبه تحليل التباين الأحادي لتقديرات حجم الأثر، في إطار نماذج الأثر الثابتة والعشوائية، وكذلك قوة ما وراء الانحدار التي تتطلب معرفة قيم المنبئات بكل دراسة في النموذج.

- **قوة نموذج التصنيف لتقديرات حجم الأثر الثابتة**

يتم هذا عندما يهتم الباحث بمقارنة تقديرات حجم الأثر بين مجموعات الدراســـة، التي تتشكل من عدة مستويات للمتغير الفئوي أو التصنيفي ولاحقاً نقدم إجراءات تقدير قيمة إحصاءات تحليل التباين الأحادي في إطار نموذج الآثار الثابتة مع تحديد إجراءات تقدير قيمة قوة هذا الاختبار.

- **قوة اختبار تحليل التباين في اتجاه واحد لتقديرات حجم الآثار الثابتة**

ويرمز لفئة تقديرات حجم الأثر بـــ $(y_1.y_2.\ldots.y_i)$ وعددها بـــ K أو لعدد الأصناف K، التي تعرف بالمتغير الوســـيط؛ مثل مســـتويات الفرق الدراســـية (ابتدائي، متوســـط، ثانوي) وللدراســـة m_i. في تحليل التباين الأحادي لتقديرات حجم الأثر نختبر هل متوسـطات المجموعات

متساوية؟ وما مقدار التجانس داخل الدراسات؟، لذلك سوف نحسب قيم إحصاء Q، وهو يشمل التجانس بين مجموعات (Q_B) والتجانس داخل المجموعات (Q_w).

$$Q_w = \sum\sum w_{ij}\left(y_{ij} - \bar{y}_i\right)$$

$$Q_w = Q_{w_1} + Q_{w_2} + \ldots\ldots Q_{w_p}$$

$$Q_B = \sum_{i=1}^{K} w_i(y_i - \bar{y})^2$$

$$Q_T = Q_B + Q_w$$

حيث إن Q هو مجموع التباين الكلي عبر كل تقديرات حجم الأثر، وسوف نبين تقدير قيمة قوة اختبار التباين (التجانس) لـ Q_w وQ_B

- **قوة اختبار التجانس أو التباين بين المجموعات (Q_B) بنموذج الآثار الثابتة:**

يركز تحليل التباين على اختبار الفرض الصفري، وهو تساوي متوسط تقديرات حجم الأثر بالمجموعات K

$$H_0: \bar{y}_1 = \bar{y}_2 = \cdots = \bar{y}_P$$

ولاختبار الفرض الصفري نحسب التجانس بين المجموعات باستخدام المعادلة

$$Q_B = \sum_{i=1}^{K} w_i(\bar{y}_i - \bar{y})^2$$

حيث $\bar{y}_i$ متوسط حجم آثار المجموعة (i)، و$\bar{y}$ المتوسط الكلي لتقديرات حجم الأثر بجميع المجموعات، w_i هو وزن المجموعة (i)

$$w_i = \sum_{j=1}^{m_i} w_{ij}$$

حيث عدد الدراسات $j = 12 \ldots . m_i$ $\quad$ $i = 12 \ldots . p$ عدد المجموعات

$$\bar{y}_i = \frac{\left(\sum_{j=1}^{m_i} w_{ij} \cdot y_{ij}\right)}{\left(\sum_{j=1}^{m_i} w_{ij}\right)}$$

$$\bar{y} = \frac{\left(\sum_{i=1}^{p}\sum_{j=1}^{m_i} w_{ij} \cdot y_{ij}\right)}{\left(\sum_{i=1}^{p}\sum_{j=1}^{m_i} w_{ij}\right)}$$

الفرض الصفري عندما يكون صحيحاً فإن متوسطات المجموعات تكون متساوية، وحيث إن قيم Q_B تتبع توزيع χ^2 بدرجة حرية (K-1) فإن ذلك لا يتحقق، وعندما تكون $Q_B > c_\alpha$، حيث c_α تساوي $100(1-\alpha)$ في المائة. وهذا يحدث عندما يوجد على الأقل متوسط مجموعة يختلف عن متوسط مجموعة أخرى. إن Q_B تتبع توزيع χ^2 اللامركزي بدرجة حرية (K-1)، وأن البارامتر اللامركزي λ_B الذي يتحدد باستخدام المعادلة:

$$\lambda_B = \sum w_i(y_i - \bar{y})$$

وأن القوة تساوي: $P = F(c_\alpha / (P-1), \lambda_B)$

حيث إن $F(c_\alpha / (K-1).\lambda_B)$ هي التوزيع المتجمع لـχ^2 بدرجة حرية $(K-1)$ و λ_Bهي قيمة البارامتر غير المركزي.

خطوات تقدير قيمة قوة اختبار تحليل التباين في اتجاه واحد للآثار الثابتة (Jackson, & Turner, 2017)

1. قرار بشأن القيمة الحرجة c_α.
2. قرار بشأن قيمة الفرق بين متوسط تقديرات حجم الأثر بالمجموعات؛ مثال ذلك الفرق المتوسط المعياري (g, d) مثلاً، أو الفرق في معاملات الارتباط.
3. حساب قيم متوسطات المجموعات $\bar{y}_1.\bar{y}_2.\ldots\bar{y}_K$
4. قرار بشأن عدد الدراسات داخل كل مجموعة $m_1.m_2\ldots.m_i$
5. احسب w_{ij} أو القيمة المشتركة لأوزان تقديرات حجم الأثر كلها، وأعط القيم المعيارية لتقديرات حجم الأثر داخل الدراسة.

