

التحليل الدمجى (ما وراء التحليل) للعلوم الاجتماعية والنفسية والتربوية

الاسس النظرية والمنهجية والتطبيقات

اعداد

ا.د. عبد الناصر السيد عامر

استاذ القياس والتقويم والاحصاء النفسى والتربوى

كلية التربية- جامعة قناة السويس

2018م-1440هـ

فهرس المحتويات

3	المقدمة	1
8	الفصل الاول: مقدمة لما وراء التحليل	-2
47	الفصل الثانى: تحديد المشكلة او موضوع البحث	-3
62	الفصل الثالث: مسح وتحديد التراث البحثى	-4
75	الفصل الرابع: تكويد خصائص الدراسات	-5
95	الفصل الخامس: حساب وتكويد حجم التأثير	-6
137	الفصل السادس: احجام تاثير متقدمة	-7
159	الفصل السابع: تصحيحات احجام التأثير	-8
173	الفصل الثامن: دمج احجام التأثير	-9
192	الفصل التاسع: نماذج تحليل ما وراء التحليل	-10
215	الفصل العاشر: نموذج التأثيرات المثبتة: تجزئة تباين حجم التأثير او التحليلات المتفاعلة	-11
236	الفصل الحادى عشر: تحيز النشر	-12
261	الفصل الثانى عشر: مداخل اجراء ما وراء التحليل	-13
281	الفصل الثالث عشر: ما وراء التحليل باستخدام البرامج الاحصائية	-14
315	الفصل الرابع عشر: كتابة تقرير نتائج ما وراء التحليل	-15
350	المراجع	-16

مقدمة

يعتبر أسلوب ما وراء التحليل احد طرق المراجعة المنظمة الكمية للادبيات البحثية فى كل التخصصات خاصة العلوم فى العلوم السلوكية والاجتماعية والطبية والبيئة وغيرها. وهو أسلوب لتقييم التراث البحثى المرتبط بظاهرة ما فى مجال ما لمعرفة نواحي القصور والقوة فى محاولة تحسين الممارسات البحثية الحالية، واعطاء الباحثين وصانعى القرارات بمعلومات جديدة وبناء ناتجة عن تلخيص النواتج الكمية لتراث بحثى ما.

ويعتبر أسلوب ما وراء التحليل البديل المفضل للاجابة عن الاسئلة البحثية للعلاقة بين متغيرين بدون اجراء الدراسات الاولية التى تعتمد على انتقاء عينة من الافراد وبناء المقاييس، انما عينته هى الدراسات السابقة من ثم فلا حاجة للطريقة النمطية لاجراء البحوث الاولية خاصة اذا كان التراث البحثى يتضمن تراكم هائل من الدراسات الاولية ذات النتائج المتعارضة وعلى ذلك فهو يسهم فى تصميم بحوث جديدة وعليه فانه طريقة منهجية جديدة لاجراء الدراسات حيث عينته هى الدراسات السابقة.

واستحدث أسلوب ما وراء التحليل كرد فعل للانتقادات الموجهة لاختبارات الدلالة الاحصائية حيث لاتساعد على صناعة قرارات فعالة ولاتساعد على تقدم العلوم النفسية ، ايضاً ناتجة للانتقادات الموجهة للطريقة التقليدية الكيفية السردية فى مراجعة الدراسات السابقة حث تودى الى استنتاجات مختلفة وتعتمد على ذاتية الباحث خاصة لنتائج الدراسات المتعارضة.

الدافع الحقيقى لاعداد هذا الكتاب هى خلو المكتبة العربية بمرجع شامل لاسلوب ما وراء التحليل انما كانت فى معظمها مقالات نظرية لا تتناول الخطوات المنهجية والاجرائية والاساليب الاحصائية المتضمنة فى أسلوب ما وراء التحليل، وهو يصلح لمستخدمى ما وراء التحليل فى كل التخصصات السلوكية والاجتماعية والطبية وغيرها.

وتوجد العديد من الممارسات الخاطئة فى الدراسات المنشورة وغير المنشورة حيث لا تتبع الخطوات والاجراءات الصحيحة فى اجراء هذا الاسلوب ولا تتناول ماهية النموذج المناسب انما

فى معظمها تعتمد على نموذج التاثرات المثبتة على الرغم من وجود اختلافات بين احجام التاثير المختلفة وهذا بدوره يؤدى الى نتائج متحيزة، ويمكن ان يعتمد الباحث على اكثر من حجم تاثير فى الدراسة الواحدة وبالتالي لم يتحقق من مسلمة الاستقلالية بين احجام التاثير وهذا من اهم افتراضيات ما وراء التحليل، وكذلك قصور فى تفسير نتائج ما وراء التحليل وعرض الاشكال والرسومات الخاصة به وكيفية تفسيرها بموضوعية بعيداً عن الذاتية.

والدراسات المنشورة لا تعتمد على تصحيح احجام التاثير من المحددات المنهجية فى الدراسة خاصة اخطاء القياس وهذا من اهم الاجراءات التى يجب تضمينها فى دراسة ما وراء التحليل لان قياسات العلوم السلوكية والنفسية غير تامة الثبات لانه قياس غير مباشر ولذلك قدمت فصل كامل عن تصحيح حجم التاثير من المحددات المنهجية كما اوضحها Hunter & Schmidt (2004) حيث تبناوا هذا الاتجاه فى دراسات ما وراء التحليل ووضعوا لهذه التصحيحات اهمية قصوى للحصول على نتائج موثوق فيها.

الهدف من الكتاب اعطاء معظم المبادئ الاساسية النظرية والفنيات المنهجية لاسلوب ما وراء التحليل بعيداً بقدر الامكان عن العمليات الحسابية والاحصائية المعقدة مستعيناً بمراجع هامة ورائدة فى هذا الشأن خاصة مراجع (2001) Lispey & Wilson و Borenstein et al. (2009) و Card (2012) حيث من افضل ما كتب فى ما وراء التحليل وتتميز بسهولة فهمها وتسلسل موضوعاتها وشموليتها لمعظم فنيات ما وراء التحليل، وقد استعنت فى معظم حسابات ما وراء التحليل بمثال لدراسة Card, Stucky, Sawalani, & Little, 2008 كما عرضه Card(2008).

وعرضت فى الكتاب كيفية اجراء ما وراء التحليل باستخدام برامج احصائية اعدت لهذا الشأن واستعنت ببرنامج Comprehensive Meta Analysis(CMA) وهو برنامج تجارى حيث يتضمن كل معظم اساليب وطرق ما وراء التحليل ويعتبر من افضل البرامج بتضمينه للتطورات الحديثة لفنيات ما وراء التحليل.

وهذا الكتاب يساعد الباحثين على اجراء دراسات ما وراء التحليل بدجة كبيرة من الفهم لطبيعة البيانات التى يتعامل معها وكيفية استخلاصها وتكويدها واختبار مدى الاختلاف بين الدراسات

وطبيعة النموذج المناسب للتحليل سواء نموذج مثبت او نموذج عشوائى وتقدير مدى تحيز الدراسات فى التحليل وكيفية تفسير نتائج التحليل بل والذهاب ابعد من ذلك من خلال مساعدة الباحث على تقييم دراسة ما وراء التحليل المنشورة فى المجالات العلمية.

ويتضمن الكتاب العديد من الفصول كالتالى:

الفصل الاول: ماهية ما وراء التحليل وتضمن مشاكل الدلالة الاحصائية والاستنتاجات الخاطئة من اختبارات الدلالة الاحصائية، الحاجة الى توليف البحوث فى كافة التخصصات، دور ما وراء التحليل فى العلوم الاجتماعية والسلوكية، مفهوم ما وراء التحليل، المواقف المناسبة لاستخدام ماوراء التحليل، خطوات ماوراء التحليل اوتوليف البحوث، مميزات ماوراء التحليل، والانتقادات الموجهة لاسلوب ما وراء التحليل.

الفصل الثانى: تحديد المشكلة او موضوع البحث وتضمن تحديد اهداف واسئلة او موضوع البحث المراد الاجابة فى ما وراء التحليل، معايير الصلاحية او تضمين واستبعاد الدراسات، تحديد نواتج البحث المتطلبة لاجراء دراسة ما وراء التحليل.

الفصل الثالث: مسح وتحديد التراث البحثى تضمن مصادر الحصول على التراث، البحثى المناسب، استراتيجية البحث الخلفى، استراتيجية البحث الامامى، وبناء قاعة بيانات للدراسات المتطلبة لما وراء التحليل.

الفصل الرابع: تكويد خصائص الدراسات ويتضمن بناء برتكول التكويد متضمناً كود الدراسة، مفاهيمها، خصائص العينة، خصائص القياس، توقيت القياس، وخصائص التصميم. وكذلك تكويد حجم التأثير وتكويد جودة الدراسة، توصيفات الدراسة وتقويم عملية التكويد.

الفصل الخامس: حساب وتكويد حجم التأثير تضمن هذا الفصل مفهوم حجم التأثير، أهمية حجم التأثير، تصنيف مؤشرات حجم التأثير، مؤشرات الفروق بين المتوسطات، مؤشرات فروق المتوسطات المعيارية، مؤشرات قوة العلاقة (معامل ارتباط بيرسون)، مؤشرات تقديرات المخاطرة (نسبة الترجيح Odds ratio)، و تفسير حجم التأثير.

الفصل السادس: احجام تأثير متقدمة تضمن هذا الفصل مؤشرات حجم التأثير لاحصاء المتغير الواحد، نسب المتغير، التباينات والانحرافات المعيارية، معامل ارتباط Point (r_{pb}) bi-serial، مقارنة القياسات القبليّة والبعدية، درجات التغير الطولى، مؤشرات تحليل الانحدار، مؤشرات حجم التأثير في اختبارات الإحصاء المتدرج، وما وراء التحليل لثبات الاتساق الداخلى.

الفصل السابع: تصحيحات احجام التأثير تضمن هذا الفصل دواعى تصحيح احجام التأثير، التحويرات وتصحيح التحيز، التصحيحات من عدم الثبات، تصحيحات الصدق الناقص، تصحيحات الثنائية التصنيفية الاصطناعية، والتصحيح من المدى المقيد.

الفصل الثامن دمج احجام التأثير تضمن هذا الفصل اشكالية تحليل ما وراء التحليل، بناء قاعدة بيانات لمجموعة مستقلة من احجام التأثير، حساب متوسط حجم التأثير، حساب المتوسط الموزون لاجسام التأثير، اختبارات الدلالة الاحصائية وفترات الثقة حول متوسط حجم التأثير، اختبار التجانس بين احجام التأثير.

الفصل التاسع: نماذج تحليل ما وراء التحليل تضمن هذا الفصل نموذج التأثيرات المثبتة، نموذج التأثيرات العشوائية، الفروق بين نموذج التأثيرات المثبتة والعشوائية فى ما وراء التحليل، العمليات الحسابية لنموذج التأثيرات العشوائية، والمقارنة بين نموذج التأثيرات المثبتة ونموذج التأثيرات العشوائية.

الفصل العاشر: نموذج التأثيرات المثبتة: تجزئة تباين حجم التأثير او التحليلات المتفاعلة تضمن هذا الفصل تأثير المتغيرات المتفاعلة التصنيفية، التحليل البعدى للمتفاعل التصنيفى، المتغيرات المتفاعلة المتصلة او ما وراء الانحدار، الاطار العام للانحدار المتعدد للمتغيرات المتفاعلة، ونماذج التأثيرات المختلطة.

الفصل الحادى عشر: تحيز النشر تضمن هذا الفصل ماهية تحيز النشر فى ما وراء التحليل، مصادر تحيز النشر، محاولة تجنب التحيز، الطرق الاحصائية لتقييم تحيز النشر، شكل القمع وتفسيره، وطرق تصحيح تحيز الاتاحية او النشر.

الفصل الثاني عشر: مداخل اجراء وراء التحليل تضمن هذا الفصل طرق دمج احجام التأثير، طرق ما وراء التحليل التي تركز على خطأ المعاينة فقط، ما وراء التحليل السيكومتري، اجراء ما وراء التحليل في ضوء مدخل (2004) Hunter & Schmidt، اجراء ما وراء التحليل لتوليف r باستخدام مدخل Hedges.

الفصل الثالث عشر: ما وراء التحليل باستخدام البرامج الاحصائية تضمن هذا الفصل عرض لاهم البرامج المستخدمة لاجراء المراجعة المنظمة للتراث البحثي والبرامج التجارية لما وراء التحليل مثل برنامج CMA والبرامج المجانية مثل MetaXL وعرض مفصل لبرنامج 3.Comprehensive Meta-Analysis(CMA) Version ..

الفصل الرابع عشر: كتابة تقرير نتائج ما وراء التحليل تضمن هذا الفصل تفسير حجم التأثير، محتوى تقرير دراسة ما وراء التحليل من العنوان والمقدمة والطريقة وعرض النتائج والتفسير والمناقشة وعرض الجداول والاشكال، تجنب المشاكل عند كتابة عرض نتائج ما وراء التحليل، قضايا في ما وراء التحليل خاصة العدد الكافي من الدراسات لاجراء ما وراء التحليل وتقييم جودة دراسة ما وراء التحليل.

والله الموفق والمستعان

أ.د. عبد الناصر السيد عامر

أستاذ القياس والتقويم والإحصاء النفسي والتربوي

رئيس قسم علم النفس التربوي - كلية التربية بالاسماعيلية

جامعة قناة السويس - مصر

adr.abdenasser@yahoo.com

الفصل الاول

ماهية لما وراء التحليل

Introduction to Meta-Analysis

مقدمة

المهمة الاساسية للعلم فى كل التخصصات هو بناء وتطوير النظرية التى تفسر ظاهرة ما، وفى كثير من الاحيان يمتلك الباحث كم هائل من الدراسات السابقة فى موضوع ما وتصبح الوظيفة الاساسية للباحث هى ادراك ما تتضمنه هذه العلاقات والوصول الى استنتاج عام واخذها فى الاعتبار عند بناء النظرية، ولادراك هذه العلاقات فمن الضرورى مراجعة هذه الدراسات للوصول الى رؤية عامة شاملة وتلخيص منظم وواضح لهذه الدراسات بما تتضمنه من عينات ومنهجيات واجراءات مختلفة.

يطلق على العلوم السلوكية بالعلوم الناعمة لانها ينقصها الاطار العام المنظم المتناسك Scientific Paradigm حيث لم تصل الى القدر الكافى من الاستقرار حيث يوجد القليل من الاتفاق حول المشكلات الرئيسية وكيفية دراستها (Kline, 2013).

واى علم يتضمن التراكمية النظرية متضمنة المفاهيم والنظريات والقوانين والتصنيفات، وكذلك التراكمية الامبريقية Empirical Cumulativity نتيجة الدراسات والبحوث وهذا متوافر الى حداً ما فى العلوم السلوكية والاجتماعية التى ينتج من دراساتها نتائج متعارضة واحياناً غير قابلة للاعادة Replicability مقارنة بنتائج العلوم الطبيعية.

وتوجد مصادر عديدة للحصول على نتائج زائفة Spurious results من اهمها انخفاض جودة المقاييس المستخدمة فى الدراسة، خطأ ادخال البيانات، خطأ نقل النتائج من المخرجات الكمبيوترية، خطأ كتابة الاحصائيات فى تقرير البحث، وخطأ المعاينة غيرها.

وكثيراً من الاكتشافات والتقدم والتطور فى المعرفة التراكمية لم يتم باستخدام دراسات البحوث الاولى انما عن طريق هؤلاء الذين يجرون ما وراء التحليل لاكتشاف المعنى الكامن او الخفى للتراث البحثى الموجود لظاهرة ما (Schmidt, 1992).

واشار (2004) Hunter & Schmidt الى ان الدرس المستخلص من المراجعة الكيفية السردية التقليدية "النتائج المتعارضة فى الادبيات البحثية لموضوع ما ربما تكون مزيفة may "be entirely artifactual .

مشاكل اختبارات الدلالة الاحصائية

يعتقد الباحث ان مستوى الخطأ من النوع الاول يظل ثابتاً ولكن بتعبير اكثر صدمة فان نسبة الخطأ اكبر من 5% الذى يتبناها الباحث، فالكثير يعتقد ان استخدام اختبارات الدلالة كقول ان يجعل نسبة الخطأ من النوع الاول عند 5% او اقل وهذا غير حقيقى (Hunter & Schmidt, 2004).

فلو ان الفرض الصفرى حقيقى فى المجتمع (لا علاقة او لافروق) وتم رفضه فى ضوء بيانات العينة، فى هذه الحالة تم الوقوع فى الخطأ من النوع الاول α Type 1 error، واذا كان الفرض الصفرى غير حقيقى على مستوى المجتمع وتم قبوله على مستوى بيانات العينة بالتالى تم الوقوع فى الخطأ من النوع الثانى β Type 11 error. ، واختبارات الدلالة الاحصائية تحاول بقدر الامكان ان تجعل الخطأ من النوع الاول تقريباً عند 5% ولكن هذا صعب المنال ولكن الشئ الوحيد الضامن لهذا هو هجر او الابتعاد عن استخدام اختبارات الدلالة الاحصائية والاعتماد على استخدام فترة الثقة Confidence intervals كبديل لاختبارات الفروض الصفريّة.

وتوجد عدة انتقادات للدلالة الاحصائية اهمها:

- حساسيتها لحجم العينة **Sensitivity to sample size** : عندما يكون حجم العينة صغيراً فان التأثيرات القوية لا تكون دالة احصائياً (ارتكاب الخطأ من النوع الثانى)، فى المقابل عندما تكون حجم العينة كبيراً فان التأثيرات الصغيرة تكون دالة احصائياً (ارتكاب الخطأ من النوع الاول). فمثلاً اذا كان معامل الارتباط $r(20) = .40$ غير دال احصائياً عند $P \geq 0.05$ لاختبار نو ذيلين، بينما $r(1000) = .07$ دال احصائياً. وعلى الجانب الاخر اذا كانت $r = .25$ فان النتائج دالة اذا كان حجم العينة $n=63$ وغير دالة اذا كان $n=61$ وعليه فان هذا يؤدى الى استنتاجات خاطئة.

- لا تخبر الباحثين عن ماذا يريدوا ان يعرفوه ولكنها تخلق اضاءة عن برهان احتمالي للتناقض بين الفرض الصفري والفرض البديل.
- **نسب الخطأ والقوة Power and error rate:** الانتقاد الاساسى لاختبارات الفروض الصفرية النسبة العالية للخطأ من النوع الثانى غير المقبولة (Schmidt, 1996)، وهذا يؤدي الى نقص القوة الاحصائية المناسبة (Cohen, 1962). والملاحظ فى التراث البحثى فى التربية وعلم النفس ان متوسط القوة فى معظم التراث البحث فى علم النفس تراوح من 0.40 الى 0.60 بل احياناً 0.20 فى بعض التخصصات (Hunter, 1997) (Maxwell, Schmidt, 1996; Sedlmeier & Gigerenzer, 1989). وتوصل (2004) الى ان وسيط القوة فى الدراسات غير التجريبية فقط حوالى 50%. وهذا مفاده ان فرصة 50% لرفض صحيح للفرض الصفري فى ضوء بيانات العينة و50% رفض خاطئ للفرض الصفري. ويشير (Kline, 2013) فى هذه الحالة من الافضل عدم اجراء البحث وجمع البيانات والاعتماد على مبدا ملك اوكتابة للعملة لتحديد رفض او قبول الفرض الصفري.

وتوصل عامر (2014) إلى أن وسيط القوة للدراسات النفسية التربوية العربية ذات حجم التأثير الصغير 0.28 ولحجم التأثير المتوسط 0.90، ولحجم تأثير كبير 0.99 و فى المجمل فان متوسط القوة الاحصائية 72. . فى المجمل فى التراث الاجنبى والعربى انخفضت القوة الاحصائية عن الحد الادنى المناسب للقوة الاحصائية وهو 0.80 (معيار Cohen)، ومن اهم العواقب المترتبة على الاعتماد على قوة منخفضة هو عدم القدرة على تفسير نتائج التراث البحثى لظاهرة ما، فاذا وجد تأثير حقيقى للظاهرة موضع الدراسة ولكن القوة الاحصائية 0.50 فان هذا يعنى ان نصف الدراسات تعطى نتائج دالة بينما النصف الاخر لا يعطى دلالة احصائية، ولو نشرت هذه الدراسات فان نصفها يعطى نتائج ايجابية والنصف الاخر نتائج سالبة وقد عبر Simmons, Nelson, & Simonsohn (2011) عن ذلك بعبارة " علم النفس الموجب والخطأ".

وعليه فالاعتماد التقليدى على اختبارات الدلالة الاحصائية فى تفسير الدراسات يقود الى تفسيرات خاطئة عن ما تعنيه نتائج الدراسات. ويرى (Schmidt, 1996) ان استخدام المدخل التقليدى

في تحليل البيانات يجعل من غير المحتمل الوصول الى استنتاجات صحيحة في معظم التخصصات. ويشير (1990) Cohen الى ان العلوم الطبيعية مثل الفيزياء والكيمياء لاتستخدم اختبارات الدلالة الاحصائية في تفسير بياناتها. ويعتقد الباحثون انه اذا كانت النتيجة دالة عند 0.05 فان احتمالية تكرار هذه النتيجة في الدراسات اللاحقة هي $1 - 0.05 = 0.95$. وهذا الاعتقاد خاطئ تماماً. ولكن احتمال الاعداد فان القوة الاحصائية للدراسة تقريباً اقل من 0.95.

الاستنتاجات الخاطئة بعد صناعة القرار عن الفرض الصفري

توجد استنتاجات خاطئة بعد رفض او الفشل في رفض الفرض الصفري وهي كالتالي (Kline, 2013):

- **خدوة الاحتمالية في مقابل الصدفة Odds against Chance Fallacy**: يوجد اعتقاد خاطئ بان القيمة الاحتمالية P تشير الى ان نتائج الدراسة تحدث نتيجة خطأ المعاينة ولذلك $P < 0.05$ تعنى يوجد احتمال اقل من 5% ان نتيجة الدراسة تكون للصدفة ويوجد مفهوم خاطئ هو ان قيم P تصنف النتائج الى تصنيفين اما الى نتيجة تكون للصدفة (قبول الفرض الصفري) ونتيجة تكون للتاثيرات الحقيقية الواقعية (رفض الفرض الصفري) وهذه الاعتقادات خاطئة. ولكن توجد عوامل اخرى تلعب دوراً في تفسير وحساب قيمة P في اطار التصميم البحثي للدراسة وعملية القياس و اساليب التحليل الاحصائي. وكثيراً من الباحثين يعتقدون ان $P < 0.05$ تعنى ان احتمال القرار اللازم لرفض الفرض الصفري هو ان الخطأ من النوع الاول اقل من 0.05.
- **خدوة الاعداد Replicability**: يوجد اعتقاد خاطئ ان النتائج الدالة احصائياً هي قابلة للاعادة اذا اجريت الدراسة على عينات مماثلة اعتقاداً ان $P < 0.05$ تعنى ان احتمالية الاعداد اكبر من 95%، ولكن (Greenwalw, Gonzalez, Harris, & Guthrie (1996) وجدوا علاقة منحنية بين قيم P ومتوسط القوة الاحصائية لدراسات عشوائية تم اعاتها على دراسات اخرى مماثلة لها بنفس العدد من احجام العينات وهذا يعنى احتمال الاعداد لنفس النتائج تقريباً 50%.
- **خدوة الحجم Magnitude Fallacy**: الاعتقاد بان القيمة المنخفضة ل P تشير الى تاثير كبير او قوة المعالجة كبيرة حيث يصف الباحث النتائج التي $P < \alpha$ بانها دالة

احصائياً بالتالى فان النتائج الدالة عند 0.01 اكثر حجماً من النتائج الدالة عند 0.05 وهذا الاستنتاج غير مضمون لان هذه الدلالة تتاثر بحجم التأثير وحجم العينة لانه يمكن الحصول على قيم صغيرة جدا لـ P لاجسام العينات الكبيرة بغض النظر عن حجم الاثر او قوة العلاقة.

- **اخدوة المعنوية Meaningfulness fallacy Fallacy:** فى ضوء اخدوة المعنى

يعتقد الباحث ان رفض الفرض الصفري يؤكد الفرض البديل وهذا يعكس خطأين. الاول رفض H_0 فى الدراسة لايمنى تحقق ان الفرض البديل، الثانى حتى لو ان الفرض البديل الاحصائى H_1 صحيح فان هذا لايمنى الفرض الجوهري من وراء H_1 حقيقى. فالدلالة الاحصائية لا تبرهن على صحة فرض معين، فبعد رفض او ابقاء الفرض الصفري لا يوجد امام الباحث الكثير ليفعله سوى اصدار التوصية.

- **اخدوة الصفر واخدوة التكافؤ Zero Fallacy and Equivalence : Fallacy:**

يوجد اعتقاد خاطئ بان الفشل فى رفض الفرض الصفري يعنى ان حجم التأثير فى المجتمع يساوى صفر. بينما اخدوة التكافؤ تحدث عند الفشل فى رفض الفرض الصفري H_0 : $\mu_1 = \mu_2$ فانه يفسر على اساس ان المجتمعين متكافئين ولكن هذا اعتقاد خاطئ لانه يمكن ان يحدث اختلاف بينهما فى التباين او شكل التوزيع.

- **اخدوة الجودة واخدوة النجاح Quality Fallacy and Success Fallacy:** الاعتقاد

بان الحصول على الدلالة الاحصائية مؤشر لجودة التصميم التجريبي او البحثى ويشير ايضاً الى نجاح الدراسة، فالدراسة ضعيفة التصميم او التى تتضمن خطأً معاينة يمكن ان تقود الى رفض غير صحيح لـ H_0 (الخطا من النوع الاول)، واحياناً الفشل فى رفض الفرض الصفري يقود الى تولد افكار ومناقشات جيدة خاصة عندما لا يتم تدعيمه من الباحثين الاخرين فى المجال وعليه وجود علم جيد وعليه نقص النتائج الايجابية من الدراسات يكون اكثر فائدة فى تقدم العلم.

- **اخدوة الفشل Failure Fallacy:** يوجد اعتقاد خاطئ بان نقص الدلالة الاحصائية

مؤشر لفشل الدراسة فى تحقيق اهدافها ولكن استخدام الطرق التحليلية غير المناسبة او القوة الاحصائية المنخفضة تسبب الخطأ من النوع الثانى ويعزو الباحث الفشل فى رفض الفرض الصفري الى تصميم الدراسة غير الجيد.

• **ادخوة التنقيح Reification Fallacy**: الاعتقاد الخاطئ بان الفشل فى الحصول على نفس النتائج من دراسات الاخرى لنفس المتغيرات هو فشل فى الوصول لنفس القرار عبر دراسات اخرى. فالنتيجة لاتعتبر قابلة للتكرار لو ان H_0 رفض فى الدراسة الاولى وليس كذلك فى الدراسة الثانية، وهذه الاشكالية تتجاهل حجم العينة، حجم التأثير، والقوة عبر الدراسات المختلفة. ومما اختلفت نتائج الدراسات حول H_0 فان هذا فى حد ذاته دليل للاعادة . Replication

• **ادخوة الموضوعية Objectivity Fallacy**: الاعتقاد بان اختبارات الدلالة هى طريقة موضوعية لاختبارات الفروض وكل طرق الاستدلال الاخرى ذاتية هى اضحوكة غير مقبولة. فهى ليست الاطار الوحيد فقط لاختبار الفروض بل يوجد التقدير البيزانى Bayesian estimation وهى بديل جيد لاختبارات الدلالة وكذلك فترت الثقة.

• **ادخوة الثنائية Sanctification Fallacy**: وتشير الى اشكالية التفكير التصنيفى

الثنائى عن متصل قيم P . فماذا لو $\alpha = 0.05$. وكانت $P = 0.049$ فى مقابل $P = 0.051$ عملياً متماثلين فى ضوء نواتج الاختبار، فالباحث يولى اهتمام للاولى (دالة) ويتجاهل الثانية (غير دالة)، وهذا التغيير فى درجة الثقة نتيجة تغير على متصل قيمة P وعلية القيم الكبيرة لـ P تناظر تغيرات صغيرة وغير دالة فى البناء التحتى للمتغيرات. لكن قيم P الاقل من 0.05 مثل 0.048 او 0.049 . فهى توصف بانها غير دالة Kline, (2013).

• **ادخوة المناعة Robustness Fallacy**: الاختبارات الاحصائية المعلمية البارامترية ليست لديها مناعة او ضلاعة ضد القيم المتطرفة او عدم تحقق مسلمات خصائص التوزيع مثل الاعتدالية خاصة للعينات الصغيرة، ولكن كثير من الباحثين يعتقدون العكس والباحثين لا يعطوا مؤشرات على مدى توافر مسلمات الاختبار المستخدم (Onwuegbuzie, 2002) وعليه الاعتماد على الدلالة الاحصائية فى تفسير الدراسات يقود الى استنتاجات خاطئة عن ماذا تعنى نتائج الدراسة، وفى الواقع فان المدخل التقليدى (اختبارات الفروض الصفرية) فى تحليل البيانات يجعل من غير المحتمل الوصول الى استنتاجات صحيحة فى معظم مجالات البحث (Schmidt, 1996).

حقيقة حدثت نقاشات كثيرة منها المؤيد لاختبارات الدلالة بعضها معارض لاستخدامها في اجراءت العمل الاحصائي ونتجة لذلك قدم Kline (2013) كتاب معنون "ما بعد اختبارات الدلالة: تجديد الاحصاء للعلوم السلوكية " قدم فيه عرض لمحددات الدلالة الاحصائية وبدائلها.

والتراث البحثي في مجال العلوم السلوكية وضع جل اهتمامه على الحصول على دلالة احصائية واهمل جانب اخر من الدقة الاحصائية وهو القوة الاحصائية على الرغم من تأكيد الجمعية النفسية الامريكية على استخدامها (APA, 2010) وهذا التجاهل للقوة الاحصائية نتجة القوة الاحصائية المنخفضة في البحوث.

وفي ضوء المراجعة السرديّة الكيفية التقليدية تمدنا بنتائج متضاربة والوصول الى استنتاج هو " اجراء المزيد من الدراسات السابقة" ولكن اي نتائج جديدة سوف تعزز او تزيد من هذا التعارض والغموض لو ان القوة منخفضة.

فقد وصف كثيراً من الباحثين الاخطأ المعرفية المرتبطة باختبارات الدلالة الاحصائية بانها نوعا من العبث او عدم التعلم المتدرب عليه Trained incapacity الذي يعوق الباحثين من فهم نتائج دراساتهم، والبعض الاخر يصفها بانها اساس الفشل التربوي.

ويشير Kline (2013) الى ان الاعتماد المتزايد على اختبارات الفروض الصفرية دمر تراثنا البحثي. وتحيز النشر للابحاث المتضمنة الدلالة الاحصائية

Publication bias for significance يقترح ان النسبة الحقيقية او الواقعية للخطأ من النوع الاول او الرفض غير الصحيح للفرض الصفرى تكون اعلى من المستويات التقليدية لمستوى الدلالة بالاحصائية المفترضة.

وحدث جدل ومناقشات مرتبطة بمدى جدوى الاعتماد على الدلالة الاحصائية، فالبعض نادى باستبعادها من البحوث المنشورة في المجلات العلمية وتمت مناقشة فكرة الاستبعاد في اعداد او

اصدرات خاصة من مجلات مثل Journal of Experimental education

(Thomposon, 1993) و Psychological Science (Shrout, 1997) وغيرها وكان

السؤال الرئيسي: "ماذا يكون الوضع بدون استخدام اختبارات الدلالة الاحصائية؟"، وقد

Armstrong (2007) النصيحة الاتية:

الكتابة لكتاب او تقارير بحوث فلا تذكر اختبارات الدلالة الاحصائية. والكتابة لمجلات فلا بد للبحث عن طرق لتقليل الاضرار المحتملة جراء اختبارات الدلالة، ويجب استبعاد كلمة دلالة لان النتائج التى ترفض الفرض الصفري ليست بالضرورة دالة وتلك التى تفشل فى رفض الفرض الصفري ليست بالضرورة غير دالة.

ونادى بعض الباحثين بهجر الدلالة الاحصائية فى الدراسات والبحوث امثال (Schmidt, 1996; Shmidt & Hunter, 1997) ، حيث اشاروا الى ان الدلالة الاحصائية اعافت تقدم المعرفة العلمية ولا تقدم اسهامات ايجابية للعلم وهى خطأ رهيب واستراتيجية علمية فقيرة واحدى الاشياء السيئة التى حدثت فى تاريخ علم النفس. وهى استخدام سئ لفترة طويلة (Cohen, 1994)، بل تقود الى صناعة قرار بحثى فقير وضعيف. وقدمت الجمعية النفسية الامريكية فى اصدارها الخامس والسادس (APA (2001, 2010) عدة توصيات فيما يخص اختبارات الدلالة الاحصائية اهمها:

- اعرض الاحصاء الوصفى مثل المتوسطات والتباينات وحجم العينة لكل مجموعة لان هذه المعلمات ضرورية لاجراء دراسات لاحقة ثانوية مثل ما وراء التحليل.
- اعرض احجام التأثير كلما امكن وعدم عرضه يعتبر قصور فى الدراسة.
- من الضرورى استخدام فترات الثقة كاستراتيجية لصناعة القرار.

وشدد (Rosnow & Rosenthal (1989 على ضرورة تقدير حجم التأثير عند تقدير للنتائج الدالة إحصائياً وغير الدالة إحصائياً وهذا عكس ما يعتقد الباحثين من أن تقدير حجم التأثير يكون للنتائج الدالة إحصائياً فقط . وأشار Cohen (1990) الى ان الناتج الاولى للبحث همقياس او اكثر لحجم التأثير وليس قيم P الاحتمالية.

علاوة على الاعتماد الكلى على اختبارات الفروض الصفرية وتجاهل القوة الاحصائية وحجم التأثير وماينتج عنه من استنتاجات خاطئة، توجد اشكالية اخرى هى ازمة القياس Measurement crisis فى العلوم السلوكية منذ عهود كثيرة حيث تفنقر البرامج الدراسية

للطلاب فى المراحل الجامعية الاولى الى المقررات السيكومترية ولا يتم تدريب طلاب الدراسات العليا فى برامج الماجستير والدكتوراة على كيفية اعداد المقاييس بالصورة المرجوة.

وبالاضافة الى محددات قيمة P الاحتمالية حيث تمدنا باجابة عن السؤال البحثى فى ضوء نعم او لا(رفض او الفشل فى رفض الفرض الصفرى) وهذا لايعطى متعة فى العمل العلمى ولا تخبرنا عن المناسبة العلمية او الدلالة الاكلينيكية او حجم التأثير وهذا ما دعا Armstrong (2007) الى القول بان اختبارات الدلالة الاحصائية لا تساعد على التقدم العلمى.

وفى الانتقادات للدلالة الاحصائية اوصى Cumming (2012) يجب على الباحثين اعطاء قليل من الاهتمام لقيمة P والتاكيد على عينة نتائج البحوث و اشار الى مصطلح الاحصاء الجديدة New statistics وهى ليست احصاء جديدة على الاطلاق انما تتكون اساسا من التاكيد على استخدام حجم التأثير وفترات الثقة وهذا ما اكدت عليه من قبل الجمعية النفسية الامريكية (APA, 2010) عن حجم التأثير " لمعرفة حجم او اهمية نتائج الدراسة يجب بقدر المكان كلما كان متاح تضمين بعض مقاييس حجم التأثير " .

والدلالة الاحصائية لاتشير مباشرة الى حجم التأثير ولان هذا المؤشر هو احصائية للعينة فهو تقدير تقريبي لحجم التأثير فى المجتمع بينما تقدير فترة الثقة يشير الى اى درجة خطأ المعاينة مرتبط بالاحصائية وعلى الرغم يتم تقدير خطأ المعاينة فى اختبارات الفروض الصفرية لكن يتم اخفائه فى حساب قيمة P ولكن مقدار من خطأ المعاينة يظهر فى ضوء الحدود الدنيا والعليا لفترات الثقة.

وعرض فترات الثقة يعكس التفكير التقديرى Estimation thinking (Cumming, 2012) واوصت الجمعية النفسية الامريكية (APA, 2010) كلما كان متاح تتم مناقشة النتائج وتفسيرها فى ضوء مستوى الدلالة و تقديرات الثقة.

وفى ضوء الحديث على محددات الدلالة الاحصائية فان التجديد الاحصائى يتطلب:

- الاقلال من استخدام اختبارات الدلالة الاحصائية كلما امكن.
- توجيه الاهتمام الى انواع اخرى من الاحصاء مثل حجم التأثير وفترات الثقة.
- التاكيد على اهمية الاعادة للدراسات Replication خاصة فى العلوم السلوكية.

واشار (2001) Cumming & Finch الى ضرورة ان يفكر الباحثين تفكيراً ماوراء تحليلاً
Meta-analytic thinking الذى يركز فى المقام الاول على تراكم الادلة المتنوعة عبر
الدراسات ومن اهم مظاهره:

- التقدير الدقيق لنتائج الدراسات السابقة.
 - يجب ان تكون دراسة الباحث اسهمت فى تراث الظاهرة فهى وحدها غير قادرة على الوصول لحل قضية اساسية، فدراسة العينات الصغيرة غيرقادرة على حل ابسط القضايا.
 - الباحث يعرض نتائج الدراسات من خلال تضمينها فى دراسات مستقبلية لما وراء التحليل.
 - تفسير النتائج الجديدة تتم بمقارنتها باحجام الاثر السابقة للظاهرة موضع الدراسة.
- والتفكير ما وراء التحليلي لا يركز فقط على الاختبارات الاحصائية انما يعتمد على احجام
التاثير عبر مجموعة من الدراسات الاولية المرتبطة وعليه فان التركيز يكون على حجم
التاثير وليس الدلالة الاحصائية فى الدراسات المفردة.

ماهية ما وراء التحليل

وفى السنوات الاخيرة زاد الاهتمام باستخدام المراجعات المنظمة فى العلوم الاجتماعية والنفسية
هذه المراجعات تمدنا بطريقة شفافة وموضوعية لتلخيص التراث البحثى لتحسين القرارات
والسياسات وتحسين تصميم الدراسات الجديدة فى المجال وهذا الكتاب يحاول ان يقدم
استراتيجيات لتخطيط وتنفيذ وتفسير احد مظاهر المراجعات المنظمة وهو اسلوب ما وراء التحليل.

فى عام 1952 اثار Hans Eysenck جدلاً واسعاً فى مجال علم النفس الاكلينكى وذلك بان
برامج العلاج النفسى ليس لها جدوى او فائدة على المرضى، وفى منتصف عام 1970 وجدت
المئات من الدراسات فى العلاج النفسى بعضها اعطى نتائج موجبة وبعضها نتائج سالبة
وبعضها نتائج صفرية، والمراجعة الكيفية لهذة الدراسات لم تستطع حل هذا الجدل القائم بمدى
جدوى برامج العلاج النفسى.

وللرد على ادعاء Eysenck قام (1977) Smith & Glass بمراجعة كمية احصائية معيارية
وذلك بحساب متوسط الفروق ابين متوسطات لمجموعات التجريبية والضابطة لـ 375 دراسة فى

مجال العلاج النفسى وسما هذه الطريقة الاحصائية لمراجعة الدراسات السابقة كميأ بماوراء التحليل Meta- analysis واستنتجا ان برامج العلاج النفسى فعالة، ولكن (Eysneck, 1978) لم يقتنع بهذه الطريقة وحاول ان يقلل من شان هذه الطريقة بقوله انها ممارسة الغباء الذائد عن الحد An exercise in mega-silliness وبالرغم من انتقاد Eysneck والآخرين لهذا الاسلوب الا انها اصبحت اكثر انتشاراً وقبولاً كطريقة لتلخيص نتائج الدراسات الامبريقية داخل العلوم الاجتماعية والسلوكية والطبية.

وفى نفس الوقت الذى طور فيه Glass طريقة ماوراء التحليل ظهرت مداخل احصائية اخرى لتوليف الابحاث على يد (Rosenthal & Rubin, 1978) فى علم النفس الاجتماعى لدراسة التأثيرات المتوقعة للعلاقات الشخصية، وعلى يد (Schmidt & Hunter, 1977) فى علم النفس الصناعى التنظيمى لدراسة مصداقية اختبارات الانتقاء الوظيفى، وكذلك ظهرت عدة كتب فى مطلع 1980 تناولت ماوراء التحليل (Glass, McGraw, & Smith, 1981; Hunter, Schmidt & Jackson, 1982; Rosenthal, 1984; Wolf, 1986) ، منذ ذلك الوقت اصبح اسلوب ماوراء التحليل طريقة للتوليف الكمي لنتائج الدراسات السابقة وتم اجراء الالاف من دراسات ماوراء التحليل فى كافة التخصصات. ثم قام (Hedges & Olkin, 1985) بتطوير الطرق الاحصائية لاسلوب ماوراء التحليل، ثم ظهر اول مجلد يتضمن عدة دراسات فى مجال توليف الابحاث Handbook of Research Synthesis على يد (Cooper & Hedges, 1994).

ومنذ ذلك الوقت اصبح ماوراء التحليل اسلوباً اكثر استخداماً وشيوعاً فى ابحاث العلوم الاجتماعية والنفسية، واصبح على الباحثين ان يكونوا اكثر الماماً بهذا الاسلوب لفهم التراث البحثى.

يمكن القول بان المنظر والمطور لاسلوب ماوراء التحليل هو Glass واول من اطلق تسمية ماوراء التحليل فى مقال عام 1976 المعنون "التحليل الاولى والثانوى وماوراء التحليل للبحوث"، ولكن مطورى الاساليب والطرق الاحصائية المتقدمة وكيفية توظيفه يرجع الى فريقين هما مجموعة Hedges وزملائه ومجموعة Hunter & Shmidt. ويرى البعض ان Hedges

ملك ما وراء التحليل على الاطلاق. ولا احد ينسى اعمال Rosenthal فى تطوير ما وراء التحليل.

ومنذ ابتدع علماء النفس اسلوب ما وراء التحليل منذ 1970 اصبح اكثر استخداماً فى كافة التخصصات وخاصة الطبية وحدثت تطورات كبيرة فى هذه المنهجية سواء بالكتب، الدراسات، والنماذج التحليلية، ومن التطورات الحديثة تنفيذ ما وراء التحليل باستخدام برامج احصائية متخصصة مثل CMA و MetaXL وكذلك فى البرمج الاحصائية العامة مثل SPSS, SAS, Stata والبرنامج الاحصائى المتقدم R .

ومصطلح Meta يعنى ما بعد Beyond وتحليلات ما راء تطلق على اى تحليل ابعد من الدراسات الاولية، ويطلق عليه البعض التحليل البعدى وهى تسمية مقبولة ولكنها تتداخل مع التحليل البعدى المصاحب لتحليل التباين لتحديد المسئول عن مسبب الدلالة الاحصائية.

محددات الدراسات الاولية Limits of Study Design

- **محددات تصميم الدراسة:** تسمح التصميمات التجريبية بعمل استدلالات عن السببية ولكن توجد اشكالية فيما يخص الصدق البيئى Ecological validity وهو يوازى الصدق الخارجى بمعنى الى اى مدى نتائج الدراسة قابلة للتطبيق فى الواقع الحقيقى، بينما التصميمات الارتباطية لها ميزة فى امدادنا بصدق بيئى افضل من التصميمات التجريبية ولكنها لا تستطيع امدادنا بادلة صادقة عن السببية، وهذه المحددات لتصميم الدراسة الاولية(مشاكل الصدق الداخلى) سوف تبقى فى دراسة ما وراء التحليل مثل اشكالية السببية للدراسات الارتباطية ومشكلة الصدق الداخلى للدراسات التجريبية.

- **محددات المعاينة Limits of Sampling:** النواتج من الدراسات الاولية التى تستخدم عينات متجانسة بغض النظر عن خصائص معينة (مثل الجنس، المستوى الاقتصادى الاجتماعى، العمر، وغيرها) تخبرنا عن فهم المجتمعات التى لها نفس خصائص عينة الدراسة وتكون النتائج قاصرة على هذه النوعية من المجتمعات، نفس القضية موجودة فى استنتاجات دراسات ما وراء التحليل ولكن من اهم مميزات

ما وراء التحليل تقييم النتائج فى ضوء خصائص العينة او الدراسات من خلال التحليلات المتفاعلة بالتالى يمكن تعميم النتائج على مجتمعات غير متجانسة.

- **محددات التشوهات المنهجية Limits of Methodological Artifacts:** هذه

التشوهات تظهر اثناء اجراء الدراسة ومنها المقاييس غير تامة الصدق والثبات، المدى المقيد لمتغيرات الدراسة، التصنيفية الثنائية للمتغيرات المتصلة، وضعف الاختيارات التحليلية. وهذه التشوهات تقلص اوتنقص احجام التأثير المقدر من الدراسات الاولية وتعطى قوة احصائية منخفضة. ولكن ما وراء التحليل يسعى الى تصحيح هذه التشوهات للوصول الى حجم التأثير الحقيقى خاصة مع استخدام مدخل (Hunter & Schmidt (1990, 2004).

- **محددات القوة الاحصائية Limits of Statistical Power:** تشير القوة

الاحصائية الى ان قدرة الدراسة للكشف عن التأثير الحقيقى الموجود بالفعل عندما يكون حقيقة موجود فى المجتمع وتتحدد القوة الاحصائية بعدة عوامل منها الخطأ من النوع الاول، الاختبار الاحصائى، حجم التأثير فى المجتمع، حجم العينة، وغيرها. ويمكن اجراء تحليل القوة الاحصائية القبلى لتحديد حجم العينة المتطلب لتحقيق قوة 0.80، ولكن تعاني الدراسات الاولية من انخفاض القوة الاحصائية وهذا بدوره لا يكشف عن التأثيرات الحقيقية فى المجتمع، وعندما يتضمن التراث البحثى دراسات فشلت فى الكشف عن التأثيرات الحقيقية نتيجة نقص القوة الاحصائية فمن الخطورة الاعتماد على هذه الدراسات لانها لا تقدم الصورة الحقيقية للتأثيرات، ولكن فى هذه المواقف فان ما وراء التحليل يدمج الدراسات منخفضة القوة مع الدراسات مرتفعة القوة فى الدراسة الواحدة وهذا من شأنه يؤى الى الوصول الى قوة احصائية مناسبة.

وعلى الرغم ان دراسات ما وراء التحليل نفسها قد تكون منخفضة القوة الاحصائية ولكنها فى المجمل افضل من الدراسات الاولية (Cohen & Becker, 2003) ، وان دراسات ما وراء التحليل اقل تائراً من انخفاض القوة الاحصائية من الدراسات الاولية.

الحاجة الى توليف البحوث فى كافة التخصصات

فى ظل تراكم الكم الهائل من البيانات البحثية، فمن الصعب على الباحثين الامام بالتطورات المتلاحقة فى تخصصه وعليه فلا مفر من استخدام اسلوب ما وراء التحليل كمحاولة للتوليف بين التراكم البحثى الهائل خاصة فى العلوم الاجتماعية والنفسية والوصول الى استنتاج عام حول هذه البحوث المتراكمة منذ عشرات السنين خاصة ان نتائجها تكون متعارضة، وقيمة التأثير لمتغيراتها تكون صغيرة نسبياً مقارنة بالعلوم الطبيعية. ويرى (Schmidt 1992) ان كثيراً من الاكتشافات والاسهامات فى المعرفة المتراكمة ليست من الدراسات البحثية الاولية ولكن باستخدام ماوراء التحليل لاكتشاف المعنى الخفى من التراث البحثى الموجود.

واحد معوقات تقدم العلوم الانسانية هو عدم القدرة على تنظيم وتوليف النواتج الامبريقية للدراسات السابقة. ويرى (Card 2012) ان البحث فى العلوم النفسية والاجتماعية فى اشد الحاجة لفحص وتنظيم التراث البحثى المتراكم لعدة عقود وليس لاجراء دراسات جديدة، فعلى سبيل المثال توجد الالاف من الدراسات التى تناولت العلاقة بين الذكاء والتحصيل، ومن هنا كانت الحاجة الملحة للمراجعة المنظمة للنتائج المتراكمة للدراسات السابقة ومحاولة دمج او توليف النتائج للوصول الى استنتاج عام لهذة الدراسات بغض النظر عن مصادر الاختلاف عبر الدراسات مثل المنهج او العينة او ادوات القياس او الثقافة او البيئة.

ويستخدم ما وراء التحليل فى كافة المجالات كالتالى (Borestein et al ., 2009):

- **الطب:** شكلت دراسات ما وراء التحليل والمراجعات المنظمة محور اهتمام المؤسسات الطبية لتقييم فعالية المعالجات الطبية فى ضوء البيانات الامبريقية المتاحة لتراكم كم هائل من التجارب. فعلى سبيل المثال طبعت (Cochrane Collabrations 2009) حوالى 3700 دراسة ما وراء التحليل لتوليف بيانات المعالجات فى كافة مجالات العناية الصحية مثل الصداع، السرطان، امراض القلب، الاكتئاب، وغيرها فى كافة المراحل العمرية. وتوجد سلسلة من المجالات الطبية تكون نتائجها فى ضوء ما وراء التحليل.
- **الصيدلة:** واجرت شركات الصيدلة سلسلة من الدراسات لتقدير فعالية عقار ما واستخدموا ما وراء التحليل لتوليف البيانات من دراسات مختلف لاعطاء اختبار اكثر قوة لدراسة تأثير العقار

وايضاً تمدنا ما وراء التحليل باطار لتقويم سلسلة من الدراسات ككل فضلاً عن النظر اليها بصورة منعزلة وايضاً تستخدم لتقويم الاعراض العكسية او الجانبية0

- **التربوية:** تم تطبيق ما وراء التحليل فى موضوعات مختلفة مثل المقارنة بين التعلم عن بعد فى مقابل التعلم المدرسى الفصلى التقليدى على التحصيل والاتجاه نحو المدرسة والاستقلالية، او العلاقة بين خصائص المعلم على التحصيل، او اثر برامج تدريسية على التحصيل وغيرها. ونتائج ما وراء التحليل لها اثر فى الممارسات و السياسات التربوية.
- **علم النفس:** نشأ ما وراء التحليل فى احضان علم النفس وتم تطوير فنياته عل يد علماء علم النفس، ويستخدم لدراسة تطور الشخصية عبر المراحل العمرية لمختلفة، وتقدير اثر برامج العنف على سلوك العدوان، ودراسة الفروق بين الجنسين فى القدرة الحسابية واللغوية والقيادة والتواصل الشفوى، وتقدير فعالية البرامج النفسية فى خفض القلق والوحدة النفسية وغيرها.
- **مجال الجريمة:** فان المؤسسات الحكومات تقدم دعم لاجراء ما وراء التحليل لتقدير الفعالية النسبية للبرامج المختلفة لتقليل سلوك الجريمة، النصب، الاحتيال، التشدد وسلوك الارهاب.
- **مجال الاعمال:** يستخدم ما وراء التحليل للتحقق من الصدق التنبؤى لاختبارات انتقاء الموظفين فى المؤسسات المختلفة والتحقق من فعالية برامج تدريب الموظفين.
- **البيئية:** يستخدم ما وراء التحليل لتقدير الاثار البيئية على التلوث البيئى بكافة اشكاله وعلى التغيرات المناخية.

مفهوم ما وراء التحليل

يعتبر اسلوب ما وراء التحليل مدخل منهجى واحصائى شامل للوصول الى الاستنتاجات من التراث البحثى واحد الطرق لاحداث الدمج والتوليف المنظم للتراث الامبريقي(الدراسات الاولية) لظاهرة ما بهدف التكامل بين النواتج للوصول الى تعميمات او حل التناقضات بين الدراسات. ويعرف ما وراء التحليل بانه التحليل الاحصائى لمجموعة كبيرة من نتائج الدراسات المفردة بهدف احداث تكامل لهذة النواتج (Glass, 1976). وماوراء التحليل طريقة لاجراء التحليل الاحصائى لنواتج البحوث لكثير من الدراسات المفردة المستقلة لموضوع محدد (Borenstein et al., 2009; Hedges & Olkin, 1985)، الطريقة المستخدمة للوصول الى حجم تاثير من الدراسات المفردة (Durlark, 1995)، التوليف الكمى لبيانات من دراسات متعددة (Leandro,)

2005)، مجموعة من الطرائق الاحصائية لتوليف النتائج الكمية من دراسات متعددة لاعطاء ملخص للمعرفة الامبريقية لموضوع محدد (Littlell, Corcoran, & Pillai, 2008))، مدخل قوى وفعال لتلخيص ومقارنة نتائج التراث البحثى لظاهرة ما (Card, 2012)، مراجعة منظمة مع اعطاء ناتج مشترك للدراسات مع دقة تقديرات عالية، (Cleophas & Zwinderman, 2017).

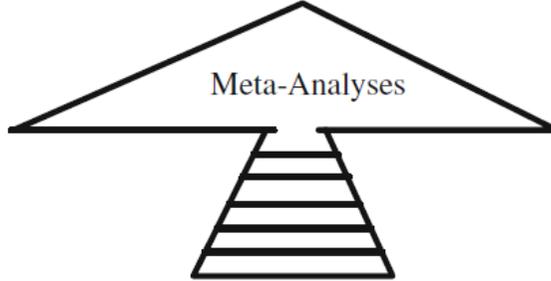
ويطلق على اسلوب ما وراء التحليل بالتوليف الكمي Quantitative synthesis ويستخدم احياناً كمرادف للتوليف البحثى، ويتضمن اسلوب ماوراء التحليل مجموعة من الاساليب الاحصائية المستخدمة فى المراجعة المنظمة للدراسات السابقة لتقليل التحيز كلما امكن. والاجراءات الاحصائية المستخدمة فى اسلوب ماوراء التحليل يمكن تطبيقها لاي مجموعة بيانات. ويهدف ماوراء التحليل الى تحديد وتقدير ثم توليف البيانات من عدد كبير من المصادر. يهدف ماوراء التحليل الى الوصول الى استنتاج وحيد وشامل لنتائج دراسات مختلفة لنفس الموضوع باستخدام نفس الطرق، ويساعد فى فهم نتائج تجارب التدخل (المنهج التجريبي وشبه التجريبي).

لا يساعد ماوراء التحليل فى الحصول المعرفة التراكمية فقط بل يعطى مؤشر واتجاه واضح عن ماهى الفجوات او الموضوعات فى التراث البحثى التى هى بحاجة لمزيد من الدراسة والبحث فى المستقبل.

وينصح (Cooper 1998) من الاقلال من مصطلح ماوراء التحليل او التحليل الاحصائى لنتائج الدراسات المتعددة واستبداله بالمراجعة المنظمة Systematic review اوتوليف البحوث Research synthesis على اساس انها عملية للبحث فى التراث وتقويم الدراسات وهكذا، ولكن (Littell et al. 2008) يرى ان المراجعة المنظمة وماوراء التحليل مفهومان ليس مترادفان ولكنها من المداخل الجيدة لتوليف البحوث ويمكن اعتبار ماوراء التحليل جزء من المراجعة المنظمة ولكن هذا ليس كذلك لكل دراسات ماوراء التحليل لان مصطلح المراجعة المنظمة يمكن ان تكون بصورة كيفية، ويرى (Card 2012) من الافضل استخدام ماوراء التحليل او المراجعة الما وراء التحليلية Meta-analytic review لانه اقل ارتباكاً وأكثر الفة

للباحثين وقد استخدمت معظم الكتب في هذا الشأن مصطلح ماوراء التحليل (انظر Borenstein et al.,2009 ;Hedges & Olkin, 1985; Hunter & Schmidt, 2004)

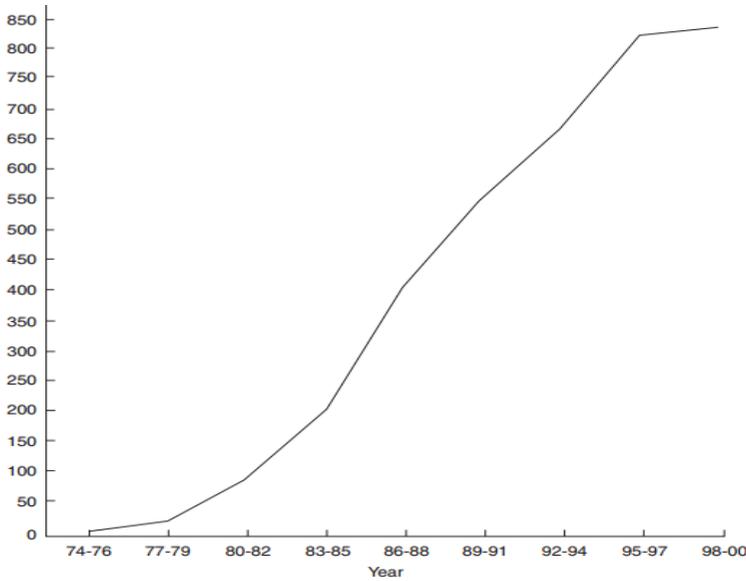
ويمكن تشبيهه كالآتي:



الشكل (1.1): ما وراء التحليل كمظلة للدراسات المفردة.

حيث يعتبر المظلة العامة لكل الدراسات المفردة.

وتزايد استخدام دراسات ماوراء التحليل من لا شئ في المدة من 1974 الى 1976 الى 835 في المدة من 1998 الى 2000 وذلك في قاعدة بيانات المعلومات النفسية PsycINFO وذلك في ملخصات او عناوين الدراسات. الشكل الاتي يوضح استخدام ماوراء التحليل منذ 1974 حتى 2000 في قاعدة البيانات النفسية PsycINFO (Hunter & Schmidt, 2004) :



الشكل (2.1): استخدام ماوراء التحليل في PsycINFO في المدة من 1974 حتى 2000 م.

الحاجة الى ماوراء التحليل في العلوم السلوكية والاجتماعية

الملاحظ تراكم كم هائل من الدراسات فى تراث علم النفس والعلوم الاجتماعية، فالحاجة اليوم ليس لمزيد من الدراسات الامبريقية التى تتناول ظاهرة ما بل لاعطاء نظرة تقييمية شاملة للكم الهائل من البيانات المتراكمة على مدى اكثر من خمسين عاماً بمعنى اعادة تحليل هذا التراث البحثى المتصارع، والاتجاه الى التكامل بين النتائج المتعارضة لمحاولة الوصول الى معرفة عامة وشاملة تتوحد فيها هذه النتائج المتضاربة فى محاولة الوصول الى بناء نظرية متكاملة.

فالدراسات ذات احجام العينات الصغيرة Small-sample studies فى البحث النفسى التى تحاول الاجابة على نفس السؤل ينتج عنها نتائج متضاربة وهذه الدراسات تعتمد على اختبارات الدلالة الاحصائية وهذا يجعل نتائج الدراسات اكثر تعارضاً لكون القرار ثنائى (نعم-لا)، ولذلك فان ما وراء التحليل يهدف الى حدوث تكامل عبر هذه الدراسات للكشف عن النمط العام للعلاقات التى تكمن فى التراث البحثى وهذا هو الاساس لبناء نظرية جديدة او تطويرها. ما وراء التحليل يصحح النتائج (التاثيرات- العلاقات) المتعارضة من خطأ المعاينة وخطأ القياس وى عوامل اخرى تؤدى الى ظهور النتائج المتعارضة.

هدف اى علم هو انتاج المعرفة المتراكمة التى تتضمن العلاقات والنظريات التى تستطيع تفسير الظواهر المرتبطة به. فعلى سبيل المثال النظريات التى تفسر كيفية تطور ونمو سمات الشخصية عند الاطفال والراشدين عبر الزمن وكيف لهذه السمات ان تؤثر فى حياتهم، او النظرية التى تفسر العوامل المسهمة فى الرضا الوظيفى او النظرية التى تفسر العوامل المسمة فى والتحصيل وغيرها.

وبدون تحديد واضح لطبيعة العلاقات بين متغيرات الظاهرة فلا نستطيع بناء نظرية متماسكة لتفسيرها، وكذلك الحصول نتائج متعارضة مثل الحصول على علاقات قوية موجبة وعلاقات قوية سالبة ولا علاقات فان هذا بدوره لا يساعد على بناء نظرية قوية متماسكة تستطيع تفسير جوانب الظاهرة المختلفة.

والمهمة الاساسية فى العلوم الاجتماعية والسلوكية هو بناء النظرية، والنظرية الجيدة هى التى لها القدرة على تفسير طبيعة العمليات والتفاعلات التى تحدث داخل الظاهرة وتفسرها بصورة سهلة

وبسيطة. والنظرية عبارة عن تفسيرات نظرية وعلاقات بين متغيرات الظاهرة وتنتج من تراكم من الاطروحات النظرية مقرونة بالتحقق الامبريقي من الدراسات.

ويلعب اسلوب ماوراء التحليل دوراً كبيراً فى الكشف عن طبيعة هذه العلاقات واستخدامها كمدخلات لاساليب احصائية متقدمة تستخدم للتحقق من العلاقات السببية مثل تحليل المسار Path analysis ونمذجة المعادلة البنائية Structural equation modeling.

على ذلك يمكن استخدام ماوراء التحليل للحصول على مصفوفة الارتباط Correlation matrix بين المتغيرات حيث كل معامل ارتباط فى المصفوفة هو ملخص لعدد كبير من الدراسات التى درست العلاقة بين المتغيرين.

وكما نعلم ان ماوراء التحليل يصحح معاملات الارتباط من عوامل التحيز مثل خطأ القياس وعلى ذلك يمكن تطبيق تحليل المسار او نمذجة المعادلة البنائية لاختبار النظريات السببية او الاستكشافية ويطلق على هذه الاساليب التى تدمج نتائج ماوراء التحليل و اسلوب تحليل المسار بـ ماوراء تحليل المسار Meta-path analysis او ماوراء نمذجة المعادلة البنائية Meta-SEM .

ومن خلال المقارنة الكمية لنواتج الدراسات المتنوعة فان ماوراء التحليل يكتشف معرفة جديدة غير واضحة فى الدراسات المفردة واحياناً يجيب على اسئلة لايمكن حلها من الدراسات المفردة، فالدراسة المفردة غير قادرة على الحكم على مقارنة فعالية Effectiveness برنامج تدريبي سواء للافراد مرتفعى ومنخفضى القدرة العقلية، ومن خلال مقارنة متوسط احصاء حجم التأثير Cohen d عبر المجموعات المختلفة للدراسات فان ماوراء التحليل يجيب عن هذا السؤال، وكذلك يمكن دراسة المتغيرات المتفاعلة Moderator variables التى احياناً يصعب دراستها من خلال الدراسات المفردة عن طريق ماوراء التحليل.

وفى الواقع لاتوجد دراسة تامة بدون اخطاء، فكل الدراسات تتضمن اخطاء قياس فى كل قياساتها حيث من الصعب القول انها تامة الصدق البنائى خاصة فى العلوم السلوكية والنفسية حيث تتداخل وتتشابك فيها المتغيرات ببعضها البعض بالتالى فان المفهوم فى حد ذاته ليس خالصاً نقياً (Hunter & Schmidt, 2004)، وبالإضافة الى خطأ القياس فان الدراسات تعانى من

خطأ المعاينة حيث تتضمن الدراسة مجموعة صغيرة من المجتمع وهذا بدوره يؤدي الى نتائج متعارضة ويبدو هذا واضحاً في العلوم السلوكية والطبيعية، وتناول Hedges (1987) باستخدام طرق ما وراء التحليل لدراسة الاختلاف لنتائج الدراسات في 13 مجال في الفيزياء الطبيعية و13 مجال بحثي في علم النفس اظهرت النتائج وجود اختلافات كبيرة عبر دراسات الفيزياء كما هو الحال في علم النفس، بمعنى ان الاختلاف في نتائج الدراسات لا يختلف من تخصص الى اخر بغض النظر عن طبيعة المفاهيم المراد قياسها.

وما وراء التحليل ليس اسلوب لاعطاء رؤية شاملة عن التراث البحثي المترام ولكنه يمدنا باتجاهات واضحة عن طبيعة الدراسات التي نحن بحاجة اليها مستقبلاً بمعنى ما نوعية الدراسات الاولية التي نحتاجها للتغلب اوسد فجوة بحثية او التغلب على بعض المحددات التي يكشف عنها اسلوب ما وراء التحليل.

ويبدو تأثير ماوراء التحليل واضح في المجالات النفسية الامريكية التي تتبع APA خاصة Psychological Bulletin فقد زاد استخدامه في الدراسات المنشورة في الفترة الاخيرة ويؤكد معظم محرري المجالات على توظيف هذا الاسلوب في دراساتهم كمدخل لمراجعة الدراسات السابقة لان هذا بدوره يحسن جودة الاستنتاجات من التراث البحثي مما ينعكس ايجابياً على بناء نظرية متماسكة واكثر قدرة على تفسير الظاهرة.

ويرى (1997) Hunt ان ماوراء التحليل اكثر تأثيراً في ابحاث المجال الطبي من العلوم النفسية والاجتماعية، ويبدو ان فعالية ماوراء التحليل اكثر في العلوم الطبية نظراً لطبيعة الابحاث التجريبية التي تعتمد على عدد محدود من الافراد قد يصل الى فرد واحد من ثم فان ماوراء التحليل يسعى الى احداث تكامل بين هذه الدراسات للوصول الى استنتاج عام.

فاذا كان باحث مهتم بدراسة اثر المتغير المستقل X على المتغير التابع Y وبدأ في جمع التراث المتاح في هذا المجال، فالباحث يصنف كل نتائج الدراسات في ثلاثة تصنيفات، الاول يتضمن الدراسات التي اثبتت وجود علاقة موجبة ودالة احصائياً، والثاني يتضمن تلك الدراسات التي توصلت الى علاقة سالبة ودالة احصائياً، والثالث يتضمن تلك الدراسات التي توصلت الى علاقة غير دالة احصائياً، وعليه فالباحث غير قادر للوصول الى استنتاج من هذه الدراسات

المنفصلة، ويرى ان هذا مدعاة لاجراء دراسته وبعد اجراء الدراسة ووصل الى وجود تأثير سالب ويفسر نتائج دراسته بان هذا يتفق مع دراسات ويتعارض مع دراسات اخرى ، ويصدر توصية باجراء مزيد من الدراسات للوصول الى تأكيد الى هذه النتيجة ويظل البحث في هذا المجال على هذه الوتيرة وهكذا دواليك وهذا لا يحدث تطور في العلم، وعليه فان الدراسات امفردة الاولية نادراً ما تحل التناقضات في العلوم الاجتماعية والنفسية (Ellis, 2010).

ويرى (Hunter & Schmidt (2004 ان الحقيقة المؤلمة ان معظم التراث البحثي اظهر نتائج متعارضة فبعضها وجد علاقات دالة احصائياً والبعض غير دالة احصائياً. ووجدت في التراث النفسى والاجتماعى تقريباً %50-50 (Cohen, 1962; Sedlmeier & Gigerenzer, 1989) بمعنى ان نصف النتائج دالة احصائياً ونصفها غير دالة احصائياً، وهذا هو الحال الطبيعى في معظم العلوم السلوكية والاجتماعية، وعليه فمن الصعب بناء نظريات ومعرفة متراكمة متماسكة، من ثم فان ماوراء التحليل يلعب دواً كبيراً في عملية البحث في العلوم النفسية والاجتماعية.

المراجعة المنظمة الكيفية السردية وماوراء التحليل هما طريقتين متميزتين لتوليف الدراسات السابقة ويستخدمان لتحديد وتحليل وتوليف الادلة الكمية من الدراسات السابقة.

المواقف المناسبة لاستخدام ماوراء التحليل

يعتبر اسلوب ماوراء التحليل شكل من اشكال البحث المسحى حيث تعتبر تقارير البحوث بمثابة عينة الافراد التى يتم اجراء المسح عليهم . ويوضع فى شكل التكويد(بروتكول المسح) عينة اومجتمع تقارير البحوث التى يتم تجميعها من التراث البحثى وكل دراسة يتم تكويدها(المقابلة فى لغة البحث المسحى) عن طريق المكود او المرمز الذى يقرأ الدراسة بعناية ويكود المعلومات المناسبة عن خصائصها ونواتجها الكمية ثم تحلل البيانات المجمعة باستخدام اساليب احصائية تقليدية لوصف نمط النواتج لمجموعة الدراسات . ويشير (Lipsey & Wilson (2001 الى ما وراء التحليل فقط احد الطرق لتلخيص وتكامل وتفسير مجموعة منتقاة من الاعمال البحثية فى مختلف التخصصات، ومن اهم المواقف التى يستخدم فيها ماوراء التحليل وهى:

- يستخدم فقط فى الدراسات البحثية الامبريقية(الكمية) وبذلك لايمكن استخدامه لتلخيص الدراسات النظرية اوالدراسات الكيفية التى تهتم بالمراجعات النظرية ودراسات السياسات العامة.
 - يستخدم مع الدراسات ذات النواتج الكمية التى تستخدم مقاييس كمية لقياس متغيراتها وتعرض الاحصائيات الوصفية والاستدلالية فى تقارير بحوثها. وهذا ربما لا ينطبق على بحوث دراسة الحالة والبحث الانثربولوجى.
 - ماوراء التحليل اسلوب لتكويد وتحليل الاحصائيات التى تلخص نواتج البحوث باستخدام اجراءت احصائية تقليدية.
- ولان ماوراء التحليل يهتم بتجميع ومقارنة النتائج من دراسات بحثية مختلفة فمن الضرورى ان تكون مقارنة هذه النتائج ذات معنى وهذا يفترض ان:
- تتعامل الدراسات مع نفس المفاهيم والعلاقات حيث تكون المقارنة بينها لها معنى وقابلة للتفسير.
 - تتعامل الدراسات مع طرق واساليب احصائية متماثلة.
- فعلى سبيل المثال لو ان مجموعة من الدراسات اهتمت بدراسة فعالية معالجات لعلاج الاكتئاب فيمكن اجراء ماوراء التحليل لهذه الدراسات لو ان المعالجات المختلفة مناسبة للمقارنة ببعضها البعض ولو ان نواتج المعالجات لها نفس الشكل مثل قياسات الاكتئاب لمجموعة ضابطة وتجريبية من المستجيبين.
- لايمكن استخدامه لدراسات تتضمن موضوعات مختلفة فى نفس ماوراء التحليل مثل دراسات تتناول برامج لعلاج الاكتئاب ودراسات تتناول الفروق بين الجنسين فى الوحدة النفسية، فيشار الى هذا باشكالية "التفاح والبرتقال" فى ماوراء التحليل بمعن استخدامه لتوليف دراسات لا تتعامل مع نفس المفاهيم او العلاقات ولكن هذا ما اتبعه Smith & Glass (1977) حيث دمج نواتج دراسات من معالجات مختلفة مثل العلاج النفسى المعرفى، العلاج السلوكى المعرفى، العلاج الجشتالتى وغيرها فى دراسة واحدة للحصول على الفعالية العامة لهذه البرامج.

بالمثل ليس من الافضل تضمين نواتج دراسات من تصميمات بحثية مختلفة التي تتضمن اساليب احصائية مختلفة حتى لو تعاملت مع نفس الموضوع. على سبيل المثال لا تتضمن دراسات لتاثير معالجات على الاكتئاب باستخدام تصميم المجموعة الضابطة والتجريبية مع دراسات من منهجية بحثية اخرى تتعامل مع الاكتئاب كدراسات الملاحظة او العلاقة. وعليه نتائج من تصميمات بحثية مختلفة بمعالجات احصائية مختلفة من الصعب دمجها فى نفس ماوراء التحليل ولكن من الافضل اجراء ماوراء تحليل منفصل للنواتج التجريبية وللنواتج دراسات العلاقة.

ماوراء التحليل فى مقابل التحليل الاولى او الثانوى

قبل تناول اسلوب ماوراء التحليل لابد من التفريق بين مصطلحات ماوراء التحليل و التحليل الاولى Primary analysis والتحليل الثانوى Secondary analysis . ويرى Glass (1976) ان التحليل الاولى يحدث عندما يقوم باحث بجمع البيانات من العينة او الشركات او اى وحدة معاينة خلال تطبيق ادوات القياس ثم تحليلها للاجابة على اسئلة بحثه او دراسته وهذا ما يحدث فى معظم ابحاث الماجستير والدكتوراة وما بعد الدكتوراة.

اما التحليل الثانوى يشير الى اعادة تحليل بيانات تم جمعها مسبقاً للاجابة عن اسئلة بحثية مختلفة او الاجابة عن على نفس الاسئلة بطريقة منهجية اخرى مثل استخدام اساليب احصائية اخرى لم تتاح عند اجراء الدراسة الاولية . ودراسات التحليل الثانوى يمكن ان يجربها نفس الباحث او باحثين اخرين مختلفين اذا كان باستطاعتهم الحصول على البيانات الخام من صاحب الدراسة الاولية.

لاحظ ان التحليل الاولى والثانوى يتطلب الحصول على البيانات الخام الاصلية للدراسة. فى المقابل فان ماراء التحليل يتطلب التحليل الاحصائى لنتائج (ليست البيانات الخام) اكثر من دراسة (ليست لدراسة واحدة). فى هذا يتضح مظهرين لماوراء التحليل:

اولاً: يتضمن نتائج الدراسات كوحدة التحليل خاصة ان النتائج تكون فى صورة احجام التاثير، والحصول على احجام التاثير لايتطلب الوصول للدرجات الخام حيث يمكن حسابها

من البيانات المطبوعة في الدراسات الاولية او الثانوية المنشورة مثل المتوسطات والتباينات واحجام العينات ودرجات الحرية.

ثانياً: يعتبر تحليل لنتائج دراسات متعددة وعليه فان الدراسة المفردة هي وحدة التحليل ويتراوح حجم الدراسات من دراستين كحد ادنى الى مئات الدراسات فاكثر اذا كان متاح.

يهدف ماوراء التحليل الى الوصول الى استنتاج من عينة من الدراسات فى المقابل فان التحليلات الاولية والثانوية تهدف للوصول الى استنتاج من عينة من الافرد، وبالتالي يعتبر ماوراء التحليل شكل من اشكال مراجعة التراث البحثى.

وعليه فان ماوراء التحليل ليس دراسة اولية او ثانوية لحل قضية معينة او للاجابة عن سؤال بحثى محدد انما فقط هو اسلوب احصائى يسعى لتكامل النتائج الدراسات المختلفة لنفس الموضوع وباستطاعته ضبط خطأ المعاينة واى اخطأ اخرى غير مرغوبة لاعطاء اساس متكامل وشامل للاستنتاجات.

ماوراء التحليل احد اشكال مرجعة الادبيات البحثية

يعتبر ماوراء التحليل احد الطرق المنهجية المستخدمة فى مراجعة التراث او الادبيات البحثية لظاهرة ما، ويعرف (2012) Card مراجعة التراث Literature review بانها توليف للتراث البحثى الموجودة بالفعل لموضوع او ظاهرة ما، وتتضمن مراجعة الدراسات السابقة والنظريات ونتائج الدراسات وممارسات معينة، وتتوع طرق توليف الدراسات السابقة فمنها (Card, 2012; Hunter & Schmidt, 2004):

- **المراجعات السردية Narrative reviews:** تركز على التفسيرات النظرية لموضوع ما والبحث فى الى اى مدى يوجد تكامل بين اطر نظرية مختلفة تفسر ظاهرة ما وهذا سائد فى مقالات مجلات مثل مجلة Psychological Reviews ويطلق عليها المراجعة السردية Narrative review وهى طريقة كيفية وصفية وفيها يحاول الباحث الوصول الى التعارضات والتوافقات بين الدراسات ولكن لو ان عدد الدراسات اكثر من 50 دراسة فمن الصعب استخدام هذه المراجعة الكيفية ومن الصعب الوصول الى استنتاجات لطبيعة القياسات والتصميمات، ويصبح الباحث مشنت فى الوصول الى استنتاج فيما يخص العلاقة

بين المتغيرات موضع اهتمام الدراس ولايستطيع الباحث احداث تكامل بين الدراسات انما يضطر الى الاعتماد على عينة محدودة من الدراسات، ونتجة لذلك يهدر الباحث كثيراً من المعلومات التي تكون مفيدة في توجيه الدراسة المزمع القيام بها لانه لايستطيع الامام بكل الدراسات المتاحة في التراث البحثي. ويرى (Cooper & Rosenthal 1980) حتى لو ان عدد الدراسات المراد مراجعتها صغيراً مثل سبعة مثلاً فان مراجعة الدراسات بالطريق الكيفية ومراجعتها كمياً يؤديوا الى استنتاجات متعارضة.

وفي المراجعة السردية فان المراجع او الباحث يقيم البحوث ويصل الى استنتاجات ولكن من محددات هذه الطريقة لتوليف البحوث هي ان كيفية الوصول الى الاستنتاج غير معلومة وغير محددة وذاتية Subjectivity ومتحيزة بدرجة كبيرة، وباستخدام هذه الطريقة فان الباحث يحاول ان يوحى للآخرين بان نتائج الدراسات متضاربة او متصارعة وغير متناسقة. ويرى (Hunter & Schmidt 2004) ان اجراءت المراجعة التقليدية غير مناسبة لتكامل النتائج المتعارضة عبر عدد كبير من الدراسات. وفي هذا يشير (Glass 1976) ان النتائج المتحصل عليها من مئات الدراسات لا يمكن ادراكها او فهمها في المراجعة السردية التقليدية فهذا لا يعد كونه الا ادراك درجات 500 اختبار بدون استخدام اساليب لتنظيم ووصف وتفسير البيانات. وفي ضوء المراجعة الكيفية السردية يصل الباحثين الى استنتاجات مختلفة من نفس مجموعة الدراسات وعليه فان التعارض في التراث ينتج عنه تعارض بين المراجعين او الباحثين.

ويوجد خطوتين لاراء المراجعة الكيفية السردية التقليدية للدراسات السابقة:

اولاً: جمع التراث المنشور.

ثانياً: توليف واستنتاج المعلومات المطلوبة وهذا التوليف يتم عن طريق الباحث ولكن يوجد درجة من التحيز الشخصي Personal bias عند الوصول الى استنتاجات خاصة اذا وجد تعارض شديد بين النتائج. كما ان اجراء مراجعة لكم هائل من الدراسات في مجال ما عملية ليست سهلة وتزداد صعوبة اذا اكانت النتائج متعارضة فكيف يمكن اجراء مراجعة مثلاً لـ 100 دراسة او حتى 50 دراسة تتضمن نتائج متعارضة؟.

فاذا اهتم باحث ما بدراسة اثر المتغير المستقل X على المتغير التابع Y ، وبدأ بجمع كل الدراسات المتاحة فى التراث وبدأ بتصنيف الدراسات السابقة الى ثلاثة تصنيفات وفقاً الى النتائج. فى التصنيف الاول وضع النتائج الموجبة الدالة احصائياً وفى الجانب الاخر وضع الدراسات السالبة والدالة احصائياً وفى المنتصف وضع الدراسات غير الدالة احصائياً، ثم يجرى الدراسة التى بصدها وتوصل الى ان المتغير X له تاثير سالب ودال احصائياً على Y ويكتب ان نتيجة الدراسة تتفق مع الدراسات الاخرى التى توصلت الى نفس النتيجة وعليه فان الباحث غير قادر على الوصول الى استنتاجات من هذه النتائج المتعارضة. فالدراسات المفردة او الاولوية غير قادرة على حل عدم الاتساق فى العلوم الاجيماعية والسلوكية.

• **المراجعات المسحية Survey reviews**: تركز على الممارسات السائدة داخل مجال ما مثل استخدام طرق معينة او الاتجاهات الحديثة فى مجال العلاج النفسى فى الدراسات المنشورة، او اجراء مسح للتعرف على طبيعة المنهجية البحثية السائدة فى ابحاث الجريمة او فى التراث البحثى لعلم النفس المدرسى.

• **المراجعة فى طرق العد للدلالة الاحصائية للتاثيرات Vote counting methods**: حيث تؤخذ الدلالة الاحصائية فى الاعتبار عند مراجعة الدراسات السابقة سواء دالة موجبة او دالة سالبة او نتائج غير دالة ويحدد عدد الدراسات الدالة احصائياً و عدد الدراسات غير الدالة ويتم الوصول الى استنتاجات، وتتضمن:

- **مدخل العد غير المنظمة Informal vote-counting**: المدخل التقليدي وتتضمن الوصول الى الاستنتاج فى ضوء معيار قاعدة الاغلبية فاذا كانت التاثيرات الموجبة اكثر من التاثيرات السالبة والتاثيرات غير الدالة يصل الباحث الى وجود استنتاج وجود تاثير موجب و لكن معرفة عدد النتائج الدالة او غير الدالة لا يخبرنا عن شيئاً عن مقدار حجم العلاقة او التاثير.

- **مدخل العد المنظم Formal vote-counting**: يستخدم التحليل الاحصائى للتكرار المتوقع للنتائج بمعنى: هل نتائج عينة من الدراسات السابقة سواء نتائج دالة احصائياً او غير دالة احصائياً تتطابق مع النتائج فى مجتمع الدراسات السابقة؟. وذلك عند مستوى الدلالة الاحصائية التقليدى 0.05. بمعنى هل اكثر من 5% من الدراسات انتجت دلالة

احصائية؟. وعلى الرغم من ان هذا النوع من المراجعة مفيد عندما لا يكون حجم التأثير للدراسة غير متاح حسابه.

- **ما وراء التحليل:** والوصول الى استنتاج فى ضوء الدلالة الاحصائية وحده غير كافى ولذلك فان ماوراء التحليل لايركز على الدلالة فى حد ذاتها بل على حجم التأثير الذى يمدنا بمعلومات عن مقدار التأثير Magnitude of the effect. وما وراء التحليل اسلوب من اساليب عديدة لتلخيص وتكامل وتفسير مجموعة منتقاة من اعمال الباحثين فى مختلف التخصصات. ومن الخطأ ان نعتبر ماوراء التحليل اختبار احصائى بسيط بل هو فى حد ذاته مجال للدراسة او منهجية بحثية شاملة يتضمن خطوات منظمة واجراءت احصائية عديدة.

والمراجعة المنظمة لدراسات موضوع ما تتطلب جمع وتحليل كل البيانات او الدراسات المنشورة وليس فقط الدراسات المهمة او المتاحة فقط.

حقيقة من الصعب القول بوجود دراسة متقنة تماماً Perfect study فكل الدراسات تعاني من بعض المحددات المنهجية وتتضمن اخطأ القياس، فإى دراسة ليست تامة صدق البناء Construct validity بالاضافة لعوامل اخرى تشوة نواتج الدراسة مثل خطأ المعاينة وعليه لا يمكن الوصول الى استنتاجات علمية من دراسة مفردة وحيدة او مجموعة فرعية من الدراسات.

وماوراء التحليل يستخدم لدمج اشكال مختلفة من البيانات الكمية فهو مناسب للبحث عن قيم النزعة المركزية او توزيع قيم لمتغير وحيد، اتجاه وقوة العلاقة بين متغيرين، تأثيرات المعالجات، وهكذا. ولكن دراسات ماوراء التحليل الاكثر شيوعاً كانت فى مجال تأثير المعالجات الاجتماعية او النفسية او التربوية او الطبية وغيرها. وتمتاز ماوراء التحليل بانها طريقة لتوليف النتائج للوصول الى استنتاجات محددة .

وفيما يلى مقارنة بين طرق مراجعة الدراسات السابقة (Koricheva & gurevitch, 2013):

الخاصية	المراجعة السرديّة	عدّ الدلالة الاحصائية	ما وراء التحليل
وضع قيود على الدراسات	لا	لا	نعم
تقيم النتائج في ضوء الدلالة	نعم	نعم	لا
ياخذ في اعتباره حجم العينة والقوة الاحصائية	لا	لا	نعم
يقدر حجم التأثير العام للدراسات	لا	لا	نعم
يقدر الدلالة الاحصائية لمتوسط حجم التأثير	لا	لا	نعم
يحلل مصادر اختلاف احجام التأثير بين الدراسات	لا	لا	نعم

خطوات ما وراء التحليل او توليف البحوث Research Synthesis

لتوضيح اهمية عملية توليف البحوث التي تتضمن احد صورها ما وراء التحليل فمن الضروري اتباع الخطوات العلمية لعملية التوليف او الدمج وهي تتضمن خمسة مراحل مناظرة لتلك المتبعة في البحوث الاولية كالتالى (Card, 2012; Cooper, 1998, Cooper & Hedges, 2009; Little et al., 1998; Kline, 2013) :

- **صياغة مشكلة البحث** Problem formulation: المرحلة الاولى فى ماوراء التحليل هو تحديد المشكلة او القضية المراد دراستها، او ما يرغب الباحث فى الاجابة عنه وتحديد المتغيرات او الابنية والمجتمع المتضمن فى دراسته ما وراء التحليل المزمع انجازها. فاذا كان هدف الباحث الوصول الى استنتاج عن السببية Causality فالباحث يعتمد على الدراسات التجريبية وربما الدراسات الطولية ودراسات تحليل المسار، ويمكن ان يستخدم التراث تعريفات اجرائية لنفس المفهوم اوريا يستخدم الباحث مصطلحات متشابهة لوصف مفاهيم مختلفة ولذلك فمن الضرورى تحديد مصطلحات وعينة الدراسة (طلاب مرحلة تعليمية او عينة صعوبات تعلم وغيرها) جيداً قبل البدء فى بحث ومراجعة الادبيات البحثية وكل ذلك فى تحديد الدراسات المتجانسة المناسبة لاجراء دراسات ماوراء التحليل.

- **البحث عن عينة الدراسات المرتبطة بموضوع البحث (الدراسات الاولية) Searching literature او Sampling:** وفى هذا الشأن يبدأ الباحث فى الحصول عن الدراسات الاكثر ارتباطاً ومن المناسب الحصول على عينة من الدراسات ممثلة تمثيلاً جيداً للتراث البحثى للظاهرة موضع الدراسة من مصادر متعددة سواء مقالات، كتب، تقارير لهيات، وغيرها. ولو ان الدراسات غير ممثلة للتراث البحثى عندئذ تكون الاستنتاجات متحيزة لواقع الظاهرة واحد التهديدات جراء ذلك هو **تحيز النشر Publication bias** ويحدث هذا للدراسات التى تفشل فى الحصول على تأثيرات دالة احصائياً فهذه الدراسات اقل احتمالية للنشر فى الدوريات العلمية ومن الصعب الحصول عليها. وللتغلب على هذه الاشكالية يحاول الباحث الحصول على الدراسات غير المنشورة مثل دراسات الماجستير والدكتوراة ودراسات المؤتمرات وتقييم ماذا كانت توجد فروق فى التأثيرات بين الدراسات المنشورة والدراسات غير المنشورة . ويوجد تهديد اخر هو ان المراجعين للتراث يعتمدون على الدراسات التى تعتمد على اللغة التى يعرفونها ويستبعدون التراث المنشور فى بلاد اخرى وثقافات اخرى، وعلى الرغم انه من غير العملى ان يعرف الباحث كل اللغات فى مجتمعات اخرى ولكن على الباحث ان يكون مدركاً لمحددات ذلك على استنتاجات الدراسة ويجب ان يتميز بالشفافية Transparency ان يذكر الاساليب التى استخدمها فى البحث عن الادبيات البحثية. ويرى (Kline 2013) لا يوجد حد ادنى مطلق للدراسات المتضمنة فى دراسة ما وراء التحليل الا انه يبدو لى ان 20 دراسة مختلفة على الاقل هو العدد المعقول

لاجراء ما وراء التحليل والفشل فى الحصول على عدد كافى من الدراسات يمثل فجوة معرفية.

● **Evaluation Quality of الدراسات السابقة لتحديد مدى تضمينها**

Studies: وهذه المرحلة تتضمن قراءة التراث والوصول الى استنتاجات عن طبيعة العلاقات او التأثيرات بين المتغيرات. وتوضيح معايير الاستبعاد Exclusion او التضمين للدراسات التى تتضمن طبيعة الابنية او العينات غير المناسبة او المجتمع المناسب لطبيعة الدراسة التى بصدد اجرائها او التصميم البحثى او طبيعة النواتج، والمعلومات المتحصل عليها فى المرحلة تفيد فى اعادة صياغة المشكلة واستبعاد الدراسات التى لم يتوفر فيها معايير التضمين Inclusion.

● **التكويد وحساب حجم التأثير Coding and Compute Effect Size:** يوجد العديد من

المداخل لكيفية الوصول الى استنتاجات عن الدراسة تتراوح من المرجعة الكيفية الى ماوراء التحليل. وباستخدام ماوراء التحليل لابد من التكويد Coding البناء لخصائص او عوامل الدراسة وحساب احجام التأثير وكذلك تحديد وزن كل دراسة فى ضوء حجم العينة وتباين الخطأ المعيارى وكذلك يمكن تصحيح احجام التأثير من المحددات المنهجية مثل ضيق المدى او الثنائية التصنيفية الاصطناعية التى تسبب تقلص لاحجام التأثير.

● **تحليل البيانات Data analysis:** بعد حساب حجم التأثير تبدأ عملية التوليف الكمي

لاحجام التأثير المختلفة باستخدام اى من نماذج ما وراء التحليل سواء المثبتة او العشوائية او المختلطة واجراء التحليلات المتفاعلة للتعرف على مدى اسهام خصائص الدراسات فى تفسير الاختلاف بين احجام التأثير .

● **وصف وتفسير وعرض النتائج Describe, Interpret, and Report the**

results: يتم عرض النتائج والرسومات والجدول ثم مناقشتها وتفسيرها وعرض التوصيات والتطبيقات العملية او الدراسات المقترحة. وفى هذه المرحلة لابد ان يتميز الباحث بالشفافية ويبتعد عن التحيز الشخصى خاصة فى حالة استخدام المراجعة الكيفية.

وقدم Ellis (2010) خطوات ما وراء التحليل فى ستة خطوات مبسطة كالاتى:

1. جمع الدراسات.

2. تكويد الدراسات.

3. حساب متوسط حجم التأثير.

4. حساب الدلالة الاحصائية لمتوسط حجم التأثير.

5. فحص الاختلاف في توزيع تقديرات احجام التأثير.

6. تفسير النتائج.

ويوجد عدة ارشادات لابد ان تؤخذ في الاعتبار اهمها:

- بالشفافية عن عملية المراجعة والقرارات التي يتخذها الباحث، ولا بد ان يعرض تفاصيل كافية حتى يستطيع باحث اخر على اعادة نتائج المراجعة.
- من المفيد عرض التقارير البحثية التي اجابت عن الاسئلة الاصلية للدراسة التي بصدد اجرائها وما الذي لم تستطيع الدراسات السابقة الاجابة عنه.
- تجنب عرض حرفى جاف للدراسات السابقة بدون وجود درجة من الاستبصار لمغزى المعلومات الخفية وهذا يتطلب درجة من التعمق في مضمون هذه الدراسات.
- وهذا يتحقق من خلال استخدام ما وراء التحليل للوصول الى معلومات قيمة من الدراسات المتعددة والوصول الى ما هو ابعد من محتوى الدراسات المفردة .