مثال 7: لتقدير قوة اختبار التجانس أو التباين بين المجموعات في إطار نموذج الآثار الثابتة، أجرى سيرين (Sirin, 2005) دراسة استخدم فيها ما وراء التحليل، ومعاملات الارتباط بين الحالة الاجتماعية والاقتصادية للطالب وبين تحصيله الأكاديمي. وطبقها على عينة من الطلاب مصنفة إلى ثلاثة مستويات من الفرق، هي: الابتدائي والمتوسط والمدرسة العليا، ونفرض أننا نريد أن يكون لما وراء التحليل قوة كافية لاستكشاف الفرق في متوسطات تقديرات حجم الآثار (انظر Ellis, 2010)

لنفترض أن متوسط تقديرات حجم الأثر بالمراحل الثلاث هي:

$$\bar{y}_1 = 0.0 \qquad \bar{y}_2 = 0.10 \qquad \bar{y}_3 = 0.5$$

وباستخدام جدول فيشر (Z_r) فإن هذه القيم تساوي على الترتيب

$$\bar{y}_{21} = 0.0 \qquad \bar{y}_{22} = 0.10 \qquad \bar{y}_{23} = 0.55$$

ونفرض أن لدينا خمس دراسات بكل مجموعة من المجموعات الثلاث، فإنه يكون لدينا 15 دراسة بتلك المجموعات، ولتقدير قيمة البارامتر اللامركزي λ_B فنحن في حاجة إلى قيمة W_{ij} وقيم فيشر Z_r.

$$W_{ij} = 1/\ V_{ij} = (n_{ij} - 3) = (N - 3)$$

حيث: $V_{ij} = 1/(n_{ij} - 3)$ $\qquad V_{ij} = 1/\ W_{ij}$

$$W_i^* = \sum_{j=1}^{5}(15 - 3) = \sum_{j=1}^{5}(12) = 60$$

والمتوسط الكلي لتقديرات حجم الأثر يساوي: $y_i^* = (0.0 + 0.10 + 0.55)/3 = 0.22$

وقد تم ذلك لأن تقديرات حجم الأثر لها الوزن نفسه والأعداد نفسها للدراسات داخل كل دراسة، أما إذا اختلفت أوزان حجوم الأثر واختلفت أعداد حجوم الأثر داخل المجموعات، فإننا نستخدم المتوسط الموزون. ويمكن تقدير قيمة البارامتر غير المركزي λ_B كما يلي:

$$\lambda_B = \sum w_{ij}^* (\bar{y}_{ij} - \bar{y}_i^*)^2$$

$$= 60(0.0 - 0.22)^2 + 60(0.10 - 0.22)^2 + 60(0.55 - 0.22)^2 = 10.302$$

والتوزيع المركزي لـــ χ^2 بدرجة حرية (K-1)= 3-1=2 تناظر الدرجة الحرجة 5.99 أو C_α عند مستوى دلالة 0.05 وقوة اختبار الفرض الصفري $H_o: \bar{y}_{21} = \bar{y}_{22} = \bar{y}_{23}$ تتحدد باستخدام المعادلة.

$$P = 1 - F(5.99/2, 10.302) = 1 - 0.17 = 0.83$$

- **قوة اختبار التجانس داخل المجموعة Q_w في إطار الآثار الثابتة** (Hedges & Pigott, 2004)

إذا وجد الباحث اختلافاً بين متوسطات تقديرات حجم الأثر بين المجموعات فإنه يوجد سؤال ثانٍ هو: هل تقديرات حجم الأثر داخل المجموعات متجانسة؟ وإذا اختلفت تقديرات حجم الأثر بين المجموعات فإننا في حاجة إلى معرفة هل تقديرات حجم الأثر داخل المجموعات لها أثر مشترك Common effect؟ ويشير هيدجز وبيجو (Hedges & Pigott, 2004) إلى اختبار التجانس

داخل المجموعات؛ مثل اختبار المطابقة لنموذج الآثار الثابتة، وبكلمات أخرى فإن نموذج تحليل التباين يجيب عن السؤال: هل تقديرات حجم الأثر داخل المجموعات متجانسة؟ ومن ثم نقدر قيمة المتوسط المشترك، ويمكن صياغة الفرض الصفري على النحو الآتي:

$$H_o: \bar{y}_{ij} = \bar{y}_i^* \qquad i = 1,2,\ldots\ldots,P \qquad j = 1,2,\ldots\ldots,m_i$$

ويكون الفرض البديل هو أنه يوجد حجم أثر واحد على الأقل في المجموعة (i)، حيث يوجد m_i تقدير حجم الأثر داخل المجموعة (i) واختبار التجانس الكلي داخل المجموعات يتحدد باستخدام المعادلة:

$$Q_w = \sum_{i=1}^{K} \sum_{j=1}^{m_i} w_{ij} (\bar{y}_{ij} - \bar{y}_1)^2 = \sum_{i=1}^{K} Q_{wi}$$

لاحظ أن هذا المجموع يمكن أن يكتب كمجموع التجانس الإحصائي داخل المجموعات Q_{wi} عندما تكون كل مجموعة متجانسة، كذلك فإن Q_w تتوزع طبقاً لـ χ^2 المركزية بدرجة حرية (عدد حجوم الأثر - عدد المجموعات أو الأصناف). ونحن نرفض الفرض الصفري عندما $Q_w > C_\alpha$ بتوزيع χ^2 بدرجة حرية (k-p). ورفض الفرض الصفري يشير إلى أن إحدى المجموعات على الأقل مختلفة جوهرياً مع الأخرى، وهذا يعني أنه يوجد على الأقل حجم أثر واحد يختلف جوهرياً عن متوسط المجموعة.

وعندما يكون الفرض الصفري زائفاً فإن Q_w تتبع توزيع χ^2 غير المركزي بدرجة حرية (k-p) أو (عدد حجوم الأثر - عدد المجموعات أو الأصناف)، ويتحدد البارامتر المركزي λ_w بالمعادلة:

$$\lambda_w = \sum_{i=1}^{p} \sum_{j=1}^{m_i} w_{ij} (\bar{y}_{ij} - \bar{y}_1^*)^2$$

كما تتحدد القوة من خلال المعادلة:

$$P_{Qw} = 1 - F(C_\alpha / (k\text{-}p) , \lambda_w)$$

حيث $F(C_\alpha / (k\text{-}p) , \lambda_w)$ هو التوزيع المتجمع لـ χ^2 اللامركزي بدرجة حرية (K-P)، وكذلك البارامتر اللامركزي λ_w، والقيمة C_α هي القيمة الحرجة المرغوب فيها لـ χ^2 المركزية.

مثال 8: لتقدير قوة اختبار التجانس داخل المجموعات في إطار نموذج الأثر الثابت، دعونا نرجع للمثال 7، حيث يوجد لدينا ثلاث مجموعات P=3 هي المدرسة الابتدائية والوسطى والعليا وتتضمن كل مجموعة 5، ولذلك يكون العدد الكلي لحجوم أثر للدراسات k=15 ومتوسط قيمة Z_r لفيشر بكل مجموعة على الترتيب $(\bar{y}_1 = 0.0, \bar{y}_2 = 0.10, \bar{y}_3 = 0.55)$، وحجم العينة داخل المجموعة

n=15 ولذلك لدينا تباين مشـــترك لحجم الأثر يســـاوي V=1/(n-3) أو V=1/(15-3) ومن ثم يكون الوزن مساوياً W=[1/(1/12)]=12 وبتحديد حجم العينة يمكننا تقدير قيمة التباين (والخطأ المعياري) لمتوسط تقديرات حجم الأثر على النحو الآتي:

$$V_i^* = \frac{1}{\sum_{j=1}^{5} w} = \frac{1}{\sum_{j=1}^{5} 12} = \frac{1}{60} = 0.017$$

الانحراف المعياري $\sqrt{V_i^*} = \sqrt{0.017}$ =0.13

ويمكننا تقدير قيمة البارامتر المركزي λ_w على النحو الآتي:

$$\lambda_w = \sum_{i=1}^{K} \sum_{j=1}^{m_i} w_{ij} \, (\bar{y}_{ij} - \bar{y}_1^*)^2$$

$$\lambda_w = \sum_{j=1}^{5} 12(0.13)^2 + \sum_{j=1}^{5} 12(0.13)^2 + \sum_{j=1}^{5} 12(4 \text{ x} 0.13)^2$$

$$\lambda_w = 5(12)\ (0.017) + 5(12)\ (0.017) + 5(12)\ (0.52)^2$$

$$= 60(0.017) + 60(0.017) + 60(0.27) = 1.02 + 1.02 + 16.20 = 18.24$$

من الحســـابات الســـابقة، ســـوف نبدل الفرق بين تقديرات حجم الأثر الفردية للدراســـات داخل المجموعة ومتوســـط مجموعاتها $(\bar{y}_{ij} - \bar{y}_1^*)^2$ باقتراح عدد الانحرافات المعيارية $(\sqrt{V_i^*})$ بين المتوسطات. ولتقدير قيمة القوة، فإننا في حاجة إلى تحديد C_α أو الدرجة الحرجة لــ χ^2 المركزية عند درجة حرية (k-p) أو (3-15) وهي تساوي 21.03 وتحسب القوة من المعادلة

$$P = 1 - F(C_\alpha / (k-p) , \lambda_w) = 1 - F(21.03/12 , 18.24) = 1 - 0.18 = 0.82$$

وهذا يعني أن لدينا قوة كافية.

نماذج تحليل التباين في اتجاه واحد لتقديرات حجم الآثار العشوائية (Jackson, & Turner, 2017)

في نموذج الآثار العشـوائية، نفترض أن تقديرات حجم الأثر بكل دراسـة تكون عشـوائية؛ ولذلك فإن متوسطات تقديرات حجم الأثر بالمجموعات تختلف، ويرجع ذلك إلى التباين الأساسي الذي يرمز له بالرمز τ^2، وأما الذي يرجع إلى تباين المعاينة داخل الدراسة؛ فترمز له بالرمز (V_i)

وإذا افترضـنا المكون نفسـه للتباين داخل الدراسـات (v_i)، فسـوف نركز على اختبار التباين بين المجموعات فقط، وكما في مثال ســـابق عدد المجموعات (P) وهي تشـــمل عدد الدراســـات K حيث $(k=m_1+ m_2+\ldots\ldots+m_p)$ وحجم الأثر في الدراســـة (j)بالمجموعة (i)

وسوف يشار إليها بالرمز y_{ij} وتباين V_{ij}. وتباين كل حجم أثر يتضمن مكونين، أحدهما يرجع إلى خطأ المعاينة بالدراسات الفردية (الداخلي) وسوف نرمز له بالرمز V_{ij}، والآخر يرجع إلى مكون التباين τ^2 وتباين تقديرات حجم الأثر العشوائية y_{ij}^* وهذا يعني أن $(V_{ij}^* = V_{ij} + \tau^2)$.

قوة اختبار التجانس بين المجموعات في إطار نموذج الآثار العشوائية (Hedges & Pigott, 2004)

إن متوسطات مجموعة الآثار العشوائية تختلف؛ ولذلك يمكن كتابة الفرض الصفري على النحو الآتي: تساوي متوسطات المجموعات $H_o: \bar{y}_1 = \bar{y}_2 = \cdots \bar{y}_P$

حيث y_i هي المتوسطات العشوائية (حيث i=1,2,….,P) بالمجموعات P، ونختبر هذا الفرض باستخدام تجانس الآثار العشوائية بين المجموعات، الذي يتحدد باستخدام المعادلة

$$Q_B = \sum_{i=1}^{p} w_{ij}^* (y_i - \bar{y}^*)^2$$

حيث w_{ij}^* هو مجموع الأوزان بالمجموعة (i)، y_i هو متوسط حجم الأثر بالمجموعة (i)، $\bar{y}^*$ هو المتوسط العام لكل تقديرات حجم الأثر العشوائية، وهذه القيم يمكن الحصول عليها باستخدام المعادلات التالية:

$$w_{ij}^* = \sum_{j=1}^{m_i} w_{ij} \quad i=1,2,\ldots..,p \;\&\; j=1,2,\ldots\ldots, m_i$$

$$\bar{y}_{ij}^* = \left(\sum_{j=1}^{m_i} (w_{ij}^*)(y_{ij}^*)\right) / \sum_{j=1}^{m_i} (w_{ij}^*)$$

$$\bar{y}_T^* = \frac{\left[\sum_{i=1}^{p} \sum_{j=1}^{m_i} (w_{ij}^*)(y_{ij}^*)\right]}{\left[\sum_{i=1}^{p} (w_{ij}^*)(y_{ij}^*)\right]}$$

عندما يكون الفرض الصفري حقيقياً؛ أي تتساوى كل المتوسطات، كما يتبع Q_B^* توزيع χ^2 بدرجة حرية (p-1) وعندما تتعدى قيمة Q_B^* القيمة الحرجة C_α، وفي هذه الحالة نفترض أنه يوجد على الأقل متوسط واحد يختلف عن باقي المتوسطات في إطار نموذج الآثار العشوائية، وفي هذه الحالة نرفض الفرض الصفري. ويلاحظ أن Q_B^* يتبع توزيع χ^2 اللامركزي بدرجة حرية (p-1) كما يحسب البارامتر المركزي λ_B^* باستخدام المعادلة:

$$\lambda_B^* = \sum_{i=1}^{p} w_i^* (y_i - \bar{y}_T)^2$$

وتتحدد قوة اختبار الفروق بين متوســطات المجموعات في إطار نموذج الأثر الثابت، باستخدام المعادلة:

$$\text{القوة } \text{Power} = 1\text{-}F(C_\alpha/(1\text{-}p)\ ,\ \lambda_B^*)$$

وما سبق هي دالة التوزيع المتجمع لـــχ^2 اللامركزية بدرجات حرية (p-1) للنقطة C_α وλ_B^* هي البارامتر اللامركزي.