يرى (Little et al. 2008) ان ما وراء التحليل يتضمن مهمة اوالكثير من المهام الاتية:

- تقدير متوسط التأثير العام.
- تقدير التباينات او الاختلافات Heterogeneity في احجام التأثير عبر الدراسات.
- تقدير التأثيرات لمجموعات فرعية من المجتمع موضع اهتمام الدراسة.
- اكتشاف التأثيرات المحتملة للمتفاعلات (خصائص الدراسات) على اختلاف احجام التأثير.
- تقدير اثر تحيز النشر وتحيز العينات الصغيرة على اداء ما وراء التحليل.

وتوجد بعض المفاهيم الخاطئة الخاصة بما وراء التحليل منها:

السائد	الحقيقة
نشأ في ابحاث المجال الطبي	نشأ في العلوم الاجتماعية والسلوكية
مناسب للدراسات التجريبية	مناسب لمعظم طرائق البحث مثل البحث الارتباطي والسببي المقارن.
يجب ان يتضمن دراسات	يمكن ان يتضمن دراسات من من تصميمات تجريبية تتضمن مجموعة ضابطة.
من تصميمات تجريبية تتضمن مجموعة ضابطة.	تصميمات عينات غير عشوائية(منهج شبه تجريبي) مثل تصميم قياس قبلي - بعدى.
يتطلب دراسات كثيرة.	يمكن استخدامه لدراستين او ثلاثة فقط.
يتطلب دراسات ذات عينات كبيرة.	يمكن استخدامه لدراسات ذات احجام عينات صغيرة وتستخدم اختبارات لتصحيح من تحيز العينة الصغيرة.
يتغلب على مشاكل الجودة(الصدق) للدراسات الاصلية	ماوراء التحليل لا يحسن من جودة الدراسات الاصلية.

مميزات ماوراء التحليل

يمتلك ماوراء التحليل مميزات ومصادر قوة عديدة كالاتى (Borenstein et al., 2009; leandro, 2005; Littlel et al., 2008; Lisepy & Wilson, 2002)

- **الموضوعية:** ماوراء التحليل يتغلب على محددات التلخيص الكيفي التقليدي للدراسات السابقة ويمدنا بطريقة فعالة لتلخيص نتائج كم هائل من الدراسات و الكشف عن علاقات لا يمكن ادراكها من الدراسات المفردة ويمدنا بشفافية غير متاحة في المراجعة السرديّة التقليديّة Narrative review لنتائج الدراسات وهذا بدوره يقلل التحيزات

والاجتهادات الشخصية. ويعتبر ماوراء التحليل طريقة منظمة وموضوعية لمعالجة معلومات من عدد كبير من نواتج الدراسات.

- **التغلب على محددات اختبارات الدلالة الاحصائية:** احد طرائق مراجعة التراث طريقة العد Vote-counting المعتمدة على عد النتائج الدالة والنتائج غير الدالة وهى تعانى من محدد اساسى وهو الاعتماد على اختبارات الدلالة الاحصائية وهى تتاثر بحجم العينة فدراسات العينات الكبيرة تعطى دلالة للعلاقات او التأثيرات الضعيفة، فالدلالة الحصائية لاتخبرنا عن قوة او حجم العلقه او التأثير ولذلك فان طريقة العد ربما تخبرنا عن دلالة احصائية موجبة او سالبة ولكنها حقيقة ليس لها دلالة عملية. كما ان دراسات العينات الصغيرة غالباً تعانى من قوة احصائية منخفضة للكشف عن التأثيرات الدالة احصائياً بمعنى تعطى نتائج غير دالة احصائياً بينما هى فى طبيعتها دالة عملياً او اكلنيكياً. وباستخدام ماوراء التحليل يمدنا بتقديرات عامة اكثر دقة نتيجة دمج الدراسات ذات احجام العينات الصغيرة والكبيرة فى دراسة واحدة وهذا من شأنه زيادة القوة الاحصائية للكشف عن التأثيرات الدالة. وبذلك يتغلب اسلوب ماوراء التحليل على محددات اختبارات الدلالة الاحصائية.

وعلى ذلك فان ماراء التحليل تميز عن اختبارات الدلالة الاحصائية بانه يتعامل مع مؤشر حجم التأثير من كل دراسة وليس الدلالة الاحصائية او قيمة P لتحديد ما اذا كان حجم التأثير متسق عبر الدراسات وكذلك تضمن كل احجام التأثير لكل الدراسات فى توليف احصائى واحد وهذا فى غاية الاهمية لحساب حجم تأثير تلخيصى، على ذلك فان ماوراء التحليل يذهب ابعد من لدلالة الاحصائية ويسمح ماوراء التحليل بتقدير مدى الاختلاف بين التأثيرات ويميز بين الاختلاف الحقيقى والاختلاف الزائف.

- **وراء التحليل اجراء مفيد فى بناء معرفة علمية صادقة** قائم على عملية منظمة لتلخيص النواتج البحثية تشمل معايير لتحديد مجتمع نتائج الدراسات وكذلك استراتيجيات بحثية منظمة لتحديد واسترجاع الدراسات وتكويد خصائص الدراسة والنواتج وتحليل البيانات للوصول الى استنتاجات صادقة.

○ قدرة ماوراء التحليل للوصول الى علاقات وتأثيرات لايمكن الوصول لها من الطرق المختلفة لمراجعة التراث. فالطرق الكيفية او السردية لنتائج الدراسات لاتخبرنا عن الفروق بين الدراسات. ومن خلال التأكيد المنظم (احد مراحل ماوراء التحليل) لخصائص الدراسات يمكن المقارنة بين الدراسات من حيث طبيعة المعالجة وتصميم البحث واجراءات القياس وبتقدير حجم التأثير لكل دراسة ودمجها معاً عبر الدراسات المختلفة يعطى تقديرات لحجم التأثير اكثر قوة احصائياً من الدراسات المفردة.

○ الحصول على تقديرات معالم افضل **Parameter Estimates**: نتجة لخطأ المعاينة يمكن لاي دراسة ان تنتج تقدير غير دقيق للتاثير خاصة فى المجتمعات الكبيرة ، فالمقاييس متعددة المفردات اكثر ثباتاً من المقياس ذو المفردة الواحدة، وبالمثل فتوليف النتائج من دراسات عديدة يعطى تقديرات اكثر دقة ومناعة لمعالم المجتمع من الدراسة المفردة .

○ اجراء التحليل المتفاعل **Moderator analysis** : المراجعة السردية لايمكن ان تكشف عن مدى تاثير نوعية المشاركون او المعالجة او تصميم الدراسة على نواتج الدراسة، ويستطيع ما وراء التحليل الكشف عن تاثير هذه العوامل او الخصائص المسئولة عن الاختلاف بين النتائج او احجام التاثير مثل خصائص العينة(عمرها، جنسها)، نوعية المعالجة، مكان اجراء الدراسة، ونوعية القياسات(ملاحظة، تقرير ذاتي)، جودة تصميم البحث، مصدر تمويل الدراسة(خاص، حكومي)، تاريخ النشر، الخلفية المهنية للباحث (Kline, 2013). ولو ان التحليل المتفاعل كشف عن عدم تاثير هذه العوامل على التاثير العام عندئذ فان نواتج ماوراء التحليل تشير الى ان النتائج لديها مناعة او صمود وقابلية لتعميمها الى مواقف اخرى، لو ان الدراسات غير متسقة بين المؤسسات او المواقف فلايد من فحص وكشف مصادر الاختلاف. وتحليل التفاعل يمكن تحديد البرامج اوالمعالجات اكثر نجاحاً. ولذلك فان تحليل المتغير المتفاعل يهدف الى تقدير ما اذا خصائص الدراسة تفسر الاختلاف بين احجام التاثير.

○ تقليل الخطأ والتحيز **Minimizing error and bias** : يمكن لما وراء التحليل تقليل خطأ المعاينة والتحيز من خلال توليف عدد كبير من الدراسات لموضوع ما.

ويرى (Littlel et al., 2008) ان من اهم مميزات اسلوب ماوراء التحليل هو التغلب على محددات المراجعة الكيفية السردية للدراسات، وكذلك يمدنا بطريقة فعالة لتلخيص نتائج كم هائل من الدراسات ويكشف عن علاقات غير مدركة فى الدراسات المفردة وتمدنا بشفافية فى مراجعة الدراسات كذلك يعطى معلومات اكثر دقة عن النتائج المختلطة عن طريق التلخيص الكمي للتاثيرات.

وكذلك يرى (leandro 2005) ان ماوراء التحليل يسهم فى التالى:

- زيادة القوة الاحصائية للتحليلات او المقارنات.
- يعطى تقدير اكثر دقة لتاثيرات المعالجات.
- دمج او توليف او تقييم نتائج الدراسات المتعارضة.
- الاجابة عن اسئلة بحثية جديدة.
- تحليل مجموعات فرعية من لعينات مختارة من دراسات مختلفة.
- تحليل الاتجاهات العامة فى دراسة موضوع ما.
- لفت الانتباه الى المجالات البحثية التى بحاجة الى دراسة وبحث.

الانتقادات الموجهة لاسلوب ما وراء التحليل

منذ نشأة اسلوب ما وراء التحليل على يد (Smith & Glass (1977) واخرون يلاقى العديد من الانتقادات او المحددات معظمها مرتبط بالتطبيقات غير المنطقية وسوء استخدام الاسلوب (Littell et al., 2008). ومن اهم الانتقادات الموجهه لهذا الاسلوب (Borenstien et al., 2009; Card, 2012; Littell et al., 2008; Lisepy & Wilson, 2002; Rosenthal & DiMatteo, 2001)، كالتالى:

- يحتاج الى جهد وخبرة عالية لفهمه وتنفيذه من جمع الدراسات وتكويدها تقدير احجام التاثير المناسبة وتحليلها احصائياً وتفسيرها وذلك مقارنة بالمراجعة الكيفية السردية، وعليه يتطلب جهد احصائى عالى لتنفيذه والالمام بفنياته ولكن كاي اسلوب يبدو صعب

عندما تقرأه اول مرة، ولكن معظم تطبيقات ما وراء التحليل لا تتطلب تحليلات معقدة، وايضاً كتابة تقرير دراسة ما وراء التحليل تحتاج الى مهارة عالية مقارنة بتقرير الدراسات الاولية

- بدون اتباع القواعد المتبعة لاسلوب ما وراء التحليل من جمع الدراسات المرتبطة وتكويدها وتحليلها تصبح عبارة عن **تدريب احصائي** لامعنى له.
- **التحديد غير المناسب للمشكلة:** عند اجراء ما وراء التحليل لابد تحديد نوعية الدراسات او المعالجات المتضمنة فى التحليل وغياب المعنى المنطقى من تضمين الدراسات المختلفة التى تتضمن معالجات وعينات وتصميمات بحثية ونواتج مختلفة يعطى نتائج من الصعب تفسيرها ويعتبر هذا من الانتقادات الكلاسيكية لما وراء التحليل.
- يشبه البعض خلط دراسات مختلفة التصميمات والعينات كالخلط بين **التفاح والبرتقال** حيث تتباين الدراسات فى الجودة المنهجية والمقاييس والعينات وغيرها من المظاهر المنهجية، ولكن محلل او مستخدم الاسلوب يعتبر ان هذا الخلطة مناسبة عندما يكون البعض مهتم بعمل سلطة فواكهة فمن غير المناسب ان يكون الباحث مهتم بالتفاح فقط او يريد ان يميز التفاح عن البرتقال. ولابد من وجود منطقية نظرية قوية لاجراء ما وراء التحليل لاعطاء نتائج ذات معنى ويمكن التغلب على ذلك بدراسة العوامل المسؤولة عن تفسير التباين بين الدراسات المختلفة.
- **تقدير غير مناسب لجودة الدراسة:** ما وراء التحليل لدراسات منخفضة الجودة لها ثبات منخفض وتصميم بحثى ضعيف ومشاكل اجرائية ينتج عنها نتائج غير صادقة. ويطلق على هذه المشكلة " زبالة دخلت، زبالة خرجت". وللتغلب على هذه المشكلة لابد من وضع معايير لانتقاء الدراسات فى ضوء تصميم البحث الجيد ونوعية المعالجات التى تتناول المتغيرات المستقلة لتضمين الدراسات فى التحليل ويسعى ما وراء التحليل ما اذا كان التنوع المنهجى يؤثر على النتائج من خلال تحليل المتفاعلات.
- **انخفاض الصدق الخارجى External validity:** كأى نوع من الدراسات فان نتائج ما وراء التحليل يمكن تعميمها فقط الى المجتمعات والمعالجات والمواقف التى تشبه المواقف التى تمت فى التجربة، فالمعالجات التجريبية التى تتم فى ضوابط منهجية صارمة وعينات متجانسة والتركيز على متغير مستقل وحيد فى الغالب وتدريب على

الجودة للمعالج او المتدرب، فالنتائج المتحصل عليها من هذه النوعية من الدراسات من الصعب تعميمها فى المواقف الطبيعية سواء فى المدرسة او الجامعة او المستشفى او غيرها. وعليه فنتائج ما وراء التحليل لهذه الدراسات تعاني من نفس المحدد وهو انخفاض القوة التعميمية فى المواقف الطبيعية.

• **اعتماده على ابحام التأثير Effect sizes:** لدمج الدراسات كمياً فمن الضرورى الحصول على قياسات تكون قابلة للمقارنة عبر الدراسات، ولذلك يحصل الباحث على مجموعة من ابحام التأثير لتحديد قوة اوحجم العلاقات بين المتغيرات. وفى ما وراء التحليل يحسب حجم التأثير لكل دراسة على حدة ثم توزن بحجم العينة ومتوسطها لتمدنا بالتاثير العام Overall effect . وفى هذا الشأن ينتقد البعض ما وراء التحليل باعتماده على حجم التأثير فقط وعلى الرغم انه يمدنا بمؤشر محكم لمتوسط التأثير عبر الدراسات ولكنه ربما يكون غير مفهوم ومتداول لمعظم الافراد. ولكن هذا المحدد يمكن التغلب عليه بوضع قياسات لهذا المؤشر لى تكون مفهومة لدى الاكاديميين وصانعى القرار. ولكن تتوافر بعض الفنيات للتحقق من مسلمة التجانس Homogeneity لتحديد ما اذا كانت ابحام التأثير من الدراسات المختلفة متجانسة من عدمه.

• **بعض فنياته الجديدة لم تصل الى درجة مصداقية عالية :** على سبيل المثال الفنيات المستخدمة لتصحيح تحيز النشر مازالت محل جدل ونقاش بين المتخصصين.

• **يتجاهل المعلومات الكيفية البناءة:** على الرغم ان ما وراء التحليل لا يمكن استخدام لتوليف البيانات الكيفية ولكن المعلومات الكيفية على درجة كبيرة من الاهمية حيث من الممكن اقتراح قضايا ومعلومات ذات مغزى واضاءات اثناء التجريب. ولكن احد المداخل لما وراء التحليل توليف الادلة الافضل Best evidence synthesis الذى يحاول ان يدمج بين فنيات المراجعة الكيفية والكمية. فالبحت الكيفى على درجة كبيرة من الاهمية حيث يمدنا برؤى وبمعلومات ثرية تسهم فى فهم الظاهرة مقارنة بالبحث الكمى الامبريقى. والتقدم العلمى يكون محدداً لو اعتمد فقط على البيانات الكمية المتحصل عليها من طرائق البحث الكمية .

• **المعاينة غير المناسبة وطرق جمع البيانات:** لا يعطى ما وراء التحليل انتباه كافى لمصادر الخطأ المحتمل والتحيز فى معاينة (عينة الدراسات) ومعايير تقييم الدراسات

الدمجة، وعدم الانتباه لثبات وصدق البيانات المستخلصة للتحليل من الدراسات. ويعتمد تحليلات ما وراء التحليل في معظمها على الدراسات المنشورة على الرغم من عدم وجود فرق بين نتائج التحليل بين الدراسات المنشورة والدراسات غير المنشورة De Smidt & (Gorey, 1997).

- **تحيز تقرير النتائج Outcome reporting bias** : دائماً يتم نشر النتائج الدالة احصائياً مقارنة بالدراسات ذات النتائج غير الدالة احصائياً (الصفيرية) وهذا يعرف بتحيز تقرير النتائج عليه فان الباحث يعتمد على حساب حجم التأثير للنتائج الدالة احصائياً التي تكون متاحة للباحث وهذا بدوره يؤدي الى استنتاجات مريكة ومتحيزة لاجسام التأثير، وهذا مرتبط بتحيز النشر حيث الدراسات الدالة نتائجها احصائياً اكثر مقبولة للنشر في الدوريات العلمية مقارنة بنظيرتها غير الدالة احصائياً بالتالي لا نستطيع الوصول الى استنتاجات صادقة لانه يعتمد على الدراسات المنشورة الدالة احصائياً.
- **التحيز في انتقاء الدراسات التي تعتمد على لغة الباحث** بينما الدراسات في البلدان الاخرى بلغات اخرى تكون غير متاحة للباحث وكذلك مدى الاتاحة Availability حيث بعض الدراسات تحتاج الى تكلفة عالية للحصول عليها وهذا غير متاح للباحث. فالباحث العربي لديه اتاحة ومرغوبية للحصول على الدراسات المكتوبة باللغة الانجليزية لانه اكثر الفة بها مقارنة بالدراسات الفرنسية والالمانية والصينية وغيرها، وهذا بدوره يقود الى استنتاجات خاطئة عن النظرية او الظاهرة موضع الدراسة واذا اعتمد التحليل على مجموعة متحيزة من الدراسات الدالة احصائياً ويتجاهل النتائج غير الدالة احصائياً بالتالي يعطى نتائج متحيزة، ولكن المؤيدون يرون ان الاسلوب يتضمن العديد من الفنيات لتشخيص هذه المشكلة وهي تحيز الدراسات ومحاولة معالجتها. وهذه المشكلة ليست قاصرة فقط على ما وراء التحليل انما ايضاً قائمة في المراجعات الكيفية للدراسات السابقة.
- **مصدقية استخلاص البيانات Data extraction**: كثيراً من الباحثين متاكدين من قدراتهم العالية لاستخلاص وتقييم المعلومات من الدراسات ولكن بدون وجود اكثر من مقدر او مكود وحساب ثبات او اتفاق المقدرين المستقلين Inter-rater agreement فان القارئ لا يستطيع التأكد من مدى مصداقية البيانات المستخلصة من الدراسات.

ووجد (2007) Gotzsche, Hrogjartsson, Maric, & Trendal الى وجود اخطأ في استخلاص بيانات 10 دراسات (37%) من 27 دراسة لما وراء التحليل، وهذه بدوره لايعطى نتائج دقيقة بدون تجنب اخطأ ادخال البيانات .

- يلخص المعالجات او تاثيرات الدراسات بمعنى يحاول ان يقلل النتائج في قيمة واحدة ويتجاهل التنوع بين الدراسات حيث اذا حدث خلل او عيب في استخلاص هذه القيمة فله عواقب خطيرة على صناعة القرار. ولكن المؤيدون يرون ان الهدف من الاسلوب هو توليف احجام التأثير وليس اعطاء ملخص لها ولو كانت هذه التأثيرات متجانسة فان التحليل يظهر تاثير لديه مناعة عبر الدراسات المتضمنة في التحليل ولكن لو ان الاختلافات جوهرية بين احجام التأثير فان الاهتمام يتحول من ملخص لحجم التأثير الى تباين احجام التأثير.
- تتعامل ما وراء التحليل مع التأثيرات الرئيسية للمتغيرات ولا تستطيع ان تتعامل مع التفاعلات بين المتغيرات في تأثيرها على المتغير التابع انما يمكن ان يتم ذلك من خلال التحليلات المتفاعلة.

الفصل الثانى

تحديد المشكلة او موضوع البحث

Topic or problem specification

مقدمة

مثل اى طريق او منهجية من طرق البحث فان الخطوة الاولى لاجراء ما وراء التحليل تبدأ بتحديد الموضوع المراد دراسته او السؤال المطلوب الاجابة عليه وهذا ضرورى لتحديد اهداف دراسة ما وراء التحليل وهذا التحديد للموضوع يعتبر كمرشد لاختيار الدراسات السابقة وتكويد المعلومات من هذه الدراسات وتحليل نتائج البيانات، ويرى (Littell et al. 2008) انه لاجراء دراسة ما وراء التحليل لابد من وجود فريق عمل لان الاسلوب يتضمن فنيات ومهارات لتنفيذة ومن اهم مهامه اتفاق المقدرين لتقليل الخطأ والتحيز فى تحديد الدراسات وتكويدها.

وتوجد اسباب لاجراء ما وراء التحليل منها:

- تقويم اتجاة ومقدار متوسط حجم التأثير.
- الكشف عن مصادر الاختلاف بين الدراسات.
- التحقق من فروض بحثية.
- مقارنة نتائج دراسات العينات الصغيرة مع العينات الكبيرة.
- اختبار فعالية برامج علاجية او تدريبيه او ارشادية وغيرها.
- تحديد الفجوات البحثية.

وحدد (Cooper 1988) اهداف مراجعة الدراسات السابقة وهى التكامل Integration المتضمن الوصول الى استنتاجات، حل التناقضات بين نتائج الدراسات، تحديد العلاقات بين النظريات، نقد الدراسات، تحديد القضايا الرئيسية. وتفيد المراجعة فى تحديد الاتجاه السائد فى الدراسات وطرائق البحث المستخدمة.

وتحديد المشكلة يتطلب تعريف وتحديد المتغيرات وطبيعة العلاقة بينهما وكيفية اختبارها امبريقياً، والمتغيرات فى العلوم الاجتماعية يجب تحديدها بطريقتين الاولى التعريف المفاهيمى او النظرى

وهي تحدد صفات المتغير وخصائصه، فمثلاً التعريف النظرى للذكاء هو القدرة على التعلم، الاستدلال، الفهم، وای نشاط عقلى اخر، فالتعريف النظرى شامل وعريض بينما التعريف الاجرائى اكثر تحديداً ويكون قابل للقياس ووصف لخصائص للاحداث الملاحظة لتحديد مدى وجود الظاهرة او المفهوم

ويرى (Lipsey & Wilson (2001 يجب ان تكون صياغة المشكلة واضحة ومباشرة واكثر تفصيلاً مثل:

هل البرامج القائمة على الوسائط المتعددة فعالة فى تحسين التحصيل؟.

هل البرامج القائمة على استراتيجيات التعلم المنظم ذاتياً لها تأثيرات فعالة فى تنمية تقدير الذات او اى نواتج اخرى مثل التحصيل او الاتجاه او التفكير الناقد وغيرها من النواتج الانفعالية والمعرفية والسلوكية؟.

لاحظ ان صياغة المشكلة وتحديدها يساعد فى تحديد طبيعة المتغيرات المستقلة والتابعة، فالمتغير المستقل موضع الاهتمام والدراسة هو خصائص البرنامج بينما المتغيرات التابعة تقدير الذات، التحصيل، الاتجاه، التفكير الناقد. وهذا التحديد يعتبر بمثابة المرشد لتحديد الكلمات المفاتيحية وبدورها المساعدة فى البحث عن التراث البحثى والدراسات السابقة.

ودائماً تصاغ عناوين ما وراء التحليل لدراسة تأثيرات المعالجات كالاتى:

اثر (المعالجة او البرنامج) لـ (المشكلة او المتغير الناتج) فى (مجتمع)

مثال: العلاج المعرفى السلوكى لتخفيف اضطرابات القلق للاطفال والبالغين

اثر (البرنامج او التدخل) مقارنة بـ (الشرط او المجموعة الضابطة) لـ (المشكلة او الناتج)

فى (مجتمع)

مثال: اثر العلاج السلوكى مقارنة بالعلاج المعرفى فى تخفيف اعراض الاكتئاب لدى نزلاء السجون.

تحديد اهداف واسئلة او موضوع البحث المراد الاجابة فى ما وراء التحليل

وتتضمن عدة خطوات اهمها:

الموضوع والمجال Topic and scope

يرى (Littell et al. (2008 ان الخطوة الاولى لاجراء توليف البحوث هى تحديد الاسئلة الجوهرية التى سوف تكون بمثابة مرشد للباحث لجمع التراث البحثى ومن الضرورى تحديد نطاق ومجال و المسلمات الاساسية لاجراء دراسات ما وراء التحليل.

والاسئلة والفروض الجيدة تأتى من مصادر عديدة منها الممارسة البحثية، الخبرة، والمراجعة المستمرة لتراث ظاهرة ما. والسؤال البحثى يجب ان يكون قابل للبحث، محدد بوضوح، ممتع، جديد، اخلاقى، ومناسب. فمثلاً هل البرنامج فعال؟، فهذا ليس سؤال محدد لان التجربة او البرنامج ليس فعالاً لكل الفئات بمعنى له شروط وظروف معينة، ولكن الصيغة الافضل للاسئلة عن تاثير برنامج او تدخل ما هو: هل المعالجة X1 اكثر فعالية من المعالجة X2 على المتغير التابع Y لمجتمع ما؟. وهذا من شأنه تحديد ما هى المعلومات التى يحاول الباحث الحصول عليها ومنها تحديد المجتمع والمعالجات والمواقف والنواتج، كذلك تحديد المكونات المستبعدة من الدراسات السابقة.

فمثلاً احد المراجعات يهدف الى تقدير تاثير البرامج المدرسية على مدى واسع من النواتج مثل كفاءة الذات، العلاقات بين الاقران، التحصيل الاكاديمى، وهكذا. ومراجعة ما وراء تحليل اخرى تركز على تاثيرات هذه البرامج المدرسية لدى التلاميذ منخفضى الدخل، ومراجعة ثالثة تركز على تاثير البرامج المدرسية على مدى محدود من النواتج او على ناتج واحد مثل التحصيل الاكاديمى. بالتالى هذه الاسئلة تعكس اطر مختلفة للبحث فى التراث.

وكلما كان الباحث اكثر وضوحاً فى تحديد الموضوع المراد دراسته فى ما وراء التحليل كلما كانت الدراسات المراجعة من التراث البحثى اكثر تجانساً ولكن عدم التحديد الدقيق للموضوع المراد دراسته يؤدى الى اتساع دائرة البحث فى التراث ويحصل على دراسات غير متجانسة، وعليه تظهر مشكلة التفاح والبرتقال فى ما وراء التحليل ولكن البعض يرى ان الهدف العام هو البحث فى سلة الفواكهة ومحاولة ايجاد العلاقات بينهم.

ومن الافضل ان تكون القضية المراد بحثها محددة المفاهيم والاسباب والعواقب وان تكون مستندة لنظرية او نموذج لان هذا يساعد الباحث في صياغة العلاقات المفترضة بين المعالجات والنواتج على المدى القصير وعلى المدى الطويل.

وما وراء التحليل يتبع احد المدخلين لتكامل الدراسات (Card, 2012):

- **دمج للدراسات Combination** : فدمج الدراسات يتضمن استخدام احجام التأثير من الدراسات الاولية لتقدير حجم التأثير الاجمالي او لتقدير مدى احجام التأثيرات وذلك لعمل استدلال حول متوسط حجم التأثير في ضوء اختبارات الدلالة او فترات الثقة، وعلى ذلك فالهدف من ما وراء التحليل الدمج تحديد متوسط احجام التأثيرات. فعلى سبيل المثال باحث يرغب في دمج العلاقات بين تقدير الذات X والتحصيل Y من دراسات ارتباطية لتحديد افضل تقدير لقوة العلاقة بين المتغيرين.

وباحث اخر يرغب في دمج دراسات تناولت اثر معالجة ما على متغير تابع ما متصل عليها من تصميمات تجريبية عالية الصدق الداخلى للوصول الى استنتاج لتحديد مدى قوة المعالجة في حدوث تحسين في المتغير التابع.

- **مقارنة الدراسات Comparison**: وهذا المدخل يتطلب وجود اختلاف او عدم تجانس **Heterogeneity** لاحجام التأثير عبر الدراسات وعلى ذلك فان الهدف من المقارنة هو تقييم ما اذا كانت احجام التأثير تختلف اعتماداً على خصائص الدراسات المكودة وذلك من خلال تحليلات ما وراء التحليل التفاعلية *Meta-analytic moderator analyses*. وعلى ذلك فالهدف من وراء التحليل المقارنة تقييم العلاقات بين احجام التأثير وخصائص الدراسات.

وفى ضوء ما وراء تحليل المقارنة، اراد باحث تقييم الى اى درجة المتغير المستقل X منبئ بالمتغير التابع Y فى الدراسات الطولية فى فترات زمنية مختلفة و متعاقبة لتقويم مدى قوة التنبؤ عبر المراحل الزمنية المختلفة.

او يريد باحث المقارنة بين مدى فعالية طريقة علاج معينة على تخفيف الاكتئاب لدى الذكور والاناث. او يرغب باحث فى مقارنة ثبات مقياس معين لعينات مختلفة لتحديد مدى جدوى او

فائدة المقياس عبر مجتمعات متنوعة، هذه امثلة لنوعية قليلة من الاسئلة البحثية التي يتناولها اسلوب ما وراء التحليل.

معايير الصلاحية او تضمين واستبعاد الدراسات Eligibility or Inclusion and Exclusion Criteria

بعد تحديد الموضوع المراد دراسته باستخدام ما وراء التحليل واتضحت الفكرة عن طبيعة النواتج البحثية المناسبة للموضوع، يبدأ الباحث فى تحديد الدراسات البحثية المراد دمجها فى ما وراء التحليل ومن اهم مظاهر هذا التحديد هو مجتمع الدراسات البحثية المناسبة. ومن الضرورى وضع معايير Criteria لتحديد الدراسات المتضمنة او المستبعدة من التحليل، وهذه المعايير فى ضوء النظرية او النموذج القائم عليه موضوع البحث التى سبق تحديدها فى الخطوة السابقة وهى مجموعة من المحكات عن ملامح او خصائص الدراسات المراد تضمينها او استبعادها، ويجب تحديدها قبل البدء بالبحث فى التراث البحثى للموضوع.

وتحديد هذه المعايير يفيد فى الاتى (Card, 2012; Littell et al., 2008):

- تمدنا بمرشد موضوعى وصادق للدراسات المراد تضمينها او استبعادها وبدورها تقلل التحيز الشخصى Subjectivity الذى يحدث لو لم تتوفر هذه المعايير، وكذلك الابتعاد عن تحيز الانتقاء للدراسات كلما امكن. واذا لم تحدد هذه المعايير مسبقاً فان تضمين الدراسات يعتمد على التفضيل الشخصى والاتاحية للدراسات وعوامل اخرى وهذا بدوره يبتعد عن الموضوعية.

- تحدد طبيعة المجتمع المراد دراسته، فاذا كان الموضوع منصب على مرحلة التعليم الثانوى فيتم استبعاد كل الدراسات دون ذلك.

- الالتزام بالشفافية Transparency التى تعتبر من اهم خصائص دراسات ما وراء التحليل مما يساعد باحثين اخرين على اعادة الدراسة مرة اخرى Replication ، فافتراض انك حصلت على 100 دراسة لاجراء ما وراء التحليل ولكن بعد تطبيق معايير التضمين والاستبعاد قررت استبقاء 60 دراسة ولو استخدم باحث اخر هذه المعايير سوف يصل الى نفس الـ 60 دراسة.

ويحدد (Lipsey & Wilson 2001) معايير صلاحية الدراسات او معايير استبعاد او تضمين الدراسات فى ما وراء التحليل بالاتى: الملامح المميزة لاهلية الدراسة، المستجيبين فى الدراسة، المتغيرات الاساسية، تصميمات البحث، المدى اللغوى و الثقافى، زمن الدراسة، نوع النشر.

ويحددها (Littell et al. 2008) بالاتى: التصميمات البحثية، المجتمعات، التدخلات او البرامج، المقارنات، والنواتج المقاسة المراد تضمينها او استبعادها.

بينما يحددها (Card 2012) بالاتى: تعريفات المفاهيم موضع اهتمام الدراسة، خصائص العينة، تصميم الدراسة، وقت انجاز الدراسة، نوع النشر، ومعلومات عن حجم التأثير.

وفبما يلى عرض لاهم هذه المعايير:

• **الملامح المميزة او التدخلات Distinguishing Features or Interventions:** ماهى الملامح المميزة للموضوع المراد تحليله؟، فلو كان سؤال ما وراء التحليل: ما فعالية برنامج او تدخل ما؟، فان الملامح الاساسى طبيعة التدخل او البرنامج وعليه لابد من تحديد خصائص البرنامج او المعالجة المستخدمة وتعريفها والاسس النظرية التى تقوم عليها والفنيات والانشطة المستخدمة ومؤهلات القائمين على تطبيقه ومدته، وبعض دراسات ما وراء تستبعد البرامج التى تقل جلساتها عن 8 او 12 (Littell et al., 2008). ولو تعامل ما وراء التحليل مع مقارنة مجموعات (مثل الفروق بين الجنسين) فان المعيار تحديد نوعية المجموعات، لو ان موضوع ما وراء التحليل العلاقة بين بنائين (Y, X) فان المعيار تحديد طبيعة الابنية وتعريفها اجرائياً، وعليه لابد من تحديد التعريفات الاجرائية Operational definition للمفاهيم المتضمنة فى التحليل. ويرى بعض الباحثين والخبراء انه يمكن اجراء ما وراء التحليل بدون التعمق فى هذا المعيار ولكن له مخاطر لان التراث البحثى خاصة فى العلوم النفسية والاجتماعية يتضمن مفاهيم لها مسميات مختلفة ولكن لها نفس المعنى او العكس مثل مفهوم الابتكار والابداع او التأجيل والتسويق والتلكؤ الاكاديمى مسميات مختلفة لنفس المعنى.

• **خصائص العينة او المشاركين Sample characteristics or participants:** تعتبر خصائص المشاركين او المستجيبين المشتق منهم البيانات فى الدراسة من اهم معايير

تضمنين او استبعاد الدراسة من التحليل، وفي الدراسات الاولية المشاركين هم الافراد. فاذا كان موضوع ما وراء منصب على مجتمع الاعاقة السمعية فلا بد ان تكون الدراسات تتضمن عينات الاعاقة السمعية، وكذلك اذا كانت موضوع ما وراء التحليل فعالية برنامج لمجتمع المرحلة الجامعية فلا بد ان تتضمن الدراسات عينة طلاب المرحلة الجامعية، ويمكن تضمين بعض الخصائص الديمجرافية مثل الجنس او الموقع او الحالة الاقتصادية الاجتماعية. فمثلاً دراسة ما وراء التحليل للذكور فقط ام للاناث فقط. ويمكن ان تكون دراسة ما وراء التحليل قاصرة على بيئة معينة مثل البيئة المصرية او السعودية او غيرها وعليه لابد ان تقتصر الباحث على الدراسات فى البيئة المصرية او السعودية او اى بيئة اخرى.

● **المتغيرات المفاتيحية او النواتج Key Variables or Outcomes** : ماهى المتغيرات الاساسية المتضمنة فى الموضوع او المشكلة المراد دراستها؟، فى الدراسات التجريبية تكون النواتج (المتغيرات التابعة) محور الاهتمام، وفى دراسات مقارنة المجموعات تكون المتغيرات المراد اجراء المقارنات عليها محور اهتمام ما وراء التحليل، فى الدراسات الارتباطية يكون المتغيرين Y, X محك اساسى لانتقاء الدراسات. ويوجد اعتبار هام لتضمنين قياسات النواتج وهو مصداقيها وثباتها. وايضاً توقيت قياس النواتج فالاعتماد على القياسات البعدية لا يخبرنا عن مدى حدث استمرارية للتحسن ولذلك يهتم بالقياسات البعدية Follow up measures.

● **تصميم البحث Research Method** : ماهى الملامح المنهجية والتصميمات البحثية التى تؤهل الدراسة لان تكون فى ما وراء التحليل؟، المحك الاساسى الذى يحدد منهج او تصميم البحث هو شكل النواتج البحثية. فى دراسات التدخل لابد من استخدام دراسات تجريبية يتم توزيع افراد العينة عشوائياً على المجموعة التجريبية والمجموعة الضابطة وان لم يكن توزيع عشوائى يصبح منهج شبه تجريبى وكذلك يعتمد المحلل على الدراسات التجريبية ذات تصميم المجموعتين وقياس قبلى بعدى ويستبعد اى دراسات اعتمدت على تصميم المجموعة الواحدة ذات القياس القبلى والبعدى ويمكن ان يستبعد الدراسات التى لم يتم التوزيع فيها عشوائى على المجموعة التجريبية والمجموعة الضابطة او يمكن الاعتماد عليها لو استخدمت قياسات قبلية للمتغير التابع او المتغيرات الاكثر ارتباطاً به. ويمكن الاعتماد على الدراسات ذات الجودة المنهجية العالية (توزيع عشوائى للمجموعات) او الدراسات شبة التجريبية ذات القياسات القبلىة للمجموعة الضابطة والتجريبية للمتغيرات الاكثر ارتباطاً لتحديد التماثل او

التكافؤ بين المجموعتين على الرغم انها تمدنا بتأثيرات عالية للبرامج والصدق الداخلى لها منخفض. ويرى (Shadish & Myers (2004) بضرورة اجراء تحليل منفصل للدراسات التجريبية ذات المجموعات العشوائية وللدراسات شبه التجريبية (توزيع غير عشوائى) واختبار الفرق بينهما باستخدام تحليل التفاعل فلو اعطت احجام تاثير متجانسة فيمكن دمج نتائج التصميمين فى ما وراء تحليل واحد. توصل (Shadish & Myers (2004) الى ان الدراسات التى اعتمدت على تصميم المجموعات غير العشوائية (شبه التجريبية) تعطى احجام تاثير أعلى من نظيرتها (التجريبية) ذات التوزيع العشوائى وهذا تحيز موجب.

وكذلك يحدد الباحث معيار التضمن للدراسات التى حددت مبررات لكيفية التغلب على مهددات الصدق الداخلى Internal validity للاستنتاج السببية بمعنى اثبات ان التحسن الحادث يتم ارجاعه الى البرنامج فقط وليس الى عوامل دخيلة فى التجربة. تصميم الدراسة وجودة التنفيذ تلعب دوراً كبيراً فى اختلاف النتائج وتقود الى استنتاجات خاطئة.

ودائماً يتم استبعاد دراسات المجموعة الواحدة قياس قبلى وبعدى لأنها معرضة بدرجة شديدة لانخفاض الصدق الداخلى لها بمعنى تكون عرضة للمتغيرات الدخيلة حيث لايمكن ضبط تاثير هذه المتغيرات وعليه لايمكن ارجاع التحسن الى المعالجة بل الى عوامل اخرى ليست فى الحسبان (Littell et al., 2008)، ولكن منهجياً يمكن اجراء دراسات ما وراء التحليل لدراسات المجموعة الواحدة ومقارنتها بدراسات ما وراء التحليل لتصميمات المجموعتين لنفس الموضوع.

وعلى ذلك لابد ان يحدد الباحث نوعية المنهج البحثى المستخدم فى الدراسات المتضمنة فى تحليل ما وراء فهل يعتمد على دراسات تجريبية فقط؟، ام دراسات شبه تجريبية فقط؟، ان كانت شبه تجريبية هل تصميم المجموعة الضابطة والتجريبية وقياس قبلى وبعدى ام تصميم المجموعة الواحدة وقياس قبلى وبعدى؟.

ويجب على الباحث ان يتبع الخطوات الاتية:

- ضع سقف ادنى من التصميمات المقبولة التى تناولت الموضوع.

- قدر مدى الاختلاف بين الدراسات المتضمنة في ضوء تصميم البحث وخصائص التنفيذ التي ربما تزيد من مخاطر التحيز.

- استخدم تحليل التفاعل لتقدير تأثيرات التصميمات المختلفة.

● **المدى اللغوي والثقافي Cultural and Linguistic Range**: لابد من الاعتماد على الدراسات في البيئة الثقافية للباحث بقدر الامكان نظراً لاختلاف المفاهيم من ثقافة الى اخرى. فمثلاً معنى مفاهيم مثل السرقة او الاحتيال او السعادة في البيئة العربية مختلف عن معناها في البيئات الاخرى، وهذا بدوره يؤدي الى دمج دراسات مختلفة من حيث التعريف الاجرائي للمتغيرات. ولكن اذا تاكد الباحث ان المفهوم له نفس المعنى عبر الثقافات المختلفة فلا مانع من التضمين ولكن من الافضل اجراء مقارنة بين الدراسات عبر البيئات الثقافية المختلفة من خلال تحليل المتفاعل.

ولابد الاخذ في الاعتبار معيار لغة النشر وهذا مرتبط بقضية تحيز النشر، فهل يعتمد الباحث على الدراسات باللغة العربية فقط ام على الدراسات العربية والمنشورة باللغة الانجليزية، هذا يتوقف على الهدف من ما راء التحليل هل تحديد فعالية البرنامج في البيئة المصرية فقط ام في البيئة العربية ام الوصول الى بناء او تطوير نظرية بالتالي على اكبر قدر من الدراسات في كل اللغات.

● **اطار الوقت Time Frame**: من المهم وضع حدود على المدة الزمنية التي تغطيها عملية البحث في التراث البحثي فمن غير المعقول ان يتناول الباحث دراسة الموضوع منذ مثلاً 1970 فهذا يمثل صعوبة كبيرة على الباحث من حيث الوقت والتكلفة. يمكن للباحث على الدراسات في المدة الزمنية من مثلاً 2000 الى 2010 او الدراسات المنشورة بعد عام 2010. ولكن يمكن اجراء ماوراء التحليل التراكمي Cumulative meta-analysis لاختبار ما اذا كان تاثير المعالجة تختلف عبر الزمن وهذا يتطلب تضمين الدراسات التي اجريت خلال كل الفترات الزمنية وايضاً يمكن استخدام مدخل التحليل المتفاعل Moderator analysis لدراسة تاثير الزمن على فعالية معالجة ما عبر فترات زمنية مختلفة.

● **نوع النشر Publication Type**: ماهي طبيعة الدراسات التي يرغب الباحث تضمينها في تحليل ما وراء؟، فمثلاً يرغب الباحث في تضمين الدراسات المنشورة في مجلات علمية محكمة

والدراسات غير المنشورة مثل ابحاث الماجستير والدكتوراة وابحاث المؤتمرات وكذلك التقارير الحكومية وغيرها. فالاعتماد على الدراسات المنشورة فقط تعتبر اشكالية كبيرة نتيجة الاحتمالية العالية لوجود تحيز النشر جراء ان الابحاث ذات النواتج الدالة تجد طريقها للنشر عكس الابحاث ذات النواتج غير الدالة التي فى احيانا كثيرة يتم رفضها للنشر. ويرى McLeod & Weiz (2004) ان الدراسات غير المنشورة اكثر جودة من نظيراتها غير المنشورة، وعليه فان تحيز النشر يمثل اشكالية خطيرة لمصادقية المراجعة للتراث البحثى.

وكلما كان الباحث اكثر مرونة فى تبنى معايير منهجية فان هذا يؤدى الى نتائج غير دقيقة ولكن وضع معايير منهجية صارمة يؤدى ايضا الى الحصول على عينة صغيرة من الدراسات ويمكن استبعاد دراسات مفيدة من التحليل وهذا من شأنه يؤدى الى الحصول على عينة من الدراسات غير ممثلة تمثيلاً جيداً لمجتمع الدراسات وبالتالي يحد من القدرة التعميمية للنتائج. واجريت العديد من دراسات ماوراء التحليل لمعرفة مدى وجود فروق بين دراسات ما وراء التحليل ذات المعايير المنهجية الصارمة بنظيرتها ذات المعايير المنهجية المرنة ولم تتضح وجود فروق (Shadish, 1995).

ويرى (Lipsey & Wilson, 2001) ان الجودة المنهجية فى تراث البحث الاجتماعى والسلوكى فقيرة، فالقارير غامضة عن توضيح القضايا والتفاصيل المنهجية والاجرائية، فباحث ما وراء التحليل يعانى من ان الدراسات لاتمدد بالمعلومات الكافية عن المعايير المنهجية التى يتبناها.

الجودة المنهجية ليست شئ بسيط انما تتضمن عوامل عديدة العينة، العشوائية، الصدق، الثبات، التصميم، التكافؤ، وغيرها. فالمعايير المفضلة لدراسة لموضوع ما فقد تختلف عن دراسة موضوع اخر ولكن المعايير التى لا خلاف عليها هى قضايا ثبات البيانات والتصميم التجريبي المستخدم.

وبعض الباحثون وضعوا اطر وقوائم لتقدير الجودة المنهجية ولكن لا يوجد اتفاق بين الباحثين حول مكونات هذه الاطر وهى تتضمن الخصائص المنهجية الجيدة التى يجب ان تتسم بها الدراسات.

وتوجد علاقة تبادلية بين قيود معايير التضمين للدراسات المتضمنة في ما وراء التحليل والملاحم المنهجية للدراسات، فكلما زاد تنوع واتساع هذه المعايير كلما حصل الباحث على مدى واسع لدراسات ذات ملامح منهجية واجرائية متنوعة.

عموماً معايير التضمين والاستبعاد المرنة والصارمة لها مميزات أهمها (Card, 2012):

- تضمين أكبر عدد من الدراسات الممكنة واستبعاد عدد محدود بينما المعايير الصارمة تستبعد عدد كبير من الدراسات وتضمن عدد قليل. والاختيارين يعكسوا نقطتي النهاية على متصل. والمعيار الأهم للتضمين هو مجتمع الدراسات التي بصدد الوصول الى استنتاجات بشأنه، وكذلك معيار الجودة المنهجية فهل الباحث على استعداد لتضمين دراسات تعاني من خلل منهجي ما وهذا من شأنه يزيد من انتقاد اسلوب ما وراء التحليل بمعنى " زباله دخلت التحليل، زباله خرجت من التحليل" ولكن المعايير الصارمة تجعل استنتاجات النتائج لمجتمع محدود.

- المعايير المرنة تجعل الدراسات أكثر تنوعاً في ملامحها المنهجية بينما المعايير الصارمة ينتج عنها عدد محدود من الدراسات ومتشابهه في ملامحها وهذا يؤدي الى قوة احصائية غير مناسبة لتقويم متوسط حجم التأثير.

- **معلومات حجم التأثير Effect Size Information:** في النهاية من المعايير الهامة توافر البيانات الكافية لحساب حجم التأثير، وفي معظم الدراسات تتوفر هذه المعلومات مثل المتوسطات والانحرافات المعيارية وحجم العينة. وإذا لم تتوفر هذه المعلومات فيمكن التواصل مع المؤلف لطلب هذه المعلومات.

ولابد من اعداد بروتوكول او اطار او خطة مكتوبة يتضمن ماهية المعايير التي يعتمد عليها باحث ما وراء التحليل حتى يسمح لمراجع اخر يكود النواتج البحثية المتطلبة لانجاز دراسة ما وراء التحليل.

ويتضمن البروتوكول الآتي:

- غلاف يتضمن عنوان دراسة ما وراء التحليل، تفاصيل الاقتباس، والتواصل مع المؤلفين.
- مقدمة (الخلفية واهداف الدراسة).

- المنهجية وتتضمن معايير التضمين والاستبعاد، طرق البحث (قواعد البيانات، الكلمات المفتاحية، استراتيجيات البحث)، استخلاص البيانات (شيت التكويد، تقدير ثبات المقدرين)، تحليل البيانات والتوليف (كيفية تنفيذ ما وراء التحليل، تقدير تحيز النشر، التحليل المتفاعل، تحليل الحساسية)، التفسير، الجداول والاشكال، والمراجع.

تحديد نواتج البحث المتطلبية لإجراء دراسة ما وراء التحليل

يرى (Lipsey & Wilson 2001) ان نواتج الدراسات الكمية تاخذ اشكال عديدة مثل:

- الفروق بين متوسطات المجموعات (الفروق بين الذكور والاناث فى الدافعية).
- العلاقات بين المتغيرات (العلاقة بين القلق والتحصيل).
- نسب لملاحظات او ظواهر معينة فى المجتمع (نسبة مدمنى الفيس بوك او نسبة انتشار الجريمة فى المجتمع).

تتعدد النواتج المراد تجميعها من الدراسات اهمها:

- **حجم التأثير:** تجمع نتائج الدراسات لاستخدامها فى ما وراء التحليل فى صورة حجم التأثير ولمقارنة النتائج بصورة لها معنى لابد ان يستخدم نفس حجم التأثير لتكويد كل النتائج حتى تكون قابلة لاستخدامها فى ما وراء التحليل، وعليه فلا بد من تحديد شكل النواتج البحثية المناسبة لما وراء التحليل وتعرض فى ضوء احصاء حجم التأثير العام واذا وجد صور مختلفة لاحجام التأثير فلا بد من تصنيف النتائج واجراء ما وراء التحليل منفصل لكل حجم تأثير وكذلك يمكن اجراء تحويلات حجم التأثير من صورة الى اخرى مثل تحويل مؤشر d الى مؤشر معامل الارتباط r .

على ذلك فالمكون العام لنوعى التحليل سواء دمج او مقارنة هو احجام التأثيرات حيث تعكس ماهية اسلوب ما وراء التحليل. وتتنوع مؤشرات حجم التأثير الاكثر استخداماً فى ما وراء التحليل منها مؤشرات فروق المتوسطات المعيارية، معامل الارتباط، مؤشر نسبة او احتمال الترجيح Odds-ratio وغيرها، وعلى الرغم من المؤشرات المختلفة لاحجام التأثير الا انها فى معظمها تعكس العلاقة بين متغيرين، والعلاقة المتحصل عليها من الدراسات الارتباطية تعكس درجة

التلازم بين متغيرين ويمكن ان تكون مقياس تقريبي للعلاقات السببية Causal relations ولكن العلاقات من الدراسات التجريبية فهي مقياس حقيقي للسببية على الرغم من مهددات الصدق الداخلي في التصميمات التجريبية.

واى من انواع العلاقات هي مقياس لحجم التأثير المستخدم في ما وراء التحليل، ويوجد العديد من مؤشرات حجم التأثير متضمنة المتوسطات، النسب، ثبات المقاييس، درجات التغير الطولى (الدراسات الطولية) ولكنها اقل استخداماً في ما وراء التحليل.

واحجام التأثير المشتقة من دراسات ارتباطية، طولية، شبة تجريبية، تجريبية تعكس مدى من الاسئلة البحثية يمكن الاجابة عليها باستخدام ما وراء التحليل (Cooper, 2009; Lipsey & Wilson, 2001).

• **احصائيات النزعة المركزية Central tendency statistics:** نواتج البحث تصف خصائص المتغيرات المقاسة لعينة وحيدة من المستجيبين وتوزيع قيم المتغيرات تلخص في ضوء مؤشرات النزعة المركزية مثل المتوسط، الوسيط، المنوال، والنسب. واذا توافرت هذه المؤشرات لنواتج الدراسات فيمكن اجراء ما وراء التحليل للدراسات. لاحظ ان هذه المؤشرات تكون لدراسات متنوعة تناولت نفس المتغير وغالباً تستخدم نفس المقياس كمتوسط الذكاء كما قيس في اختبار ستانفورد- بينيه . ويمكن حساب مؤشرات النزعة المركزية لدراسات استخدمت مقاييس مختلفة لنفس المفهوم شرط تحديد التعريف الاجرائي للمتغير عبر كل الدراسات.

• **مقارنات المجموعات Group contrasts:** هذا النوع من نواتج البحوث يتضمن متغير واحد فاكثر تم قياسه لمجموعتين فاكثر من الافراد ثم تتم المقارنة بين المجموعات ويتم عرض المتوسطات او النسب لكل مجموعة. وفي هذا الاطار يوجد شكلين لمقارنة المجموعات وهي موضع اهتمام ما وراء التحليل:

- **البحوث التجريبية او المحاولات الاكلينيكية Experimental or clinical trials:**

حيث تتم المقارنة بين مجموعة تعرضت لمعالجة ما تجريبية (مجموعة تجريبية) بمجموعة لم تتلقى معالجة او معالجة تقليدية (مجموعة ضابطة) حيث تكون المقارنة على قيم المتغير الناتج (التابع) وتفسر على انها تأثير المعالجة.

- **بحوث الفروق بين المجموعات** Group differences research: تتم المقارنة في ضوء خصائص المجموعة مثل المقارنة بين الذكور والاناث في التحصيل او اى متغيرات اخرى تابعة او المقارنة بين طلاب ذى صعوبات التعلم ومنخضى التحصيل باقرانهم العاديين فى القدرة القرائية، وهذه النوعية من الدراسات تتبع من دراسات المنهج السببى المقارن .Causal comparative method.

وهذه النوعية من الدراسات سواء التجريبي او مقارنة المجموعات هى شائعة فى التراث البحثى لما وراء التحليل وخاصة البحوث التجريبية التى تهدف الى تقدير فعالية المعالجة او التدخل (Lipsey & Wilson, 2001; Rosenthal, 1994). وعندما تتم المقارنة بين ثلاث مجموعات فاكتر فيتم دراسة كل مجموعتين معاً على حدة نظراً لان الفنيات المستخدمة للمجموعات الثلاثة معاً غير متاحة (Lipsey & Wilson, 2001) .

- **العلاقة بين المتغيرات** Association between variables: هذا النوع من نواتج البحوث يعكس التباير عبر الافراد لمتغيرين لتحديد درجة العلاقة بينهما، كالعلاقة بين ماوراء التعلم وما وراء الذاكرة، ويعتبر معامل الارتباط مؤشر لحجم التاثير او العلاقة بين الحالة الاقتصادية الاجتماعية للأسرة وتحصيل الرياضيات فى المرحلة الثانوية، وتعتبر مؤشرات كاي تربيع او معامل فاي او لمدا λ او نسبة Odds-ratio مؤشرات حجم التاثير، وفى هذه الفئة من البحث يوجد تصنيفين عريضين من اهتمامات ماوراء التحليل هما:

- **بحوث القياس** Measurement research: وهذه النوعية من بحوث العلاقة تكون منصبة على تقدير خصائص ادوات القياس مثل حساب معامل الارتباط بين درجات الاختبار واعادته Test-retest كمؤشر لمعامل الثبات، واستخدام معامل الارتباط لتقدير الصدق التنبؤى Predictive validity بين درجات اختبار القدرات والتحصيل اللاحق فى الجامعة ويستخدم للتحقق من صدق اختبار الاستعدادات او القدرات.

- **بحوث الفروق الفردية** Individual differences research: وهذه النوعية من بحوث الارتباط هى الاكثر استخداماً وهى تهتم بدراسة التباير بين سمات مختارة للافراد

كالعلاقة بين ادمان الانترنت والوحدة النفسية، او العلاقة بين التنشئة الاسرية والعنف المدرسى، او العلاقة بين وقت المذاكرة درجات الطلاب فى الرياضيات وغيرها. وتعتبر دراسات ما وراء التحليل لاجاث العلاقات او الارتباطات شائعة الاستخدام فى التراث وذلك لتوافر العديد من مؤشرات احجام التأثير المناسبة (Hunter & Schmidt, 2004; Rosenthal, 1994)، وتعتبر قيمة معامل الارتباط مؤشراً لحجم التأثير. وتعتبر الاشكال المختلفة من الدراسات سواء الارتباطية او الفارقة تعتبر مادة خصبة لدراسات ما وراء التحليل وما دون ذلك فمن الصعب اجراء دراسات ما وراء التحليل له باستخدام مؤشرات حجم التأثير مثل نواتج دراسات التى استخدمت تحليل التباين المتدرج MANOVA او الانحدار المتعدد او التحليل التمييزى او دراسات التحليل العاملى او دراسات نمذجة المعادلة البنائية (Lipsey & Wilson, 2001) Structural Equation Modeling.

الفصل الثالث

مسح وتحديد التراث البحثي

Scanning and Identifying Research Literature

مقدمة

بعد التحديد الجيد لمعايير التضمين والاستبعاد للدراسات المناسبة لما وراء التحليل التي حددت طبيعة مجتمع الدراسات، فإن الخطوة التالية هي البحث عن الدراسات الامبريقية التي تتوافق مع اطار المعاينة ومن المحتمل ان يتضمن مجتمع الدراسات عينة ممثلة بدرجة كافية، بالاضافة الى ان المحلل الاحصائي لما وراء غالباً يرغب في تقسيم الدراسات الى تصنيفات مختلفة مع توافر عدد كافي من الدراسات في كل تصنيف ليمح باجراء التحليل والمقارنة لذلك فان الباحث يحاول تحديد واسترجاع كل دراسة في المجتمع وليس عينة من هذا مجتمع الدراسات.

ما وراء التحليل هي دراسة ملاحظة Observational Study قائمة على نتائج الدراسات المتاحة للعلاقة بين المتغيرين في التراث البحثي، وعليه فالباحث مهتم في المقام الاول بعض قائمة من بعض اشكال نواتج الدراسات التي تمثل احصائيات احجام التأثير التي يعتمد عليها محلل ما وراء التحليل.

اجراء عملية البحث

تتضمن عملية البحث عن الدراسات عدة خطوات كما حددها Reed & Baxter (2009) كالتالي:

- **تحديد استراتيجية البحث:** وضع خطة لعملية البحث في التراث البحثي تتضمن السؤال البحثي او الموضوع المراد البحث فيه وتوصيفات هذا الموضوع من حيث عينته ومنهجه ومقاييسه وغيرها، وماهي قواعد البيانات المراد البحث فيها، وما الفترة الزمنية المراد البحث فيها.
- **خطوات عملية البحث:** وتتضمن الخطوات الاتية:
 - **تحديد الموضوع:** متضمناً المتغيرات، العينات، والمنهجية.

- **تحديد المصطلحات او الكلمات المفاتيحية:** من الافضل تضمين مترادفات المفهوم او المصطلح للحصول على اكبر عدد من الدراسات.
- **بروفيل البحث:** ما طبيعة الربط المنطقي بين المفاهيم مثل وضع AND او علامة (+) وتستخدم للحصول فقط على الدراسات التي تتضمن المفهومين فقط، اما OR تستخدم للبحث فى مفهومين بنفس المعنى بهدف توسيع دائرة البحث مثل Program or intervention او High school or secondary school ، اما الاختيار NOT يستخدم لاستبعاد عينة معينة مثل Creativity and intelligence not children .
- **البحث الرئيسى:** يتضمن تحديد قواعد البيانات الرئيسية الالكترونية والحصول على المصادر او الدراسات الرئيسية.
- **تقويم الدراسات:** وهذه خطوة تتضمن تقويم الدراسات من حيث ارتباطها بالموضوع وهذه عملية صعبة وتحتاج جهد ووقت.

مصادر الحصول على التراث البحثى المناسب

ولكن قبل عرض استراتيجيات الحصول على التراث البحثى فمن الضرورى التميز بين مفهومين هما الاستدعاء Recall والدقة Precision (Card, 2012) فالاستدعاء هو نسبة الدراسات التي تم الحصول عليها من تلك الدراسات التي يجب استدعائها والتي توافر فيها معايير التضمين والاستبعاد وهذا لا يمكن تحقيقه لان الباحث لايعرف كم عدد الدراسات الموجودة بالفعل فى التراث البحثى للظاهرة. اما الدقة هي نسبة الدراسات التي توافرت فيها معايير التضمين او الاستبعاد.

ومخاطر الاستدعاء غير الدقيق فقد كثير من الدراسات كان يجب تضمينها وهذا يؤدي الى نقصان القوة الاحصائية والحصول على نتائج متحيزة خاصة لو ان الدراسات المستبعدة اختلفت كثيراً عن الدراسات المتضمنة.

اما مخاطر عدم الدقة هو تضيع الوقت فى مصادر الاسترجاع وقراءة الدراسات التي لا تتضمن فى ما وراء التحليل، والهدف من استراتيجية البحث تحقيق استرجاع للدراسات كلما امكن بدون نقصان الدقة وهذا يحقق قوة عالية باقل قدر من التحيز.

وتوجد العديد من مصادر الحصول على الدراسات المناسبة لما وراء التحليل مثل الملخصات، قواعد البيانات المتخصصة، المجلات المفهرسة، المجلات المناسبة، المؤتمرات والندوات، الخبراء في المجال، والتقارير الحكومية. وفيما يلي عرض لبعض هذه المصادر:

قواعد البيانات الالكترونية

توجد العديد من قواعد البيانات المتاحة Databases على الانترنت في معظم المكتبات الجامعية وهذا يجعل مهمة البحث في التراث البحثي عملية سهلة مقارنة بالوضع في اواخر القرن الماضي وهذا بدوره ييسر من اجراء دراسات ما وراء التحليل، وقواعد البيانات متاحة في كل التخصصات ففي الاقتصاد (Econlit) وفي التربية (ERIC) وفي الطب (Medline) وفي علم النفس (PsycINFO) وفي علم الاجتماع (Sociological Abstracts).

وهذه القواعد تغطي مدى واسع من الدراسات في هذه التخصصات وتعتبر ادوات بحث رئيسية في ما وراء التحليل وتعتبر الخطوة الاولى في عملية البحث ولكنها ليست كافية لكنها تعتبر اداة البحث جيدة، وعملية البحث هدفها الحصول على ملخص شامل بقدر الامكان للتراث البحثي للموضوع ولذلك من الضروري البحث في قواعد بيانات عديدة. ويوجد ثلاث اعتبارات لاستخدام قواعد البيانات حددها Card (2012) كالتالي:

اولاً: ما الدراسات المتضمنة والمستبعدة؟: توجد اعتبارين قبل استخدام قاعدة البيانات:

- **طبيعة الدراسات المتضمنة:** هو ما هي الدراسات المراد تضمينها في دراسة ما وراء التحليل من قاعدة البيانات؟، للاجابة عن هذا السؤال يتطلب قراءة توثيق قاعدة البيانات ومعرفة المجالات التي تهتم بها، فبعض قواعد البيانات تهتم بالاعمال غير المنشورة مثل رسائل الماجستير والدكتوراة مثل Proquest فاذا كانت معايير التضمين الاعتماد على الدراسات غير المنشورة فيجب البحث في هذه القواعد.

- **مدى شمولية قاعدة البيانات للدراسات المنشورة:** فقواعد البيانات دائماً تتضمن الدراسات المنشورة في مجلات معروفة وعامل التأثير Impact factor لها عالي وكذلك تتضمن الكتب المنشورة من دور نشر زائفة الصيت ولا تتضمن كتب من دور نشر غير

معروفة، وعلى ذلك فالاعتماد على قاعدة البيانات تتضمن دراسات دون اخرى يؤدي الى الحصول على عينة متحيزة من الدراسات.

ثانياً: **الكلمات المفتاحية Key words** : بعد تحديد قاعدة البيانات المناسبة للحصول على الدراسات فلا بد من تحديد مصطلحات البحث Search term (الكلمات المفتاحية) التي تعتمد عليها في عملية البحث بعناية وبدورها تساعد على ان تكون عملية البحث محددة وفعالة وهذه المصطلحات يمكن ان تكون المجتمع، طبيعة البرنامج التدخلي، والنواتج. ويمكن البحث بدمج اكثر من مصطلح باستخدام حروف مثل علامة "*" خاصة في قاعدة البيانات PsycINFO اوعلامة "&" للربط بين العديد من الكلمات المفتاحية ولا بد ان تكون الكلمات المفتاحية مستندة على المصطلحات المألوفة ومحددة للموضوع او الظاهرة موضع الدراسة في تراث الظاهرة.

فمثلاً عنوان دراسة ما وراء التحليل: اثر برنامج مستند على عادات العقل لتنمية التفكير المنظومي فان الكلمات المفتاحية: عادات العقل والتفكير المنظومي واحياناً كلمة برنامج.

واذا كنت بصدد البحث في قواعد البيانات الاجنبية فراعى الكلمات المترادفة مثل مفهوم المراهقة adolescence وقد تستخدم بمعنى Teen وعليه يتم البحث كالاتى (or Teen) *adolescence) و كلمة البرنامج يمكن ان نبحث عنها بالمترادفات الاتية:

(Program* or service* or* treatment* or intervention* or prevent*)

ولذلك يجب استشارة المتخصصين لتحديد مترادفات المصطلح. فالبحث في قاعدة البيانات لعنوان "برنامج لخفض الاكتئاب لدى المراهقين" فان الكلمات المفتاحية كالاتى:

(Adolescence* or teen) And depression* and (Program* or service* or treatment* or intervention* or prevent*)

وهذا البحث اكثر تحديداً لانه يمدنا بالدراسات التي تتضمن المراهقة والاكتئاب والبرنامج معاً.

او "Adolescence, depression, and intervention"

بالتالى يسترجع الدراسات التي تضمنت هذه الكلمات معاً.

فكلمة * or او ؟ تستخدم للبحث باستخدام كلمات مختلفة لها نفس المعنى فى التراث البحثى.

وكلمة "و" and تستخدم للبحث للعلاقات بين مفهومين مثل عادات العقل ودافعية الانجاز، يمدنا بالدراسات التى تتضمن العلاقة بينهما.

معظم برامج البحث تتضمن كلمة لا "Not" اى استبعاد الدراسات التى تتضمن كلمة او مصطلح معين. ولتحديد الدراسات المستبعدة يمكن استخدام كلمة Not. ولمحدودية مجال عملية البحث يمكن وضع العينة موضع اهتمام الدراسة ككلمة مفاتيحية.

فيما يلى امثلة لكلمات مفاتيحية:

- نوع العينة: المراهقين، المراهقة، الذكور، الاناث، اطفال الاحداث، طلاب المرحلة الثانوية او المدارس العليا، تلاميذ المرحلة الابتدائية، المعاقين سمعياً.....الخ.
- نوع التدخل **Type of intervention** او البرنامج: برنامج تدريبى، برنامج علاجى، برنامج ارشادى، برنامج تحدى،الخ
- نوع البحث **Type of Research**: تقويم، تاثيرات، فعالية، تجربة، تجريبى مجموعة ضابطة، ارتباطى،.....الخ

ثالثاً: المحاذير **Cautions**: تعتبر قواعد البيانات ادوات فعالة وقوية للحصول على الدراسات المناسبة للتحليل ولكن يوجد ثلاثة محاذير لابد وضعها فى الاعتبار:

- لابد ان يدرك الباحث ماهى طبيعة الدراسات المتضمنة وغير المتضمنة فى قاعدة البيانات فلو لم تتضمن الدراسات غير المنشورة او المجالات ضعيفة عامل التأثير عندئذ يؤدى الى نقصان الدراسات السابقة فى ما وراء التحليل. وهذا النقصان مهدد لاسلوب ما وراء التحليل لانه يؤدى الى انخفاض القوة الاحصائية وحدوث اشكالية تحيز النشر خاصة لو ان الدراسات غير المتضمنة تختلف عن الدراسات فى التحليل. ولتجنب هذه المشكلة يجب تحديد قواعد بيانات بديلة والبحث عن استراتيجيات بحث بديلة مثل البحث اليدوى.
- معظم قواعد البيانات تكون متخصصة فمنها متخصص فى العلوم الاجتماعية ومنها متخصص فى العلوم الطبية وهكذا، ولكنها يمكن ان تتضمن اعمال او دراسات فى تخصصات اخرى وعلى ذلك لاتعتمد على قاعدة بيانات وحيدة فى تخصصك. فمثلاً العدوان الاجتماعى

Social aggression لا يظهر في علم النفس (قاعدة بيانات PsycINFO) فقط انما يمكن ان يكون في تخصصات الجريمة، التربية، الطب، الصحة العامة، والاجتماع. ولذلك ينصح Card (2012) بضرورة البحث في قاعدة او اثنين من قواعد البيانات من تخصصات اخرى.

- البحث في قواعد البيانات يعطى قائمة من الدراسات المتاحة ولكن لا يعطى معلومات عن كم الدراسات غير المتاحة وتبقى امكانية ان بعض الدراسات العالية الجودة لم نحصل عليها نتيجة عدم تناظرها لمعايير البحث. ويمكن حل هذه الاشكالية باستخدام مصطلح اكثر عمومية مثل العدوان بدلاً من العدوان الاجتماعي ثم اجراء مسح للدراسات لمعرفة مدى توافر دراسات اضافية ومناسبة.

وللحصول على دراسات مناسبة لطبيعة الموضوع فلا تقتصر على البحث في قواعد البيانات الالكترونية ولكن لابد من الاعتماد على البحث التقليدي اليدوي Hand searching خاصة في المجالات المتخصصة وبحثها يدوياً باسترجاع الصفحة لكل محتواها وهذا ضرورى لان احياناً بعض الدراسات غير متضمنة في قواعد البيانات الالكترونية وقد يلجا الباحث الى البحث اليدوي عندما لا يحصل على عدد كافي ومناسب من الدراسات الاجراء ما وراء التحليل.

وفى واقع البيئة العربية فانه توجد قواعد عربية للتراث فى البيئة العربية فى كل التخصصات كما اتاحت معظم الجامعات العربية قواعد بيانات من خلال عالم المعرفة او المنظومة وهذه القواعد توفر قواعد بيانات فرعية مثل Academic Search Ellit و EBSCO host و ProQuest و Ovid و Science Direct وغيرها. وتوجد مصادر عديدة للحصول على الدراسات مثل الشبكة العنكبوتية World Wide Web والمؤتمرات، المجالات العلمية، المكتبات، المجالات الالكترونية، وقواعد البيانات، ملخصات الابحاث العالمية، مؤشر الاستشهاد العلوم الاجتماعية، وغيرها).

مجلدات المراجع البيبلوغرافية Bibliographical Reference volumes

المجلدات المرجعية هي اعمال مطبوعة تمدنا بمعلومات مشابهة كتلك التى تمدنا بها قواعد البيانات الالكترونية مثل العناوين، المؤلفين، والمستخلصات. تتضمن قائمة من الدراسات

يمكن الحصول عليها عن طريق الكلمات المفتاحية. هذه المجلات يتم نشرها عن طريق الجمعيات البحثية الكبيرة مثل الجمعية النفسية الأمريكية American Psychological Association (APA) فهي تطبع المستخلصات الابحاث النفسية Psychological abstracts من 1927 حتى 2006 وكذلك مستخلصات ابحاث علم الاجتماع Sociological abstracts ولكن البحث في هذه المجلات ليس بقوة البحث في قوعد البيانات ونادراً ما يعتمد عليها في دراسات ما وراء التحليل لانها احياناً لا تتضمن كل المعلومات المطلوبة لاجراء التحليل ولكنها احياناً مفيدة لانها تتضمن دراسات لتتضمنها قواعد البيانات الالكترونية.

الدراسات غير المنشورة Unpublished Studies

احد المخاطر في دراسات ما وراء التحليل هو تحيز النشر ولكن كيف للباحث ان يتجنب ويقيم هذه الاشكالية من خلال البحث وتضمين الدراسات غير المنشورة في ما وراء التحليل، وتعتبر قواعد البيانات الالكترونية احد المصادر الحصول على رسائل الماجستير و الدكتوراة غير المنشورة ويضيف (Card 2012) ثلاثة مصادر اخرى للحصول على الدراسات غير المنشورة كالتالى:

- **دراسات المؤتمرات:** من خلال البحث بعنوان المؤتمر فيمكن عرض ملخصات الابحاث وبالتواصل مع المؤلفين يمكن الحصول على نسخ هذه الدراسات ولكن من الافضل التواصل مع اصحاب الدراسات مباشرة بعد المؤتمر بمدة قصيرة وليست بعدها بسنوات كثيرة. وللحصول على المؤتمر ابحث بمعلومات تفصيلية مثل عنوانه وتاريخ انعقاده ومكانه، وتواصل مع الباحثين المشتركين في البحث ان تعذر التواصل مع الباحث الرئيسي (اكثر من باحث). ويعتقد Card (2012) ان ابحاث المؤتمرات على درجة كبيرة من الاهمية ومصدر جيد للدراسات غير المنشورة الا انه توجد محددات ومحاذير تؤخذ في الاعتبار اهمها ان البحث في المؤتمرات يجب ان يكون بناءً ومنظماً Systematic فيجب بذل الجهد للحصول عن كتاب المؤتمر المتضمن الدراسات، ويجب ان تتأكد ان الدراسة تتضمن المعلومات التي تحتاج اليها في تكويد الدراسة وكذلك معلومات لحساب حجم التأثير. وتوجد بعض قواعد البيانات تتضمن اعمال المؤتمرات مثل ERIC في التربية.

- الدراسات الممولة من جهات هيئات Funding Agency studies: من اهم مصادر الحصول على دراسات غير منشورة هي الابحاث التي تمولها مؤسسات او هيئات خاصة او حكومية ولكن يرى البعض ان هذه النوعية من الدراسات تعاني من تحيز نتائجها.
- ملخصات رسائل الدكتوراة العالمية **Dissertation Abstract Online**: تمدنا بملخصات لمعظم رسائل الدكتوراة فى الولايات المتحدة وكندا وهذه الرسائل تكون مناسبة لدراسات ما وراء التحليل.
- التراث **Grey Literature**: التراث البحثى المنتج من الحكومات والمصانع والمؤسسات التجارية سواء كانت مطبوعة او الكترونياً ولا يخضع للنشر التجارى.

استراتيجية البحث الخلفى **Backward Searches**

بعد جمع وتراكم مجموعة كبيرة من الدراسات لتطبيق معايير التضمين الممكنة فى ما وراء التحليل يبدأ الباحث فى عملية التكويد لهذه الدراسات ولكن لابد ان يقرأ محتوى الدراسات كاملة خاصة المنهجية المستخدمة والنتائج التى تتضمن معظم المعلومات المراد تكويدها وكذلك يمكن الاستشهاد ببعض الدراسات المذكورة فى الدراسات المراجعة والمناسبة للموضوع المراد درسته وايضاً من الضرورى قراءة دراسات ما وراء التحليل السابقة التى اجريت فى نفس الموضوع ان وجدت او موضوعات قريبة لموضوع بحثك. وعملية البحث فى الدراسات المناسبة التى وردت فى الدراسات المراجعة تسمى البحث الخلفى وهذا المدخل مفيد فى تحديد الدراسات القديمة ولكن يوجد عيب فى هذا المدخل هو احتمال التحيز حيث ان الدراسات التى تعطى نتائج مرغوبة (دالة احصائياً) اكثر احتمالية لتضمينها فى الدراسة المراجعة مقارنة بالدراسات غير المرغوبة (النتائج التى لا تعطى دلالة احصائية). وعلى الرغم من قضية التحيز الا ان استراتيجية البحث الخلفى تمدنا بمصدر جيد للبحث خاصة انها احياناً تمدنا بدراسات تفشل مصادر البحث الاخرى امدادنا بها.

استراتيجية البحث الامامى **Forward Searches**

بينما البحث الخلفى يحاول الوصول للدراسات المذكورة فى الدراسات التى جمعتها فان البحث الامامى محاولة ايجاد الدراسات التى استشهدت بالدراسات التى جمعتها ويمكن اجراء

ذلك من خلال قواعد بيانات متخصصة مثل مؤشر استشهاد العلوم الاجتماعية Social Science Citation Index او قاعدة بيانات علم النفس PsycINFO فهي تمتلك هذه الخاصية. ولإجراء البحث الامامى ادخل معلومات الدراسة التي تعتبر مناسبة لما وراء التحليل والتي تمتلكها في قاعدة البيانات وسوف يبين لك الدراسات الاخرى التي استشهدت بها لان هذه الدراسات يمكن ان تساعد الباحث في الحصول على دراسات جديدة ويمكن اجراء ذلك في اى قاعدة بيانات شرط ادخال الكلمات المفاتيحية التي استخدمتها الدراسة. ويمكن اجراء البحث الامامى لكل الدراسات المتضمنة في التحليل بعد تطبيق معايير التضمين. ويرى Card (2012) ان هذه الاستراتيجية البحثية غالباً تكون محدودة في دقتها.

التواصل مع الباحثين والخبراء في المجال Communication with Researchers and Experts

المدخل الاخير للحصول على الدراسات هو استشارة الخبراء او الباحثين في مجال دراسة الموضوع المراد اجراء ما وراء التحليل فيه، ويمكن استشارة الخبير في مدى مناسبة معايير التضمين او الاستبعاد وقائمة الدراسات التي حددتها للتحليل وسؤاله عن ما كنا بحاجة لتضمين دراسات اضافية، ويفضل سؤال الخبراء الذين لهم اسهامات ورؤى عن الموضوع وذلك لاثراء دراسة ماوراء التحليل، ويمكن التواصل مع الخبراء من خلال الايميل او البريد المدون في الدراسات المنشورة لهم. وللمزيد عن اليات البحث في التراث البحثي والحصول على الدراسات السابقة يفضل الرجوع الى (Cooper 1998).

مصدقية البحث: هل اجراءات البحث كافية؟

بغض النظر عن طرق البحث المختلفة في التراث التي اعتمد عليها الباحث، فالسؤال الذي من الضروري التأكد منه هل اجراءات البحث في التراث مناسبة وكافية؟. يمكن التأكد من مناسبة مدى كفاية البحث في التراث من خلال الطرق الاتية (Card 2012):

الاولى: هل عينة الدراسات التي حصلنا عليها ممثلة تمثيلاً جيداً لمجتمع الدراسات ام متحيزة؟، ويمكن التأكد من ذلك من خلال احد اجراءات ما وراء التحليل للتأكد من اشكالية تحيز النشر.

الثانى: هل عينة الدراسات المتحصل عليها كافية لاعطاء قوة احصائية مناسبة لتقييم الفروض التى بصدد التحقق منها او هل كافية لاعطاء مدى ضيق من فترات الثقة لاحجام التأثير؟.

الثالث: هل وجد الباحث كل الدراسات المتاحة فى المجال ام فقدت دراسات كان يجب تضمينها فى ما وراء التحليل؟.

التساولين الاول والثانى يؤثروا على جودة الاستنتاجات الامبريقية لدراسة ما وراء التحليل. اما التساؤل الثالث اقل اهمية فيما يخص الاستنتاج من ما وراء التحليل. وللجابة على هذه التساؤلات هو بذل المزيد من الجهد للحصول على اكبر قدر من الدراسات الممثلة لمجتمع الدراسات كلما امكن.

ويقدم Card (2012) عدة اقتراحات للتأكد من مصداقية عملية البحث:

- اجراء البحث فى التراث البحثى باستخدام عدة طرائق للبحث مثل قواعد البيانات المختلفة والمجلات البحثية والتواصل مع الباحثين و المؤتمرات غيرها.
- عمل قائمة من الدراسات المتحصل عليها من استراتيجيات البحث المختلفة وعرضها على بعض الباحثين المتخصصين فى موضوع الدراسة لفحصها لمعرفة مدى مناسبتها فى ضوء معايير التضمن والاستبعاد، ولو اقرروا بمدى كفايتها ابدأ فى اجراء ما وراء التحليل ولو اشاورا الى غياب بعض الدراسات فيجب اعادة استراتيجيات البحث مرة اخرى.
- خذ قائمة الدراسات التى استبقيت عليها وابدأ استراتيجيات البحث الخلفى والامامى ومن الافضل البدء بالبحث الخلفى لانه يستغرق وقتاً اقل، ويمكن الحصول على بعض الدراسات الاضافية، واذا لم تحصل على المزيد من الدراسات فابدأ باجراء استراتيجية البحث الامامى فى الدراسات التى ابقيت بعد تحقيقها معايير التضمن والاستبعاد على ان تبدأ بالدراسات القديمة اولاً الى ان تصل الى نقطة عندها لاتمدنا هذه الاستراتيجية بدراسات جديدة عندئذ اوقف اجراء هذه الاستراتيجية.

بعد ذلك ابدأ فى ترميز الدراسات ولكن يمكن اجراء استراتيجية البحث الخلفى اثناء القراءة المتعمقة فى الدراسات وتساعد فى الحصول على المزيد من الدراسات الاضافية واذا شعرت

انك لا تحصل على مزيد من الدراسات بنهاية عملية التكويد فانه يعطى درجة من الثقة ان عملية البحث كانت تقريباً شاملة وممثلة للتراث. وعلى الرغم من هذه الثقة فى شمولية عملية البحث فى الادبيات البحثية فان (2012) Card ينصح بخطوتين اضافيتين:

- اجلس فى المكتبة او على الانترنت وحاول ان تطلع على جدول محتويات المجالات الاكثر ارتباطاً بالموضوع فى السنوات الخمسة الاخيرة مثلاً واذا لم تحصل على مزيداً من الدراسات فان هذا يزيد الثقة فى شمولية ومصداقية عملية البحث واذا وجدت المزيد من الدراسات فانك بحاجة للمزيد من البحث للحصول على الدراسات الغائبة والمناسبة لطبيعة الموضوع ويمكن استخدام كلمات مفاتيحية او مصطلحات مختلفة غير التى استخدمتها من قبل.
- بعد الطمأنينة بان البحث فى التراث كافى ارسل قائمة الدراسات التى حصلت عليها من التراث مرة اخرى الى الخبراء فى التخصص ليقوموا بمدى كفاية ومناسبة قائمة الدراسات لما وراء التحليل وعليه ذلك فالخبراء لهم دور فعال فى عمل البحث.

بناء قاعة بيانات للدراسات المتطلبة لما وراء التحليل

بعد جمع الدراسات السابقة فان الخطوة التالية هى كيفية تنظيم عرض الدراسات، فعملية البحث والحصول على الدراسات السابقة تستغرق وقتاً وجهداً الا ان عملية التنظيم تيسر على الباحث ادراك العلاقات بين هذه الدراسات.

والتنظيم الجيد يتطلب وضع اطار تنظيمى محدد يتضمن مكونات اساسية مثل اولاً: معايير التضمين والاستبعاد التى تستخدمها لتقييم الدراسات المجمعة وفيها تحدد الدراسات المستبعدة فى ضوء كل معيار وعلى ذلك توضح للقارئ مبررات استبعاد دراسة ما مما يوفر الشفافية. ثانياً: يجب ان توجد قائمة محددة لطرائق البحث فى التراث البحثى مثل PsycINFO و ERIC مثلاً حتى تتيح للاخرين اعادة عملية البحث. ويمكن الاستعانة بالدراسات النظرية لانها مفيدة فى تحديد دراسات اضافية وكذلك فى تفسير نتائج ما وراء التحليل.

ويمكن البحث في المحركات البحثية باستخدام الكلمات المفاتيحية خاصة في Google scholar او Yahoo. ولايفضل البعض الاستعانة بها على اساس انها محركات بحث غير اكااديمية.

وعملية البحث تتكون من جزئين:

- ايجاد المراجع المنشورة التي تلبى معايير التضمين.
- الحصول على نسخ من هذه الدراسات لقراءتها وتكويدها في ضوء محكات التضمين او الاستبعاد.

ما وراء التحليل هي دراسات مسحية ملاحظة قائمة على النتائج المتاحة من الدراسات السابقة، ومن الضروري كيفية تحديد الدراسات المتضمنة في التحليل وفي هذا الجانب تظهر مشكلة المعاينة ويحاول الباحث الحصول على عينة من الدراسات ممثلة تمثيلاً جيداً لمجتمع الدراسات. ويبدأ الباحث في جمع الدراسات المرتبطة بموضوع التدخل او البرنامج للمتغير التابع المراد اجراء ماوراء التحليل عليه، وفي الشأن لايد من تحديد المعالجات موضع الاهتمام، المجتمعات، النواتج، التصميمات البحثية المناسبة، وغيرها.

ويبدأ الباحث في الحصول الدراسات المتاحة التي تتضمن على العينات المتاحة من المعالجات وعينات من بعض المجتمعات وبعض التصميمات بينما توجد دراسات غير متاحة مازالت في النشر او صعب الحصول عليها نتيجة التكلفة او في ثقافة ولغة اخرى غير مألوفة للباحث، وهذا يشابه البيانات المفقودة في البحوث الاولية حيث يحدث ان ينسحب افراد من التجربة او التطبيق او يرفض بعض المستجيبين الاجابة عن مفردات الاستبيان.

والمبادئ الاساسية للمعاينة تشير الى ان العينات المتاحة Convenience samples من غير المحتمل ان تكون ممثلة تمثيلاً جيداً لمجتمع الدراسات الكبيرة، فالبحث في قاعدة البيانات PsycINFO ليس ضماناً للحصول على قدر كافي من الدراسات المتطلبة لماوراء التحليل، وتزداد هذه المشكلة في واقع البحث في البيئة العربية لعدم توافر قواعد البيانات الكافية المتضمنة للتراث البحثي خاصة في العلوم السلوكية والاجتماعية. وينصح (Littell et al. 2008).

بالوصول على عينة ممثلة لمجتمع الدراسات وتجنب اشكالية تحيز النشر او اى مصادر تحيز اخرى مرتبطة.

ولامانع من الحصول على عينة طبقية من الدراسات للموضوع حيث تفيد فى دراسات مجموعات فرعية، فمثلاً اراد باحث دراسة فعالية برنامج تدخلى لتخفيف العسر القرائى فلا مانع من جمع دراسات لبرامج فى المرحلة الابتدائية والمرحلة الاعدادية.

ولتنظيم جمع الدراسات السابقة المرتبطة يفضل بعض الباحثون وضع خرائط منظمة Systematic maps للتراث البحثى لاعطاء نظرة لعدد وانواع الدراسات التى تناولت دراسة الموضوع، وتساعد هذه الخرائط او الاطر فى تحديد وتصنيف الدراسات فى ضوء مجتمعها، الموقع، تصميم البحث وغيرها من مظاهر البحث المختلفة.

الفصل الرابع

تكويد خصائص الدراسات

Coding Study Characteristics

مقدمة

اجراء ما وراء التحليل يتطلب تقدير حجم التأثير ثم اجراء احصاء استدلالى حول متوسط حجم التأثير وهذا يتطلب اعداد بروتكول تكويد Coding Protocol الذى يحدد المعلومات المطلوب استخلاصها من الدراسات.

وعملية تكويد البيانات من الدراسات الاولية غالباً تكون عملية صعبة ومتعبة وتحتاج وقت كبير وهى من العمليات الحرجة والبناءة فى دراسة ما وراء التحليل (Hunter & Schmidt, 2004)، وعن طريق عملية التكويد يستخلص المكود بيانات ومعلومات وخصائص عن الدراسات يمكن ان تفيد فى اجراء التحليل المتفاعل Moderator analysis بمعنى اجراء تحليلات فى ضوء خصائص الدراسة مثل تحليل للذكور وتحليل للاناث وغيرها من المتغيرات التصنيفية وهذا يفيد فى التنبؤ بالتباين او الاختلاف فى احجام التأثير عبر الدراسات المختلفة، ومن خلال تكويد خصائص الدراسات فمن المهم وصف مكونات الدراسة لمعرفة مدى مناسبتها لاجراء ما وراء التحليل وكذلك معرفة الاستنتاج العام الذى توصلت اليه عينة الدراسات وكذلك تحديد مدى وجود فجوات فى التراث البحثى للظاهرة مثل استخدام نوعية من المقاييس دون غيرها او التركيز على نوع معين من العينات دون غيرها، وعليه فان تكويد خصائص الدراسات يفيد فى اجراء التحليل المتفاعل او وصف عينة الدراسات، والمكود يقرأ تقرير الدراسة ثم يملئ بروتكول التكويد بالمعلومات المناسبة من الدراسة.

وتكويد دراسات ما وراء التحليل مثل تكويد الدراسات المسحية فمن المهم اعداد استبيان التفريغ بعناية وتدريب المكودين على ملئ الاستبيان وتقدير صدق وثبات البيانات. وفى هذا الجزء التالى نعرض لكيفية اعداد بروتكول التكويد والاجراءات المستخدمة لاستخدامه لتكويد الدراسات وتدريب المكودين.

بناء بروتكول التكويد Developing A coding Protocol

فى ضوء دراسات ما وراء التحليل لا توجد صيغة موحدة لعملية تكويد الدراسات بل تختلف من باحث الى اخر وحسب اهداف الدراسة ويتضمن بروتكول التكويد خصائص الدراسات المراد اجراء ما وراء التحليل لها وهذا يتوقف عل طبيعة المعلومات فى الموضوع او الظاهرة المراد دراستها.

وبرتكول التكويد يهدف الى توثيق الاجراءات المستخدمة لاستخلاص المعلومات من الدراسات واعطاء توجيه لكيفية ترجمة المعلومات الكيفية السردية لتقرير الباحث الى صيغة بنائية موضوعية كمية (Wilson, 2009).

ويحدد (Lipsey & Wilson, 2001) الهدف من التكويد هو بناء قاعدة بيانات للتحليل الاحصائى ومن الافضل استخدام مفردات ذات النهاية المغلقة كلما امكن ويحاول الباحث ان يحدد مسبقاً الاختيارات المحتملة لكل مفردة فى استبيان التكويد حيث تسهل على المكود ملئ مفردات الاستبيان او القائمة مثل مفردة حجم العينة، ولكن توجد مواقف لايسطيع الباحث تحديد بدائل مفردات الاستبيان وعليه فهو مضطر الى بناء مفردات ذات نهاية مفتوحة وفى هذه الحالة يلجأ الباحث الى مجموعة من التصنيفات الفرعية او الابعاد ولكن هذا يستغرق وقت وجهد من الباحث او المكود. كلما كان المكود مدرباً كلما زادت قدرته على تجميع المفردات فى اطر او تصنيفات متقاربة.

ويفضل ان يكون بروتكول التكويد بسيط وقصير كلما امكن ومنظم فى اجزاء واضحة، ومن المهم ايضاً اعطاء تعريف كامل وواضح لكل مفردة وارشادات لكيفية تكويدها لمساعدة المكود للتحكم فى الحالات الغامضة وغير المعتادة. وفى الممارسة العملية بناء بروتكول تكويد يدوى Coding manual حيث يترجم ماهى المعلومات المراد الحصول عليها من كل مفردة وكذلك اعطاء توصيف للبدائل المختلف للاجابة واعطاء ارشادات للحالات التى بها غموض. وكتاب التكويد اليدوى يكون بمثابة الدليل المرشد للمكود عندما يواجهه صعوبات.