خطوات تقدير قيمة قوة النموذج التصنيفي للآثار العشوائية (Hedges & Pigott, 2004)

1) تحديد القيمة الحرجة للدلالة الإحصائية (C_α).
2) اتخاذ قرار بشـــأن قيمة الفرق بين متوســـطات المجموعات مصـــفوفة حجوم الأثر المســـتخدمة، مثل فروق المتوســـط المعيارية، اتخاذ قرار بشـــأن عدد الانحرافات المعيـارية التي تكون مهمة جوهريا بـالنســـبة لمعامل الارتبـاط وقرار الفرق في معاملات الارتباط المهمة.
3) إنشاء قيم متوسطات المجموعات ($\theta_1, \theta_2, \ldots, \theta_i$) المناظرة للفروق بين متوسطات المجموعات في الخطوة 2.
4) تقدير عدد الدراسات داخل كل مجموعة أو تصنيف $m_1, m_2, \ldots, m_p$.
5) تقدير قيمة V القيمة المشتركة لتباين المعاينة لكل حجم أثر، أعطى القيم المعيارية أو حجم عينة الدراسة الداخلي.
6) تقدير قيمة τ_x^2 عندما يكون مســـتوى التباين المعتمد على V بدرجة مرتفعة وهو يساوي ($\tau^2 = 3V$)، وعندما يكون درجة متوسطة ($\tau^2 = V$)، بدرجة منخفضة ($\tau^2 = \frac{1}{3}V$).
7) احسب قيم الوزن المشترك $W = 1/(V+\tau^2)$.

قوة اختبار التجانس أو التباين بين المجموعات في إطار نماذج الآثار العشوائية

لتقدير قيمة قوة لاختبار متوســط تقديرات حجم الأثر بين المجموعات، نحتاج إلى اتخاذ قرار بشـأن كمية التباين التي نريدها من الدراسـات. وهذا يعني أننا نحتاج إلى قيمة τ^2 باسـتخدام قاعدة تعتمد على هيجنز وطومسون (Higgins, & Thompson, 2002) وفيها نختار قيمة τ^2 التي تعتمد V؛ حيث يأخذ القيم الثلاث الآتية: تباين مرتفع ($\tau^2 = 3V$) وهي تناظر $I^2 = 75\%$،

تبـاين متوســط (τ^2=V) وهي تنـاظر $I^2 = 50\%$، تبـاين منخفض (τ^2=$\frac{1}{3}$ V) وهي تناظر$I^2 = 25\%$.

مثال 9: لتقدير قوة اختبار التباين بين المجموعات في إطار نماذج الآثار العشــوائية، وبالرجوع للمثال 8، حيث يوجد 3 مجموعات أو مراحل دراســية، ومتوســطات تقديرات حجم الأثر بها $\bar{y}_1 = 0.0$ $\bar{y}_3 = 0.50$, $\bar{y}_2 = 0.1$, وباسـتخدام تحويل فيشـر لمعاملات الاترباط فإن هذه المتوســطات تناظر $\bar{y}_{12} = 0.0$ $\bar{y}_{32} = 0.55$ $\bar{y}_{22} = 0.1$, , على الترتيب. ولنفرض أن لدينا 5 دراســات بكل مجموعة، فإن ذلك يعني أن عدد الدراســات الكلية بالمجموعات الثلاث k=5(3) وإذا كان حجم العينة بكل مجموعة 15 ولتقدير قيمة البارامتر اللامركزي λ_B^* فإننا نحتاج في البداية إلى تقديم قيمة مكون التباين المشــترك τ^2 بين الدراســات معتمدين على قيمة التباين المشترك V (Hedges & Pigott, 2004)

وطبقاً لتحويل فيشر Z_r لمعاملات الارتباط فإن القيمة المشتركة لتباين المعاينة داخل المجموعات في هذا المثال يساوي V=1/(n-3)=1/12 أو (0.083) وفي حالة درجة مرتفعة من التباين الذي يساوي τ^2= 3V =3(0.083)=0.25،وفي حالة الدرجة المتوسطة للتباين فإن τ^2= V =0.083، وفي حالة الدرجة المنخفضــة للتباين فإن τ^2=$\frac{1}{3}$ V= $\frac{1}{3}$(0.083)= 0.028، ويرمز لتباين تقديرات حجم الأثر العشوائية V وهي تساوي V^2=0.083+τ^2.

وعندما نفرض درجة منخفضــة من التباين، فإن قيمة وزن متوســطات تقديرات حجم الأثر العشوائية تعطى بالمعادلة:

$$W_i^* = 1/\sum_{i=1}^{5} \frac{1}{V+\tau^2} = 1/\sum_{i=1}^{5} \frac{1}{(0.083+0.028)}$$

$$= 1/\left[\sum_{i=1}^{5} 1/(0.111)\right] = 1/\left(\sum_{i=1}^{5} 9.01\right) = 1/5(9.01) = 0.022$$

وبتحديد قيمة W_i^* يمكننا تقدير قيمة البارامتر اللامركزي λ_B^* من خلال المعادلة التالية:

$$\lambda_B^* = \sum_{i=1}^{3} w_i^* (y_i - \overline{y^*})^2$$

$$= 0.022\ (0\text{-}0.22)^2 + 0.022\ (0.1\text{-}0.22)^2 + 0.022(0.55\text{-}0.022)^2$$

$$= 0.022(0.048) + 0.022(0.014) + 0.022(0.11) = 0.004$$

وقوة اختبار التجانس بين المجموعات تتحدد باستخدام الدالة التالية:

$Power = 1-F(C_\alpha/(p-1) , \lambda_B^*)$ بدرجة حرية P-1=2

وحيث أن C_α بتوزيع χ^2 المركزي $C_\alpha= C_{0.05}= 5.99$

القوة في حالة التباين المنخفض

$Power = 1- F(5.99/2, 0.004) = 1-0.95 =0.05$

وفي حالة التباين المرتفع فإن قيم V^* و w_i^* و λ_B^* في إطار نموذج الآثار العشوائية تساوي:

$V^*= 0.083+0.25=0.33$

$W_i^* = 1/ \sum_{i=1}^{5} 1/(V + \tau^2) =1/ 5(3.033) = 1/(15.01) =0.067$

وباستخدام قيمة w_i^* يمكننا تقدير قيمة البارامتر اللامركزي λ_B^* كما يلي

$\lambda_B^* = \sum_{i=1}^{3} w_i^* (y_i - \bar{y}^*)^2$

$= 0.067 (0 – 0.22)^2 +0.067 (0.1- 0.22)^2 +0.067 (0.55- 0.22)^2$

$= 0.067(0.048) +0.067(0.14) +0.067(0.11) = 0.012$

ومن توزيع χ^2 عند درجات حرية 2 ومستوى دلالة 0.05 و $C_\alpha= 5.99$

$Power = 1- F(5.99/2, 0.012)$

المراجع

أبو حطب، فؤاد، وصادق، آمال. (1991). *مناهج البحث وطرق التحليل الإحصائي في العلوم النفسية والتربوية والاجتماعية*. مكتبة الأنجلو المصرية.

أبو علام، رجاء محمد. (1993). *علم النفس التربوي*. دار القلم.

أبو علام، رجاء محمد. (2006). حجم أثر المعالجات التجريبية ودلالة الدلالة الإحصائية. *المجلة التربوية، 78*، 1 - 153.

أبو هاشم، السيد محمد. (2004). مؤشرات التحليل البعدي لبحوث فعالية الذات في ضوء نظرية باندورا. *مجلة كلية التربية، جامعة الملك سعود، 238*، 1 - 88.

الخضر، عثمان. (2022). المراجعات المُمنهجة في بحوث علم النفس في العالم العربي. *مجلة العلوم الاجتماعية* [قيد الطبع].

سكران، السيد عبد الدايم. (2006). ما وراء التحليل كمنهج وصفي تحليلي لتجميع نتائج البحوث وتكاملها في مجال التربية وعلم النفس. *مجلة كلية التربية بالزقازيق. 53*، 1 - 38.

الشربيني، زكريا. (2011). *التحليل الإحصائي الماورائي: دراسة على بحوث عن الفعالية الذاتية في ضوء بعض المتغيرات*. الأنجلو المصرية.

شريف، نادية محمود. (1993). المنهج البعدي للتحليل كأسلوب لمتابعة نتائج البحوث والدراسات النفسية والتربوية. *المجلة المصرية للتقويم التربوي، 1*، 155 - 186.

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

الكناني، ممدوح (2020). *الإحصاء النفسي والتربوي والجداول الإحصائية لاختبار الفروض*. مكتبة الفلاح.

عبد الحميد، محمد. (1987). أسلوب التحليل البعدي لنتائج البحوث والدراسات السابقة. حولية كلية التربية، *جامعة قطر، 5*، 336 - 391.

عودة، أحمد، والخليلي، خليل. (1998). الإحصاء للباحث في التربية والعلوم الإنسانية. *دار الفكر للنشر والتوزيع*.

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

نصار، يحيى حياتي. (2006). استخدام حجم الأثر لفحص الدلالة العملية للنتائج في الدراسات الكمية. *مجلة العلوم التربوية والنفسية، 7*، 45 - 54.

Aert, R. C. M, van Assen, M. A. L. M., & Viechtbauer, W. (2019). Statistical properties of methods based on the Q-statistic for constructing a confidence interval for the between-study variance in meta-analysis. *Research Synthesis Methods, 10*(2), 225 - 239.

Borenstein, M. (2000). The shift from significance testing to effect size estimation. In A.S. Bellack & M. Hersen (eds.), *Comprehensive Clinical Psychology* (Volume 3, pp. 313 - 349). Pergamon.

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

Brockhaus, A.C., Bender, R., & Skipka, G. (2015). The Peto odds ratio viewed as a new effect measure: derivation and application. In: *Filtering the information overload for better decisions.* Abstracts of the 23rd Cochrane Colloquium; 2015 3 - 7 Oct; Vienna, Austria. John Wiley & Sons.

Carlson, M., Marcus-Newhall, A., & Miller, N. (1990). Effects of situational aggression cues: A quantitative review. *Journal of Personality and Social Psychology, 58*(4), 622 - 633.

Cohen, J. (1977). Eta-squared and partial eta-squared in fixed factor ANOVA designs. *Educational and Psychological Measurement, 33,* 107 - 112.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed)*. Lawrence Erlbaum Associates.

Cooper, H., & Findley, M. (1982). Expected effect sizes: Estimates for statistical power analysis in social psychology. *Personality and Social Psychology Bulletin, 8,* 168 - 173.

Cortina, J. M., & Nouri, H. (2000). *Effect size for ANOVA designs.* Sage.

Cumming, G. (2012). Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis. Routledge/Taylor & Francis Group.

DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled clinical trials, 7*(3), 177 - 188.

Dickersin, K., & Min, Y. I. (1993). Publication bias: the problem that won't go away. *Annals of the New York Academy of Sciences, 703,* 135 - 148.

Daniels, K. (2019). Guidance on conducting and reviewing systematic reviews (and meta-analyses) in work and organizational psychology. *European Journal of Work and Organizational Psychology, 28*(1), 1 - 10.

DeCoster, J. (2009). *Meta-analysis notes.* Retrieved from http:// www.stat-help.com/ notes.html

Durlak, A., Lipsey, M. (1991). A Practitioner's guide to Meta-Analysis. *American journal of community psychology, 19*(3), PP 291 - 332.