ويحدد (Card, 2012) اعتبارين يمكن تطبيقهما لمراجعات ما وراء التحليل هما:

- اسئلة البحث موضع الاهتمام: مثل اجراء الدراسات الاولية من تحديد المتغيرات فان التخطيط لما وراء التحليل يتطلب تحديد السؤال البحثى الذى على اساس يتم تكويد خصائص الدراسة او توصيفات الدراسة Study descriptors مثل (الطريقة او المنهج، القياسات، العينات، المتغيرات المستقلة والتابعة، نوعية المعالجة، الاجراءات وغيرها)، فلو ان سؤال البحث منصّباً على متوسط احجام التأثير عبر الدراسات (دمج الدراسات) فالباحث ليس بحاجة سوى الى معلومات مثل حجم التأثير وحجم العينة او اى بيانات تستخدم فى حساب حجم التأثير، وفى هذه الحالة لانحتاج الا الى عدد محدود من خصائص الدراسات. وفى المقابل اذا كان هدف الباحث مقارنة الدراسات لتحديد ما اذا كانت دراسات بملاح معينة تعطى احجام تأثيرات اكبر من احجام تأثيرات لدراسات بملاح مختلفة وبالتالي فالباحث بحاجة الى معلومات كثيرة لخصائص الدراسات لتكويدها.

ويوجد نوعية من الاسئلة البحثية تتضمن متغيرات متفاعلة، فبعض محلى ما وراء التحليل يرغبون فى التنبؤ بعدم التجانس فى احجام التأثير من خصائص الدراسات المكودة، وعليه فان الباحث بحاجة الى تكويد عدد كبير من الدراسات خاصة التى يتوافر لها معيارين الاول: خصائص الدراسات تكون متماثلة فى معظمها وتجنب غياب البيانات خاصة للمتغير المراد التنبؤ به، والثانى: تنوع خصائص الدراسات خاصة عبر على الاقل بعض الدراسات وهذا التنوع ضرورى عبر الدراسات لمعرفة مدى التلازم او التباين لاجسام التأثيرات، وعليه يتم ادخال خصائص الدراسات المكودة فى نموذج انحدارى تنبؤى لاكتشاف العلاقات بين هذه الخصائص والتباين فى احجام التأثيرات.

- الاعتبار الخاص بمظاهر الدراسة: خصائص الدراسة تخبرنا بتفاصيل عن مجال البحث، وعموماً يجب ان يوضع فى الاعتبار اربعة انواع اساسية لاجراء ما وراء التحليل فى العلوم الاجتماعية: خصائص العينة، عملية القياس، التصميم البحثى، والمصدر. ويمكن عرضها كالاتى (Card, 2012; Valentine, 2009):

المظهر العام	المظاهر الخاصة	امثلة
العينة	اجراءات المعاينة	نوعية المجتمع، العينة الممثلة، العينة العشوائية، البلد.....الخ.
خصائص القياس	مصدر البيانات	الملاحح الديموجرافية الاجتماعية، الذكاء، الموقع...الخ تقرير ذاتي(الاستبيان - المقاييس)، ملاحظة، مقابلة وغيرها.
عملية القياس	قياسات معينة	ملاحظة منظمة ام غير منظمة، تطبيق فردي ام جماعي، مقيد بوقت اوغير مقيد الصورة كاملة ام مختصرة، مقياس مترجم ام من البيئة او الثقافة.
خصائص التصميم	نوع التصميم البحثي	المجموعة الواحدة وقياس قبلي وبعدي، تصميم المجموعتين قياس قبلي وبعدي.
خصائص مميزة للتصميم	نوع المعالجة التي تلقتها المجموعة	الضابطة، الفترة الزمنية بين القياسات الطولية.
المصدر	النشر	دراسة منشورة او غيرمنشورة، جودة النشر
عام النشر	التمويل	السنة، تاريخ جمع البيانات. ممولة ام غير ممولة، جهة التمويل.
خصائص الباحثين	جودة الدراسة	الجامعة، الجنس، الجنسية.
الصدق الداخلي	الصدق الخارجي	معاينة عشوائية، ضبط المتغيرات الدخيلة، التزييف، الانسحاب.
		العينات من المجتمع الممثلة له، التجريب

يناظر الواقع.

الصدق البنائي ثبات القياسات (معايير التضمين) ،
خصائص القياسات.

وفيما يلي عرض لاهم خصائص الدراسات (Card, 2012; Littell et al. 2008; :Lipsey& Wilson, 2002; Wilson, 2009)

أولاً: كود الدراسة: يتم وضع كود الدراسة رقمياً او حرفياً واحياناً يتم تحديد مؤلف الدراسة وسنة النشر.

ثانياً: المفاهيم Constructs: دائماً تركز الدراسات التجريبية على متغير تابع واحد فهي تسعى الى تنمية مهارات التواصل او جودة العلاقات الاجتماعية او تقدير الذات او مفهوم الذات وغيرها من المتغيرات الناتجة، اما الدراسات الارتباطية تركز على العلاقة بين متغيرين مثل العلاقة بين سلوك الجريمة والحالة الاقتصادية الاجتماعية او العلاقة بين ماوراء التعلم والتحصيل، وعليه يقرر محلل ماوراء التحليل ماهى المتغيرات موضع اهتمام دراسة لتكويدها. فمثلاً ماوراء التحليل لدراسات برامج القراءة العلاجية فهي تنصب على متغير ناتج وحيد هو القدرة القرائية على الرغم لو ان بعض الدراسات تبنت مفاهيم اخرى مرادفة مثل الاتجاه نحو القراءة يمكن تضمينها.

ثالثاً: خصائص العينة Sample: تتضمن هذه الخصائص اجراءات المعاينة والخصائص الديموجرافية، فاجراءات المعاينة المفضل تكويدها تتضمن: **المجتمع** الذى اشتقت منه سواء طلاب جامعة، مرحلة ثانوية، مدمنين المخدرات، نزلاء المستشفيات النفسية، نزلاء السجون بقضايا الارهاب والتطرف وغيرها، **نوعية العينة** سواء عشوائية بسيطة ام طبقية، عينة متاحة لموقع عمل الباحث، مقصودة عمدية، **بلد العينة** مثل مصر، السعودية، المغرب... وغيرها. **والخصائص الديموجرافية** مثل الجنس، مكان المعيشة (ريف - بدو - حضر)، الديانة، الحالة الاقتصادية الاجتماعية، العمر، وادى توصيف اخر للعينة مثل الذكاء. و**حجم العينة** يعتبر من اهم البيانات فى دراسة ما وراء التحليل حيث يساعد فى حساب الخطأ المعيارى

وكذلك تحديد العينات الفرعية كلما امكن. والباحث ليس بصدد تكويد كل خصائص العينة مالم يكن لها دور في ما وراء التحليل.

وتقدر احجام التأثير ماوراء التحليل من علاقات او تاثيرات محسوبة من قياسات مطبقة على عينة من المستجيبين او المشاركين واحياناً يتم عرض النتائج في ضوء مجموعات فرعية فمثلا تعرض العلاقة بين التحصيل والتنشئة الاسرية لعينة الريف ولعينة المدينة على حدة اذا كان هذا موضع اهتمام دراسة ماوراء التحليل حيث تقدر احجام التأثير للذكور وللاناث كلاً على حدة مثلاً لدراسة الفروق بينهما لتحديد مدى استجابة المجموعتين للمعالجة. وفي اعداد بروتكول التكويد من الضروري تحديد ما اذا كانت احجام التأثير تقدر للعينة الاجمالية ام تقدر لعينات فرعية.

رابعاً: خصائص القياس Measurement: من الضروري التميز بين المفاهيم الاجرائية او القياسات Operational definitions فى الدراسة عن المفاهيم النظرية Theoretical contrasts لانه توجد مقاييس مختلفة تقيس نفس المفهوم، فمثلاً فى دراسة لتقدير المشاكل السلوكية للاطفال يوجد قياسات مختلفة مثل تقديرات الاخصائى النفسى وتقديرات الوالدين والتقديرات الذاتية.

فى العلوم الاجتماعية والنفسية توجد مداخل عديدة لعملية القياس وقياسات عديدة متخصصة للمتغيرات المتضمنة فى ما وراء التحليل. ولذلك من الضرورى تكويد خصائص القياسات للمتغير او المتغيرين فى ما وراء التحليل. ومن المظاهر الهامة التى من الضرورى تكويدها هو مصدر البيانات المقاسة مثل تقرير ذاتى، تقرير من مصدر اخر مثل الوالدين او المعلم او ملاحظة الباحث او غيرها، ملامح عملية القياس مثل ملاحظة مفتوحة ام مغلقة، مقابلة فردية ام جماعية، اختبار سرعة محدد بوقت ام اختبار قوة غير محدد بوقت. ويمكن فى بعض الحالات تثبيت المقياس المستخدم فى الدراسة كاجراء ما راء التحليل لاثر العلاج المعرفى على الاكتئاب كما قيست بقائمة بيك Beck depression inventory ويمكن استبعاد الدراسات التى اعتمدت على مقاييس ذات جودة منخفضة.

وتوجد اشكالية فيما يخص المواقف التي تستخدم اكثر من مقياس لقياس نفس المفهوم في نفس الدراسة ولكن المدخل الافضل هو ترميز كل مقياس على حدة ولكن هذا يؤدي الى عدم تماثل عبر الدراسات، والافضل الاعتماد على الدراسات التي اعتمدت على المقياس الاكثر استخداماً، واحياناً يكود محلل ما وراء التحليل نتائج الدراسات التي اعتمدت على مقاييس متعددة لنفس المفهوم ثم يقدر متوسط حجم التأثير.

ومن المهم تضمين الخصائص السيكومترية مثل معاملات الصدق والثبات كمعايير للتضمين فاذا كان معامل الثبات 0.8 فيكتب في ملف الادخال بنفس القيمة ويفضل تحديد نوعية طريقة حساب الثبات ونوعية معامل الثبات هل الفا كرونباخ ام معامل ارتباط بيرسون ومن الطبيعي ان يعتمد الباحث على دراسات عالية الجودة لا يقل عن 0.70.

خامساً: وقت القياس Time of measurement: يمكن قياس المفهوم على نفس المجموعة او مجموعات فرعية من الافراد في اكثر من وقت. ففي دراسات المعالجات يوجد قياسات قبلية ثم بعدية ثم قياس تتبعي. في الدراسات الطولية تجمع البيانات من نفس الافراد على نفس المقياس في فترات زمنية متتالية، وعليه يقرر الباحث ما اذا كان يقدر ويكود احجام تأثير في كل اوقات القياسات (قبلي- بعدى- تتبعي) ولكن في الدراسات التجريبية يكود الباحث حجم التأثير للنتائج الناتج في القياس البعدى.

سادساً: خصائص التصميم Design Characteristics: لابد من تكويد خصائص التصميم مثل طبيعة التصميم المستخدم في الدراسة سواء تصميم المجموعة الواحدة وقياس قبلي وبعدي او تصميم مقارنة المجموعات شبة التجريبية او تصميمات التجريبية العشوائية (الحقيقية)، وايضاً تكويد طبيعة المجموعة الضابطة كونها تلقت معالجة تقليدية او لم تتلقى معالجة على الاطلاق. او الدراسات الارتباطية المحددة بالتصميمات عبر عرضية Cross-sectional designs.

سابعاً: خصائص المصدر Source characteristics: توجد بعض المعلومات الخاصة بمصدر النشر مثل مسمى المجلة او طبيعة النشر دراسة منشورة او غير منشورة وذلك لتقييم تحيز النشر، وسنة النشر في المجلة او المؤتمر او بحث دكتوراة وذلك لدراسة التحليل

المتفاعل لمعرفة اتجاهات احجام التأثير عبر فترات زمنية مختلفة، ومصدر تمويل الدراسة لان هذا من اكبر العوامل التى تؤدى احياناً الى تحيز النتائج، وخصائص الباحث مثل الجنس او الثقافة او المؤسسة التى ينتمى اليها لدراسة تحيز الباحثين بمعنى مدى اختلاف احجام التأثير للدراسات التى اجراها الذكور عن الدراسات التى اجريت عن طريق الاناث.

تكويد حجم التأثير Effect Size Coding

اي ناتج احصائى فى اى دراسة منتقاة فى ما وراء التحليل يجب ان تعرض فى ضوء قيمة حجم التأثير المصاحب، وتوجد العديد من المعادلات لحساب حجم التأثير وتتنوع حسب الاختبار الاحصائى المستخدم والتكويد لحجم التأثير تتضمن حساب قيمته لكل ناتج بحثى او ناتج الاختبار الاحصائى للمتغير موضع الاهتمام باستخدام صيغ ومعادلات احصائية. وتوجد العديد من البرامج الاحصائية الحاسوبية المستخدمة لتقدير حجم التأثير مثل SPSS , STATA, SAS, G-Power وغيرها من البرامج. وطور (Shadish, Robinson, & Lu (1999) برنامج حاسوبى Effect Size Calculator ومتاح فى الموقع (<http://www.assess.com/>). والمعلومات المراد تكويدها المرتبطة بحجم التأثير تقع فى عدة تصنيفات متداخلة هي:

- المعلومات التى تصف **طبيعة المتغيرات** المراد تقدير حجم التأثير لها مثل المفهوم المراد قياسه وتعريفه الاجرائى والطبيعة الاحصائية للمتغير سواء تصنيفى او رتب او متصل.
- **توقيت قياس المتغير** مثل قياسات قبلية وبعديّة او قياسات تتبعية.
- حساب حجم التأثير **للعينة الاجمالية** ام **لعينات فرعية** مثل حسابه لعينة الاناث فقط ولايمكن تضمين حجم التأثير للعينة الاجمالية وللعينات الفرعية فى تحليل واحد لانهما معتمدين احصائياً ولكن يمكن تضمين احجام الاثر للذكور وللاناث فى تحليل واحد لانهما مستقلين.
- عملية التكويد تتطلب **بيانات احصائية** عن حجم الاثر مثل المتوسطات، النسب، احجام العينة، التباينات، الانحرافات المعيارية، الارتباطات القبلية والبعديّة وقيمة P الاحتمالية... الخ. وكذلك يتضمن بروتوكول التكويد معلومات احصائية مطلوبة لعمل تصحيحات او تعديلات لحجم الاثر من عدم الثبات، التصنيفية، والمدى المقيد

للمتغيرات. وتوجد معلومات احصائية اضافية مثل الدلالة الاحصائية للنتائج، الاختبار الاحصائي المستخدم، طبيعة البيانات الغائبة وفيما يلي ملخص للمعلومات الاحصائية (Lipsey & Wilson, 2001):

(المتغير / المفهوم)، (التوقيت)، (العينات الفرعية)، (حجم العينة)، (المتوسطات او النسب)، (الانحرافات المعيارية او التباينات)، (اجراء حساب حجم التأثير)، (ثبات المتغيرات)، (فترات الثقة لحجم التأثير)، (تصنيفية المتغيرات)، (المدى الضيق للمتغيرات)، (الاختبار الاحصائي المستخدم ودلالته الاحصائية)، (رقم الصفحة فى تقرير البحث المقدر فيها حجم التأثير)

تأكيد جودة الدراسة Coding Study quality

ينصح البعض ان يتضمن ماوراء التحليل تأكيد جودة الدراسة ويمكن ادخال عامل الجودة كمتغير فى التحليل باستخدام التحليل المتفاعل لاحجام الاثر ويرى (Card 2012) ان هذا التضمن لعامل الجودة يمثل اشكالية وذلك لانه ذات طبيعة متعددة الابعاد وليس ذات بعد واحد كما يعتقد البعض لان الجودة تتضمن مظاهر عديدة خاصة فى ضوء صدق البحث سواء داخلى او خارجى او بنائى Construct او صدق الاستنتاج الاحصائى Statistical Conclusion Validity وداخل كل نوع من انواع الصدق الاربعة توجد مظاهر عديدة Shadish, Cook, (2002) & Campbell:

- **الصدق الداخلى Internal validity**: يشير الى اى درجة ان تصميم الدراسة او المتغير المستقل يسمح باستنتاج السببية، بمعنى ان التغير الحادث فى المتغير التابع يرجع ويرجع فقط الى المعالجة وليس الى متغيرات اخرى دخيلة. يعتبر تصميم الدراسة من اكثر الامور تأثيراً فى الصدق الداخلى فالدراسات التجريبية (المعاينة العشوائية) اكثر صدقاً داخلياً من الدراسات شبه التجريبية (لا عشوائية).

- **الصدق الخارجى External validity**: يشير الى اى درجة نواتج الدراسة قابلة للتعميم لمجتمع الدراسة وكذلك لعينات مختلفة لنفس المجتمع وشروط اخرى غير الشروط التى اجريت فيها التجربة ولطرق قياس اخرى لنفس المفهوم موضع الدراسة. والدراسات الاكثر صدقاً خارجياً

هى التى اعتمدت على عينات عشوائية من مجتمع الدراسة وهذا الشرط يكاد يكون نادراً فى الدراسات النفسية والاجتماعية. من الضرورى فى دراسات ماوراء التحليل تكويد ماذا كانت الدراسات المجمعـة حققت هذا المعيار من عدمة، ولكن اجراء ما وراء التحليل لدراسات مفردة محدودية الصدق الخارجى يـؤدى الى استنتاجات عالية الصدق الخارجى وهذا من اهم مميزات ماوراء التحليل.

- **صدق المفهوم Construct validity**: يشير الى اى درجة المقاييس المستخدمة فى الدراسة تقيس البناء النظرى الذى ينوى الباحث قياسه. ولكن هذا المصطلح يشير فى هذا الصدد الى خصائص القياس وهى الصدق والثبات للقياسات ولذلك يجب تكويد ثبات المقاييس للقياسات المتضمنة للمتغيرين فى التحليل لاجراء تصحيحات احجام الاثر، وكذلك يمكن تضمين قياسات الصدق لان من معايير التضمين ان يكون التعريف الاجرائى للمفهوم واضح ويمكن تصحيح الصدق غير الدقيق من خلال اساليب ماوراء التحليل، خاصة حينما يقدر من خلال معامل الارتباط(معامل الصدق) بين المقياس المستخدم فى الدراسة ومقياس معيارى متفق عليه لنفس المفهوم ويطلق عليه الصدق المحكى.

- **صدق الاستنتاج الاحصائى**: يشير الى صدق الاستنتاجات الاحصائية فيما يخص العلاقة بين المسبب المفترض والتاثير او الناتج من التجربة.

وجودة الدراسة بناء تحتى ومتعدد المظاهر من تصميم الدراسة وانتقاء العينة وطبيعة المعالجة الاحصائية ويمكن اخذها فى الاعتبار كمتغير وسيط فى ما وراء التحليل.

ويقترح Card (2012) توصيف المظاهر المختلفة لجودة الدراسة المهمة فى المجال اوال تخصص واعتبارها متغيرات متفاعلة لاحجام الاثر بين الدراسات.

ماذا عن الدراسات ضعيفة المنهجية؟.

كثير من الخبراء يستبعد الدراسات ضعيفة المنهجية من دراسة ما وراء التحليل ولكن Hunter & Schmidt (2004) يرى ان هذا غير مناسب ولا يفضل استبعادها لان عدم المناسبة المنهجية Methodological inadequacies لا يمكن اختبارها امبيريقياً ولا توجد دراسة بدون محددات وقصورات منهجية على الرغم ان هذه المحددات لاتسبب دائماً تحيز فى النتائج.

الاختلافات المنهجية بين الدراسات تسبب اختلاف النتائج عبر الدراسات نتيجة على الأقل لخطأ المعاينة في الدراسات ويمكن اختبار عدم المناسبة المنهجية من خلال اعتبارها متغير متفاعل ويمكن دراسة اثرها في دراسة ما وراء التحليل واذا لم توجد فروق بين الدراسات عالية المنهجية والدراسات ضعيفة المنهجية على ذلك لا توجد اشكالية عدم المناسبة المنهجية، وعليه يجب تضمين الدراسات الضعيفة المنهجية والجيدة المنهجية ثم دراسة الفروق بينهما ويكون القرار في النهاية في ضوء الدراسات عالية الجودة واذا لم يوجد فروق بينهما (الضعيفة - الجيدة) يكون التحليل على كل الدراسات، وعليه فلا بد من تقييم المناسبة المنهجية امبريقياً ولا يتم الحكم عليها في ضوء التقييمات الذاتية الشخصية حيث توصل (Cooper 1998) ان معامل الاتفاق بين المقدرين للحكم على المناسبة المنهجية على افضل الاحوال 0.50.

على ذلك لابد من تضمين سؤال هام في دراسة ما وراء التحليل لدراسة الفروق بين الدراسات المناسبة والدراسات غير المناسبة لتحديد مدى تضمين الدراسات الضعيفة منهجياً في دراسة ما وراء التحليل.

توصيفات الدراسة Study Descriptors

كل الخصائص المكونة المتحصل عليها من الدراسة يطلق عليها مواصفات الدراسة وطرح (Lipsey & Wilson 2001) اطار شامل لمفردات بروتكول التكويد، ولكن من الافضل لبناء هذا البروتكول اجراء مسح لعينة من الدراسات المجمعمة لتحديد طبيعة المعلومات المراد تكويدها، وهذا الاطار يتضمن مجموعة من التصنيفات لتحديد مواصفات الدراسة كالاتي:

- **المظاهر الاساسية للدراسة Substantive aspects**: مثل طبيعة المشاركين في العينة، المعالجات المستخدمة، والواقع الثقافي. وهذه المتغيرات تسمح باختبار ما اذا كان يوجد اختلاف في احجام الاثر عبر المعالجات المختلفة او عبر الثقافات المختلفة او عبر عينات مختلفة.
- **الطريقة والاجراءات Method and procedures**: وهذا النوعية من المعلومات المنهجية تسهم في دراسة العلاقة بين التنوع المنهجي ونتائج الدراسة حيث يكشف عن التحيز في النتائج في ضوء الممارسات المنهجية المختلفة، وفي اغلب الاحوال الاختلاف في نتائج الدراسات يرجع

الى الفروق المنهجية بين الدراسات، وعليه اذا لم يتم تكويد المتغيرات المنهجية فان الباحث يصل الى استنتاجات غير صحيحة.

- توصيفات المصدر Source descriptors: ومتغيرات المصدر تتعلق بالاطار العام للدراسة، عملية النشر، خصائص الباحث او المؤلفين، تاريخ النشر، وغيرها. هذا يفيد مثلاً في ما اذا يوجد اختلاف لاحجام الاثر عبر التخصصات المختلفة (علم النفس في مقابل الصحة النفسية). ويمكن تلخيص توصيفات الدراسة كالاتي:

القضايا الاساسية

مصدر العينة

- الخصائص الديموجرافية (المستوى الاقتصادي الجماعي، العمر، الجنس، المستوى التعليمي، مكان المعيشة وغيرها).
- الخصائص الشخصية (القدرات المعرفية، السمات الشخصية).
- الخصائص التشخيصية او ملامح خاصة (نزول سجن، صعوبة تعلم، مصاب بمرض معين، معاق سمعياً وغيرها).
- المتغير المستقل (التدخل او المعالجة او البرنامج).
- النظرية القائمة عليها الدراسة او البرنامج.
- خصائص المنظمة او المؤسسة (عمرها، حجمها، الهيكل الاداري).
- خصائص القائم بتطبيق البرنامج ومؤهلاته.

الطريقة والاجراءات

- اجراءات المعاينة (عشوائية - عمدية).
- تصميم المسح (البريد، التليفون، المقابلة، دراسة طولية، دراسة عبر عرضية، دراسة مستقبلية، دراسة ارشيفية، دراسة تاريخية).
- التآكل او الانسحاب للعينة من التجربة.
- مهددات الصدق الداخلي (العوامل الدخيلة في التصميم).

- القوة الاحصائية.
- جودة القياسات(الصدق والثبات).
- كيفية تحليل البيانات.
- طبيعة المجموعة الضابطة (لم تتلقى معالجة او تلقت معالجة بديلة).
- طريقة توزيع افراد العينة على مجموعات التصميم.
- دور المجرب او الباحث فى المعالجة او البرنامج.

توصيفات المصدر

شكل النشر(مجلة، كتاب، مجلد، رسالة دكتوراة، تقرير حكومى او مؤسسى)، تاريخ النشر، بلد النشر واللغة، تمويل الدراسة والجهة المانحة، خصائص الباحث (الجنس، التخصص، والدرجة العلمية.

ومن الضرورى تكويد هذه المتغيرات وتختلف عملية التكويد حسب الهدف من دراسة ماوراء التحليل وطبيعة الدراسات المتضمنة فى التحليل.

تقويم عملية التكويد Evaluatig Coding Process

بعد الانتهاء من تحديد خصائص الدراسة المراد تكويدها فان الخطوة التالية اقرأ بعناية الدراسات لتكويد المعلومات منها وبطبيعة الحال المعلومات المكودة فى ضوء خصائص الدراسة وحجم الاثر ولا بد من التأكد من عملية الحكم التى يطبقها المكود عندما يتفاعل مع الدراسة لان هذه العملية تختلف من مكود الى مكود اخر، ومن دراسة الى دراسة اخرى، وهذا بالطبع يؤثر فى تحليل بيانات ماوراء التحليل ويوجد مبدأين اوقضيتين على درجة كبيرة من الاهمية مرتبطين ببروتكول التكويد يحددهم (Card, 2012; Lipsey & Wilson, 2002) بالاتى:

- الشفافية والاعادة للتكويد **Transparency & Replicability of Coding**: عند

كتابة التقرير النهائى لدراسة ماوراء التحليل يجب اعطاء تفاصيل كافية عن عملية التكويد حتى يعرف القارئ كيفية عمل قرارات التكويد(الشفافية)، وكذلك اذا ما طبق باحث اخر نفس قواعد التضمين هل يصل الى نفس الاستنتاج او القرارات التكويدية؟. ولتحقيق هذين المبدأين فمن الضرورى اعطاء وصف كامل لكيفية تكويد كل دراسة كميأ،

فبعض خصائص الدراسة تكود بطريقة مباشرة ولا تتطلب جهد من المكود ودائماً لا تختلف من مكود الى اخر مثل متوسط عمر العينة، ولكن عندما تكون خصائص الدراسة اقل وضوحاً حيث تتطلب احكام تكويد عالية الاستدلال فمن الضروري كتابة تقرير كامل عن عملية التكويد للتأكد من الشفافية والاعادة وهذا بدوره يسهم فى تقدير ثبات المكودين، فعلى سبيل المثال لتكويد طبيعة القياسات المستخدمة وتصميمات الدراسات فيجب كتابة قيم مختلفة لهذا الكود التصنيفى.

- **ثبات التكويد Reliability of coding**: لتقويم عملية الاعادة لنظام التكويد فمن الضرورى تقدير ثبات المكود الذى بدوره يأخذ بعدين هما:
 - **اتساقه عبر ذاته Intra-coder reliability**: اى اتساق نفس المكود من مناسبة الى اخرى بمعنى ان يكود الباحث مجموعة فرعية من الدراسات مرة تلو الاخرى بوجود فترة زمنية بينهما يمكن ان تكون كحد ادنى خمسة عشر يوماً او باقل تقدير اسبوع ويقارن بين النتائج ولكن الباحث يكون على دراية بتكويد الدراسات فى المرة الاولى وعليه يحدث تضخم لقيمة الثبات واحد الاسباب لتقدير ثبات المكود عبر الزمن التحقق من معرفة مدى وجود تغير دراماتيكى اثناء عملية التكويد عبر الزمن. وفى دراسات ما وراء التحليل الصغيرة ربما يوجد مكود واحد لكل الدراسات واجراء اتساق بين المكودين يصبح اقل اهمية ولكن كلاً من النوعين ضروريين لاجراء ماوراء التحليل.
- **الاتساق بين مكودين مختلفين Inter-coder**: حيث يكود باحثان او مكودان نفس العينة من الدراسات المجمعة بصورة مستقلة ويفضل ان تكون عينة الدراسات من 20 الى 50 دراسة للحصول على ثبات دقيق نسبياً ولكن كثير من دراسات ماوراء التحليل لا تتضمن هذا العدد الكبير من الدراسات وفى هذه الحالة يجب استخدام كل الدراسات التى استوفت معايير التضمين والاستبعاد ويجب عدم معرفة بان مكود اخر يقوم بعملية التكويد لان المعرفة تزيد من عملية الحذر واليقظة فى التكويد ويعطى تقدير عالى Overestimate للثبات الحقيقى.
- وباستخدام مدخلى تقدير ثبات المكودين فمن المفيد تقدير الثبات لكل خصائص دراسة تم تكويدها على حدة وايضاً تقدير ثبات كل متغير اوخاصية عبر كل الدراسات. وتوجد العديد من المؤشرات لتقدير ثبات لمكودين من اهمها:

- نسبة الاتفاق (AR) Agreement rate: يقيس نسبة الدراسات لدى المكودين

اونفس المكود عبر مناسبتين قدروا نفس الكود التصنيفي. ويقدر كالاتي:

$$AR = \frac{\text{عدد الاتفاقات}}{\text{عدد الدراسات}}$$

- عدد الاتفاقات : عدد الدراسات التي اخذت نفس الكود للخاصية.

- عدد الدراسات: عدد الدراسات التي كودها المكودين على خاصية الدراسة.

فمثلا في دراسة ما وراء التحليل تضمنت 20 دراسة وتضمن نظام التكويد خاصة نوعية

العينة سواء عشوائية او غير عشوائية وعرضت على مكودان فاتفق تقدير المحكمان على

نوعية العينة لـ 15 دراسة بينما اختلفت الاستجابة لـ 5 دراسات وعليه فان معامل

الاتفاق:

$$AR = \frac{15}{20} \times 100 = 75\%$$

ويرى البعض ان الحد الادنى لقبول الثبات هي 70% . وهو بسيط في حسابه ويعتبر

مؤشراً لثبات التكويد ولكن توجد بعض المحددات هي قضية التحيز فيمكن ان يكون احد

المكودين متحيز لبديل معين وعليه يوجد احتمال لحدوث حكم او تقدير بالصدفة او

الخطأ.

- **المعامل كابا لكوهين Cohen Kappa(K):** هذا المعامل يعتبر مؤشر لتقدير ثبات

المقدين Inter-rater reliability للاكواد التصنيفية(نعم- لا او مناسب- غير

مناسب) ويعالج قضية التحيز التي يعاني منها معامل الاتفاق

ويقدر المعامل Kappa من المعادلة الآتية:

$$K = \frac{\sum fo - \sum fe}{N - \sum fe}$$

حيث: $\sum fo$ مجموع التكرارات الملاحظة في الخلايا القطرية، $\sum fe$ مجموع التكرارات

المتوقعة في الخلايا القطرية، N حجم العينة او عدد الدراسات.

ويتم حسابه من جدول اقتران ثنائي مثل حسابات كا تربيع. وهذا الاحصاء مفيد كمؤشر لثبات

التكويد للخصائص الدراسة والاستجابة على مستويات اسمية تصنيفية. ويرى (Card 2012) اذا

كانت الاستجابة على مقياس ترتيبي تدرج خماسي مثلاً فيمكن استخدام مؤشر كابا الموزون،

والمحدد الرئيسي لهذا المعامل هو يتطلب عدد كبير من الدراسات تتراوح من 40 الى 50 دراسة للحصول على تقدير دقيق للثبات.

- **معامل ارتباط بيرسون (Pearson Correlation(r):** يكون مفيد اذا كانت بعض خصائص الدراسة تكود على مقياس رتبي متعدد التصنيفات. واذا كان التكويد للمفردة عبر كل الدراسات على مقياس رتبي متعدد عبر مناسبتين متتاليتين لنفس المكوود يمكن استخدام احد الاختبارات لتحديد الفروق بين القياسات التكويدية باستخدام اختبار T المرتبطة.

بناء وتنظيم نظام او بروتوكول التكويد

بعد تحديد الخصائص المراد تكويدها، فالخطوة التالية التخطيط لتكويد الدراسات، والمرشد لعمل التخطيط هو بروتوكول التكويد الذى يتضمن وسيط التكويد الشيت يدوى او الكمبيوتر الذى يفرغ فيه المعلومات المكوودة بالاضافة الى كتاب التكويد الذى يتضمن تعليمات عملية التكويد. وفى النهاية نحصل على قاعدة بيانات قابلة للتحليل فى برامج ما وراء التحليل.

التكويد اليدوى Manual Coding

كتاب التكويد اليدوى هو مجموعة مفصلة من التعليمات تصف نوعية المعلومات المطلوب استخلاصها من البحوث والدراسات فى صورة كمية لادخالها البرنامج الكمبيوترى لما وراء التحليل، وهذا الكتاب التكويدى يخدم ثلاثة اهداف الاول يمد المكوود بمرشد لنقل المعلومات او خصائص الدراسة الى وسيط التكويد (الشكل الورقى)

وهذه المعلومات فى كتاب التكويد تستخدم فى المواقف النموذجية كذلك فى المواقف التى تتضمن اشكاليات معينة. **الثانى** يهدف الى توحيد المعلومات المتحصل عليها عبر المكوودين بالتالى اى حدوث اتساق عن طريق تعليمات واضحة وشفافة لما يقوم به كل مكوود. **الثالث** يعتبر مفتاح التكويد اليدوى بمثابة وثيقة لعملية التكويد التى تعتبر كالمُرشد لقاعة بيانات ما وراء التحليل. ويعد مفتاح التكويد اليدوى الاولى قبل تكويد الدراسات واى معلومات يمكن ان تضاف الية اثناء تكويد الدراسات وعليه فهو قابل للتعديل بالحذف او الاضافة.

وفيما يلى مثال التكويد اليدوى:

مسمى الدراسة: تكتب مسماها حسب APA

1- **رقم الدراسة:** يحدد رقم محدد لكل دراسة.

2- **ما نوع النشر للدراسة؟**

-كتاب - مجلة - دكتوراة او ماجستير - مؤتمر - تقري حكومى

3- ما سنة النشر؟.....

توصيفات العينة

4- متوسط عمر العينة:....

5- جنس العينة:

- ذكر - انثى - خليط

6- المرحلة الدراسية او نوعية العينة:

- رياض اطفال - ابتدائى - اعدادى - ثانوى

- جامعى - دراسات عليا - لم يحدد

وتختلف بدائل هذه الفقرة حسب موضوع البحث فاذا كان البحث مهتم بمدمن المخدرات فتكون

تصنيفات العينة كالتالى:

- حشيش - افيون - ترامدول - بانجو - تامول - لم يحدد

7- طريقة انتقاء العينة:

- عشوائية بسيطة - عشوائية طبقية - عمدية - لم يحدد

8- حجم العينة:

9- بيئة العينة:

- ريفية - حضرية - بدوية - ساحلية

توصيفات تصميم البحث

10- نوعية المنهج المستخدم:

- تجريبى - شبة تجريبى - سببى مقارن - ارتباطى - لم يحدد

11- وحدة المعاينة:

- طلاب - فصل دراسى - نزلاء سجون - اطفال شوارع - لم يحدد

12- توزيع افراد العينة على مجموعات التصميم:

- عشوائية بسيطة - عشوائية طبقية - غير عشوائية - لم يحدد

13- هل تحقق من التكافؤ فى الاختبارات القبلىة؟:

- نعم - لا

14 - نوعية التصميم التجريبي:

- المجموعة الواحدة قبلى بعدى
- المجموعتين وقياس قبلى بعدى وعشوائية
- المجموعتين وقياس قبلى بعدى وعشوائية
- المجموعتين وقياس بعدى
- المجموعات الثلاثة وقياس قبلى وبعدى
- تصميم سولمون ذو الاربع مجموعات

15- نوعية المعالجة او التدخل او البرنامج:

- ارشادى - علاجى - تدريبي

فاذا كان البرنامج لتخفيف الاكتئاب:

- معرفى - سلوكى - سلوكى معرفى - اسرى

16- مدة البرنامج بالاسبوع:

17- النظرية القائم عليها البرنامج:

18- طبيعة المعالجة التى تلقىها المجموعة التجريبية:

- معالجة تجريبية اخرى - معالجة تقليدية - لم تتلقى اى معالجة - لم يحدد

تأكيد حجم التأثير

لتأكيد حجم التأثير لكل دراسة يجب تأكيد المفردات الاتية:

1- كود او رقم الدراسة.

2- عدد حجم التأثير فى الدراسة الواحدة.

توصيفات المتغير التابع

3- نوع حجم التأثير:

- للقياسات القبلية - القياسات البعدية - القياسات التتبعية

4- المتغير او المتغيرات الناتج:

بيانات حجم التأثير

5- نوعية البيانات لحساب حجم التأثير:

- متوسطات وانحرافات المعيارية - قيمة T - قيمة F

- قيمة كاي تربيع - التكرارات والنسب للمتغيرات التصنيفية

6- رقم الصفحة الذى قدرت فيه حجم التأثير فى البحث:.....

حجم العينة:

المجموعة الضابطة:.....

المجموعة التجريبية:.....

المتوسطات والانحرافات المعيارية

متوسط المجموعة التجريبية:.....

متوسط المجموعة الضابطة:.....

الانحراف المعيارى للمجموعة التجريبية.....

الانحراف المعيارى للمجموعة الضابطة.....

النسب والتكرارات

عدد افراد المجموعة التجريبية اللذين حققوا نجاح او تحسن

عدد افراد المجموعة الضابطة اللذين حققوا نجاح او تحسن

نسبة افراد المجموعة التجريبية اللذين حققوا تحسن

نسبة افراد المجموعة الضابطة اللذين حققوا تحسن

اختبارات الدلالة

- قيمة اختبار T

- قيمة اختبار F (df=1)

- قيمة كا تربيع

حساب حجم التأثير

- قيمة حجم التأثير:

- قيمة حجم التأثير:

- كبيرة - متوسطة - ضعيفة.

قاعدة البيانات لما وراء التحليل

ناتج عملية التكويد يجب ان تترجم الى ملف بيانات الكترونى لاجراء ما وراء التحليل

كالآتى:

كود الدراسة	المؤلفين	تاريخ النشر	حجم العينة	العمر	الجنس	المرحلة	انتقاء العينة
01	عامر	2003	30	17.5	1	4	1
02	الهادي	2006	40	13	3	3	2
03	القطراوي	2009	50	16.8	2	4	3

فالمغيرات او الاعمدة تمثل خصائص الدراسة المكودة اما الصفوف تمثل الدراسات، وقاعدة البيانات تتضمن كل خصائص الدراسة من العينة والمنهج والطريقة وحجم التأثير وعليه فان هذا الجدول يتضمن عدد كبير من الاعمدة.

الفصل الخامس

حساب وتكويد حجم التأثير

Computing and Coding Effect Size

مقدمة

الخطوة المهمة فى ما وراء التحليل تكويد او قياس النواتج البحثية على مقياس كمي ومن اهم النواتج البحثية موضع اهتمام الدراسة العلاقات بين بنائين او متغيرين.

ويقوم اسلوب ما وراء التحليل على احصاء حجم الاثر حيث يعتبر المادة الخام ومن اهم المعلومات المستخلصة من الدراسات المتضمنة فى التحليل. ويشير Borenstein et al. (2012) الى ان مصطلحى تاثيرات المعالجة Treatment effects و احجام التأثير Effect sizes تستخدم بمعانى مختلفة بين الباحثين، فيستخدم مصطلح حجم التأثير فى الطب او المعالجات الطبية ليشير الى تاثير المعالجة ويتضمن احصائيات مثل نسب المخاطرة Risk ratios او فروق المخاطرة Risk difference، بينما فى العلوم النفسية والاجتماعية يستخدم مصطلح حجم التأثير ليشير الى فروق المتوسطات المعيارية او الارتباطات حيث انها اكثر استخداماً فى دراسات ما وراء التحليل فى الدراسات النفسية والاجتماعية.

ويعتبر مصطلح حجم التأثير اكثر مناسبة عندما يستخدم كمؤشر لتكميم العلاقة بين متغيرين او الفروق بين مجموعتين فى المقابل مصطلح تاثير المعالجة مناسب فقط كمؤشر لتكميم اثر معالجة او تدخل، وعليه فان دراسة الفروق بين الذكور والاناث يسمى حجم الاثر بينما الفروق بين المجموعة الضابطة والتجريبية يطلق عليها اما حجم الاثر او تاثير المعالجة (Borenstein, et al., 2009).

الدلالة الاحصائية وحجم التأثير

قبل التعرض لمفهوم حجم التأثير يجب الاشارة الى ان حجم التأثير والدلالة الاحصائية مفهومان مختلفان، ويرى Card (2012) وعلى الرغم ان حجم التأثير يقدر لنتائج الاختبارات الدالة احصائياً ويتاثر باحتمالية الحصول على نتائج دالة

احصائياً (القوة الاحصائية) غير انه من المفيد التفريق بين حجم التأثير والدلالة الاحصائية وفيما يلي الفروق بينهما:

الجدول (1.5): الفروق بين الدلالة الإحصائية وحجم التأثير.

حجم التأثير	الدلالة الإحصائية
<ul style="list-style-type: none"> • تجيب عن ما حجم هذا التأثير أو قوة العلاقة؟ • المظهر المكمل لاختبارات الفروض الكلاسيكية. • تعطي القيمة التطبيقية للقرار. • القرار في ضوءها علي متصل كمي. • لا تتأثر بحجم العينة. • يتم تقديرها قبل إجراء الدراسة وأيضا بعد إجرائها. • يمكن إجراء اختبارات الفروض بدونها. • قيم d إذا كانت > 0.2 حجم تأثير ليس له قيمة عملية ومن 0.2 إلي 0.49 حجم تأثير صغير ومن 0.5 الي 0.79 حجم تأثير متوسط و 0.8 فأكثر حجم تأثير كبير. 	<ul style="list-style-type: none"> • تجيب عن هل التأثير الملاحظ حقيقي أم يعود للصدفة؟ • المظهر الأولي لاختبارات الفروض الكلاسيكية. • لا تعطي القيمة التطبيقية للقرار. • القرار في ضوءها قطبي ثنائي (رفض - قبول). • تتأثر بحجم العينة. • يتم تقديرها مسبقا قبل إجراء البحث أو من خلال البيانات. • لا يمكن إجراء اختبارات الفروض إلا بدونها. • قيم p إذا كانت > 0.001 دليل قاطع للدلالة ومن 0.001 إلي 0.01 تعطي دليل جوهري و من 0.01 إلي 0.05 دليل مقبول واكبر من 0.05 لا دليل علي الدلالة الإحصائية

افترض ان باحث اراد ان دراسة ما اذا كانت المجموعتين (ذكور مقابل اناث او المجموعة التجريبية مقابل المجموعة الضابطة) تختلفان على متغير التحصيل، وقام بجمع البيانات من خمسة افراد في كل مجموعة ($N=10$) ونتائج اختبار T المستقلة

حيث: $p=.013$, $T(8)=3.16$ ان المجموعة الثانية لها متوسط اعلى ثم نشرت الدراسة.

واراد باحث اخر اعادة نتائج نفس الدراسة، وجمع بيانات لثلاثة افراد فى كل مجموعة ($N=6$) وقيمة $p=.071$, $T(4)=2.45$ ، وتوصل الى استنتاج الى عدم اختلاف بين المجموعتين، وعليه فشل الباحث الثانى فى اعادة نتائج الدراسة الاولى، وعلى ذلك حدث تعارض.

وحاول باحث ثالث حل هذا التعارض من خلال اجراء نفس الدراسة وتضمنت كل مجموعة 10 افراد ($N=20$) وتوصل الى دلالة احصائية: $p=.00016$, $T(18)=4.74$ والدلالة الاحصائية اكثر قوة.

هذا المثال يوضح محددات الاعتماد على اختبارات دلالة الفروض الصفرية فى مقارنة النتائج عبر الدراسات، لكن الاعتماد على حجم التأثير يجنب الباحث محددات الاعتماد على الدلالة الاحصائية وتحاول الذهاب ابعد من السؤال هل يوجد اختلاف؟، الى ما هو مقدار هذا الاختلاف؟. بالتالى فان الدلالة الاحصائية تقدم لنا دليل على وجود فرق بين المعلم اوالعلاقة الحقيقية بين المتغيرات ولكن لاتخبرنا عن حجم هذه الفروق او العلاقة. والنتائج الدالة احصائياً ربما لاتكون كبيرة او هامة او تكون ضعيفة ايضاً.

مفهوم حجم التأثير

أخذ مفهوم حجم التأثير (ES) Effect size مسميات مختلفة فى التراث البحثي منها قوة العلاقة Strength of association، قوة الارتباط Strength of relation ، مقدار التأثير Magnitude of effect ، حجم التباين المفسر Explained variance ، المناسبة العملية Practical Relevance ، والدلالة العملية.

ويري (2011) Richardson أن مصطلح تأثير effect مرتبط باستخدام معالجات مختلفة فى البحث التجريبي ولكن المسمي الأكثر انتشاراً فى التراث هو حجم التأثير.

ولاهداف ما وراء التحليل من الضروري تعريف حجم التأثير كمؤشر لاتجاه وقوة العلاقة بين متغيرين. ويعرف بأنه تقدير لدرجة وجود الظاهرة في المجتمع (Cohen, 1988) ، أو تكميم لحجم الفروق أو قوة العلاقة بين متغيرين فأكثر وبصفة عامة حجم النواتج (Synder & Lawson, 1993). ويعرفها (Olejnik & Algina (2003) بأنه مؤشر معياري وتقدير للمعلم ومستقل عن حجم العينة ويكمم مقدار أو حجم الفروق أو العلاقة بين المتغيرات في المجتمع. ويعرفها (Grissom & Kim (2005) هي مقياس الى اي درجة الفرض الصفري خاطئ، بمعنى الى اي درجة تختلف النتائج عن قيمة المعلم في المجتمع المحددة في الفرض الصفري. ويعرفها Kelly & Preacher (2012) تمثيل كمي لمقدار وجود الظاهرة من اجل الاجابة عن سؤال بحثي، وهو مسمى تعبيرى يناظر او يرسم البيانات، الاحصائيات، او المعالم كميأ لتعرض مقدار الظاهرة، وعرفها (Littell et al.(2008) بانها مقياس لقوة (مقدار) واتجاه العلاقة بين متغيرين، وبراها (Lipsey & Wilson (2001) بانها احصاء يستخدم لعرض نتائج الدراسة في ضوء الاتجاه والقوة. وبراها (Kline (2013) بانها احصائية Statistic في العينة او معلم Parameter في المجتمع هدفه تكميم الظاهرة موضع الاهتمام.

ويختلف تحديد مفهومه اعتماداً على تصميم الدراسة، ففي الدراسات التجريبية يشير الى مقدار التاير للمتغير المستقل في المتغير التابع، في الدراسات الارتباطية يشير الى درجة التغير او الارتباط بين المتغيرين.

ويعطي حجم التأثير مؤشراً للدلالة العملية لنتائج الاختبارات الإحصائية أو بصفة عامة لنتائج الدراسة (Kirk, 1996)، وكلما زادت قيمة حجم التأثير فان ذلك يعني تواجد الظاهرة بدرجة كبيرة. اي هو مقدار او حجم النتيجة التي حدثت او الموجودة بالفعل في المجتمع.

وحجم التأثير الجيد يمتلك الخصائص الاتية (Kelley & Preacher, 2012) :

- مقياس متري كمي مناسب للاجابة عن سؤال البحث.
- مستقل عن حجم العينة بمعنى لا يتاثر بحجم العينة.

- يمتلك خصائص احصائية جيدة مثل غير متحيز ومتسق، وذو كفاءة بمعنى له اقل عرضة لخطأ المعاينة.
- يمكن التعبير عنه في ضوء فترات الثقة.

أهمية حجم التأثير

لحجم التأثير أهمية كبيرة في مجال الدراسات كالاتي:

- طريقة لتكميم فعالية البرنامج أو التجربة علي متصل.
 - يساعد الباحثين علي تفسير النتائج.
 - يستخدم قبل إجراء الدراسة في مرحلة التخطيط حيث يرشدنا في تقدير حجم العينة (Cohen, 1994).
 - يساعد علي التخلص من عيوب الدلالة الإحصائية أي أنه مكمل للدلالة الإحصائية عند صناعة القرار (Thompson, 1998) .
 - مؤشر لمدي تطبيق نتائج الدراسة في الواقع الحقيقي عكس الدلالة الإحصائية.
 - يستخدم لدمج او مقارنة نتائج الدراسات لمجال معين وذلك لعدم تأثرها بحجم العينة أو ظروف الدراسة وهو ما يطلق عليه بأسلوب ما وراء التحليل أي إنها المادة الخام لأسلوب ما وراء التحليل (Ledesma, Macbeth & Dekohan, 2009; Prentice & Miller , 1992)
 - تشكل مكون هام في مجال نشر البحوث في المجالات العلمية حيث تشترط معظم الدوريات العلمية تقدير حجم التأثير كشرط للنشر ., (Kirk, 2001; Ledesma et al ., 2009)
 - تعتبر مكون أساسي في تحليل أو تقدير قوة الاختبار الإحصائية قبل وبعد إجراء الدراسة (Cohen, 1988; 1990; 1994).
- وعلي ذلك يتضح الفوائد العديدة من تقدير حجم التأثير وهي إعطاء القيمة التطبيقية للنتائج، واستخدامها في تقدير حجم العينة اللازمة لتحقيق مستوي قوة احصائية معينة، وايضاً الدور الهام الذي تلعبه في تطبيق أسلوب ما وراء التحليل للدراسات في مجال معين.

وعلى الرغم من أهميتها إلا أن استخدامها محدود في الممارسات البحثية.

فسمح (1995) Stout & Ruble 153 دراسة في ثلاث مجالات للتربية المحاسبية في عام 1994 وجد 20% من الدراسات قدرت حجم التأثير وبالنسبة لنماذج تحليل التباين وجد ان 8.5% من الدراسات قدرت حجم التأثير و 11 دراسة قدرت R^2 ودراستين قدرت مؤشر w^2 ودراستين قدرت مؤشر η^2 ويوجد أربع دراسات من العشرين التي استخدمت تحليل الانحدار قدرت R^2_{adj} وأخيرا دراستين من 153 دراسة استخدمت مؤشر استدلالي لحجم التأثير وهو ω^2 .

وقام (1996) Kirk باستخدام حجم التأثير في أربعة مجالات تابعة للرابطة النفسية الأمريكية (APA) في عام 2001 واتضح أن نسبة المقالات التي قدرت حجم التأثير (سواء حجم تأثير فروق معيارية أو قوة علاقة) تراوحت من 12% لمجلة علم النفس التجريبي إلى 77% لمجلة علم النفس التطبيقي وأن 60% من أحجام التأثير المقدره مؤشر R^2 .

وقيم Keselman, Huberty, Lix, Olejnik, Cribbie, Donahue, Kowalchuk, و Lowman, Petoskey, Keselman & Levin (1998) مدى استخدام حجم التأثير لـ 226 دراسة في مجال التربية وعلم النفس في 17 مجلة منشورة في الفترة من 1994 حتى 1995 واتضح أن 16 دراسة قدرت حجم التأثير بنسبة 7.1%.

وفحص Fiddler, Cumming, Thomason, pannuzzo, Smith, Fyffe, و Edmonds, Harrington, Schmitt (2005) استخدام حجم التأثير في مجلة علم النفس الإرشادي الاكلينيكي في الفترة من 1993 إلى 2001 واستخدمت مؤشرات حجم التأثير المعيارية لاختبارات ANOVA و X^2 وهي علي التوالي 32% و 16%.

وحلل (2008) Zientek, capraro & Capraro 217 اختباراً إحصائياً في 174 دراسة منشورة قبل 2005 ووجد ان 39% من الدراسات قدرت حجم التأثير وأن 94% من هذه الدراسات اعتمدت علي معاملات الارتباط ومعاملات الانحدار كمؤشر لحجم التأثير و 7% استخدمت حجم التأثير في الاختبارات الفارقة مثل مؤشر d.

وحلل (2009) Alhija & Levey 183 اختباراً إحصائياً في 99 دراسة في مجالات نفسية وتربوية في الفترة من 2003 إلى 2004 و توصلوا إلي أن معظم الدراسات اعتمدت علي مؤشرات r و R و χ^2 كمقياس لقوة العلاقة (86%) وأن (57%) من الدراسات الفارقة قدرت حجم التأثير باستخدام مؤشرات d و μ^2 و μp^2 ويوجد تفسير محدود لمؤشرات حجم التأثير في هذه الدراسات.

وأجرى (2010) Sun, Pan, & Wang مسحاً لـ 1243 دراسة منشورة في 14 مجلة علمية تربوية ونفسية في الفترة من 2005 حتى 2007 ووجد أن 49% من المقالات قدرت حجم التأثير و فقط 57% منها فسرت قيمة حجم التأثير والمؤشر الأكثر استخداماً هو R .

وراجع (2011) McMillan أربع مجلات من إصدارت APA في الفترة من 2008 إلى 2010 متضمنة 417 دراسة وتوصل إلي أن 74% من الدراسات قدرت حجم التأثير وكانت أكثر ثلاثة مؤشرات استخداماً هي R^2 و η^2 و d ويوجد استخدام محدود جداً لمؤشر (g) Hedges و Odds ratio (OR) و Cohen (f) و ω^2 و أن 20% من الدراسات تضمنت تفسيراً لمؤشر حجم التأثير وأشار إلي زيادة استخدام حجم التأثير من 1997 حتى 2010.

وشدد (1989) Rosnow & Rosenthal على ضرورة تقدير حجم التأثير عند تقدير للنتائج الدالة إحصائياً وغير الدالة إحصائياً وهذا عكس ما يعتقد الباحثين من أن تقدير حجم التأثير يكون للنتائج الدالة إحصائياً فقط.

تصنيف مؤشرات حجم التأثير Effect size Taxonomy

تتاول العديد من الباحثين تصنيفات مختلفة لمؤشرات حجم لتأثير أهمها:

- تصنيف (Kirk, 1996; Richardson, 1996) وفيه قسمت مؤشرات حجم التأثير إلى:
أ- مؤشرات الفروق المعيارية بين المتوسطات : وتتضمن مؤشر كوهين d وجلاس Δ (دلتا) وهيدجز (g) وهو تحدد حجم الفروق لیتضمن متغير تصنيفي مستقل ومتغير تابع متصل (اختبار T المستقلة).

ب- مؤشرات قوة العلاقة أو التباين المفسر مثل مؤشرات R^2 ، η^2 (ايتا)، ω^2 (اوميغا)، ε ايبسلون، r^2 ، φ^2 (فاي) وغيرها. وهي مؤشرات تحدد نسبة التباين المفسر في المتغير التابع نتيجة التغير الحادث في المتغير المستقل.

• صنفها (Fren & Monroe (1996) إلى:

أ. مؤشرات الفروق بين المتوسطات : تتضمن مؤشرات d و Δ و g .

ب. مؤشرات العلاقة وتتضمن:

- المؤشرات الارتباطية: وتتنوع حسب طبيعة البيانات المحللة وأشهرها معامل

ارتباط بيرسون (r)، معامل ارتباط سيرمان (ρ)، معامل ارتباط Point- (r_{pd})

biserial، معامل ارتباط فاي (φ)، معامل ارتباط كرامير (V) Crammer

، وأخيراً مؤشر η وهي صيغة تعميمية لـ r وغيرها.

- مؤشرات التباين المفسر: تستخدم في تصميمات تحليل التباين مثل η^2 ، ω^2

، R^2_{adj} ، اختبار Interclass (Icc) وغيرها.

• صنفها (Vacha-Haase & Thompson (2004) إلى:

- مؤشرات الفروق بين المجموعات: هي تحدد حجم الفروق بين مجموعتين مثل

مؤشرات d ، Δ ، و g .

- مؤشرات قوة العلاقة: وهي تقدر حجم التباين المفسر بين متغيرين فأكثر مثل مؤشر r ،

η^2 ، w^2 ، ε^2 ، R^2 ، φ^2 ، وغيرها.

- التقديرات المصححة: مثل مربع معامل الارتباط المتعدد المصحح R^2_{adj} ، ω^2_{adj} ، η^2_{adj}

وغیرها وهي مصححة من خطأ المعاينة Sampling error الناتج عن العينات الصغيرة.

- تقديرات المخاطرة Risk estimates: هذه المؤشرات تقارن بين مجموعتين فأكثر في

المخاطرة النسبية لمخرج أو ناتج معين وهي أكثر شيوعاً في أبحاث المجال الطبي مثل

مؤشر المخاطرة النسبية (RR) Relative Risk أو مؤشر Odds ratio (OR).

• وصنفت (Fidler et al. (2005 إلى:

أ. مؤشرات الفروق المعيارية (d ، Δ ، g) في مقابل مؤشرات التباين المفسر (η^2)

، ω^2 ، r^2 ، R^2 .

ب. المؤشرات غير المصححة (η^2 ، ω^2 ، g و R^2) في مقابل المؤشرات المصححة (R^2_{adj} ، ω^2_{adj} ، q^2_{adj} ، g^2_{adj}).

فالمؤشرات غير المصححة متحيزة تحيز موجب Positively biased مثل مؤشر η^2 في حين المؤشرات المصححة تعالج هذا التحيز مثل η^2_{adj} .

• **صنفها** (Kampense, Dyba, Hannay, & Sjoberg (2007) الي عائلتين هما:

أ. عائلة d: تتضمن مؤشرات d ، Δ ، g وغيرها.

ب. عائلة r: وتتضمن مؤشرات r ، Γ_{pd} ، φ ، η^2 ، ω^2 .

• **صنفها** (Stout & Ruble (1995) إلى:

أ. مؤشرات تحليل الارتباطات البسيطة مثل r ، σ ، φ وغيرها.

ب. مؤشرات المقارنات البسيطة: التصميمات ذوي المعالجة او العامل الواحد مع مجموعتين فأكثر ω^2 ، η^2 ، d ، g وغيرها.

ج. مؤشرات المقارنات المعقدة: التصميمات العاملة التي بها أكثر من معالجة او عامل مثل η^2 ، ω^2 .

د. مؤشرات تحليل الانحدار: مثل مؤشر R^2 ، R^2_{adj} ، β وغيرها.

• **صنفها** (Trusty, Thompson & petrocelli (2004) في ضوء الاختبار الإحصائي إلى:

أ. مؤشرات حجم التأثير في ضوء تحليل التباين والتحليلات المرتبطة به مثل مؤشر η^2 ، ω^2 وأيضاً مؤشر Wilks lambda في MANOVA ومؤشر d في اختبار (T).

ب. حجم التأثير للانحدار المتعدد مثل R^2 ، β ، R^2_{adj} .

ج. حجم التأثير لجداول (X^2) مثل Crammer V ، ϕ^2 ، Cohen Kappa ،

Tau Goodman ، Kruskal gamma

وانتقد (2003) Breugh تصنيف مؤشرات حجم التأثير إلي مؤشرات الفروق

بين المتوسطات ومؤشرات التباين المفسر كما أوردها (Kirk 1996; Richardson,

(1996) وقال أن التصنيف بهذا الشكل يؤدي إلى استنتاجات خاطئة حيث يمكن تحويل مؤشر d إلى r والعكس صحيح.

ويري (Rosenthal, Rosnow & Rubin (2000) ان مقاييس التباين المفسر تعطي

تقديرات منخفضة under-estimating لأهمية العلاقة.

• **صنف (2003) Breugh مؤشرات حجم التأثير في ضوء مستويات القياس للمتغيرات المستقلة والتابعة إلى:**

أ- في حالة المتغير المستقل تصنيفي Dichotomus والتابع تصنيفي فانه يستخدم مؤشر فاي ϕ ، نسبة المخاطرة (RR) Relative Risk ، نسبة الترجيح او الاحتمال Odds Ratio وغيرها ولكن أكثرهم شيوعا هو فاي (ϕ) .

ب- في حالة المتغير المستقل تصنيفي بمستويين والتابع متصل: يستخدم مؤشر Δ , g , rbp , d وغيرها.

ت- في حالة المتغير المستقل متصل والتابع تصنيفي: يستخدم الانحدار اللوغاريتمي او اللوجستك logistic Regression ومؤشرات logged Odds و Odds ratio وهي تعطي خلال مخرج SPSS.

ث- في حالة المتغير المستقل متصل والتابع متصل: يستخدم مؤشرات r ، r^2 ، μ^2 .

ج- في حالة متغير تصنيفي بأكثر من مستويين والتابع متصل فانه يستخدم مؤشر η^2 و ω^2 وفي حالة تصميمات عاملية بأكثر من عامل يمكن استخدام η^2 partial و ω^2 partial.

• **وتصنف مقاييس حجم التأثير إلى :**

أ. مقاييس معيارية: هي مقاييس حرة من وحدات القياس ويتم تقديرها في ضوء الانحراف المعياري للبيانات ويمكن لهذه المقاييس مقارنتها ببعضها أو علي الأقل تحويل احدهما للأخر وبندرج تحتها كل المؤشرات السابقة.

ب. مقاييس غير معيارية: تقدر في ضوء وحدات خام بدون الأخذ في الاعتبار الانحراف المعياري للبيانات وهي غير مستقلة (حرة) عن مستوي القياس مثل

الفروق بين المتوسطات ($\bar{X}_1 - \bar{X}_2$) أو الفروق في الوسيط أو الفروق في النسب المئوية وغيرها.

وفيما يلي عرض لثلاثة أنواع رئيسية لاحجام التأثير:

مؤشرات الفروق بين المتوسطات

فروق المتوسطات غير المعيارية Unstandardized Mean Difference

عندما تحدث مقارنة بين مجموعتين (ضابطة - تجريبية، ذكور - اناث) على متغير متصل ما (التفكير الابداعي - الاكتئاب) فان احصاء حجم التأثير يمكن التعبير عنه بالفروق بين متوسطى المجموعتين ويسمى احصاء حجم التأثير غير المعيارى او الخام Unstandardized effect size، ولكن يفترض ان تستخدم الدراسات نفس المقياس فمثلاً هدفت الدراسات الى دراسة فعالية برامج الالعب التعليمية فى تنمية الذكاء، واعتمدت الدراسات على مقياس ستانفورد بينيه فى هذه الحالة يستخدم الفروق الخام بين المتوسطات مؤشر لحجم التأثير بفاعلية او يستخدم للقياسات فى المجال الطبى التى لها مقياس موحد ومعروف مثل ضغط الدم او يستخدم لمقاييس واسعة الانتشار مثل مقياس Beck للاكتئاب او تقدير الذات لـ Rosenberg وغيرها.

ويقدر حجم التأثير كالاتى (Lipsey & Wilson, 2001; Borenstien et al., 2009):

$$ES_{um}(D) = \bar{x}_1 - \bar{x}_2 = M_1 - M_2$$

• \bar{x}_1 ، M_1 متوسط المجموعة الاولى

• \bar{x}_2 ، M_2 متوسط المجموعة الثانية

ولتقدير الخطأ المعيارى لفروق المتوسطات غير المعيارية لاستخدامها فى تصحيح او موازنة

حجم التأثير كالاتى (Card, 2012; Lipsey & Wilson, 2001):

$$SE_{um} = S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \approx \frac{2S_p}{\sqrt{N_{total}(equal\ sample\ size)}}$$

• S_p الانحراف المعيارى المدمج اوالمشترك Pooled للمجموعتين.

• n_1 ، n_2 حجمى المجموعتين الاولى والثانية.

• N_{total} حجم العينة الكلى عند تساوى حجم العينة فى كل مجموعة.

ويقدر الانحراف المعيارى المدمج للمجموعتين كالاتى:

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

• S_1^2 تباين قيم المجموعة الاولى.

• S_2^2 تباين المجموعة الثانية.

مثال: افترض ان دراسة تجريبية هدفت الى المقارنة بين المجموعة التجريبية والمجموعة الضابطة فى العدوان وكان متوسط المجموعة الاولى $(\bar{x}_1) = 103$ ومتوسط المجموعة الثانية $(\bar{x}_2) = 100$ والانحراف المعيارى للعينتين $S_1 = 5.5$ ، $S_2 = 4.5$ وحجم العينة للعينتين متساوى $n_2 = n_1 = 50$ وعليه فان حجم التأثير:

$$SE_{um} = 103 - 100 = 3$$

والانحراف المعيارى المدمج كالاتى:

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

والخطأ المعيارى لفروق المتوسطات:

$$SE_{um} = S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} = 5.0249 \sqrt{\frac{1}{50} + \frac{1}{50}} = 1.0050$$

الاجراءات الحسابية: تتطلب حساب حجم التأثير من خلال حساب الفرق بين المتوسطين مباشرة وتوافر الانحراف المعيارى للمجموعتين واحجام عينة. ولكن بفرض ان بعض الدراسات لم تذكر احد هذه المؤشرات خاصة الانحراف المعيارى فى هذه الحالة يمكن حساب الانحراف المعيارى من قيمة الاختبار مباشرة مثل T و F.

مؤشرات فروق المتوسطات المعيارية Standardized Mean Difference

وهي تحدد مقدار الفروق بين متوسطات مجموعتين على درجات المتغير التابع الذي ليس له نفس التعريف الاجرائى عبر عينات كل الدراسات (مثلاً مفهوم الذات او العنف او ادمان الفيس بوك)، وهي أكثر ملائمةً للتصميم التجريبي ذو المجموعتين الضابطة والتجريبية لدراسة فعالية المعالجات او البرامج، ويستخدم ايضاً فى الدراسات غير التجريبية التي تهدف الى دراسة الفروق بين الذكور والاناث على متغيرات تابعة وهذه النوعية من الدراسات مشتقة من منهجية البحث السببى المقارن ويمكن اعتبار حجم التأثير كمؤشر للعلاقة بين متغير تصنيفى بمجموعتين ومتغير متصل ومن أهم المؤشرات الاكثر استخداماً فى تحليلات ما وراء التحليل:

مؤشر كوهين d (Cohen (1969, 1988): أكثر مؤشرات حجم الفروق شيوعاً لسهولة

تفسيره ومعياريته ويستخدم بكثرة في دراسات ما وراء التحليل والتصميمات التجريبية وشبه التجريبية ويتحدد بالمعادلة الآتية:

$$d = \frac{\bar{x}_1 - \bar{x}_2}{Sd_{pooled}}$$

- \bar{x}_1 متوسط المجموعة التجريبية، \bar{x}_2 متوسط المجموعة الضابطة.
- Sd_{pooled} الانحراف المعياري المشترك للعينة

ويقدر Sd_{pooled} كالآتي:

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

ويفضل استخدام هذا المؤشر إذا توفر التوزيع اعتدالي للمتغير التابع (المتصل) وايضاً تساوي تباين المجموعتين. وتوجد إشكالية عند حساب مؤشر d في حالة المقارنات المتعددة المصاحبة لتحليل التباين (توكي وشيفية وغيرها) حيث غالباً لا يوجد تجانس بين تباينات المجتمعات عند مقارنة كل مستويين ببعضها البعض (Ferguson, 2009).

مؤشر جلاس Δ (دلتا) لـ (Glass (1976 : اقترحه Glass (1976) كمؤشر لحجم التأثير مشابهة لمؤشر d وتتحدد قيمته بالمعادلة الآتية:

$$\Delta (g_{Glass}) = \frac{\bar{x}_1 - \bar{x}_2}{S1}$$

• S1 الانحراف المعياري لمجتمع المجموعة الضابطة.

ونلاحظ أن Glass استعان بالانحراف المعياري لمجتمع المجموعة الضابطة، في حين اعتمد

Cohen علي الانحراف المعياري المشترك للمجموعتين في العينة. ويفترض مؤشر Δ (g_{Glass}) تساوي تباينات المجموعتين الضابطة والتجريبية ($S^2c = S^2e$) وفي حالة عدم تساوي فان مؤشر دلتا يكون متحيز تحيز موجب أي يعطي تقدير عالي لحجم التأثير (Ledesma, Macbeth, & Dekohan, 2009).

ولكن هذا المؤشر اقل تطبيقاً في الدراسات التي لا تتضمن مجموعة ضابطة مثل المقارنة بين الذكور والإناث (Ferguson, 2009). ورغماً عن ذلك ففي الدراسات التجريبية فان الانحراف المعياري للمجموعة التجريبية يتاثر بالمعالجة فربما يوجد تنوع في درجات المجموعة التجريبية اكثر من نظيرتها الضابطة وفي بعض المواقف العكس (خفض العدوان او الاكتئاب) وفي هذه الحالة من الافضل استخدام الانحراف المعياري للمجموعة الضابطة لانها لاتتأثر بالمعالجة واقد استخدم Smith & Glass(1977) هذه الاستراتيجية في دراسة ما وراء التحليل للكشف عن فعالية برامج العلاج النفسى.

ويمكن استخدام مؤشر $\Delta(g_{Glass})$ في حالة تصميم المجموعة الواحدة وقياس قبلي وبعدي حيث يستخدم اختبار T المرتبطة ويكون المقام هو الانحراف المعياري للقياس القبلي (Lipesy & Wilson, 2001).

ومن اهم محدداته لا يمكن حسابه من نتائج اختبارات الدلالة مباشرة مثل حسابه من قيمة اختبار T مباشرة، ويعتمد على مجموعة واحدة فقط لحساب الانحراف المعياري وهذا اقل دقة من حسابه للمجموعتين، ولهذا السبب فهو اقل استخداماً في دراسات ما وراء التحليل مقارنة بمؤشرى d, g.

مؤشر هيدجز (g) لـ (Hedges (1981): عدل Hedges (1981) مؤشر دلتا لـ Glass ، حيث تخلص من محدداته واقترح إحلال الانحراف المعياري للمجموعة الضابطة بالانحراف المعياري المشترك لمجتمع العينتين وتحدد قيمة g من المعادلة الآتية (Cooper & Hedges, 1994; Hedges, 1981; Hedges & Olkin, 1985)

$$g(ES_{sm}) = \frac{\bar{x}_1 - \bar{x}_2}{S_{Pooled}}$$

• S_{Pooled} الانحراف المعياري المدمج لمجتمع العينتين.

وتحدد قيمة S_P كالآتي:

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

والخطأ المعياري لهذا المؤشر:

$$SE_g = \sqrt{\frac{n_1+n_2}{n_1n_2} + \frac{g^2}{2(n_1+n_2)}}$$

ويفترض مؤشر g ان توزيع المتغير التابع المتصل اعتدالي التوزيع. ويرى (Durlak (2009) أن مؤشري d , g متحيزان تحيز موجب عندما تكون احجام العينات صغيرة وخاصة عندما يكون حجمها اقل من 20 (Hedges, 1981).

ونقح (Hedges & Olkin (1985) هذا المؤشر للتخلص من التحيز الموجب خاصة في حالة العينات الصغيرة اقل من 20 مثلاً ولذلك لو تضمنت دراسات ما وراء التحليل دراسات ذات احجام عينات صغيرة فيجب تطبيق التصحيح الآتي (Hedges & Olkin, 1985; Lipsey & Wilson, 2001):

$$g_{adj} = g \left(1 - \frac{3}{4df - 1}\right) = g \left(1 - \frac{3}{4N - 9}\right)$$

والخطأ المعياري لمؤشر g كالآتي (Lipsey & Wilson, 2001):

$$SE_{g_{adj}} = \sqrt{\frac{n_1 + n_2}{n_1 n_2} + \frac{(g)^2}{2(n_1 + n_2)}} \approx \frac{4 + g^2}{N_{\text{total}}(\text{equal sample size})}$$

- g حجم التأثير المصحح من صغر حجم العينة.
- N_{total} حجم العينة الكلى عند تساوى احجام العينتين.

ويستخدم الجزء الاول من المعادلة عندما تكون حجم عينة كل مجموعة معروف بينما الجزء الثاني عندما لا تكون حجم كل عينة غير معروف ويفترض تساوى العينتين. ويعتبر المؤشر اكثر استخداماً فى دراسات ما وراء التحليل وربما يتفوق على مؤشر d .

مثال تطبيقي: ذكر تقرير احدى الدراسات التجريبية لتنمية مفهوم الذات ان حجم المجموعة التجريبية = 25 بينما حجم عينة الضابطة = 30، ومتوسط التجريبية = 132.3 ومتوسط الضابطة = 127.8، والانحراف المعياري للتجريبية = 9.8 والانحراف المعياري للضابطة = 10.4. ولحساب حجم التأثير لابد من حساب الانحراف المعياري المدمج كالاتى:

$$S_{\text{pooled}} = \sqrt{\frac{(25 - 1) \times 9.8^2 + (30 - 1) \times 10.4^2}{25 + 30 - 2}} = 10.08$$

ويقدر حجم التأثير من صيغة Hedges كالاتى:

$$g(ES_{sm}) = \frac{\bar{x}_1 - \bar{x}_2}{S_{\text{pooled}}} = \frac{132.3 - 127.8}{10.08} = 0.45$$

والمؤشرات الثلاثة متفقة فى البسط الفروق بين المتوسطات والاختلاف فى الصيغ الثلاثة لـ g ، d ، Δ فى المقام فاعتبرها Cohen الانحراف المعياري المشترك للعينة (n) حيث الانحراف المعياري:

$$sd = \sqrt{\frac{(x - \bar{x})^2}{n}}$$

والانحراف المعياري للعينة هو تقدير متحيز للانحراف المعياري للمجتمع خاصة للعينات الصغيرة.

واعتمد Hedges علي الانحراف المعياري المشترك للمجتمع (n-1) حيث الانحراف المعياري:

$$s = \sqrt{\frac{(x - \bar{x})^2}{n - 1}}$$

في حين اعتمد Glass علي الانحراف المعياري لمجتمع المجموعة الضابطة لانه يعطى افضل تقدير مقارنة بالانحراف المعياري للمجموعة التجريبية.

ويرى (Hedges & Olkin, 1985) انه في حالة العينات الكبيرة فان المؤشرات الثلاثة يكون لها نفس الاداء ولكن مؤشر g أفضل في حالة العينات الصغيرة خاصة مع استخدام الصيغة المنقحة للمؤشر ويتفق معهم Ferguson (2009) في أن مؤشر g اكثر معيارية من مؤشري Δ , d, ولكن الواضح ان مؤشر g اقل تحيزاً من مؤشر Δ عندما $s_1^2 = s_2^2$ (Ledesma et al., 2009).

فالمؤشرات الثلاثة تقيس حجم التأثير (حجم الفروق) عندما يكون المتغير المستقل تصنيفي بمستويين والمتغير التابع متصل، وعلي ذلك تستخدم في حالة اختبار T المستقلة أو اختبار (F) للمقارنة بين مجموعتين ويمكن استخدامهم في حالة اختبار T المرتبطة.

والملاحظ لمعادلتى حساب d, g من الممكن ان تكون قيمتهما سالبة او موجبة اعتماداً على ما اذا كان متوسط المجموعة الاولى او الثانية اعلى وهذه جودة مرغوبة عندما يتضمن تحليل ما وراء التحليل اتجاهات متناقضة من التأثيرات.

وإذا كان الانحرافات المعيارية لمجمعتي المجموعتين غير متساوي فمن الافضل حساب حجم التأثير من قيم الاختبار مباشرة. ولكن يمكن اعتبار المجموعة التجريبية او الذكور المجموعة الاولى عبر كل الدراسات المتضمنة في ما وراء التحليل. ولاحظ ان فروق المتوسطات المعيارية يمكن ان تأخذ اي قيمة بمعنى انها غير محددة بالحدود ± 1.00 مثل معامل الارتباط r، حيث قيمة g يمكن ان تكون $g = \frac{3-5}{1} = -2.00$ ولذلك فهي غير قابلة للتفسير ولكن Cohen(1969) قدم اقتراحات لتفسيرها حيث $d = 0.20$ حجم تأثير ضعيف او صغير، $d = 0.50$ حجم تأثير متوسط، $d = 0.80$ حجم تأثير كبير. ويجب ان تاخذ في الاعتبار ان هذه المعايير ليست ثابتة انما تختلف باختلاف التخصص والمجال البحثي.

في ضوء ذلك فهذه المؤشرات انسب للاختبارات التي تعالج التصميمات التجريبية ذو المجموعتين التي تستخدم T وعلي الرغم من ان مؤشر g يفضل عند البعض علي مؤشر d، ولكن Kline (2013) افترض انه لو أن النسبة بين الانحرافين المعياريين للمجموعتين أكبر من أربعة فإن حجم التأثير يقدر مرتين باستخدام كل انحراف معياري علي حدة ويتم مناقشة مدي الاختلاف بين العينتين.

حساب مؤشري d , g من الاحصائيات مباشرة

قيمة اختبار F, T

تقدر قيمة d , g من قيمة اختبار (T) للعينات المستقلة إذا كان $n_1 = n_2$ كالاتي:

$$d, g = \frac{2T}{\sqrt{N}}$$

حيث df درجات الحرية : $df = n_1 + n_2 - 2$

وفي حالة $n_1 \neq n_2$ فإن المعادلة (3) تصبح كالاتي (Borenstein 2009):

$$d, g = T \sqrt{\frac{n_1 + n_2}{n_1 n_2}}$$

• T قيمتها المحسوبة سواء بالسالب او الموجب

مثال: ذكر تقرير دراسة دراسة تجريبية ان قيمة $T = 1.68$ لصالح المجموعة التجريبية وحجم

العينة التجريبية = 10 وحجم العينة الضابطة = 12 وعليه فان حجم التأثير:

$$g = 1.68 \sqrt{\frac{10 + 12}{10 \times 12}} = 1.68 \sqrt{\frac{22}{120}} = .72$$

ويمكن تقدير قيمة حجم التأثير **d , g** لاختبار T المرتبطة:

$$g, d = \frac{T_{dependent}}{\sqrt{N}}$$

• $T_{dependent}$ القيمة المحسوبة.

ويمكن تقدير قيمة d , g في حالة استخدام اختبار F للمقارنة بين مجموعتين غير متساويتين كالآتي (Borenstein 2009) :

$$g, d = \sqrt{\frac{F_{(1,df)}(n_1+n_2)}{n_1 n_2}} \approx 2 \sqrt{\frac{F_{(1,df)}}{N}}$$

- $F_{(1,df)}$ نسبة F من درجة حرية = 1 للبسط ، df درجة الحرية للمقام $(N-2)$.
- n_1, n_2 احجام العينات للمجموعتين الاولى والثانية على التوالي.
- N حجم العينة الكلى.

ويمكن تقدير قيمة d في حالة استخدام اختبار F للمقارنة بين مجموعتين متساويتين كالآتي:

$$g, d = 2 \sqrt{\frac{F}{N}}$$

مثال: ذكر تقرير لدراسة تجريبية للفروق بين المجموعة التجريبية والضابطة باستخدام تحليل التباين الاحادى ANOVA حيث $F=13.1$ بدرجتى حرية 1 للبسط و 18 للمقام وعليه فان حجم العينة:

$$N= 98+2= 100$$

وحجم التأثير كالآلى:

$$g = 2 \sqrt{\frac{13.1}{100}} = 0.72$$

وتقدر d , g من قيمة اختبار F المرتبطة كالآلى:

$$d = \sqrt{\frac{F_{repeated}}{N}}$$

تقدر مؤشر d , g من خلال اختبار χ^2

بعض الدراسات تقارن بين مجموعتين على نسب درجات المتغير التابع بعد تصنيفه وفي هذه الحالة يمكن حساب مؤشر حجم التأثير لفروق المتوسطات المعيارية من خلال حساب احصاء χ^2 ثم يحول الى مؤشر g كالتالى (Card 2012):

$$d = 2 \sqrt{\frac{\chi^2_{(1)}}{N - \chi^2_{(1)}}}$$

تقدير حجم التأثير من مستوى القيمة الاحتمالية P

فى بعض الاحيان يذكر الباحث ان اختبار T اظهر دلالة احصائية والقيمة الاحتمالية $P=0.03$ التى حددت فى الدراسة بالتالى لم يذكر الباحث سوى قيمة P وفى هذه الحالة يقدر حجم التأثير كالتالى (Card, 2012):

$$g = \frac{2Z}{\sqrt{N}}$$

حيث: $p=0.03$ تناظر $Z=2.14$ (فى جداول Z)، و $p > .05$ تناظر $Z=1.96$.

ويرى (Lipsey & Wilson 2001) لحساب قيمة T المناظرة لقيم P يوجد جداول خاصة كالتالى:

df	p-value (Two-tailed)								
	.80	.50	.20	.10	.05	.02	.01	.002	.001
1	0.325	1.000	3.078	6.314	12.706	31.821	63.657	318.309	636.621
2	0.289	0.816	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.277	0.765	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.271	0.741	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.267	0.727	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.265	0.718	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.263	0.711	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.262	0.706	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.261	0.703	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.260	0.700	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.260	0.697	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.259	0.695	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.259	0.694	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.258	0.692	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.258	0.691	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.258	0.690	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.257	0.689	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.257	0.688	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.257	0.688	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.257	0.687	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.257	0.686	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.256	0.686	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.256	0.685	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.256	0.685	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.256	0.684	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.256	0.684	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.256	0.684	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.256	0.683	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.256	0.683	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.256	0.683	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	0.255	0.681	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	0.254	0.679	1.296	1.671	2.000	2.390	2.660	3.232	3.460
120	0.254	0.677	1.289	1.658	1.980	2.358	2.617	3.160	3.373
*	0.253	0.674	1.282	1.645	1.960	2.326	2.576	3.090	3.291

SOURCE: Computer generated using SPSS 6.0 for Windows.

مثال: في دراسة ما ذكر الباحث ان قيمة اختبار T دالة احصائياً عند $p = 0.037$ وان حجم العينة الاولى = 10 وحجم العينة الثانية = 12 ودرجات الحرية $df = 20$ وبالكشف في الجدول السابق نجد ان القيمة 0.037 تقع بين 2.086 و 2.528 وعليه فان متوسطهما $T = 2.278$ وقيمة حجم التأثير كالتالي:

$$g = 2.278 \sqrt{\frac{10 + 12}{10 \times 12}} = 2.278 \sqrt{\frac{22}{120}} = 0.98$$

وايضاً يمكن حساب قيمة اختبار F لمجموعتين من القيمة الاحتمالية P وذلك من حساب قيمة T ثم تربيعها حيث:

$$F = T^2$$

تقدير قيمة g من مؤشر كوهين d كالآتي:

$$d = g \sqrt{\frac{N}{df}}$$

• N حجم العينة ، df درجات الحرية

تقدير g او d من قيمة معامل ارتباط بيرسون كالتالي (Card, 2012):

$$g = \left(\frac{r}{\sqrt{1 - r^2}} \right) \sqrt{\frac{df(n_1 + n_2)}{n_1 n_2}}$$

مؤشرات قوة العلاقة:

هي المؤشرات البديلة لفروق المتوسطات المعيارية لتقدير حجم التأثير وتنقسم إلى المؤشرات الارتباطية ومؤشرات التباین المفسر.

المؤشرات الارتباطية: تقيس هذه المؤشرات قوة العلاقة بين متغيرين فأكثر وهي كالتالي:

معامل ارتباط بيرسون (r): يقيس درجة التباین المشترك أو قوة العلاقة الخطية بين متغيرين

متصلين. وتقدر قيمته كالآتي:

$$r_{xy} = \frac{N \sum XY - \sum X \sum Y}{\sqrt{(N(\sum X^2 - (\sum X)^2))(\sum Y^2 - (\sum Y)^2)}}$$

• $\sum XY$ مجموع حاصل ضرب X في Y.

• $\sum X$ مجموع قيم X ، $\sum Y$ مجموعة قيم Y.

- ΣX^2 مجموع مربعات X ، $(\Sigma X)^2$ مربع مجموع X .
- ΣY^2 مجموع مربعات Y ، $(\Sigma Y)^2$ مربع مجموع Y ، N حجم العينة.

ويكون أدائه أفضل إذا كان توزيع البيانات اعتدالياً والعلاقة بين المتغيرين خطية. والارتباط الموجب يعني ان الافرد مرتفعي الدرجات على المتغير X لهم درجات مرتفعة على المتغير Y بينما الارتباط السالب الافراد مرتفعي الدرجات على المتغير X لهم درجات منخفضة على المتغير Y .

وتعتبر قيمة معامل الارتباط مؤشراً لحجم التأثير، وهو مؤشر أكثر معيارية وأسهل في تفسيره نتيجة اعتماده علي معايير في تحديد قوة العلاقة ويقع مداه من 1.0 (موجب) إلى -1.0 (سالب) عكس مؤشر كوهين " d " فليس له معايير واضحة لتفسيره (Rosenthal et al., 2000)

وتفسير قيمة r كما وضحها Cohen (1988) كمؤشر لحجم التأثير إذا كانت $r = \pm 0.10$ حجم تأثير ضعيف، $r = \pm 0.30$ حجم تأثير متوسط، $r \geq \pm 0.5$ حجم تأثير كبير، ولكن عليك الحذر من ان هذه المعايير ليست ثابتة في كل التخصصات او المجالات.

ويفسر r من خلال مربع قيمته (r^2) ويطلق عليه معامل التحديد Coefficient of determination، وعلى ذلك يكون مربع معامل الارتباط ضعيفاً إذا كانت $r^2 = 0.01$ ، متوسط $r^2 = 0.09$ ، وكبير $r^2 \geq 0.25$. فمثلاً إذا كان $r = .50$ فان مربع معامل الارتباط $r^2 = .25$ بمعنى ان احد المتغيرين فسر 25% من تباين المتغير الثاني. وعلى الرغم من مميزات r إلا انه يعطي تقديراً متحيزاً موجباً Positively bias خاصة في حالة العينات الصغيرة (Wang & Thompson, 2006).

وتباين الخطأ او الخطأ المعياري لمعامل الارتباط تقريباً (Borenstien et al., 2009):

$$V_r = \frac{(1 - r^2)^2}{N - 1}$$

حيث N حجم العينة، r معامل الارتباط.

ومعظم تحليلات ما وراء التحليل لاتجرى على معامل الارتباط ولذلك يحول معامل الارتباط قبل دمج احجام التأثير عبر الدراسات المختلفة الى تحويل او مقياس فيشر (Z_r) Fisher

transformation (ليس اختبار الدلالة الاحصائية Z) ويجرى التحليل على القيم المحورة،

ويقدر Z_r كالتالى (Card, 2012; Lipsey & Wilson, 2001):

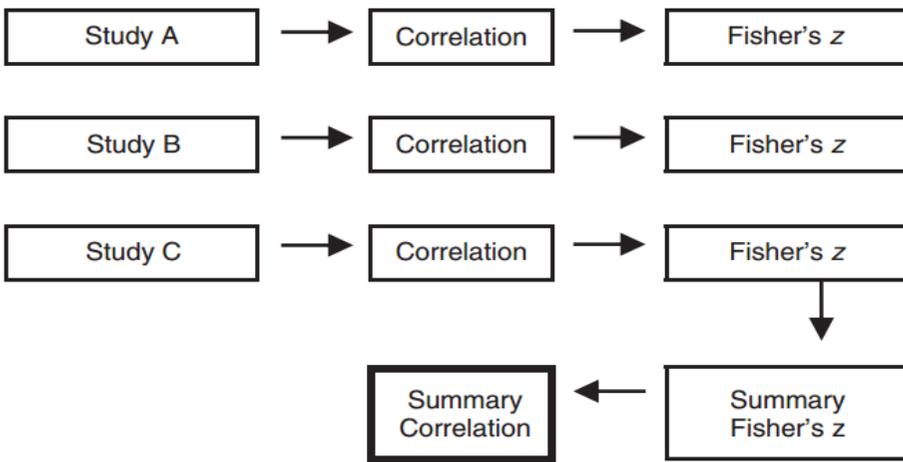
$$(ES_{Z_r})Z_r = \frac{1}{2} \ln\left(\frac{1+r}{1-r}\right)$$

حيث \ln او \log_e اللوغاريتم الطبيعي.

وعليه فان الدراسات التى تعتمد على معامل الارتباط كمؤشر لحجم التأثير فى الدراسات الارتباطية فى ما وراء التحليل يجب تحويله الى الدرجة او تحويل Fisher ثم تحويل القيم النهائية لفisher الى معامل الارتباط لتفسير النتائج.

ويحول معامل ارتباط بيرسون الى تحويل فيشر فى ماوراء التحليل لان التوزيع العينى لمعامل الارتباط فى المجتمع يكون ملتوى ناحية اليمين للمعاملات السالبة وناحية اليسار للمعاملات الموجبة وعندما يكون قريب من الصفر فإن التوزيع العينى يكون تقريباً اعتدالى وتتحسن الاعتدالية بزيادة حجم العينة ولان التوزيع العينى يكون ملتوى كلما زادت قيمة معامل الارتباط فانه من غير المناسب استخدام التوزيع الملتوى لاختبارات الفروض حول معامل الارتباط فى المجتمع ρ إلا إذا كانت $\rho = 0$ وهذا أدى بـ Fisher لاجراء تعديل فيشر Z لانه توزيع منظم Symmetric (Lipsey & Wilson, 2001, Hedges & Olkin, 1985).

ويمكن تشبيه ذلك بالآتى (Borenstein, et al., 2009):



الشكل (1.5): تحويل معاملات الارتباط الى تحويل فيشر.

وتباين تحويل Z_r :

$$V_z = \frac{1}{N - 3}$$

حيث N حجم العينة.

والخطأ المعياري لـ Z_r :

$$\begin{aligned} SE_{Z_r} &= \sqrt{V_z} \\ &= \sqrt{\frac{1}{N - 3}} \end{aligned}$$

وكلما زادت حجم العينة فان مقام المعادلة السابقة يزيد وبالتالي تنقص قيمة الخطأ المعياري. ويجب ادخال الخطأ المعياري لتحويل فيشر الى قاعدة البيانات اثناء عملية التكويد لاستخدامة فى عملية موازنة احجام التأثير، وعليه عند استخدام تحويل فيشر لانستخدم تباين معامل الارتباط انما نعتمد على تباين الدرجة Z_r .

ولان الدرجة Z_r اقل قابلية للتفسير لعدم امتلاكها معايير واضحة مثل معايير معامل الارتباط فتحول نتائج الدرجة فيشر Z_r الى معامل الارتباط من خلال التالى Hedges & Olkin, (1985):

$$r = \frac{e^{2Z_r} - 1}{e^{2Z_r} + 1}$$

حيث r معامل ارتباط بيرسون، Z_r تحويل فيشر المقابل الى r.

فعلى سبيل المثال فى دراسة ما قدرت معامل الارتباط بـ 0.50 مع حجم عينة 100 بالتالى فان تحويل فيشر Z_r كالاتى:

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

وتباين الدرجة Z_r :

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

والخطأ المعياري للدرجة Z_r :

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

ولتحويل الدرجة Z_r الى معامل الارتباط:

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

مثال: ارد باحث دراسة العلاقة بين ادمان الانترنت والوحدة النفسيه بالتالي جمع الدراسات الت تناولت هذه العلاقة من التراث فى البيئه العربيه وهنا يعتبر معامل الارتباط مؤشر لحجم التأثير ثم يحول الى تحويل Fisher وبعد حساب المتوسط العام لتحويل Z يحول مرة اخرى الى معامل الارتباط.

حساب قيمة معامل الارتباط من قيم الاختبارات

بعض الدراسات الاولية المنشورة تتضمن نتائج الاحصائيات الاستدلالية دون حساب حجم التأثير ويفضل فى كثير من تحليلات ما وراء التحليل التعبير عن معظم الاحصائيات فى ضوء معامل الارتباط r وهذا الموقف يحدث لعدة اسباب منها:

اولاً: تتضمن الدراسات الدلالة الاحصائية لمعامل الارتباط بدون قيمة معامل الارتباط وهذا الموقف نادر الحدوث.

ثانياً: بعض الباحثين يميلون الى تقسيم المتغير المتصل الى تصنيفين مثل مرتفع ومنخفض التحصيل او مرتفعى ومنخفضى الدافعية ثم المقارنة بينهما على المتغير التابع باستخدام اختبار T او F ، ولكن هذا تقسيم اصطناعى وبدوره يؤدي الى تقلص Attenuation قيمة حجم التأثير (Hunter & Schmidt, 2004)، ولذلك من المفيد حساب قيمة معامل الارتباط بل وتصحيحة من التقسيم الاصطناعى للمتغيرات المتصلة.

ثالثاً: بعض الدراسات تقسم كلا من المتغيرين المتضمنين فى معامل الارتباط الى تصنيفين ثم تحلل المتغيرين من خلال جداول الاقتران باستخدام احصاء χ^2 وفى هذه الحالة فالباحث بحاجة الى تصحيح قيمة معامل الارتباط.

وتقدر قيمة r كمؤشر لحجم التأثير من قيمة اختبار T المستقلة كالتالى:

$$r = \sqrt{\frac{T^2}{T^2 + df}}$$

• T قيمة الاختبار المحسوبة

• df= N-2= n1+ n2 -2 درجات الحرية:

وتقدر قيمة r من اختبار F لعينتين مستقلتين و $n_1 \neq n_2$ كالتالى:

$$r = \sqrt{\frac{F}{F + df}}$$

• df درجات الحرية للمقام او الخطأ N-C، حجم العينة، C عدد المجموعات ويساوى 2.

ويقدر قيمة r من اختبار T المرتبطة كالتالى:

$$r = \sqrt{\frac{T^2}{T^2 + df}}$$

• df درجات الحرية (df =N-1)

ويقدر قيمة r من اختبار F لعينتين مرتبطتين كالتالى:

$$r = \sqrt{\frac{F_{\text{المرتبطة}}}{F_{\text{المرتبطة}} + df}}$$

والبعض ضد عدم دمج نتائج اختبارات العينات المستقلة مع نتائج اختبارات العينات المرتبطة، وينصح (2001) Lipsey & Wilson باجراء ماوراء تحليل منفصل لنتائج الاختبارات المستقلة (تصميم بين المجموعات او الافراد) عن نتائج الاختبارات المرتبطة (تصميم داخل المجموعات او الافراد)، ولكن (2012) Card يرى لاجراء فصل لابد من اجراء تحليل التفاعل باعتبار نوعية المنهجية متغير تصنيفى (عينات مستقلة فى مقابل عينات مرتبطة) واذا اتضح فروق فى احجام التأثير يجرى ما وراء التحليل منفصل.

وتقدر r من احصاء χ^2 كالتالى:

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

• N حجم العينة

حساب معامل الارتباط r من الاحصائيات الوصفية

بعض الدراسات تعرض الاحصائيات الوصفية مثل المتوسطات والانحرافات المعيارية واحجام العينات للمتغير تابع للمجموعتين ويمكن استخدامها لحساب r من خلال حساب قيمة فروق المتوسطات المعيارية ثم تحويلها الى r من خلال حساب مؤشر d او g ثم تحويلهما الى r من المعادلة التالية (Card, 2012):

$$r = \sqrt{\frac{g^2 n_1 n_2}{g^2 n_1 n_2 + (n_1 + n_2) df}}$$

بعض الدراسات البحثية تعرض النتائج فى ضوء جدول اقترانى لمتغيرين تصنيفين 2×2 وفى هذا الموقف يتم حساب قيمة معامل الارتباط فاي ويفسر على انه معامل ارتباط بيرسون r .

حساب معامل الارتباط من المستويات الاحتمالية (P) لاختبارات الدلالة

فى بعض الدراسات البحثية يذكر الباحث القيمة الاحتمالية دون عرض قيم الاختبارات الاحصائية، وقد يحدث هذا نتيجة عرض غير مناسب لنتائج الاختبارات مثل F, T . ويمكن حساب حجم التأثير من قيمة P الاحتمالية لاختبار ذو ذيلين من خلال البحث فى جدول الانحراف الاعتدالى المعيارى (Z).

فعلى سبيل المثال فى دراسة لاختبار ذو ذيلين قيمة $P=0.03$ وتكون لاختبار ذو ذيل واحد $P=0.016$ ، وبالبحث فى جدول Z فى كتب الاحصاء المتخصصة فان $Z=2.14$ ولكن هذه الجداول لاتتضمن القيم الصغيرة جداً لـ P ، ولذلك يجب البحث عن قيم Z المناظرة لـ P فى برامج الكمبيوتر مثل Excel من خلال تحديد مقلوب التوزيع التراكمى الاعتدالى المعيارى.

وبعد حساب Z المناظرة لـ P يمكن حساب حجم التأثير من خلال Z كالتالى:

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

فى كثير من الدراسات البحثية لا يذكر الباحث قيمة P بل يحدد مدى من الاحتمالات مثل $P < 0.05$. او اى قيم اخرى، وينصح (Card 2012) باستخدام بديلين فى هذه الحالة الاول: الاتصال المباشر بالمؤلف للحصول على معلومات تفيد فى حساب حجم التأثير مثل قيمة الاحصاء الاستدلالي F, T او القيمة الاحتمالية P ولكن هذ البديل احياناً غير عملى نتيجة عدم استعداد المؤلف للتعاون او عدم القدرة على التواصل مع المؤلف. فى هذه المواقف فان البديل الثانى حساب الحد الادنى لحجم التأثير المتلازم باعلى احتمال. بكلمات اخرى اذا كانت قيمة $P < 0.05$ وحجم العينة 100 ويمكن وضع مسلمة هى اعتبار $P=0.05$ وحساب قيمة Z المقابلة وهى 1.96، وعليه يمكن حساب r:

$$r = \frac{Z}{\sqrt{N}} = \frac{1.96}{\sqrt{100}} = 0.20$$

ويمكن اعتبار قيمة r هى الحد الادنى لحجم التأثير الفعلى فى الدراسة. فاذا كانت $P=0.03$ فان $r=0.22$ ، لو $P=0.01$ فان $r=0.26$.

معامل ارتباط سبيرمان الرتبى (ρ)

هو حالة خاصة من معامل ارتباط بيرسون ويستخدم عندما تكون المتغيرات ترتيبية، ويستخدم اختبار كندل تاو ومعامل ارتباط جاما لقياس حجم التأثير بين أزواج الرتب (Ferguson, 2009).

معامل ارتباط فاي (φ)

هو حالة خاصة من معامل ارتباط بيرسون ويستخدم كمقياس لحجم التأثير عندما يكون المتغيرين تصنيفين (2×2) حيث يستخدم اختبار كا² (χ^2) (Fern & Monroe, 1996) وتتحدد قيمة كالأتي:

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

ويستخدم مؤشر ϕ بكثرة في دراسات المجال الطبي ولكن يجب ان نتعامل معه بحذر ويرى (1993) Tatsuoka بأنه يمكن التعبير عن حجم التأثير بمربع ϕ :

$$ES(\phi)^2 = \frac{\chi^2}{N}$$

وإذا كانت قيمتها 0.01 فأكثر فهي ذات دلالة عملية.

مثال توضيحي لحسابه

اراد باحث دراسة العلاقة بين طبيعة العمل (يعمل - لا يعمل) والسعادة (سعيد - غير سعيد) لـ 40 فرد مشارك في الدراسة وفيما يلي توزيع افراد العينة :

المجموع	الحالة الوظيفية			
	يعمل	لا يعمل		
A	b	a	سعيد	السعادة
20	6(10)	14(10) *		
B	D	c	غير سعيد	
20	14(10)	6(10)		
N=40	D	C		المجموع
	20	20		

* بين القوسين التكرار المتوقع

ويقدر من المعادلة التالية:

$$r\phi = \phi = \frac{ad-bc}{\sqrt{ABCD}}$$

- a , b , c , d مجموع التكرارات في الخلايا.
- A , B , C , D مجموع التكرارات في الصفوف والاعمدة .

وتتحدد درجات الحرية كالتالى:

$$df = (c-1)(r-1) = (2-1)(2-1) = 1$$

وعليه تقدر ϕ كالتالى:

$$= \frac{14 \times 14 - 6 \times 6}{\sqrt{20 \times 20 \times 20 \times 20}} = 0.40$$

ولتحديد دلالة معامل ارتباط ϕ ، فأنتنا نحول معامل ارتباط ϕ الى احصاء χ^2 كالتالى:

تقدر قيمة χ^2 كالتالى:

$$\begin{aligned} \chi^2 &= \frac{(14-10)^2}{10} + \frac{(6-10)^2}{10} + \frac{(6-10)^2}{10} + \frac{(14-10)^2}{10} \\ &= 1.60 + 1.60 + 1.60 + 1.60 = 6.40 \end{aligned}$$

إذاً معامل ϕ كالتالى:

$$\phi = \sqrt{\frac{\chi^2}{N}} = \sqrt{\frac{6.40}{40}} = 0.4$$

وعلى ذلك فان ϕ و χ^2 مرتبطين ويمكن استخدام قيمة احدهما لتقدير قيمة الاخر. ويستخدم مؤشر ϕ كقوة للعلاقة بين المتغيرين التصنيفيين 2×2 . ومعامل الارتباط ϕ مماثل لحساب لمعامل ارتباط بيرسون r بين متغيرين تصنيفيين Y, X .

معامل ارتباط كرامير (V) Crammer (V)

يستخدم اذا زادت مستويات احد المتغيرات التصنيفية عن 2 مثل جداول χ^2 (3×3) أو (3×3) أو (2×2) أو غيرها وتتحدد بالاتي:

$$V = \sqrt{\frac{\chi^2}{N(K-1)}}$$

• K الحد الادنى من الصفوف أو الأعمدة.

ويري كثيراً من الباحثين أن التعبير عن حجم التأثير في ضوء مؤشر r أفضل نظراً لسهولة تفسيره وتفوقه على مؤشر d نظراً لأن مؤشر d ليس له معايير واضحة لحجم التأثير الصغير والمتوسط والكبير (Ledesma et al., 2009; Rosenthal et al., 2000)، وقد عدد McGrath & Meyer (2006) مميزات تحويل المؤشرات d و g و Δ إلى r بالاتي:

- العلاقة بين r والقوة الإحصائية للاختبار مباشرة وواضحة.
- أكثر مرونة ويمكن تقدير لأدنى مستوي قياس للمتغيرات سواء متصلة او منفصلة وهذا مفيد في دراسات ما وراء التحليل.
- مشترك في كل الاختبارات التي تتبع النماذج الخطية العامة.
- له معايير وأسس في التفسير واضحة.

ويمكن تقدير مؤشر r من خلال مؤشر g او d كالتالي (Rosenthal et al. 2000):

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

وإذا كانت $n_1 = n_2$ تصبح المعادلة السابقة كالتالي:

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

ويمكن ان تقدر قيمة r من المؤشر g كالتالي:

$$r = \sqrt{\frac{g^2 n_1 n_2}{g^2 n_1 n_2 + (n_1 + n_2) df}}$$

وتقدر قيمة r من قيمة الاختبار الإحصائي T كالتالي (Rosenthal et al. 2000):

$$r = \sqrt{\frac{T^2}{T^2 + df}}$$

ويقدر قيمة r من اختبار F لمجموعتين (Fern & Monroe 1996) كالتالي:

$$r = \sqrt{\frac{F}{F + df_{within}}}$$

• $df_{within} = N - C$ و C عدد المجموعات.

مؤشرات تقديرات المخاطرة

تستخدم في التصميم البحثي الذي يتضمن متغيراً مستقلاً تصنيفي بمستويين ومتغير تابع تصنيفي بمستويين، وتستخدم بكثرة في أبحاث المجال الطبي، وهو مؤشر مفيد للاقتران Contingency او العلاقة Association بين متغيرين تصنيفين Dichotomous (تحسن- لا تحسن، ناجح- راسب)، ولتوضيح مؤشرات المخاطرة يمكن عرض الجدول الآتي:

المجموعة	لا تحسن	تحسن
ضابطة	a	b
تجريبية	c	d

• a ، b ، c ، d تكرارات.

وتعتمد تقدير مؤشرات المخاطرة علي نسبة الأفراد في كل خلية. وكلمة Odds للحدث تعنى احتمالية او ترجيح حدوثه (الدرجة التاكيدية او التحسن) مقسوماً على احتمال حدوث البديل (الدرجة السلبية او اللاتحسن)، وهو احصاء لحجم التأثير تتم فيه مقارنة المجموعتين في ضوء نسبة احتمال Odds الحالة او الحدث (مرض، نجاح، التعرض لمخاطر) ويمكن التعبير عنه كالاتي:

$$\text{Odds} = \frac{P}{1-P} \text{ (احتمال وقوع الحدث)}$$

• P نسبة الافراد في العينة الذين لديهم الخاصية المرغوبة او التحسن. او احتمال الحدث.

فمثلاً خاصية مثل العدوانية لدى الاطفال حيث تصنف الى عدواني - لاعدواني، ونسبة الاطفال الذين لديهم عدوانية P وبالتالي يقدر احتمال او Odds ان يكون الطفل عدواني باحتمال $\left(\frac{p}{1-p}\right)$.

فمثلاً احتمال نجاح المعالجة 0.25 وعليه فان النسبة Odds للنتائج:

$$Odds = \frac{.25}{1 - .25} = .33$$

نسبة odds للنجاح نتيجة التعرض للمعالجة او البرنامج او الدواء هي 1 الى 3 بمعنى حالة نجاح في مقابل ثلاثة فشل. بينما احتمال نجاح المعالجة هي 1 في كل 4 حالات. افترض ان احتمال النجاح للمجموعة الضابطة 0.20 بالتالي فان نسبة Odds للنجاح او الشفاء او الاجتياز للمجموعة التي لا تتلقى معالجة هي:

$$Odds \text{ (الضابطة)} = \frac{.20}{1 - .20} = .25$$

بالتالي تحدث مقارنة للاحتمالات او Odds عبر المجموعتين، وهذه المقارنة تسمى نسبة Odds Ratio(OR) كالآتي:

$$OR = \frac{\text{Odds تجريبية}}{\text{Odds ضابطة}} = \frac{.33}{.25} = 1.33$$

فان احتمال الشفاء او النجاح او التحسن للمجموعة التجريبية يعادل 1.33 مرة اكبر من احتمال التحسن في المجموعة الضابطة.

ويمكن تقدير النسبة Odds كمؤشر لحجم التأثير من الصيغة التالية Lispey & Wilson (2001):

$$ES_{OR} = \frac{ad}{bc} = \frac{P_a P_d}{P_b P_c} = \frac{P_a/P_b}{P_c/P_d}$$

- d, c, b, a التكرارات في الخلايا.
- P_d ، P_c ، P_b ، P_a النسب في الخلايا.

والقيمة 1 تشير الى لا علاقة، القيمة المحصورة بين 0.0 و 1.0 علاقة سالبة، القيمة اكبر من 1.0 علاقة موجبة.

وكل تحليلات ما وراء التحليل فى ضوء نسبة الاحتمال OR كمؤشر لحجم التأثير تتم فى ضوء لوغاريتم نسبة Odds لان توزيع اللوغاريتمات يكون تقريباً اعتدالى بمتوسط 0.0 وانحراف معيارى 1.83 ، وتوجد ميزة اخرى جراء التعامل مع اللوغاريتمات هو يمكن حساب الانحراف المعياري له كالاتى:

$$ES_{LOR} = \log_e(ES_{LOR})$$

- ES_{LOR} حجم التأثير لـ OR اللوغاريتمى.
- \log_e التحويل اللوغاريتمى.

ويقدر الخطأ المعياري للوغاريتم حجم التأثير كالاتى:

$$SE_{LOR} = \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}$$

ثم تحول النتائج النهائية الى نسبة Odds كالاتى:

$$ES_{OR} = e^{ES_{LOR}}$$

- e اس base اللوغاريتم الطبيعي وهو تقريباً 2.718

وتوجد اشكالية خاصة باحتمال Odds كمؤشر لحجم التأثير فى ما وراء التحليل عندما يكون تكرارات احد الخلايا صف، وللتغلب على هذه الاشكالية يضاف 0.50 الى تكرارات الخلايا ولكن هذا يسبب تحيز .

مثال: اراد باحث اجراء ما وراء التحليل للدراسات التى هدفت الى دراسة فعالية برامج للتوعية بالاثار المترتبة بادمان المخدرات، واختار الباحث الدراسات التى اعتمدت على المقارنة بين المجموعة الضابطة والمجموعة التجريبية، وكان المتغير التابع نسبة او عدد الطلاب الذين تكون لديهم وعى بادمان المخدرات فى مقابل الذين لم ينمى لديهم الوعى بينما المتغير المستقل المجموعة التجريبية فى مقابل المجموعة الضابطة.

مثال: فاذا افترضنا ان بيانات متغيرين تصنيفين وهو المعالجة بالاسبرين (ياخذ - لا ياخذ) والاصابة بالازمة القلبية (الازمة - لا ازمة) كالاتى (Howell, 2013) :

		الناتج		
		لا تحسن	تحسن	
		ازمة قلبية	ازمة قلبية	
تناول الاسبرين	11037	10933(b)	104(a)	(تجريبية)
لا تناول اسبرين	11034	10845(d)	189(c)	(ضابطة)
	22071	21778	293	Σ

وفى الجدول السابق % 0.94 (104/11037) من الافراد فى المجموعة التجريبية (مجموعة الاسبرين) يعانون من ازمة قلبية وحدث لهم تحسن و (189/11034) % 1.71 فى المجموعة الضابطة(معالجة اخرى) يعانون من ازمة قلبية وحدث لهم تحسن.

تقديرات المخاطرة Risk estimates

تحدد مدى المخاطرة بمعنى اى فرد ياخذ او لا ياخذ الاسبرين سوف يعانى من ازمة قلبية. وهو مؤشر للكشف عن حجم التأثير لتناول الاسبرين وتوجد عدة مؤشرات لتقديرات المخاطرة اهمها:

فرق المخاطر (RD)Difference Risk

هو الفرق بين النسبتين الاتيتين:

$$RD = \frac{A}{(A+B)} - \frac{C}{(C+D)}$$

$$= 1.71\% - 0.94 = 0.77\%$$

اي حوالى ثلاثة ارباع النسبة المئوية هي فروق بين المعالجتين بالتالى الفرق فى المخاطرة بين الافراد الذين يتناولوا الاسبرين والآخرين الذين لم يتناولوا الاسبرين حوالى ثلاثة ارباع من الواحد الصحيح بالتالى لم تصل 1.0% وبمعنى ان حجم المخاطرة ليس كبير، بالتالى فان تناول الاسبرين له مخاطر محدودة. بالتالى نتوقع عدم وجود فروق كبيرة بين المعالجتين.

نسبة المخاطرة (RR) او المخاطرة النسبية Relative risk

وهى النسبة بين المخاطرتين:

$$RR = \frac{\text{المخاطرة بدون اسبرين}}{\text{المخاطرة مع الاسبرين}} = \frac{[A/(A+B)]}{[C/(C+D)]} = \frac{1.71\%}{0.94\%} = 1.819$$

وهذا المؤشر يحسب من خلال خارج قسمة النسبتين لمقدار التحسن فى المعالجتين او المستويين للمتغير الاسمى، وبالتالى المخاطرة فى حدوث ازمة قلبية جراء عدم تناول الاسبرين يعادل 1.8 مرة اعلى مما لو اخذ الاسبرين. وقيمة $RR = 0$ تشير الى لا فروق فى المخاطرة بين المجموعتين، القيمة اقل من 1.00 تشير الى مخاطرة اقل للمجموعة الضابطة مقارنة بالمجموعة التجريبية، والقيمة 2.00 تشير الى ان المجموعة الضابطة لديها احتمال مخاطرة مضاعف من المجموعة التجريبية. فمثلاً نسبة المخاطرة بالنسبة للمدخنين للاصابة بسرطان بالرئة 23 مرة وهذا يعنى ان الذكور المدخنين معرضين بالاصابة بالرئة باحتمال 23 مرة مقارنة باقرانهم غير المدخنين. ويشير Ferguson (2009) ان قيمة RR بين 1.00 و 2.00 ليست ذات اهمية عملية كبيرة، وتختلف تفسيرها حسب مجال اهتمام الدراس، فربما تكون لها معنى اذا كان حد المخاطرة 1.0% وقد تكون لها معنى اذا كان حد المخاطرة 10% (0.1) وهذا المؤشر يفضل استخدامه لبيانات تصنيفية خاصة مع متغيرين اسميين.

نسبة الاحتمال او الترجيح Odds Ratio(OR)

احتمال Odds للمجموعة التجريبية هي عدد الذين اصابوا بازمة قلبية وحدث لهم تحسن مقسوما على العدد الكلى من الافراد فى هذه المجموعة التى تأخذ الاسبرين وهى الاحتمالية الاتية:

$$\frac{A}{A+B} = \frac{104}{11037} = 0.0094$$

بينما Odds للاصابة بالازمة القلبية فى المجموعة التى لا تتلقى الاسبرين هى عدد الذين اصيبوا بازمة قلبية وحدث لهم تحسن مقسوماً على عدد الذين لم الذين لم يحدث لهم تحسن فى نفس المجموعة

وتقدر نسبة Odds Ratio (OR) كالتالى:

$$OR = \frac{\text{Odds}_{\text{لااسبرين}}}{\text{Odds}_{\text{اسبرين}}}$$

$$OR = \frac{1.17}{.0094} = 1.83$$

بالتالى احتمالية حدوث ازمة قلبية بدون تناول اسبرين هى اعلى 1.83 مرة من احتمالية حدوث ازمة قلبية مع تناول الاسبيرين. وكما هو ملاحظ ان قيمة المخاطر النسبية (RR) تتقارب وتعطى نتائج متماثلة مع OR .

ويفضل (Howell (2013) استخدام احصاء OR لانه يمكن حسابه فى مواقف معينة يصعب فيها حساب نسبة المخاطرة الحقيقية RR. ويمتاز احصاء OR بميزة على درجة كبيرة من الاهمية وهو ان يمكن حساب Log (OR) ليعطى احصاء مفيد فى مواقف مختلفة ويمكن ان تكون قيمته قريبة من قيمة معاملات الانحدار اللوجسية. او تستخدم فى النماذج الخطية اللوغارتمية Log-linear models، و اذا كانت قيمة $OR \leq 3$ يعتبر حجم تأثير قوى نسبيا. وعرض (Ferguson (2009) حدود لحجم التأثير لبيانات العلوم الاجتماعية لمؤشرى المخاطرة النسبية (RR) ونسبة Odds كالتالى: الحد الادنى لحجم التأثير العملى 2.00 ، حجم تأثير متوسط 3.00 ، حجم تأثير كبير 4.0 فاكثر.

واخيراً علق (Howell (2013) على مؤشرات حجم التأثير الفروق (d) ومؤشرات العلاقة (r) لاحصاء χ^2 لجداول 2×2 فيقول أننى غير راضى باستخدام مؤشرات (r، ρ) كرامير) لانها ليست لها تفسيرات ذات معنى فى معظم المواقف البحثية. فمثلاً احتمال ان تتعرض للازمة القلبية اذا لم تاخذ الاسبرين هى اعلى 1.83 مرة من الاصابة بالازمة القلبية لو لم تاخذ

الاسبرين فهذه العبارة لها معنى مفهوم لدى معظم الناس، ولكن ماذا تعنى العلاقة بين تناول الاسبرين والازمة القلبية 0.034 فانها لا تخبرنا عن شىء مفيد لنا. او ان اخذ الاسبرين يسهم فى تفسير 0.1% من تباين حدوث الازمة القلبية وينصح بعدم الاعتماد على مؤشرات العلاقة كمقياس لحجم التأثير لاحصاء χ^2 او لجدول 2×2 ما لم يوجد سبب قوى لهذا الاستخدام .

تفسير حجم التأثير

يوجد القليل من التوجهات لتفسير حجم التأثير ولا توجد معايير واضحة ومتفق عليها لكل مؤشر لتفسيره ومن أهم هذه المعايير:

معايير Cohen (1988)

قدم كوهين معايير لتفسير معظم مؤشرات حجم التأثير كالتالى:

جدول (2.5): معايير Cohen (1988) لتفسير مؤشرات حجم التأثير.

المؤشر	صغير	متوسط	كبير
d, Δ , g	0.2	0.5	≥ 0.8
مؤشر f	0.10	0.25	≥ 0.40
مؤشر r	0.10	0.30	≥ 0.50
R^2_{adj} , r^2 , R^2 , W^2 , η^2 , ε^2	0.01	0.06	≥ 0.14
OR	1.5	2.5	4.3

معايير وردت في Ferguson (2009)

الجدول (3.5): معايير وردت في Ferguson (2009) لتفسير حجم التأثير.

المؤشر	صغير	متوسط	كبير
الفروق بين المجموعات d, Δ, g	0.41	1.15	2.70
مؤشر العلاقة (r, R, ϕ, σ, B)	0.20	0.50	0.80
التباين المفسر $r^2, R^2, \omega^2, \mu^2, \varepsilon^2$	0.04	0.25	0.64
تقديرات المخاطرة (RR, OR)	2.0	3.0	4.0

ووضع Lipsey (1998) معايير لمؤشرات الفروق بين المجموعات وهي القيمة 0.15 صغير، 0.45 متوسط، و 0.90 فأكثر كبير. ووضع Hinkle, Wiersma, & Jurs (1994) معايير لمعامل ارتباط r 0.30 منخفض، 0.70 متوسط، و 0.90 فأكثر كبير.

وفيما عرض للحدود المتفق عليها لمؤشرات حجم التأثير:

الجدول(4.5): حدود حجم التأثير لمؤشرات بعض الاختبارات.

الاختبار	المؤشر	صغير	متوسط	كبير
T المستقلة	d, Δ, g	.20	.50	.80
الارتباط	(r^2) r	(.01).10	(.09).30	(.25).50
جدول الاقتران	W, V, φ, c	.10	.30	.50
ANOVA	f	.10	.25	.40
	ω^2, μ^2	.01	.06	.14
الانحدار المتعدد	f^2	.02	.15	.35
	R^2	.02	.13	.26

وهذه المعايير أو الإرشادات ما هي الا اجتهادات بحيث لا تطبق بطريقة جامدة (Cohen, 1994) ويتفق معه (Thompson 1998) حيث أشار الى انه في حالة التعامل مع حجم التأثير في ضوء معايير جامدة فإننا نعيد تجربة الدلالة الإحصائية ونكون غير موفقين مرة أخرى في التعامل مع القرارات الإحصائية، وأشار (Pedhazur & Schmelkin 1991) بأن القيمة العالية لحجم التأثير عند باحث ما في وقت ما لتخصص ما تختلف عند باحث آخر في تخصص آخر. ومن الخطأ القول ان حجم التأثير الكبير في القيمة ذو دلالة عملية وان حجم التأثير الصغير في القيمة يعتبر تافهة وليس له اي معني (Prentice & Miller, 1992)، ويعتمد تفسير حجم التأثير أيضاً علي الظروف التي تتم فيها الدراسة (Dulark, 2009). ويؤكد (Stout & Ruble 1990) على وجود عوامل اخري عند تفسير حجم التأثير مثل تحليل الكلفة - المنفعة كمحك خارجي للحكم ما اذا كان حجم التأثير ذو دلالة عملية ام لا.

وأوضح Iversson, Anderson, Johnson & Lindwall (2012) لمساعدة الباحثين علي تفسير المعنوية (الدلالة الاكلينكية) يتم استخدام مؤشرات عديدة منها حدود الثقة لإحجام التأثير Confidence Interval (CI) واحتمالية الافضيلة Comon language Probability Superiority (PS) من خلال مؤشر .effect Size (CL).

وأشار Hill, Bloom, Black & Lipsey (2007) إلى ان تفسير حجم التأثير للدراسة لابد ان يتم في ضوء مرحلة النمو للمجتمع المستهدف (روضة - ابتدائي - ثانوي - جامعي)، وفي ضوء المحتوي المقاس، فتفسير حجم التأثير لمحتوي الرياضيات مختلف عن تفسير حجم التأثير لمحتوي اللغة العربية، وأيضاً يفسر حجم التأثير علي أساس المجموعات الفرعية فمثلا تفسيره لموضوع ما علي الذكور يختلف عن تفسيره عند الإناث، وأيضاً بالنسبة للموقع الجغرافي وذلك نظراً لوجود فروق بين المجتمعات، ويفسر أيضاً في ضوء التراث البحثي المشابهة للظاهرة موضوع الدراسة وهذا يفيد تماماً مجال دراسات ما وراء التحليل.

ويؤكد Frohlich, Emrich, Pieter & Stark (2009) علي ان الدراسات الامبريقية أوضحت أن حجم التأثير يجب ان لا يفسر بصورة جامدة استاتيكية Statically بل بصورة ديناميكية dynamically ويكون وظيفة لمدة البرنامج والخبرة السابقة وأبحاث المجال، ويمكن ان يكون لأحجام التأثير الصغيرة معني وهذا يتوقف علي مجال دراسة الظاهرة.

وأشار Hill et al. (2007) إلى أن وجود قاعدة وحيدة لتفسير حجم التأثير مثل معايير Cohen وغيرها لا تكون مناسبة لكل المجالات.

الفصل السادس

احجام تاثير متنوعة ومتقدمة

Advanced and Effect Sizes

مقدمة

فى الفصل السابق تعرضنا لاهم ثلاثة انواع من حجم التاثير وهى معامل الارتباط r ، مؤشرات فروق المتوسطات المعيارية (d , g)، ومؤشر نسبة Odds(OR) بالضافة الى احجام تاثير اخرى، ولكن توجد مجموعة اخرى من مؤشرات حجم التاثير تناسب اسئلة بحثية معينة لا تناسب معها المؤشرات السابقة ولكنها اقل استخداماً فى تراث ما وراء التحليل مثل المؤشرات المستخدم مع نواتج دراسة ذات متغير وحيد او مؤشرات مستخدمة مع تحليلات الاحصاء المتدرج مثل تحليل الانحدار وتحليل التباين المتدرج MANOVA او فى التحليلات بحوث الطولية Longitudinal Researches .

مؤشرات حجم التاثير لاحصاء المتغير الواحد

توجد القليل من دراسات ما وراء التحليل تتعامل مع متغير وحيد واحصائيات المتغير الواحد تتضمن:

- وصف للقيمة النموذجية او المنتصفية وتتمثل فى مؤشرات النزعة المركزية مثل المتوسط، الوسيط، والنوال للمتغير المتصل.
- مقدار الاختلاف بين قيم المتغير متمثلة فى مقاييس التشتت مثل المدى، الانحراف المعيارى، والتباين للمتغير المتصل.
- نسب الافراد فى اى من تصنيفات المتغير التصنيفى.

من النادر اجراء دراسة ما وراء التحليل لمتغير وحيد ولكن لاجراء هذا النوع من التحليل فلا بد من توفر شرطين اساسين (Lipsey & Wilson, 2001):

- كل الدراسات التى تتضمن المتغير يكون لها نفس التعريف الاجرائى للمتغير او يوجد تشابه بدرجة كبيرة للقياسات عبر الدراسات مثل الدراسات اعتمدت على اختبار ذكاء محدد ومعروف مثل ستانفورد بينيه. ولكن يمكن اجراء ما وراء التحليل لنفس المفهوم ولكن بتعريفات اجرائية مختلفة مثل الدرجات على اختبارات تحصيلية مختلفة ولكن بعد معايرتها اى التعبير عنها فى ضوء الدرجات المعيارية حتى تتم المقارنة بينها.
- من الضرورى تحديد حجم التأثير المناسب الذى يتضمن المعلومات موضع الاهتمام وكذلك تحديد الخطأ المعيارى المرتبط بحجم التأثير.

متوسط المتغير Mean

ما وراء التحليل لمتوسطات المتغير الواحد له قيمة كبيرة خاصة:

- **دمج** الدراسات للحصول على تقدير اكثر دقة لقيمة المتوسط لمتغير ما عبر الدراسات المختلفة خاصة عندما تتضمن هذه الدراسات احجام عينات صغيرة او تنوع التعريفات الاجرائية للمفهوم وذلك بعد معايرتها.
- **المقارنة** بين متوسطات الدراسات لتحديد الاسباب المحتملة لاختلاف المتوسطات عبر الدراسات لاسباب منهجية مثل خصائص العينة (الموقع او العمر او غيرها).

ويوجد محدد اعتبارى مهم عند اجراء ما وراء التحليل لمتوسطات الدراسات المختلفة هو ان يكون بنفس وحدة القياس. ويوجد مدخلين للتعامل مع هذه القضية:

- التعامل مع الدرجات المعيارية (Lipsey & Wilson, 2001).
- تحويل متوسطات الدراسات ذات المقاييس المختلفة الى وحدة قياس عامة، فمثلاً دراسة تضمنت مقياس تصحيحه على مقياس 0-4، ودراسة اخرى تضمنت مقياس تصحيحه 1-10 فمن الممكن تحويل متوسط احد المقياسين ليكافئ متوسط المقياس الاخر كالتالى (Card, 2012):

$$X_2 = (X_1 - Min_1) \left(\frac{Max_2 - Min_2}{Max_1 - Min_1} \right) + Min_2$$

- X_2 الدرجة المكافئة على المقياس الثانى.
- X_1 الدرجة على المقياس الاول التى ترغب فى تحويلها.
- Max_1, Min_1 الدرجة الدنيا والقصى فى المقياس الاول.
- Max_2, Min_2 الدرجة الدنيا والقصى فى المقياس الثانى.

ويبدو ان تطبيق هذه المعادلة عملية صعبة نظراً لان الدراسات المنشورة غالباً لا تتضمن القيم الدنيا و القصى للمتغير وعلى ذلك فان الحل الذى طرحه Lipsey & Wilson (2001) اكثر واقعية.

ويرى (2012) Card حتى لو كانت المقاييس المختلفة فى الدراسات استخدمت نفس التدرج التصحيحي مثل تدرج (0 - 4) فمن الممكن ان يختلف طبيعة التدرج من دراسة الى اخرى، ففى دراسة يكون دائماً، غالباً، احياناً، نادراً، على الاطلاق، بينما فى دراسة اخرى ولامرة، مرة، 2-3 مرة، 4-6 مرات، 7 فاكتر، فى هذه الحالة ليس من المعقول اجراء ما وراء التحليل للدمج او المقارنة لهذة النوعية من نتائج الدراسات.

ويقدر المتوسط كمؤشر لحجم التأثير كالاتى:

$$ES_m = \bar{x} = \frac{\sum x_i}{N}$$

- $\sum x_i$ مجموع درجات افراد العينة.
- N حجم العينة.

وليس من الضرورى حساب المتوسط حيث عادة يكون معروض فى الدراسة. وفى بعض الدراسات يكون معروض الجدول التكرارى حيث 5 افراد حاصلين على الدرجة التحصيلية 10، و 7 افراد حاصلين على الدرجة 15 وهكذا وعلى ذلك يقدر المتوسط من الجدول التكرارى كالاتى:

$$ES_m = \bar{x} = \frac{\sum Xf}{\sum f}$$

- $\sum Xf$ مجموع حاصل ضرب الدرجة فى التكرار المقابل.
- $\sum f$ مجموع التكرارات (حجم العينة).

ولابد من حساب الخطأ المعياري لاستخدامه فى عملية الموازنة فى ما وراء التحليل وكذلك لاجراء اختبارات الدلالة الاحصائية وحساب فترات الثقة ويقدر كالاتى
(Lipsey & Wilson, 2001):

$$SE_m = \frac{S}{\sqrt{N}}$$

- S الانحراف المعياري المعياري
- N حجم العينة

وبعد حساب المتوسط والخطأ المعياري لكل متوسط فان الباحث يبدأ فى اجراء ماوراء التحليل.

نسب المتغير Proportions

دائماً تعرض احصائيات المتغير الواحد فى ضوء نسب مئوية لعدد الافراد لكل تصنيف من تصنيفات المتغير، وتعتبر هذه النسب مؤشراً لحجم التأثير وتتراوح قيمتها من 0.0 الى 1.0. فمثلاً نسبة الافراد مدمنى الانترنت فى عينة الدراسة 40%. راعى ان النسبة تكون لنفس المتغير عبر الدراسات المختلفة. ويوجد مدخلين لحساب حجم التأثير فى هذه الحالة حددهما Lipsey & Wilson(2001) بالاتى:

اولاً: الاعتماد على النسب المئوية مباشرة وتقدر كالتالى:

$$ES_p = P = \frac{K}{N}$$

- K عدد افراد فى التصنيف موضع اهتمام الدراس.
- N حجم العينة.

بينما الخطأ المعياري للنسب كالاتى:

$$SE_p = \sqrt{\frac{P(1-p)}{N}}$$

واثبتت دراسات المحاكاة لتقديرات متوسط مجتمع النسب عبر الدراسات المختلفة يكون مناسب في ماوراء التحليل ولكن يعطى تقديرات منخفضة لفترات الثقة حول متوسط حجم التأثير (النسب) وتقديرات عالية لدرجة عدم التجانس عبر الدراسات خاصة عندما تكون النسب المقدرة اقل من 0.20 او اعلى من 0.80 ولو ان متوسط النسب عبر الدراسات المختلفة بين 0.20 الى 0.80 هو موضع اهتمام دراسة ما رواء التحليل فان استخدام النسب مباشرة كحجم التأثير يكون مناسب (Lipsey & Wilson, 2001)

ثانياً: تحويل النسب المئوية الى لوغاريتمات اذا كان الهدف من وراء التحليل دراسة الاختلاف حول متوسط النسب فمن الافضل استخدام الطريقة اللوغاريتمية حيث يحول النسب الى اللوغاريتمات وتعامل كأنها حجم تأثير ثم تحول النتائج النهائية الى نسب للمساعدة في التفسير، وعكس النسب فان قيم اللوغاريتمات تأخذ اى قيمة رقمية.

ويقدر حجم التأثير اللوغاريتمى كالاتى (Card, 2012, Lipsey & Wilson, 2001):

$$ES_{log} = L = \ln\left(\frac{P}{1-P}\right)$$

• P نسبة الافراد فى التصنيف موضع الاهتمام

ويقدر الخطأ المعياري للوغايتم النسب كالتالى:

$$SE_{log} = \sqrt{\frac{1}{NP} + \frac{1}{N(1-P)}}$$

وتتم تحليلات ما وراء التحليل باستخدام اللوغاريتم بعد موازنة الخطأ المعياري ثم تحول النتائج النهائية اللوغاريتمية (متوسط حجم التأثير اللوغاريتمية) الى نسب مرة اخرى حتى يسهل تفسيرها كالتالى:

$$P = \frac{e^{logit}}{e^{logit} + 1}$$

• e اساس اللوغاريتم الطبيعي وتقريباً 2.7183

مثال: افترض ان دراسة ما وراء التحليل لتحديد نسبة ادمان الانترنت بين طلاب المرحلة الثانوية فى مناطق مختلفة فيجرى التحليل على النسب مباشرة او تحويلها اللوغاريتمية المقابل كاحصائيات لحجم التأثير ثم موازنة احجام التأثير عن طريق الخطأ المعياري ثم التحقق من مدى تجانس من خلال احصاء التجانس (انظر الفصل الثامن) ثم حساب المتوسط الموزون واختبار الدلالة الاحصائية حوله وحساب فترات الثقة.

التباينات والانحرافات المعيارية Variances and Standard Deviations

قليل من دراسات ما وراء التحليل تستخدم الانحراف المعياري او التباين (مربع الانحراف المعياري) كمؤشرات لحجم التأثير. ويقدر الانحراف المعياري من الدرجات الخام كالتالى:

$$S = \sqrt{S_x^2} = \sqrt{\frac{(X_i - \bar{X})^2}{N-1}}$$

• X_i درجة فرد ما.

• \bar{X} متوسط الدرجات.

• N حجم العينة.

هذه المعادلة تقدير غير متحيز للانحراف المعياري فى المجتمع، ولكن عند حساب الانحراف المعياري فى العينة يكون المقام N .

ولاجراء ما وراء التحليل لدمج او مقارنة الانحرافات المعيارية او التباينات عبر الدراسات فلا بد من حساب الخطأ المعياري لاستخدامة في الموازنة كالتالى:

$$SE_s = \frac{S}{\sqrt{2N}}$$

بينما الخطأ المعياري للتباين:

$$SE_{s^2} = \frac{S^2}{2N}$$

ودراسات ما وراء التحليل التى تعتمد على الانحرافات المعيارية لابد ان تعتمد على نفس المقياس لنفس المفهوم. ويقترح (Pigott & Wu (2008) ان لا تجرى ما وراء التحليل لانحرافات المعيارية لدراسات تتضمن احجام عينات اقل من 25.

احصاء المتغيرين

تناولنا فى الفصل السابق احجام التأثير الخاصة بمتغيرين سواء اختبار فروض حول فروق متغيرين للقياسات المستقلة تحت ما يسمى بمقارنة المجموعات Group contrast وفى هذا الجزء نتناول:

معامل ارتباط Point bi-serial (r_{pb}) (يسمى أحياناً معامل الارتباط الثنائي الأصيل): أحياناً تكون النواتج البحثية شكل من العلاقة بين متغير تصنيفى ثنائى بطبيعته (ذكر - انثى، تجريبية - ضابطة) ومتغير متصل (درجات الاختبار او المقياس) وهذا النوعية من العلاقات يمكن عرضها فى اما صورة حجم تأثير فروق المتوسطات المعيارية (g, d) او فى صورة مع معامل ارتباط يسمى معامل الارتباط الثنائى ولكنه معرض لضيق مدى المتغير التصنيفى، وعليه لا يتوفر لة خصائص مرغوبة خاصة عندما تكون النسبة 80-20 او اقل، ولذلك فان مؤشرات فروق المتوسطات مفضلة عليه لاجراء ما وراء التحليل.

ومعامل الارتباط الثنائى حالة خاصة من r ويطبق عندما يكون المتغير المستقل تصنيفى بمستويين والتابع متصل وتتحد قيمته كالتالى:

$$r_{pb} = \frac{\bar{Y}_1 - \bar{Y}_2}{S_Y} \sqrt{\frac{n_1 n_2}{N}}$$

- \bar{Y}_1 متوسط درجات المتغير التابع للمجموعة الاولى ($X = 0$).
- \bar{Y}_2 متوسط درجات المتغير التابع للمجموعة الثانية ($X = 1$).
- S_Y الانحراف المعياري للمتغير التابع للعينة الكلية.

ويستخدم في حالة التصميم التجريبي الذي يتعامل مع مجموعتين وقياس بعدي ومتغير متصل تابع واحد ومربع r_{pb} هي تقريباً قيمة η^2 . ويمكن تقدير r_{pb} لتحليل التباين ANOVA لمجموعتين:

$$r_{pb}^2 = \frac{SS_{treatment}}{SS_{total}} = \frac{SS_c}{SS_t}$$

ولاغراض التفسير يمكن تحويل مؤشر هيدجز g الى معامل ارتباط r_{pb} كالتالى:

$$r_{pb} = \frac{g}{\sqrt{(1/P)(1 - P) + g^2}}$$

• P نسبة الافراد فى المجموعة الاولى.

• $1 - P$ نسبة الافراد فى المجموعة الثانية.

ولو عدد الافراد فى كل مجموعة متساوى: $1 - P = P$ فان الصيغة السابقة تصبح كالتالى:

$$r_{pb} = \frac{g}{\sqrt{4 + g^2}}$$

ويمكن حساب مؤشر g من معامل الارتباط r_{pb} كالتالى (Lipsey & Wilson, 2001):

$$g = \frac{r_{pb}}{\sqrt{P(1 - P)(1 - r_{pb}^2)}}$$

وفى حالة تساوى نسب المجموعتين تكون الصيغة السابقة كالتالى:

$$g = \frac{2r_{pb}}{\sqrt{(1 - r_{pb}^2)}}$$

مقارنة القياسات القبلية والبعدية Pre-post contrast

هذا التصميم يتضمن مقارنة متغيرين متصلين عبر زمن القياس ودائماً تتم المقارنة لمتوسط اونسب نفس المتغير عبر قياستين متتاليتين لدراسة التغير او التحسن عبر الزمن لدرجات العينة مثل قياس التحسن فى تحصيل التلاميذ بعد نهاية العام الدراسى مقارنة ببدايته. ولتحليل هذه النوعية من الدراسات لابد ان يتضمن نفس المفهوم ونفس التعريف الاجرائى.

وتستخدم الفروق غير المعيارية بين متوسطى القياس القبلى والبعدى مباشرة كمؤشر لحجم التأثير ولكن اذا تضمن القياس البعدى والقبلى مقياس مختلفة التعريف الاجرائى فلا بد من استخدام فرق المتوسطات المعيارية كمؤشر لحجم التأثير وفيما يلى عرض لمؤشرات حجم التأثير لهذه النوعية من الدراسات:

كسب المتوسطات غير المعيارية Unstandardized Mean Gain: يستخدم عندما تكون كل نتائج الدراسات تستخدم نفس المفهوم بنفس التعريف الاجرائى مثل مقياس Beck للاكتئاب او مقياس العوامل الخمسة الكبرى للشخصية والبيانات المتحصل عليها بداية ونهاية العام الدراسى على اختبار الرياضيات او القراءة المعيارى. ويقدر حجم التأثير كالاتى:

$$ES_G = \bar{X}_{T2} - \bar{X}_{T1} = G$$

- \bar{X}_{T2} متوسط القياس البعدى.
- \bar{X}_{T1} متوسط القياس القبلى.
- G الفرق بين القياس القبلى والبعدى ويسمى الكسب

ويقدر الخطأ المعيارى لهذا المؤشر كالتالى (Lipsey & Wilson, 2001):

$$SE_G = \sqrt{\frac{2S_p^2(1-r)}{N}} = \sqrt{\frac{S_G^2}{N}}$$

- S_p^2 مربع التباين المدمج
- r معامل ارتباط بيرسون بين القياس القبلى والبعدى.
- N حجم العينة.

- S_G^2 تباين درجات الكسب (الفرق بين القبلى والبعدى) ويقدر كالاتى:

$$S_G^2 = \frac{S_{T1}^2 + S_{T2}^2}{2}$$

- S_{T1}^2 تباين درجات القياس القبلى.

- S_{T2}^2 تباين درجات القياس البعدى.

والاحصائيات الاستدلالية من الفروق قبلى وبعدى يمكن الحصول عليها من خلال اختبار T المرتبطة او تحليل التباين للقياسات المرتبطة.

درجات التغير الطولى Longitudinal Change Scores

التغيرات الطولية على درجة كبيرة من الاهمية فى الدراسات النمائية حيث تعطى كثيراً من الانتباه للتغيرات التى تحدث للمتغيرات او الظاهرة عبر الزمن وتدرس من خلال تصميمات طولية طبيعية Naturalistic Longitudinal Designs وايضاً فى الدراسات التجريبية وشبه التجريبية قياسات قبلية وبعديّة فيمكن استخدامها لدراسات ما وراء التحليل للدمج او المقارنة بين الدراسات.

وفى هذه الدراسات يتم التعامل مع الزمن كمتغير مستقل تصنيفى وتعرض التغير فى الدرجات كمتوسط معيارى اذا استخدمت الدراسات مقاييس مختلفة لقياس المتغير التابع اومتوسط درجات التغير غير المعيارية لو اعتمدت الدراسات على نفس المقياس.

ويقدر حجم التأثير لمتوسط درجات التغير المعيارية كالاتى (Lipsey & Wilson, 2001):

$$g_{change} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2 + S_2^2}{2}}}$$

- \bar{X}_1, \bar{X}_2 متوسطات فى التطبيق الاول والثانى عبر الزمن.

- S_1^2, S_2^2 تباين درجات التطبيق الاول والثانى.

وهذه المعادلة مشابهة لمعادلة حجم التأثير g.

ويمكن حساب g_{change} من قيمة اختبار T المرتبطة او F المرتبطة لمجموعتين كالاتى (Card, 2012):

$$g_{change} = \frac{T_{dependent} \sqrt{2(1-r)}}{\sqrt{N}} = \sqrt{\frac{F_{repeated} \sqrt{2(1-r)}}{N}}$$

• $T_{dependent}$ ، $F_{repeated}$ قيمة الاختبار المحسوبة.

• N حجم العينة.

• r معامل ارتباط بيرسون بين قياسات الفوج الاول والثاني.

وهذه المعادلة تسمح بحساب g_{change} من المستويات الاحتمالية P .

وبالاضافة الى حجم التأثير فلا بد من حساب الخطأ المعياري لـ g_{change} لاجراء الموازنة

لاحجام التأثير واجراء اختبارات الدلالة الاحصائية كالتالي:

$$SE_{g_{change}} = \sqrt{\frac{2(1-r)}{N} + \frac{g_{change}^2}{2N}}$$

• r معامل الارتباط (الاستقرار) للمتغير عبر الزمن.

• N حجم العينة.

ويستخدم هذا المدخل مع اختبارات T المرتبطة و ANOVA للقياسات المتكررة لمجموعتين

وعلى ذلك يمكن استخدام مع لدراسات التجريبية التي تتضمن مجموعة ضابطة وتجريبية.

مؤشرات نسبة التباين المفسر لتصميمات تحليل التباين احادى الاتجاه والمتعدد

هي المؤشرات التي تقيس نسبة التباين المفسر في المتغير التابع نتيجة المتغير المستقل ومن أهم

هذه المؤشرات:

مؤشر مربع ايتا (η^2): هذا المؤشر مقروناً بتصميم تحليل التباين ويقدر كالتالي

(Richardson, 2011):

$$\eta^2 = \frac{SS_b}{SS_T}$$

• SS_b مجموع المربعات بين المجموعات و SS_T مجموع المربعات الكلي.

وهذا المؤشر يعطي تقديراً للعلاقة بين متغير المستقل والمتغير التابع في العينة

(Richardson, 2011; Tatsuoka, 1993). ويرى Ferguson (2009) أن η^2

هو مؤشر لقوة العلاقة بين المتغيرات مثل مؤشرات r ، R ، R^2 ، ϕ ، r_{pb} ، ولكن مؤشر

η^2 مفضل عن مؤشر r^2 عندما تكون العلاقة بين المتغيرات شبه منحنية Curvilinear وتكون قيمته أعلى قليلاً من r^2 (Tatsuoka, 1993).

وأشار Richardson (2011) إلى ان الفرق بين η^2 و r^2 مساوي لمربع معامل الارتباط المتعدد وفي حالة اذا كان المتغير المستقل بمستويين فإن قيمة η^2 مساوية لمؤشر r_{pb} (Tatsuoka, 1993). ويمكن تقدير مؤشر η^2 من قيمة اختبار F كالتالى:

$$\eta^2 = \frac{(C - 1) - F}{(C - 1)(F + (N - C))}$$

• C عدد مستويات المتغير المستقل.

لا تتم مقارنة مؤشر η^2 عبر الدراسات المختلفة لانه متحيز تحيز موجب ويكون التحيز فى ادناه مع احجام العينات الكبيرة (Grissom & Kim,2005; Olejnik & Algina, 2003; ولذلك تم تصحيح مؤشر η^2 كالتالى:

$$\eta^2 = \frac{SS_b}{SS_w + SS_T}$$

حيث SS_w مجموع مربعات داخل المجموعات (الخطأ).

ويمكن تقدير مؤشر η^2 من اختبار T كالتالى (Richardson, 2011):

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

• N حجم العينة الكلى (n_1+n_2)

والعلاقة بين r و η كالتالى:

$$\eta^2 = r^2$$

ويمكن تقدير η^2 من مؤشر d لكوهين كالتالى:

$$\eta^2 = \frac{d^2}{d^2 + 4}$$

ويري (2009) Ferguson أن الفرق بين مؤشري η^2_p و η^2_p ضئيل جداً مع العينات الكبيرة وأحياناً يخطئ الباحثون ويعتبروا ان مؤشر η^2_p هو نفسه مؤشر η^2 وذلك في مخرج SPSS (Levine & Hullet, 2002).

مؤشر ابيسون ε^2 لـ Kelly (1935): يقدر هذا المؤشر نسبة التباين المفسر لتصميمات ANOVA ويعالج محددات مؤشر η^2 (Fern & Monroe, 1996)، وهو صيغة مصححة لمؤشر η^2 ويتحدد بالصيغة الآتية (Ferguson 2009, Fern & Monroe, 1996):

$$\varepsilon^2 = \frac{SS_b - (K - 1)MS_w}{SS_T}$$

• MS_w متوسط مجموع مربعات الخطأ او داخل المجموعات.

ويعتبر المؤشر ε اقل تحيزاً من η^2 وقيمه أقل منه أو تساويها وهو مؤشر لتقدير حجم التأثير في المجتمع، لذلك فهو مؤشر استدلالي (Richardson, 2011).

مؤشر (f) Cohen: يقدر نسبة التباين المفسر في حالة اختبار تحليل التباين الأحادي ويمكن تقدير قيمة f كالتالي (Cohen, 1988):

$$f = \sqrt{\frac{\omega^2}{1-\omega^2}}$$

ويمكن تقدير مؤشر f من قيمة اختبار F كالتالي:

$$f = \sqrt{\frac{(C - 1)(F - 1)}{NC}}$$

والعلاقة بين مؤشر f و η^2 كالتالي (Grissom & Kim, 2005):

$$f = \frac{\eta}{\sqrt{1 - \eta^2}}$$

مؤشر اوميغا 2 (ω^2) Hayes (1963): يعتبر أكثر مؤشرات قياس نسبة التباين المفسر انتشاراً وذلك في التصميمات ذات العامل أو المعالجة الواحدة في ANOVA (Fern & Monroe, 1996)، وأكثر تفضيلاً من مؤشر η^2 لتقدير حجم التأثير. وعلي الرغم من

أهمية مؤشري ω^2 و ε^2 كتصحيح لمؤشر η^2 إلا أن Kirk (2001) يري محدودية استخدامهم في الدراسات ولكن Ferguson (2009) يري أن مؤشري ω^2 و ε^2 أكثر استخداماً في تصميمات ANOVA ، وأن مؤشر ω^2 مثل مؤشر ε^2 يقدر حجم التأثير في المجتمع ولذلك فهو مؤشر استدلالي عكس مؤشر η^2 .

وتحدد قيمة مؤشر ω^2 كالتالي (Olejnik & Algina, 2000):

$$\omega^2 = \frac{SS_b - MS_w}{SS_T - MS_w}$$

بينما عبر عنها (Fern & Monroe, 1996; Grissom & Kim, 2005) بالتالي:

$$\omega^2 = \frac{SS_b - (k - 1)MS_w}{SS_T + MS_w}$$

ويقدر مؤشر ω^2 من اختبار T كالتالي (Hayes, 1988):

$$\omega^2 = \frac{T^2 - 1}{T^2 + n_1 + n_2 - 1}$$

ويقدر مؤشر ω^2 من قيمة اختبار F للمقارنة بين مجموعتين كالتالي:

$$\omega^2 = \frac{(c - 1)(F - 1)}{(c - 1)(F - 1) + NC}$$

ويعتبر مؤشر ω^2 أقل تحيزاً من مؤشري η^2 و ε^2 ، ويمكن حسابه للتصميمات البحثية الأكثر تعقيداً. ويوصي باستخدام الصيغة المصححة لمؤشر ω^2 وهي كالتالي (Olejnik & Algina, 2000):

$$\omega_{partial}^2 = \frac{SS_b - df_b MS_w}{SS_b + (N - df_b) MS_w}$$

• df_c درجات الحرية بين المجموعات وتساوي C-1.

وقيمة ω_p^2 هي تقريباً نفس قيمة ω^2 في التصميمات الأحادية العامل ولكنهما يختلفا في حالة التصميمات المعقدة (Fern & Monroe, 1996).

والعلاقة بين المؤشرات الثلاثة (Kline, 2013):

$$\hat{\eta}^2 > \hat{\epsilon}^2 > \hat{\omega}^2$$

معامل ارتباط **Interclass Correlation (r_{ICC})**: يستخدم كمؤشر لقوة العلاقة في حالة التصميمات العاملية وهو أكثر فعاليةً في حالة تصميم القياسات المتكررة والتصميم الذي يكون فيه تأثير المتغير المستقل عشوائي (نماذج التأثيرات العشوائية) وليست التأثيرات المثبتة (Fern & Monroe, 1996) ، وتتحدد قيمته كالتالي (Olejnik & Algina, 2000):

$$r_{ICC} = \frac{C(MS_C - MS_E)}{SS_t + MS_C}$$

وعبر عنه (Fern & Monroe (1996) بالاتي:

$$r_{ICC} = \frac{SS_C(C - 1)MS_e}{SS_t - (n - 1)MS_e}$$

مداخل بديلة

مؤشر **Common language effect Size (CLES)**: اقترح McGraw & Wong

(1992) مؤشر حجم التأثير (CLES) ويعبر عن حجم التأثير في ضوء احتمالات وهو يعطي مقدار الفروق معبراً عنه باحتمالات بمعنى ان احتمال درجات فرداً في المجموعة التجريبية تكون اعلي من درجات فرداً اخر في المجموعة الضابطة.

ولتوضيح قيمة هذا المؤشر في حالة المقارنة بين الذكور والإناث علي متغير ما متصل، فإذا كان متوسط الذكور 69.7 والانحراف المعياري 2.8 ومتوسط الإناث 64.3 و S=2.6 فإن قيمة مؤشر CLES=0.92 وهذا يعني أن احتمال الذكور أطول من الإناث بنسبة 92%.

وهذا المؤشر يقى نفسه من عدم توفر الاعتدالية (Breaugh, 2003) ويقدر هذا المؤشر من خلال تقريب اختبار Z كالتالي (Ledesma et al., 2009):

$$z = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{S_e^2 + S_c^2}}$$

وهذا المؤشر يستخدم للمقارنة بين متوسطين ولكن (Dunlap 1994) وسع استخدام هذا المؤشر ليناسب معامل ارتباط بيرسون.

مؤشر نسبة الكفاءة السببية (CER) Causal Efficacy Ratio : لاحظ Prentice & Miller (1992) أن لحجم التأثير وظيفة مرتبطة بقوة المعالجة للمتغير التابع وأكد Abelson (1995) ان المعالجة يتم تجاهلها بدرجة كبيرة أثناء تقدير حجم التأثير، واقترح تقدير مؤشر CER ونسبته تعكس حجم التأثير الخام وهو الفرق بين متوسط المجموعتين علي المتغير التابع مقسوما علي حجم السببية Cause Size (الفرق بين متوسط المجموعتين علي المتغير المستقل).

العلاقات المتدرجة Multivariate Relationships

كثيراً من الدراسات تتضمن اكثر من متغيرين مثل اساليب الانحدار المتعدد، تحليل المسار، التحليل العامل، نمذجة المعادلة البنائية، وغيرها . اجراء ما وراء التحليل لهذة النوعية من الاساليب عملية صعبة، ولكن يمكن اجراء ما وراء التحليل على الارتباطات من الدراسات. وفيما يلي عرض:

مؤشرات تحليل الانحدار:

إذا اعتبرنا متغيرين منبئيين (مستقلين) متصلين X_1 و X_2 ، ومتغير تابع متصل Y ، فإن معادلة الانحدار اللامعيارية تكون كالاتى:

$$\hat{Y} = b_1 X_1 + b_2 X_2 + a$$

• \hat{Y} الدرجة المتنبأ لها.

• b_1 ، b_2 معاملات الانحدار غير المعيارية.

• a ثابت الانحدار.

اما معادلة الانحدار المعيارية كالاتى:

$$Z_{\hat{Y}} = \beta_1 Z_{X_1} + \beta_2 Z_{X_2}$$

• β_1 ، β_2 معاملات الانحدار المعيارية.

• Z_{X_1} ، Z_{X_2} الدرجات المعيارية.

وإذا اهتم الباحث بدراسة ما وراء التحليل لمعادلة التنبؤ بين المتغيرين Y, X1 مع ضبط المتغير الثالث X2 فان معامل الانحدار المعيارى يكون مؤشراً لحجم التأثير ويقدر كالتالى:

$$\beta_1 = \frac{r_{YX1} - r_{YX2}r_{X1X2}}{1 - r_{X1X2}^2}$$

- r_{YX1} معامل الارتباط بين Y, X1.
- r_{YX2} معامل الارتباط بين Y, X2.
- r_{X1X2} معامل الارتباط بين X2, X1.

ويفضل استخدام معامل الارتباط الجزئي Partial r كمؤشر لحجم التأثير إذا تم عزل تأثير المتغير الدخيل، ويهدف الى تقدير معامل الارتباط بين متغيرين بعد ضبط تأثير المتغير الثالث الدخيل على كلاً من المتغيرين، و يستخدم للكشف عن التضخم او التقلص الخادع لمعامل ارتباط بين متغيرين نتيجة تدخل متغير ثالث (Z) ويقدر كالاتي:

$$r_{xy.z} = \frac{r_{xy} - (r_{xz})(r_{yz})}{\sqrt{1 - r_{xz}^2}}$$

- $r_{xy.z}$ ، معامل الارتباط الجزئي بين y,x.
- r_{yx} معامل ارتباط بين y, x.
- r_{xz} معامل ارتباط بين z,x.
- r_{yz} معامل ارتباط بين z,y.

وفي هذه الحالة فان Partial r و β أفضل من مؤشر r للتعبير عن حجم التأثير ويفضل استخدام R^2_{adj} كمؤشر لحجم التأثير.

وبعد حساب معامل الارتباط الجزئي يجرى ما وراء التحليل على تحويل فيشر ثم تحول النتيجة النهائية الى معامل الارتباط.

ولاجراء الموازنة فى ما وراء التحليل لايد من حساب الخطأ المعيارى لمعامل الانحدار المعيارى كالتالى (Cohen, Cohen, West, & Aiken, 2003):

$$SE_{\beta_1} = \frac{S_Y}{S_X} \sqrt{\frac{1}{1 - R_X^2} \sqrt{\frac{1 - R_Y^2}{N - P - 1}}}$$

والخطأ المعياري لتحويل Z_{sr} كالتالى:

$$SE_{Z_{sr}} = \frac{1}{\sqrt{N - P - 2}}$$

- S_X ، S_Y النحراف المعياري للمتغيرين Y, X .
- R_X^2 التباين المفسر فى المتغير X عن طريق متغيرات اخرى فى النموذج (R_{XZ}^2 فى حالة وجود متغير ثالث Z بجانب Y, X).
- R_Y^2 التباين المفسر فى Y عن طريق كل المتغيرات فى النموذج X وكل المتغيرات الاخرى الدخيلة Z .
- N حجم العينة.
- P عدد المتغيرات المنبئة بما فيها X وكل المتغيرات الاخرى الدخيلة Z .

ولاجراء ما وراء التحليل للدراسات التنبؤية بين متغيرين يجب ان تتضمن كل الدراسات نفس المتغيرين ونفس المتغيرات الدخيلة او المستقلة الاخرى، فاذا كان هدف الدراسة هو بناء معادلة تنبؤ بين سلوك الجريمة(تابع) والتنشئة الاسرية(مستقل) فيمكن ان يعتمد على دراسات تتضمن متغير ثالث اخر مثل عدد سنوات ادمان المخدرات، ولكن تضمين دراسات تتضمن متغير ثالث مختلف يجعل عملية دمج الدراسات عمليه صعبة ولكن يمكن التغلب على ذلك من خلال (Card, 2012):

اولاً: استخدام معامل الارتباط الجزئى من معاملات ارتباط بيرسون.

ثانياً: الاعتماد مباشرة على معاملات الارتباط بين المتغيرين كمؤشرات لحجم التأثير وهذا الاختيار اكثر مرونة.

مؤشرات حجم التأثير في اختبارات الإحصاء المتدرج

التعامل مع الظواهر النفسية يفترض أن تتعدد فيها المتغيرات التابعة والمستقلة، ويوجد القليل من الدراسات التي تناولت مؤشرات حجم التأثير أو تحليل القوة في اختبارات الإحصاء المتدرج، فعلي سبيل المثال الفرض الصفري في تحليل التباين الأحادي يفترض تساوي المتوسطات لكل المجموعات علي المتغير التابع، بينما في تحليل التباين المتدرج MANOVA يفترض تساوي متوسطات المجموعات لكل المتغيرات التابعة.

ويوجد في التراث مقاييس لتقدير حجم التأثير لاختبارات الإحصاء المتعدد، ففي الانحدار المتعدد يستخدم مؤشر معامل الارتباط المتعدد R^2 وتصحيحه R^2_{adj} . وتم اقترح مؤشراً بديلاً لمؤشر d يستخدم في MANOVA ويرمز له بالرمز D^2 وذلك في حالة تعدد المتغيرات التابعة وأكد (2009) Ferguson علي ضرورة الاهتمام بمؤشر D^2 ويسمي distance Mahalanobis ويحسب من اختبارات الدلالة الإحصائية لاختبار MANOVA .

ومؤشر D^2 مستقل عن حجم العينة ولتفسير D^2 وضع Stevens (1980) القيمة 0.25 ضعيف، 0.50 متوسط، و 1.00 فأكثر كبير.

كما يمكن اعتبار إحصاء الدلالة الإحصائية في تحليل التباين المتدرج (MANOVA) و هو Wilks Lambda (Λ) مؤشر لحجم التأثير، وتحدد قيمته كالتالي (Tatsuoka, 1993):

$$\Lambda = \frac{|S_w|}{|S_t|}$$

- S_w مجموع المربعات داخل المجموعات.
- S_t مجموع المربعات الكلي.

وكذلك يوجد مؤشر $\eta^2_{Mult\ adj}$ في اختبار MANOVA وتحدد قيمته بالاتي:

$$\eta^2_{multi-adj} = \eta^2_{multi} - \frac{(p^2 + q^2)(1 - \eta^2_{multi})}{3N}$$

- P عدد المتغيرات التابعة، q عدد درجات الحرية للفروض.

واشار (1993) Tatsuoka الي استخدام مؤشر $\eta^2_{multi-adj}$ للاختبارات المتدرجة وذلك لان مؤشر η^2_{Mult} متحيز تحيز موجب ،اما مؤشر ω^2 لـ Hayes ويمكن التعبير عنه بالصيغة الآتية:

$$\omega^2_{mult} = \frac{(N - K) - (N - 1)\Lambda}{(N - K) + \Lambda}$$

وقد عبر عنها (2000) Olejnik & Algina بالصيغة الاتية:

$$\omega^2_{Mul} = 1 - \frac{N\Lambda}{df_{error} + \Lambda} \quad , \quad df_{error} = N - K$$

وان df درجات الحرية وتساوي N-K ولتصحيح مؤشر ω^2_{Mul} ادخلت عليه تعديلات لأنه مؤشر متحيز وللتخلص من ذلك التحيز يوجد تعديل (Tatsouka, 1973):

$$\omega^2_{multi adj} = \omega^2_{multi} - \frac{P^2 + (K - 1)^2}{3N} (1 - \omega^2_{multi})$$

بالإضافة إلي مؤشرات D^2 ، η^2 ، ω^2 يوجد مؤشرات أخرى لتقدير حجم التأثير ويوصي باستخدامها وهي مؤشرات هوتلنج Hotelling و Pillais و Lawley trace & Bartlet و Trace وهذه المؤشرات الثلاثة يمكن الحصول عليها من خلال برنامج SPSS.

ما وراء التحليل لثبات الاتساق الداخلي

يهدف مقياس الاتساق الداخلي او الثبات الداخلي للمقياس السى تحديد درجة التناسق بين مفردات المقياس، وتوجد العديد من مقاييس الاتساق الداخلي ولكن اكثرهم انتشاراً المعامل الفا كرونباخ ويقدر فى ضوء عدد المفردات فى المقياس j ومتوسط الارتباطات بين هذه المفردات (\bar{r}) كالاتى:

$$\alpha = \frac{j \bar{r}}{1 + (j - 1) \bar{r}}$$

متى يتم اجراء ما وراء التحليل لمقاييس الاتساق الداخلي؟، يرى (2012) Card موقفين لاستخدام الاتساق الداخلي فى ما وراء التحليل هما:

الأول: عند تصحيح الثبات في الدراسات من نقص الثبات Unreliability .

الثاني: عندما يهدف الباحث الى الوصول الى متوسط المعامل الفا كرونباخ لدراسات مختلفة

استخدمت نفس المقياس وهذا في غاية الاهمية خاصة للمقاييس المعروفة في دراسة ظاهرة معينة مثل مقياس كاتل للشخصية او اختبار ستانفورد بينية او العوامل الخمسة الكبرى للشخصية وغيرها لتقييم استخدام هذه المقاييس في دراسة الظاهرة. وتوجد طرق مطروحة في التراث لاجراء ما وراء التحليل لمعامل الفا كرونباخ منها ما قدمه Rodriguez & Maeda (2006) حيث تم تحويل المعامل الفا لمؤشر لحجم التأثير كالتالى:

$$ES_{\alpha} = \sqrt[3]{1 - \alpha}$$

• α معامل الثبات في الدراسة.

والخطأ المعياري لهذا التحويل لمعامل الفا كالتالى:

$$SE_{ES_{\alpha}} = \sqrt{\frac{18j(N-1)(1-\alpha)^2}{(j-1)(9N-11)^2}}$$

• j عدد المفردات في المقياس في الدراسة.

• N حجم العينة.

• α معامل الثبات في الدراسة.

وبعد حساب متوسط معامل الاتساق المحور الفا يحو الى معامل الفا مرة اخرى حتى يسهل

تفسيره كالتالى:

$$\alpha = 1 - ES_{\alpha}^3$$

• ES_{α} معامل الفا المحور في ما وراء التحليل.

واخيراً توجد مجموعة من الاشكاليات او التحديات عن اجراء ما وراء التحليل لهذه النوعية من احجام التأثير التي عرضناها في هذا الفصل يحددها (Card, 2012; Lipsey & Wilson, 2001):

- عدم توافر البيانات المتوفرة لتقدير حجم التأثير، فمثلاً في الدراسات الطولية لا تتوفر للباحث معامل الارتباط بين مرتى القياس للمتغير عكس معامل الاتساق الداخلى الذى

تتوفر البيانات المطلوبة لحساب حجم التأثير. وعدم توفر البيانات اللازمة لحساب حجم التأثير يتطلب الاتصال المباشر بمؤلفي هذه الدراسات.

- التحدى الثانى هو عدم الاتساق فى الطرق التحليلية وتقرير احجام التأثير المناسب، فمثلاً فى تحليل الانحدار عند حساب حجم التأثير حيث تتضمن الدراسات متغيرات منبئات او تغايرات متنوعة غير المتغيرين موضعى الاهتمام وتتضمن الدراسات طرق تحليلية متنوعة فبعضها اعتمد على طريقة التحليل Enter وبعضها اعتمد على Stepwise.

ايضاً عند اجراء ما ما وراء التحليل على التحليل العاملى الاستكشافى حيث يعتمد على تشبعات المفردات بالعوامل كمؤشر لحجم التأثير فيوجد اختلاف واضح فى الطرق التحليلية فى ضوء طرق التقدير او الاستخلاص(المكونات الاساسية- طرق التحليل العاملى) وطرق التدوير(مائل- متعامد) وكيفية تحديد عدد العوامل، وبكل تأكيد يمكن ان تختلف النتائج بين هذه الطرق التحليلية المختلفة وهذا يمثل تحدى كبير امام الباحث عند دمج هذه النوعية من الدراسات.

ولهذا فان استخدام هذه المؤشرات فى تراث ما وراء التحليل قليل وبفضل الاعتماد على مؤشرات التقليدية مثل OR, g, r ، وقبل اجراء هذه النوعية من الدراسات يجب التأكد من توافر المعلومات او البيانات اللازمة لحساب حجم التأثير.

الفصل السابع

تصحیحات احجام التأثير

Corrections to Effects Sizes

مقدمة

قبل اجراء ما وراء التحليل لدمج احجام التأثير فمن الضروري اجراء تصحيح لها من المحددات المنهجية المرتبطة بالدراسات، وايضاً للحصول على خصائص احصائية مرغوبة لهذه احجام التأثير مثل توافر الاعتدالية. وكما تعلم ان توزيع معامل ارتباط بيرسون ملتوى سالب او موجب ولهذا تم تحويله الى تحويل فيشر للحصول على التوزيع الاعتدالي وكذلك تحويل مؤشر Odds Ratio الى اللوغاريتمات، ويوجد العديد من التصحيحات او التعديلات على احجام التأثير للتخلص من قضية التحيز خاصة للعينات الصغيرة. وحدد (2004) Hunter & Schmidt العديد من هذه التصحيحات لمعالجة محددات او قصورات منهجية فى الدراسات الاولية.

دواعى تصحيح احجام التأثير

ويرى (2004) Hunter & Schmidt ان هذه التصحيحات تهدف الى:

- اعطاء تقديرات اكثر دقة لاحجام التأثير من محددات الدراسات سواء فى عملية القياس او المحددات المنهجية مثل حجم العينة.
- زيادة التجانس بين احجام التأثير عبر الدراسات المختلفة وذلك لان التجانس شرط ضرورى لاجراء اختبارات الفروض حول المتوسط العام لاحجام التأثير.

والهدف من وراء التحليل للارتباطات وصف التوزيع الفعلى للبناء موضع الدراسة للارتباطات بين متغير مستقل ومتغير تابع، ولو اجريت الدراسات بدون عيوب او محددات فان توزيع الارتباطات يستخدم بطريقة مباشرة لتقدير التوزيع الحقيقى فى المجتمع ولان الدراسات لا تجرى بدقة تامة فان التناظر او التماثل بين الارتباطات المقاسة من الدراسات لا تعكس تماماً الارتباطات الفعلية الحقيقية فى المجتمع .

وتوجد بعض المحددات او الاخطاء فى الدراسات تحول الى ان تكون دراسات العلوم السلوكية والاجتماعية تامة الاتقان لان عملية القياس تكون غير مباشرة بالتالى تتضمن قدراً معيناً من الاخطاء، بعض هذه المحددات يمكن التغلب عليها او تصحيحها والبعض لا يمكن تصحيحه، وتصحيح احجام التأثير للدراسات يتطلب معرفة حجم العينة، المتوسطات والانحرافات المعيارية، تقديرات الثبات وغيرها.

حدد (Hunter & Schmidt (2004) اهم المحددات فى الدراسات التى يمكن ان تغير حجم معامل الارتباط او التأثير فى الدراسة عن حجمها الحقيقى فى المجتمع كالاتى:

1. خطأ المعاينة Sampling error : فصدق الدراسة يعتمد على كيفية انتقاء العينة من المجتمع سواء اختيار عشوائى او مقصود.

2. خطأ القياس فى المتغير التابع Measurement error: فصدق الدراسة يكون منخفض عن الصدق الحقيقى اذا تضمنت قياسات المتغير التابع اخطأ قياس عشوائية.

3. خطأ قياس فى المتغير المستقل: الاختبار غير تام الثبات.

4. التصنيفية الثنائية المصطنعة للمتغير التابع المتصل او المتغير المستقل المتصل Dichotomization: مثل تقسيم متغير التحصيل الى مرتفعى ومنخفضى التحصيل او عدد ساعات استخدام الفيس بوك " اكثر من ثلاث ساعات" " اقل من ثلاث ساعات".

5. المدى المقيد Restricted Range للمتغيرات المستقلة والتابعة المتصلة.

6. الابتعاد عن قياس المفهوم الحقيقى للمتغيرات المستقلة والتابعة بمعنى اختلاف البنية العاملة للمقاييس عن البناء النظرى المفترض.

7. اخطاء التحليل الاحصائى: مثل عدم الدقة فى ترميز وادخال البيانات، اخطأ حسابية وتحليلية، اخطأ فى قراءة المخرجات الكمبيوترية، اخطأ مطبعية، وغيرها.

8. تباين يرجع الى العوامل الخارجية الدخيلة التى تؤثر فى طبيعة العلاقة بين المتغيرين مثل الخبرة، الظروف الاقتصادية، الاجتماعية، وغيرها.

ويرى (Hunter & Schmidt, 2004) ان هذه المحددات المنهجية والمرتبطة بتصميم الدراسة ماعدا الاخطاء المطبعية واخطاء النقل او الكتابة.

ويؤيد البعض اجراء التصحيحات لاحجام التأثير للدراسات ومن اشد المؤيدين لهذا الاتجاه (Hunter & Schmidt, 1990, 2004; Rubin,1990) حيث يرون ان الدراسات الاولية التى ينتج عنها احجام تاتير تعتمد على قياسات غير دقيقة Imperfect Measures للمفاهيم المراد دراستها، وهذه الدقة ترجع الى نقص الثبات والصدق غير التام للمقاييس وعدم وضوح اجراءات ادارة عملية القياس، وهذه الاشكالية اكثر وضوحاً فى المقاييس النفسية والاجتماعية حيث تتعدد قياسات نفس المفهوم وتختلف فيما بينها فى النظرية والمحتوى.

كذلك توجد اشكالية كبيرة وهى خطأ المعاينة احد مسببات التعارض الكبير بين نتائج الدراسات فى كل التخصصات خاصة فى العلوم الاجتماعية، فى ظل الاعتماد على الدلالة الاحصائية التى تركز على دلالة معامل الارتباط وليس مقداره او حجمه فمثلاً معامل ارتباط 12. لحجم عينة 500 دال احصائياً.

ويرى المؤيدون للتصحيح ان المحددات السابقة تؤدى الى الحصول على احجام تاتير متباينة فيما بينها واستخدام التصحيح يقلل هذه الفروق بين احجام التاتير وهذا يعتبر من اهم متطلبات ما وراء التحليل.

بينما المعارضون لاجراء التصحيحات لاحجام التاتير يرون ان الهدف من هذا الاسلوب هو تقييم التراث البحثى لظاهرة ما بدون تجميل للواقع البحثى بكل ما فيه من محددات منهجية او قياسية بمعنى اظهار نقاط الضعف والقوة فى الدراسات لمحاولة توجيه نظر الباحثين على تجنبها وهذا يتفق مع هدف وراء التحليل وهو التكامل او التوليف بين نواتج الدراسات للوصول الى استنتاج اوفهم للظاهرة وهذا ما قام به مبدع هذا الاسلوب (Glass,1976).

وبعض المحدثون لما وراء التحليل لا يحبذون اجراء هذه التصحيحات فيرى Rosenthal (1991) ان الهدف من وراء التحليل " ان يعلمنا ماذا يجب عمله وليس ما هو الافضل الذى يجب فعله وحذر من ان هذه التصحيحات لا تعطى احجام تاتير دقيقة حيث عندما نصح حجم

التأثير من عدم او نقص الثبات يعطى ارتباطات تزيد عن الواحد الصحيح وفي هذه الحالة كيف نفسره.

كما ان اجراء هذه التصحيحات يضيف صعوبة وتعقيد على اجراءات ما وراء التحليل، ويتطلب هذا تفسير هذه التصحيحات وكيفية اجرائها في تقرير الدراسة. فتخيل اجراء هذه التصحيحات لكل دراسة اكثر من مرة في ضوء المحددات المنهجية المختلفة سواء اشكاليات العينة، نقص الثبات، نقص الصدق، وغيرها. ولكن المؤيدون يرون لا ضرر في ذلك اذا كان هذا يزيد جودة نتائج دراسة ما وراء التحليل.

التحويلات وتصحيح التحيز Transformation and Bias Corrections

من اهم احصائيات حجم التأثير فروق المتوسطات المعيارية، معامل الارتباط، احتمال Odds وكلها يمكن تحويلها او تصحيحها من التحيز، فمثلاً حجم التأثير للفروق بين المتوسطات تعاني من تحيز لاحجام العينات الصغيرة لذلك تم تصحيحها والحصول على *Gadjusted* وكذلك يتم تحويل معامل الارتباط ونسبة Odds لصيغ اكثر مناسبة للتحليل ثم تحور مرة اخرى الى الصيغ الاصلية.

المتطرفات Outliers

الهدف من ما وراء التحليل الوصول الى تلخيص منطقي للنواتج الكمية لتراث بحثي من الدراسات المرتبطة بظاهرة ما وهذا الاستنتاج العام لا يكون دقيق اذا تم تضمين احجام تأثير متطرفة بمعنى انها تكون متناقضة بدرجة كبيرة عن مجمل الدراسات التي تتضمن في ما وراء التحليل، فوجود احجام التأثير المتطرفة يؤثر على قيم المتوسطات والتباينات والاختفاء المعيارية في ما وراء التحليل، ولذلك من الضروري فحص توزيع احجام التأثير من خلال عدة اساليب منها شكل BoxPlot لتشخيص القيم المتطرفة لاحجام التأثير واستبعادها او تصحيحها لجعلها اقلها تطرفاً.

ويمكن استخدام المتوسط المهدب Trimmed Mean لاستبعاد القيم المتطرفة او يمكن استبعاد حجم التأثير الذي يزيد عن 2 او 3 انحرافات معيارية من المتوسط العام لكل احجام التأثير.

تصحیحات Hunter & Schmidt

اقترح Hunter & Schmidt (2004) العديد من التصحيحات للمحددات المنهجية في الدراسات الاولية مثل عدم الثبات، الصدق الضعيف للمقاييس، التصنيفية الثنائية الاصطناعية للمتغيرات المتصلة، المدى المقيد، وغيرها. ولكن قبل عرض هذه التصحيحات والعمليات الحسابية الخاصة به فلا بد من عرض الصيغة العامة لهذه التصحيحات كالآتي Card, (2012):

$$ES_{adjusted} = \frac{ES_{observed}}{a}$$

- $ES_{adjusted}$ حجم التأثير المصحح من المحددات المنهجية في الدراسة.
- $ES_{observed}$ حجم التأثير المقاس او الملاحظ (O, g, d).
- a معامل التصحيح النهائي من كل المحددات المنهجية في الدراسة.

وعلى ذلك فان حجم التأثير المصحح هو دالة لحجم التأثير المقاس ومعامل التصحيح لكل محدّدات الدراسة a وهذا المعامل عبارة حاصل ضرب معاملات التصحيح للمحددات المنهجية كالآتي:

$$a = a_1 * a_2 * a_3 * \dots * a_i$$

وقيم هذه التصحيحات تنقص كلما زادت الجودة المنهجية للدراسة.

وتصحیحات احجام التأثير تتطلب تصحيحات للاخطاء المعيارية، والصيغة العامة لتصحيح الاخطاء المعيارية من محدّدات الدراسة كالتالي:

$$SE_{adjusted} = \frac{SE_{observed}}{a}$$

- $SE_{adjusted}$ الخطأ المعياري المصحح.
- $SE_{observed}$ الخطأ المعياري المقاس غير المصحح.
- a معامل التصحيح الكلي لكل الدراسات

التصحیحات من عدم الثبات Unreliability

هذا التصحيح من عدم الثبات لقياسات المتغيرات المتضمنة لحساب حجم التأثير مثل المتغيرين Y, X . ويشير عدم الثبات الى الخطأ غير المنتظم Nonsystematic error أي أنه لا يحدث بانتظام مثل الحالة المزاجية الوقتية فيعطي المستجيب استجابات مرغوبة إذا كان في حالة مزاجية طيبة وإستجابات غير مرغوبة إذا كان في حالة مزاجية سيئة. والثبات مؤشر لمدى خلو القياسات من اخطأ القياس وتوجد العديد من طرق تقدير الثبات اهمها:

أولاً: الاختبار وإعادته Test – Retest (الاتساق عبر الزمن): وهو إلى أي مدى حصلنا على نفس النتائج من تطبيق نفس الأداة على نفس المجموعة من المفحوصين تحت نفس الشروط عبر الزمن ومعامل الثبات يسمى في هذه الحالة بمعامل الاستقرار Stability Coefficient ونحصل على معامل الاستقرار بحساب معامل الإرتباط (r) بين الدرجات في التطبيق الأول والدرجات في التطبيق الثاني لنفس الأداة على نفس الأفراد.

ثانياً: معامل الإتساق الداخلي Internal consistency: الاجراء الآخر لتقدير معامل الثبات للقياسات هو من تطبيق الاختبار مرة واحدة ومعامل الثبات المتحصل يسمى معامل الإتساق الداخلي ونحصل على معامل الثبات للقياسات من تطبيق الاختبار او المقياس مرة واحدة وفي هذه الطريقة يفضل ان يكون الاختبار أحادي البعد وليس متعدد الأبعاد ومن اهم طرق الاتساق الداخلي طريقة الفاكروبناح (α) ابتدعها كرونباخ (1951) ويسمى معامل الثبات في هذه الحالة بمعامل الاتساق الداخلي (α) وهي تصلح لأي نوع من الإختبارات سواء الموضوعية او المقالية او اختبار مكون من اسئلة موضوعية ومقالية. ومعامل الفا كرونباخ يتحدد بالمعادلة الآتية:

$$\alpha = \frac{k}{k-1} \left[1 - \frac{\sum S_i^2}{S_T^2} \right]$$

• k عدد الأسئلة.

• $\sum S_i^2$ مجموع تباينات درجات مفردات الإختبار

• S_T^2 تباين المجموع الكلي.

ثالثاً: ثبات المقدرين Inter-rater reliability: يقيم من خلال الاتفاق بين العديد من المقدرين ونحصل عليه اما عن طريق معامل الارتباط بين مجموعة من الدرجات المتصلة عن طريق مقدرين او متوسط معاملات الارتباطات بين العديد من المقدرين او من خلال المعامل كبا Cohen Kappa(K) اذا كانت تقارير المقدرين تصنيفية ثنائية(نعم - لا).

وتتراوح قيم معاملات الثبات الثلاثة (r, α, K) من 0.0 كحد ادنى الى 1.0 كحد اقصى والعلاقة بين الثبات وعدم الثبات كالنالى:

$$\text{Reliability} = 1 - \text{Unreliability}$$

ويعتبر الثبات من اهم العوامل التى تؤثر فى مقدار وحجم التأثير فى الدراسة، فاذا كان الثبات مرتفع بمعنى قريب من الواحد الصحيح وليكن 0.80 لقياسات المتغيرين فى الدراسة فانه من المتوقع ان تكون العلاقة المقاسة بين المتغيرين علاقة غير متحيزة وتعكس العلاقة الفعلية فى المجتمع لان القياسات تكون خالية تقريباً من اخطأ القياس، وعلى العكس لو تضمنت قياسات احد المتغيرين او كلاهما اخطأ قياس(ثبات منخفض) فان حجم التأثير المقدر لا يمثل حجم التأثير الحقيقى فى المجتمع ويكون تقدير زائف وهذا يتناقض مع الهدف من اسلوب ما وراء التحليل وهو الوصول الى حجم التأثير الحقيقى للدراسات المختلفة التى تناولت نفس المتغيرين لان معامل الارتباط لا يكون بين الابنية التحتية للمتغيرات انما بين اخطأ قياس.

ولو قدر مقدار عدم الثبات للمتغيرين نستطيع تقدير حجم تقلص حجم التأثير وتصحيه وهذه تعتبر اولوية قصوى لاجراءات ما وراء التحليل لانه يتطلب التعامل مع احجام التأثير الفعلية من الدراسات الاولية.

فمثلا اذا كان لدينا متغيرين Y, X فى دراسة ما ومعامل الثبات للمتغيرين r_{xx} و r_{yy} فانه يمكن تقدير معامل التصحيح بين ابنية Y, X كالاتى (Card, 2012; Hunter & Schmidt, 2004):

$$a_{\text{unreliability}} = \sqrt{r_{xx}r_{yy}}$$

- $a_{\text{unreliability}}$ معامل التصحيح من عدم الثبات.
- r_{xx}, r_{yy} معامل الثبات للمتغيرين Y, X .

تقدير حجم التأثير الحقيقي عن طريق قسمة حجم التأثير المقاس على قيمة التصحيح. وبالمثل يمكن تقدير الخطأ المعياري المصحح عن طريق قسمة الخطأ المعياري المقاس على معامل التصحيح $a_{unreliability}$.

وفيما يلي توضيح لهذا من دراسة ما وراء التحليل لـ Card et al.(2008) لدراسة للعلاقة بين العدوان والرفض من الآخرين وتوصلت الى ان معامل ارتباط بيرسون $r = 0.19$ ومعامل الاتساق الداخلي للمتغيرين على التوالي 0.82 ، 0.81 ولتقدير معامل التصحيح من عدم الثبات كالاتي:

$$a_{unreliability} = \sqrt{r_{xx}r_{yy}} = \sqrt{0.82 * 0.81} = 0.815$$

وعلى ذلك حجم التأثير المصحح من عدم الثبات كالاتي:

$$r_{adjusted} = \frac{r_{observed}}{a} = \frac{0.19}{0.815} = .23$$

وحساب الخطأ المعياري لتحويل فيشر لمعامل الارتباط غير المصحح 0.19 وكانت قيمته 0.0498 لعينة $N=407$ وتم تصحيح الخطأ المعياري كالاتي:

$$SE_{adjusted} = \frac{SE_{observed}}{a} = \frac{0.0498}{0.815} = 0.061$$

لاحظ زادت قيمة حجم التأثير وانخفضت قيمة الخطأ المعياري بعد عملية التصحيح.

ويمكن استخدام معادلة تصحيح حجم التأثير من عدم الثبات لمتغير واحد وليكن X فقط دون المتغير الاخر Y وذلك لو افترض الباحث ان احد المتغيرين يقاس بدون خطأ قياس، وعليه التعويض بثبات Y بالقيمة واحد على اساس انه تام الثبات او غير مهتم سوى بمتغير واحد من متغيرات الدراسة.

وفي حالة التعامل مع الابنية التحتية للمتغيرات في التحليل العاُملي التوكيدي ونمذجة المعادلة البنائية فلا حاجة لعملية التصحيح على اساس ان العلاقة بين العومل تكون نقية من خطأ القياس وتعتبر مؤشرات لحجم التأثير.

تصحيات الصدق الناقص Imperfect Validity

يشير الصدق الى التداخل الشديد بين المحتوى المقاس والمحتوى الفعلى او ان يقيس الاختبار المحتوى الذى ينوى قياسه بالفعل. ميز (2012) Card بين الصدق والثبات، فالثبات يشير الى قابلية التكرار Repeatability عبر مفردات او مقدرين، او مناسبات القياس، والثبات المرتفع يشير الى وجود التناسق او التناظر او الارتباط المرتفع بين المفردات او المقدرين المختلفين، فالثبات لا يخبرنا عن ما اذا كانت اداة القياس تقيس بالفعل البناء التحتى للمفهوم الذى ننوى قياسه بالفعل ولكن يخبرنا فقط ان نفس السمة تقاس باتساق عبر مرات التطبيق، فى المقابل فان الصدق يشير الى الاتساق Consistency او التوافق بين السمة المقاسة والسمة الحقيقية الفعلية، فهل يقيس اختبار الذكاء بالفعل ظاهرة الذكاء؟، او هل يقيس مقياس العوامل الخمسة الكبرى للشخصية بالفعل البناء التحتى للشخصية؟، هذة اسئلة يجيب عليها الصدق. الصدق والثبات ظاهرتان مستقلتان، فالاختبار يمكن ان يكون عالى الثبات ومنخفض الصدق او عالى الصدق ومنخفض الثبات.