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

Fan, X., & Chen, M. (2001). Parental involvement and students' academic achievement: A meta-analysis. *Educational Psychology Review, 13,* 1 - 22.

Frey, B. (2018). *The SAGE encyclopedia of educational research, measurement, and evaluation* (Vols. 1 - 4). Sage.

Gopalakrishnan, S., & Ganeshkumar, P. (2013). Systematic Reviews and Meta-analysis: Understanding the Best Evidence in Primary Healthcare. *Journal of family medicine and primary care, 2*(1), 9 - 14.

Glass, G. V. (1986). Primary, secondary, and Meta-Analysis of Research. *Educational Researcher, 5,* 3 - 8.

Glass, G.V., McGaw, B., & Smith, M.L. (1981). *Meta-analysis in Social Research.* Sage.

Grissom, R. J., & Kim, J. J. (2005). *Effect sizes for research: A broad practical approach.* Lawrence Erlbaum Associates Publishers.

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

Hedges, L.V. (1983). A Random Effects Model for Effect Size. *Psychological Bulletin, 93,* 388 - 395.

Hedges, L.V. & Pigott, T.D. (2001). The power of statistical tests in Meta-Analysis. *Psychological Methods, 6,* 17 - 203.

Hedges, L.V., & Pigott, T.D. (2004). The power of statistical tests for moderators in meta-analysis. *Psychological Methods, 9,* 426 - 445.

Hedges, L.V., & R.J. Vivea, R.J. (1998). Fixed- and random-effects models in meta-analysis. *Psychological Methods, 3*(4), 486 - 504.

Hedges, L.V., Gurevitch, J., & Curtis, P. (1999). The meta-analysis of response ratios in experimental ecology. *Ecology, 80,* 1150 - 1156.

Higgins, J.P., Thompson, S.G., Deeks, J.J., Altman, D.G. (2003). Measuring inconsistency in Meta-Analyses. *BMJ, 327,* 60 - 557.

Higgins, J.P., & Thompson, S.G. (2002). Quantifying heterogeneity in a Meta-Analysis. *Statistics in Medicine, 21,* 1539 - 1558.

Holmes, C. (1983). Effect Size Estimation in Meta-Analysis. *Journal of Experimental education, 51,* 106 - 109.

Huedo-Medina, T. B., Sánchez-Meca, J., Marín-Martínez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I^2 index? *Psychological Methods, 11*(2), 193 - 206.

Huque, M. F. (1988). Experiences with meta-analysis in NDA submissions. In *Proceedings of the Biopharmaceutical Section of the American Statistical Association* (Vol. 2, No. 1, pp. 28 - 33).

Hunter, J.L., Schmidt, F., & Jackson, G. (1986). Meta-Analysis: Cumulating Research Findings Across Studies *Sage Publications: Beverly Hills, 1982, 176 pp. Educational* Researcher, 15(8), 20 - 21.

Hunter, J.L. & Schmidt, F. (2004). *Methods of Meta-Analysis.* Sage.

Jackson, D., & Turner, R. (2017). Power analysis for random-effects meta-analysis. *Research synthesis methods, 8*(3), 290 - 302.

Johnson, B. T., Mullen, B., & Salas, E. (1995). Comparison of three major meta-analytic approaches. *Journal of Applied Psychology, 80*(1), 94 - 106.

Kerby, D. S. (2014). The Simple Difference Formula: An Approach to Teaching Nonparametric Correlation. *Comprehensive Psychology.* https://doi.org/10.246611/.IT.3.1

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

LeCroy, C. W., & Krysik, J. (2007). Understanding and Interpreting Effect Size Measures. *Social Work Research, 31*(4), 243 - 248.

Lenhard, W. & Lenhard, A. (2016). Computation of effect sizes. Retrieved from: https://www.psychometrica.de/effect_size.html. Psychometrica. DOI: 10.13140/RG.2.2.17823.92329

Lyons-Ruth, K., Easterbrooks, M. A., & Cibelli, C. D. (1997). Infant attachment strategies, infant mental lag, and maternal depressive symptoms: Predictors of internalizing and externalizing problems at age 7. *Developmental Psychology, 33*(4), 681 - 692.

Sánchez-Meca, J., Marín-Martínez, F., & Chacón-Moscoso, S. (2003). Effect-size indices for dichotomized outcomes in meta-analysis. *Psychological Methods, 8*(4), 448 - 467.

Sánchez-Meca, J,, & Marín-Martínez, F. (2011). Meta–Analysis in psychological research, *International Journal of psychological Research, 3*(1), 151 - 163.

Melby-Lervag, M. & Hulme, C. (2012). Is working memory training effective? A meta-analytic review. *Developmental Psychology, 49*(2), 270 - 291.

Nouri, H., & Greenberg, R.H. (1995). Meta-Analytic procedures for estimation of effect sizes in experiments using complex analysis of variance. *Journal of Management, 21,* 801 - 812.

Redfield, D. L., & Rousseau, E. W. (1981). A Meta-Analysis of Experimental Research on Teacher Questioning Behavior. *Review of Educational Research, 51*(2), 237 - 245.

Riley, R. D., Higgins, J. P., & Deeks, J. J. (2011). Interpretation of random effects meta-analyses. *BMJ, 342*.

Robins, J., Breslow, N., & Greenland, S. (1986). Estimators of the Mantel-Haenszel variance consistent in both sparse data and large-strata limiting models. *Biometrics, 42*(2), 311 - 323.

Rosenthal, R. (1991). *Meta-Analytic Procedures for Social Research.* Sage.

Rosenthal, R. (1994). Parametric Measure of Effect Size. In H. Cooper & L. V. Hedges (Eds.), *The Handbook of Research Synthesis.* Sage.

Rosenthal, R., & Rosnow, R. (2003). Effect sizes experimenting psychologist. *Canadian Journal of experimental psychology, 57* (3), 221 - 237.

Rosenthal, R., & Rubin, D. (1982). Comparing effect sizes of independent studies. *Psychological Bulletin, 99,* 400 - 406.

Rosenthal, R., & Rubin, D. (1986). Meta-Analysis procedures for combining studies with multiple effect sizes. *Psychological Bulletin, 99,* 400 - 406.

Rothstein, R., & Borenstein, M. (2014). *Basic Aspects of Meta-Analysis.* City university of New York.