وصدق المقياس (X) لتقدير المفهوم او البناء التحتى (T) هو r_{XT} عندما يكون المقياس تام الثبات. ولو ان حجم التأثير موضع الاهتمام فى ما وراء التحليل هو العلاقة بين البناء الحقيقى T ومتغير اخر Y فان معامل الارتباط موضع الاهتمام r_{TY} ، وبالتالي فالعلاقة المقاسة بين المقياس X والمتغير Y هى مساوية لحاصل ضرب صدق المقياس والعلاقة بين البناء التحتى مع المتغير الاخر Y كالاتى:

$$r_{XY} = r_{XT} * r_{TY}$$

ولتحديد العلاقة بين البناء التحتى T والمتغير Y كالاتى (2012) Card:

$$r_{TY} = \frac{r_{XY}}{r_{XT}}$$

وعلى ذلك التصحيح من نقص الصدق كالاتى:

$$a_{validity} = r_{XT}$$

• r_{XT} معامل الصدق.

حيث ان r_{XT} يمثل معامل الصدق او الارتباط الصحيح بين المقياس X والمفهوم T . لكن يوجد تحديين لهذا التصحيح الاول: ان هذا التصحيح يفترض ان جزء خاص للمقياس X وليس جزء من المفهوم T ولا يرتبط مع المتغير الاخر Y وهذا جزء غير ثابت وغير صادق نتيجة تباين الطريقة Method variance طريقة تقديم المقياس. الثانى: صعوبة الحصول على على معامل الصدق r_{XT} حيث لا يذكر فى تقرير الدراسات ولذلك فان اجراء هذا التصحيح يكون فى غاية الصعوبة.

تصحيات الثنائية التصنيفية الاصطناعية Artificial Dichotomization

يقصد بها تقسيم المتغير المتصل بطبيعته الى متغير تصنيفى بمستويين مثل تقسيم متغير الذكاء الى مرتفعى ومنخفضى الذكاء وهذا من شأنه تقليص او انكماش Attenuation معامل الارتباط بين هذا المتغير بالمتغيرات الاخرى، ودائماً يتم تقسيم المتغير المتصل فى ضوء معايير غير متفق عليها مثل المتوسط او الوسيط او وحدة انحراف معيارى فوق المتوسط. ويمكن ان تتضمن الدراسة متغير تصنيفى اصطناعى او متغيرين تصنيفين وهذه التصنيفية ذات بناء تحتى متصل وليست تصنيفية حقيقية (ذكر - انشى).

وتصحيح حجم التأثير من تصنيفية احد المتغيرين عملية مباشرة وهذا يتطلب معرفة الاعداد او نسب الافراد فى كل تصنيف وللتصحيح تستخدم المعادلة الاتية (Hunter & Schmidt, 2004):

$$a_{dichotomization} = \frac{\phi(c)}{\sqrt{PQ}}$$

• P ، Q النسب فى كل مجموعة.

• $\phi(c)$ التناسق الاعتدالى عند النقطة C

حيث $\phi(c)$ هى التناسق او الترتيب الاعتدالى عند النقطة C التى تقسم التوزيع الاعتدالى الطبيعى الى النسب P ، Q وهذا المؤشر غير مألوف ولكن يمكن عرض قيمه من الجدول الاتى (Card 2012):

Split	$\Phi(c)$	$\alpha_{\text{dichotomization}}$	Split	$\Phi(c)$	$\alpha_{\text{dichotomization}}$
.50 / .50	.3989	.7979	.25 / .75	.3178	.7339
.49 / .51	.3988	.7978	.24 / .76	.3109	.7279
.48 / .52	.3984	.7975	.23 / .77	.3036	.7215
.47 / .53	.3978	.7971	.22 / .78	.2961	.7148
.46 / .54	.3969	.7964	.21 / .79	.2882	.7076
.45 / .55	.3958	.7956	.20 / .80	.2800	.6999
.44 / .56	.3944	.7946	.19 / .81	.2714	.6917
.43 / .57	.3928	.7934	.18 / .82	.2624	.6830
.42 / .58	.3909	.7920	.17 / .83	.2531	.6737
.41 / .59	.3887	.7904	.16 / .84	.2433	.6637
.40 / .60	.3863	.7886	.15 / .85	.2332	.6530
.39 / .61	.3837	.7866	.14 / .86	.2226	.6415
.38 / .62	.3808	.7844	.13 / .87	.2115	.6290
.37 / .63	.3776	.7820	.12 / .88	.2000	.6156
.36 / .64	.3741	.7794	.11 / .89	.1880	.6010
.35 / .65	.3704	.7766	.10 / .90	.1755	.5850
.34 / .66	.3664	.7735	.09 / .91	.1624	.5674
.33 / .67	.3621	.7702	.08 / .92	.1487	.5480
.32 / .68	.3576	.7666	.07 / .93	.1343	.5262
.31 / .69	.3528	.7628	.06 / .94	.1191	.5016
.30 / .70	.3477	.7587	.05 / .95	.1031	.4732
.29 / .71	.3423	.7544	.04 / .96	.0862	.4398
.28 / .72	.3366	.7497	.03 / .97	.0680	.3989
.27 / .73	.3306	.7448	.02 / .98	.0484	.3458
.26 / .74	.3244	.7395	.01 / .99	.0267	.2679

مثال: دراسة ما تناولت العلاقة بين العدوان ورفض الاقران حيث قام الباحث بتقسيم متغير العدوان الى تصنيفين لعينة مكونة من 491 حيث بلغ حجم العينة غير العدوانية (83.9%) 412 وحجم عينة العدوانية (16.1%) 79 وبالمبحث في الجول السابق بـ $\frac{16}{84}$ يتضح ان قيمة:

$$\phi(c) = 0.243$$

وعليه فان التصحيح من التصنيفية الثنائية المصطنعة كالاتى:

$$\alpha_{\text{dichotomization}} = \frac{\phi(c)}{\sqrt{PQ}} = \frac{0.243}{\sqrt{0.839 \times 0.161}} = \frac{0.243}{0.368} = 0.664$$

وبما ان معامل الارتباط غير المصحح من التصنيفية الثنائية 0.16 وتم تحويلها الى تحويل فيشر والخطأ المعياري لتحويل فيشر 0.0453 وعليه فان معامل الارتباط المصحح من التصنيفية كالاتى:

$$r_{adjusted} = \frac{0.16}{0.664} = 0.24$$

والخطأ المعياري المصحح كالاتى:

$$SE_{adjusted} = \frac{0.0453}{0.664} = 0.0682$$

كل هذه التصحيحات عندما يكون متغير واحد فقط تصنيفى ولكن اذا كان المتغيرين ذات تصنيفية اصطناعية فان عملية التصحيح لمعامل الارتباط تكون اكثر تعقيداً وتكون معامل التصحيح للمتغيرين معاً كالاتى (Card, 2012):

$$Aa_{XY,dichotomizatin} = a_{X,dichotomizatin} \times a_{Y,dichotomizatin}$$

• a_X ، a_Y معامل التصحيح من التصنيفية لكل متغير على حدة.

اي ان تصحيح المتغيرين هو حاصل ضرب تصحيح المتغيرين. و اشار Hunter & Schmidt (2004) الى امكانية اجراء هذا التصحيح للمتغيرين عندما:

- يتم تقسيم احد المتغيرين فى ضوء الوسيط ومعامل الارتباط المصحح اقل من 0.70.
- تقسيم احد المتغيرين حيث $0.40 < P < 0.60$ ومعامل الارتباط المصحح اقل من 0.50.
- كلاً من المتغيرين تقسيمهما $0.20 < P < 0.80$ ومعامل الارتباط اقل من 0.40.

لو استوفت هذه الشروط فان التقريب يكون دقيق بمعنى التحيز اقل من 10%، ولكن لو

حدث تضخم لمعامل الارتباط المصحح بصورة كبيرة فانه يفضل استخدام معامل الارتباط الرباعى.

تصحيح من المدى المقيد Range Restriction

تقدير معامل الارتباط للعلاقة بين متغيرين متصلين ذات مدى مقيد او محدود لا يعطى صورة حقيقية لمعامل الارتباط فى المجتمع لانه يحدث انكماش لمعامل الارتباط. ويحدث لحجم التأثير تقلص لو ان المتغير قائم على مدى صغير فى عينة الدراسة مقارنة بمجتمع العينة، على سبيل المثال اراد باحث دراسة العلاقة بين الذكاء والاداء الوظيفى ولكن العينة المتضمنة فى الدراسة

فقط الافراد فوق متوسطى الذكاء وعليه فان معامل الارتباط المحسوب من العينة اصغر من نظيره فى المجتمع حيث الافراد متنوعى القدرة العقلية.

الخطوة الاولى لتصحيح احجام التأثير من المدى المقيد لاحد المتغيرين هو تحديد الانحراف المعياري للمتغير فى المجتمع ثم تحديد الانحراف المعياري فى الدراسة الاولى بالنسبة الى الانحراف المعياري فى المجتمع المرجعي Reference. والنسبة بين الانحراف المعياري فى الدراسة الى الانحراف المعياري المرجعي كالاتى:

$$u = \frac{S_{study}}{S_{reference}}$$

مثال: ذكرت دراسة ما ان الانحراف المعياري للذكاء لعينة ما 10 مع انحراف معياري فى المجتمع 15 وعليه فانه نستطيع مقارنة مدى المتغير فى العينة مقارنة بمدى المجتمع كالاتى:

$$u = \frac{S_{study}}{S_{reference}} = \frac{10}{15} = .67$$

وفى ضوء هذه النسبة فان معامل التصحيح من المدى المقيد كالاتى (Hunter & Schmidt, 2004):

$$a_{range} = \sqrt{u^2 + r^2(1 - u^2)}$$

- u درجة قيد المدى.
- r معامل الارتباط.

المظهر الخاص بهذا التصحيح ان قيمة التصحيح a يمكن ان تكون اكبر من الواحد الصحيح. وهذا يحدث اذا كان مدى المتغير فى العينة اكبر من مدى المتغير فى المجتمع.

وحجم التأثير المصحح كالاتى:

$$r_{adjusted} = \frac{(u)(r)}{\sqrt{u^2 - 1)r^2 + 1}}$$

• u نسبة الانحراف المعياري للدراسة (المقيد) الى الانحراف المعياري (غير المقيد) او المستهدف.

• معامل الارتباط بين المتغيرين (حجم التأثير).

ويرى (Lipsey & Wilson (2001) انه يمكن استخدام هذه الصيغة لتصحيح حجم التأثير من المدى الكبير في العينة اكبر من نظيره المستهدف في المجتمع وفي هذه الحالة فان النسبة بين الانحراف المعياري للدراسة والانحراف المعياري المستهدف في المجتمع تكون اكبر من الواحد الصحيح.

وتصحيح الخطأ المعياري المصحح يكون كالتالي:

$$SE_{adjusted} = \frac{(u)(r)(SE)}{r \times \sqrt{u^2 - 1}r^2 + 1}$$

• SE الخطأ المعياري غير المصحح.

واثناء جمع الدراسات الاولية توجد العديد من الاعتبارات اهمها تكويد الاتساق الداخلى للمقاييس المستخدمة في الدراسات لتحديد الثبات المنخفض او المرتفع للدراسات، وكذلك اذا اعتمدت الدراسات على نفس او مقاييس مشابهة للمتغير فيجب حساب الانحرافات المعيارية عبر الدراسات لتحديد ما اذا كانت بعض الدراسات تعاني من المدى المقيد، والهدف من التصحيحات السابقة محاولة الوصول الى العلاقة النقية الفعلية بين المتغيرات بحيث تناظر العلاقة بين الابنية التحتية لهذه المتغيرات

الفصل الثامن

دمج احجام التأثير

Combining Effect Sizes

مقدمة

بعد جمع الدراسات وحساب احجام التأثير ثم اجراء التصحيحات يبدأ الباحث فى دمج هذه احجام التأثيرات فى ما وراء التحليل لحساب حجم التأثير العام عبر الدراسات المختلفة للوصول الى الاستنتاج العام عن الظاهرة او المتغيرات موضع الدراسة. وعملية الدمج تكون سهلة لو ان الدراسات اتبعت نفس المنهجية والعينات والقياسات والتصميمات ونفس المفاهيم النظرية.

ومن الاسئلة الرئيسية التى يحاول ما وراء التحليل الاجابة عليها الاول: ما هو حجم التأثير النموذجى للعلاقة بين المتغيرين موضع الدراسة من عينة من الدراسات التى تناولت المتغيرين؟، **الثانى**: هل يوجد اختلافات بين احجام التأثير المتضمنة فى التحليل بمعنى انها احجام تأثير متجانسة؟، والاجابة على السؤال الثانى فى غاية الاهمية لتحديد طبيعة نوع النموذج المستخدم لحساب حجم التأثير العام وايضاً ضرورية لتحديد مصادر التنوع والاختلاف بين احجام التأثير من الدراسات المختلفة من خلال استراتيجية التحليل المتفاعل Moderator analysis.

يتضمن ما وراء التحليل تصنيفين من المتغيرات (Lipsey & Wilson, 2001):

الاول: احجام التأثير للمتغيرات التابعة موضع الاهتمام(التحليل الرئيسى).

الثانى: المتغيرات الوصفية او خصائص الدراسات التى ولدت احجام التأثير والتى تؤسس للمتغيرات المستقلة فى التحليل(التحليل الثانوى).

وتحليل البيانات يتضمن وصف توزيعات المجموعة المنتقاة لاحجام التأثير من حساب المتوسطات والتباينات لاحجام التأثير ثم بعد ذلك بحث العلاقة بين احجام التأثير والمتغيرات الوصفية او خصائص الدراسات من خلال استخدام تحليل التباين ANOVA او ما وراء الانحدار.

اشكالة تحليل ما وراء التحليل

يعتبر الفرد او الشخص هو وحدة التحليل فى العلوم السلوكية فى الدراسات البحثية باختلاف منهجياتها المختلفة بينما وحدة التحليل فى ما وراء التحليل هى الدراسة البحثية وهنا توجد اشكاليات عديدة اهمها:

الاول: الدراسات فى ما وراء التحليل يمكن ان تنتج اكثر من حجم تأثير حيث يمكن للدراسة الواحدة ان تتضمن علاقة بين المتغيرين للذكور وللاناث او علاقة بين متغيرين على فترات زمنية مختلفة ولكن الاعتماد على اكثر من حجم تأثير لنفس الدراسة يسبب الاعتمادية الاحصائية Statistically dependence وهذا اختلال لمسلمة الاستقلالية اللازمة لتحليل البيانات وهذا يسبب خطأ جوهرى للتحليل الاحصائى الاستدلالي لاحجام التأثير. والاعتماد على اكثر من حجم تأثير للدراسة يسبب زيادة فى حجم عينة احجام التأثير المعتمدة وهذا يشوه الخطأ المعيارى.

الثانى: الدراسات المجمعة لحساب احجام التأثير تتضمن احجام عينات مختلفة من الافراد وبما ان حساب احجام التأثير يعتمد على المتوسطات والانحرافات المعيارية و الارتباطات، وعليه فان الخصائص الاحصائية لاحجام التأثير تعتمد فى المقام الاول على حجم العينة، فحجم التأثير القائم على حجم عينة كبير يتضمن خطأ معاينة صغير وبالتالي يكون تقديره اكثر دقة وثبات من ذلك حجم التأثير القائم على حجم عينة صغير، وبالتالي اذا تم اجراء التحليل الاحصائى لما وراء التحليل بدون الاخذ فى الاعتبار حجم التأثير كاننا عالجننا كل الدراسات كما لو كانت لها نفس حجم العينة ويكون لهما نفس الاسهام لنتائج ما وراء التحليل، ولكن احجام التأثير المقدره من احجام عينات كبيرة يجب ان تلعب دوراً اكبر من نظيرتها المقدره من احجام عينات صغيرة لانها اقل خطأ معاينة، وعليه لا يجب معالجة احجام التأثير بنفس الدرجة من الاهمية فى التحليل الاحصائى ولكن يجب ان تعطى وزن Weight اكبر لاحجام التأثير من الدراسات ذات العينات الكبيرة فى الحسابات الاحصائية من تلك التى تعتمد على احجام عينات صغيرة ولذلك لابد من اجراء التحليل الموزون Weighted analysis.

ويمر تحليل البيانات فى ما وراء التحليل بعدة خطوات:

أولاً: اجراء تصحيحات لاحجام التأثير من المحددات المنهجية للدراسات (الفصل السابع).

ثانياً: حساب متوسط حجم التأثير وتحليل توزيع احجام التأثير: وتعتبر هذه من اهم الخطوات لاجراء ما وراء التحليل وتتضمن الخطوات الاتية:

1. انشاء قاعدة بيانات لمجموعة من احجام التأثير المستقلة.

2. حساب المتوسط الموزون Weighted mean.

3. حساب فترات الثقة والدلالة الاحصائية حول المتوسط الموزون.

4. اختبار التجانس Homogeneity بين احجام التأثير المختلفة.

وإذا لم يوجد تجانس بين احجام التأثيرات بمعنى اختلافات بينهم فمن الضروري اعادة التحليل من الخطوة 2 حتى 4 باستخدام نموذج احصائي مختلف الى جانب تحليلات اضافية.

بناء قاعدة بيانات لمجموعة مستقلة من احجام التأثير Creating an Independence Set of Effect Size

كما سبق واشرنا ان انه يجب ان تكون احجام التأثيرات مستقلة عن بعضها البعض بمعنى لا يجب تضمين اكثر من حجم تأثير من نفس الدراسة لنفس العينة من الافراد، ويرى البعض ان عدم تضمين احجام التأثير للعينات الفرعية فى الدراسة (مثلاً ذكور - اناث)، وكذلك احجام تأثير من دراسات مختلفة لنفس الفريق البحثى (Wolf, 1990). ولكن دراسات ما وراء التحليل يتم جمعها بناءً على المفهوم المتضمن فى الدراسة ويمكن ان تتضمن الدراسة اكثر من مفهوم وعليه يمكن الاعتماد على اكثر من حجم تأثير فى الدراسة التى تتناول مفاهيم مختلفة وهذا نادراً ما يحدث لانه لا يتناسب مع منهجية ما وراء التحليل التى تركز على متغيرين، ولكن يمكن اجراء ما وراء التحليل لاجحام تأثير لدراسات تتضمن اكثر من متغيرين مثل دراسات الانحدار المتعدد وتحليل التباين المتدرج وهذا يتم من خلال ما وراء التحليل المتدرج Multivariate meta-analysis.

ولو تضمنت الدراسة أكثر من حجم تأثير للمفهوم بمعنى اعتمدت على تعريفات اجرائية مختلفة فلا يجب تضمينها جميعها في نفس التحليل كما لو كانت بيانات مستقلة. اذا تضمنت الدراسة احجام تأثير متعددة فيمكن اختصارها باحدى الطريقتين (Lipsey & Wilson, 2001):

الاولى: حساب متوسطهما و التعامل مع المتوسط كحجم تأثير واحد.

الثاني: الاكتفاء بتضمين حجم تأثير واحد واستبعاد بقية احجام التأثير الاخرى والاختيار يتم في ضوء مدى مناسبة التعريف الاجرائي في الدراسة مع التعريفات الاجرائية في كل الدراسات الاخرى وجودة عملية القياس واذا تساوت هذه المعايير فيتم الاختيار عشوائى.

حساب متوسط حجم التأثير The mean effect size

قبل حساب متوسط حجم التأثير فان اعطاء وزن متساوى لكل حجم تأثير فيه شئ من التجاهل لحقيقة ان الدراسات ليست بنفس القدر من الجودة وان كلها يمكن ان تسهم بنفس القدر من الاسهام في حساب متوسط حجم التأثير. في الحقيقة توجد دراسات اكثر جودة ودقة من حيث الخصائص المنهجية مثل اجراءات الضبط وحجم العينة الكافي من نظيراتها الاقل جودة من حيث اقل ثبات و حجم عينة صغير بالتالى فمن غير المقبول ان تتساوى مع الدراسات عالية الجودة في حساب حجم التأثير العام. من هنا تاتي اهمية عملية الموازنة Weighting للدراسات المتضمنة في التحليل بما يتناسب مع دقتها وجودتها وحجم العينة.

مثال (Card, 2012): افترض ان دراسة تتضمن عينة 10 افراد وتوصلت الى ان العلاقة بين Y, X تساوى 0.20 (تحويل فيشر = 0.203)، والدراسة الثانية اعتمدت على عينة 10000 فرداً و معامل الارتباط بين Y, X 0.30 (تحويل فيشر = 0.310)، ولكن قبل حساب متوسط العلاقة بين المتغيرين في الدراستين فمن المهم التامل في دقة تقديرات احجام التأثير للدراستين، فالدراسة الاولى اعتمدت على عينة 10 افراد والخطأ المعياري لتحويل فيشر Z_r لها كالاتى:

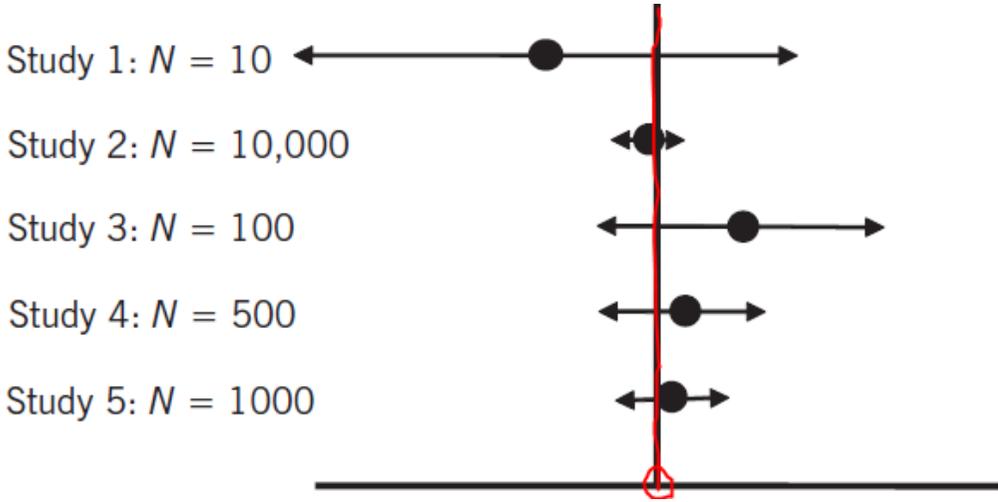
$$SE_{Zr} = \frac{1}{\sqrt{N-3}} = \frac{1}{\sqrt{10-3}} = 0.378$$

في حين ان الخطأ المعياري لـ Z_r للدراسة الثانية:

$$SE_{zr} = \frac{1}{\sqrt{N-3}} = \frac{1}{\sqrt{10000-3}} = 0.010$$

ولاحظ ان الخطأ المعياري في حالة استخدام عينات كبيرة قليل مقارنة بنظيره في حالة عينات صغيرة، وعليه فان احجام التأثير المتولدة من عينات كبيرة اكثر دقة وتقترب من حجم التأثير الحقيقي الفعلي في المجتمع. والخطا المعياري يتناسب عكسياً مع حجم العينة.

والشكل التالي يبين مفهوم الدقة لاحجام التأثير:



مدى احجام التأثير في المجتمع

الشكل (1.8): عرض لمفهوم الدقة لاحجام التأثير.

الشكل يعرض خمسة دراسات متفاوتة حجم العينات وتختلف فيما بينها في الدقة لاحجام التأثير في المجتمع. فالخط المستقيم في المنتصف (الاحمر) يمثل حجم التأثير الحقيقي في المجتمع.

فالدراسة الاولى انتجت حجم تأثير الممثل بالدائرة على يسار الخط الراسي وهي تعتبر اقل من حجم التأثير الحقيقي ولها خطأ معياري كبير لان حجم العينة 10 صغير نسبياً وهذا يؤدي الى فترة ثقة كبيرة او واسعة (الخط الافقي)، وعليه لو اعتمد الباحث على هذه الدراسة فانها تعطى تقدير منخفض لحجم التأثير الحقيقي في المجتمع.

اما الدراسة الثانية (N=10000) تتضمن حجم عينة كبير جداً فان حجم التأثير في هذه الدراسة (الدائرة السوداء) يكون قريب جداً من حجم التأثير الحقيقي في المجتمع ومدى فترات الثقة يكون

ضيقاً لأن هذه الدراسة لها خطأ معيارى صغير جداً. بالتالى تمتلك دقة عالية فى تقدير حجم التأثير الحقيقى فى المجتمع.

وعلى ذلك يجب اعطاء وزن واهمية للدراسة الثانية اكثر من الدراسة الاولى فى التحليل الاحصائى لما وراء التحليل للوصول الى حجم التأثير الحقيقى فى المجتمع.

ولكن الباحث لا يمكن الحصول على حجم التأثير الحقيقى فى المجتمع لمعرفة مدى اقتراب او ابتعاد احجام التأثير فى الدراسات من حجم التأثير الحقيقى ولكن الهدف الاسمى من اجراء ما وراء التحليل هو محاولة الوصول افضل تقدير لحجم التأثير، وعلى ذلك فان الهدف من عملية الموازنة فى ما وراء التحليل اعطاء وزن اكبر لبعض الدراسات التى لها اخطاء معيارية صغيرة ودقة صغيرة (احجام عينات كبيرة) مقارنة بنظيرتها التى لها اخطاء معيارية كبيرة ودقة كبيرة (احجام عينات صغيرة).

وتتم عملية الموازنة من خلال اجراء احصائى منطقى ومقبول وهو وزن احجام التأثير عن طريق مقلوب او معكوس تباينات الاخطأ المعيارية. فاذا كانت وزن الدراسة i هو w_i ، والخطأ المعيارى لحجم التأثير فى الدراسة SE_i ، فان وزن الدراسة هو مقلوب او معكوس تباين الخطأ المعيارى كالاتى:

$$w_i = \frac{1}{SE_i^2}$$

• SE_i الخطأ المعيارى لحجم التأثير المقدر للدراسة i

مثال: قاعدة بيانات لدراسة ما وراء التحليل لـ 22 دراسة للعلاقة بين العدوان والرفض (Card, 2012)

الجدول (1.8): بيانات دراسة ما وراء التحليل لـ 22 دراسة لـ Card et al. (2008)

Sample size (N)	Effect size (r)	Transformed ES (Z_r)	Standard error (SE)	Weight (w)	wES ($w * Z_r$)	wES ² ($w * Z_r^2$)
228	.525	.583	.0693	208.12	121.27	70.66
491	.198	.201	.0579	297.81	59.74	11.98
65	.311	.322	.1325	56.95	18.34	5.90
458	.554	.624	.0484	427.07	266.55	166.37
929	.161	.162	.0346	835.96	135.72	22.03
904	.336	.349	.0347	831.95	290.69	101.57
74	.396	.419	.1223	66.89	28.02	11.73
151	.617	.721	.0855	136.66	98.48	70.97
150	.557	.628	.0845	139.90	87.90	55.23
590	.575	.655	.0426	550.97	361.12	236.69
180	.039	.039	.0831	144.75	5.70	0.22
139	.358	.375	.0892	125.71	47.14	17.68
132	.049	.049	.0922	117.70	5.81	0.29
60	.000	.000	.1386	52.05	0.00	0.00
839	.326	.339	.0381	689.52	233.68	79.19
262	-.048	-.048	.0642	242.93	-11.73	0.57
266	.454	.489	.0636	246.94	120.85	59.15
209	.253	.258	.0715	195.52	50.54	13.06
314	.160	.162	.0589	288.73	46.71	7.56
881	.477	.519	.0351	810.71	420.41	218.01
517	.469	.509	.0455	482.45	245.37	124.79
228	.572	.651	.0702	202.92	132.06	85.94
				7,152.21	2,764.36	1,359.60

وهذا الجدول يتضمن:

- حجم العينة لكل دراسة Sample size.
- معامل ارتباط بيرسون وهو مؤشر لحجم التأثير Effect Size.
- تحويل فيشر لكل معامل ارتباط Z_r .
- الخطأ المعياري لمعامل التحويل SE_{Zr} لفisher فمثلاً الدراسة الاولى حجم العينة 228 وعليه فان الخطأ المعياري لتحويل فيشر كالتالي:

$$SE_{Zr} = \frac{1}{\sqrt{N - 3}} = \frac{1}{\sqrt{228 - 3}} = 0.0693$$

وهكذا لكل دراسة.

• الوزن لكل حجم التأثير w_i وهى كالاتى:

$$w_{study1} = \frac{1}{SE_i^2} = \frac{1}{(0.0693)^2} = \frac{1}{0.00480249} = 208.2253$$

وهكذا لكل دراسة.

• حاصل ضرب وزن الدراسة فى حجم التأثير الدراسة $w_i * z_r$ فالدراسة الاولى:

$$w_1 * z_{r1} = 208.2253 \times 0.583 = 121.3953$$

• حاصل ضرب مربع حجم التأثير للدراسة فى وزن $w_i * z_r^2$ فالدراسة الاولى:

$$w_1 * z_{r1}^2 = 208.2253 \times 0.583^2 = 70.7735$$

وهذه الدراسات تم تصحيحها من محددين اساسين فى هذه الدراسات وهى التصنيفية الثنائية الاصطناعية وعدم الثبات اذا كان مناسب وفى ضوء البيانات المتاحة من الدراسات.

حساب المتوسط الموزون لاحجام التأثير Weighted Mean Effect Size

تتعدد مؤشرات النزعة المركزية فمنها المنوال Mode وهو القيمة الاكثر حدوثاً او شيوعاً ولايفضل على الاطلاق استخدامه فى تحليلات ما وراء التحليل لعرض القيمة النموذجية المنتصفية لاحجام التأثير. اما الوسيط Median وهو القيمة المنتصفية بعد ترتيب القيم تنازلياً او تصاعدياً من السهل تحديدها لانها تعتمد على قيمة واحدة فى التوزيع اذا كان عدد القيم فردى او قيمتين اذا كان عدد القيم زوجى. فالوسيط فى بيانات الجدول السابق لاحجام التأثير $r = 0.35$ ومن اهم مميزات الوسيط عدم تاثره بالالتواء لاحجام التأثير لانه يستبعدا من طرفى التوزيع وهذا عكس المتوسط الاكثر تاثراً بالالتواء. ولكن نظراً لاعتماد الوسيط على قيمة او قيمتين فلا ياخذ فى اعتباره كل احجام التأثير الموزونة للدراسات ولذلك يعتبر الوسيط مؤشر مكملاً للمتوسط.

ومقياس النزعة المركزية الاكثر استخداماً فى دراسات ما وراء التحليل المتوسط Mean وهو مجموع القيم مقسوماً على عددها وهو يتضمن كل احجام التأثير الموزونة.

ويقدر المتوسط الموزون لحجم التأثير للدراسات المختلفة من الاوزان w_i واحجام التأثير ES_i من كل دراسة i كالاتى (Card, 2012; Lipsey & Wilson, 2002):

$$\overline{ES}(M_{ES}) = \frac{\sum(w_i ES_i)}{\sum w_i}$$

- w_i وزن كل دراسة i .
- ES_i حجم التأثير لكل دراسة i .
- \overline{ES} او M_{ES} متوسط حجم التأثير.

وعليه فان متوسط حجم التأثير للدراسات هو مجموع حاصل ضرب حجم التأثير لكل دراسة فى وزنها مقسوماً على مجموع اوزان الدراسات. لو كان وزن كل الدراسات مساوياً للواحد الصحيح او كان وزن كل الدراسات متساوى فان المتوسط يكون مجموع احجام التأثير مقسوماً على عددها. والمعادلة السابقة تطبق على اى نوع من احجام التأثير سواء معامل الارتباط وتحويره Z_r حيث يقدر المتوسط الموزون لتحوير فيشر ثم يحول مرة اخرى الى معامل الارتباط، ومؤشرات فروق المتوسطات المعيارية مثل d , g ومؤشر Odds Ratio.

مثال: فى الجدول السابق المتضمن 22 دراسة يتم حساب المتوسط الموزون كالاتى:

$$\begin{aligned} \overline{Z}_r &= \frac{\sum(w_i Z_{ri})}{\sum w_i} \\ &= \frac{\sum(.583 \times 208.12) + .201 \times 297.81 + \dots + .509 \times 482.45 + .651 \times 202.92}{\sum(208.12 + 297.81 + \dots + 482.45 + 202.92)} \\ &= \frac{2764.36}{7152.21} = 0.387 \end{aligned}$$

وعليه فان معامل ارتباط بيرسون المقابل لقيمة متوسط معامل التحوير فيشر \overline{Z}_r كالتالى:

$$r = \frac{e^{(2 \times 0.387)} - 1}{e^{(2 \times 0.387)} + 1} = 0.368$$

وعليه يوجد معامل ارتباط موجب من الدرجة المتوسطة بين العدوان ورفض الاقران.

اختبارات الدلالة الاحصائية وفترات الثقة لمتوسط حجم التأثير

حساب متوسط حجم التأثير ليس كافي انما لابد من اختبار الدلالة الاحصائية وتقدير فترات الثقة حول هذا المتوسط التي تمثل مدى من احجام التأثير في المجتمع من المحتمل ان تقع فيه قيمة المتوسط المحسوبه من الدراسة. على سبيل المثال 95% فترات الثقة من 0.05 الى 0.45 حول حجم التأثير تشير الى ان احتمال 95% ان متوسط حجم التأثير في المجتمع يقع بين هاتين القيمتين، واذا لم تتضمن فترات الثقة القيمة صفر بالتالي فان متوسط حجم التأثير دال احصائياً عند مستوى دلالة احصائية 0.05 لـ 95% فترة ثقة.

ولا تتم اختبارات الفروض الصفرية او الدلالة الاحصائية بدون حساب الخطأ المعياري لمتوسط حجم التأثير ويقدر كالاتي (Hedges & Olkin, 1985):

$$SE_{ES} = \sqrt{\frac{1}{\sum w_i}}$$

• w_i وزن كل دراسة

مثال: في الجدول السابق يقدر الخطأ المعياري كالاتي:

$$SE_{\bar{Z}_r} = \sqrt{\frac{1}{\sum w_i}} = \sqrt{\frac{1}{7152.21}} = 0.118$$

لاحظ ان الدراسات لو تضمنت احجام عينات كبيرة بالتالي فان مجموع w_i يكون كبير بالتالي فالخطأ المعياري يكون صغير والعكس صحيح.

وبعد حساب الخطأ المعياري يعتمد عليه الباحث في اجراء الاستدلال احصائي وحساب فترات الثقة من خلال اختبار ولد Wald او اختبار Z (لا تخط بين اختبار Z للدلالة الاحصائية وتحويل فيشر Z_r) من خلال قسمة قيمة المعلم (المتوسط الموزون) مقسوماً على الخطأ المعياري كالاتي:

$$Z = \frac{\bar{ES}}{SE_{ES}}$$

- \overline{ES} متوسط حجم التأثير لكل الدراسات.
- $SE_{\overline{ES}}$ الخطأ المعياري لمتوسط حجم التأثير.

وهذا الاختبار يتم تقيمه وفقاً للتوزيع الاعتنالي المعياري لاختبار Z وقيمتها الجدولية او الحرجة عند مستوى دلالة احصائية $\alpha = 0.05$ هي 1.96 ($p \leq 0.05$) ولمستوى دلالة احصائية 0.01 هي 2.58 ($p \leq 0.01$) ، وعليه لو كانت قيمة Z المحسوبة اكبر من القيمة الجدولية او الحرجة بالتالي توجد دلالة احصائية بمعنى ان العلاقة او الفروق بين المتغيرين دالة احصائياً .

ويكون الفرض الصفري والبديل في هذه الحالة:

$$H_0: \overline{ES} = 0$$

$$H_A: \overline{ES} \neq 0$$

- \overline{ES} متوسط حجم التأثير في المجتمع.

وتقدر فترات الثقة من الخطأ المعياري كالاتي (Hedges & Olkin, 1985):

$$\overline{ES}_{lower} = \overline{ES} - Z_{(1-\alpha)} SE_{\overline{ES}}$$

$$\overline{ES}_{upper} = \overline{ES} + Z_{(1-\alpha)} SE_{\overline{ES}}$$

- \overline{ES}_{lower} حد الثقة الادنى.
- \overline{ES}_{upper} حد الثقة الاعلى.
- \overline{ES} متوسط حجم التأثير.
- $Z_{(1-\alpha)}$ القيمة الجدولية عند 0.05 او 0.01 لاختبار ذو ذيلين.
- $SE_{\overline{ES}}$ الخطأ المعياري لمتوسط حجم التأثير.

مثال: بالرجوع الى جدول البيانات فان قيمة اختبار Z كالاتي:

$$Z = \frac{\overline{ES}}{SE_{\overline{ES}}} = \frac{0.387}{0.118} = 32.70, P < .001$$

بالتالى فان متوسط حجم التأثير دال احصائياً بمعنى اكبر من الصفر (القيمة فى الفرض الصفرى) لان القيمة المحسوبة 32.70 اكبر من القيمة الجدولية 1.96 ، وعليه يوجد معامل ارتباط موجب من النوع المتوسط ودال احصائياً بين العدوان ورفض الاقران. ولحساب 95% فترات الثقة كالاتى:

$$E\bar{S}_{lower} = 0.387 - 1.96 \times 0.0118 = 0.363$$

$$E\bar{S}_{upper} = 0.387 + 1.96 \times 0.0118 = 0.410$$

ثم تحول قيمتى فترتى الثقة لتحويل فيشر الى معامل الارتباط المقابل لتصبح الحد الادنى $r = 0.348$ وهكذا للحد الاعلى $r = 0.388$ ، بالتالى فترتى الثقة 95% (0.348 , 0.388)، وعليه فان القيمة صفر المفترضة فى الفرض الصفرى لا تقع فى مدى فترات الثقة وعليه توجد دلالة احصائية.

وتكتب نتائج ما وراء التحليل فى الدراسة كالاتى: اتضح من نتائج ما ما وراء التحليل لـ 22 دراسة تناولت العلاقة بين العدوان والرفض ان متوسط معامل الارتباط 0.368 وهى دالة احصائياً حيث انها اكبر من الصفر ($P < .01$)، وان 95% فترات الثقة حول متوسط معامل الارتباط تراوحت من 0.348 الى 0.388

اختبار التجانس بين احجام التأثير Homogeneity among Effect Sizes

السؤال المهم المطروح هو ما اذا كانت احجام التأثير المختلفة التى قدر منها المتوسط العام تقدر نفس حجم التأثير فى المجتمع؟، وهذا السؤال يهتم بمدى توافر التجانس بين احجام التأثيرات المختلفة. فى التوزيع المتجانس فان تشتت او اختلاف احجام التأثير حول متوسطها لا يكون اكبر مما توقعناه من خطأ المعاينة المرتبط بعينات الافراد فى الدراسات الاولية التى قدر منها احجام التأثير، وعليه فان التوزيع المتماثل لاحجام التأثير للدراسات لا يختلف عن متوسط حجم التأثير الحقيقى فى المجتمع.

الاختبار الاحصائى الذى يرفض الفرض الصفرى للتجانس يشير الى ان الاختلاف او عدم التجانس بين احجام التأثير اكبر مما توقعناه من خطأ المعاينة، وعليه فان حجم التأثير لا يقدر متوسط حجم التأثير العام فى المجتمع، بكلمات اخرى وجود فروق او انحرافات بين

احجام التأثير فانها لا ترجع فقط الى خطأ المعاينة بل ترجع الى مصادر اخرى ربما تكون مرتبطة بخصائص الدراسات المختلفة في هذه الحالة لا يوجد تجانس ولكن تباين واختلافات.

انظر الى الشكل (1.8) تلاحظ ان كل الدراسات لها حجم تأثير واحد متضمناً في الخط الراسي الذي يعرض حجم التأثير العام في المجتمع في هذا الموقف يوجد تجانس وان اى ابتعاد لحجم التأثير في الدراسة (الدائرة) يكون متوقع نتيجة خطأ المعاينة وليس لاي مصدر اخر وفي هذا الموقف يوجد تجانس.

ولكن وجود اختلافات اكبر من خطأ المعاينة في هذه الحالة لا يصلح الخط المستقيم في الشكل (1.8) لعرض احجام التأثير انما توجد عدة بدائل للتعامل مع هذا التوزيع غير المتجانس نتناولها لاحقاً.

وتقيم الدلالة الاحصائية للتجانس لاحجام التأثير من خلال احصاء Q ويسمى اختبار test Hedges Q للتجانس وهذا الاختبار يعرض مدى الاختلاف في احجام التأثير بين الدراسات المختلفة المتضمنة في ما وراء التحليل ويقدر من كالتالي (Card, 2012; Hedges & Olkin, 1985; Lipsey & Wilson, 2001):

$$Q = \sum w_i (ES_i - \overline{ES})^2 = \sum (w_i ES_i^2 - \frac{(\sum w_i ES_i)^2}{\sum w_i})$$

- w_i وزن الدراسة i .
- ES_i حجم التأثير للدراسة i .
- \overline{ES} متوسط حجم التأثير الموزون العام لكل الدراسات.

درجات الحرية لاختبار Q:

$$df = k - 1$$

- K عدد الدراسات في ما وراء التحليل.

ويعتبر مربع انحراف حجم التأثير لكل دراسة عن متوسط حجم التأثير لكل الدراسات (الجزء الاول من المعادلة) هو افضل تقدير لحجم التأثير في المجتمع وهو الخط الراسى فى الشكل (1.8).

ولكن كيف يمكن استنتاج وجود تجانس من عدمه؟، يعامل احصاء Q فى ضوء توزيع كاي تربيع χ^2 درجات حرية k-1 وعلى ذلك يجب البحث فى جدول χ^2 لاستخراج قيمتها الجدولية واذا كانت:

$$\chi^2_{\text{الجدولية}} \geq \text{قيمة } Q \text{ المحسوبة}$$

نرفض الفرض الصفرى بمعنى عدم وجود تجانس او وجود اختلاف بين احجام التأثير اكثر مما توقعناه نتيجة خطأ المعاينة.

مثال: فى الجدول (1.8) تم حساب:

$$\sum w_i = 7152.21$$

$$\sum w_i ES_i^2 = 1359.60$$

$$\sum w_i ES_i = 2764.36$$

وبالتعويض فى المعادلة:

$$Q = \sum (w_i ES_i^2 - \frac{(\sum w_i ES_i)^2}{\sum w_i}) = 1359.60 - \frac{(2764.36)^2}{7152.21} = 291.17$$

حيث درجات الحرية: $df = k - 1 = 22 - 1 = 21$ وبالبحث فى جدول χ^2 الاتى:

الجدول (2.8): القيم الحرجة لـ χ^2 لدرجات الحرية ومستوى الدلالة الاحصائية.

df	p=.1	p=.05	p=.01	p=.005	p=.001	df	p=.1	p=.05	p=.01	p=.005	p=.001
1	2.706	3.841	6.635	7.879	10.828	45	57.505	61.656	69.957	73.166	80.077
2	4.605	5.991	9.210	10.597	13.816	50	63.167	67.505	76.154	79.490	86.661
3	6.251	7.815	11.345	12.838	16.266	55	68.796	73.311	82.292	85.749	93.168
4	7.779	9.488	13.277	14.860	18.467	60	74.397	79.082	88.379	91.952	99.607
5	9.236	11.070	15.086	16.750	20.515	65	79.973	84.821	94.422	98.105	105.988
6	10.645	12.592	16.812	18.548	22.458	70	85.527	90.531	100.425	104.215	112.317
7	12.017	14.067	18.475	20.278	24.322	75	91.061	96.217	106.393	110.286	118.599
8	13.362	15.507	20.090	21.955	26.124	80	96.578	101.879	112.329	116.321	124.839
9	14.684	16.919	21.666	23.589	27.877	85	102.079	107.522	118.236	122.325	131.041
10	15.987	18.307	23.209	25.188	29.588	90	107.565	113.145	124.116	128.299	137.208
11	17.275	19.675	24.725	26.757	31.264	95	113.038	118.752	129.973	134.247	143.344
12	18.549	21.026	26.217	28.300	32.909	100	118.498	124.342	135.807	140.169	149.449
13	19.812	22.362	27.688	29.819	34.528	110	129.385	135.480	147.414	151.948	161.581
14	21.064	23.685	29.141	31.319	36.123	120	140.233	146.567	158.950	163.648	173.617
15	22.307	24.996	30.578	32.801	37.697	130	151.045	157.610	170.423	175.278	185.571
16	23.542	26.296	32.000	34.267	39.252	140	161.827	168.613	181.840	186.847	197.451
17	24.769	27.587	33.409	35.718	40.790	150	172.581	179.581	193.208	198.360	209.265
18	25.989	28.869	34.805	37.156	42.312	160	183.311	190.516	204.530	209.824	221.019
19	27.204	30.144	36.191	38.582	43.820	170	194.017	201.423	215.812	221.242	232.719
20	28.412	31.410	37.566	39.997	45.315	180	204.704	212.304	227.056	232.620	244.370
21	29.615	32.671	38.932	41.401	46.797	190	215.371	223.160	238.266	243.959	255.976
22	30.813	33.924	40.289	42.796	48.268	200	226.021	233.994	249.445	255.264	267.541
23	32.007	35.172	41.638	44.181	49.728	210	236.655	244.808	260.595	266.537	279.066
24	33.196	36.415	42.980	45.559	51.179	220	247.274	255.602	271.717	277.779	290.556
25	34.382	37.652	44.314	46.928	52.620	230	257.879	266.378	282.814	288.994	302.012
26	35.563	38.885	45.642	48.290	54.052	240	268.471	277.138	293.888	300.182	313.437
27	36.741	40.113	46.963	49.645	55.476	250	279.050	287.882	304.940	311.346	324.832
28	37.916	41.337	48.278	50.993	56.892	300	331.789	341.395	359.906	366.844	381.425
29	39.087	42.557	49.588	52.336	58.301	350	384.306	394.626	414.474	421.900	437.488
30	40.256	43.773	50.892	53.672	59.703	400	436.649	447.632	468.724	476.606	493.132
35	46.059	49.802	57.342	60.275	66.619	450	488.849	500.456	522.717	531.026	548.432
40	51.805	55.758	63.691	66.766	73.402	500	540.930	553.127	576.493	585.207	603.446

وبالبحث في الجدول السابق بدرجات حرية 21 وبمستوى دلالة احصائية مثلاً 0.01 فان

$$\chi^2_{\text{الجدولية}} = 38.932 \text{ اذا:}$$

I^2 الجدولية (38.932) $Q >$ المحسوبة (291.71) وعليه يتم رفض الفرض الصفري بمعنى وجود اختلاف بين احجام التأثير اى عدم تجانس بينهم بمعنى ان احجام التأثير كلها لا تقدر حجم التأثير الحقيقي فى المجتمع وعلى ذلك فان احجام التأثير اختلفت عن حجم التأثير الحقيقي اختلافاً اكبر مما توقعناه نتيجة خطأ المعاينة. ويشير Card(2012) الى تحذيرين عند تفسير نتائج Q وهما:

الاول: هذا الاختبار يمدنا بمعلومات عن احتمال وجود او عدم وجود تجانس فقط لاحجام التأثير ولا يخبرنا عن مقدار Magnitude هذا الاختلاف او عدم التجانس . Heterogeneity

الثانى: يجب الاخذ فى الاعتبار القوة الاحصائية لهذ الاختبار ونقص هذه القوة للاختبار فيجب ان يكون الباحث على حذر فى تفسير النتائج غير الدالة احصائياً كدليل للتجانس(الفرض الصفري).

بدائل اختبار التجانس بين احجام التأثير

مدخل Hunter & Schmidt

افترض مدخل بديل لاختبار عد التجانس بين احجام التأثير لا يعتمد على اختبارات الدلالة. هذا المخل يجرى تبين حجم التأثير المقاس من الدراسة الى مكونين هما جزء يعزى الى خطأ المعاينة لعينات الافراد فى الدراسة والثانى يعزى الى الفروق بين الدراسات. يفترض ان يكون التوزيع متجانس لو ان 75% او اكثر من الاختلافات بين احجام التأثير ترجع الى خطأ المعاينة وهذه قاعدة افتراضية متعارف عليها وليست نقطة قطع صارمة.

مؤشر I^2

على الرغم ان احصاء Q والدلالة الاحصائية المرتبطة به من المؤشرات المفيدة لتحديد مدى وجود تجانس من عدمه بين احجام التأثير فى ما وراء التحليل الا انه لا يشير الى كيفية ان يكون هذا الاختلاف بين احجام التأثير. ومن المؤشرات المفيدة لتحديد التجانس فى ما وراء التحليل مؤشر I^2 ويفسر كنسبة مئوية للاختلاف بين احجام التأثير الموجود بين الدراسات

بالنسبة الى الاختلاف الكلى بين احجام التأثير (Higgins & Thompson, 2002; Huedo-Medina, Sanchez-Meca, Marin-Martinez, & Botella, 2006) ويتم تقديره كالاتى:

$$I^2 = \frac{\tau^2}{\tau^2 + \sigma^2} = \frac{Q - (k - 1)}{Q} \times 100, \text{ when } Q > (k - 1)$$

- τ^2 الاختلاف المقدر بين الدراسات.
- σ^2 الاختلاف داخل الدراسات.
- Q الاحصاء المستخدم لحساب الدلالة الاحصائية للتجانس.
- k عدد الدراسات.

والجزء الايمن من المعادلة يعتمد على احصاء Q وعدد الدراسات k وقيمة I^2 تكون تقريباً عند الصفر عندما تكون Q اقل من المتوقع تحت الفرض الصفري للتجانس. والمقام مكون من Q وهو عبارة عن التباين الكلى بين احجام التأثير بينما البسط يتكون من الاختلاف الكلى بين احجام التأثير مطروحاً منه الاختلاف المتوقع جراء خطأ المعاينة.

مؤشر I^2 هو النسبة بين الاختلاف بين الدراسات (الاختلاف الكلى مطروحاً من الاختلاف نتيجة خطأ المعاينة) بالنسبة الى الاختلاف الكلى بين الدراسات ويعبر عنه فى ضوء نسبة مئوية على مقياس من 0.0% الى 100% وعلى ذلك فان هذا المؤشر يعبر عن مقدار الاختلاف بين احجام التأثير فى ما وراء التحليل ويستخدم لمقارنة الاختلافات عبر الدراسات المختلفة لما وراء التحليل، وهذا المؤشر اقل استخداماً فى دراسات ما وراء التحليل كما ليس له معايير متفق عليها لتحديد مقدار الاختلاف سواء صغير او متوسط او كبير، ولكن Huedo-Medin et al. (2006) اقترحوا المعايير الاتية:

$I^2 \approx 0\%$ تجانس بين احجام التأثير.

$I^2 \approx 25\%$ قدر صغير من الاختلاف بين احجام التأثير.

$I^2 \approx 50\%$ قدر متوسط من الاختلاف.

$I^2 \approx 75\%$ قدر كبير من الاختلاف.

مثال: يقدر مؤشر I^2 لبيانات لـ 22 دراسة كالآتي:

$$I^2 = \frac{Q - (k - 1)}{Q} \times 100 = \frac{291.17 - (22 - 1)}{291.17} = 92.8\%$$

وهذا يعبر عن اختلافات كبيرة جداً بين احجام التأثير.

عدم الاستقلالية بين احجام التأثير

في هذا الفصل عرضنا لاهمية الاستقلالية لاجسام التأثير ومن المفضل الاعتماد على حجم تأثير واحد من الدراسة لان حساب اكثر من حجم تأثير من نفس الدراسة يحطم مسلمة الاستقلالية. في ما وراء التحليل يجب الاعتماد على العشوائية بقدر الامكان في عينة الدراسات المتضمنة في ما وراء التحليل نظراً لاستخدام اختبارات الفروض الاحصائية حول المتوسط الموزون لاجسام التأثير.

اجسام التأثير المتعددة من القياسات المتعددة

احد مصادر اجسام التأثير في الدراسة ان الدراسة يمكن ان تتضمن مقاييس مختلفة لنفس المفهوم او البناء في الدراسة، فمثلاً في الدراسات الـ 22 اعتمد احد المؤلفين على العلاقة بين العدوان والرفض من الاقران من وجهة نظر زملائه ($r = .556$) وايضاً من وجهة نظر المعلمين ($r = .338$)، واحد الدراسات اعتمدت على دراسة هذه العلاقة على فترات زمنية مختلفة مثلاً على مستوى تلاميذ الصف الثاني وعلى مستوى تلاميذ الصف الرابع.

في هذه المواقف يرى (Card 2012) يوجد اختياريين للحصول على حجم تأثير وحيد:

الاول: اختيار حجم التأثير الذي يقع في اهتمام الدراسة التي يقوم بها الباحث وهذا في ضوء معايير التضمين والاستبعاد التي اعتمد عليها الباحث ففي الدراسة الاولى يضع الباحث معيار تقدير العلاقة في ضوء تقديرات الاقران وليس تقدير المعلم. ومن الضروري وضع المعيار لصناعة القرار قبل اجراء ما وراء التحليل وليس في ضوء حجم التأثير.

الثانى: وهو الاكثر استخداماً فى الدراسات وهو اخذ متوسط احجام التأثير المتعددة فى نفس الدراسة لنفس المتغيرين موضع الاهتمام واستخدام متوسط احجام التأثير كمؤشر وحيد لحجم التأثير المتولد من الدراسة.

احجام التأثير المتعددة من مجموعات فرعية من المشاركين فى الدراسة

المصدر الاخر لاحجام التأثير المتعددة من نفس الدراسة هو احجام التأثير من المجموعات الفرعية من الدراسة (ذكور - اناث، المجموعات التجريبية المتعددة) الامر يتوقف على معايير التضمنين او الاستبعاد هل تهتم الدراسة بالعينة الاجمالية ام العينات الفرعية، فاذا كان الاهتمام بالعينات الفرعية بالتالى يجب حساب متوسط احجام التأثير من هذه المجموعات الفرعية ويجد اعتبارين فى هذه الحالة هما:

الاول: عند حساب متوسط المجموعات الفرعية فيجب الاعتماد على مجموع العينات للمجموعتين عند حساب الخطأ المعياري فى الدراسة.

الثانى: يجب حساب المتوسط الموزون لان العينات الفرعية الكبيرة تسهم بقدر اكبر فى المتوسط من العينات الصغيرة.

احجام التأثير من التقارير العديدة لنفس الدراسة

المصدر الثالث لعدم الاستقلالية لاحجام التأثير فى ما وراء التحليل هو ان نفس الدراسة يعاد نشرها اكثر من مرة فقد تكون رسالة دكتوراة او ماجستير ثم يعاد نشرها فى كتاب او دراسة فى مجلة علمية متخصصة، وقد تكون دراسة معروضة فى مؤتمر ثم يعاد نشرها فى مجلة متخصصة او تختلف احجام التأثير فى نفس الدراسة نتيجة لان الباحث قد يعتمد على مقياس فى الدراسة الاولى مختلف عن مقياس نفس الدراسة فى مكان النشر الاخر، فى هذه الحالة يجب اخذ متوسط احجام التأثير لنفس الدراسة ذات المقاييس المختلفة ومكان النشر المختلف وهذ نادر الحدوث فى دراسات ما وراء التحليل.

الفصل التاسع

نماذج تحليل ما وراء التحليل

Analysis Models for Meta-analysis

مقدمة

في الفصل السابق تناولنا مدخل لكيفية حساب متوسط حجم التأثير واختبار الدلالة الاحصائية وحساب فترات الثقة حول المتوسط وهذا المدخل يفترض التجانس بين احجام التأثير بمعنى ان الاختلافات بين احجام التأثير المقدره من الدراسات وحجم التأثير الحقيقي في المجتمع ترجع الى فقط خطأ المعاينة، وهذا المدخل في تحليل ما وراء التحليل يسمى نموذج التأثيرات المثبتة Fixed-effects model ويسميه Hedges (2009) بالنموذج الاستدلالي الشرطي Conditional Inference Model ولا يفضل استخدام هذا النموذج اذا ثبت ان الفروق او الاختلافات بين احجام التأثير ترجع الى المتغيرات المتفاعلة Moderator variables او خصائص الدراسات بالاضافة الى خطأ المعاينة. واذا وجدت اختلافات بين احجام التأثير اكبر مما نتوقعة من خطأ المعاينة ونتجة لخصائص الدراسات فان نموذج تحليل ما وراء التحليل يسمى نموذج التأثيرات العشوائية Random effects mode ويسمى النموذج غير الشرطي Unconditional Model.

ويعتبر اختيار نموذج التحليل المكون الاحصائي الاساسي لما وراء التحليل وهذا يعكس الهدف من التحليل، واختيار نموذج التحليل يتوقف على عدة عوامل منها الهدف من التحليل، نوعية البيانات المتاحة من الدراسات، معالجة الاختلاف بين احجام تأثير الدراسات، التوزيع الاحصائي لتقديرات احجام التأثير، الاختيار بين الطرق الكلاسيكية والطرق البيزانية (Mengersen, Schmid, Jennions, & Gurevitch, 2013)

وعرضنا في الفصل السابق اختبار Q لفحص مسلمة التجانس بين احجام التأثير المقاسة من الدراسات المختلفة واذا تحققت يستخدم نموذج التأثيرات المثبتة ولكن الدلالة الاحصائية تعني رفض الفرض الصفرى بمعنى اختلافات بين احجام التأثير اكبر مما توقعناه جراء خطأ

المعاينة في الدراسات، وهذا الاختلاف يلقي بتحدى كبير على استخدام النموذج المثبت لانه عند حساب المتوسط الموزون في الفصل السابق حدث موازنة لكل حجم تأثير في ضوء الخطأ المعياري الذي يأخذ في اعتبارة اخطاء المعاينة فقط وهذا كافي لاجراء التحليلات الاحصائية.

ولكن في ظل عدم التجانس بين احجام التأثير نتيجة مصدر اخر للاختلافات غير خطأ المعاينة بالتالي اجراء موازنة في ضوء الخطأ المعياري غير كافي فلا بد من اجراء موازنة لمصادر الاختلافات الاخرى بين احجام التأثير للدراسات، وعليه فان نموذج التأثيرات المثبتة غير ملائم لتحليل توزيع احجام التأثيرات غير المتجانسة.

كما ان تحليلات ما وراء التحليل تتعامل مع عينة صغيرة نسبياً من احجام التأثيرات او الدراسات تتراوح في معظمها من 15 دراسة الى 25 دراسة وهذا من شأنه ان تكون القوة الاحصائية لاختبار Q غير كافية لرفض الفرض الصفري حتى لو وجدت مصادر اختلاف بين الدراسات غير خطأ المعاينة وعليه فاذا لم تتحقق مسلمة نموذج التأثيرات المثبتة فان الباحث امامه ثلاث اختيارات حددها (Lipsey & Wilson (2001 بالتالي:

الاول: افترض ان الاختلافات بين احجام التأثير ترجع الى مصادر اختلاف عشوائية اخرى غير خطأ المعاينة لعينات الافراد في الدراسات وهذه المصادر لا يمكن تحديدها. وفي هذه الحالة يتبنى الباحث نموذج التأثيرات العشوائية Random effects model حيث يفترض ان التباين او الاختلاف بين احجام التأثير يرجع الى مكون او خطأ عشوائى اخر غير خطأ المعاينة ويجب تضمين هذا الخطأ الى النموذج الاحصائى لتحليل احجام التأثيرات، وهذا الخطأ الاضافى يتصرف مثل خطأ المعاينة حيث يؤدي الى ابتعاد او انحراف حجم التأثير المقاس عن حجم التأثير الحقيقى في المجتمع وعليه فان النموذج العشوائى يتعامل مع مصدرين للخطأ هما:

- خطأ معاينة على مستوى الدراسات Study Level حيث ان عينة الدراسات غير ممثلة تمثيلاً جيداً لمجتمع دراسات الظاهرة التي تتم عليها دراسات ما وراء التحليل.

- خطأ معاينة على مستوى الافراد او الاشخاص Subject level حيث ان عينة الافراد فى كل دراسة غير ممثلة تمثيلاً جيداً للمجتمع.

ولذلك لابد من اجراء تصحيح للمتوسط مرتين وليس مرة واحدة كما عرضنا فى الفصل السابق(من خطأ معاينة الافراد) ولا بد من تصحيح كل حجم تاثير من خطأ المعاينة للافراد وخطأ المعاينة للدراسات، وهذا يتطلب اجراءات احصائية مختلفة عما عرضناه فى تحليل نموذج التاثيرات المثبتة فى الفصل السابق متمثلة فى نموذج التاثيرات العشوائية.

الثانى: استخدم **نموذج التاثيرات المثبتة** وافترض مسلمة هو ان الاختلاف غير خطأ معاينة الافراد يكون منتظم Systematic مشتق من فروق غير محددة بين الدراسات، وهذا يتطلب تحليل الاختلاف بين احجام التاثير فى ضوء خصائص الدراسات التى تولدت منها احجام التاثير وفى هذه الحالة يفترض الباحث ان وجود اختلافات فى توزيع احجام التاثيرات نجمت عن مصادر غير اخطاء المعاينة للافراد ويكون هذا الاختلاف غير عشوائى والاختلافات الكبيرة بين احجام التاثير ترجع الى المتغيرات المتفاعلة مثل الخصائص المنهجية (الضبط - المعالجات المختلفة- التصميمات البحثية- الاجراءات- حجم العينة- نوعية العينة، وغيرها) وكذلك خصائص العينة وهذه العوامل تجعل حجم التاثير منخفض او كبير فى الدراسات. وفى هذه الحالة اجراء موازنة Weighting لحجم التاثير فى ضوء خطأ المعاينة للافراد يكون مقبول ومناسب. والاجراءات الاحصائية لدراسة الاختلاف بين احجام التاثير للدراسات المختلفة يسمى **نموذج التاثيرات المثبتة: تجزئة تباين احجام التاثير Fixed effects model: Partitioning effect size variance**

الثالث: افترض ان التباين بين احجام التاثير يرجع الى مصادر اخرى غير خطأ المعاينة هى مكونة من مصدرين احدهما عوامل منتظمة يمكن تحديدها (خصائص الدراسات) والآخر مصادر عشوائية، و الاجراءات الاحصائية لتحليل ما وراء التحليل تسمى **نموذج التاثيرات المختلطة Mixed effects model** حيث ان الاختلاف بين احجام التاثير لا يرجع الى خطأ المعاينة فقط بل الى مكونين اخرين احدهما منتظم(خصائص الدراسات) والآخر عشوائى(خصاً معاينة الدراسات من مجتمع الدراسات التى تناولت العلاقة بين المتغيرين). وهذا النموذج يتشابهة مع نموذج التاثيرات المثبتة: تجزئة تباين احجام التاثير فى

تحليل خصائص الدراسات كمتغيرات متفاعلة ولكنه يختلف عنه في انه يسمح بتضمين المكون العشوائى فى التحليل كتباين بواقى ليصبح فى النموذج بعد تقدير المكون المنتظم، بكلمات اخرى يفترض ان تباين احجام التأثير غير خطأ المعاينة يرجع الى مكونين احدهما منتظم حيث يعرض العلاقة بين خصائص الدراسات وحجم التأثير، والاخر يمثل الخطأ العشوائى نتيجة خطأ معاينة للدراسات الممثلة لمجتمع الدراسات التى تناولت المتغيرين ويجب تضمينه فى الوظيفة الموازنية **Weighting function** عند تحليل حجم التأثير حيث يعاد حساب المتوسط الموزون وفترات الثقة بعد اسخدام اوزان مختلفة وليس وزن واحد كما فى النموذج المثبت.

نموذج التأثيرات المثبتة **Fixed effects model**

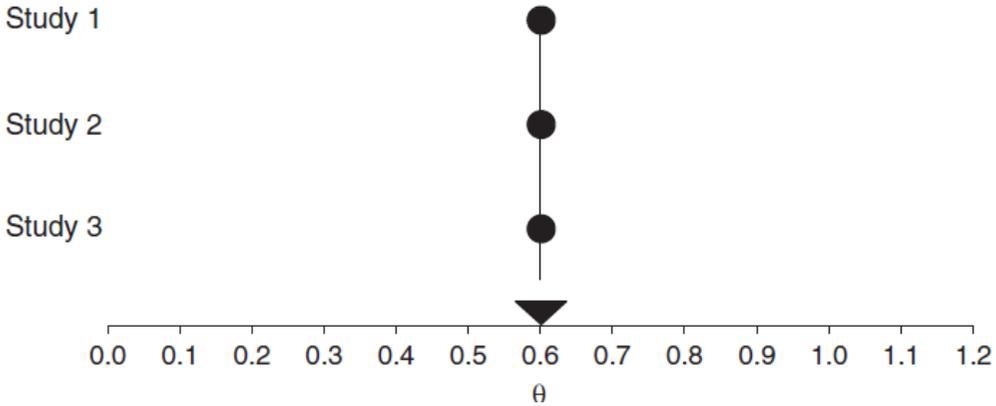
مصطلح التأثيرات المثبتة يشير الى حقيقة ان مستويات المعالجة فى التجربة تعتبر مثبتة ومصدر الاختلاف هى عينات او مجموعات الافراد والمتغير التابع. ويفترض هذا النموذج ان حجم التأثير المقدر من الدراسة يناظر حجم التأثير الحقيقى فى المجتمع مع وجود خطأ عشوائى نتيجة عوامل الصدفة المرتبطة بخطأ المعاينة فى الدراسة الناتج عن الاختيار غير العشوائى للعينة من مجتمع الدراسة وايضاً نتيجة ان العينة غير ممثلة تمثيلاً جيداً لمجتمع العينة وكذلك لصغر حجم العينة. ولاسباب نظرية ربما لا يعتقد الباحث ان هذه الافتراضات مناسبة او متوفرة لتوزيع احجام التأثير. فعلى سبيل المثال يعتقد بوجود فروق عشوائية بين الدراسات ترجع الى اختلافات فى الاجراءات والمواقع المختلفة التى اجريت فيها الدراسات وغيرها من العوامل التى تكون ابعد من خطأ المعاينة فى الدراسات، وربما يكون الاختلافات ترجع الى عينة الدراسات فى ما وراء التحليل حيث لا تكون ممثلة لمجتمع الدراسات تمثيلاً جيداً وهذا يودى الى خطأ معاينة خاص بالدراسات بالاضافة الى خطأ المعاينة الخاص بعينات الافراد فى الدراسة.

ويرى (Borenstein et al. (2009 ان النموذج المثبت فى ما وراء التحليل يفترض ان احجام التأثير للدراسات تقيس نفس حجم التأثير الحقيقى فى المجتمع بمعنى ان كل العوامل التى تؤثر فى حجم التأثير واحدة فى كل الدراسات ولذلك فان حجم التأثير هو نفسه (مثبت) عبر كل الدراسات والاختلاف عن حجم التأثير الحقيقى يرجع الى خطأ المعاينة.

وعلى ذلك فان نموذج التأثيرات المثبتة يفترض ان كل الدراسات فى ما وراء التحليل تاتى من نفس المجتمع وهذا يعنى ان حجم التأثير الحقيقى يكون هو نفسه عبر كل الدراسات وفى هذه الحالة يفترض وجود تجانس بين احجام التأثير المختلفة وان مصدر الاختلاف بين حجم التأثيرات يرجع الى مكونات من داخل الدراسات مثل خطأ المعاينة او خطأ القياس (Hegger, 2006). ويهتم هذا النموذج بالتباين داخل الدراسة نتجة اختلاف المعاينات ولا يهتم بالتباين بين الدراسات ويفترض هذا النموذج (Kline, 2013):

- وجود مجتمع واحد متجانس من الدراسات مع حجم تأثير حقيقى وحيد.
- اختلاف حجم التأثير للدراسة عن حجم التأثير الحقيقى فى المجتمع يرجع الى التباين داخل الدراسة بمعنى الى خطأ المعاينة.

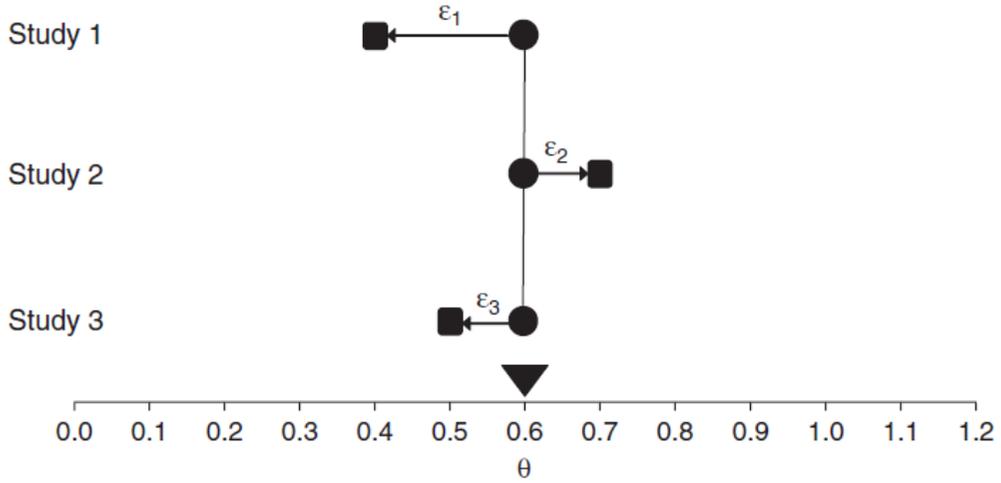
ويمكن توضيح ذلك بالشكل الاتى:



الشكل(1.9): نموذج التأثيرات المثبتة: التأثيرات الحقيقية.

فاذا افترضنا ان حجم التأثير العام او الحقيقى فى المجتمع هو 0.60 (كما معروض فى المستطيل على المحور الافقى فى الاسفل $\overline{0.6}$) ويرمز له بالرمز θ وحجم التأثير الحقيقى لكل دراسة معروض على الخط الراسى (دائرة). ويفترض نموذج التأثيرات المثبت ان حجم التأثير الحقيقى يجب ان يكون 0.60 وهذا معروض فى الدوائر ● التى تقع مباشرة على الخط المستقيم، ولكن هذا صعب حدوثه فى الواقع التطبيقى الامبريقى للدراسات لان حجم التأثير

يختلف من دراسة لآخرى نتيجة خطأ المعاينة في كل دراسة فلو كان هذا الخطأ مساوياً للصفر في الدراسة فان حجم التأثير لها يكون مساوياً لحجم التأثير في الحقيقي في المجتمع. ولكن في الممارسة العملية فان حجم العينة في كل دراسة يكون محدد بحجم معين ولذلك فيوجد قدراً من خطأ المعاينة يجعل حجم التأثير في الدراسة ليس هو نفسه حجم التأثير الحقيقي وقد تزداد قيمته او تنقص. ويمكن توضيح ذلك بالشكل الاتي:



الشكل (2.9): نموذج التأثيرات المثبتة في ظل تأثير اخطأ المعاينة للدراسات.

لاحظ في هذا الشكل ان حجم التأثير الحقيقي لكل دراسة ما زال 0.60 (الدوائر) بينما حجم التأثير المقاس (المربعات) \leftarrow اختلف عن حجم التأثير الحقيقي من دراسة الى اخرى.

ففي الدراسة الاولى خطأ المعاينة $(\epsilon_1) = -0.20$ وعليه فان حجم التأثير المقاس:

$$ES_1 = 0.60 - 0.20 = 0.40$$

وفي الدراسة الثانية خطأ المعاينة $(\epsilon_2) = 0.10$ وعليه فان حجم التأثير المقاس:

$$ES_2 = 0.60 + 0.10 = 0.70$$

وفي الدراسة الثالثة خطأ المعاينة $(\epsilon_3) = -0.10$ وعليه فان حجم التأثير المقاس:

$$ES_3 = 0.60 - 0.10 = 0.50$$

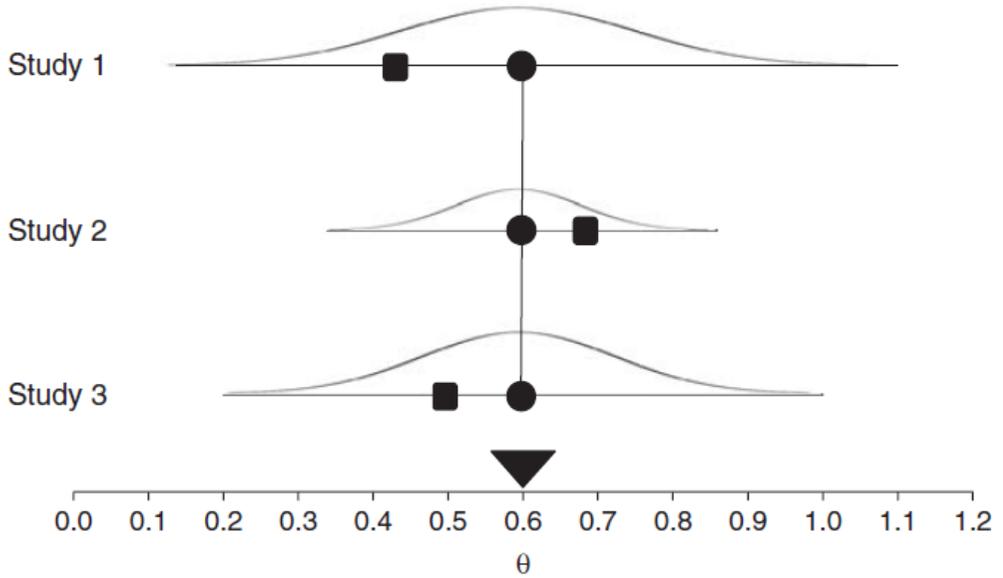
بصفة عامة فان حجم التأثير المقاس في كل الدراسة ES_i كالآتي:

$$ES_i = \theta + \epsilon_i$$

• θ حجم التأثير العام الحقيقي فى المجتمع.

• ε_i ابيسلون خطأ المعاينة للدراسة i .

لاحظ ان الدراسة الاولى ان حجم التأثير نقص بمقدار وحدتين وفى الدراسة الثانية زاد بمقدار وحدة وفى الدراسة الثالثة نقص بمقدار وحدة وعليه فان الخطأ فى اى دراسة عشوائى ويمكن تقدير التوزيع العينى لهذه الاخطاء كالاتى:



الشكل (3.9): التوزيع العينى لاختأ المعاينة فى النموذج المثبت.

وفى هذا الشكل نلاحظ ان توزيع اجسام التأثير المقاسة حول حجم التأثير الحقيقى فى ضوء المنحنى الطبيعى، فى الدراسة الاولى حيث حجم عينة صغير فمن المتوقع ان يقع حجم التأثير المقاس على جانبى حجم التأثير الحقيقى فى مدى من 0.20 الى 1.00، فى المقابل فى الدراسة الثانية حيث حجم العينة كبير نسبياً فمن المتوقع ان يقع حجم التأثير المقاس فى مدى ضيق من 0.40 الى 0.80.

نموذج التأثيرات العشوائية Random Effects Model

فى نموذج التأثيرات المثبتة افترضنا ان حجم التأثير الحقيقى نفسه فى كل الدراسات ولكن هذه المسلمة من الصعب تحقيقها. فلمجموعة من الدراسات المراد دمجها فى ما وراء التحليل فانها ليست متشابهة او متماثلة تماماً ان تعطى نفس حجم التأثير للعلاقة بين المتغيرين، فهذه

الدراسات لها منهجيات واجراءات وتصميمات واحجام عينات مختلفة بل مقاييس مختلفة، على ذلك فان الدراسات مختلفة في خصائصها وقيمة حجم التأثير مختلف من دراسة الى اخرى نتيجة.

فمثلاً الدراسات التي تتناول اثر العلاج السلوكى المعرفى لتخفيف الاكتئاب فان حجم التأثير يكون قوى فى الدراسات التي تتضمن عينة شديدة الاكتئاب مقارنة بدراسة اخرى تضمنت عينة متوسطة الاكتئاب كذلك يكون حجم التأثير كبير فى الدراسات التي تضمنت عدد من الجلسات اكبر من نظيرتها اقل جلسات، وعلى ذلك فان الدراسات المختلفة تنتج احجام تأثير مختلفة.

وفى الدراسات التي تتضمن برامج التدخل التربوى النفسى فان مقدار حجم التأثير يعتمد على المصادر المتاحة سواء المستوى الاقتصادى الاجتماعى للعينة او حجم الفصل او عمر العينة او مستوى ذكاء الافراد او عوامل اخرى تختلف من دراسة الى اخرى، وعلى ذلك ففى البرامج النفسية التربوية الاجتماعية يوجد كثيراً من التغيرات او المتغيرات الدخيلة التي تؤدى الى الاختلافات بين احجام التأثير للدراسات عبر الدراسات المختلفة التي تتناول نفس المتغيرات لنفس الظاهرة.

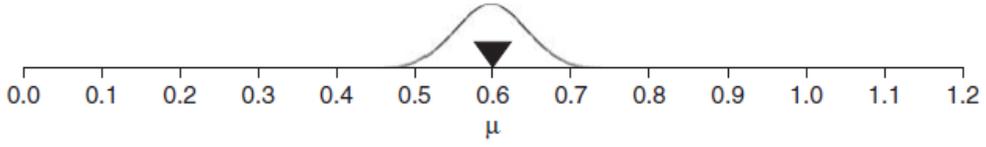
ويرى (Borenstein et al. (2009 ان احد الطرق للتعامل مع هذه الاختلافات عبر الدراسات هو استخدام نموذج التأثيرات العشوائية حيث يتم افتراض ان التأثيرات الحقيقية تتوزع اعتدالى وليس اجبارها كلها على خط راسى كما فى النموذج المثبت بالتالى لا يوجد حجم تأثير حقيقى لكل الدراسات انما يختلف من دراسة الى اخرى.

ويفترض هذا النموذج (Kline (2013:

• يوجد توزيع لمجتمع احجام التأثير بمعنى وجود حجم تأثير حقيقى لكل دراسة وليس حجم تأثير حقيقى واحد لكل الدراسات.

• يوجد مصدرين لتباين الخطأ احدهما تباين داخل الدراسات Within-studies Variance نتيجة خطأ المعاينة والآخر التباين بين الدراسات Between-studies Variance الذى يختص بتوزيع كل مجتمعات احجام التأثير حول متوسط حجم التأثير العام للمجتمعات. وهذا التباين فى مجمله من مصادر عشوائية لا يمكن تحديدها.

ويمكن توضيح ذلك بالشكل التالى:



الشكل (4.9): توزيع احجام التأثير الحقيقية في نموذج التأثيرات العشوائية.

وفي الشكل السابق فان متوسط كل احجام التأثير الحقيقية 0.60 ولكن احجام التأثير للدراسات تتوزع اعتدالياً حول هذا المتوسط وطول او مسافة المنحنى تقترح ان معظم احجام التأثير الحقيقية تقع في المدى من 0.50 الى 0.70.

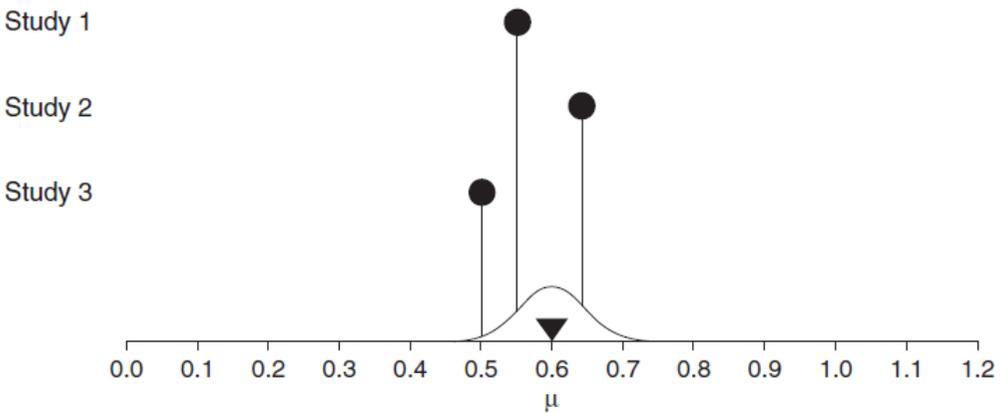
مثال: افترض ان دراسة ما وراء التحليل لثلاثة دراسات وان احجام التأثير الحقيقية لهذه الدراسات هي:

$$\theta_1 = 0.50$$

$$\theta_2 = 0.55$$

$$\theta_3 = 0.65$$

ولو ان حجم اعينات في الدراسات الثلاثة كبيراً كبيراً كافياً فان خطأ المعاينة يكون صفر ويكون حجم التأثير المقدر هو نفسة حجم التأثير الحقيقي لهذه الدراسة ويمكن التعبير عن ذلك بالشكل الاتي (Borenstien et al., 2009):



الشكل (5.9): التأثيرات الحقيقية في نموذج التأثيرات العشوائية.

في هذا الشكل لاحظ ان حجم التأثير الحقيقي العام في المجتمع (μ) يساوي 0.60 (على الخط الافقى) بينما حجم التأثير الحقيقي للدراسة الاولى 0.50 بالتالي حدث ابتعاد لحجم التأثير

للداسة الاولى عن حجم التأثير الحقيقى العام فى المجتمع بحوالى 0.10 وهذا يمثل نوع من الخطأ لاحظ ان هذا الخطأ لا يرجع الى خطأ المعاينة لاننا افترضنا ان خطأ المعاينة مساوى للصفر .

ولكن حجم العينة فى كل دراسة ليس ممثلاً للمجتمع تمثيلاً كافياً وعليه فان خطأ المعاينة لا يساوى صفر . فاذا كان حجم التأثير الحقيقى للدراسة θ_i فان حجم التأثير المقاس فى الدراسة ES_i يكون اكبر او اصغر من حجم التأثير الحقيقى θ_i نتيجة خطأ المعاينة. ففى الشكل (7.8) الدراسة الثالثة لها حجم تاثير حقيقى 0.50 وخطأ المعاينة -0.10- بالتالى حجم التأثير المقدر للدراسة 0.40.

وفى الشكل (6.9) يتضح ان المسافة بين المتوسط العام الحقيقى فى المجتمع (μ) وحجم التأثير المقاس فى الدراسة يتكون من جزئين هما:

- المسافة بين المتوسط العام لحجم التأثير الحقيقى (μ) وحجم التأثير الحقيقى للدراسة θ_i ويطلق على هذا الخطأ او الفرق زيتا ζ_i .

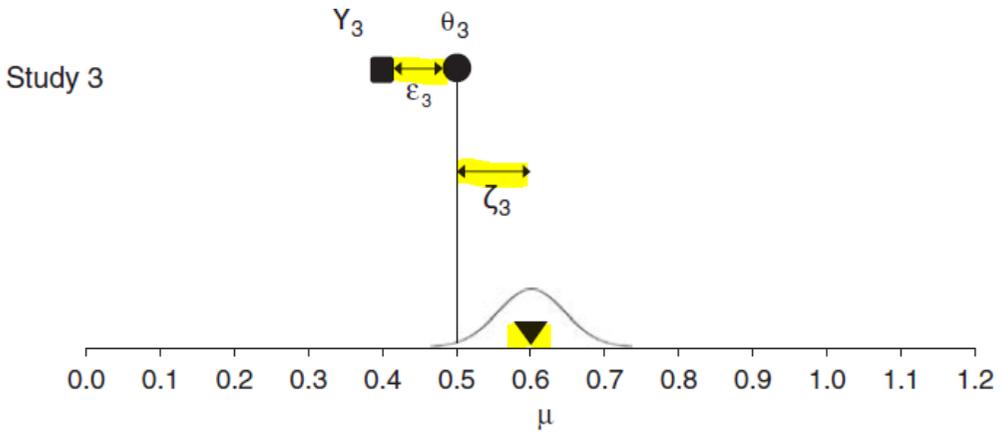
- المسافة بين حجم التأثير الحقيقى للدراسة θ_i وحجم التأثير المقاس للدراسة ES_i وهذا الفرق نتيجة خطأ المعاينة فى الدراسة ويطلق عليه الخطأ ابيسلون ε_i .

وعلى فان حجم التأثير المقاس فى الدراسة كالتالى:

$$ES_i = \mu + \zeta_i + \varepsilon_i$$

- ES_i حجم التأثير المقاس فى الدراسة i .
- μ متوسط توزيع احجام التأثير فى المجتمع.
- ζ_i (زيتا) انحراف او ابتعاد ثابت لحجم التأثير المقاس عن متوسط حجم التأثير الحقيقى.
- ε_i (ابيسلون) انحراف شرطى لحجم التأثير الحقيقى للدراسة نتيجة خطأ المعاينة.

ويمكن توضيح ذلك بالشكل الاتى (Borenstien et al., 2009):



الشكل (6.9): حجم التأثير الحقيقي والمقاس في نموذج التأثيرات العشوائية.

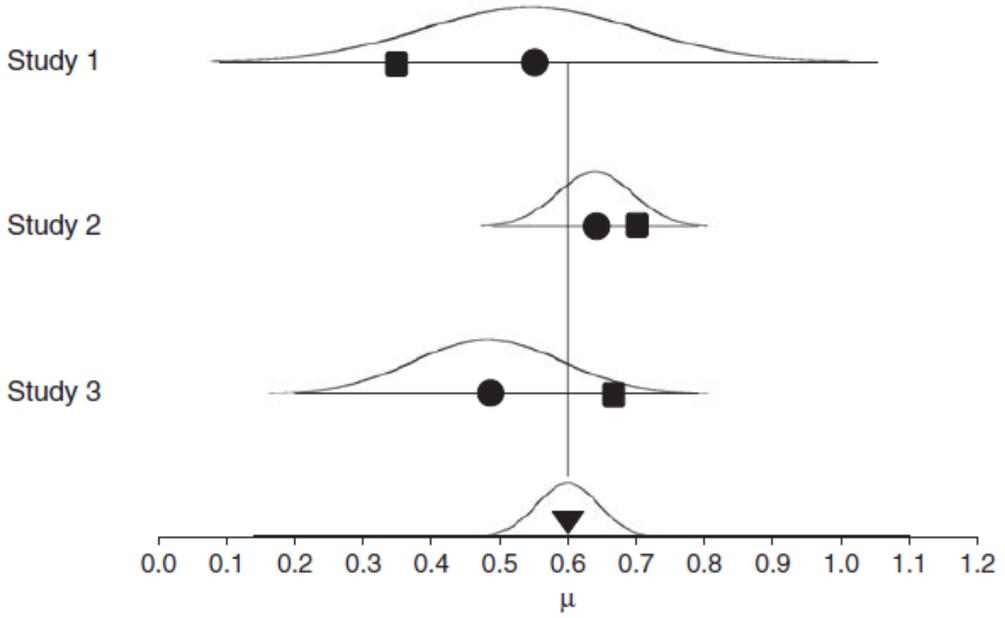
فالدراسة الثالثة متوسط حجم التأثير الحقيقي لمجتمع الدراسات $(\mu) = 0.60$ ، بينما حجم التأثير الحقيقي في الدراسة $\theta_3 = 0.50$ ، وعليه فان المسافة بين متوسط حجم التأثير الحقيقي الحقيقي للدراسات وحجم التأثير الحقيقي للدراسة -0.10 وهذا الخطأ نشأ نتيجة خطأ المعاينة لعينة الدراسات في ما وراء التحليل.

والفرق بين حجم التأثير الحقيقي للدراسة θ_3 وحجم التأثير المقاس $(Y_3)ES_3$ يساوى -0.10 وهذا يرجع الى خطأ المعاينة للأفراد داخل الدراسات وعليه فان الفرق بين متوسط حجم التأثير الحقيقي للدراسات وحجم التأثير المقاس في الدراسة يساوى -0.20

وعليه فان:

$$(Y_3)ES_3 = 0.60 - 0.10(\zeta_3) - 0.10(\epsilon_3) = 0.40$$

والمسافة من متوسط حجم التأثير الحقيقي للدراسات μ (المستطيل) الى حجم التأثير الحقيقي للدراسة θ_i (الدوائر) يعتمد على الانحراف المعياري لتوزيع التأثيرات الحقيقية عبر الدراسات ويطلق عليه تاو τ او تباينها τ^2 ويكون توزيعها اعتدالي كالتالى:



الشكل (7.9): توزيع احجام التأثير الحقيقية للدراسات في نموذج التأثيرات العشوائية.

لاحظ ان احجام التأثير الحقيقية للدراسات توزعت عشوائياً حول متوسط التأثيرات الحقيقية في نموذج التأثيرات العشوائية بينما في نموذج التأثيرات المثبتة تم اجبار كل احجام التأثير ان تقع على الخط الراسى.

الفروق بين نموذج التأثيرات المثبتة والعشوائية في ما وراء التحليل

توجد العديد من الفروق المفاهيمية، التحليلية عند حساب المتوسط الموزون ودلالاتها الاحصائية، والتفسيرية بين النموذجين كالاتى (Card, 2012; Kisamore & Brannick, 2008; Mengersen et al., 2013):

الفروق المفاهيمية النظرية Conceptual Differences

في النموذج المثبت الفرق بين حجم التأثير الوحيد الحقيقي في المجتمع θ وحجم التأثير المقاس في الدراسة نتيجة خطأ المعاينة ϵ_3 ، ويتلامس حجم التأثير الحقيقي للدراسات مع حجم التأثير الحقيقي في المجتمع، اى ان متوسط حجم التأثير ثابت لكل الدراسات.

بينما في النموذج العشوائى حجم التأثير الحقيقي للدراسة θ لا يتلاقى او يتلامس مع متوسط حجم التأثير الحقيقي العام للدراسات في المجتمع μ بل يتوزع حوله بانحراف معيارى τ ، وعليه

فان متوسط حجم التأثير العام للدراسات μ بانحراف معيارى تاو τ وليس حجم تأثير وحيد فى المجتمع θ كما فى النموذج المثبت، ويفترض ان كل حجم تأثير مقاس يختلف عن متوسط مجتمع احجام التأثير فى ضوء خطأ المعاينة على مستوى الافراد بالاضافة الى قيمة اخرى للتباين من مصادر اخرى نتيجة الاختلاف بين الدراسات ويفترض انها تتوزع عشوائياً . وعليه فان تباين حجم التأثير المقاس يتكون من مكونين (Lipsey & Wilson, 2001) :

- تباين مرتبط بخطأ المعاينة للافراد على مستوى الدراسة.
- تباين مرتبط بالتاثيرات العشوائية الاخرى نتيجة الاختلاف بين عينة الدراسات.

ويمكن التعبير عن التباين المرتبط بحجم التأثير المقاس كالاتى:

$$S_T^2 = S_\theta^2 + S_i^2$$

- S_θ^2 التباين نتيجة مصادر عشوائية او نتيجة الاختلاف بين الدراسات.
- S_i^2 التباين المرتبط بخطأ المعاينة للافراد.

الفروق التحليلية او الحسابية Analytic differences

حجم التأثير المقاس فى الدراسة فى النموذج المثبت هى حاصل جمع:

- حجم التأثير الحقيقى الوحيد فى المجتمع θ .
- الخطأ ε نتيجة خطأ المعاينة فى الدراسات.

بينما حجم التأثير المقاس فى النموذج العشوائى حاصل جمع:

- μ متوسط توزيع احجام التأثير فى المجتمع.
- ζ_i (زيتا) انحراف او ابتعاد ثابت لحجم التأثير المقاس عن متوسط حجم التأثير الحقيقى نتيجة خطأ معاينة الدراسات فى ما وراء التحليل.
- ε_i (ايبسلون) انحراف شرطى لحجم التأثير الحقيقى للدراسة نتيجة خطأ المعاينة الافراد فى الدراسات.

ما وراء التحليل باستخدام النموذج المثبت يهدف الى الوصول الى استنتاجات شرطية Conditional Inferences حول مجموعة محددة من الدراسات المتضمنة في التحليل فقط وليس على مجتمع الدراسات التي تناولت المتغيرين بمعنى " ان متوسط حجم التأثير لـ 15 دراسة....." ولا يعمم على مجتمع الدراسات التي تناولت المتغيرين ويمكن القول ان نتائج نموذج التأثيرات المثبتة وصفية .

بينما ما وراء التحليل باستخدام نموذج التأثيرات العشوائية يهدف الى الوصول لاستنتاجات غير مشروطة Unconditional Inferences ليست قاصرة على عينة الدراسات في التحليل بل الى مجتمع الدراسات المحتملة التي تناولت دراسة العلاقة بين المتغيرين بمعنى ان " تراث الدراسات للعلاقة بين المتغيرين Y, X توصل الى ان حجم التأثير او العلاقة بين المتغيرين....." ، ويمكن القول ان نتائج نموذج التأثيرات العشوائية استدلالية.

ولذلك فان نموذج التأثيرات العشوائية اكثر مناسبة في الدراسات النفسية التربوية الاجتماعية نتجة تراكم كم هائل من الدراسات للوصول الى استنتاجات اكثر تعميمية وليست قاصرة على عينة الدراسات في التحليل.

وعلى ذلك فان نموذج التأثيرات العشوائية لايفترض ان الدراسة مشتقة من نفس المجتمع كما هو الحال في نموذج التأثيرات المثبتة بل انها تمثل مجتمع فريد في حد ذاتها، وفي هذه الحالة يوجد مصدرين للاختلافات احدهما من الدراسة كما هو الحال في نموذج التأثيرات المثبتة والاخر من الاختلافات بين الدراسات ولذلك يسمى بنموذج عدم التجانس (Egger, 2006, Mengersen et al., 2013)

العمليات الحسابية لنموذج التأثيرات العشوائية

تناولنا في الفصل الثامن تحليل نموذج التأثيرات المثبتة من حساب المتوسط الموزون لاحجام التأثير في ضوء الخطأ المعياري، وحساب اختبار التجانس وفترة الثقة والدلالة الاحصائية باستخدام اختبار Z . اما عن العمليات الحسابية لنموذج التأثيرات العشوائية فانها تختلف عن النموذج المثبت نظراً لوجود مكون اخر للتباين يتمثل في المكون العشوائى ϵ_i ، وعليه فلا بد من

اجراء موازنة لكل حجم تأثير فى ضوء تعديلين احدهما خاص بالخطأ المعيارى المرتبط بخطأ المعاينة للافراد والآخر مرتبط بالمكون العشوائى الذى ربما يرجع فى جزء منه الى خطأ معاينة الدراسات المتضمنة فى ما وراء التحليل.

ويرى (Card 2012) ان ماوراء التحليل للنموذج العشوائى يتضمن اربع خطوات عامة هي كالاتى:

اولاً: تقدير التجانس او عدمه بين احجام التأثير للدراسات من خلال اختبار Q.

ثانياً: تقدير التباين بين احجام التأثير فى المجتمع.

ثالثاً: استخدام التباين لاحجام التأثير فى المجتمع للحصول على اوزان التأثيرات العشوائية Random-effects weights لاحجام التأثير المقاسة فى الدراسات.

رابعاً: استخدام اوزان التأثيرات العشوائية لتقدير متوسط التأثيرات العشوائية لاحجام التأثير والاختطأ المعيارية المرتبطة بها لاستخدامها فى حساب الدلالة الاحصائية وفترات الثقة.

الخطوة الاولى: تقدير التجانس من عدمه بين احجام التأثيرات: يقدر التجانس من عدمه من خلال اختبار Q كالاتى:

$$Q = \sum w_i (ES_i - \overline{ES})^2 = \sum (w_i ES_i^2 - \frac{(\sum w_i ES_i)^2}{\sum w_i})$$

- w_i وزن الدراسة i .
- ES_i حجم التأثير للدراسة i .
- \overline{ES} متوسط حجم التأثير الموزون العام لكل الدراسات.

درجات الحرية لاختبار Q:

$$df = k - 1$$

- K عدد الدراسات فى ما وراء التحليل.

وفى الفصل السابق تم تقدير اختبار $Q = 291.17$ وبمقارنتها بقيمة احصاء كاي تربيع الجدولية اتضح ان قيمتها اكبر من القيمة الجدولية لكاي تربيع، وعليه رفض الفرض الصفري بالتالى توجد دلالة احصائية بمعنى لا يوجد تجانس او الاختلاف بين احجام التأثير اكبر مما هو متوقع نتيجة خطأ المعاينة بين احجام التأثير لـ 22 دراسة فى ما وراء التحليل. وهذا الاستنتاج يفرض على الباحث استخدام نموذج التأثيرات العشوائية وليس النموذج المثبت لان النموذج المثبت يتطلب تجانس بين احجام التأثير المتضمنة فى دراسة ما وراء التحليل.

الخطوة الثانية: تقدير تباين مجتمع احجام التأثير Population Variability : وعدم التجانس بين احجام التأثير لا يرجع فقط الى خطأ المعاينة بل ايضاً الى كون عشوائى بسبب انحرافات حقيقية عن المتوسط العام لحجم التأثير الحقيقى لمجتمعات حجم التأثير ويتم تقدير تباين احجام التأثير الحقيقية للدراسات من خلال معلم τ^2 وهو التباين احجام التأثير الحقيقية للدراسات ولكن الباحث لا يعرف احجام التأثير الحقيقة ولكن يمكن تقدير تباينها كالاتى
:(Borenstien et al., 2009; Card, 2012)

$$\tau^2(v_{\theta}) = \frac{Q - (K - 1)}{(\sum w_i) - \frac{(\sum w_i^2)}{(\sum w_i)}}$$

- Q احصاء التجانس.
- K عدد الدراسات او احجام التأثير.
- w_i وزن كل دراسة من خطأ المعاينة.

ولو ان قيمة هذه المعادلة بالسالب فان تباين مجتمع احجام التأثير صفر وهذا يحدث لو ان توزيع احجام التأثير متجانس وهذا يتطابق مع نموذج التأثيرات المثبتة بالتالى شروط النموذج العشوائى لم تتحقق.

ولتقدير تباين مجتمع احجام التأثير الحقيقية للدراسات الـ 22 فى الجدول الاتى(انظر : Card, 2012):

الجدول (1.9): بيانات الـ 22 دراسة وحساباتها فى النموذج العشوائى.

Sample size (N)	Age (years)	Effect size (Z _r)	Standard error (SE)	Weight (w)	wES	wES ²	w ²	w*	w*ES
228	9.2	.583	0.0693	208.12	121.27	70.66	43315.0	21.94	11.51
491	9.4	.201	0.0579	297.81	59.74	11.98	88693.4	22.66	4.49
65	4.5	.322	0.1325	56.95	18.34	5.90	3242.9	17.14	5.34
458	8.0	.624	0.0484	427.07	266.55	166.37	182388.9	23.20	12.85
929	14.7	.162	0.0346	835.96	135.72	22.03	698829.8	23.83	3.84
904	7.5	.349	0.0347	831.95	290.69	101.57	692141.3	23.83	8.00
74	6.0	.419	0.1223	66.89	28.02	11.73	4474.6	17.95	7.11
151	9.5	.721	0.0855	136.66	98.48	70.97	18675.3	20.80	12.84
150	16.0	.628	0.0845	139.90	87.90	55.23	19570.9	20.87	11.62
590	9.0	.655	0.0426	550.97	361.12	236.69	303565.1	23.48	13.51
180	4.8	.039	0.0831	144.75	5.70	0.22	20953.4	20.98	0.83
139	3.6	.375	0.0892	125.71	47.14	17.68	15802.7	20.52	7.35
132	4.1	.049	0.0922	117.70	5.81	0.29	13853.1	20.30	1.00
60	4.6	.000	0.1386	52.05	0.00	0.00	2709.6	16.67	0.00
839	14.5	.339	0.0381	689.52	233.68	79.19	475442.6	23.69	7.73
262	8.7	-.048	0.0642	242.93	-11.73	0.57	59016.0	22.28	-1.07
266	9.5	.489	0.0636	246.94	120.85	59.15	60977.2	22.31	10.12
209	15.5	.258	0.0715	195.52	50.54	13.06	38226.3	21.79	5.51
314	9.0	.162	0.0589	288.73	46.71	7.56	83367.0	22.61	3.63
881	8.0	.519	0.0351	810.71	420.41	218.01	657255.5	23.81	11.35
517	8.0	.509	0.0455	482.45	245.37	124.79	232756.5	23.34	10.94
228	9.0	.651	0.0702	202.92	132.06	85.94	41174.7	21.88	12.52
				7152.21	2764.36	1359.60	3756432	475.89	161.02

وبتقدير قيمة تباين مجتمع احجام التأثير الحقيقية كالاتى:

$$\tau^2 = \frac{Q - (K - 1)}{(\sum w_i) - \frac{(\sum w_i^2)}{(\sum w_i)}} = \frac{291.17 - (22 - 1)}{7152.21 - \frac{3756432}{7152.21}} = \frac{270.17}{6627} = 0.0408$$

الخطوة الثالثة: حساب اوزان التأثيرات العشوائية: بعد تقدير تباين مجتمع احجام التأثير فان الخطوة التالية حساب اوزان التأثيرات العشوائية لكل دراسة. المنطق من تقدير الاوزان مماثل لما تم حسابه من موازنة او تصحيح حجم التأثير فى نموذج التأثيرات المثبتة من عدم الدقة لحجم التأثير من الخطأ المعيارى نتيجة خطأ المعاينة للحصول على حجم تأثير اكثر دقة وفى هذا

التصحيح فان حجم التأثير المصحح او الدقيق هو دالة وظيفية نتيجة حجم التأثير المقاس وانحراف المعاينة فى كل دراسة.

وفى النموذج العشوائى فان حجم التأثير يكون وظيفة لحجم التأثير المقاس بالاضافة الى مكونين من عدم الدقة Imprecision وليس مكون واحد كما فى نموذج التأثيرات المثبتة هما:

- ζ_i (ابساى) انحراف او ابتعاد ثابت لحجم التأثير المقاس عن متوسط حجم التأثير الحقيقى نتيجة خطأ معاينة الدراسات فى ما وراء التحليل، ويقدر من خلال تباين مجتمع احجام التأثيرات τ^2 .
- ε_i (زيتا) انحراف شرطى لحجم التأثير الحقيقى للدراسة نتيجة خطأ المعاينة الافراد فى الدراسات، ويقدر من خلال الخطأ المعيارى SE_i .

وعلى ذلك فان اوزان التأثيرات العشوائية تتكون من تباين مجتمع احجام التأثير بين الدراسات نتيجة عوامل مختلفة والخطأ المعيارى الخاص بكل دراسة.

وتقدر اوزان التأثيرات العشوائية كالاتى:

$$w_i^* = \frac{1}{\tau^2 + SE_i^2}$$

- w_i^* اوزان التأثيرات العشوائية.
- τ^2 تباين مجتمع احجام التأثيرات.
- SE_i الخطأ المعيارى المرتبط بحجم التأثير لكل دراسة.

وعلى ذلك يتم تصحيح حجم التأثير لكل دراسة من خلال اوزان التأثيرات العشوائية. بالرجوع الى جدول البيانات فى جدول (1.9) تم حساب w_i^* لكل دراسة ثم حساب حاصل ضرب وزن التأثير العشوائى $w_i^* \times$ حجم التأثير ES_i لكل دراسة.

وفى الجدول (1.9) بالنسبة للدراسة الاولى فان التصحيح فى ضوء نموذج التأثيرات المثبتة كالاتى:

$$w_1 = \frac{1}{SE^2} = \frac{1}{(0.0693)^2} = 208.12$$

ولكن فى النموذج العشوائى تم حساب وزن جديد كدالة لكلاً من تباين مجتمع احجام التأثير 0.0408 والخطأ المعيارى لكل دراسة SE=0.0693 وعليه يصبح تصحيح حجم التأثير للدراسة كالاتى:

$$w_i^* = \frac{1}{\tau^2 + SE_i^2} = \frac{1}{0.0408 + 0.0693^2} = 21.94$$

لاحظ ان قيمة تصحيح او وزن حجم التأثير للدراسة الاولى فى نموذج التأثيرات المثبتة هى 208.12 بينما قيمة التصحيح فى ضوء نموذج التأثيرات العشوائية 21.94.

وبالنظر الى اوزان التأثيرات العشوائية لـ 22 دراسة (العمود الثانى يميناً) فى الجدول تلاحظ الاتى:

- قيمة هذه الاوزان w_i^* اصغر بكثير من اوزان التأثيرات المثبتة w_i وهذا يعنى ان مجموعها عبر الدراسات يكون اصغر من نظيراتها فى النموذج المثبت وهذا يعنى ان الخطأ المعيارى لمتوسط حجم التأثير يكون كبير.
- على الرغم ان كل الدراسات لها تقريباً نفس الترتيب النسبى فى ضوء نموذج التأثيرات العشوائى او المثبتة (الدراسة التى لها اعلى وزن على احدهما تكون لها اعلى وزن على النموذج الاخر) الا ان التناقض فى الاوزان عبر الدراسات اقل بالنسبة للنموذج العشوائى مقارنة بالنموذج المثبت.

الخطوة الرابعة: حساب متوسطات التأثيرات العشوائية والاستدلال حولها: الخطوة الاخيرة فى تحليل نموذج التأثيرات العشوائية تقدير متوسط حجم التأثير ثم اجراء اختبارات الدلالة الاحصائية وحساب فترات الثقة حول هذا المتوسط. ويقدر متوسط حجم التأثير تماماً مثل متوسط حجم التأثير فى النموذج المثبت مع استبدال الوزن w_i بالوزن w_i^* وتصبح المعادلة كالتالى (Card, 2012; Lipsey & Wilson, 2001):

$$\overline{ES} = \frac{\sum(w_i^* ES_i)}{\sum w_i^*}$$

ويتطبيق هذه المعادلة على الـ 22 حجم التأثير في الجدول:

$$\bar{z}_r = \frac{\sum(w_i^* ES_i)}{\sum w_i^*} = \frac{161.02}{475.89} = 0.338$$

لاحظ انه تم تحويل حجم التأثير معامل الارتباط r الى تحويل فيشر Z_i ثم قدر متوسط تحويلات فيشر لكل الدراسات \bar{z}_r وقيمة 0.338 ثم تعاد هذه القيمة مرة اخرى الى متوسط معامل الارتباط وهي $\bar{r} = 0.326$. انظر الفصل(الثامن) نموذج التأثيرات المثبتة حيث ان قيمة $\bar{z}_r = 0.387$ بينما $\bar{r} = 0.387$.

اي ان متوسط حجم التأثير في النموذج العشوائى اقل من متوسط حجم التأثير في النموذج المثبت وهذا يرجع الى انه تم تصحيح حجم التأثير في النموذج العشوائى من نوعين من الخطأ بينما في النموذج المثبت تم تصحيحها من خطأ واحد فقط وهو خطأ المعاينة.

ويقدر الخطأ المعياري لمتوسط حجم التأثير كالاتى (Card, 2012; Lipsey & Wilson, 2001):

$$SE_{\bar{z}_r} = \sqrt{\frac{1}{\sum(w^*)}} = \sqrt{\frac{1}{475.89}} = 0.0458$$

لاحظ ان الخطأ المعياري في النموذج العشوائى اكبر من الخطأ المعياري في النموذج المثبت 0.0118(انظر الفصل الثامن).

وتقدر الدلالة الاحصائية من خلال اختبار Z (للدلالة وليس تحويل فيشر):

$$Z = \frac{\bar{z}_r}{SE_{\bar{z}_r}} = \frac{0.338}{0.0458} = 7.38, P < 0.01$$

ويتضح ان قيمة Z اكبر من 1.96 (قيمة Z الجدولية عند مستوى دلالة احصائية 0.05 لاختبار ذو ذيلين) او اكبر من 2.58 (قيمة Z الجدولية عند مستوى دلالة احصائية 0.01 لاختبار ذو ذيلين)، على ذلك يرفض الفرض الصفرى اى توجد علاقة ارتباطية موجبة ودالة احصائياً بين العدوان ورفض الاقران.

لاحظ ان قيمة Z للنموذج العشوائى اقل بكثير من نظيرتها للنموذج المثبت 32.70(انظر الفصل الثامن)، وهذا يرجع الى ان الخطأ المعياري(المقام) فى النموذج العشوائى(0.0458) اكبر من الخطأ المعياري فى النموذج المثبت(0.0118) وهذا ادى الى انخفاض قيمة Z فى النموذج العشوائى عن نظيراتها فى النموذج المثبت.

وتقدر فترات الثقة من الخطأ المعياري كالتالى(Hedges & Olkin, 1985):

$$\overline{ES}_{lower} = \overline{ES} - Z_{(1-\alpha)}SE_{\overline{ES}}$$

$$\overline{ES}_{upper} = \overline{ES} + Z_{(1-\alpha)}SE_{\overline{ES}}$$

- \overline{ES}_{lower} حد الثقة الادنى.
- \overline{ES} متوسط حجم التأثير.
- $Z_{(1-\alpha)}$ القيمة الجدولية عند 0.05 او 0.01 لاختبار ذو ذيلين.
- $SE_{\overline{ES}}$ الخطأ المعياري لمتوسط حجم التأثير.

ولكن قبل حساب فترات الثقة تحول قيمة متوسط تحويل فيشر الى متوسط معامل الارتباط. ويقدر فترات الثقة CI 95% كالتالى:

$$\overline{r}_{lower} = \overline{r} - Z_{(1-\alpha)}SE_{\overline{r}} = 0.326 - 1.96 \times .0458 = .2363$$

$$\overline{r}_{upper} = \overline{r} + Z_{(1-\alpha)}SE_{\overline{r}} = 0.326 + 1.96 \times .0458 = .349$$

المقارنة بين نموذج التأثيرات المثبتة ونموذج التأثيرات العشوائية

قدم Borenstien et al.(2009) مقارنة جيدة ومستفيضة بين النموذجين فى النموذج المثبت نفترض ان حجم التأثير الحقيقى هو نفسة لكل الدراسات وان متوسط حجم التأثير المقدر من الدراسات هو تقدير لحجم التأثير الحقيقى فى المجتمع وان السبب الوحيد للاختلاف بين احجام التأثير هو خطأ المعاينة(عينات الدراسات)، فى حين فى النموذج العشوائى نفترض ان حجم التأثير الحقيقى يختلف من دراسة الى اخرى وان الدراسات المتضمنة فى التحليل هى عينة عشوائية لاحجام التأثير للظاهرة موضع الدراسة وان ملخص حجم التأثير هو تقدير لمتوسط احجام التأثير.

فى النموذج المثبت مصدر الخطأ او عدم التاكيد هو نتيجة خطأ المعاينة بينما فى نموذج التاثيرات العشوائية يوجد مصدرين للاخطأ هما خطأ المعاينة بالاضافة الى خطأ ناتج عن الفروق بين الدراسات، وعليه فان الاخطأ المعيارية وفترات الثقة وتباين احجام التاثير فى النموذج العشوائى اكبر او اوسع من نظيراتها للنموذج المثبت.

اما الفرض الصفري فى النموذج المثبت هو ان حجم التاثير فى كل دراسة يساوى صفر بينما فى النموذج العشوائى متوسط حجم التاثير لكل الدراسات يساوى صفر وبالتاكيد هذا ينعكس على عملية تعميم نتائج النموذجين.

ويفضل استخدام نموذج التاثيرات المثبتة اذا تحقق الشرطين:

- الدراسات المتضمنة فى التحليل متماثلة او متجانسة.
- الهدف من التحليل حساب حجم التاثير العام للدراسات المتماثلة دون تعميم النتائج لمجتمعات غير مجتمع عينة الدراسات.

فلو ارادت شركة ادوية التحقق من فاعلية عقار ما واستخدمت مئات المرضى من خلال المجموعة التجريبية والضابطة ولان الشركة تجرى الدراسة على 100 مريض فى نفس الوقت فالشركة تجرى عشرة دراسات فى نفس الوقت على العينة وفى هذه الحالة فان الدراسات العشرة متماثلة لان العينة لها خصائص متشابهة ونفس المعالجة(العقار) ونفس القياسات ونفس الفريق البحثى، وعليه فان هذه النوعية من الدراسات متماثلة او متجانسة(الشرط الاول) وتقوم على نفس المجتمع فان الهدف من دراسة ما وراء التحليل دراسة مدى فاعلية العقار على نفس العينة وتعمم النتائج على نفس المجتمع وليس مجتمعات مرضى اخرى(الشرط الثانى). ولكن يبدو هذا التجانس فى دراسات العلوم الطبية ولكنه بعيد المنال فى دراسات العلوم النفسية والاجتماعية.

وعلى النقيض من المثال السابق حيث يجمع الباحث بيانات او دراسات من سلسلة من الدراسات فى التراث البحثى لظاهرة ما لباحثين مستقلين باجراءات وعينات ومنهجيات وبرامج مختلفة بمعنى ان الدراسات ليست متجانسة او متكافئة ولذلك لا يمكن افتراض وجود حجم تاثير عام.

واذا كان عدد ت الدراسات صغير جداً فى ما وراء التحليل فان تقدير التباين بين الدراسات τ^2 يكون غير دقيق وعلى الرغم صلاحية استخدام نموذج التاثيرات العشوائية ولكن المعلومات

المتطلبه لتنفيذة تكون ناقصة وفي هذه الحالة امام الباحث عدة خيارات , Borenstein et al., (2009):

الاول: عرض احجام التأثير بصورة منفصلة ولا تعرض ملخص التأثير بالتالى لايمكن الوصول الى استنتاجات حول حجم التأثير وفترات الثقة المرتبطة به، وهنا يلجأ الباحث الى عرض الدراسات فى ضوء طريقة العد كطريقة لمراجعة الدراسات السابقة وربما يصل الى استنتاج خاطئ.

الثانى: اجراء ما وراء التحليل باستخدام نموذج التأثيرات المثبتة حيث يعطى تحليل وصفى للدراسات المتضمنة فى التحليل.

ومعظم الباحثون فى العلوم الاجتماعية يتبنوا مدخل Hedges & Olkin (1985) و Rosenthal & Rubin(1982) للتاثيرات المثبتة لسهولة فهمة وبساطة ولكن يؤدى الى تقدير متحيز لحجم التأثير الحقيقى لان البيانات الحقيقية لاحجام التأثير لا تقيس نفس حجم التأثير انما تتنوع باختلاف مجتمع عينات الدراسة وخصائص الدراسات او ما يعرف بالمتغيرات المتفاعلة (Hagger, 2006)، واثبتت دراسات المحاكاة ان باستخدام نموذج التأثيرات المثبتة لاحجام التأثير تولدت من مجتمعات مختلفة يؤدى الى تضخم الخطأ من النوع الاول قبول وجود حجم تاثير مفترض بينما فى الواقع لا يوجد تاثير حقيقى وعلى ذلك يجب تبنى نموذج التأثيرات العشوائية لـ Hunter & Schmidt (1990, 2004) او Hedges & Vevea (1998) (Field, 2003).

راجع Hunter & Schmidt (2000) طرق ما وراء التحليل المستخدمة فى الدراسات فى مجلة Psychological Bulletin ووجدوا ان الـ 21 دراسة استخدمت كلها نموذج التأثيرات المثبتة متبوعة بالتحليلات المتفاعلة ولم تستخدم اى من الدراسات نموج التأثيرات العشوائية وحذروا من ان استخدام نموذج التأثيرات المثبت لانه يضخم نسبة الخطأ الاول بنسبة اعلى من 11% لدراسات ما وراء التحليل التى تتضمن حجم عينة 25 دراسة، واعلى من 28% للدراسات التى تتضمن حجم عينة اكثر من 100 دراسة وفى هذا الشأن يتفق مع Field (2003).

الفصل العاشر

نموذج التأثيرات المثبتة: تجزئة تباين حجم التأثير او التحليلات المتفاعلة

Fixed Effects Model: Partitioning Effect Size Variance or Moderator analyses

مقدمة:

عندما لا يوجد تجانس بين احجام التأثير فمن الضروري دراسة مصادر هذا الاختلاف من خلال التحليلات المتفاعلة Moderator analyses حيث بدلاً من تفسير التباين بين احجام التأثير نتجة مصادر عشوائية غير ملاحظة فيمكن عزو هذا الاختلاف الى مصادر غير عشوائية منتظمة Systematic sources التي تتضمن المتغيرات او خصائص الدراسة التي كودت في بروتكول التكويد وتسمى في هذه الحالة متغيرات مستقلة مثل حجم العينة، طبيعتها، نوع التصميم البحثي، سنة النشر، وغيرها من خصائص الدراسات. واحياناً دراسة اثر خصائص الدراسة على حجم التأثير تكون اكثر اهمية للباحث من حساب متوسط حجم التأثير وهذا يعتمد على اسئلة البحث.

وهذه التحليلات تحاول تفسير الاختلاف بين احجام التأثير عبر الدراسات المختلفة باستخدام خصائص الدراسة كمنبئات Predictors.

ويعرف (Baron & Kenny (1986) التفاعل بانه اختلاف العلاقة بين متغيرين على مستويات مختلفة من المتغير المتفاعل بمعنى دراسة ما اذا كانت العلاقة بين Y, X (حجم التأثير) تختلف او تعتمد على مستويات المتغير المتفاعل (احد خصائص الدراسة).

يقدم ما وراء التحليل طريقة لتناول امكانية وجود المتغير المتفاعل، والمتغير المتفاعل يستخدم لتجزئة الدراسات الى مجموعات فرعية، واجراء ما وراء التحليل لكل مجموعة منفصلة واذا وجدت فروق كبيرة بين متوسطات المجموعات فمن المحتمل نقصان التباين داخل الدراسات الفرعية، ويظهر ما وراء التحليل كم متبقى من التباين الى عوامل اخرى، ولو اظهر ما وراء التحليل فروق

صغيرة عبر مجموعة الدراسات المتضمنة فى التحليل فان تأثير المتغيرات المتفاعلة يكون قليلاً جداً.

والمتغيرات المتفاعلة او خصائص الدراسات يمكن ان تكون تصنيفية (نوع المقياس المستخدم) او متصله (متوسط عمر العينة)، ويمكن فحص تأثير المتغيرات التصنيفية او المتغيرات المتصلة على حجم التأثير او كلاهما معاً فى تحليل واحد.

تأثير المتغيرات المتفاعلة التصنيفية Categorical Moderators

تقوم تأثير المتغيرات التصنيفية المتفاعلة على حجم التأثير مشابهة لتحليل التباين الاحادى فى الدراسات الاولية (Hedges, 1982)، حيث يجرى ANOVA التباين فى درجات الافراد او احجام التأثير (او اى وحدة تحليل) الى تباين بين Between وتباين داخل Within المجموعات، والتالى فان تحليل التفاعل فى ما وراء التحليل يجرى الاختلاف بين الدراسات الى بين وداخل مجموعات الدراسات او احجام التأثير فى ضوء المتغير المستقل (احد خصائص الدراسات التصنيفية) بمعنى مقارنة بين مجموعات مختلفة من احجام التأثير فى ضوء مستويات المتغير التصنيفى مثلاً طبيعة التصميم التجريبي (تجريبى - شبه تجريبى)، طبيعة التصميم البحثى (معاينة عشوائية - غير عشوائية)، طبيعة التدخل او المعالجة (فرى - جماعى او معرفى - سلوكى)، جنس العينة (ذكر - انثى)، نوع المرحلة العمرية (ابتدائى - اعدادى - ثانوى - جامعى)، وغيرها من خصائص الدراسات المكودة.

فاذا وجدت دلالة احصائية بين مستويات المتغير التصنيفى فان اختلاف متوسط حجم التأثير عبر المجموعات لا يرجع فقط الى خطأ المعاينة بل الى عوامل اخرى منها المتغير التصنيفى موضع التحليل فى ANOVA .

واحصاء التجانس او الاختلاف Q الكلى الاتى:

$$Q_{total} = \sum w_i (ES_i - \bar{ES})^2 = \sum (w_i ES_i^2 - \frac{(\sum w_i ES_i)^2}{\sum w_i})$$

ويجزئ ANOVA هذا الاحصاء الى جزئين احدهما اختلاف بين مجموعات احجام التأثير $Q_{between}$ اي الاختلاف نتجة المتغير التصنيفي والآخر داخل مجموعات احجام التأثير بمعنى البواقي Residuals لمتغيرات او لخصائص اخرى غير متضمنة في التحليل Q_{within} .

و $Q_{between}$ مجموع المربعات الموزونة لمتوسط احجام التأثيرات حول المتوسط العام بينما Q_{within} مجموع المربعات الموزونة لاجسام التأثير المفردة داخل كل مجموعة او تصنيف حول متوسط المجموعة.

وعليه فان التباين الكلي كالآتي (Card, 2012):

$$Q_{total} = Q_{between} + Q_{within}$$

- Q_{total} التباين الكلي بين كل احجام التأثير.
- $Q_{between}$ التباين بين المجموعات لاجسام التأثير.
- Q_{within} التباين داخل كل مجموعة لاجسام التأثير.

وإذا اتضح اختلاف بين المجموعات بمعنى رفض الفرض الصفري فان الاختلاف بين احجام التأثير اكبر مما توقعناه نتجة خطأ المعاينة وعليه يوجد علاقة بين المتغير المتفاعل التصنيفي واحجام التأثير للدراسات والعكس صحيح.

والطريقة المباشرة لتقدير التباين بين المجموعات $Q_{between}$ هو تقديرها من خلال التباين الكلي Q_{total} والتباين داخل المجموعات Q_{within} كالآتي:

$$Q_{between} = Q_{total} - Q_{within}$$

ويقدر التباين داخل المجموعة g كالتالي (Card, 2012; Lipsey & Wilson, 2001):

$$Q_g = \sum w_i (ES_i - \overline{ES}_g)^2 = \sum (w_i ES_i^2 - \frac{(\sum w_i ES_i)^2}{\sum w_i})$$

$$df_g = K_g - 1$$

(تم تغيير Q_{within} الى Q_g نتجة لان المعادلة تتضمن w_i وزن كل دراسة)

- w_i وزن كل دراسة.

- ES_i حجم التأثير في كل دراسة.
- \overline{ES}_g متوسط حجم التأثير عبر الدراسات في المجموعة g .
- K_g عدد الدراسات في المجموعة g .

فاذا تضمن التحليل ثلاثة مجموعات وكل مجموعة تتضمن عدد معين من احجام التأثير فان هذه المعادلة تقدر ثلاث مرات لكل مجموعة على حدة .

وتقدر التباين داخل المجموعات من حاصل جمع التباين داخل كل مجموعة g كالاتى:

$$Q_{within} = \sum_{g=1}^G Q_g = \sum_{1} Q_1 + Q_2 + Q_3 + \dots + Q_g$$

$$df_{within} = \sum_{g=1}^G df_g = k - G$$

- G عدد المجموعات او تصنيفات المتغير المتفاعل.
- Q_g التباين داخل كل مجموعة.
- df_{within} درجات الحرية داخل المجموعات.
- df_g درجات الحرية داخل المجموعة g .
- k عدد الدراسات في كل المجموعات.
- G عدد المجموعات.

وبعد حساب التباين الكلى Q_{total} والتباين داخل المجموعات Q_{within} يمكن حساب التباين بين مجموعات الدراسات $Q_{between}$. والدلالة الاحصائية للتباين بين المجموعات تقيم من خلال قيمة احصاء $Q_{between}$ بالنسبة الى درجات الحرية بين المجموعات:

$$df_{between} = G - 1$$

والتعامل مع احصاء $Q_{between}$ فى ضوء توزيع F^2 بالبحث فى جدول F^2 عن القيمة الجدولية ومقارنتها بالقيمة المحسوبة لـ $Q_{between}$ فاذا كانت: $Q_{between}$ المحسوبة \leq قيمة

² الجولية يرفض الفرض الصفري وعليه ان يوجد اختلاف بين المجموعات فى حجم التأثير عبر المجموعات .

مثال: فى جدول (1.9) ل 22 دراسة للعلاقة بين العدوان ورفض الاقران كما اوضحنا سالفاً ان:

$$Q_{21} = 291.17, P < .01$$

ويفسر هذا الاختلاف اما فى ضوء مصادر عشوائية باستخدام نموذج التأثيرات العشوائية او فى ضوء مصادر منتظمة باستخدام تحليل التفاعل اعتماداً على خصائص الدراسات.

وفى المثال السابق ارد الباحث اجراء تحليل التفاعل فى ضوء المصادر المختلفة لقياس عدوان العلاقات اما فى ضوء الملاحظات (المجموعة الاولى)، تقارير الاباء (المجموعة الثانية)، تقارير الاقران (المجموعة الثالثة)، تقارير المعلمين (المجموعة الرابعة) اى ان المتغير التصنيفى او المتفاعل هو طبيعة القياس والهدف من اجراء دراسة ما وراء التحليل هو تحديد ما اذا كانت العلاقة بين العدوان ورفض الاقران تختلف عبر طرق القياس المختلفة للعدوان.

وعلى تم الحصول على 27 حجم تأثير من الدراسات 22 وتم تقسيم الـ 27 حجم تأثير الى اربع مجموعات كالاتى (Card, 2012):

الجدول (1.10): بيانات دراسة ما وراء التحليل لـ 22 دراسة في ضوء مجموعات مختلفة لطريق

القياس .

Method/study	Sample size (N)	Effect size (r)	Transformed ES (Z _r)	Standard error (SE)	Weight (w)	wES (w * Z _r)	wES ² (w * Z _r ²)
<u>Observation (k = 3)</u>							
Ostrov (under review) ^a	139	.181	.183	.0897	124.18	22.72	4.16
Ostrov & Crick (2007)	132	.049	.049	.0922	117.70	5.81	0.29
Ostrov et al. (2004) ^b	60	.000	.000	.1386	52.05	0.00	0.00
			<u>Within-group sums (Σ)</u>		293.94	28.53	4.45
<u>Parent report (k = 1)</u>							
Blachman (2003)	228	.525	0.583	.0693	208.12	121.27	70.66
			<u>Within-group sums (Σ)</u>		208.12	121.27	70.66
<u>Peer report (k = 17)</u>							
Crick & Gropeter (1995)	491	.198	.201	.0579	297.81	59.74	11.98
Crick et al. (1997)	65	.443	.476	.1347	55.15	26.24	12.48
Geiger (2003)	458	.554	.624	.0484	427.07	266.55	166.37
Hawley et al. (2007)	929	.161	.162	.0346	835.96	135.72	22.03
Henington (1996)	904	.336	.349	.0347	831.95	290.69	101.57
Johnson (2003)	74	.531	.591	.1222	66.92	39.55	23.37
Leff (1995)	151	.617	.721	.0855	136.66	98.48	70.97
Miller (2001)	150	.557	.628	.0845	139.90	87.90	55.23
Murray-Close & Crick (2006)	590	.575	.655	.0426	550.97	361.12	236.69
Nelson et al. (2005)	180	.090	.090	.0831	144.75	13.04	1.17
Pakaslahti & Keltikangas-Järvinen (1998)	839	.326	.339	.0381	689.52	233.68	79.19
Phillipsen et al. (1999)	262	-.048	-.048	.0642	242.93	-11.73	0.57
Rys & Bear (1997)	266	.556	.627	.0642	242.84	152.16	95.34
Salmivalli et al. (2000)	209	.253	.258	.0715	195.52	50.54	13.06
Tomada & Schneider (1997)	314	.313	.324	.0580	296.76	96.01	31.06
Werner (2000)	881	.477	.519	.0351	810.71	420.41	218.01
Werner & Crick (2004)	517	.469	.509	.0455	482.45	245.37	124.79
			<u>Within-group sums (Σ)</u>		6447.89	2565.46	1263.90
<u>Teacher (k = 6)</u>							
Crick et al. (1997)	65	.167	.168	.1305	58.68	9.88	1.66
Johnson (2003)	74	.074	.075	.1223	66.83	4.98	0.37
Nelson et al. (2005)	180	-.011	-.011	.0900	123.53	-1.39	0.02
Ostrov (under review) ^a	139	.513	.567	.0887	127.22	72.13	40.90
Rys & Bear (1997)	266	.338	.352	.0631	250.96	88.40	31.14
Tomada & Schneider (1997) ^b	314	.000	.000	.0597	280.47	0.00	0.00
			<u>Within-group sums (Σ)</u>		907.70	174.00	74.09
			<u>Total sums (Σ)</u>		7857.64	2889.26	1413.09

لحساب Q_{total} تم استخدام ثلاث مجاميع عبر الـ 27 دراسة (اسفل الجدول):

$$Q_{total} = \sum (w_i ES_i^2) - \frac{(\sum w_i ES_i)^2}{\sum w_i} = 1413.09 - \frac{(2889.26)^2}{7857.64} = 350.71$$

اما حساب التباين لاحجام التأثير داخل المجموعة الاولى التى تتضمن ثلاث دراسات اعتمدت على قياس العدوان فى ضوء الملاحظة كالاتى:

$$Q_{within-observation} = \sum (w_i ES_i^2) - \frac{(\sum w_i ES_i)^2}{\sum w_i}$$

$$= 4.45 - \frac{(28.53)^2}{293.94} = 1.68$$

وباستخدام نفس المعادلة قدر التباين لاحجام التأثير داخل المجموعة الثانية التى تتضمن حجم تأثير واحد وهو قياس العدوان من وجهة نظر الاباء:

$$Q_{within-observation} = 0.00$$

وتباين داخل مجموعة تقدير العدوان من وجهة نظر المعلمين:

$$Q_{within-teacher} = 40.73$$

وتباين داخل مجموعة تقدير العدوان من وجهة نظر الاقران:

$$Q_{within-peer} = 243.16$$

إذا مجموع التباينات الاربعة داخل المجموعات كالاتى:

$$Q_{within} = \sum_{g=1}^G Q_g = 1.68 + 0.00 + 234.16 + 40.73 = 285.57$$

التباين بين المجموعات الاربعة:

$$Q_{between} = Q_{total} - Q_{within} = 350.71 - 285.57 = 65.14$$

وبما ان توزيع $Q_{between}$ له توزيع χ^2 بدرجات حرية:

$$df_{between} = G - 1 = 4 - 1 = 3$$

وبالكشف فى جدول η^2 بدرجات حرية 3 عند مستوى دلالة احصائية 0.01 يتضح ان قيمة η^2 الجدولية=11.345 وعلية فان:

$$(65.14)Q_{between} > (11.345)\eta^2$$

بالتالى نرفض الفرض الصفري بمعنى توجد فروق دالة احصائية فى متوسط حجم التأثير بين انواع القياسات الاربعة بمعنى يوجد تأثير للمتغير المتفاعل.

ويادخال بيانا الجدول السابق فى برنامج SPSS وباجراء تحليل التباين الاحادى لدراسة الفروق بين المجموعات الاربع والنتائج كالاتى:

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
r	Between Groups	.397	3	.132	4.034	.019
	Within Groups	.755	23	.033		
	Total	1.152	26			
Zr	Between Groups	.497	3	.166	3.808	.024
	Within Groups	1.000	23	.043		
	Total	1.497	26			

يتضح وجود فروق دالة احصائياً بين المجموعات الاربعة فى حجم التأثير سواء r او تحويل فيشر حيث $r \downarrow P=0.019$ او $Zr \downarrow P=0.024$ عند مسوى دلالة احصائية 0.05 وعلية فان الاختلاف بين احجام التأثير يرجع الى نوعية قياس العدوان.

التحليل البعدى للمتفاعل التصنيفى Follow up analysis to a Categorical Moderator

اذا تكون المتغير المتفاعل التصنيفى من مستويين فان عملية التفسير فى تحديد اى مستوى مسبب للدلالة عن طريق حساب متوسط حجم التأثير داخل كل مجموعة وتكون الدلالة لصالح المتوسط الاكبر.

ولكن اذا كان المتغير التصنيفى باكثر من مستويين فان العملية تكون اكثر صعوبة لان الدلالة الاحصائية باستخدام اختبار Omnibus F بين المجموعات تشير الى ان على الاقل بعض المجموعات تختلف عن بعضها البعض ولكن غير واضح اى هذه المجموعات التى احدثت الدلالة الاحصائية. وفى هذه الحالة لابد من اجراء اختبار بعدى او تتبعى مشابهة للتحليل البعدى Post hoc فى ANOVA وهذا التحليل البعدى فى اطار ما وراء التحليل يتم فى ضوء مدخلين كالتالى (Card, 2012):

المدخل الليبرالى Liberal Approach

هذا المدخل لا يحاول التحكم فى تضخم الخطأ من النوع الاول حيث يسعى الى اجراء العديد من المقارنات بين كل زوج من المجموعات او مستويات المتغير المتفاعل التصنيفى مثل اجراء اختبار T بين المجموعة الاولى والثانية، وبين الاولى والثالثة وهكذا.

ولاجراء هذه المقارنات اتبع الاتى:

1. حساب التباين او الاختلاف الكلى بين المجموعات Q_{total} للدراسات المتضمنة فى المجموعتين.
2. حساب مجموع التباين داخل المجموعات Q_{within} فقط لدراسات المجموعتين.
3. حساب التباين بين المجموعتين $Q_{between}$ بـ $df=1$ لان عدد المجموعات 2 ثم بالبحث فى جدول F^2 عن القيمة الجدولية وبمقارنتها بقيمة $Q_{between}$ يصل الباحث الى القرار بوجود اختلاف بين احجام التأثير فى المجموعتين من عدمه.

ويكرر هذا الاجراء عند المقارنة بين اى مجموعتين من مجموعات او مستويات المتغير التصنيفى المتفاعل وعليه فان اجمالى عدد المقارنات يكون فى ضوء العلاقة:

$$\text{عدد المقارنات} = \frac{G(G - 1)}{2}$$

فاذا كان عدد مستويات المتغير التصنيفى المتفاعل 4 فان عدد المقارنات:

$$= \frac{4(4-1)}{2} = \frac{12}{2} = 6$$

وهذه الطريقة تماماً مثل اختبار فروق الدلالة الدنيا ليفشر Fisher Least Significance Difference (LSD) المستخدم فى المقارنات البعدية فى ANOVA وهذا الاختبار يعانى من محددات اهمها تضخم الخطأ من النوع الاول. واذا وجدت مقارنات عديدة فان قيمة التضخم للخطأ من النوع الاول هى:

$$\alpha_E = 1 - (1 - \alpha)^c$$

- α مستوى الدلالة الاحصائية.
- c عدد المجموعات

المدخل المحافظ Conservative Approach

اجراء المقارنات البعدية عن طريق سلسلة من كل المقارنات المحتملة بين اى زوج من المجموعات مثل الطريقة الليبرالية ولكن يتم استخدام تصحيح بيرونيفرونى Bonferroni ، فعند عمل ثلاثة مقارنات وكل مقارنة باحتمال $\alpha=0.5$ (الخطأ من النوع الاول) فان احتمال ان لا يزيد الخطأ من النوع الاول عن $3 \times 0.05 = 0.15$

وعليه اذا كانت α للمقارنة = 0.05 فعند اجراء ثلاثة مقارنات فان α لكل مقارنة هى:

$$\alpha = \frac{0.05}{3} = 0.0167$$

والقضية هنا ان جداول T لا تعطى قيم جدولية عند مستوى الفا 0.017 ويمكن تنفيذ ذلك من خلال برنامج الكمبيوتر حيث يعطى قيمة P.

ولكن لهذا المدخل محددات اهمها نقص القوة الاحصائية فاذا كان عدد المجموعات 3 فبدلاً من استخدام الفا 0.05 تكون قيمتها 0.0167 بالتالى تزداد صعوبة رفض الفرض الصفرى بالتالى تضخم الخطأ من النوع الثانى. وهذا المحدد يعتمد على حجم عينة احجام التأثير فتزداد هذة الاشكالية اذا كان حجم العينة صغير وتقل حدتها اذا كان حجم عينة احجام التأثير كبيرة.

في المثال السابق يمكن اجراء التحليل البعدى Post hoc باستخدام اختبار توكى Tukey او
 تصحيح Bonferroni وكانت النتائج كالتالى:

Warnings

Post hoc tests are not performed for r because at least one group has fewer than two cases.

Post hoc tests are not performed for Zr because at least one group has fewer than two cases.

لم يعطى المقارنات المتعددة لان احد المجموعات يتضمن حالة واحدة وهذا لقياسات العدوان عن طريق الالباء وباستبعاد هذه المجموعة واجراء التحليل مرة اخرى والنتائج كالتالى:

Post Hoc Tests

Multiple Comparisons								
Dependent Variable		(I) measure	(J) measure	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
							Lower Bound	Upper Bound
r	Tukey HSD	1.00	3.00	-.30698039*	.11346660	.033	-.5911389-	-.0228219-
			4.00	-.10716667-	.12812183	.685	-.4280268-	.2136935
		3.00	1.00	.30698039*	.11346660	.033	.0228219	.5911389
			4.00	.19981373	.08604031	.073	-.0156602-	.4152876
		4.00	1.00	.10716667	.12812183	.685	-.2136935-	.4280268
			3.00	-.19981373-	.08604031	.073	-.4152876-	.0156602
	Bonferroni	1.00	3.00	-.30698039*	.11346660	.038	-.5999531-	-.0140077-
			4.00	-.10716667-	.12812183	1.000	-.4379794-	.2236461
		3.00	1.00	.30698039*	.11346660	.038	.0140077	.5999531
			4.00	.19981373	.08604031	.088	-.0223438-	.4219713
		4.00	1.00	.10716667	.12812183	1.000	-.2236461-	.4379794
			3.00	-.19981373-	.08604031	.088	-.4219713-	.0223438
Zr	Tukey HSD	1.00	3.00	-.34154902*	.13060904	.040	-.6686380-	-.0144601-
			4.00	-.11816667-	.14747836	.706	-.4875021-	.2511688
		3.00	1.00	.34154902*	.13060904	.040	.0144601	.6686380
			4.00	.22338235	.09903920	.083	-.0246451-	.4714098
		4.00	1.00	.11816667	.14747836	.706	-.2511688-	.4875021
			3.00	-.22338235-	.09903920	.083	-.4714098-	.0246451
	Bonferroni	1.00	3.00	-.34154902*	.13060904	.046	-.6787838-	-.0043142-
			4.00	-.11816667-	.14747836	1.000	-.4989583-	.2626250
		3.00	1.00	.34154902*	.13060904	.046	.0043142	.6787838
			4.00	.22338235	.09903920	.102	-.0323386-	.4791033
		4.00	1.00	.11816667	.14747836	1.000	-.2626250-	.4989583
			3.00	-.22338235-	.09903920	.102	-.4791033-	.0323386

* The mean difference is significant at the 0.05 level.

يتضح ان مسبب الدلالة هو بين المجموعة الاولى (قياسات الملاحظة) مع المجموعة الثالثة (قياسات الاقران) حيث $P=0.033$ و $P=0.038$ و Benferonni لذلك لحجم التأثير r . وهذه الدلالة لصالح المجموعة الثالثة حيث المتوسطات كالتالى:

Homogeneous Subsets

		r	
measure	N	Subset for alpha = 0.05	
		1	2
Tukey HSD ^{a,b}	1.00	3	.0766667
	4.00	6	.1838333
	3.00	17	.3836471
Sig.			.603

Means for groups in homogeneous subsets are displayed.

المتغيرات المتفاعلة المتصلة **Continous Moderators** او ما وراء الانحدار **Meta-regression**

بالمثل كما استخدمنا ANOVA كطريقة لاختبار تاثير المتغير التصنيفى الوحيد مثل نوع المعالجة لتفسير الاختلاف فى توزيع احجام التأثير، فكثيراً من خصائص الدراسات المكودة متغيرات متصلة وبدورها تكون مرتبطة بحدوث اختلاف فى احجام التأثير المقاسة فى للدراسات مثل **عمر العينة**، **دخل الاسرة**، **عدد افراد الاسرة**، **عدد جلسات البرنامج**، **مدة البرنامج**، وغيرها. وهذه بمثابة متغيرات متفاعلة متصلة ويتم تقييم تاثير المتغيرات المتصلة على حجم التأثير من خلال تحليل الانحدار الموزون **Weighted Regression analysis** حيث يعتبر المتغيرات المتصلة المتفاعلة بمثابة متغيرات مستقلة او منبئات **Predictors** واحجام التأثير متغير تابع، وقيمة معامل الانحدار الدالة احصائياً تدل على ان المتغير المستقل او المتغير المتفاعل يرتبط او يؤثر على العلاقة بين المتغيرين Y, X .

وفى هذه الحالة يستخدم الانحدار الموزون **Weighted Least Square Regression** حيث يتم وزن الانحدار من خلال معكوس التباين w (وزن كل حجم تاثير بمعكوس التباين)، وهذا الموازنة لها ثلاث تطبيقات:

الاول: الحصول على احجام تاثير اكثر دقة.

الثاني: متوسط المربعات في مخرج الانحدار (تحليل ل ANOVA في برامج مثل SPSS , SAS) تمثل الاختلاف بين احجام التأثير التي ترجع الى الاتحاد الخطى للمتغير المتفاعل المتصل. وتستخدم هذه القيمة لتقويم الدلالة الاحصائية لنموذج الانحدار.

الثالث: هذا الوزن او التصحيح لاحجام التأثير يؤثر في الاخطأ المعيارية لمعاملات الانحدار.

ويوجد مؤشرين للمطابقة العامة لنموذج الانحدار الموزون احدهما احصاء Q نتيجة للانحدار وتسمى $Q_{regression}$ ، والاخر Q نتيجة للخطأ او البواقي $Q_{residual}$ يعكسان التباين الكلي في نموذج الانحدار، وجدول تحليل التباين في مخرج الانحدار يتكون من مجموع المربعات التي تتضمن مجموع مربعات الانحدار ومجموع مربعات الخطأ وكلاهما يتم تقييم دلالتها من خلال احصاء كاي تربيع، ومجموع مربعات الانحدار بدرجات حرية تساوى عدد المتغيرات المنبئة في النموذج P بينما مجموع مربعات البواقي بدرجات حرية (K-P-1) حيث K عدد احجام التأثير في التحليل.

ولو كانت $Q_{residual}$ دالة احصائياً فان الاختلاف بين احجام التأثير لا يرجع الى خطأ المعاينة فقط بل الى عوامل اخرى ابعد من ذلك، بكلمات اخرى حتى لو استبعدنا الاختلاف بين احجام التأثير نتيجة المتغير المنبئ المتضمن في نموذج الانحدار فانه يوجد ايضاً عدم تجانس بين احجام التأثير نتيجة متغيرات وعوامل اخرى لم تتضمن في النموذج.

مثال (Card, 2012): بالنظر لمثال السابق عرضه في الفصل (الثامن) لماوراء التحليل لـ 22 دراسة للعلاقة بين العدوان ورفض الاقران، فاعتبر الباحث ان العمر Age متغير متفاعل متصل واراد معرفة ماذا كان للعمر مرتبط بالاختلاف بين احجام التأثير فسجل متوسط العمر بالسنوات لعينة الدراسة في بروتكول الدراسة بالتالي اصبحت معادلة الانحدار كالاتي:

$$Z_r = B_0 + B_1(Age) + e$$

- Z_r حجم التأثير المتنبأ به (لاحظ ان هذا ليس الدرجة المعيارية Z انما تحويل فيشر المناظر لمعامل الارتباط r).
- B_0 ثابت الانحدار (المناظر لـ a).
- B_1 معامل الانحدار اللامعيارى.

- خطأ القياس وهو انحراف حجم التأثير المقاس عن متوسط حجم التأثير الحقيقي في المجتمع.

وإذا ادخلت هذه البيانات برنامج SPSS في ضوء ثلاثة متغيرات هي Z_r متغير تابع والعمر متغير مستقل ومعامل التصحيح أو الوزن w وبإجراء الانحدار باستخدام طريقة Weighted Least Squares (WLS). والنتائج كالتالي:

وننتائج جدول تحليل التباين ANOVA:

- مجموع المربعات لنموذج الانحدار SS_{model} أو $SS_{regression} = 9.312$
- مجموع مربعات البواقي $SS_{error} = 281.983$ أو $SS_{residual}$
- متوسط مجموع المربعات للبواقي $MS_{residual} = 14.099$

ومن جدول معاملات الانحدار يتضح ان:

- معامل الانحدار غير المعياري $B_1 = -0.0112$ بخطأ معياري $SE = 0.0138$.
- قيمة الثابت $B_0 = 0.496$

وما يهمنا دلالة $SS_{regression}$ حيث الدلالة تعنى وجود عدم تجانس أو اختلاف بين احجام التأثير جراء المتغير المتفاعل وفي هذه الحالة العمر.

وقيمة لـ $SS_{regression}$ تكتب في تقارير دراسات ما وراء التحليل بـ $Q_{regression}$ وتقيم دلالتها في ضوء احصاء كاي تربيع بدرجات حرية تساوى عدد المنبئات $df = 1$ وبالبحث في جدول كاي تربيع بدرجات حرية 1 فان قيمتها $= 9.312$ وهي دالة احصائياً عند مستوى دلالة احصائية 0.05، وعليه يرفض الفرض الصفري بالتالي اختلاف العلاقة بين العدوان والرفض الاقران (حجم التأثير) ترجع الى المتفاعل العمر.

والخطأ المعياري المصاحب لمعامل الانحدار غير دقيق ولا بد من تصحيحه كالآتي:

$$SE_{adjusted} = \frac{SE_{output}}{\sqrt{MS_{residual}}} = \frac{0.0138}{\sqrt{14.099}} = 0.00368$$

- SE_{output} الخطأ المعياري في المخرج البرنامج.
- $MS_{residual}$ متوسط مجموع المربعات للبقاوى $(\frac{SS_{residual}}{df_{residual}})$

وتتم تقييم الدلالة الاحصائية لمعامل الانحدار كالاتى:

$$Z = \frac{B_1}{SE_{adjusted}} = \frac{-0.0112}{0.00368} = -3.05$$

وهى دالة احصائياً عند 0.01 حيث تقيم فى ضوء اختبار Z وقيمة 2.58 لاختبار ذو ذيلين وقيمة 1.96 لاختبار ذو ذيلين عند مستوى 0.05. ولاحظ ان الدلالة الاحصائية لنموذج الانحدار $SS_{regression}$ ولمعامل الانحدار متماثلين حيث:

$$Z^2 = \chi^2_{(df=1)} = (-3.05)^2 = 9.31$$

الاطار العام للانحدار المتعدد للمتغيرات المتفاعلة

فى المثال السابق اجرى تحليل الانحدار البسيط لمتغير وحيد منبئ ومتغير وحيد تابع فالدلالة الاحصائية تخبرنا عن تاثير متغير واحد ولكن ماذا لو تضمن التحليل متغيرات منبئة عديدة فمن الضرورى فحص معاملات الانحدار للمتغيرات المنبئة وتقيم دلالتها الاحصائية والاختأ المعيارية المصاحبة لها.

حقيقة تقيم تاثير المتغيرات المتصلة المتفاعلة (خصائص الدراسات) فى ما وراء التحليل من خلال نموذج تحليل الانحدار يطلق عليه ما وراء تحليل الانحدار Meta-regression وسوف نتناوله من عدة محاور:

النموذج الفارغ لحساب متوسط حجم التاثير والاختلاف Empty Model

هو ذلك النموذج الذى يتم فيه انحدار المتغير التابع بدون اى متغير منبئ ولكن مع وجود الثابت Constant (حيث القيمة 1 لكل الحالات) ويمكن عرضه كالتالى (Card, 2012):

$$ES_i = B_0(1) + e_i$$

- ES_i حجم التاثير للدراسة مثل r او تحويل فيشر.
- B_0 ثابت نموذج الانحدار.

• e_i الخطأ الناتج عن انحراف حجم التأثير للدراسة عن متوسط حجم التأثير العام.

واجراء الانحدار الموزون لاحجام التأثير المنبئة عن طريق الثابت يعطى معلومات عن متوسط حجم التأثير الموزون او المصحح واختبار الاختلاف.

مثال: لدراسة ما وراء التحليل للعلاقة بين العدوان ورفض الاقران لـ 22 دراسة تم اجراء الخطوات الاتية:

الاول: ادخال احجام التأثير للدراسات فى ضوء تحويل فيشر Z_r ووزن الدراسة w فى البرنامج الاحصائى SPSS او اى برنامج اخر.

الثانى: انشاء متغير اخر هو الثابت واعطاؤه القيمة 1 لكل الدراسات.

الثالث: انحدار المتغير التابع حجم التأثير بعد موازنتها بـ w على الثابت وعليه يستخدم الثابت كمتغير منبئ.

الناتج Output:

• معامل الانحدار غير المعيارى $B=0.387$ وهذا يمثل متوسط حجم التأثير عبر الدراسات المختلفة.

• الخطأ المعيارى المرتبط بمعامل الانحدار المعيارى ويتم تصحيحه كالتالى:

$$SE_{adjusted} = \frac{SE_{output}}{\sqrt{MS_{residual}}} = \frac{0.0440}{\sqrt{113.865}} = 0.0118$$

يستخدم الخطأ المعيارى المعدل فى تقييم الدلالة الاحصائية وحساب فترات الثقة.

• مجموع مربعات البواقي $SS_{residual} = 291.17$ وهذا يمثل قيمة احصاء Q للاختلاف

بين احجام التأثير ويقوم فى ضوء كاي تربيع بدرجات حرية $(K-1)21$ وهذه النتائج مشابهة لخطوات حساب متوسط حجم التأثير واجراء اختبارات الفروض حول المتوسط وقيمة اختبار عدم التجانس Q لاحجام التأثير عبر الدراسات المختلفة للنموذج التاثيرات المثبتة فى الفصل(الثامن).

واجرى (2008) Cheung تحليل المتغيرات المتصلة المتفاعلة في ماوراء التحليل في اطار
نمذجة المعادلة البنائية من خلال كتابة: Meta-Analysis: A Structural Equation
Modeling Approach ولكن عرض هذا الدمج يقع خارج نطاق هذا الكتاب.

نماذج التأثيرات المختلطة Mixed-Effects models

بعد نمذجة معظم العوامل اوخصائص الدراسات المؤثرة على تباين احجام التأثير باستخدام تحليل
التباين او تحليل الانحدار الموزون يبقى توزيع احجام التأثير غير متجانس او مختلف، وهذا
يشير الى ان مسلمة نموذج التأثيرات المثبتة التي تجزئ التباين بين احجام التأثير الى مصدرين
احدهما تباين منظم(خصائص الدراسات) وخطأ المعاينة الافراد في الدراسات لايمكن الدفاع عنه
او مازالت غير قائمة(Lipsey & Wilson, 2001)، ويفترض نموذج التأثيرات المختلطة ان
التأثيرات بين متغيرات الدراسات مثل نوع المعالجة تكون منتظمة ولكن يوجد تأثيرات عشوائية
غير مقاسة لتوزيع حجم التأثير بالاضافة الى خطأ المعاينة، بمعنى ان الاختلاف في توزيع حجم
التأثير يرجع الى:

- مصادر منتظمة(خصائص الدراسات) تسبب فروق بين احجام التأثير ويتم نمذجتها
باستخدام تحليل التباين او تحليل الانحدار.
- خطأ المعاينة على مستوى عينات الدراسات.
- مكون عشوائي اضافي اخر.

ومطابقة نموذج التأثيرات المختلطة لبيانات حجم التأثير مشابهة الى طريقة مطابقة نموذج
التأثيرات العشوائية. وفي النموذج المختلط تباين التأثيرات العشوائية قائم على تباين البواقي
 $Q_{residual}$ مقارنة بالتباين الكلي. ويقدر تباين المكون العشوائي في النموذج المختلط من خلال
جبر المصفوفات(للمزيد انظر: Card, 2012).

السؤال المطروح باستمرار ايهما اكثر مناسبة نموذج التأثيرات المثبتة ام نموذج التأثيرات المختلط
لتوزيع احجام تأثير مختلفة؟، النموذج المثبت اكثر قوة احصائية لاجراء العلاقات المتفاعلة مع
احجام التأثير واكثر احتمالية لتحديد الفروق المنتظمة بين الدراسات ولكن هذا له مخاطر هو ان
النموذج المثبت له مخاطر عالية من الخطأ من النوع الاول، وعلى العكس فان نموذج التأثيرات

المختلط له مستويات عالية الدقة من نسب الخطأ من النوع الاول ولكن مستويات اقل من القوة الاحصائية فى تحديد التأثيرات المتفاعلة.

ويطلق على النموذج المختلط بالانماذج العشوائية الشرطية *Conditionally Random models* حيث تجمع بين تحليل المتفاعلات التى عرضت فى هذا الفصل مع تقدير تباين مجتمع احجام التأثيرات (نموذج التأثيرات العشوائية) التى عرضناها فى الفصل (التاسع)

ويرى (Card 2012) ان هذه النماذج مفيدة عندما تقيم المتغيرات المتفاعلة فى ما وراء التحليل فى اطار نموذج التأثيرات المثبتة من خلال الدلالة لـ $Q_{residual}$ فى اطار تحليل الانحدار او Q_{within} فى اطار تحليل التباين، مع الوصول الى تعميم النتائج من خلال نماذج التأثيرات العشوائية.

ونماذج التأثيرات المختلطة تتم فى اطار تحليل الانحدار العام الذى يعتبر تحليل التباين حالة خاصة منه مع اضافة مكون اضافى وهو تباين احجام التأثير فى المجتمع.

والمعادلة العامة لنموذج التأثيرات المختلطة كالتالى (Card, 2012):

$$ES_i = B_0 + B_1X_1 + B_2X_2 + \dots + \zeta_i + \varepsilon_i$$

- ES_i حم التأثير للدراسة i .
- B_0 ثابت الانحدار.
- B_1, B_2 معاملات الانحدار غير المعيارية للمتغيرات المتفاعلة فى التحليل.
- X_1, X_2 المتغيرات المتفاعلة (خصائص الدراسات).
- ζ_i زيتا الخطأ الناتج عن انحراف حجم التأثير المقاس للدراسة عن القيمة المتنبأ بها فى معادلة الانحدار المتضمنة المتغيرات المتفاعلة.
- ε_i انحراف حجم التأثير فى الدراسة عن متوسط حجم التأثير الحقيقى العام نتيجة خطأ المعاينة.

اي النماذج انسب لتحليل احجام التأثير في ما وراء التحليل؟.

عرضنا نموذج التأثيرات المثبتة في مقابل نموذج التأثيرات العشوائية، ونموذج التأثيرات المثبتة مع تحليل المتفاعل في مقابل نموذج التأثيرات المختلطة، فالسؤال المطروح اي من هذه النماذج يفضل استخدامها في ما وراء التحليل؟.

يرى (Card 2012) بانه يوجد خمسة اعتبارات لتفضيل نموذج على نموذج اخر ويمكن عرضهما حسب اهميتهما من الاكثر اهمية الى الاقل اهمية كالتالى:

- نوع الاستنتاجات المراد الوصول اليها من دراسة ما وراء التحليل.
- مدى وجود اختلافات غير مفسرة بين احجام التأثير في دراسة ما وراء التحليل.
- القوة الاحصائية للنموذج المستخدم.
- مدى وجود قيم متطرفة Outliers لاحجام التأثير.
- درجة التعقيد الرياضى لاجراء التحليل.

المعيار الاكثر اهمية فى استخدام اي من النماذج الاربعة هي **طبيعة الاستنتاجات المراد الوصول اليها**. ففي النموذج المثبت تكون الاستنتاج قاصراً فقط على عينة الدراسات المتضمنة فى ما وراء التحليل (هذه الدراسات اوضحت.....)، بينما فى النموذج العشوائى يكون الاستنتاج اكثر تعميماً (اوضح التراث البحثى للظاهرة ان.....)، وعليه فى سبيل نتائج اكثر تعميماً يكون نموذج التأثيرات العشوائية اكثر تفضيلاً .

الاعتبار الثانى وهو الاختلاف او التباين غير المفسر Heterogeneity بين احجام التأثير ويقصد بالتباين غير المفسر:

- فى غياب التحليل المتفاعل بمعنى عدم تضمين خصائص الدراسات التى ربما تكون مسببة للاختلاف بين الدراسات فان الدلالة الاحصائية لاحصاء Q تشير الى ان الاختلاف بين احجام التأثير لا يرجع الى خطأ المعاينة فقط بل الى عوامل ابعد من ذلك.
- باجراء النموذج المثبت مع تضمين التحليل المتفاعل يجب فحص الاختلاف داخل المجموعات (Q_{within} فى ANOVA) او عدم التجانس او اختلاف البواقي ($Q_{residual}$ فى تحليل الانحدار) فلو كان الاختبارين ذو دلالة احصائية فيمكن استنتاج ان الاختلاف او عدم

التجانس بين احجام التأثير لا يمكن تفسيره عن طريق خصائص الدراسات فيفضل استخدام النموذج المختلط. ويوصى (2012) Card بان لايعتمد الباحث على هذا الاعتبار فقط لتحديد النموذ لان احصاء Q احصاء استدلالي حيث دراسات ما وراء التحليل التي تتضمن عدد كبير من الدراسات نحصل على دلالة احصائية لـ $Q_{residual}$ وهى فى طبيعتها اختلافات تافهة ولكن لكبر حجم العينة نحصل على دلالة احصائية بينما فى الدراسات التي لها احجام عينات صغيرة نحصل على عدم دلالة احصائية لـ $Q_{residual}$ بينما توجد دلالة احصائية.

والاعتبار الثالث القوة الاحصائية للنماذج المثبتة فى مقابل النماذج العشوائية او للنماذج المثبتة مع المتغيرات المتفاعلة فى مقابل النماذج المختلطة، فالقوة الاحصائية فى ما وراء التحليل تعتمد على عدة عوامل منها عدد الدراسات او احجام التأثير فى التحليل، احجام العينات فى الدراسات، مدى تصحيح احجام التأثير من المحددات المنهجية للدراسات، مقدار تباين مجتمع احجام التأثير، متوسط احجام التأثير فى المجتمع، وغيرها.

والنماذج المثبتة اكثر قوة احصائية من النماذج العشوائية عما يكون تباين احجام التأثير فى المجتمع تساوى صفراً (τ^2) او اكبر من الصفر. بالمثل تحليل النماذج المثبتة مع المتفاعلات اكثر قوة احصائياً من النماذج المختلطة. وعلى ذلك فان النماذج المثبتة اكثر قوة احصائياً من نظيرتها خاصة عندما يكون الاختلاف بين احجام التأثير غير دال احصائياً.

اما وجود قيم متطرفة لاحجام التأثير او الاخطأ المعيارية فمن الافضل استخدام النماذج العشوائية مقارنة بالنماذج المثبتة. والقيم المتطرفة تؤثر بصورة مباشرة على متوسط حجم التأثير فى النموذج المثبت وهذا بدوره يحرك هذا المتوسط ليقع داخل فترات الثقة بينما احجام التأثير المتطرفة تؤثر فى تباين المجتمع τ^2 اكثر من تأثيرها على متوسط حجم التأثير فى النموذج العشوائى.

النوع الاخر من التطرف يتضمن الدراسات التي تتضمن احجام عينات كبيرة جداً ولان حجم العينة مرتبط بالخطأ المعيارى وهذا بدوره يؤثر على اوزان الدراسات فى نماذج التأثيرات المثبتة فالدراسات ذات احجام العينات الكبيرة تكون اكثر وزناً من نظيراتها ذات احجام العينات الصغيرة.

فى ضوء ذلك مع وجود تطرف سواء لاحجام التأثير او احجام العينات كبيرة فى الدراسات يفضل استخدام نموذج التأثيرات العشوائية او المختلطة.

اما بالنسبة لمعيار التعقيدات الحسابية فان نموذج التأثيرات المثبتة لحساب متوسط حجم التأثير او لتقويم المتفاعلات (خصائص الدراسات) اكثر بساطة من نموذج التأثيرات العشوائية ونموذج التأثيرات المختلطة. وعلى الرغم ليس هذا بالمعيار الجوهرى لان الباحث لا يجب ان يهتم بالبساطة على حساب الدقة لتقدير متوسط حجم التأثير خاصة اذا لم تتوفر مسلمات نموذج التأثيرات المثبتة وهو على الاقل وجود تجانس بين احجام التأثير.

حقيقة استخدام نموذج التأثيرات العشوائية مع وجود تجانس بين احجام التأثير او استخدام نموذج التأثيرات المثبتة مع وجود اختلاف بين احجام التأثير يعتبر اختيار خاطئ لنموذج تحليل ما وراء التحليل. وتكون نتائج النموذج المثبت والنموذج العشوائى متقاربة او متشابهة عندما يوجد اختلاف قليل بين مجتمع احجام التأثيرات.

وبالنظر الى طبيعة الدراسات ومتغيرات الظواهر النفسية والاجتماعية حيث ان احجام التأثير المقاسة لا تناظر احجام التأثير الحقيقية فى المجتمع، ونظراً لتعقيد هذه الظواهر وتشابك متغيراتها وتفاعلها فان المفاهيم النظرية والاحصائية تحتم استخدام على الاقل نموذج التأثيرات العشوائية للوصول الى استنتاجات لها قدرة تعميمية. ولكن عندما تكون هدف الدراسة فحص تأثير المتغيرات المتفاعلة على اختلاف احجام التأثير يفضل استخدام نموذج التأثيرات المثبتة مع تحليل المتفاعل ولكن الاختيار الاكثر افضلية هو استخدام نموذج التأثيرات المختلطة ولكن هذا النموذج يتسم بتعقيدات حسابية معقدة وعليه فان اختيار نموذج التأثيرات المثبتة مع المتفاعلات اختيار مقبول يتميز بالبساطة مقارنة بالنموذج المختلط.

الفصل الحادى عشر

تحيز النشر

Publication Bias

مقدمة

يحدث تحيز النشر عندما تكون عينة الدراسات المتضمنة فى ما وراء التحليل غير ممثلة تمثيلاً جيداً لمجتمع الدراسات المنشورة للعلاقة بين المتغيرين موضع اهتمام دراسة ما وراء التحليل وخاصة عندما تتباين نتائج الدراسات المتضمنة فى التحليل عن الدراسات التى تناولت الظاهرة المراد دراستها وفى هذه الحالة توجد مخاطرة فى الوصول الى استنتاجات خاطئة عن التراث البحثى للعلاقة بين المتغيرين.

ماهية تحيز النشر فى ما وراء التحليل

من اهم الانتقادات الموجهة لاسلوب ما وراء التحليل هى ان الدراسات المتاحة للتحليل تمثل عينة متحيزة للدراسات الموجودة ويطلق عليها بتحيز النشر ولكنه ليس المصدر الوحيد للتحيز ويفضل (Hunter & Schmidt, 2004) تسميته تحيز الاتاحية Availability bias بينما يسميه البعض تحيز الاسترجاع او تحيز الانتقاء.

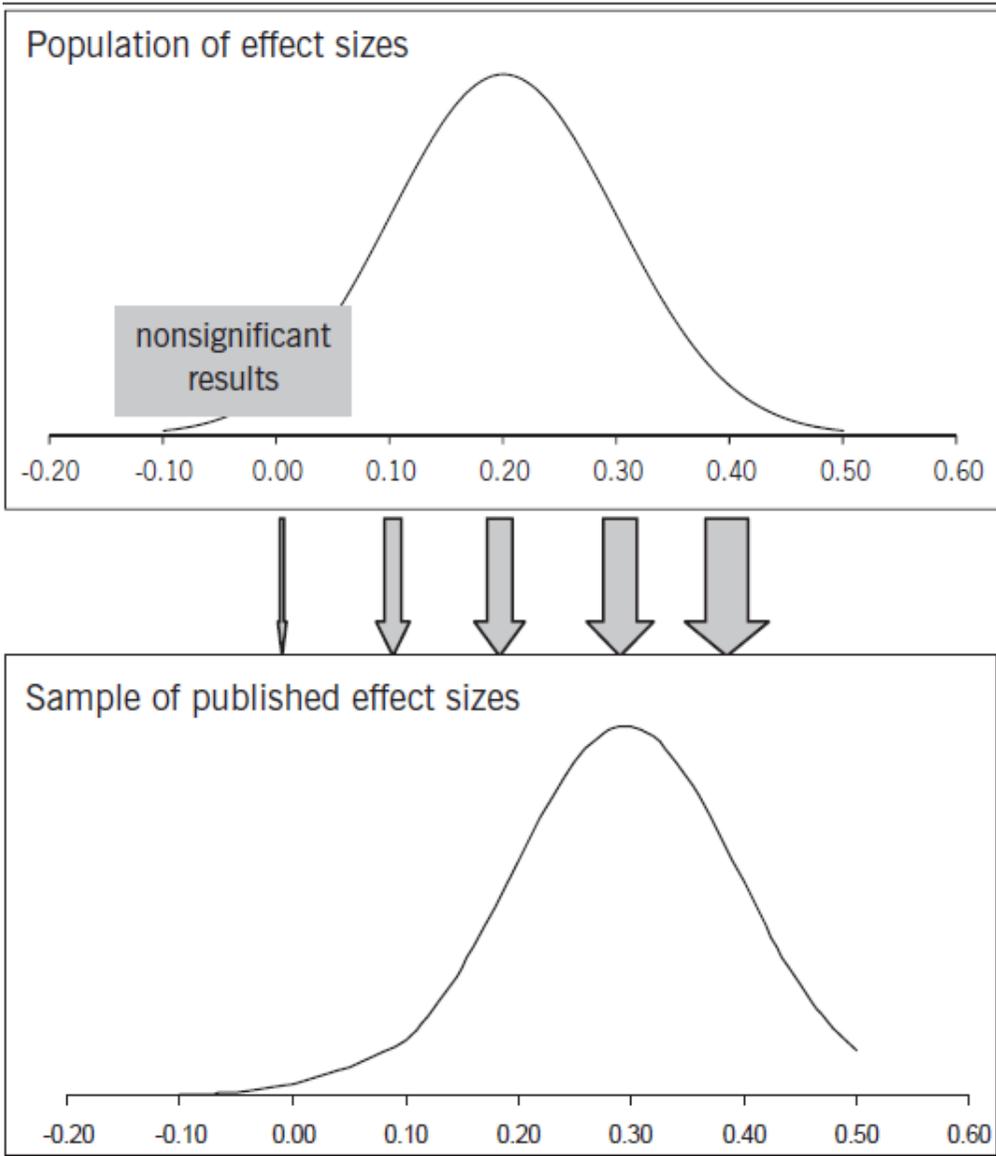
ما وراء التحليل هو توليف رياضى دقيق للدراسات المتضمنة فى التحليل فلو ان الدراسات تمثل عينة متحيزة لتراث الظاهرة فان متوسط حجم التأثير يكون متحيز (Borenstein et al., 2009) وعليه فان متوسط حجم التأثير يكون تقدير مرتفع عن حجم التأثير الحقيقى فى المجتمع ويعتبر تحيز النشر قضية محتملة فى كل المنهجية البحثية سواء البحث الكيفى او الدراسات الكمية والمراجعات السردية والمراجعات الكمية مثل ما وراء التحليل.

فلو اهتم الباحث فى دراسة ما وراء التحليل بالدراسات المنشورة فقط فيوجد مخاطرة كبيرة بوجود تحيز فى الاستنتاجات التى يصل اليها الباحث لانه يوجد العديد من الدراسات غير المنشورة ذات النتائج غير الدالة احصائياً التى لا تقبل للنشر فى المجالات العلمية لانها تهتم فى المقام الاول

بالدراسات ذات النتائج الدالة احصائياً ولكن اذا تضمنت دراسة ما راء التحليل النتائج الدالة وغير الدالة احصائياً بكل تأكيد تتغير طبيعة الاستنتاجات فى دراسة ما وراء التحليل، وعلى ذلك تنشأ مشكلة تحيز النشر نتيجة عدم الحصول على النتائج غير الدالة احصائياً او ان النتائج الدالة احصائياً تكون غير منشورة مثل دراسات الماجستير والدكتوراة.

ويشير (Card 2012) الى ان احتمالية النتائج الصفرية(عدم دلالة احصائية) او السلبية(العلاقات السالبة) هى اقل احتمالية للنشر من النتائج ذات التأثيرات الموجبة وهذا يرجع الى الباحثين اقل رغبة او دافعية لنشر تلك النتائج او ان المجالات تكون اقل تحملاً لنشر تلك النتائج، وقد يحدث تحيز النشر نتيجة معايير التضمين التى يضعها الباحث حيث يشترط ان تكون الدراسات المتضمنة فى التحليل منشورة فقط. ولكن لو ان نتائج الدراسات غير المنشورة اختلفت عن الدراسات المنشورة على ذلك من المتوقع حدوث تحيز لنتائج ما وراء التحليل. وتأثير تحيز النشر هو ان التراث البحثى للظاهرة موضع الدراسة لا يكون ممثلاً للدراسات التى اجريت فى هذا الموضوع ومن المحتمل ان تظهر النتائج فى التراث حجم تأثير قوى مقارنة بالدراسات الاخرى المستبعدة من النشر خاصة الدراسات ذات النتائج الصفرية.

ويمكن توضيح ذلك بالشكل التالى (Card, 2012):



الشكل (1.11): تأثير تحيز النشر على حجم التأثير.

فى الجزء الاعلى يوضح توزيع حجم التأثير من مجتمع الدراسات ويلاحظ ان احجام التأثير للدراسات تتمركز حول 0.20 وهذا المجتمع يتضمن نتائج سالبة وصفرية. بينما فى الجزء الاسفل يظهر توزيع حجم التأثير لعينة من الدراسات المنشورة ويلاحظ ان التوزيع يتحرك نحو اليمين المنحنى واصبح متوسط حجم التأثير 0.30 ، وعليه لو ان دراسة ما وراء التحليل اعتمدت على عينة الدراسات المنشورة واهملت الدراسات غير المنشورة فان هذا يسبب تضخم و زيادة متوسط حجم التأثير عن قيمتها الحقيقية لمجتمع الدراسات وهذا من شأنه يودى الى استنتاجات خاطئة عن العلاقة بين المتغيرين موضع دراسة ما وراء التحليل. وعليه فلو ان عينة الدراسات فى ما وراء التحليل متحيزة فان صدق النتائج يكون موضع تهديد وشك.

ويعرف (2005) Dickersin الاتجاه للنشر او عدم النشر اعتماداً على اتجاهاً وطبيعة نتائج الدراسة.

ويوجد ادلة قوية لوجود تحيز النشر في العلوم الاجتماعية والطبية في الدراسات الملاحظة والتجريبية، فعندما لا يصل الباحث الى نتائج دالة احصائياً من التجربة التي انجزها يحجم على نشرها لان النتيجة تبدو غريبة ولا تتفق مع الاخرين وهذا يمكن ان يكون اكثر وضوحاً في المجال الطبي حيث يكون الباحث متحمساً لنشر النتائج الموجبة لانها تدل على النجاح بينما النتائج الصفرية او السلبية تدل على الفشل. ويعتبر هذا من الازغاء الجسيمة في مختلف التخصصات وذلك لان النتيجة الصفرية يمكن ان تثير نقاش وحوار واستفسارات علمية في التخصص اكثر من النتيجة الدالة.

والنتائج السالبة تستخدم لتشير الى كلاً من:

- اتجاه النتائج بمعنى النتائج تكون لصالح المجموعة التي تتلقى المعالجة التقليدية (الضابطة).
- نقص الدلالة الاحصائية للنتائج بغض النظر ما اذا كانت موجبة او سالبة او صفرية.

ومن المحتمل ان تكون الدلالة احصائية وحجم التأثير للدراسات المنشورة اعلى من الدراسات غير المنشورة، وعلى ذلك فان حجم التأثير المقدر من دراسات ما وراء التحليل ان يكون متحيز تحيز موجب Upward (Hunter & Schmidt, 2004). وهذا نتيجة ان الباحثين لا يستعينون بالدراسات غير المنشورة نتجة عدم اتاحة الفرصة للحصول عليها.

وفي ضوء ما قام به Rosenthal (1984) بفحص 12 دراسة ما وراء التحليل في موضوعات مختلفة لتحديد ما اذا كان حجم التأثير يختلف باختلاف المصدر وبناءً على مئات احجام التأثير وجد لا فروق بين متوسط احجام التأثير بين الدراسات المنشورة في مجلات وتلك غير المنشورة، وان متوسط d اعلى للدراسات غير المنشورة بمقدار 0.08 بينما وسيط d اعلى للدراسات المنشورة بمقدار 0.05.

ويأخذ تحيز النشر مسميات عديدة منها اشكالية بـ File-drawer problem بمعنى ان الدراسات ذات النتائج الصفرية او السالبة يهملها باحث ما وراء التحليل، وتأخذ مسمى

Dissemination Bias ويرى (Rothstein, Sutton, & Borenstein, 2005) ان هذا المصطلح اكثر شمولية من Publication Bias وبغض النظر عن المصطلح المستخدم فان احد مصادر هذا التحيز موجود فى التراث الرمادى غير الواضح Gray Literature وهو تراث بين النتائج الصفرية والتقارير الفنية الصادرة عن مؤسسات حكومية او تجارية او صناعية او تقارير المؤتمرات ولا تكون صادرة عن ناشرين تجارين ويكون تحت النشر او مطبوع الكترونياً. ويرى (Hopewell, Clarke, & Mallett (2005) ان الدراسات التقييمية فى العلوم الاجتماعية والصحية اوضحت وجود فروق بين نتائج الدراسات المنشورة وتلك الصادرة عن مؤسسات حكومية وخاصة ومؤتمرات. وتوجد قواعد بيانات تتضمن هذه النوعية من الدراسات مثل:

- Dissertation Abstract Online : حيث تتضمن رسائل الدكتوراة فى الولايات المتحدة الامريكية منذ 1861.
- Economic & Social Data Service: ارشيف بيانات بالمملكة المتحدة يتضمن بيانات كمية وكيفية فى الاقتصاد والاجتماع.
- Science Citation Index: يتضمن مستخلصات فى المجال الطبى.
- Sociological Abstract: تغطى رسائل الدكتوراة ومستخلصات المؤتمرات وفصول كتب منذ 1963 فى العلوم السلوكية والاجتماعية.
- Social Services Abstracts & Info Net: تتضمن رسائل الدكتوراة فى العمل والساسات الاجتماعية من 1980.

ومعظم الدراسات التقييمية المسحية فى الفترة من 1970 حتى 1980 تشير الى ان الدراسات ذات النتائج التى ترفض الفرض الصفرى وهى اكثر قابلية للنشر، ومن خلال مسح ومراجعة لدراسات وجد ان 28% من الدراسات التجارب الاكلنيكية لم تقبل للنشر لان نتائجها سالبة (Diskersin, Chan, Chalmers, Sacks, & Smith, 1987).

واستقصى (Cooper, Deneve, & Charlton (1997) حوالى 159 دراسة مقبولة وفقاً لاختلافات البحث فى الفترة من 1986 حتى 1988 فى قسم علم النفس بجامعة ميسورى، وجد من 117 دراسة اجرت تحليل بيانات واختبارات فروض حوالى 62% من الدراسات توصلت الى

نتائج دالة احصائياً وان 50% من الدراسات ذات النتائج الدالة احصائياً قوبلت للنشر في مؤتمرات وان 7% من النتائج غير الدالة احصائياً تم نشرها، وان 74% من الدراسات الدالة احصائياً نشرت في كتب او مجلات مقارنة بـ 4% من الدراسات غير الدالة نشرت في كتب او مجلات.

وراجعت (2006) Rothstein حوالي 95 دراسة ما وراء التحليل في مجلة Psychological Bulletin فى الفترة من 1995 الى 2005 ليقوم مدى تضمين الدراسات غير المنشورة او دراسات التراث البحثى Grey، فوجدت 23 دراسة ما وراء التحليل لم تتضمن دراسات غير منشورة.

ومصادقية المراجعة المنظمة تعتمد بدرجة كبيرة على النتائج المجمعة من الدراسات، والهدف من البحث الجيد فى التراث الحصول على قائمة شاملة للدراسات غير المتحيزة التى تكون مناسبة للاجابة على الاسئلة البحثية.

مصادر تحيز النشر

وتوجد مصادر للتحيز تؤثر فى عملية البحث كالتالى (Hopewell et al., 2005):

- **تحيز النشر:** الدراسات ذات النتائج الدالة احصائياً اكثر احتمالية للنشر من الدراسات ذات النتائج الصفرية والسالبة.
- **تحيز الوقت Time-Lag bias:** الدراسات الدالة احصائياً تنشر بسرعة من نظيرتها الصفرية والسالبة.
- **تحيز اللغة Language bias :** الدراسات الدالة احصائياً اكثر احتمالية للنشر باللغة الانجليزية.
- **تحيز الاستشهاد Citation bias :** الدراسات الدالة احصائياً اكثر احتمالية للاستشهاد بها من الباحثين الاخرين.
- **تحيز الاتاحة Availability bias:** تضمين الدراسات التى من السهل الحصول عليها.
- **تحيز التكلفة Cost bias :** تضمين الدراسات المتاحة بدون تكلفة او الاقل تكلفة.

- **تحيز الالفة Familiarity bias:** تضمنين الدراسات الاكثر الفة للباحث وهذا يرتبط بعامل اللغة او الثقافة.

محاولة تجنب التحيز

توجد عدة طرق لتجنب التحيز للدراسات المجمعفة فى ما وراء التحليل اهمها:

- **البحث فى قواعد البيانات المتخصصة والشاملة Electronic Database:** الخطوة الاولى فى عملية البحث هى البحث فى قواعد البيانات الالكترونية وتختلف من تخصص الى اخر. فى المجال الطبى توجد قاعدة Cohanace Central Register of Controlled Trials(CENTRAL) تضم تسجيل لكل التجارب او المحاولات التدخلية فى المجال الصحى وكذلك ملخصات المؤتمرات وتجارب تدخلية من لغات اخرى غير الانجليزية وهى اكثر شمولية من قواعد بيانات طبية اخرى مثل MEDLINE. وفى مجال التربية وعلم النفس يوجد مركز مصادر المعلومات التربوية Educational Resources Information Center(ERIC) وقاعدة بيانات PsycINFO والملخصات السوسولوجية Sociological Abstract.

واجرى (Turner, Boruch, Laveberg, Schoeneberger, & De Moya (2004) عملية البحث اليدوى فى مقابل البحث فى قواعد البيانات للبحث فى موضوع ما فى 12 مجلة تربوية او مجلة مرتبطة و اشارت النتائج المتحصل عليها من قواعد البيانات تمثل ثلث الدراسات التى حصل عليها من خلال البحث اليدوى.

- **البحث فى عروض المؤتمرات Conference Proceedings:** دراسات المؤتمرات تعتبر مصدر جيد لتحديد الدراسات المتضمنة فى ما وراء التحليل وملخصات دراسات المؤتمرات عادة لا تتضمن فى قواعد البيانات الالكترونية وتوجد عوامل عديدة تعوق نشر دراسات المؤتمرات منها دلالة النتائج. ويرى (Hopewell et al.(2005) ان الدراسات فى مجملها توصلت الى ان 44.5% ان الدراسات التى تعرض فى المؤتمرات العلمية كملخصات يتم نشرها كاملة.

- **الاتصال مباشرة مع الباحثين Contact with researchers** : توجد بعض الدراسات لا تطبع باى شكل من الاشكال ومن المهم تضمينها فى ما وراء التحليل ويمكن ان تطبع الدراسة ولكن لظروف ما لا يستطيع الباحث الحصول عليها، وعليه يمكن التواصل شخصياً مع المؤلف تليفونياً او من خلال البريد الالكترونى او اى وسيلة اخرى وتضمين هذه الدراسات من شأنه ان يقلل التحيز.
 - **البحث فى تسجيلات البحوث Searching research registers**: تعتبر تسجيلات البحوث مصدر جيد للحصول على الدراسات ويوجد تسجيلات البحوث القومية مثل مركز البحوث التربوية او مركز القياس والتقويم او مركز البحوث العلمية التى تتضمن تسجيلاً للبحوث والدراسات على المستوى القومى.
 - **البحث فى الانترنت**: يعتبر الانترنت مصدر جيد للحصول على الدراسات خاصة غير المنشورة فى مجلات متخصصة ويوفر الانترنت كثيراً من المواقع web sites التى تتضمن كثيراً من الدراسات والضرورية فى ما وراء التحليل.
- كل المصادر السابقة تقلل بقدر الامكان التحيز فى انتقاء الدراسات السابقة فى ما وراء التحليل.

الطرق الاحصائية لتقييم تحيز النشر

يوجد ستة مداخل للتحكم فى تحيز النشر فى اطار ما وراء التحليل ويمكن عرض بعض هذه المداخل فى اطار المثال الذى تم عرضه لـ 22 دراسة للعلاقة بين العدوان ورفض الاقران بين الاطفال والمراهقين كما ورد فى (Card 2012).

فى الفصل(الثامن) تم عرض نتائج نموذج التاثيرات المثبتة حيث:

- متوسط حجم التاثير(معامل الارتباط) = 0.368 والخطأ المعياري = 0.0118 .
 - قيمة اختبار Z لدلالة متوسط حجم التاثير = 23.70 (P < .01).
 - فترات الثقة تراوحت من 0.348 الى 0.388.
- والمطلوب تقييم ما اذا تحيز النشر يهدد او يؤثر فى النتائج والاستنتاج حول متوسط حجم التاثير.