Schwarzer, R. (1989). *Meta-Analysis programs.* Berlin Institute of Psychology.

Sirin, S. R. (2005). Socioeconomic Status and Academic Achievement: A Meta-Analytic Review of Research. *Review of Educational Research, 75*(3), 417 - 453.

Steyn, H.S. (2020). *Non-parametric tests with effect sizes.* http://natural-sciences.nwu.ac.za/sites/natural-sciences.nwu.ac.za/files/files/SDK/non-pts.pdf

Stouffer, S. A., Suchman, E. A , DeVinney, L.C., Star, S.A., Williams, R.M. Jr (1949). *Adjustment During Army Life.* Princeton University Press.

Sutton, A. J. (2009). Publication bias. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (pp 435 - 452). Russell Sage Foundation.

Tabachnick, B. G., & Fidell, L. S. (2018). *Using multivariate statistics* (7th ed.). Pearson.

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

Thompson, K. N., & Schumacher, R. E. (1997). An evaluation of Rosenthal and Rubin's binomial effect size display. *Journal of Educational and Behavioral Statistics, 22,* 109 - 117.

Valentine, J. C., Pigott, T. D., & Rothstein, H. R. (2010). How many studies do you need? A primer on statistical power for meta-analysis. *Journal of Educational and Behavioral Statistics, 35*(2), 215 - 247.

Yetkiner, Z.E., Capraro, R.M., Zientek, L.R., & Thompson, B. (2004). *Effect Size and Confidence Interval.* Reporting Practices in Mathematics Education Texas A&M University.

Abstract

Purpose: The study aims to work out a theoretical introduction to the definition of effect size and meta-analysis, how to do it, and its importance through various examples. It starts with discussing the various methods for estimating the value of effect size, its average, and variance across different studies. The study presents the importance of moving beyond the statistical significance to the practical significance and estimating the effect size, through addressing the concept of effect size, its interpretation, and the factors affecting the accuracy of its estimation. It also explains the types of effect size, and methods for estimating its indicators, the mathematical transformations between them. **Method:** Because it is necessary to distinguish between the model of fixed effect and that of random effect when doing a meta-analysis study, the research presents methods for estimating the statistical properties of effect sizes in light of the fixed and random effects models. The study moves to present an important topic when doing a meta-analysis study, which is the comparison between the subgroups in the statistical properties of the effect size. And then, the study presents the concept of meta-analysis and its objectives, importance, and steps. The study ends with presenting the power of the statistical tests included in the meta-analysis. **Results:** The study presents numerous real-life examples and illustrations to clarify the presented concepts and to assist those researchers interested in planning and implementing their meta-analysis study.

Keywords: effect size, meta-analysis, fixed effects model, random effects, model.

The Author:

Prof. Hesham F. Gadelrab

- Ph.D. Wayne State University, USA, 2004.
- Director of Research and Statistical Consultation Unit – Kuwait University.
- Associate Professor of Quantitative Psychology – Kuwait University.
- Professor of Quantitative Psychology – Mansoura University.
- Experience Researcher and Applied Statistician – Madrid University, 2007.
- Statistical Consultant – Wayne State University, 2004

Some Publications:

A- Books:

1- Ben Zayed, S., Gani, A.B., **Gadelrab, H.F.,** & Bin Othman, M.K. (2020). Decades Trend of emergency Department System Operations, S. Ben Zayed et al. (Eds.), *Operational Management in Emergency Healthcare.* Springer.

2- **Gadelrab, H.F.** (2017). *Keys to Statistical Analyses.* Dar Alelm.

3- Morad, S., Hadi, F., & **Gadelrab, H.F.** (2016). *Inferential Statistics in Behavioral Sciences.* Dar Alketab Alhadith.

4- **Gadelrab, H.F.** (2010). *Evaluating the Fit of Structural Equation Models: Sensitivity to Specification Error and Descriptive Goodness-of-fit Indices.* LAP Lambert Academic Publishing,

5- Kasem, M., Allam, S., Alhousseni, **H., Gadelrab, H.F.** (2011). *Quality in Education: A Value-Added Approach.* Dar Alfekr Alarabie.

B- Research:

1- Zine-El-Abiddine, F., Al-Tammemi, A.B., **Gadelrab, H.F.,** Lin, C-Y, Aljaberi, M.A., Alhuwailah, A., Roubi, M.L. (2021). Arabic COVID-19 psychological distress scale: Development and initial validation. *BMJ Open, 11 :e046006. doi:10.1136/ bmjopen-2020-046006.*

2- **Gadelrab, H.F.,** Alkhadher, O., Almoshawah, S., …& Slimene, S. (2020). Organizational justice in Arab Countries: Investigation of the Measurement and Structural Invariance. *Cross-cultural Research, 54*(1), 3 - 27.

3- **Gadelrab, H.F.** (2019). An Investigation of Differential Relationships of Implicit and Explicit Aggression: Validation of an Arabic Version of the Conditional Reasoning Test for Aggression. *Journal of Personality Assessment, 101*(6), 609-620.

4- **Gadelrab, H.F.** & Alkhadher, O. (2017). To Translate or to Develop a Measure? The Case of a New Arabic Measure of Organizational Justice. *International Journal of Selection and Assessment. 25*(1), 83 - 91.

5- Alkhadher, O. & **Gadelrab, H.F.** (2016). Organizational Justice Dimensions: Validation of an Arabic Measure. *International Journal of Selection and Assessment. 24*(4), 337- 351.

The Author:

Prof. Mamdouh Abdelmonem Elkenani

- Ph.D., Mansoura University.
- Former Dean of the Faculty of Education, Mansoura University.

Some Publications:

A- Books:

1- Learning Psychology.
2- Learning Styles.
3- Creativity Principles.
4- Psychological Measurement.
5- Childhood Problems.
6- Creativity Climate.
7- Individual Differences in Intelligence and Intellectual Abilities.
8- Creativity Psychology.
9- Rearing a Creative Child.
10- Psychological Assessment.
11- Creativity Psychology in Organizations.
12- Childhood Creativity.
13- Research in Childhood Creativity.

B- Research Topics:

1- Self-Learning and Attitudes Towards Learning.
2- Creativity and Critical Thinking.
3- Creative Environment.
4- Leaning and Achievement Motivation.
5- Personality Traits.
6- Impulsivity.
7- Human Needs.
8- Locus of Control.
9- Children's Creativity.
10- Most Common Statistical Mistakes in Research.