الجدول(1.11): جدول البيانات لـ 22 دراسة للعلاقة بين العدوان ورفض الاقران لتقييم تحيز النشر.

Study	Sample size (N)	Effect size (r)	Transformed ES (Z_r)	Standard error (SE)	Published (1 = yes)	v^*	ES*	z	1/SE
Blachman (2003)	228	.525	.583	.525	0	.005	2.87	7.57	14.43
Crick & Grotpeter (1995)	491	.198	.201	.198	1	.003	-3.28	3.42	17.26
Crick et al. (1997)	65	.311	.322	.311	1	.017	-0.49	2.35	7.55
Geiger (2003)	458	.554	.624	.554	0	.002	5.06	11.45	20.67
Hawley et al. (2007)	929	.161	.162	.161	1	.001	-6.90	4.65	28.91
Henington (1996)	904	.336	.349	.336	1	.001	-1.14	9.69	28.84
Johnson (2003)	74	.396	.419	.396	0	.015	0.27	3.24	8.18
Leff (1995)	151	.617	.721	.617	0	.007	3.94	7.22	11.69
Miller (2001)	150	.557	.628	.557	0	.007	2.89	6.59	11.83
Murray-Close & Crick (2006)	590	.575	.655	.575	1	.002	6.57	13.50	23.47
Nelson et al. (2005)	180	.039	.039	.039	1	.007	-4.22	0.47	12.03
Ostrov (under review) ^a	139	.358	.375	.358	0	.008	-0.13	4.02	11.21
Ostrov & Crick (2007)	132	.049	.049	.049	1	.008	-3.69	0.54	10.85
Ostrov et al. (2004) ^b	60	.000	.000	.000	1	.019	-2.80	0.00	7.21
Pakaslahti & Keltikangas-Jarvinen (1998)	839	.326	.339	.326	1	.001	-1.32	8.57	26.26
Phillipsen et al. (1999)	262	-.048	-.048	-.048	1	.004	-6.89	-0.75	15.59
Rys & Bear (1997)	266	.454	.489	.454	1	.004	1.65	7.13	15.71
Salmivalli et al. (2000)	209	.253	.258	.253	1	.005	-1.82	3.54	13.98
Tomada & Schneider (1997)	314	.160	.162	.160	1	.003	-3.90	2.73	16.99
Werner (2000)	881	.477	.519	.477	0	.001	3.99	13.57	28.47
Werner & Crick (2004)	517	.469	.509	.469	1	.002	2.78	10.30	21.96
Zalecki & Hinshaw (2004)	228	.572	.651	.572	1	.005	3.82	8.15	14.24

فالاعدة الخمسة الاولى هي المؤلف، حجم العينة، حجم التأثير r ، حجم التأثير في ضوء تحويل فيشر Zr ، والخطأ المعياري لحجم التأثير المحور SE ، اما الاعدة الباقية تتضمن البيانات المستخدمة في تشخيص تحيز النشر فيما يلي طرق تشخيص التحيز كالتالى:

تحليلات المتفاعل Moderators Analyses

من اهم الطرق لتقييم التأثير المحتمل لتحيز النشر هي تضمين دراسات غير منشورة (الماجستير او الدكتوراة او ابحاث المؤتمرات) فى دراسة ما وراء التحليل وتقييم ما اذا كانت الدراسات غير المنشورة تعطى احجام تأثير صغيرة مقارنة بنظيراتها غير المنشورة من خلال تحليل المتفاعل باعتبار ان نوعية النشر متغير تصنيفى بمستويين (منشور فى مقابل غير منشور). ولكن لاجراء هذا التحليل لابد من:

- ان تتضمن دراسة ما وراء التحليل عدد كافى من الدراسات غير المنشورة للوصول الى استنتاجات صادقة عن الفروق بين نوعى الدراسات.
- من الضرورى الاخذ فى الاعتبار خصائص او ملامح اخرى للدراسات المنشورة فى مقابل غير المنشورة مثل الجودة المنهجية (الصدق الداخلى المرتفع فى مقابل المنخفض للتصميم التجريبي) وجودة المقاييس (مقاييس ثابتة فى مقابل غير ثابتة) فيجب ضبط الفروق بين الدراسات المنشورة فى مقابل غير المنشورة من خصائص الدراسات المحتملة التى يمكن ان تحدث فروق بين الدراسات.

ولاجراء هذا النوع من التحليل تكود الدراسات المنشورة بكود (1) والدراسات غير المنشورة بكود (0) كما واضح فى العمود السادس، وبلغ عدد الدراسات المنشورة ($K=15$) وعدد الدراسات غير المنشورة ($K=7$) وعلى الرغم من التفاوت بين نوعى الدراسات الا انه من الممكن اجراء التحليل للمقارنة بينهما وباجراء اختبار تحليل التباين F يتضح ان قيمة اختبار Q بين المجموعات :

$$Q_{between}(df=1) = 77.47, P < .001$$

وعليه يتضح فروق دالة احصائياً فى متوسط حجم التأثير بين الدراسات المنشورة والدراسات غير المنشورة. وعليه توجد قضية تحيز للنشر.

وبفحص متوسط حجم التأثير في كل مجموعة يتضح انها للدراسات غير المنشورة (0.51) اعلى من متوسط الدراسات المنشورة (0.31) وهذا يتعارض الى حداً ما مع الدراسات التقييمية، وعليه فان الدراسات غير المنشورة لها احجام تأثير مرتفعة وهذا عكس مما توقعناه لان الدراسات غير المنشورة تكون نتائجها صفرية بالتالي حجم التأثير لها منخفض ولكن في المثال السابق كانت الدراسات غير المنشورة هي رسائل دكتوراة حيث المنهجية والاجرائية اكثر جودة مقارنة باغلبية الدراسات المنشورة في المجالات العلمية المحكمة.

ولكن باجراء اختبار T المستقلة للفروق بين الدراسات المنشورة والدراسات غير المنشورة كانت النتائج كالتالي:

Group Statistics

	VAR00005	N	Mean	Std. Deviation	Std. Error Mean
Zr	.00	7	.55271	.122872	.046441
	1.00	15	.28220	.217545	.056170
r	.00	7	.49771	.093051	.035170
	1.00	15	.26340	.192733	.049764

Independent Samples Test

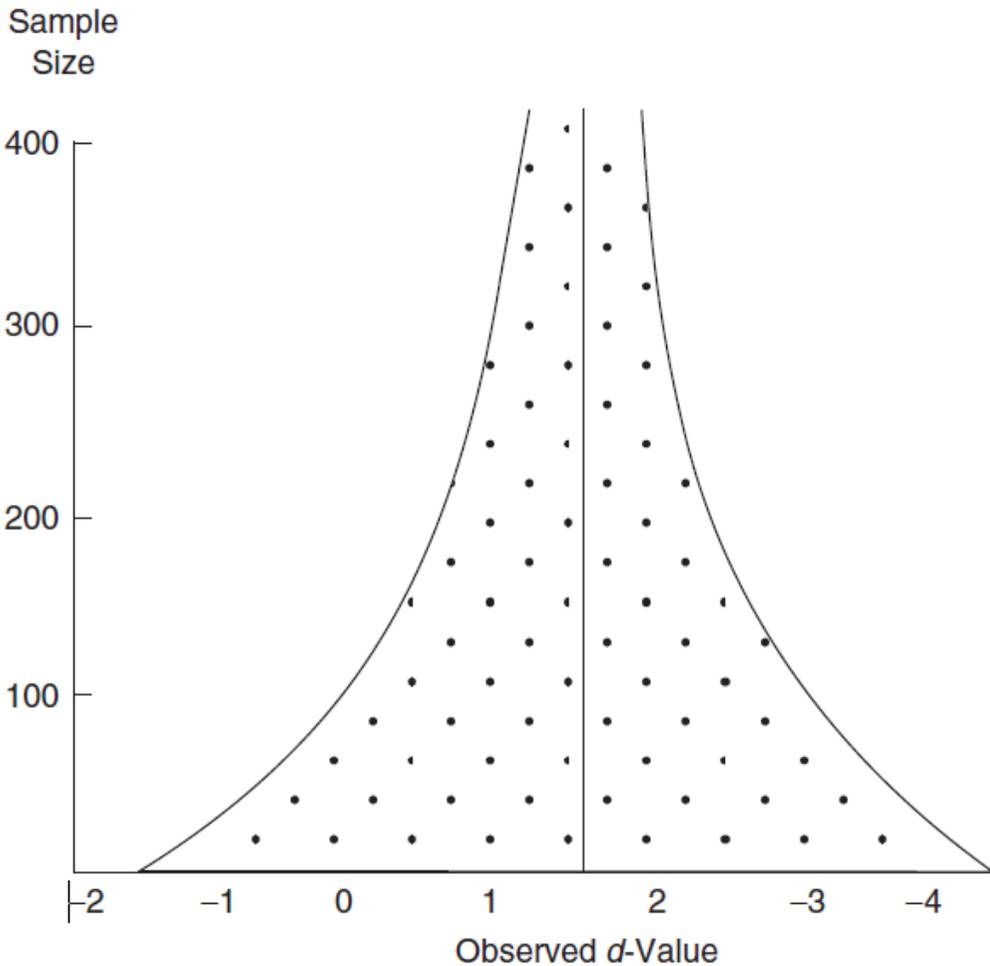
		Levene's Test for Equality of Variances				
		F	Sig.	t	df	Sig. (2-tailed)
Zr	Equal variances assumed	2.938	.102	3.045	20	.006
	Equal variances not assumed			3.712	18.984	.001
r	Equal variances assumed	4.621	.044	3.027	20	.007
	Equal variances not assumed			3.845	19.896	.001

حيث قيمة T لتحويل فيشر 3.045 وهي دالة عند مستوى دلالة احصائية 0.01 وعليه يوجد تحيز للنشر. كما ان متوسط حجم التأثير للدراسات غير المنشورة في ضوء r من النوع الكبير الحجم 0.4977 (تقريباً 0.50)، اما للدراسات المنشورة 0.263 من النوع الضعيف، وهذا يخالف الاعتقاد السائد بان الدراسات المنشورة لها حجم تأثير اكبر من الدراسات المنشورة. وبالتالي عدم تضمين الدراسات غير المنشورة ينقص من حجم التأثير للعلاقة بين المتغيرين في ما وراء التحليل.

شكل القمع Funnel Plots

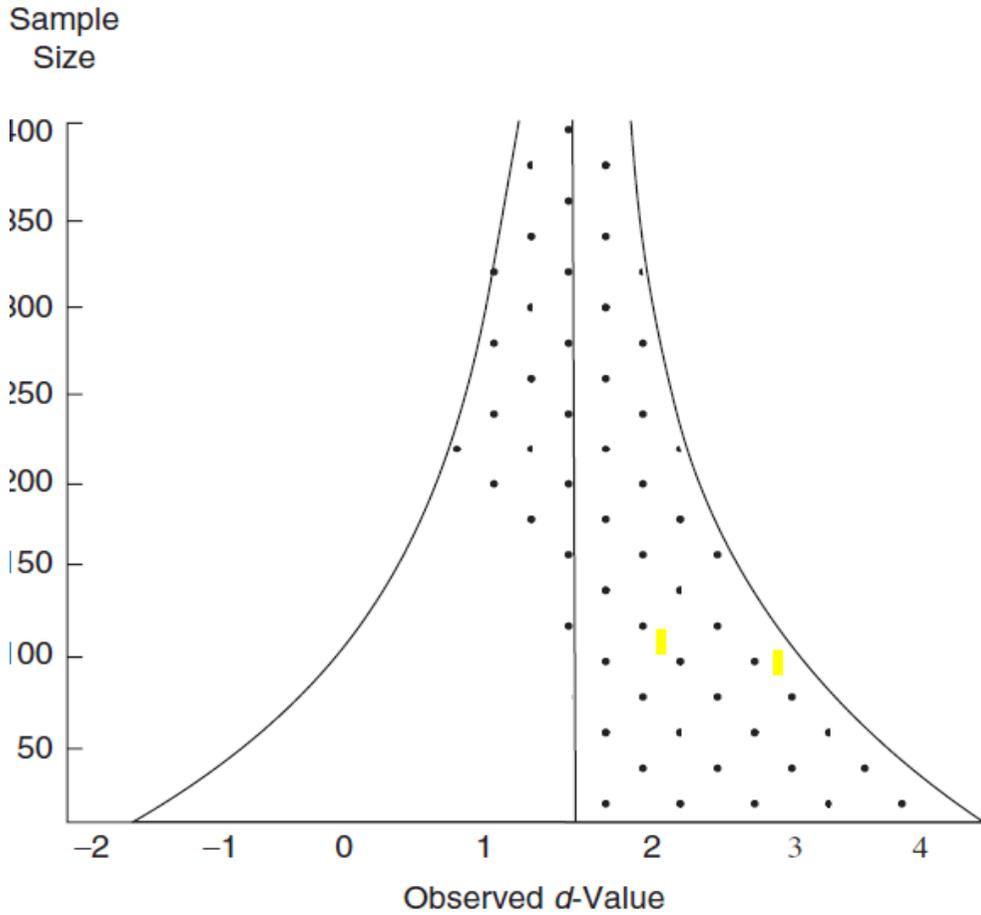
قدم (Light & Pillemer, 1984) رسم بسيط لتشخيص تحيز النشر وإى تحيزات أخرى، و شكل القمع عرض او مخطط بصرى بسيط لفحص تحيز النشر وإى تحيزات أخرى فى ما وراء التحليل ويستخدم بصورة متزايدة فى تراث الدراسات الاجتماعية والطبية.

وهذا الطريقة تعتبر فى حالة غياب تحيز النشر فانه يتوقع ان يكون متوسط حجم التأثير هو نفسة للدراسات ذات احجام العينات الصغيرة وللدراسات ذات احجام العينات الكبيرة، بينما اذا حدث اختلاف كبير فى الدراسات ذات احجام العينات الصغيرة فانه يرجع الى خطأ المعاينة وفى غياب التحيز فان الرسم يأخذ شكل القمع المقلوب كالاتى (Hunter & Schmidt, 2004):



الشكل (2.11): مثال يوضح شكل القمع المنتظم ليبدل على عدم تحيز.

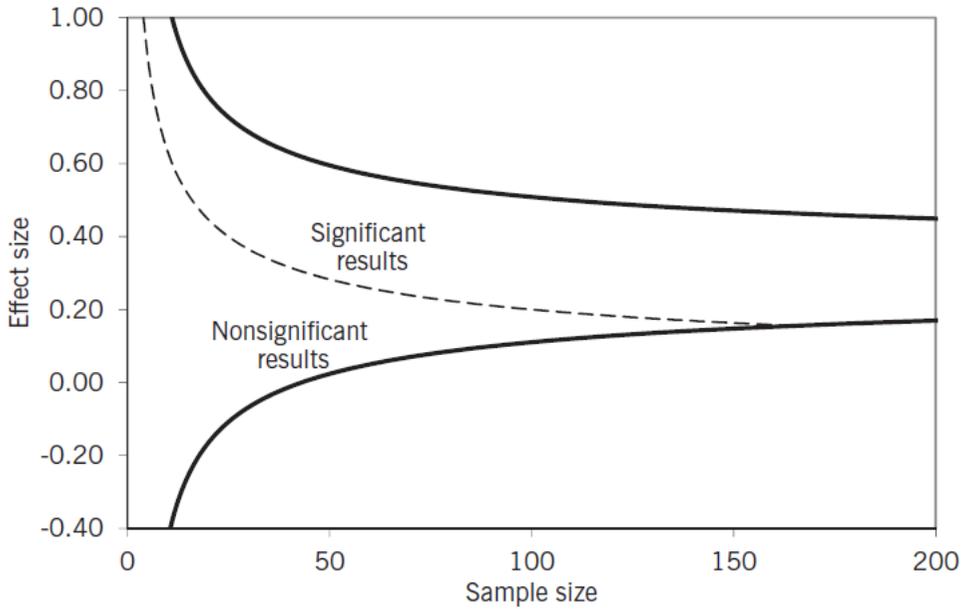
حيث ان متوسط حجم التأثير يكون نفسه بغض النظر عن حجم عينة للدراسة، ولكن لو وجد تحيز للنشر بمعنى ان الدراسات الصغيرة التي من المفترض ان لا تظهر دلالة احصائية وتظهر حجم تأثير صغير ويكون موقعها على الناحية اليسرى اسفل القمع ولكنها تكون العكس حيث تظهر احجام تأثير كبيرة مقارنة بالدراسات الكبيرة ويكون موقعها على الناحية اليمنى من المنحنى ويكون شكل القمع كالتالى:



الشكل(3.11): شكل القمع غير المنتظم ليدل على حيز النشر.

وعرض القمع عرض بياني حيث للعلاقة بين تأثيرات المعالجات او احجام التأثير المقدره من الدراسات المفردة بالنسبة الى حجم العينة ومسمى رسم القمع قائم على حقيقة ان الدقة فى تقدير تأثير المعالجة يزيد كلما زادت حجم عينة الدراسات المتضمنة فى التحليل.

يمكن عرضه كالتالى(Card, 2012):



الشكل (4.11): مثال افتراضى لعرض شكل القمع.

فالشكل السابق يوضع حجم التأثير Zr او r او OR على المحور الصادى Y وحجم العينة N على المحور الافقى السينى X او العكس والخطان الاسودين يتضمنوا 95% من النقاط او الاحداثيات حول متوسط r (0.30) او متوسط تحويل فيشر Zr (0.31).

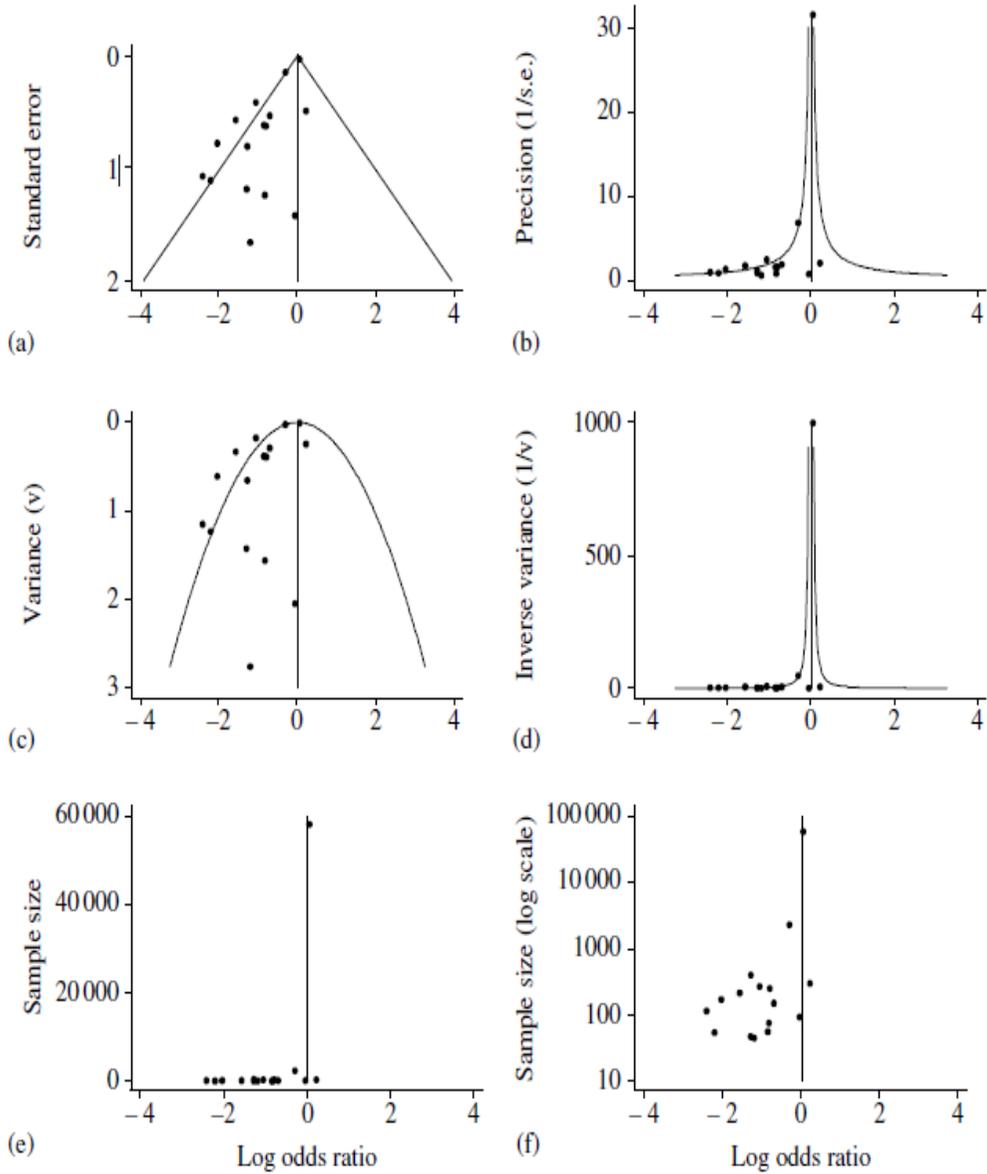
والشكل المنتظم للقمع يعنى عدم وجود تحيز النشر ولكن وجود تحيز للنشر يؤثر فى شكل القمع. لاحظ ان الخط المنقوط فى وسط الشكل فالمنطقة اعلى هذا الخط تتضمن الدراسات الدالة احصائياً بينما المنطقة اسفل الخط تتضمن الدراسات غير الدالة احصائياً. فلو وجد تحيز النشر فان بعض النقاط او الاحداثيات تقع داخل منطقة عدم الدلالة الاحصائية وهذا يحول شكل القمع من الشكل المنتظم الى الشكل غير المنتظم.

ويرى (Card 2012) ان عدم انتظامية شكل القمع لا ترجع فقط الى تحيز النشر فلو توقعنا ان الدراسات ذات احجام اعيان الصغيرة تعطى حجم تأثير قوى فان عدم انتظام القمع لا يرجع الى تحيز النشر وفى هذه الحالة يجب دراسة الفروق بين الدراسات الصغيرة والدراسات الكبيرة حجم العينة لضبط هذا التأثير قبل انشاء شكل القمع. بينما يرى Egger, Smith, Schneider, & Minder (1997) ان عدم انتظامية شكل القمع نتيجة عوامل اخرى مثل وجود اختلاف حقيقى بين الدراسات نتيجة كثافة وشدة البرنامج التدخلى او خصائص اخرى للدراسات مثل حجم العينة

او خصائصها او الاجراءات التجريبية او التصميم المنهجي الضعيف للدراسات الصغيرة او الاجراءات التحليلية غير المناسبة.

فى عرض شكل القمع يفضل عرض احجام التأثير التى لها توزيع اعتدالى على المحور الصادى او السينى مثل تحويل فيشر لمعامل الارتباط Zr او مؤشر g لهيدجز او لوغاريتم Odds ratio وعرض لوغاريتم حجم العينة على المحور السينى او الصادى وذلك للتعامل مع تنوع احجام العينات الكبيرة جداً واحجام العينات الصغيرة ويمكن وضع الخطأ المعيارى او اوزانه W_i اوبوضع معكوس تباين الخطأ المعيارى $1/SE^2$ على المحور السينى وهذا الاختيار يفضل استخدامه عند تحليل حجم التأثير لوغاريتم Odds ratio، ويوصى (Card (2012 بعرض شكل القمع فى ضوء الاشكال المختلفة للمحور السينى سواء حجم العينة او لوغاريتمها او الخطأ المعيارى او وزن الدراسة لان هذا يعطى الصورة كاملة عن تقييم شكل القمع.

وفيما يلى عرض شكل القمع فى ضوء المؤشرات السابقة لحجم العينة حيث تم وضعها على المحور الراسى بينما حجم التأثير على المحور الاقوى (Sterne, Becker, & Egger (2005):



الشكل (5.11): عرض شكل القمع في ضوء صيغ مختلفة لحجم العينة.

حيث النقاط تمثل احجام التأثير من الدراسات المفردة لـ 16 دراسة كانت نواتجها متغير تصنيفي (تحسن - لا تحسن) وحجم التأثير هو مؤشر Odds ratio والخطوط المنحنية او القطر الرأسى يمثل 95% فترات الثقة حول متوسط حجم التأثير وتم عرض شكل القمع في ضوء لوغاريتم Odds ratio على المحور الافقى فى كل الاشكال بينما تم تمثيل المحور الراسى فى ضوء صور عديدة لحجم العينة كالاتى:

فى الشكل (a): تم تمثّل حجم العينة فى ضوء الخطأ المعياري $\frac{1}{\sqrt{N}}$ على المحور الرأسى فالدراسات ذات احجام العينات الكبيرة لها خطأ معيارى صغير وتقع اعلى العرض البيانى بينما

الخطوط الرأسية تظهر 95% فترات الثقة حول المتوسط وفي غياب الاختلاف بين الدراسات فان 95% من الدراسات تقع داخل القمع المحدد بالخطوط المستقيمة. ولكن في الشكل (a) توجد بعض النقاط تقع على يسار الخط وهذا يعنى ان مؤشر حجم التأثير Odds ratio اقل مما توقعناه وهذا يشير الى وجود تحيز في الدراسات.

وفي الشكل (b) العرض البياني لـ 16 دراسة مع عرض مؤشر الدقة ($\frac{1}{SE}$) على المحور الراسى فالدراسات الكبيرة العينات لها خطأ معيارى منخفض وعليه مستوى دقة عالية واطهر المنحنى ايضاً وجود بعض النقاط تقع على يسار الخط وعليه يوجد فرق بين الدراسات الصغيرة والكبيرة الحجم في احجام التأثير ولكن العرض البياني لا يعطى شكل القمع وتم عرض مضغوط للدراسات الصغير اسفل المنحنى.

وفي الشكل (c) تم تمثيل حجم العينة في ضوء تباين الخطأ المعيارى (SE^2) ولم يعطى الشكل البياني شكل القمع وتم عرض مضغوط للدراسات الصغيرة اسفل المنحنى ولا يوصى باستخدامه. وفي الشكل (d) تم التعبير عن حجم العينة في ضوء معكوس تباين الخطأ المعيارى وهذا الاختيار لم يعطى شكل القمع ايضاً .

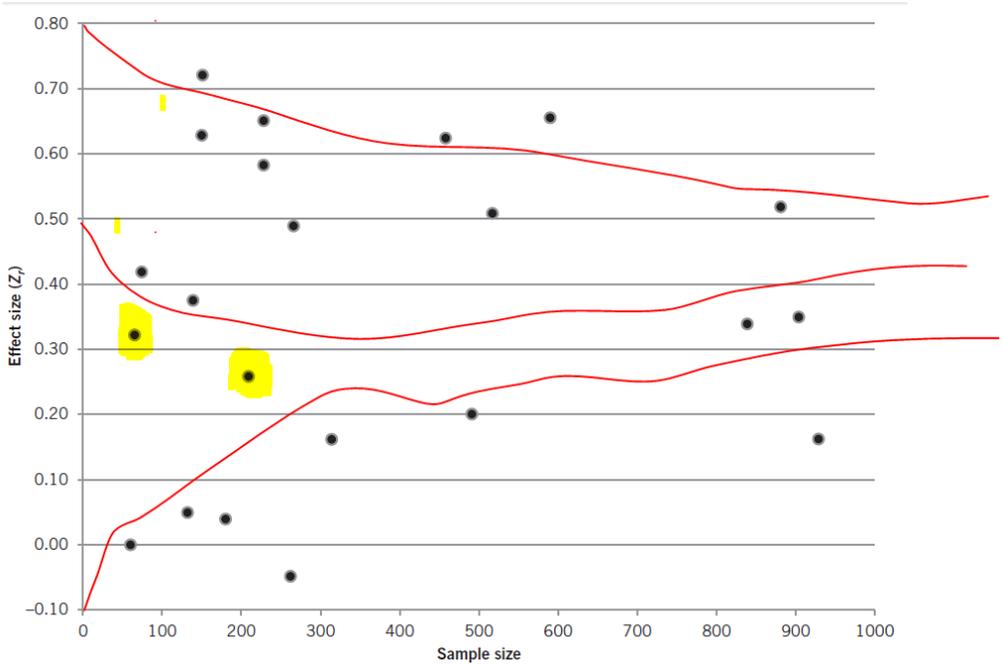
في الشكل (e, f) تم تمثيل حجم العينة مباشرة او التعبير عن حجم العينة في ضوء لوغاريتم حجم العينة لا يوصى باستخدامهما على الاطلاق خاصة اذا وجد تنوع في احجام العينات من الكبيرة جداً الى العينات الصغيرة جداً.

في كل الاشكال الاخرى مهما اختلفت صور تمثيل حجم العينة اظهرت ان بعض النقاط تقع على يسار الخط مما يشير الى عدم انتظامية القمع مما يدل على وجود تحيز في الدراسات المنتقاة في ما وراء التحليل.

ويرى (Sterne et al. 2005) من الافضل الاعتماد على الخطأ المعيارى او مؤشر الدقة ($\frac{1}{SE}$) ووزن الدراسة او تباين الخطأ المعيارى (SE^2) او معكوس تباين الخطأ المعيارى $\frac{1}{SE^2}$ لمؤشر لحجم العينة ويمكن تمثيله على المحور الافقى او الراسى ولكن تمثيله في ضوء الخطأ المعيارى

يعتبر الاختيار الامثل في كل الظروف حيث يسهل فحص شكل القمع بصرياً كما هو الحال في الشكل (a).

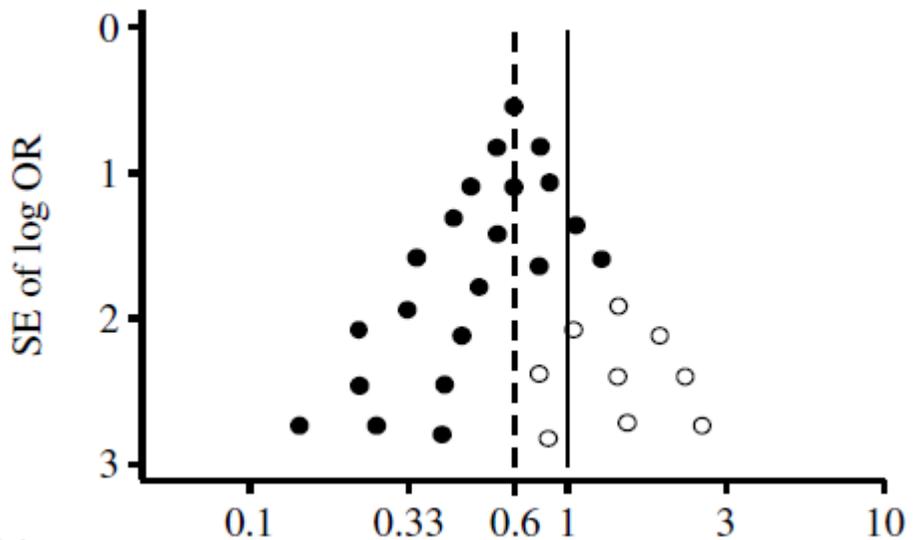
مثال: للجدول (1.8) لـ 22 دراسة وبالعرض البياني لشكل القمع كالتالي (Card,2012):



الشكل (6.11): العرض البياني لشكل القمع لـ 22 دراسة في دراسة ما وراء التحليل.

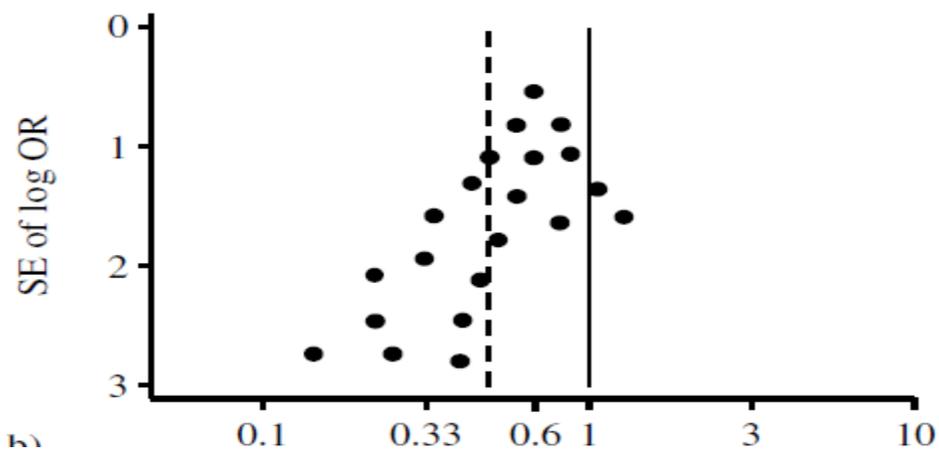
تم عرض حجم التأثير في ضوء تحويل فيشر Zr على المحور الصادي (الراسي) بينما حجم العينة على المحور السيني (الافقي)، الاستنتاج من هذا الشكل يدل على الانتظامية الى درجة كبيرة في شكل القمع حيث ان الدراسات ذات احجام العينات الصغيرة كانت نتائجها غير دالة احصائياً حيث حدث ضيق واضح في شكل القمع من اليسار الى اليمين ولكن يرى باحث اخر ان بعض النتائج غير الدالة احصائياً للعينات الصغيرة وقعت اسفل الخط المنتصف (المظلمة بالاصفر) فقد يوجد تحيز بدرجة بسيطة وعلى ذلك فالحكم على توزيع احجام التأثير في شكل القمع يخضع للتحيز او التفسير الشخصي خاصة عندما يكون حجم عينة الدراسات المتضمنة في ما وراء التحليل صغير فهذا يجعل من الصعب الوصول استنتاجات صادقة.

عندما يوضع مقياس حجم الدراسة (حجم التأثير) على المحور الراسي فان النتائج من الدراسات الصغيرة تكون في قاع العرض البياني مع انتشار ضيق بين الدراسات الكبيرة. وفي غياب التحيز بمعنى عندما تقدر الدراسات نفس حجم التأثير الحقيقي فان العرض البياني يشبه القمع المنتظم كما في الشكل الاتي (Sterne et al., 2005)



الشكل (7.11): القمع المنتظم مع غياب التحيز (الدوائر غير المظللة تشير الى الدراسات الصغيرة اوضحت لا تأثيرات مفيدة).

اذا وجد تحيز على سبيل المثال لان الدراسات الصغيرة اظهرت تأثيرات غير دالة احصائياً (الدوائر المفتوحة) في الشكل السابق تبقى غير منشورة فيظهر شكل القمع منتظم مع فجوة في قاع القمع على الناحية اليمنى كما في الشكل الاتي (Sterne et al., 2005) :



الشكل (8.11): العرض البياني لقمع غير منتظم الشكل لاحجام تأثير مفترضة.

الشكل السابق لتوزيع احجام التأثير في ضوء عرض القمع غير المنتظم نتيجة تحيز النشر حيث ان الدراسات ذات احجام العينات الصغيرة لم تتضمن في التحليل.

وراجع (Light et al. 1994) حوالى 74 دراسة ما وراء التحليل في الفترة من 1985 الى 1991 في مجلة Psychological Bulletin وجدوا دراسة واحدة لما وراء التحليل عرضت شكل القمع لتشخيص قضية تحيز النشر.

تم عرض شكل القمع باعتباره اداة مفيدة لتشخيص تاثيرات الدراسات الصغيرة في ما وراء التحليل، ولكن توجد مسببات اخرى لحدوث عدم انتظامية شكل القمع غير تحيز النشر، والفحص البصرى لشكل القمع ضرورى لفهم طبيعة البيانات في ما وراء التحليل ولكن لزيادة الموضوعية في الحكم على عدم انتظامية شكل القمع توجد العديد من الاختبارات الاحصائية الاستدلالية اهمها:

اسلوب تحليل الانحدار

بمنطق شكل القمع يمكن فحص تحيز النشر باستخدام تحليل الانحدار من خلال انحدار احجام التأثير (متغير تابع) على احجام العينات (متغير مستقل)، ففي وجود ارتباط بين احجام التأثير واحجام العينات فان هذا يدل على عدم انتظامية القمع وهذا يدل على وجود تحيز النشر.

ففي حالة متوسط حجم الاثير الموجب، يظهر تحيز النشر عندما تعطى الدراسات ذات احجام العينات الصغيرة احجام تاثيرات اكبر من الدراسات ذات احجام العينات الكبيرة وفي هذه الحالة يكون معامل الارتباط سالب بين حجم العينة وحجم التأثير، بينما في حالة متوسط حجم التأثير السالب، يوجد تحيز النشر عن طريق العلاقة الموجبة لان الدراسات ذات الصغيرة تعطى احجام تاثير سالبة مقارنة بالدراسات ذات احجام العينات الكبيرة، بينما غياب العلاقة بين حجم التأثير وحجم العينة يشير الى انتظامية شكل القمع وعليه غياب تحيز النشر.

وتوجد طرق احصائية عديدة للتأكد من عدم انتظامية شكل القمع كما عرضها Stern & Egger(2005):

الاول: معامل ارتباط الرتب المصحح Adjusted rank correlation بين احجام التأثير والاطفاء المعيارية ولتقدير هذا التحليل لكل دراسة i وتباين حجم التأثير للدراسة من متوسط حجم التأثير v_i^* وحجم التأثير المعيارى للدراسة ES_i^* فان خطأ التباين المصحح للتقدير حجم التأثير:

$$v_i^* = SE_i^2 - \frac{1}{\sum \frac{1}{SE^2}}$$

وحجم التأثير المصحح:

$$ES_i^* = \frac{ES_i - M_{ES}}{\sqrt{v_i^*}}$$

- M_{ES} متوسط حجم التأثير.
- SE_i^2 مربع الخطأ المعيارى للدراسة.
- $\sum \frac{1}{SE^2}$ مجموع معكوس تباينات الاخطاء المعيارية عبر كل الدراسات.

بعد حساب هذه المتغيرات او المعالم يمكن تقدير معامل ارتباط كندل الرتبى بين ES_i^* و v_i^* فالدلالة الاحصائية لمعامل الارتباط تدل على عدم انتظامية شكل القمع وعليه يوجد تحيز للنشر.

والمدخل الشامل هو طريقة Egger Linear Regression وفيه يتم انحدار الانحراف الطبيعى المعيارى لحجم التأثير لكل دراسة ($Z_i = \frac{ES_i}{SE_i}$) على الدقة (مقلوب الخطأ المعيارى $\frac{1}{SE_i}$) وتكون معادلة الانحدار كالتالى (Stern & Egger, 2005; Card, 2012):

$$Z_i = B_0 + B_1 Precision + e_i$$

وانحدار $\frac{ES_i}{SE_i}$ على $\frac{1}{SE_i}$ من المتوقع ان يكون صفر والدلالة الاحصائية لهذا الثابت تعنى عدم انتظامية شكل القمع ووجود تحيز النشر.

والطرق الاستدلالية التى تقيم عدم انتظامية شكل القمع مثل معامل الارتباط الرتبى لكندل او تحليل الانحدار هى اكثر مصداقية من الفحص البصرى لشكل القمع حيث تقلل من الذاتية

لفحص البصرى ووهى قائمة على الدلالة الاحصائي لكن استخدام مدخل الانحدار ل Egger يتطلب على الاقل 17 دراسة لاجرائه وهذا الحجم من خلال دراسات المحاكاة كما تتطلب طريقة معامل ارتباط الرتب على الاقل 40 دراسة. وعندما يكون تحيز النشر متوسط فان هذا يتطلب على الاقل من 50 الى 60 دراسة لاراء تحليل الانحدار وعلى الاقل 150 دراسة لاجراء معامل ارتباط الرتب لكندل.

مثال: لبيانات 22 دراسة للعلاقة بين العدوان ورفض الاقران حيث تم حساب v_i^* و Z_i و ES_i^* و $\frac{1}{SE}$:

الجدول (1.11): بيانات الـ 22 دراسة وقيم v_i^* و Z_i و ES_i^* و $\frac{1}{SE}$

Sample size (N)	Effect size (r)	Transformed ES (Z_r)	Standard error (SE)	Published (1 = yes)	v^*	ES^*	z	1/SE
228	.525	.583	.525	0	.005	2.87	7.57	14.43
491	.198	.201	.198	1	.003	-3.28	3.42	17.26
65	.311	.322	.311	1	.017	-0.49	2.35	7.55
458	.554	.624	.554	0	.002	5.06	11.45	20.67
929	.161	.162	.161	1	.001	-6.90	4.65	28.91
904	.336	.349	.336	1	.001	-1.14	9.69	28.84
74	.396	.419	.396	0	.015	0.27	3.24	8.18
151	.617	.721	.617	0	.007	3.94	7.22	11.69
150	.557	.628	.557	0	.007	2.89	6.59	11.83
590	.575	.655	.575	1	.002	6.57	13.50	23.47
180	.039	.039	.039	1	.007	-4.22	0.47	12.03
139	.358	.375	.358	0	.008	-0.13	4.02	11.21
132	.049	.049	.049	1	.008	-3.69	0.54	10.85
60	.000	.000	.000	1	.019	-2.80	0.00	7.21
839	.326	.339	.326	1	.001	-1.32	8.57	26.26
262	-.048	-.048	-.048	1	.004	-6.89	-0.75	15.59
266	.454	.489	.454	1	.004	1.65	7.13	15.71
209	.253	.258	.253	1	.005	-1.82	3.54	13.98
314	.160	.162	.160	1	.003	-3.90	2.73	16.99
881	.477	.519	.477	0	.001	3.99	13.57	28.47
517	.469	.509	.469	1	.002	2.78	10.30	21.96
228	.572	.651	.572	1	.005	3.82	8.15	14.24

بادخال البيانات برنامج SPSS او اى برنامج احصائى وبحساب معامل ارتباط الرتب لكندل بين احجام التأثير ES_i^* و تباين الاخطاء المعيارية v_i^* اتضح ان قيمته -0.07 ($P=0.67$) ويتضح عدم دلالتها احصائياً.

وبحساب مدخل Egger للانحدار من خلال انحدار قيم Z (العمود الثامن) فى الجدول قيم على قيم العمود التاسع $\frac{1}{SE}$ يتضح ان $p=0.62$ لمعامل الانحدار غير دالة احصائياً وعلى ذلك فان نتائج معامل ارتباط كندل والانحدار اشارت الى عدم وجود تحيز للنشر بالتالى فان شكل القمع منتظم.

طرق تصحيح تحيز الاتاحية او النشر

منذ منتصف 1980 حتى الان طور الاحصائيون طرق احصائية لتصحيح نتائج ما وراء التحليل من تاثيرات تحيز النشر ولكن هذه الطرق معقدة احصائياً وهذه الطرق قائمة على قيم P الاحتمالية للدراسات الاولية حيث تفترض ان احتمالية النشر تعتمد على قيمة P للدراسة وهذا ليس شرط فقط انما توجد خصائص اخرى للدراسات تؤثر فى امكانية النشر مثل الجودة المنهجية، سمعة المؤلفين، حجم التأثير، وغيرها وهذا يمكن ان يجعل هذه الطرق غير واقعية (Duval & Tweedie, 2000). وهذه الطرق تفترض عدة مسلمات ربما او ربما لا تتحقق فى البيانات الحقيقية (Hedges, 1992).

وعرض (Hunter & Schmidt (2004) مراجعة وعرض ملخص لهذه الطرق كالاتى:

الطريقة الاصلية لـ Hedges- Olkin (1985)

وتفترض هذه الطريقة مسلمة ان كل الدراسات الدالة احصائياً تكون مطبوعة بغض النظر عن اتجاه الدلالة والدراسات غير الدالة احصائياً غير مطبوعة، وان الدراسات الدالة تاخذ وزن 1 بينما الدراسات غير الدالة تاخذ وزن 0.0. حيث لو كان الفرض الصفرى حقيقى فان عدد الدراسات موجبة التأثير يكون مساوى لعدد الدراسات سالبة التأثير من ثم يكون

متوسط حجم التأثير (r, d) يكون صفر بالتالى لا يوجد تأثير لتحيز النشر فى متوسط حجم التأثير. وقام Hedges & Olkin(1985) باستخدام هذه الطريقة مع نموذج التأثيرات المثبتة وهذا نادراً ما يناسب البيانات البحثية الواقعية وهذه الطريقة تتكون من عمليات حسابية عديدة للمزيد يمكن الرجوع الى Hedges & Olkin(1985)

طريقة Iyengar-Greenhouse (1988)

وسع Hedges & Olkin (1988) الطريقة الاصلية لـ Hedges & Olkin (1985) حيث اعدو صياغة مسلمة هى ان كل الدراسات الدالة احصائياً تكون مطبوعة بغض النظر عن اتجاه الدلالة ولكن افترضوا ان كل الدراسات غير الدالة احصائياً لها احتمالات مختلفة من النشر مع قيم P مختلفة.

طريقة Begg-Mazumdar (1994)

افترض Begg & Mazumdar (1994) ان الدراسات ذات احوام العينات الصغيرة يمكن نشرها لو اعطت حجم تأثير d كبير الحجم، بينما الدراسات ذات احوام العينات الكبيرة الحجم تطبع بغض النظر عن حجم r, d لان هذه النتائج تكون دالة احصائياً وهذه المسلمة تكون وراء فلسفة شكل القمع حيث الدراسات الغائبة تكون صغيرة حجم التأثير. واقترحوا معامل ارتباط الرتب بين حجم التأثير والخطأ المعياري فالقيمة 0.0 تدل على عدم وجود تحيز للنشر بينما القيمة الموجبة تدل على تحيز النشر وهذا الاختبار هو تكميم لشكل القمع تقادياً للتفسير الذاتى لشكل القمع.

العمل اللاحق لـ Hedges وزملائه

عمم Hedges(1992) طريقة Iyengar & Greenhouse بتضمينه لاوازن الدراسات التى تعكس احتمالات النشر القائمة على نواتج البحث فى ضوء قيمة P وفى هذا التعديل يمكن استخدامه لنموذج التأثيرات العشوائية(للمزيد انظر: Hedges,1992).

طريقة التهذيب والملئ لـ Duval-Tweedie (2000) Trim-Fill

الطرق السابقة معقدة رياضياً ونادراً ما تستخدم في الممارسات العملية لذلك اقترح Duval & Tweedie (2000) طريقة بسيطة ولابارامترية قائمة على خصائص توزيع Wilcoxon وهي تقدر عدد الدراسات الغائبة التي من المفترض انها تقع على الناحية اليسرى من شكل القمع وعرضوا ثلاثة مقدرات لابارامترية لعدد الدراسات الغائبة ويعطوا تقديرات متشابهة، وفي هذه الطريقة يتم ملئ الجانب الايسر من الدراسات الغائبة لخلق مرآة متكافئة من الجانب الايمن من المنحنى وهذه الطريقة تصحح نتائج ما وراء التحليل من تاثيرات تحيز النشر، وهي ليست طريقة لتحديد ما اذا كان تحيز النشر موجود من عدمه. واذا لم يوجد تحيز نشر فان عدد الدراسات يكون صفراً وهي تصلح للنموذج العشوائي وللنموذج المثبت. و يمكن اعتبار هذا الاختبار كتحليل للحساسية Sensitivity analysis لدراسة تاثير الدراسات الغائبة على النتائج الملاحظة واسهل حسابياً من الطرق السابقة.

الفصل الثانى عشر

مداخل اجراء وراء التحليل

مقدمة

يوجد العديد من المداخل لاجراء ما وراء التحليل وكل طريقة قائمة على فلسفة ونظرية تحدد طبيعة دمج احجام التأثير وفهم النتائج. ويرى (Hunter & Schmidt 2004) ان النظرية الكاملة للبيانات تتضمن فهم خطأ المعاينة، خطأ القياس، المدى الضيق، التصنيفية الثنائية وتأثيرها، اخطأ البيانات، وعوامل طارئة او صدفية اخرى تشوه البيانات الخام التى نراها فى الدراسات البحثية، والفهم النظرى يتطلب دراسة كيفية تاثير هذه العوامل على البيانات وكيفية تصحيح تأثيراتها.

استخدام ما وراء التحليل فى ظل عدم الفهم للبيانات المراد تحليلها وعدم تصحيحها من المحددات المنهجية المختلفة يؤدى الى نتائج متحيزة. ويرى Hunter & Schmidt (2004) ان كثير من طرق ما وراء التحليل لا تصحح البيانات من هذه المحددات خاصة اخطأ القياس.

ومعظم طرق ما وراء التحليل تتعامل مع قضية خطأ المعاينة وتتجاهل خطأ القياس على اعتبار ان القياسات تامة القياس وهذا بعيد المنال فى الدراسات البحثية خاصة فى العلوم السلوكية والاجتماعية والعكس صحيح عندما تعاملت مجلدات القياس مع خطأ القياس تجاهلت خطأ المعاينة وافترضت ان العينات احجامها كبيرة.

والقضية الاخرى الجوهرية فى النظرية الكاملة للبيانات وجود احتمالية كبيرة جداً ان مجتمعات احجام التأثير تختلف من دراسة الى اخرى، ولذلك فان مسلمة ان معالم الدراسات متماثلة عبر كل الدراسات فى ما وراء التحليل (مسلمة نموذج التأثيرات المثبتة) غير واقعية واحتمال كبير ان تكون خاطئة، ونظرية البيانات تدعم بقوة نموذج التأثيرات العشوائية وليس نموذج التأثيرات المثبتة.

طرق دمج احجام التأثير

تناولنا لما وراء التحليل كان محدد بتراكم احجام التأثير وليس الدلالة الاحصائية عبر الدراسات المختلفة، ولكن توجد اطر ومفاهيم ونظريات مختلفة لاجراء ما وراء التحليل باستخدام باحثين او مجموعات عمل عبر عقود من الزمان كالاتي (Hunter & Schmidt, 2004):

قام Thorndike (1933) بجمع مجموعة كبيرة لمعاملات الثبات باستخدام الاختبار واعادته لاختبار ستانفورد - بينيه لـ 36 دراسة وصحح التباين المقاس لمعاملات الارتباط من تأثير خطأ المعاينة، وتوصل الى ان ليس كل الاختلافات الملاحظة يمكن تفسيرها عن طريق خطأ المعاينة ولكن تفسر في ضوء الفترة الزمنية بين الاختبار واعادته.

قام Ghiselli (1955) بجمع معاملات الصدق لعدد كبير من الدراسات لانواع مختلفة من الاختبارات القبول للوظائف المختلفة والاداء فيها وعرض النتائج في ضوء قيم الوسيط ولكنه لم يحلل تباينات معاملات الارتباط.

قام Rosenthal (1961) باجراء حساب متوسط عدد كبير من معاملات الارتباط لعدد كبير من الدراسات التي تناولت العلاقة بين متغيرين.

قام Bloom (1964) بحساب متوسط معاملات الارتباط ليلخص نتائج عدد كبير من الدراسات التي تناولت العلاقة بين السمات الانسانية والقدرات.

هؤلاء الباحثين وغيرهم لم يكونوا ملمين بمنهجية ما وراء التحليل كطريقة لحل مشكلة تكامل نواتج الدراسات لاعطاء معرفة تراكمية حتى عام 1976 عندما قدم Glass (1976) اولى الاجراءات المنهجية السليمة لاجراء ما وراء التحليل ثم قام Schmidt & Hunter (1977) بطباعة اول مقال لهم في منهجية ما وراء التحليل. وأشار Glass الى ان تسميته هي اسلوب تحليل التحليلات Analysis of analyses ، واحد الاسباب لتسميته بهذه التسمية ليفرق بينه وبين دراسات التحليل الثانوي، وهو تراكم كمي وتحليل لاحجام التأثير ولاحصائيات وصفية اخرى عبر الدراسات ولا يتطلب البيانات الاصلية للدراسات.

وتقع طرق او مداخل ما وراء التحليل فى ثلاثة تصنيفات عريضة كالتى:

طرق ما وراء التحليل الوصفية الخالصة (مدخل Glass)

يعطى صورة وصفية للحال الذى عليه التراث البحثى ولكنه لا يحلل او يصحح احجام التأثير من المحددات او النواقص المنهجية التى تشوه نتائج الدراسات. والسؤال الرئيسى هل العلاج النفسى بغض النظر عن نوعه له اثر على انواع معينة من النواتج المختلفة حيث تدمج الدراسات ذات المتغيرات المستقلة المختلفة (انواع مختلفة من العلاج) على متغيرات تابعة مختلفة، ونتيجة لذلك ظهرت مقولة الدمج بين التفاح والبرتقال. ويمتلك مدخل Glass ثلاثة خصائص اولية:

1. التاكيد القوى على احجام التأثير وليس الدلالة الاحصائية واعتقد ان الهدف من عملية التوليف للدراسات وصفى وليس استدلالى حيث المظهر الهام للاحصاء الوصفية هو تحديد مقدار التأثيرات ووظف تقديرات معامل ارتباط بيرسون و تقديرات d كمؤشرات لحجم التأثير، ومن اهم نواتج هذا المدخل هى متوسط والانحراف المعيارى لاحجام التأثير المقاسة عبر الدراسات.

2. اعتبار ان تباين احجام التأثير كقيمة ظاهرية سطحية وافترض هذا المدخل ان التباين الملاحظ فى احجام التأثير حقيقى وواقعي ولا بد من تفسيرها ولم يهتم بتباين خطأ المعاينة فى احجام التأثير وان التفسير للاختلاف بين احجام التأثير يرجع الى خصائص الدراسات مثل جنس او عمر العينة، طول المعالجة، وغيرها.

3. مدخل امبريقى قوى لتحديد اى من مظاهر الدراسات يجب تكويدها واختبارها واعتبارها متغيرات متفاعلة.

ومن اهم الانتقادات لهذا المدخل:

1. تعتبر حجم التأثير وحدة التحليل فالدراسات القائمة على عينة بحث وحيدة تعطى تقديرات عديدة لاحجام التأثير ولكن Glass وزملائه ضمن كل احجام التأثير فى الدراسة الواحدة فى التحليل، وهذا يخل بمسئمة الاستقلالية الاحصائية وله انعكاس على مصداقية

الاختبارات الاحصائية الاستدلالية المطبقة فى ما وراء التحليل مثل اختبار الدلالة الاحصائية حول متوسط حجم التأثير وهذا الانتقاد فى محلة، ولكن الحقيقة ان Glass لم يكن هدفه من توليف الدراسات فى المقام الاول استدلالى انما وصفى، ولكن تحطم الاستقلالية له تأثير متحفظ على نواتج حجم التأثير ويمكن ان يؤدى الى تقديرات عالية لتباين احجام التأثير ولكن ليس لها تأثير على متوسط حجم التأثير.

2. دمج كل الدراسات فى ما وراء التحليل بغض النظر عن الجودة المنهجية للدراسات.

3. دمج دراسات ذات متغيرات مستقلة متنوعة او برامج علاجية متنوعة فى تحليل واحد ولم يتناول الفروق بين متوسطات النواتج المختلفة للمتغيرات المستقلة المختلفة، فمثلاً تناول (Smith & Glass 1977) عشرة انواع مختلفة من البرامج العلاجية، ولكن هذا الانتقاد تجاهل حقيقة ان طريقة Glass تتضمن اجراء تحليلات منفصلة لما وراء التحليل لكل متغير منفصل كخطوة لاحقة ولكن هذا يتوقف على الهدف من الدراسة، هل الفعالية النسبية لكل برنامج على حدة ام الفعالية للبرامج ككل.

4. خلط مقاييس مختلفة لمتغيرات مستقلة متنوعة فمثلاً فى دراسات التدخلات النفسية التربوية فان قيم d تكون لمقاييس مختلفة مثل الاتجاهات، المعتقدات، السلوكيات، والتحصيل الاكاديمى حيث تضمن فى نفس دراسة ما وراء التحليل. وهذا النقد فى محله لانه من الصعب تفسير النتائج.

ويرى (Hunter & Schmidt 2004) ان هذه الانتقادات ليست مشتقة من الطبيعة الطرق الاحصائية لـ Glass انما فى طريقة تطبيق هذه الطرق بوسطة Glass وزملائه وانهم لديهم رؤية مختلفة عن الهدف من وراء التحليل وهو رسم صورة عامة وشاملة عن التراث البحثى.

طرق ما وراء التحليل لتاثيرات الدراسة كاستجابة للانتقادات السابقة

حاول (Bangert & Drowns 1986) تطوير طريقة اعتماداً على طريقة Glass وتختلف عنها فى طرق عديدة كالاتى:

الأول: تضمين حجم تأثير واحد فقط لكل دراسة وهذا يساعد على توافر مسلمة الاستقلالية الاحصائية في ما وراء التحليل، فلو ان الدراسة تضمنت قياسات عديدة لمتغيرات تابعة فيتم التركيز على مفهوم واحد فقط في الدراسة.

الثاني: هذا الاجراء يجعل الصورة واضحة امام الباحث للحكم على بعض مظاهر جودة الدراسة ويستبعد الدراسات ضعيفة الجودة المنهجية.

كما انه اعتمد على تصحيح احجام التأثير من خطأ المعاينة فقط واعتقد ان الهدف من ما وراء التحليل وصف نتائج الدراسات الملاحظة ويجب عدم تصحيحها من اى مشاكل بحثية اخرى خاصة بالتصميم. واعتمد على متوسط حجم التأثير الموزون.

طرق ما وراء التحليل التي تركز على صحيح احجام الاثير من خطأ المعاينة فقط: طرق

Hedges وطرق Rosenthal وطرق Bar-Bones

كثيراً من المحددات تؤدي الى الاختلاف بين احجام التأثير عبر الدراسات ومن هذه المحددات تباين خطأ المعاينة ويقبل Glass وطريقة Bangert & Drowns الاختلاف الناتج عن تباين خطأ المعاينة كاختلاف حقيقي ومسلم به ولكن مدخلين من ما وراء التحليل تذهب ابعد من طرق Glass وتحاول ضبط تباين خطأ المعاينة هما:

ما وراء التحليل القائم على اختبار التجانس: طرق Hedges's وطرق Rosenthal's

هذه الطرق قائمة على تجانس احجام التأثير المتضمنة في الدراسة وهذا المدخل تبناه مجموعة من الباحثين امثال Hedges & Olkin (1985) و Rosenthal & Rubin (1982) افترضوا ان اختبار كاي تربيع لتحديد ما اذا كانت نواتج الدراسة اكثر اختلافاً مما هو متوقع نتيجة خطأ المعاينة فقط واذا كانت قيمة كاي تربيع غير دالة احصائياً فان احجام التأثير او معاملات الارتباط في المجتمع ثابت عبر كل الدراسات وليسنا بحاجة الى متغيرات متفاعلة لتفسير الاختلافات، واستخدام اختبار كاي تربيع للتجانس لتقدير ما اذا كانت النواتج لمجموعة من الدراسات تختلف اكثر مما هو متوقع من تباين خطأ المعاينة.

ولكن اختبار كاي تربيع للتجانس يمتلك قوة احصائية منخفضة للكشف عن الاختلاف بين الدراسات التي لا ترجع الى خطأ المعاينة (Hedges & Piggot, 2001) ، وعلى ذلك يستنتج الباحث ان الدراسات متجانسة ولكنها حقيقة غير متجانسة ويستخدم الباحث نموذج التأثيرات المثبتة فى كل الاحوال ويفترض ان الاختلاف بين الدراسات يساوى صفراً وهذا يؤدي الى تقديرات منخفضة للاخطاء المعيارية المرتبطة بالمتوسط وبدوره يؤدي الى فترات ثقة ضيقة حول متوسط حجم التأثير وهذا يؤدي الى تضخم الخطأ من النوع الاول، وعلى ذلك يوصى باستخدام نموذج التأثيرات العشوائية بدلاً من نموذج التأثيرات المثبتة .

وعلى ذلك فان معظم دراسات ما وراء التحليل المنشورة فى مجلة Psychological Bulletin اعتمدت على نموذج التأثيرات المثبتة لـ (Hedges & Olkin, 1985) على الرغم انهما عرضوا نموذج التأثيرات العشوائية. واعتمد (Hedges & Olkin, 1985) على التحليل المتفاعل المصاحب لنموذج التأثيرات المثبتة فى حالة وجود دلالة احصائية لاختبار الاختلاف.

ما وراء التحليل لـ Bare-Bones

المدخل الثانى لاجراء ما وراء التحليل هو محاولة ضبط فقط خطأ المعاينة ويطبق لاحجام التأثير منها قيم d لكوهين واى مؤشرات اخرى شرط ان يكون الخطأ العيارى لها معروف. فيقدر متوسط حجم التأثير ثم يقدر تباين معاملات الارتباطات ثم تباين خطأ المعاينة وطرحه من التباين المقاس لمعاملات الارتباط ولوكانت النتيجة صفراً فان قيمة r او d هى ملخص لكل الدراسات فى ما وراء التحليل. واذا لم تكن النتيجة صفراً فان الجذر التربيعى للتباين المتبقى هو مؤشراً للتباين المتبقى حول متوسط حجم التأثير بعد استبعاد تباين خطأ التباين.

ولان وجود محددات منهجية مثل خطأ القياس يجب تصحيحها فيمكن اعتبار طريقة Bare bones & غير كاملة وغير مرضية.

ما وراء التحليل السيكومتري: تصحيح من المحددات المنهجية المتعددة

النوع الثالث من ما وراء التحليل هو ما وراء التحليل السيكومتري وهذه الطرق لا تصحح احجام التأثير من خطأ المعاينة فقط (المحدد غير المنتظم) انما من المحددات المنهجية

الأخرى مثل أخطاء القياس وتقلص المدى، الثنائية التصنيفية، نقص الصدق، الثبات، وغيرها ويطلق على هذه المحددات منتظمة لأنها تتسبب في حدوث اختلافات بين الدراسات وتخلق تحيزات سلبية Downward منتظمة في نتائج الدراسات. فمثلاً أخطاء القياس تسبب تحيزات سلبية منتظمة لكل معاملات الارتباط.

وطرق ما وراء التحليل السيكمترية لاتصحح أحجام التأثير من الاختلافات نتيجة المحددات المنهجية فقط عبر الدراسات إنما من التحيزات السلبية ويأخذ في حسابه المحددات المنهجية والاحصائية ومن أهم من تبني هذا المدخل (Hunter & Schmidt, 1990, 2004) ولم يحولوا معامل الارتباط إلى تحويل فيشر لأنه يعطى تقدير متحيز موجب لمتوسط معامل الارتباط وأقل دقة من معامل الارتباط غير المحور (Hall & Brannick, 2002). وارى من الطرق السابقة تستند إلى نظرية البيانات وهذه النظرية تحدد طرق ما وراء التحليل لتحليل البيانات، ولكن النظرية الشاملة المتكاملة تتضمن فهم خطأ المعاينة، خطأ القياس، تحيز المعاينة، التصنيفية المصطنعة، أخطاء البيانات، وغيرها .

وبلغة سيكمترية فإن نظرية البيانات تتضمن أولاً عملية كيفية تأثير هذه العوامل على البيانات وتسمى نموذج التقلص Attenuation ، وثانياً عملية تصحيح البيانات من هذه العوامل ويسمى نموذج عدم التقلص Disattenuation. ومن أهم العوامل التي يجب أن تؤخذ في الاعتبار في ما وراء التحليل هي خطأ المعاينة وخطأ القياس لأن الدراسات تتضمن عينات صغيرة الحجم وغير ممثلة للمجتمع تمثيلاً دقيقاً، وخطأ القياس لا مقاييسنا ليست تامة الثبات.

عرض مثال تطبيقي (في: Ellis, 2010)

اهتم باحث بإجراء ما وراء التحليل للعلاقة بين ادمان الانترنت والتحصيل وجمع ثلاثة دراسات والبيانات كالاتي:

الجدول (1.12): بيانات ثلاثة دراسات للعلاقة بين ادمان الفيس بوك والتحصيل.

الباحث	r	P	N	ثبات المتغير التابع
A (2009)	-0.48	0.02	80	0.70
B (2014)	-0.58	0.001	112	0.92
C (2017)	0.05	0.33	32	-

• r معامل ارتباط بيرسون.

• p القيمة الاحتمالية للخطأ من النوع الاول.

• N حجم العينة.

Hunter & Schmidt (1990, 2004) إجراء ما وراء التحليل في ضوء مدخل

حساب متوسط حجم التأثير الموزون

الطريقة الاسهل والاسواء لحساب متوسط حجم التأثير كالاتى:

$$M_{ES}(\bar{r}) = \frac{-0.48 \pm 0.58 + 0.05}{3} = -0.337$$

لاحظ ان هذا المتوسط اصغر من حجم التأثير او معامل الارتباط للدراستين A, B وهذا يعنى

ان المتوسط متحيز ويتضح ان حجم التأثير للدراسة C مختلف عن الدرستين A, B.

ولكن المدخل الاكثر دقة ومنطقية وهو اعطاء وزن او اهمية للدراسة بما يتناسب مع حجم العينة

فمن غير المنطقى ان دراسة تتضمن عينة 112 ان تتساوى مع دراسة عينتها 32 وعلى ذلك

فلا بد من حساب متوسط حجم التأثير الموزون من خلال ضرب كل حجم تأثير في حجم عينها

مقسوماً على حجم العينة الكلى كالاتى:

$$\bar{r} = \frac{\sum n_i r_i}{\sum n_i}$$

وعليه فان متوسط حجم التأثير الموزون كالتالى:

$$= \frac{(80 \times -0.48) + (112 \times -0.58) + (32 \times 0.05)}{80 + 112 + 32}$$

$$\bar{r} = \frac{(-38.4) + (-65.0) + (1.6)}{224} = -0.454$$

الملاحظ ان المتوسط الموزون -0.454 اكبر في قيمته المطلقة من المتوسط غير الموزون - 0.337 وهو قريب من قيمة حجمى التأثير للدراستين B,A بمعنى ان المتوسط الموزون اكثر دقة في التعبير عن حجم التأثير الفعلى للدراسات.

تصحيح حجم التأثير من اخطاء القياس او الثبات المنخفض

للحصول على احجام تأثير اكثر دقة يجب تصحيحها من عدم الثبات للمقاييس فالمقياس المستخدم فى قياس المتغير التابع فى الدراسة A اقل من ثبات المقياس المستخدم فى الدراسة B وهذا يودى الى انكماش حجم التأثير عن قيمته الحقيقية نتيجة اخطاء القياس ولتصحيح حجم التأثير من خطأ القياس من خلال المعادلة الاتية:

$$ES_{adjusted} = \frac{r}{\sqrt{\alpha}}$$

• α المعامل الفا كرونباخ.

وعليه فان حجم التأثير المصحح للدراسة A كالتالى:

$$ES_A(r)_{adjusted} = \frac{-0.48}{\sqrt{0.70}} = -0.574$$

وحجم التأثير المصحح للدراسة B:

$$ES_B(r)_{adjusted} = \frac{-0.58}{\sqrt{0.92}} = -0.605$$

اما الدراسة الثالثة لم يذكر الباحث معامل الثبات الفا كرونباخ ولذلك يقدر على متوسط

الثبات فى الدراستين B, A:

$$= \frac{0.70 + 0.92}{2} = 0.81$$

وحجم التأثير المصحح للدراسة الثالثة C كالتالى:

$$ES_C(r)_{adjusted} = \frac{0.05}{\sqrt{0.81}} = 0.056$$

ويمكن اجراء التصحيح فى ضوء عدم الثبات المتغيرين فى الدراسة كالاتى:

$$ES(r)_{adjusted} = \frac{r}{\sqrt{\alpha_1 \alpha_2}}$$

وعليه فان متوسط حجم التأثير الموزون المصحح من خطأ القياس او نقص الثبات:

$$\begin{aligned} &= \frac{(80 \times -0.574) + (112 \times -0.605) + (32 \times 0.056)}{80 + 112 + 32} \\ &= \frac{(-45.92) + (-67.76) + (1.79)}{224} \\ &= -0.50 \end{aligned}$$

لاحظ ان متوسط حجم التأثير الموزون تغيرت قيمة من 0.337 الى 0.454 نتيجة تصحيحة من خطأ المعاينة والى 0.50 نتيجة تصحيحة من خطأ المعاينة وخطأ القياس، وبالتالي نستطيع ان ننق فى النتيجة الثالثة لانها اقل تحيز.

حساب الدلالة الاحصائية لمتوسط حجم التأثير

لحساب الدلالة الاحصائية لمتوسط حجم التأثير (متوسط معامل الارتباط) يوجد مدخلين هما:

المدخل الاول: تحويل معامل الارتباط الى تحويل فيشر Z_r وتحديد الدلالة الاحصائية ولكن (Hunter & Schmidt (1990, 2004) اجرؤا ما وراء الحليل على قيم معامل الارتباط مباشرة دون تحويله الى تحويل فيشر وفى هذه الحالة لايد من حساب الخطأ المعيارى المرتبطة بمتوسط حجم التأثير والخطأ المعيارى لوصف انتشار التوزيع العينى لاحجام التأثير فى المجتمع بمعنى انه نوع خاص من الانحراف المعيارى وفى هذا المثال فان التوزيع العينى يتكون من ثلاثة احجام تأثير، والبعض يفضل مصطلح التباين كتعبير عن الخطأ المعيارى ولكن المصطلح الاكثر استخداماً هو الجذر التربيعى للتباين. وتباين عينة معاملات الارتباطات v_r تقدر من مجموع حاصل ضرب مربع الفرق بين كل حجم التأثير للدراسة ومتوسط حجم التأثير مضروباً فى حجم العينة ثم يقسم الناتج على حجم العينة الكلى كالاتى (Ellis(2010):

$$v_r = \frac{\sum n_i (r_i - \bar{r})^2}{\sum n_i}$$

$$= \frac{(80 \times (-.574 - -.50)^2 + (112 \times (-.605 - -.500)^2 + (32 \times (.056 - -.500)^2)}{80 + 112 + 32}$$

$$v_r = \frac{0.400 + 1.232 + 9.888}{224} = 0.51$$

ويتم تقدير الخطأ المعياري كالآتي:

$$SE_{\bar{r}} = \sqrt{\frac{v_r}{K}} = \sqrt{\frac{0.051}{3}} = 0.130$$

وتقدر الدلالة الاحصائية في ضوء اختبار تحويل معامل الارتباط الى الدرجة المعيارية Z كالآتي:

$$Z = \frac{\bar{r}}{SE_{\bar{r}}}$$

$$= \frac{0.500}{0.130} = 3.85$$

وقيمة Z الجدولية لمستوى دلالة احصائية 0.05 لاختبار ذو ذيلين 1.96 ولمستوى 0.01 هي 2.58 ، وبما ان قيمة Z المحسوبة اكبر من قيمتها الجدولية، وعلده يرفض الفرض الصفرى بالتالى فان النتائج دالة احصائياً.

المدخل الثانى: تقدير فترات الثقة: تقدر 95% فترات الثقة من المعادلات الآتية:

$$CI_{95\% \text{ lower}} = \bar{r} - Z_{\alpha/2} SE_{\bar{r}} = -.500 - (1.96 \times .13) = -.755$$

$$CI_{95 \text{ upper}} = \bar{r} + Z_{\alpha/2} SE_{\bar{r}} = -.500 + (1.96 \times .13) = -.245$$

وبما ان قيمة المعلم في الفرض الصفرى هي صفر، وبما ان القيمة صفر لا تقع بين حدى الثقة الادنى والاعلى، وعليه فان النتائج دالة احصائياً.

زيادة الخطأ المعياري يجعل فترات الثقة اكثر اتساعاً مما يجب ان تكون عليه.

اختبار الاختلاف او عدم التجانس بين توزيع تقديرات حجم التأثير

اتساع فترات الثقة يشير الى احتمال ان توزيع احجام التأثير غير متجانسة وهذا يعنى ان احجام التأثير لا تتمركز حول متوسط وحيد لاحجام التأثير فى المجتمع وتتنوع حول متوسطات عديدة لاحجام التأثير وتقدر اختبار عدم التجانس من خلال اختبار Q كالاتى:

$$\begin{aligned} Q &= \sum (n_i - 1) (r_i - \bar{r})^2 \\ &= ((80 - 1) \times (-.574 - -.50)^2) + ((112 - 1) \times (-.650 - -.500)^2) \\ &\quad + ((32 - 1) \times (.056 - -.500)^2) \\ &= (79 \times .005) + (111 \times .011) + (31 \times .309) \\ &= 0.395 + 1.221 + 9.579 \\ &= 11.195 \end{aligned}$$

وتختبر الدلالة الاحصائية من خلال اختبار كاي تربيع بدرجات حرية K-1 حيث K عدد احجام التأثير، وبما ان قيمة اختبار كاي تربيع لدرجات حرية 2 ومستوى دلالة احصائية 0.05 هي 5.991 ، بالتالى فان قيمة Q المحسوبة اكبر من 5.991 ، وعليه نرفض الفرض الصفرى بوجود اختلاف بين احجام التأثير، وهذا يتطلب اما اجراء تحليلات المتفاعلة لمعرفة اى من خصائص الدراسات التى احدثت هذه الدلالة او تحليل ما وراء التحليل فى ضوء نموذج التأثيرات العشوائية. ولكن لبياننا هذا المثال لا يمكن اجراء تحليل التفاعل نظراً لقلّة الدراسات وعدم اتاحة الفرصة لعمل مجموعات فرعية من الدراسات او احجام التأثير.

تفسير النتائج

يتضمن التفسير الاجابة عن الاسئلة الاتية: ما دلالة قيمة متوسط حجم التأثير فى ضوء معايير كوهين؟، ماهى الفئة المستفيدة من هذه النتيجة وكيف تستفيد منها؟، ماذا اضافت هذه النتيجة الى التراث البحثى؟، هل النتائج دعمت او تناقضت ما هو موجود فى التراث البحثى؟ ماهى محددات الدراسة ان وجدت؟.

اجراء ما وراء التحليل لتوليف r باستخدام مدخل Hedges وزملائه

(Hedges, 1981,1992; Hedges & Olkin,1985; Hedges & Vevea, 1998)

نموذج التأثيرات المثبتة : باستخدام مدخل Hedges et al لتوليف احجام التأثير يتم فى ضوء

نموذج التأثيرات المثبتة كالاتى:

الاول: يتم تحويل احجام التأثير الخام r الى تحويل فيشر Z_r كالاتى:

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

ويتم تقدير تباين الخطأ المعياري كالاتى:

$$v_i = \frac{1}{n_i - 3}$$

ووزن الدراسة w_i مقلوب تباين الخطأ المعياري كالاتى:

$$w_i = n_i - 3$$

الجدول (2.12): بيانات ما وراء التحليل فى ضوء البيانات الخام ونموذج التأثيرات المثبت والتاثيرات العشوائية.

Raw study data			Fixed-effects sums			Random-effects sums		
n	r	z	w	wz	wz^2	w^2	w^*	w^*z
80	-.48	-.523	77	-40.27	21.06	5,929	11.35	-5.94
112	-.58	-.662	109	-72.21	47.84	11,881	11.86	-7.86
32	.05	.050	29	1.45	0.07	841	9.12	0.46
			215	-111.03	68.97	18,651	32.34	-13.34

لحساب متوسط حجم التأثير الموزون لـ Z_r كالاتى:

$$\begin{aligned} \bar{z}_r &= \frac{\sum w_i z_i}{\sum w_i} = \frac{\sum (-0.48 \times 77) + (-0.58 \times 109) + (0.05 \times 29)}{77 + 109 + 29} \\ &= \frac{-111.03}{215} = -0.52 \end{aligned}$$

وتباين المتوسط الموزون:

$$v = \frac{1}{215} = 0.005$$

والخطأ المعياري للمتوسط كالاتي:

$$SE_{\bar{z}} = \sqrt{0.005} = 0.071$$

وحساب اختبار Z (ليس تحويل فيشر) للدلالة الاحصائية كالاتي:

$$Z = \frac{\bar{z}_r}{SE_{\bar{z}}} = \frac{0.52}{0.071} = 7.32$$

ووفترات الثقة كالاتي:

$$CI_{95\% \text{ lower}} = \bar{z} - Z_{\alpha/2} SE_{\bar{z}} = -0.52 - (1.96 \times 0.071) = -0.66$$

$$CI_{95\% \text{ upper}} = \bar{z} + Z_{\alpha/2} SE_{\bar{z}} = -0.52 + (1.96 \times 0.071) = -0.38$$

ويتم رفض الفرض الصفري بمعنى وجود علاقة ارتباطية سالبة ودالة احصائياً بين التحصيل وادمان الانترنت.

واختبار الاختلاف بين احجام التأثير كالاتي:

$$Q = \sum w_i (z_i - \bar{z})^2 = wz^2 - \frac{(wz)^2}{w}$$
$$= 68.97 - \frac{(-111.03)^2}{215} = 11.63$$

يتضح وجود دلالة احصائية بمعنى اختلاف بين احجام التأثير وعلى ذلك فالمدخل الافضل لتحليل بيانات هو:

نموذج التأثيرات العشوائية

لاجراء نموذج التأثيرات العشوائية لابد من تقدير التباين داخل الدراسات v_i (نتجة خطأ المعاينة) والتباين بين الدراسات τ^2 وتقدر قيمة τ^2 كالآتى:

$$\tau^2(v_\theta) = \frac{Q - (K - 1)}{(\sum w_i) - \frac{(\sum w_i^2)}{(\sum w_i)}}$$

$$= \frac{11.63 - (3 - 1)}{215 - \frac{18651}{215}} = \frac{9.63}{128.25} = 0.075$$

حساب الاوزان او مقلوب تباين الخطأ المعيارى كالآتى:

$$w_i^* = \frac{1}{v_i^*}$$

حيث v_i^* كالآتى:

$$v_i^* = (v_i + 0.075)$$

$$= (0.005 + \text{خطأ المعاينة للعينات}) + 0.075$$

$$= 0.08$$

ويتم حساب المتوسط الموزون كالآتى:

$$\bar{z}_r^* = \frac{\sum w_i^* z_i}{\sum w_i^*} = \frac{-13.34}{32.34} = -0.41$$

وتباين المتوسط الموزون كالآتى:

$$v_i^* = \frac{1}{\sum w_i^*} = \frac{1}{32.34} = 0.31$$

والخطأ المعيارى للمتوسط الموزون لمعامل الارتباط المحور فى نموذج التأثيرات العشوائية:

$$SE_{\bar{z}_r^*} = \sqrt{0.031} = 0.176$$

اختبار Z للدلالة الاحصائية كالآتى:

$$Z = \frac{\bar{z}_r^*}{SE_{\bar{z}_r^*}} = \frac{0.41}{0.176} = 2.33$$

لاحظ ان الاختبار دال احصائياً عند 0.05 وليس دال احصائياً عند مستوى دلالة احصائية 0.01 وهذا يختلف عن نتائج النموذج المثبت.

وفترات الثقة:

$$CI_{95\% \text{ lower}} = \bar{z}_r^* - Z_{\alpha/2} SE_{\bar{z}_r^*} = -0.41 - (1.96 \times 0.176) = -0.66$$

$$CI_{95\% \text{ upper}} = \bar{z}_r^* + Z_{\alpha/2} SE_{\bar{z}_r^*} = -0.41 + (1.96 \times 0.176) = -0.38$$

ولتفسير النتائج يجب تحويل تحويل فيشر الى معامل الارتباط كالاتى:

$$\bar{r} = \frac{e^{2\bar{z}} - 1}{e^{2\bar{z}} + 1}$$

بالتعويض بقيمة \bar{z} فى النموذج المثبت كالاتى:

$$\bar{r} = \frac{e^{2\bar{z}} - 1}{e^{2\bar{z}} + 1} = \frac{e^{2 \times -0.52} - 1}{e^{2 \times -0.52} + 1} = -0.48$$

بالتعويض بقيمة \bar{z}_r^* فى النموذج العشوائى:

$$\bar{r} = \frac{e^{2\bar{z}} - 1}{e^{2\bar{z}} + 1} = \frac{e^{2 \times -0.41} - 1}{e^{2 \times -0.41} + 1} = -0.39$$

وعليه فان نتائج ما وراء التحليل فى ضوء معامل الارتباط كالاتى:

النموذج المثبت: $-0.48(CI_{95}: -0.57, -0.36)$

النموذج العشوائى: $-0.39(CI_{95}: -0.64, -0.07)$

تم تحليل نتائج الدراسات الثلاثة باستخدام مدخل (Hunter & Schmidt, 2004) ومدخل Hedges et al ويوجد فروق مختلفة بين المدخلين وفيما يلي معادلات المدخلين:

الجدول (3.12): المعادلات البديلة المستخدمة في المدخلين (Ellis, 2010).

Output	Generic	Hunter and Schmidt		Hedges et al.	
		d	r	d	r
Weighted mean ES	$\overline{ES} = \frac{\sum w_i ES_i}{\sum w_i}$	$\bar{d} = \frac{\sum w_i d_i}{\sum d_i}$	$\bar{r} = \frac{\sum n_i r_i}{\sum n_i}$	$\bar{d} = \frac{\sum w_i d_i}{\sum w_i}$	$\bar{z} = \frac{\sum w_i z_i}{\sum w_i}$
Variance of sample ESs	$v_s = \frac{1}{\sum w_i}$	$v_d = \frac{\sum w_i (d_i - \bar{d})^2}{\sum w_i}$	$v_r = \frac{\sum n_i (r_i - \bar{r})^2}{\sum n_i}$	$v_s = \frac{1}{\sum w_i}$	$v_z = \frac{1}{\sum w_i}$
z score	$z = \frac{ \overline{ES} }{SE_{\overline{ES}}}$	$z = \frac{ \bar{d} }{SE_{\bar{d}}}$	$z = \frac{ \bar{r} }{SE_{\bar{r}}}$	$z = \frac{ \bar{d} }{SE_{\bar{d}}}$	$z = \frac{ \bar{z} }{SE_{\bar{z}}}$
Homogeneity statistic	$Q = \sum w_i (ES_i - \overline{ES})^2$	-	$\chi^2_{k-1} = \sum \frac{(n_i - 1)(r_i - \bar{r})^2}{(1 - \bar{r}^2)^2}$	$Q = \sum w_i (d_i - \bar{d})^2$	$Q = \sum w_i (z_i - \bar{z})^2$

- ES حجم التأثير.
- \overline{ES} متوسط حجم التأثير (r, d, zr).
- K عدد الدراسات او احجام التأثير.
- Wi وزن الدراسة.
- SE الخطأ المعياري.

مصدر المعادلات: (Hedges & Vevea, 1998; Hunter & Schmidt, 2004; Lipsey & Wilson, 2001).

والجدول يعرض الفروق بين المدخلين في اربع معادلات اساسية هي متوسط حجم التأثير الموزون، التباين الخطأ المعياري، اختبار Z للدلالة، واختبار تقدير الاختلاف بين احجام التأثير. فالصيغة العامة او المعيارية Generic تم عرضها في العمود الاول اما الاعمدة الباقية توضح صيغ حسابات مداخل (Hedges & Vevea, 1998) و Hunter & Schmidt (2004). والفروق بين هذه المداخل كالتالي:

الاول: معامل الارتباط او حجم التأثير بين المدخلين حيث حول Hedges معاملات الارتباط الى تحويلات فيشر Z_r في حين اعتمد Hunter & Schmidt على معاملات الارتباطات مباشرة دون تحويل. فالارتباطات المحورة تصحح متوسط معامل الارتباط من التحيز السالب الصغير ولكن في نفس الوقت تسهم بقدر صغير من التحيز الموجب للنتائج. والاختيار بين تحليل معامل الارتباط -0.454 (خطأ المعاينة) او -0.50 (خطأ المعاينة وخطأ القياس) وتحليل تحويلات معامل الارتباط (المثبت -0.48 والعشوائي -0.39) وهذا تحيز سالب يعتمد على مدى تفضيل الباحث للتحيزات الموجبة والسالبة المحدودة للنتائج.

الثاني: التباين في البيانات الذي بدوره يؤثر على الاوزان التي تعطى للتقديرات وتستخدم في حساب تباين المعاينة والاختفاء المعيارية و فترات الثقة. فعند دمج معاملات الارتباط في ضوء مدخل H.S يكون قائم على حجم العينة لكل دراسة بينما الاوزان في ضوء مدخل Hedges في ضوء $N-3$ ، ولكن الشئ المفيد في مدخل H.S هو اجراء تصحيحات لاحجام التأثير في ضوء محددات الدراسة مثل نقص الثبات والصدق والتصنيفية الثنائية بينما في مدخل Hedges للنموذج العشوائي يدخل تصحيح اخر للتباين بين احجام التأثير نتيجة عينة الدراسات وهذا يفسر الاختلاف في النتائج كالاتي:

مدخل Hunter & Schmidt المصحح في ضوء خطأ المعاينة فقط:

$$-0.454(CI_{95}: - 0.693 \text{ to } - 0.215)$$

مدخل Hunter & Schmidt المصحح في ضوء تصحيح من نقص الثبات وخطأ المعاينة:

$$-0.500 (CI_{95}: - 0.755 \text{ to } - 0.245)$$

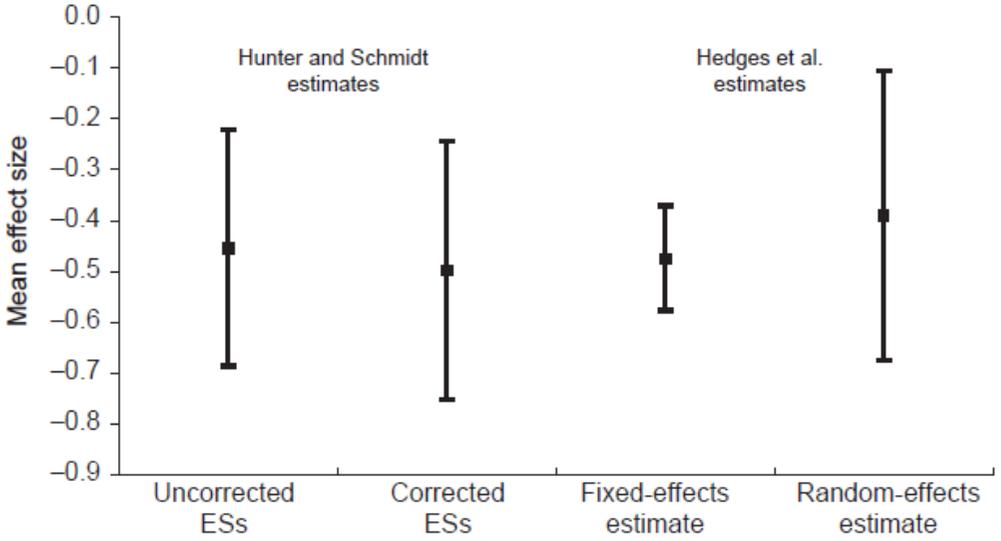
مدخل Hedges et al (التاثيرات المثبتة):

$$-0.478 (CI_{95}: - 0.572 \text{ to } - 0.365)$$

مدخل Hedges et al (التاثيرات العشوائية):

$$-0.391 (CI_{95}: - 0.639 \text{ to } - 0.068)$$

ويمكن عرض المتوسطات للمداخل الاربعة في الشكل الاتي(Ellis(2010):



الشكل (1.12): عرض متوسط احجام التأثير في المداخل الاربعة.

ويلاحظ الاختلاف الواضح بين المتوسطات في المداخل الاربعة وتساءل (2010) Ellis ماهى النتائج الاكثر دقة؟، هذا السؤال نال اهتمام كثير من الباحثين امثال (Hall & Field, 2005; Brannick, 2002; Schulze, 2004) ، ففي ضوء دراسة المحاكاة لـ (2005) Field توصل

الى ان طرق Hedges وزملائه تنتج تقديرات دقيقة لفترات الثقة بينما طريقة Hunter & Schmidt تعطى تقديرات دقيقة لتقديرات المتوسط ويعطى مدى ضيق لفترات الثقة عما هو يجب ان يكون بمعنى انه يستبعد حجم تأثير حقيقى من مدى حجم التأثير المقدر من خلاله وهذا يتفق

مع نتائج المثال السابق عرضه حيث اوسع مدى لحجم التأثير من خلال طريق Hedges باستخدام نموذج التأثيرات العشوائية حيث اكبر بنسبة 12% من فترات الثقة من خلال طريقتى Hunter & Schmidt. واوصى (2005) Field بان نضع معظم جهودنا فى تحليلات ما وراء

التحليل فى ضوء مدخل Hunter & Schmidt وفى نهاية هذه المناقشة نؤكد على ما اشار اليه (2004) Schulze بان بعض المداخل افضل من الاخرى لمهام مختلفة ولكن الطريقة الوحيدة الافضل هى مجموعة من الاجراءات يجب توليفها. وقام (2002) Hall & Brannick

بمقارنة نموذج التأثيرات العشوائية فى ضوء مدخلى (1998) Hedges & Vevea و(1990) Hunter & Schmidt خصوصاً فيما يخص بمدى جدوى تحويل معامل الارتباط الى احصاء فيشر التى اعتمد عليها Hedges، واستخدموا طريقة المحاكاة لمونت كارلو توصلوا

الى فروق طفيفة فى متوسط حجم التأثير وبعض الفروق الملاحظة فى فترات الثقة لصالح مدخل

Hunter & Schmidt

الفصل الثالث عشر

ما وراء التحليل باستخدام البرامج الاحصائية

Meta-analysis with Statistical Software

مقدمة

الحسابات التقليدية لما وراء التحليل التي تعتمد على حساب المتوسط الموزون و اختبار التجانس واختبار الدلالة الاحصائية قد لا تتطلب استخدام برامج كمبيوترية ولكن التطبيقات المتقدمة ودراسات ما وراء التحليل ذات احجام العينات الكبيرة تتطلب استخدام هذه البرامج لاعطاء نتائج اكثر دقة في اقل وقت ممكن. وتوجد العديد من البرامج الكمبيوترية العامة مثل SAS ، S-plus ، StatDirect ، StatXact ، و R غيرها لانجاز ما وراء التحليل ولكن هذه البرامج غالية الثمن. وتوجد العديد من البرامج القوية المتخصصة في دقة تقديراتها مثل MetaAnalyst ، RevMan ، MIX ، MetaXL ، CMA ، BUGS وغيرها.