Doi 10.34120/0757-044-640-001
Monograph (640)

Effect Size and Meta-Analysis

Mamdouh A. Elkenani
Department of Educational Psychology
College of Education - Mansoura University
Egypt

Hesham F. Gadelrab
Director of Research and Statistical
Consultation Unit
Kuwait University

ANNALS OF THE ARTS AND SOCIAL SCIENCES - KUWAIT UNIVERSITY

Editorial Board

- **Prof. Taghreed Alqudsi - Ghabra**
  Editor - in - Chief
  Department of Information Studies
  Kuwait University
  Taghreed.alqdsi@ku.edu.kw
- **Prof. Baqer Salman Alnajjar**
  Bahrain University - Bahrain
  Visiting Prof. Sociology and Social
  Service Department - Kuwait University
  baqer.alnajjar@ku.edu.kw
- **Prof. Numan M. Jubran**
  Irbid University - Jordan
  Visiting Prof. History Department
  Kuwait University
  numan.jubran@ku.edu.kw
- **Prof. Adnan Y. Alotoum**
  Psychology Department - Yarmouk
  University - Jordan
  atoum@yu.edu.jo
- **Prof. Mashael Al-Hamly**
  English Department and Literature
  Kuwait University
  mashael.alhamly@ku.edu.kw
- **Dr. Abdullah Mohamed Aljasmy**
  Philosophy Department
  Kuwait University
  abdulla.hasan@ku.edu.kw
- **Dr. Ibraheem Nagy Al-Hadban**
  Department of Political Science
  Kuwait University
  ibrahim.alhada@ku.edu.kw
- **Dr. Ahmed Mubarak AlHasem**
  Geography Department
  Kuwait University
  ahmed.hassam@ku.edu.kw
- **Dr. Asad H. Al-Saleh**
  Arabic Department
  Indiana University
  alsaleha@indiana.edu.kw
- **Maha Ibrahim Al-Msad**
  Editorial Manager
  Kuwait University
  maha.almsad@ku.edu.kw

Advisory Board

- **Prof. Ibrahim Al-Sa'afin**
  Department of Arabic Language
  and Literature - Jordan University
- **Prof. Hamdi Hasan Abul Enein**
  Faculty of Mass Communication
  Misr International University
- **Prof. Sari Hanafi**
  President of the International Sociological
  Association - American University- Beirut
  sh41@aub.edu.lb
- **Prof. Abdullah Al-Walee'i**
  Geography Department
  King Saud University
- **Prof. Ma'moun Fandi**
  Director of London Institute of
  Strategic Studies - United Kingdom
- **Prof. Suliman Hawamdeh**
  Regents Professor Information Science
  University of North Texas
  Suliman.hawamdeh@unt.edu
- **Prof. Maha Zaraket**
  Faculty of Information
  Lebanese University
  Maha.zaraket@ul.edu.lb

ANNALS OF THE ARTS AND SOCIAL SCIENCES

Published by the Academic Publication Council - Kuwait University

All Rights Reserved

REFEREED SCIENTIFIC QUARTERLY IN THE HUMANITIES, SOCIAL SCIENCES & LITERATURE

Volume 44, 2024

Rates:

Kuwait: K.D. 0.500, Bahrain: BD 1, Qatar: RQ 10, Emirates: DH 10,
Saudi Arabia: RS 10.

Cost per issue in Arab Countries: Equivalent to one US dollar

Cost per issue in other Countries: Equivalent to three US dollar

● Annual Subscription (individuals):

1 Year: Kuwait: 4 K.D., Arab Countries: 6 K.D., Foreign Countries 22 dollars.
2 Year: Kuwait: 7 K.D., Arab Countries: 10 K.D., Foreign Countries 37 dollars.
3 Year: Kuwait: 10 K.D., Arab Countries: 14 K.D., Foreign Countries 52 dollars.
4 Year: Kuwait: 13 K.D., Arab Countries: 18 K.D., Foreign Countries 67 dollars.

● Annual Subscription (institutions):

1 Year: Kuwait: 22 K.D., Arab Countries: 22 K.D., Foreign Countries 90 dollars.
2 Year: Kuwait: 37 K.D., Arab Countries: 37 K.D., Foreign Countries 150 dollars.
3 Year: Kuwait: 53 K.D., Arab Countries: 53 K.D., Foreign Countries 210 dollars.
4 Year: Kuwait: 67 K.D., Arab Countries: 67 K.D., Foreign Countries 270 dollars.

All correspondence and enquiries must be addressed to:

Editor

ANNALS OF THE ARTS AND SOCIAL SCIENCES
P.O.BOX 17370 EL-KHALDIAH - KUWAIT 72454
Tel: 24830256 - Fax: 24830256
ISSN 1560-5248 Key title: Hawliyyat Kulliyyat al-Adab
E-mail: aass@Ku.edu.Kw
http:// apc. kuniv.edu.kw/AASS/

تتوفر نصوص البحوث كاملة لدى:
EBSCO Publishing Products
دار المنظومة: www.mandumah.com
UIrich's Periodicals Directory 263117

The Publications of The Academic publiction Council

journal of the Social Sciences 1973, Kuwait Journal of Science and Engineering 1974, journal of the Gulf and Arabian Peninsula Studies 1975, Authorship Translation and Puplication Committee 1976, Journal of law 1977, Annals of the Arts and Social Sciences 1980, Arab Journal for the Humanities 1981, The Sducational Journal 1983, Journal of Sharia and Islamic Studies 1983, Arab Journal of Administrative Sciences 1991.

ANNALS OF THE
ARTS AND SOCIAL SCIENCES

A Refereed Academic Quarterly, Academic Publication Council - Kuwait University

حجم الأثر وما وراء التحليل

Effect Size and Meta-Analysis

Mamdouh A. Elkenani
Department of Educational Psychology
College of Education - Mansoura University
Egypt

Hesham F. Gadelrab
Director of Research and Statistical
Consultation Unit
Kuwait University

جامعة الكويت
KUWAIT UNIVERSITY

Academic
Publication Council

ISSN: 5248 - 1560

Monograph 640 - March 44

1445 A.H / March 2024