اهم البرامج المستخدمة لاجراء المراجعة المنظمة للتراث البحثي

يوجد العديد من البرامج لاجراء ما وراء التحليل يمكن عرضها كالآتي (Littell et al. 2008 ; Schmid, Stewart, Rothstein, Lajeunesse, & Gurevitch, 2001)

برنامج RevMan

هو اختصار لـ Cochrane Collaborations Review Manager يتم تشغيله في نظام الويندوز وصمم هذا البرنامج لاجراء المراحل المختلفة من عملية المراجعة المنظمة من بناء البروتوكول، كتابة تقرير البحث، المراجع، والتحليلات الوصفية الاساسية. وما وراء التحليل ويتضمن صيغ للجداول المستخدمة لوصف الدراسات المستبعدة والمتضمنة في التحليل ومقارنات ونواتج متعددة على نفس البيانات ويقدر مؤشرات Odds ratio ونسب المخاطرة RR للبيانات التصنيفية الثنائية ومؤشر هيدجز المعدل g للنواتج المتصلة، ويمكن ان يطابق البيانات للنموذج التاثيرات المثبتة والعشوائية ويقدر اختبارات الاختلاف Q و I^2 ويعطى شكل Forest وشكل القمع وغيرها من الجداول والاشكال. ويمكن تحميله مجاناً من:

برنامج EPPI-Reviewer

تم تطويره عن طريق مركز EPPI وهي مؤسسة تجارية ويقوم بإجراء معظم مراحل المراجعة المنظمة للتراث البحثي ويسمح بنقل الاستشهاد من معظم قواعد بيانات العلوم الاجتماعية باستخدام الكلمات المفتاحية ويستخلص البيانات من الدراسات المتضمن في التحليل ويقدر ثبات المقدرين باستخدام معامل كابا لكوهين ويقدر مؤشرات OR و RR و RD لجداول الاقتران الثنائية ويقدر مؤشرات فروق المتوسطات المعيارية ويعطى احصائيات ما وراء التحليل للنموذج التاثيرات المثبتة ولنموذج التاثيرات العشوائية ويعطى شكل Forest Plot ويمكن الاطلاع عليه في الموقع:

<http://www.eppi.org.uk/>

برنامج TrialStat SRS

هذا البرنامج تجارى بمعنى غير متاح مجاناً ويتضمن قرارات لمدى صلاحية او تضمين الدراسات واستخلاص البيانات ويسمح بتحميل الاستشهادات والمستخلصات من برنامج ادارة المراجع ويسمح بعمل اشكال لمسح واستخلاص البيانات ويقدر ثبات المقدرين باستخدام المعامل كابا لكوهين ويمكن اخذ النتائج من هذا البرنامج الى برنامج Excel ويمكن الاطلاع عليه فى :

www.trialstat.com

اجراء ما وراء التحليل فى البرامج الاحصائية العامة

معظم البرامج الاحصائية العامة مثل SAS, SPSS, Stata, R, S-plus تقدر الاحصائيات الاساسية لما وراء التحليل عن طريق الباحث وبكتابة اوامر فرعية Macros يمكن اجراء العمليات الحسابية المعقدة والاشكال فى هذه البرامج.

فقام David Wilson بوضع اوامر فرعية لبرنامجى SAS, SPSS لتنفيذ التحليلات المختلفة لما وراء التحليل مثل تقدير متوسط حجم التاثير فى ضوء النموذج المثبت او

النموذج العشوائى وحساب معكوس تباين اوزان الدراسات واختبار Q للاختلاف وتحليلات المتفاعلات المشابهة لـ ANOVA وما وراء الانحدار وهذا التطبيق متاح فى كتاب Practical Meta-Analysis لـ Lipsey & Wilson(2001) وهذه الاوامر متاحة فى الموقع:

<http://mason.gmu.edu/dwilsonb/ma.html>.

برنامج Stata

كثيراً من الاوامر الفرعية المتاحة فى البرنامج تسمح باجراء ما وراء التحليل ومن السهل استخدامها وتسمح هذه الاوامر باجراء معكوس تباين الخطأ المعيارى واختبارات الاختلاف وشكل القمع واختبارات الدلالة الاحصائية لعدم انتظام شكل القمع وهذه الاوامر متاحة فى:

<http://www.stata.com/support/faqs/stat/meta.html>

ويمكن تحميلها من الموقع:

www.stata.com

البرامج التجارية لما وراء التحليل

برنامج Comprehensive Meta-Analysis(CMA)

يعتبر من البرامج الفريدة والمفضلة لاجراء ما وراء التحليل حيث صمم خصيصاً لهذ المجال ويجرى فى نظام التشغيل ويندوز ومن خلال الشاشة الافتتاحية يتم ادخال بيانات الدراسة باكثر من 100 صيغة، ويسمح باخذ البيانات من Excel ويتم ادخال المجموعات الفرعية المستقلة والنواتج وزمن اجراء الدراسة ويسمح بتحليل البيانات المختلفة سواء متصلة او تصنيفية ثنائية والارتباطات ويسمح باجراء التحليلات المتفاعلة والتحليل فى ضوء النموذج المثبت او النموذج العشوائى او النموذج المختلط واجراء ما وراء الانحدار وتحليل المجموعات الفرعية وتحليل الحساسية من خلال حذف دراسة ما واجراء التحليل على باقى الدراسات . ويعطى اختبارات تشخيص تحيز النشر والاختلاف بين احجام التأثير ويعطى رسومات واشكال عالية الجودة ولكنه لا يجرى ما وراء التحليل البيزانى، وهذا البرنامج مكلف

جداً فالنسخة الكاملة تكلف حوالى 1000 دولار لكل عام لكل نسخة. ويمكن تحميله للتدريب لمدة تصل الى عشرة ايام من الموقع:

www.meta-analysis.com

برنامج MetaWin

هو برنامج تجارى اخر وابتدعه باحثين فى مجال البيئه للبيانات البيئية ومن السهل استخدامة لاجراء ما وراء التحليل وهذا يعطى احجام تاثير مختلفة مثل مؤشر هيدجزع المعدل و OR و RR وتحوير فيشر لمعامل الارتباط ويعطى تحليل النموذج المثبت والنموذج العشوائى وتحليل الاختلاف وتحليل المجموعات الفرعية وما وراء الانحدار. ويسمح باجراء تحويلات مؤشرات حجم التاثير الى مؤشرات اخرى مثل مؤشر كوهين d الى r او العكس. وهذا البرنامج يتميز باعطاء طرق اعادة المعاينات (اجراءات Bootstrap) ولكن لا يجرى تحليل المجموعات الفرعية ويعطى الاشكال المختلفة مثل شكل القمع والمدرج التكرارى وغيرها وهذا البرنامج متاح فى:

<http://www.metawinsoft.com/>

برنامج MetaStat

يعمل فى بيئه الـ DOS ويسمح بمساعدة الباحث على تكويد الملامح المختلفة للدراسات واحجام التاثير ويعطى احصائيات التجانس ونسب التباين ويعرض البيانات فى اشكال مختلفة. ويمكن تحميله مجاناً من:

<http://ericae.net/meta/metastat.htm>.

برنامج EasyMA

هو برنامج مجاناً لاجراء ما وراء التحليل ويعمل فى بيئه DOS لمساعدة الباحث على اجراء ما وراء التحليل للبيانات التصنيفية الثنائية من التجارب الاكلنيكية باستخدام النموذج المثبت والنموذج العشوائى ويمدنا برسومات واشكال مختلفة ولكن من محدداته اقتصاره على البيانات التصنيفية الثنائية ا ومتاح فى:

<http://www.spc.univ-lyon1.fr/easyma.dos/>.

برنامج Mix

وهو اختصاراً لـ Meta-Analysis with Interactive eXplanation ومتاح مجاناً ويعمل في بيئة Excel وهذا البرنامج مناسب لتعلم ما وراء التحليل لانه يقدم امثلة بيانات من كتب معروفة في مجال ما وراء التحليل ويمكن تحليلها باستخدام مداخل عديدة ومختلفة، ويقدر حجم التأثير للبيانات المتصلة والتصنيفية ويقدر احصائيات ما وراء التحليل في ضوء النموذج المثبت والنموذج العشوائى ويستخدم اشكال وجداول Excel ويعرض شكل Forest وشكل القمع ومحدود في امكانياته. ومتاح فى:

<http://www.mix-for-meta-analysis.info/>.

برنامج MedCalc

من البرامج المتخصصة فى اجراء ما وراء التحليل ويقدم التحليلات المختلفة من حساب حجم التأثير واختبارات التجانس وتحيز النشر واستخدام النموذج المثبت والنموذج العشوائى ويعرض الرسومات والاشكال المختلفة ومتخصص فى ما وراء التحليل فى الدراسات الطبية.

برنامج MetaXL

هو برنامج فرعى لما وراء التحليل يعمل من خلال برنامج Excel ويعتبر من افضل البرامج لاعطاء تقديرات دقيقة لتقديرات ما وراء التحليل ويطبق طرق واساليب ما وراء التحليل المتاحة فى البرامج الاحصائية العامة مثل Stata ويتضمن تطبيقات عديدة منها نموذج التأثيرات المثبتة والعشوائية ونماذج تأثيرات الجودة للبيانات المتصلة والثنائية وتحليل ما وراء التحليل التراكمية واختبارات التجانس وتوجد منه اصدارات عديدة اهمها Version 5.3 وهو متاح مجاناً فى الموقع: www.epigear.com وعند تحميله على الكمبيوتر يتم التعامل معه من خلال برنامج Excel حيث تظهر ايقونة MetaXL فى شريط القوائم. ويتضمن طريقة جديدة لتقدير التجانس بين الدراسات تسمى اختلاف التباين المقلوب Invert variance heterogeneity (IVhet) وتعتمد على الاحتمالات التقريبية.

برنامج MetaAnalyst

يعمل فى بيئة Windows ومن البرامج المتقدمة لاجراء ما وراء التحليل حيث يتضمن التطبيقات الحديثة لما وراء التحليل ويمكن تشغيله من خلال لغة برنامج R ويقبل البيانات من برنامج Excel ويجرى عدد كبير من التحليلات لاحجام التأثير للبيانات المتصلة والتصنيفية وينفذ تحليل الحساسية ونموذج التأثيرات المثبتة من خلال مقلوب تباين الخطأ المعيارى ويجرى التحليل باستخدام طريقة Mantel-Haenszel وطريقة الترجيح الاقصى والتحليلات البيزانية وما وراء الانحدار والتحليل للمجموعات الفرعية ويجرى دراسات المحاكاة. ويمدنا البرنامج بالعديد من الاشكال مثل شكل القمع وForest plot ويعتبر من اقوى برامج ما وراء التحليل مثل برامج CMA وMetaXL ومتاح مجاناً من الموقع:

http://tuftscaes.org/meta_analyst/.

وفيما يلى مقارنة بين اهم هذه البرامج كالتالى (Schmidt et al.(2001):

جدول(1.13): مقارنة بين بعض برامج ما وراء التحليل.

المقارنة	Re Man	Metaanalyst	MI X	MetaWin	CM A
تحميله	مجاناً	مجاناً	مجاناً	60\$	مبلغ ما
يقبل بيانات من برامج اخرى	لا	نعم	نعم	نعم	نعم
ما وراء الانحدار	لا	نعم	نعم	نعم	نعم
التاثيرات المثبتة	نعم	نعم	نعم	نعم	نعم
التاثيرات العشوائية	نعم	نعم	نعم	نعم	نعم
النماذج متعددة المستويات	لا	نعم	لا	لا	لا
النماذج البيزانية	لا	نعم	لا	لا	لا

نعم	نعم	نعم	نعم	لا	التحليل التراكمي
نعم	نعم	نعم	لا	لا	اختبار تحيز النشر
نعم	نعم	نعم	نعم	نعم	البيانات التصنيفية
نعم	نعم	نعم	نعم	نعم	البيانات المتصلة
لا	لا	لا	نعم	لا	البيانات المترجة
نعم	نعم	نعم	نعم	نعم	شكل Forest
نعم	نعم	نعم	نعم	نعم	شكل القمع

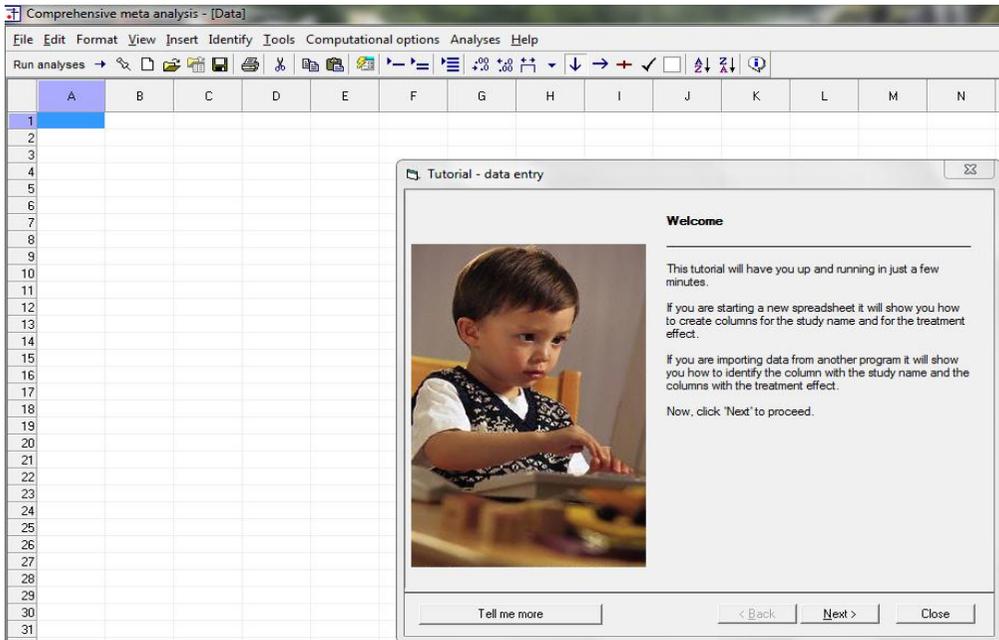
برنامج 3 Comprehensive Meta-Analysis(CMA) Version 3

تم تطوير هذا البرنامج على يد مجموعة من الباحثين هم : Borenstein,

Harris و مجموعة كبيرة من الهيئة الاستشارية امثال Hedges, Higgins, Rothstein و Matthias Egger و Mark Lipsey وغيرهم. ويتم تنزيل البرنامج على جهاز الكمبيوتر على سطح المكتب او في قائمة كل البرامج.

افتح البرنامج من خلال Start ثم All Programs ثم اضغط Comprehensive

Meta-Analysis تظهر الشاشة الاتية:

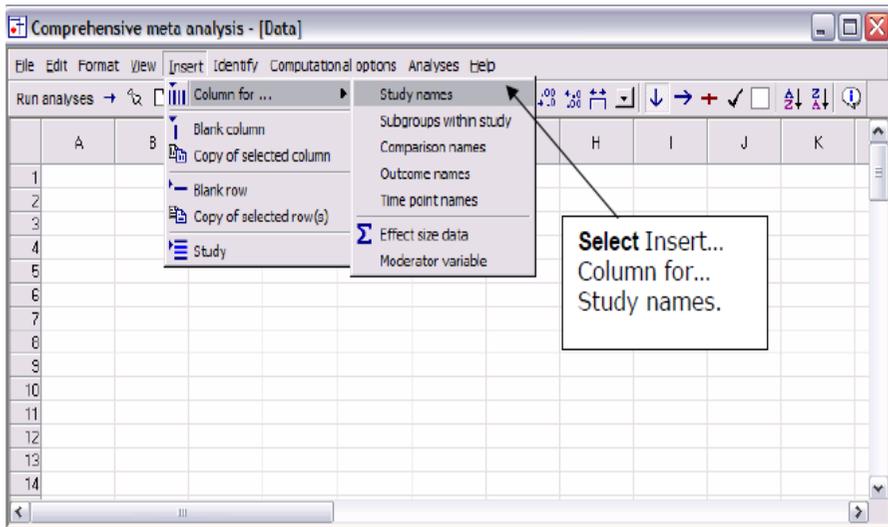


اضغط Close.

الجزء الاول: ادخال البيانات الاساسية والتحليل

لادخال البيانات انشاء عمود لاسماء الدراسات كالتالى:

Insert.....Colum for..... Study names

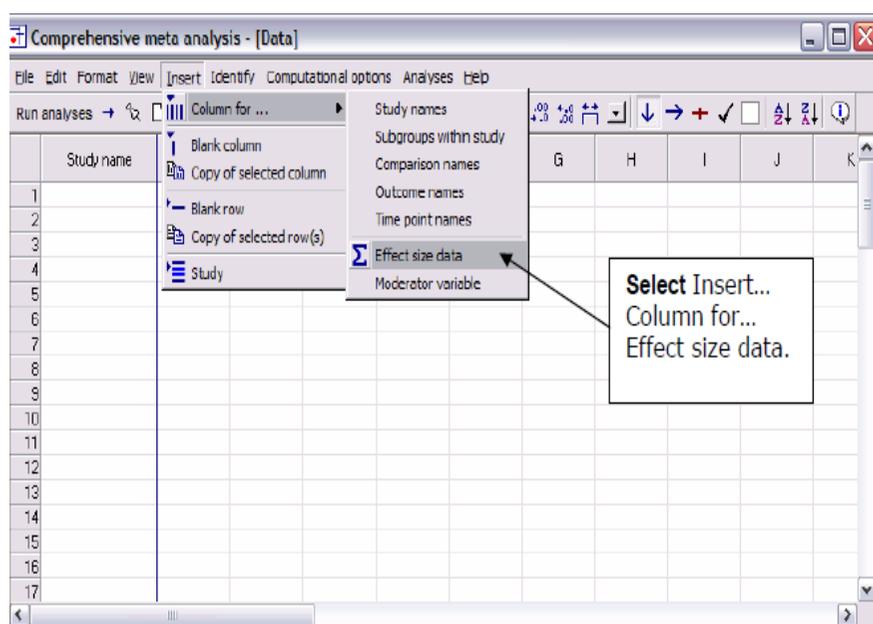


يقوم البرنامج بحشر اسماء الدراسات او مؤلف الدراسات كالتالى:

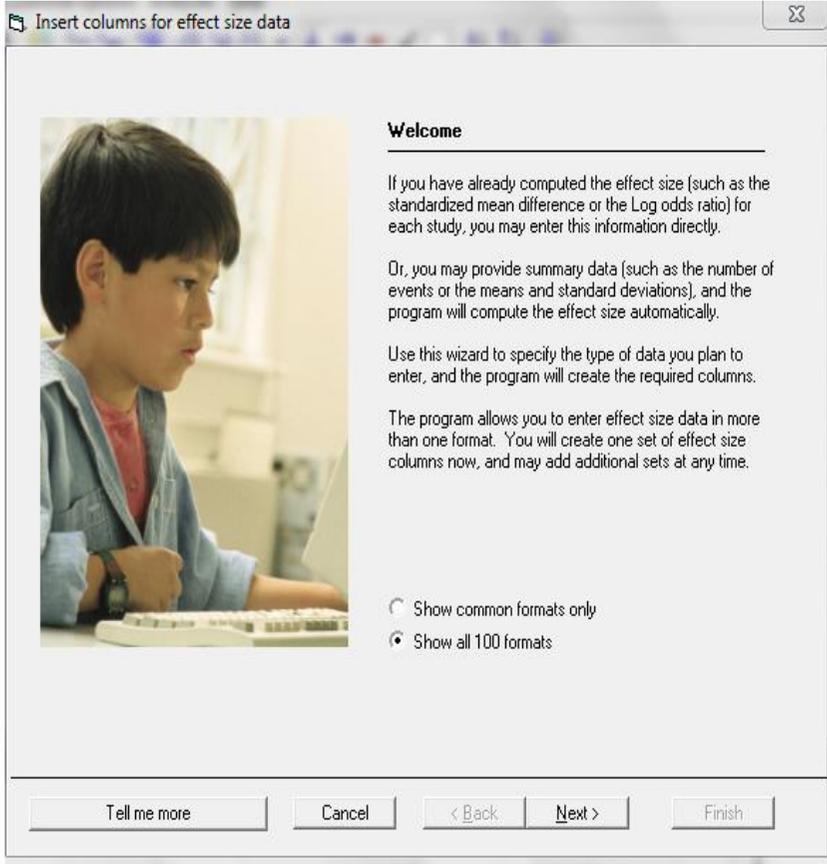
	Study name	B	C	D	E	F	G
1							
2							
3							
4							
5							
6							
7							
8							
9							
10							
11							
12							

انشاء اعمدة لبيانات حجم التأثير كالتالى:

Insert.....Column for..... Effect Size data



تظهر الشاشة الاتية:



هذه الشاشة تمدنا بخيارات لادخال بيانات حجم التأثير كالاتي:

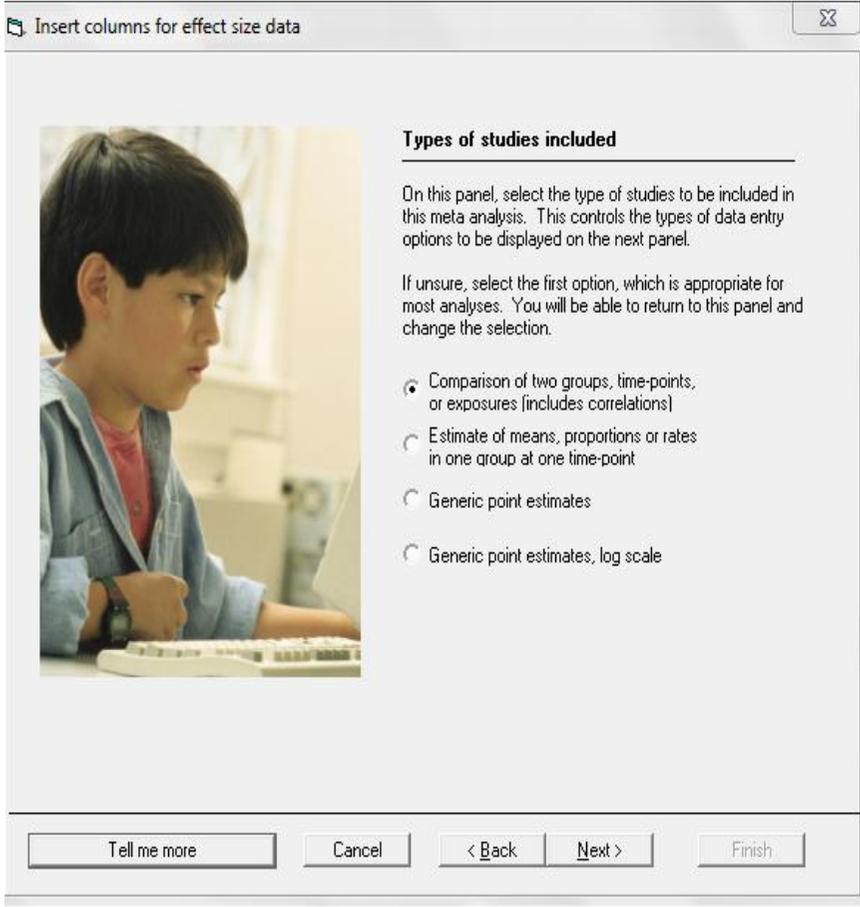
الاول: لو قدر الباحث حجم التأثير مباشرة مثل مؤشرات فروق المتوسطات المعيارية او لوغاريتم Odds ratio لكل دراسة فادخلها مباشرة في عمود حجم التأثير.

الثاني: اذا ادخلت البيانات التلخيصية مثل عدد الاحداث او الحالات او المتوسطات والانحرافات المعيارية للمجموعات فان البرنامج يقدر احجام التأثير مباشرة من تلك البيانات.

من الشاشة السابقة حدد نوعية طبيعة ادخال البيانات المراد ادخالها ثم يقوم

البرنامج بانشاء الاعمدة المطلوبة لذلك. ويسمح البرنامج بادخال البيانات باشكال عديدة.

اضغط على اختيار **Next >** تظهر الشاشة الاتية:



هذه الشاشة تسمح للباحث اختيار نوعية بيانات حجم التأثير كالتالي:

- المقارنة بين مجموعتين او التدخلات وتتضمن العلاقات.
- تقدير المتوسطات، النسب، والتقديرات لمجموعة واحدة.
- التقديرات اللوغاريتمية.

- اضغط على البديل الاول المتضمن المقارنة بين مجموعتين او العلاقات ثم اضغط Next

تظهر الشاشة الاتية:

- Continuous (means)
 - Unmatched groups, post data only
 - Unmatched groups, pre and post data
 - One group (pre-post) and matched groups
 - Computed effect sizes

وتتضمن المجموعات المستقلة للبيانات البعدية فقط، بيانات المجموعة غير المتزاوجة وقياس قبلي وبعدي، وبيانات المجموعة الواحدة قبلي وبعدي والمجموعات المتزاوجة، وحجم التأثير.

وإذا ضغط الباحث على اختيار المجموعات المستقلة وقياس بعدي فقط تظهر الشاشة

الخيارات الآتية:

- Unmatched groups, post data only
 - Mean, SD and sample size in each group
 - Difference in means, common SD, and sample size
 - Cohen's d (standardized by pooled within-groups SD) and sample size
 - Means, sample size, and t-value
 - Difference in means, sample size, and t-value
 - Sample size and t-value
 - Means, sample size, and p-value
 - Difference in means, sample size, and p-value
 - Sample size and p-value

وعلى الباحث ان يختار شكل البيانات المراد ادخالها سواء المتوسط والانحراف المعياري SD وحجم العينة في كل مجموعة، او مؤشر d لـ Cohen وحجم العينة او حجم العينة وقيمة P الاحتمالية وغيرها.

– معامل الارتباط: Computed effect sizes : قيم معامل الارتباط مباشرة. بالضغط على

Computed effect size تظهر اختيارات ادخال معامل الارتباط كالآتي:

- Correlation
 - Computed effect sizes
 - Correlation and sample size
 - Correlation and standard error
 - Correlation and variance
 - Fisher's Z and sample size
 - Fisher's Z and standard error
 - Fisher's Z and variance
 - Correlation and t-value
 - t-value and sample size for correlation
 - p-value and sample size for correlation

• التقديرات:

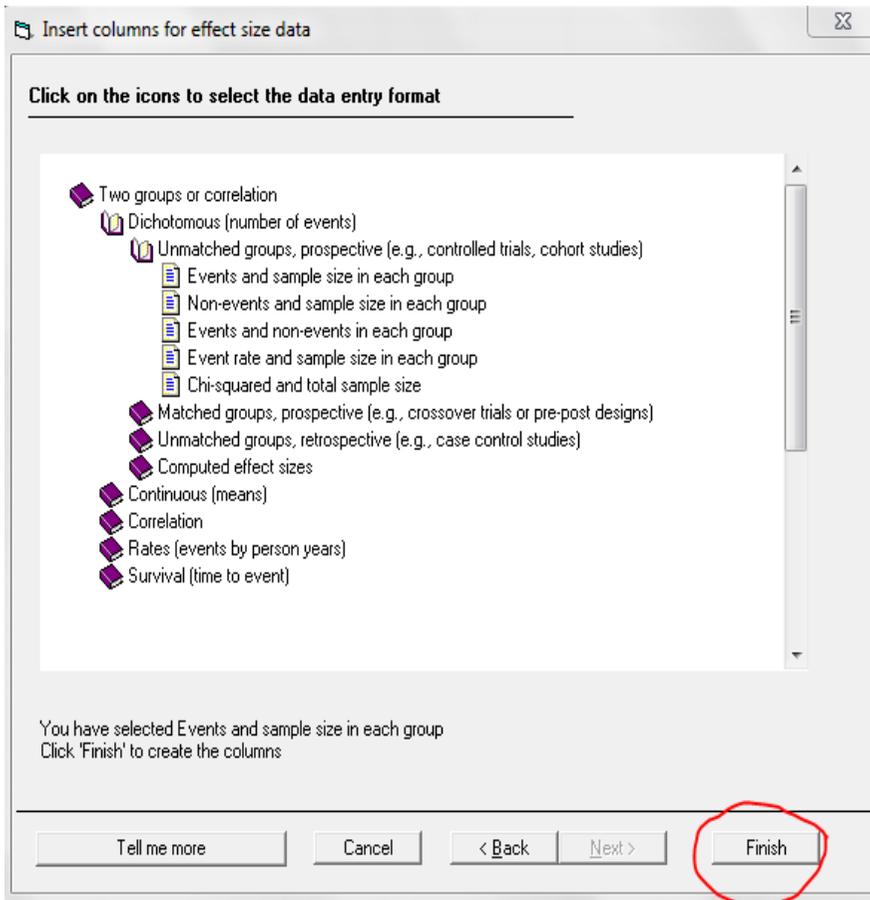
-  Rates (events by person years)
-  Unmatched groups
-  Computed effect sizes

• البيانات الطولية:

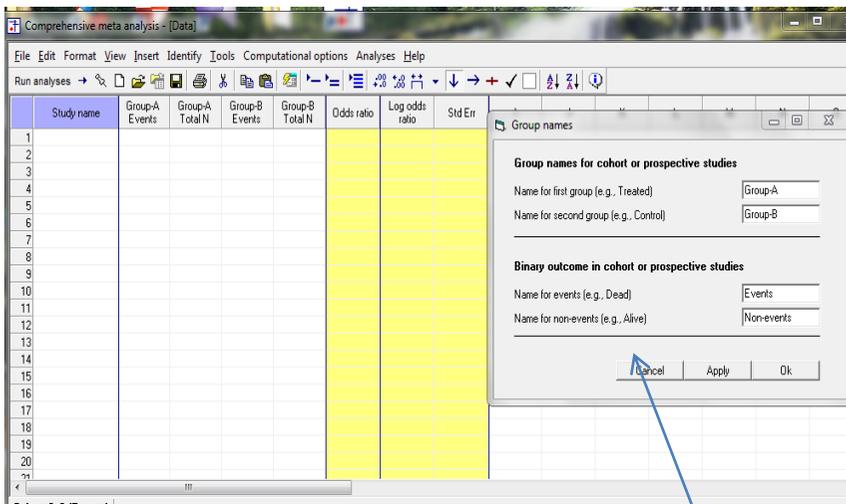
-  Survival (time to event)
-  Computed effect sizes

فاذا اهتم الباحث بدراسة ما وراء التحليل لدراسات تجريبية حيث المتغير المستقل والتابع تصنيفي ثنائى كما فى الدراسات الطبية وفى هذه الحالة فان مؤشر حجم التاثير Odds ratio(OR) فعلى الباحث اتباع الاتى:

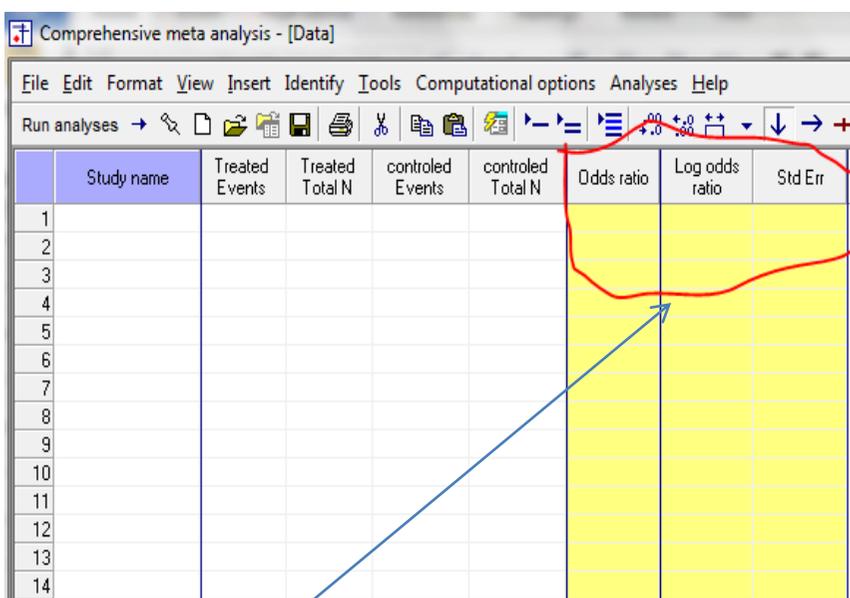
- اضغط Dichotomous ثم unmatched groups ثم اختار Events and sample size in each group size in each group تظهر الشاشة الاتية:



- اضغط Finish تظهر الشاشة الاتية المتضمنة الاعمدة لادخال البيانات:



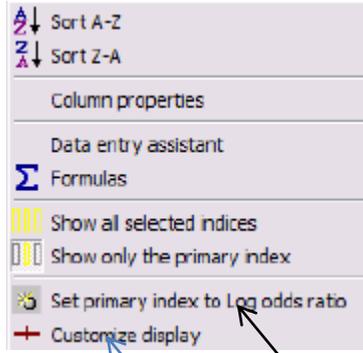
وفي المربع الفرعي يمكن تغيير مسمى المجموعات فبدلاً من Group A يمكن تسميتها بالمجموعة المعالجة Treated والمجموعة Group B يمكن تسميتها بالمجموعة الضابطة Control ثم اضغط Apply تظهر الشاشة الآتية:



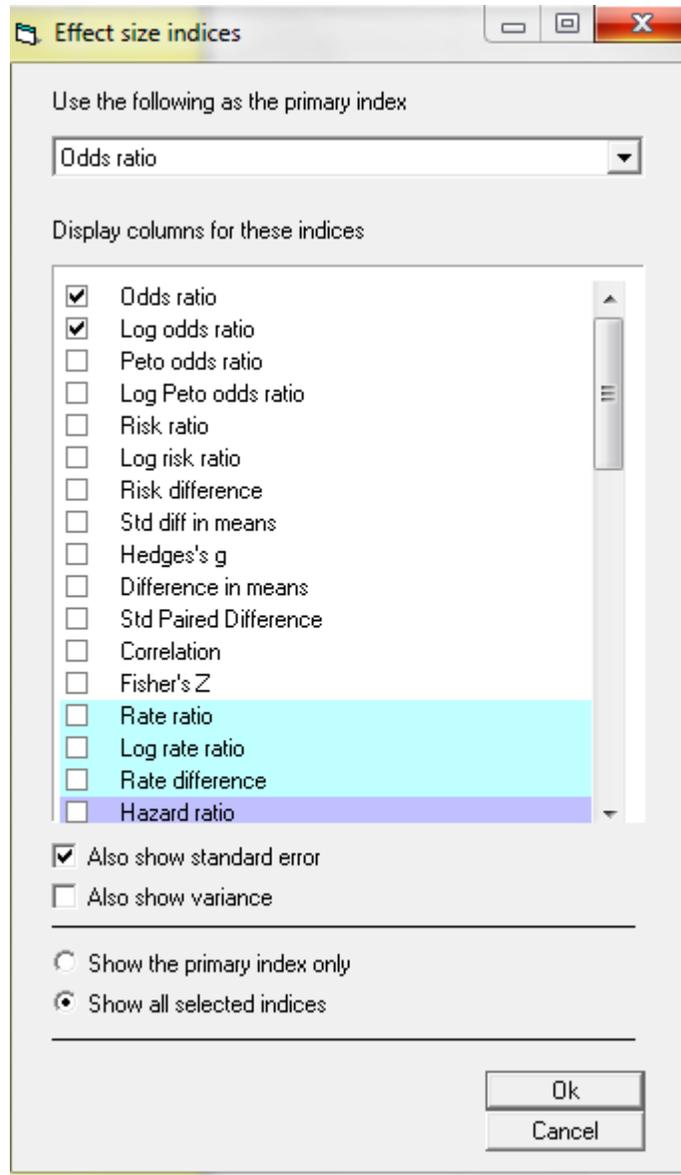
يحسب البرنامج حجم التأثير ويعرضها في الاعمدة الثلاثة.

ويمكن استعراض مؤشرات اخرى لحجم التأثير من خلال:

- اضغط كليك يمين على عمود Odds ratio تظهر قائمة كالآتي:



- اضغط set primary index او Customized Display تظهر المؤشرات الاتية:



- ابدأ بادخل البيانات كما فى الشكل الاتى واضغط على خانة Odds ratio يقوم بحساب مؤشرات حجم التأثير ويعرضها فى الاعمدة الصفراء الثلاثة :

	Study name	Treated Events	Treated Total N	controled Events	controled Total N	Odds ratio	Log odds ratio	Std Err
1	A	4	123	11	139	0.391	-0.939	0.598
2								
3								
4								
5								

ولإظهار معادلات حساب مؤشرات حجم التأثير اضغط مرتين بالموس على خانة او المربع 0.391 في عمود Odds ratio تظهر الشاشة الفرعية الآتية:

Data entry assistant

Data entry **Odds ratio**

Starting with
Cells in 2x2 table

Where cells are given as
A = Treated Events
B = Treated Total N - Treated Events
C = controled Events
D = controled Total N - controled Events

A = 4
B = 123 - 4 = 119
C = 11
D = 139 - 11 = 128

LogOddsRatio = $\text{Log}((A * D) / (B * C))$
LogOddsVariance = $(1 / A + 1 / B + 1 / C + 1 / D)$
LogOddsSE = $\text{Sqr}(\text{LogOddsVariance})$
OddsRatio = $\text{Exp}(\text{LogOddsRatio})$

LogOddsRatio = $\text{Log}((4 * 128) / (119 * 11)) = -0.939$
LogOddsVariance = $(1/4 + 1/119 + 1/11 + 1/128) = 0.357$

[< Home >](#)

ثم يقوم الباحث بادخال بيانات دراسة جديدة ويقوم بالضغط على مربع احصاء Odds ratio لحساب مؤشرات حجم التأثير كالاتى:

Comprehensive meta analysis - [Data]

File Edit Format View Insert Identify Tools Computational options Analyses Help

Run analyses →

	Study name	Treated Events	Treated Total N	controled Events	controled Total N	Odds ratio	Log odds ratio	Std Err
1	A	4	123	11	139	0.391	-0.939	0.598
2	B	5	150	10	200	0.655	-0.423	0.559
3								
4								

اجراء التحليل

ويقوم الباحث بادخال باقى بيانات دراسة ما وراء التحليل واجراء التحليل كالاتى:

Comprehensive meta analysis - [Data]

File Edit Format View Insert Identify Tools Computational options Analyses Help

Run analyses →

	Study name	Group-A Events	Group-A Total N	Group-B Events	Group-B Total N	Odds ratio	Log odds ratio	Std Err
1	A	4	123	11	139	0.391	-0.939	0.598
2	B	180	1541	372	1451	0.384	-0.958	0.100
3	C	8	2545	10	629	0.195	-1.634	0.476
4	D	505	88391	499	88391	1.012	0.012	0.063
5	E	29	7499	45	7277	0.624	-0.472	0.239
6	F	17	1716	65	1665	0.246	-1.401	0.275
7	G	186	50634	141	27338	0.711	-0.341	0.112
8	H	5	2498	3	2341	1.563	0.447	0.731
9	I	27	16913	29	17854	0.983	-0.017	0.268
10								
11								

- اضغط على Run analyses او اضغط على قائمة Analysis لتنفيذ التحليل كالاتى:

Comprehensive meta analysis - [Analysis]

File Edit Format View Computational options Analyses Help

Data entry Next table High resolution plot Select by ... Effect measure: Odds ratio

Model	Study name	Statistics for each study					Odds ratio and 95% CI				
		Odds ratio	Lower limit	Upper limit	Z-Value	p-Value	0.01	0.10	1.00	10.00	100.00
	A	0.391	0.121	1.262	-1.571	0.116					
	B	0.384	0.316	0.466	-9.627	0.000					
	C	0.195	0.077	0.497	-3.429	0.001					
	D	1.012	0.894	1.146	0.190	0.849					
	E	0.624	0.391	0.996	-1.976	0.048					
	F	0.246	0.144	0.422	-5.102	0.000					
	G	0.711	0.571	0.886	-3.046	0.002					
	H	1.563	0.373	6.548	0.611	0.541					
	I	0.983	0.582	1.661	-0.065	0.948					
Fixed		0.723	0.661	0.790	-7.136	0.000					

Fixed Random Both models

Basic stats One study removed Cumulative analysis Calculations

لاحظ انه عرض الاحصائيات الخاصة بكل دراسة تحت الاعمدة المسمى Statistics for each study متضمنة مؤشر حجم التأثير Odds ratio و 95% فترات الثقة بينما الصف الاخير يبين الاحصاء التلخيص (متوسط حجم التأثير) حيث باستخدام النموذج المثبت فان متوسط Odds ratio = 0.723 بفترات ثقة -0.66 الى 0.709 وقيمة اختبار Z للدلالة الاحصائية لمتوسط حجم التأثير -7.00 وهى اكبر من 1.96 اى ان العلاقة بين المتغيرين دالة احصائياً بمعنى للمعالجة اثر فى احداث حسن.

وعلى الناحية اليمنى عرض شكل Forest وهذا الشكل يوضح الحد الاعلى والادنى لفترات الثقة ويعطى انطباع بمدى ابتعاد الدراسات عن بعضها البعض. وهذه الشاشة تتضمن عدة ايقونات منها:

عرض ملخص للنتائج

- اضغط Next table تظهر الجدول الاتى:

Comprehensive meta analysis - [Analysis]									
File Edit Format View Computational options Analyses Help									
← Data entry		↔ Next table		⚙ High resolution plot		📄 Select by ...		+ Effect measure: Odds ratio	
Model		Effect size and 95% interval			Test of null (2-Tail)				
Model	Number Studies	Point estimate	Lower limit	Upper limit	Z-value	P-value			
Fixed	9	0.723	0.661	0.790	-7.136	0.000			
Random	9	0.556	0.373	0.828	-2.887	0.004			

Heterogeneity				Tau-squared		
Q-value	df (Q)	P-value	I-squared	Tau Squared	Standard Error	Variance
95.605	8	0.000	91.632	0.274	0.237	0.056

في هذا الجدول تم عرض اختبار عدم التجانس بين الدراسات Q وكانت قيمته 95.60 ويتم اختبار دلالتها في ضوء اختبار كاي تربيع لدرجات حرية (9-1=8) ومستوى دلالة 0.50 وهي دالة احصائياً حيث $P=0.000$ وعليه يوجد اختلاف بين احجام التأثير بالتالي فان استخدام النموذج المثبت غير مناسب والافضل الاعتماد على نتائج نموذج التأثيرات العشوائية.

عرض اوزان الدراسات

يمكن عرض اوزان الدراسات من خلال الضغط على ايقونة:



 تعرض احصائيات الدراسات المفردة.

 تعرض الاحصائيات الملخصة.

 تعرض الاحصائيات الاساسية المخفية.

 تعرض الاعداد غير المرئية.

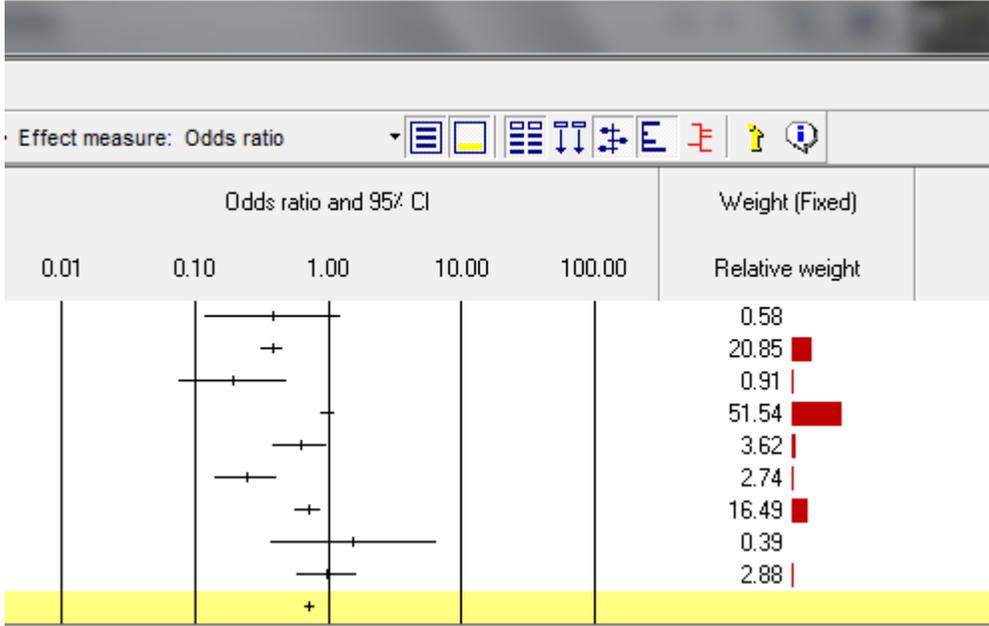
 تعرض شكل Forest.

تعرض اوزان الدراسات.

تعرض البواقي.

تعرض تفاصيل.

وبالضغط على تعطى الاوزان النسبية للدراسات كالآتي:

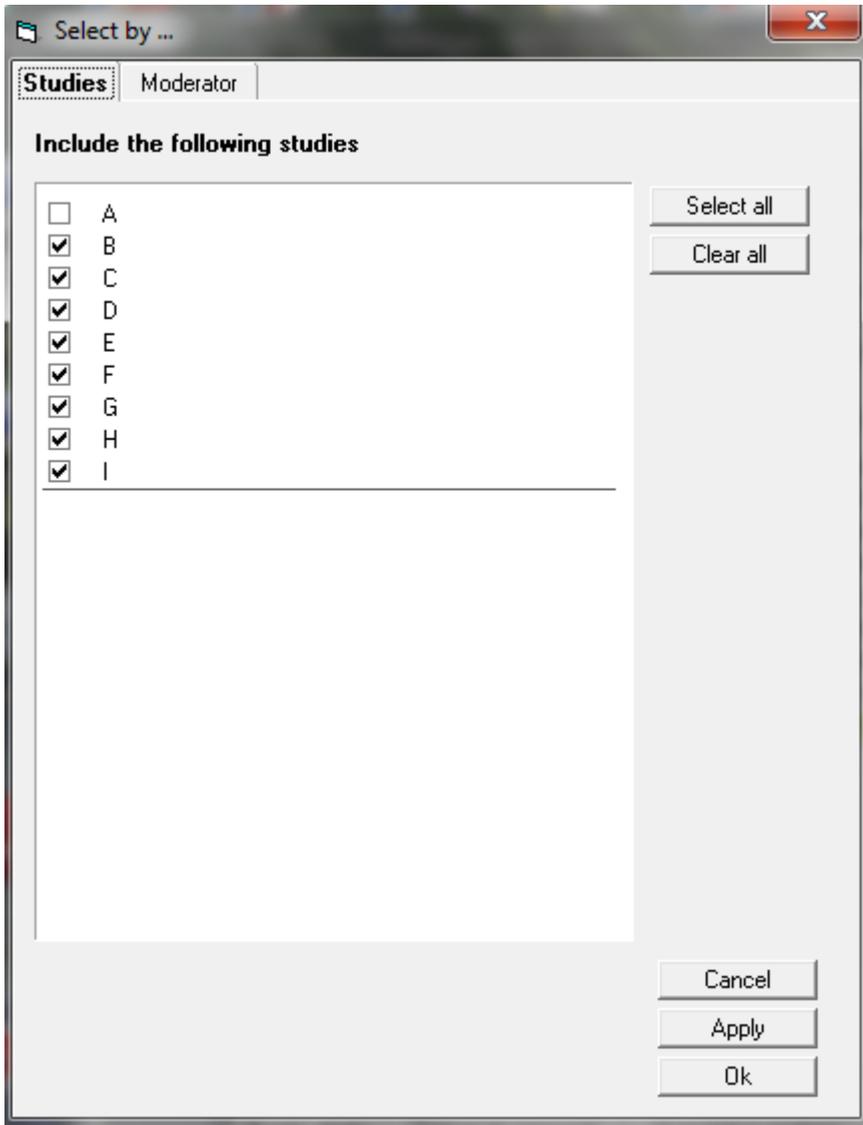


عرض نتائج ما وراء التحليل مع حذف احد الدراسات (تحليل الحساسية)

- اضغط على Select case:



يعطى الجدول الآتي:



ثم اضغط على الدراسة A يستبعدا من التحليل.

التحليل التراكمي Cumulative Meta-analysis

هو الية لعرض سلسلة منفصلة من تحليلات ما وراء التحليل في جدول او شكل واحد حيث تدمج الدراسة الاولى مع الثانية في تحليل ثم الدراسات الثلاثة الاولى في تحليل وهكذا، وهذا الدمج يتم في ضوء بعض العوامل مثل التسلسل الزمني او الجودة المنهجية وهذا يبين كيفية تغير متوسط حجم التأثير عبر فترات زمنية ويمكن اعتبار هذا التحليل اداة تربوية للحكم على مدى تطور الممارسات البحثية التربوية عبر فترات زمنية متعاقبة Borenstein et al.,2009).

من خلال الشريط الاسفل يوجد:



اضغط عليه يعطى النتائج الاتية:

Model	Study name	Statistics for each study					Odds ratio and 95% CI					Weight (Fixed)		
		Odds ratio	Lower limit	Upper limit	Z-Value	p-Value	0.01	0.10	1.00	10.00	100.00	Relative weight	Weight (Fixed)	
	A	0.391	0.121	1.262	-1.571	0.116							0.58	
	B	0.384	0.316	0.466	-9.627	0.000							20.85	
	C	0.195	0.077	0.497	-3.429	0.001							0.91	
	D	1.012	0.894	1.146	0.190	0.849							51.54	
	E	0.624	0.391	0.996	-1.976	0.048							3.62	
	F	0.246	0.144	0.422	-5.102	0.000							2.74	
	G	0.711	0.571	0.886	-3.046	0.002							16.49	
	H	1.563	0.373	6.548	0.611	0.541							0.39	
	I	0.983	0.582	1.661	-0.065	0.948							2.88	
Fixed		0.723	0.661	0.790	-7.136	0.000								

يقوم البرنامج بعرض النتائج متراكمة حيث الصف الثاني يعرض نتائج الدراساتين الاولتين السابقتين والصف الثالث يعرض نتائج الدراسات الثلاثة السابقة وهكذا.

عرض الحسابات التفصيلية

- اضغط على Calculations اسفل الشاشة يعطى الحسابات التفصيلية الاتية:

Comprehensive meta analysis - [Analysis]

File Edit Format View Computational options Analyses Help

Data entry Next table High resolution plot Select by ... Effect measure: Odds ratio

Model	Study name	Calculations (Fixed)											
		Point	Study Variance	Tau ² Within	Tau ² Between	Total Variance	IV-Weight	W	T*W	T ² *W	W ²	W ³	C
A		-0.939	0.357	0.000	0.000	0.357	2.800	2.800	-2.628	2.467	7.841	21.955	319.901
B		-0.958	0.010	0.000	0.000	0.010	100.956	100.956	-96.729	92.678	10192.196	1028967.44	319.901
C		-1.634	0.227	0.000	0.000	0.227	4.405	4.405	-7.197	11.758	19.405	85.480	319.901
D		0.012	0.004	0.000	0.000	0.004	249.566	249.566	3.000	0.036	62283.004	15543697.1	319.901
E		-0.472	0.057	0.000	0.000	0.057	17.551	17.551	-8.280	3.906	308.034	5406.279	319.901
F		-1.401	0.075	0.000	0.000	0.075	13.259	13.259	-18.578	26.032	175.795	2330.830	319.901
G		-0.341	0.013	0.000	0.000	0.013	79.839	79.839	-27.213	9.276	6374.340	508923.935	319.901
H		0.447	0.534	0.000	0.000	0.534	1.872	1.872	0.836	0.373	3.505	6.561	319.901
I		-0.017	0.072	0.000	0.000	0.072	13.960	13.960	-0.242	0.004	194.871	2720.333	319.901
		5.303	1.349	0.000	0.000	1.349	484.208	484.208	-157.031	146.531	79568.991	17092159.9	319.901

Fixed Random

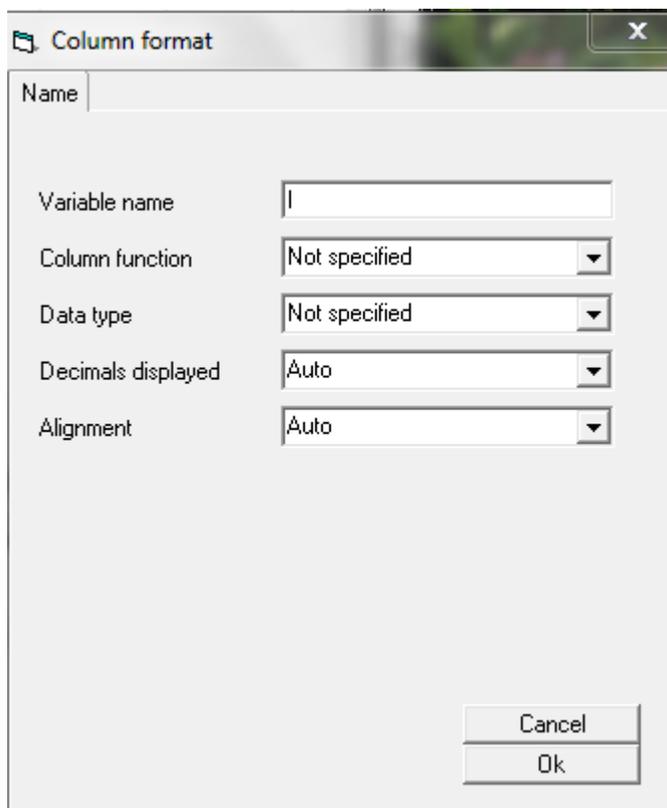
Basic stats One study removed Cumulative analysis Calculations

Q	Q df	I ²	B	K	Summary Point	Summary Variance	Group T ²	Group T ² Variance
95.605	8.000	91.632	2879.790	9.000	-0.324	0.002	0.274	0.058
95.605	8.000	91.632	2879.790	9.000	-0.324	0.002	0.274	0.058
95.605	8.000	91.632	2879.790	9.000	-0.324	0.002	0.274	0.058
95.605	8.000	91.632	2879.790	9.000	-0.324	0.002	0.274	0.058
95.605	8.000	91.632	2879.790	9.000	-0.324	0.002	0.274	0.058
95.605	8.000	91.632	2879.790	9.000	-0.324	0.002	0.274	0.058
95.605	8.000	91.632	2879.790	9.000	-0.324	0.002	0.274	0.058
95.605	8.000	91.632	2879.790	9.000	-0.324	0.002	0.274	0.058
95.605	8.000	91.632	2879.790	9.000	-0.324	0.002	0.274	0.058
95.605	8.000	91.632	2879.790	9.000	-0.324	0.002	0.274	0.058

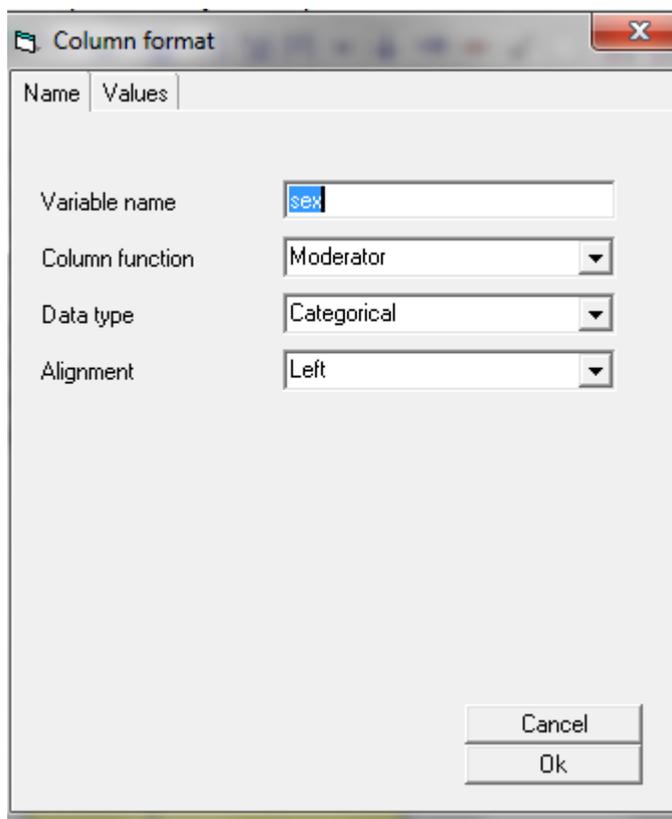
اجراء التحليلات المتفاعلة

لاجراء التحليلات المتفاعلة باستخدام متغير تصنيفي (خاصية الدراسة) وليكن جنس العينة لابد من انشاء عمود المتغير المتفاعل كالتالى:

- اضغط مرتين متتاليتين بالموس على اى عمود لايضمن بيانات تظهر الشاشة التالية:



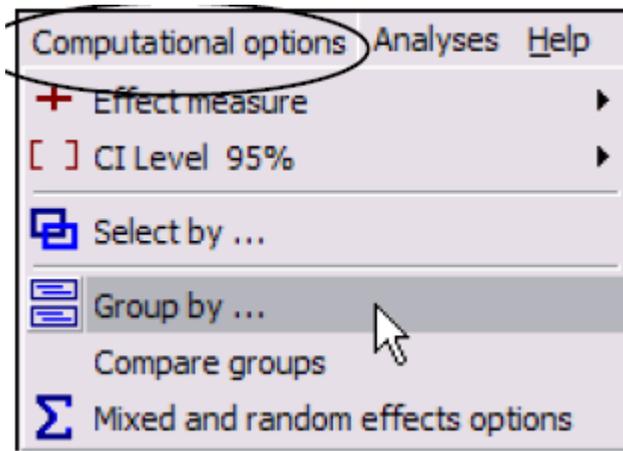
- امام Column function اضغط على السهم لاسفل واختار Moderator امام data type
- اكتب مسمى المتغير التصنيفي وليكن الجنس Sex.



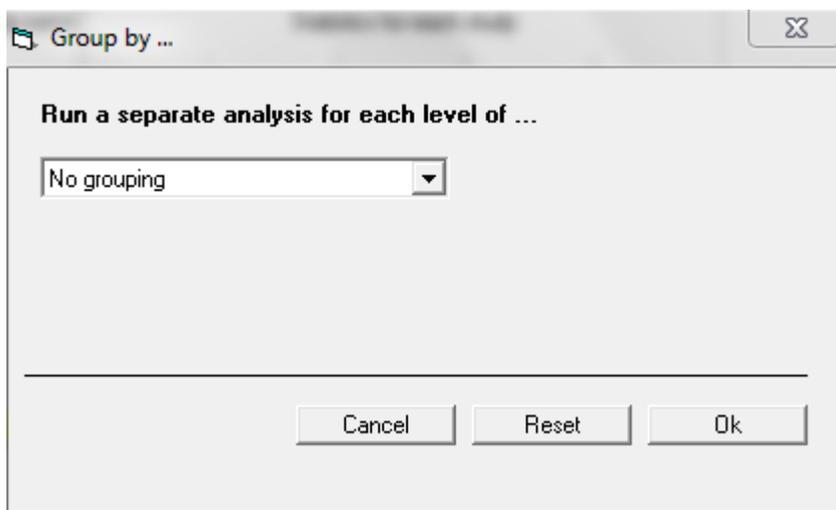
- اضغط OK يظهر مسمى المتغير فى شاشة البيانات.
- ابدأ فى ادخال بيانات المتغير المتفاعل كما يلى:

	Study name	Group-A Events	Group-A Total N	Group-B Events	Group-B Total N	Odds ratio	Log odds ratio	Std Err	sex
1	A	4	123	11	139	0.391	-0.939	0.598	M
2	B	180	1541	372	1451	0.384	-0.958	0.100	F
3	C	8	2545	10	629	0.195	-1.634	0.476	M
4	D	505	88391	499	88391	1.012	0.012	0.063	F
5	E	29	7499	45	7277	0.624	-0.472	0.239	M
6	F	17	1716	65	1665	0.246	-1.401	0.275	F
7	G	186	50634	141	27338	0.711	-0.341	0.112	M
8	H	5	2498	3	2341	1.563	0.447	0.731	F
9	I	27	16913	29	17854	0.983	-0.017	0.268	M
10									

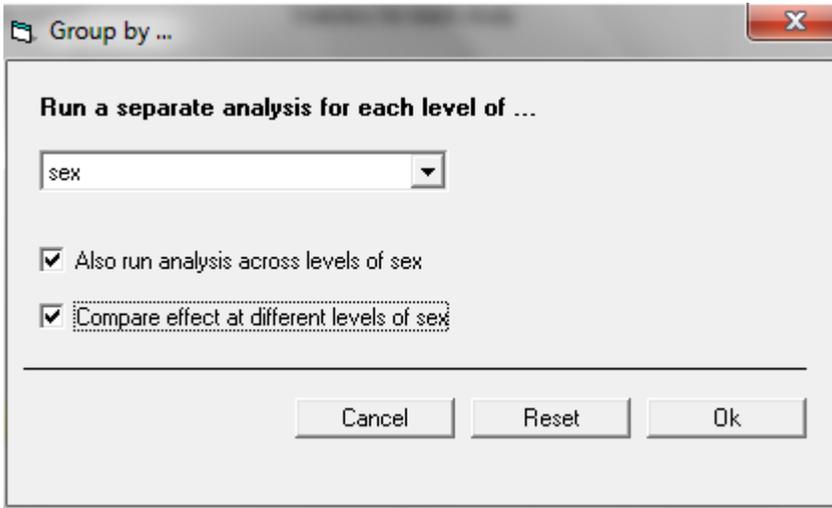
- اضغط Run analyses ويظهر المخرج اضغط على قائمة Computational options تظهر الشاشة الاتية:



- اضغط على Group by يعطى المربع التالى:



- اضغط على السهم الاسفل فى امام No grouping واختار المتغير المتفاعل Sex ونشط الايقونتين فى المربع وهذا يعنى اجراء التحليل لكل مستوى من مستويات المتغير التصنيفى وكذلك المقارنة بين المستويات المختلفة ويعطى الاختلاف بين وداخل المجموعات. وتكون المربع كالاتى:



- اضغط OK يظهر المخرج الاتي:

Model	Group by sex	Study name	Statistics for each study					Odds ratio and 95% CI		
			Odds ratio	Lower limit	Upper limit	Z-Value	p-Value	0.50	1.00	2.00
	female	B	0.384	0.316	0.466	-9.627	0.000			
	female	D	1.012	0.894	1.146	0.190	0.849			
	female	F	0.246	0.144	0.422	-5.102	0.000			
	female	H	1.563	0.373	6.548	0.611	0.541			
Fixed	female		0.737	0.665	0.817	-5.829	0.000			
	male	A	0.391	0.121	1.262	-1.571	0.116			
	male	C	0.195	0.077	0.497	-3.429	0.001			
	male	E	0.624	0.391	0.996	-1.976	0.048			
	male	G	0.711	0.571	0.886	-3.046	0.002			
	male	I	0.983	0.582	1.661	-0.065	0.948			
Fixed	male		0.681	0.569	0.815	-4.184	0.000			
Fixed	Overall		0.723	0.661	0.790	-7.136	0.000			

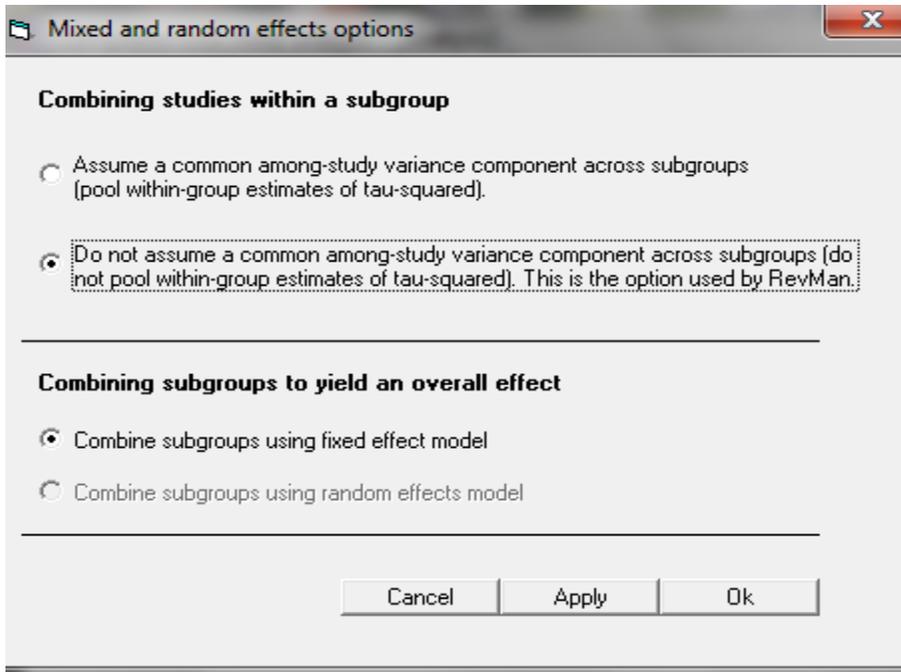
لاحظ ان المخرج اعطى احصائيات كل مجموعة على حدة المظللة بالدائرة بالاضافة الى احصائيات المجموعتين معاً (المشار اليه بالسهم).

اختيار نموذج الحسابات او التحليل

لاختيار نموذج التحليل:

- اضغط Computational options واختار Mixed and random effects

تظهر الشاشة الاتية:



- اضغط Ok .

عرض الاحصائيات الاضافية لتحليل المجموعة

- اضغط على اختيار Next table يظهر الاحصائيات الاضافية لكل مجموعة بالاضافة الى

قيم التباين بين المجموعات وداخل المجموعات كالآتي:

Comprehensive meta analysis - [Analysis]

File Edit Format View Computational options Analyses Help

Data entry Next table High resolution plot Select by ... Effect measure: Odds ratio

Groups	Effect size and 95% interval				Test of null (2-Tail)		Heterogeneity			
	Group	Number Studies	Point estimate	Lower limit	Upper limit	Z-value	P-value	Q-value	df (Q)	P-value
Fixed effect analysis										
female	4	0.737	0.665	0.817	-5.829	0.000	85.137	3	0.000	
male	5	0.681	0.569	0.815	-4.184	0.000	9.903	4	0.042	
Total within							95.040	7	0.000	
Total between							0.565	1	0.452	
Overall	9	0.723	0.661	0.790	-7.136	0.000	95.605	8	0.000	
Mixed effects analysis										
female	4	0.559	0.263	1.190	-1.509	0.131				
male	5	0.607	0.415	0.887	-2.580	0.010				
Total between							0.036	1	0.850	
Overall	9	0.597	0.425	0.838	-2.983	0.003				

Fixed Random Both models

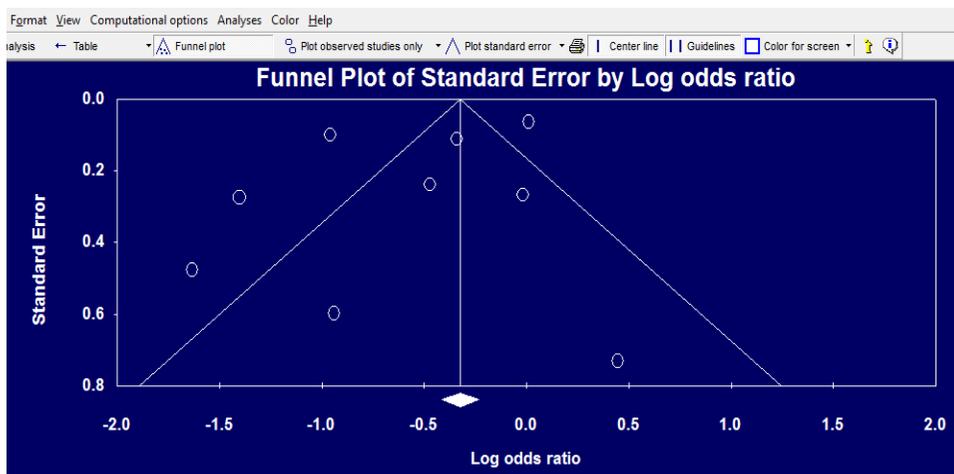
Basic stats Calculations

Tau-squared

I-squared	Tau Squared	Standard Error	Variance	Tau
96.476	0.492	0.620	0.385	0.701
59.606	0.098	0.130	0.017	0.313
91.632	0.274	0.237	0.056	0.523

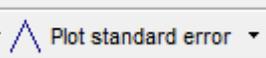
اجراء شكل القمع لتشخيص تحيز النشر

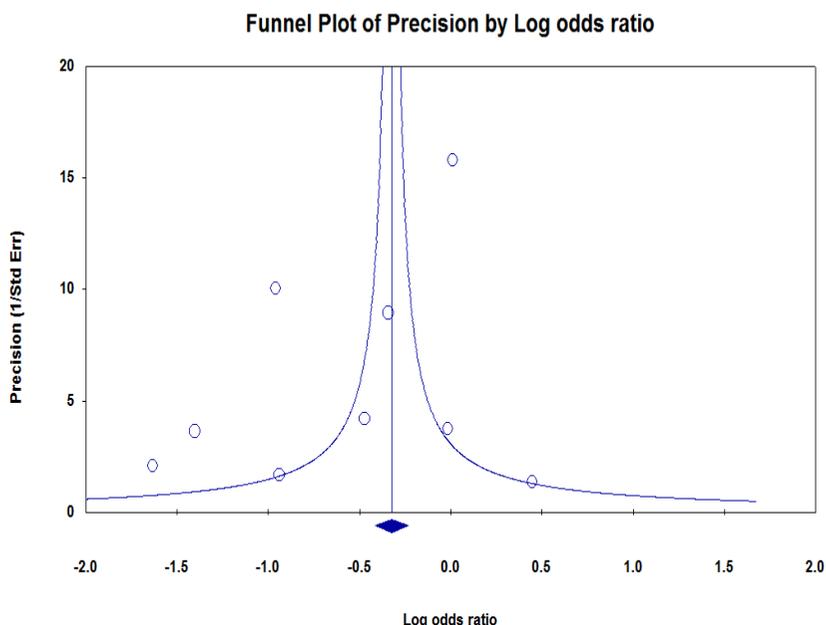
- اضغط على قائمة Analysis واختار Publication bias يعطى الشكل الاتي:



وتم عرض شكل القمع في ضوء الخطأ المعياري و لوغاريتم Odd ratio ويتضح من الشكل وجود تحيز نشر نتيجة وقوع احداثيات عديدة خارج القمع وايضاً ان احجام التأثير لا تتوزع اعتدالياً حول القمع. واذا وقعت معظم الدراسات اسفل الشكل على جانب واحد من الشكل اكثر من جانب اخر فهذا يدل على عدم انتظامية شكل القمع بالتالي وجود تحيز للدراسات.

ويمكن عرض الشكل في ضوء الدقة كالتالي:

اختيار  -الضغط على Plots standard error اعلى الشكل
Plot precision يظهر الشكل الاتي:



ويمكن تغيير لون الشكل من خلال الايقونة المظللة بالاصفر Color of screen.

وتوجد طرق لتقدير التحيز النشر باستخدام اجراء احذف واملى Trim and Fill وفكرته قائمة على شكل القمع ولتنفيذه:

- اضغط على قائمة View فى الشكل السابق واختار Trim and Fill يظهر المخرج الاتى:

Duval and Tweedie's trim and fill

	Fixed Effects			Random Effects			Q Value
	Studies Trimmed	Point Estimate	Lower Limit Upper Limit	Point Estimate	Lower Limit Upper Limit		
Observed values		0.72303	0.66141 0.79039	0.55612	0.37341 0.82823	95.60464	
Adjusted values	0	0.72303	0.66141 0.79039	0.55612	0.37341 0.82823	95.60464	

Look for missing studies where?

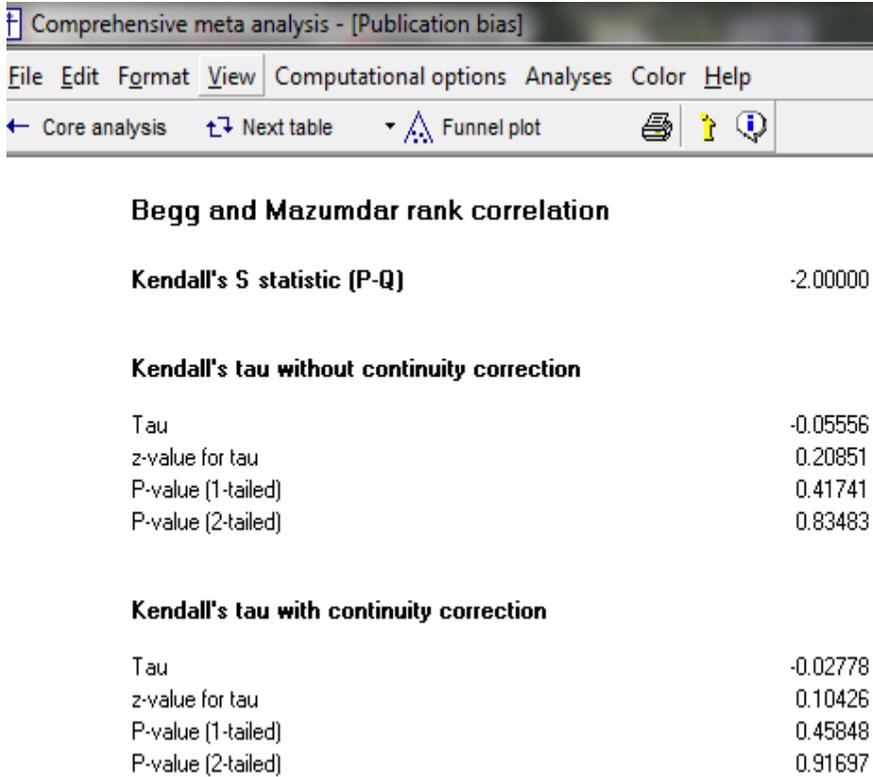
- Not specified
- To left of mean
- To right of mean

وبالضغط على الايقونة اعلى الشاشة يقدم تقرير وتفسير للنتائج. ولايوجد تاثير للدراسة الغائبة على تغير الشكل واعطى متوسط حجم التأثير وفترات الثقة لنموذج التأثيرات المثبتة و العشوائية.

معامل الارتباط الرتبي لتشخيص التحيز

وهو احد طرق الاستدلال الاحصائي لتشخيص التحيز ويقدر من خلال معامل ارتباط كندال تاو b بين احجام التأثير والاختفاء المعيارية المشتقة من احجام العينات يقدر كالتالى:

- ضغط على قائمة View ثم اختار Rank correlation test يظهر المخرج التالى:



The screenshot shows a software window titled "Comprehensive meta analysis - [Publication bias]". The menu bar includes File, Edit, Format, View, Computational options, Analyses, Color, and Help. The toolbar shows "Core analysis", "Next table", "Funnel plot", and other icons. The main content area displays the following results:

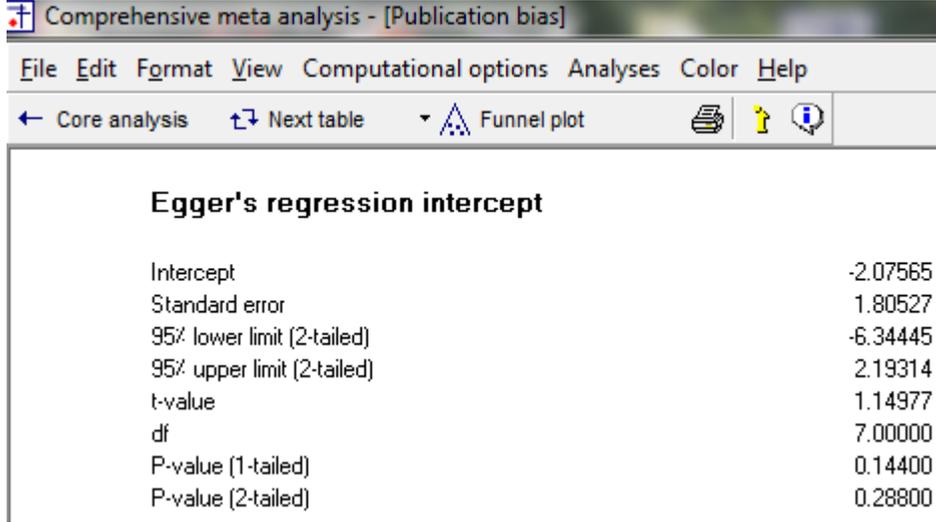
Begg and Mazumdar rank correlation	
Kendall's S statistic (P-Q)	-2.00000
Kendall's tau without continuity correction	
Tau	-0.05556
z-value for tau	0.20851
P-value (1-tailed)	0.41741
P-value (2-tailed)	0.83483
Kendall's tau with continuity correction	
Tau	-0.02778
z-value for tau	0.10426
P-value (1-tailed)	0.45848
P-value (2-tailed)	0.91697

ويتضح عدم الدلالة الاحصائية حيث $P=0.458$ وعليه لا يوجد تحيز للنشر وهذا يتعارض مع تحليل شكل القمع وذلك لان اختبار كندال ليس الاختبار الصادق فى هذه الحالة نظراً لصغر حجم عينة الدراسات ونقص القوة الاحصائية للاختبار ولا يمكن الاعتماد عليه فقط كدليل للحكم على تحيز النشر.

مدخل Egger Regression

هذا المدخل يعتمد على التنبؤ باحجام التأثير من خلال مؤشر الدقة (مقلوب الخطأ المعياري) ويتم تنفيذه كالتالى:

- اضغط على قائمة View ثم اختار Regression test يظهر المخرج الاتى:



The screenshot shows the 'Egger's regression intercept' results in the Comprehensive meta analysis software. The results are as follows:

Parameter	Value
Intercept	-2.07565
Standard error	1.80527
95% lower limit (2-tailed)	-6.34445
95% upper limit (2-tailed)	2.19314
t-value	1.14977
df	7.00000
P-value (1-tailed)	0.14400
P-value (2-tailed)	0.28800

ويتم الحكم على وجود تحيز من خلال Intercept(B0) وقيمة -2.07 بفترات ثقة (-) (6.344, 2.1931) وقيمة P الاحتمالية 0.14400 وهى غير دالة احصائياً وعليه لا يوجد تحيز للنشر انظر قيمة P لمعامل الارتباط الرتبى 0.458.

الفصل الرابع عشر

كتابة تقرير نتائج ما وراء التحليل

Writing Meta-Analytic Results

مقدمة

بعد الانتهاء من دراسة ما وراء التحليل من تخطيط وبحث في التراث البحثي وتجميع الدراسات وتكويد خصائصها وحساب حجم التأثير ثم اجراء ما وراء التحليل من دمج او مقارنة احجام التأثير تاتي مرحلة عرض النتائج وتفسيرها في تقرير الدراسة حتى تكون قابلة للفهم والاستفادة منها من الباحثين الاخرين.

ولكن قبل عرض كتابة نتائج دراسة ما وراء التحليل فمن الضروري التاكيد على انه لا توجد طريقة واحدة صحيحة لكتابة نتائج الدراسة ولكن يمكن ان تختلف من باحث الى اخر بل من تخصص الى تخصص اخر، فعند كتابة النتائج لابد ان يكون في ذهنك لماذا اجريت دراسة ما وراء التحليل؟، وهذا يذكرنا باهداف ما وراء التحليل ومن اهم هذه الاهداف هو التكامل الكمي لنتائج الدراسات لظاهرة ما او متغيرين ما. وهذا التكامل ينبثق من هدفين فرعيين هما (Card, 2012):

الاول: احد مظاهر التكامل هو التعميم من الدراسات الاولية المختلفة المظاهر من حيث العينات والخصائص والمنهجيات والاجراءتها والملاحم الاخرى للدراسات الاولية للوصول الى استنتاج عام من هذه الدراسات وربما يكون التعميم غير جوهري فربما تختلف الدراسات عن بعضها البعض في بعض الخصائص وهذا ما يتم التحقق منه من خلال تحليلات التفاعل.

الثاني: التكامل بين الدراسات يتضمن محاولة لحل الصراعات او التناقضات بين النواتج المتضاربة، وهذا التعارض ياتي من تعارض نتائج اختبارات الفروض الصفرية وفي هذه الحالات فان ما وراء التحليل يركز على حساب متوسط حجم التأثير عبر الدراسات المختلفة وليس الوصول الى استنتاجات فيما يخص الدلالة الاحصائية. وفي حالات اخرى فان النتائج المتعارضة في الحقيقة ليست متعارضة انما نتيجة اخطاء المعاينة ويتم التحقق من ذلك من

خلال دراسة مدى اختلاف احجام التأثير لمعرفة ما اذا كانت النتائج متعارضة نتيجة خطأ المعاينة ام الى عوامل اخرى.

تفسير قيم حجم التأثير

احد الصعوبات فى فهم نتائج ما راء التحليل هو الطبيعة التفسيرية غير المحددة لاحصائيات حجم التأثير المقدره من الدراسات الاولية فتختلف معايير القيمة الصغيرة والمتوسط و الكبيرة لاحجام التأثير المختلفة، فلا توجد حدود معيارية متفق عليها بين الباحثين بل تختلف معايير التفسير من مؤشر الى اخر فمعايير تفسير مؤشرات مثل g, d تختلف عن معايير تفسير مؤشر Odds ratio ولكن يبدو الامر ليس بهذا الارتباك بالنسبة لمعايير تفسير معامل الارتباط فالامر اكثر استقراراً والفة من المعايير السابقة ولذلك يفضل بعض خبراء ما وراء التحليل تحويل كل مؤشرات حجم التأثير المختلفة الى معامل الارتباط لسهولة تفسيره ووجود معايير واضحة له.

ومن اهم حدود القطع لتفسير احجام التأثير فرق المتوسطات المعيارية هو الذى وضعة (Cohen, 1968, 1988) فى ابحاث العلوم السلوكية كالاتى:

حجم التأثير الصغير $ES \leq 0.20$

حجم التأثير المتوسط $ES = 0.50$

حجم التأثير الكبير $ES \geq 0.80$

بينما قيم حدود القطع لمعامل الارتباط التى تناظر حدود القطع او معايير مؤشر (Lipsey & Wilson, 2001):

حجم التأثير الصغير $r \leq 0.10$

حجم التأثير المتوسط $r = 0.25$

حجم التأثير الكبير $r \geq 0.40$

وقام Lipsey & Wilson (1993) بإنشاء توزيع لمتوسطات احجام التأثير لاكثر من 300 دراسة ما وراء التحليل للبرامج التدخلية النفسية، السلوكية، والتربوية وتم تقسيم توزيع احجام التأثير الى الارباعيات وكانت حدود القطع لمؤشرت فرق المتوسطات المعيارية مثل g, d كالآتى:

حجم التأثير الصغير $ES \leq 0.30$

حجم التأثير المتوسط $ES = 0.50$

حجم التأثير الكبير $ES \geq 0.67$

تنظيم عرض دراسة ما وراء التحليل

نتائج ما وراء التحليل البسيط هو عرض متوسط حجم التأثير ويمكن ان يكون مصحوباً ببعض تحليلات التفاعل لخصائص رئيسية فى عينة الدراسات المتضمنة فى ما وراء التحليل. و اشار Cooper (1988) الى ان عرض مراجعات التراث او الادبيات تنظم فى ثلاثة اجزاء كالتالى:

- **تاريخياً Historically:** عرض الدراسات او التقدم فى المجال عبر الزمن مثلاً من الاقدم الى الاحدث.
- **مفاهيمياً Conceptually:** عرض الدراسات فى ضوء خدمتها لفكرة عامة او قضية معينة دون الالتزام بالتسلسل الزمنى للنشأة.
- **منهجياً Methodological:** الدراسات المتشابهة فى المنهج او مداخل القياس يتم عرضها وتنظيمها معاً.

وتعرض نتائج دراسات ما وراء التحليل اما منهجياً او مفاهيمياً، فمثلاً مفاهيمياً عند دراسة العلاقة بين التحصيل وقلق الامتحان فيمكن اجراء دراسة ما وراء التحليل للفروق بين الجنسين فى التحصيل او فى قلق الامتحان او دراسة ما وراء التحليل للمتغيرين فى المرحلة الثانوية او الابتدائية، اما منهجياً يمكن اجراء ما وراء التحليل للدراسات الارتباطية واجراء ما وراء التحليل للدراسات التجريبية.

تعتبر دراسات ما وراء التحليل من الدراسات السابقة في مجال التخصص الذي يستفيد منها الباحثين في المجال ولا بد ان تكون المصطلحات والمفاهيم النظرية اكثر الفة للباحثين او الجمهور غير المتخصص في المجال مثل صانعي السياسات سواء التعليمية او التربوية او الطبية او الخبراء في المجالات المختلفة، وعليه عند كتابة تقرير دراسة ما وراء التحليل راعى فئة غير المتخصصين للاستفادة من نتائج الدراسة في النواحي التطبيقية العملية لكن هذا التبسيط او التوضيح لا يكون على حساب النواحي الفنية لاسلوب ما وراء التحليل، وهذا يتوقف على طبيعة المجلات العلمية التي تنشر الدراسة والقواعد التي تتبناها فربما تحد من قدرة الباحث في عرض المزيد من التوضيحات او الحسابات الفنية الخاصة بالاسلوب ولكن بقدر المكان راعى التوازن بين العرض باسهاب مفصل مما يدعو الى الملل والاختصار الذي يخفى كثير من المعانى.

محتوى تقرير دراسة ما وراء التحليل

توجد عدة اعتبارات عند كتابة نتائج تقرير ما وراء التحليل اهمها (Card, 2012):

- الاول: الالتزام بالاسلوب كتابة لاي دراسة في اي تخصص من حيث الطول النموذجي ومعايير الكتابة المنطق عليها وفقاً لقواعد الجمعية النفسية الامريكية.
- الثاني: كتابة تقرير دراسة ما وراء التحليل لا يختلف باى شكل من الاشكال عن كتابة تقرير الدراسات الاولية، فالهدف هو عرض امبريقي مستند الى الاسس النظرية الذي بدوره يسعى الى اضافة معرفة لها معنى للمجال او للموضوع الذي بصدد دراسته.
- وفيما يلي عرض لاهم مكونات تقرير ما وراء التحليل وهو لا يختلف عن مكونات تقارير البحوث الاجتماعية والنفسية (Card, 2012, Clarke, 2009):

العنوان Title

عنوان دراسة ما وراء التحليل يجب ان يكون دقيق وواضح ومحدد يظهر فيه الهدف من اجراء الدراسة ويتناول قضية بحثية للعلاقة بين متغيرين او مفهومين الذي من خلالهما سيتم حساب حجم التأثير ومن المفضل وضع الصيغة الاتية في العنوان:

"مراجعة ما وراء التحليل" او "مراجعة كمية" او "دراسة ما وراء التحليل للعلاقة بين..... " او "التوليف البحثي" لكى يبين للقارئ ان هذه دراسة ما وراء التحليل. فمثلاً وضع عنوان دراسة ما وراء التحليل كالاتى:

"العلاقة بين الامن النفسى والتنشئة الاسرية"

يعتقد الباحث ان هذه دراسة اولية يقوم فيها الباحث ببناء مقاييس وتطبيقها على عينة الافراد، ولكن محتوى الدراسة ليس كذلك انما هى دراسة ما وراء التحليل تقوم عينة من الدراسات السابقة التى تناولت دراسة المتغيرين. وعليه من الافضل صياغة العنوان كالاتى:

"ما وراء التحليل للعلاقة بين الامن النفسى والتنشئة الاسرية"

او "العلاقة بين الامن النفسى والتنشئة الاسرية: ما وراء التحليل"

"التوليف البحثي لاثر العلاج النفسى المعرفى لتخفيف قلق الامتحان لدى طلاب الصف الثالث الثانوي المعرضين للضغوط الاكاديمية"

المقدمة Introduction

هى نفسها كما فى الدراسات الاولية لابد من عرض الهدف من الدراسة واعطاء خلفية للنظرية والدراسات السابقة والدافع وراء انجاز الدراسة و عرض اسئلة البحث او الفروض البحثية لدراسة ما وراء التحليل، وكذلك خصائص الدراسات المتضمنة فى التحليل التى يوليها الباحث قدراً من الاهتمام والدراسة، وكيفية واليات عمل المعالجة المستخدمة.

وفى ضوء معايير انتقاء الدراسات السابقة فى التحليل لابد من عرض مجتمعات الدراسات، عيناتها، مقاييسها، تصميمها، واى خصائص اخرى قد تلعب دوراً فى نتائج ما وراء التحليل لان هذه المتغيرات قد يكون لها دوراً فى اجراء التحليلات المتفاعلة.

الطريقة او المنهجية المستخدمة فى ما وراء التحليل تكون متفردة او مختلفة عن نظيرتها فى الدراسات الاولية ويحدد (2012) Card اربعة ملامح عامة للمنهجية فى دراسة ما وراء التحليل كالاتى:

- **اجراءات او استراتيجيات الحصول على الادبيات البحثية Literature Search**
Procedures: جودة دراسة ما وراء التحليل تتاثر بمدى تضمين عدد كبير من الدراسات المتضمنة فى التحليل ومدى تمثيلها لمجتمع الدراسات التى تناولت العلاقة بين المتغيرين ومن الضرورى وصف الاجراءات التى استخدمها الباحث فى الحصول على الدراسات السابقة والمصادر المختلفة التى اعتمد عليها الباحث سواء قواعد البيانات او المحركات البحثية او البحث اليدوى فى المكتبات، ويجب عرض الكلمات المفاتيحية المستخدمة والعمليات الاجرائية فى كتابة الكلمات المفاتيحية مثل and او or وكذلك عرض تاريخ اخر بحث فى قواعد البيانات سواء فى المجالات او المؤتمرات ويجب اعطاء تبرير لعملية البحث هذه للحصول على الدراسات المنشورة وايضاً الوسائل التى اعتمد عليها الباحث فى الحصول على الدراسات غير المنشورة.
- **معايير انتقاء الدراسات فى التحليل:** جودة نتائج واستنتاجات ما وراء التحليل يعتمد فى المقام الاول على الدراسات التى تمثل مجتمع التراث البحثى، فيجب تحديد مجتمع الدراسات سواء طلاب مرحلة معينة او مجتمع فئة عمرية معينة مثل الطفولة او المراهقة وغيرها واعتبارها معيار لتضمين او استبعاد الدراسات فى التحليل، ويجب تحديد معايير التضمين قبل تكويد الدراسات، وقد يتبنى الباحث معايير مثل الدراسات فى فترة زمنية معينة او الدراسات التى استخدمت منهجية معينة او الدراسات التى استخدمت تصميم تجريبى معين او الدراسات المنشورة او الدراسات التى تتميز بخصائص سيكومترية معينة مثل الصدق والثبات، وهذه المعايير تحدد اطار المعاينة للدراسات.
- **تكويد خصائص الدراسات وحساب حجم التأثير:** فى ضوء معايير التضمين او الاستبعاد، يبدأ الباحث فى تكويد خصائص الدراسات التى تشكل عينة الدراسات

المتضمنة فى التحليل ومن الضرورى وصف عملية التكويد للقارئ. وتوجد ثلاثة مظاهر لعملية تكويد الدراسات يجب وصفها بالتفصيل هى:

- **تكويد خصائص الدراسات:** يجب تكويد اكبر قدر من خصائص الدراسات التى بحاجة اليها والتى ليس بحاجة اليها ويجب اعطاء وصف تفصيلى لخصائص الدراسات، وفى هذا الجزء من الدراسة يجب اعطاء منطقية لتضمين هذه الخصائص. وعلى الباحث عمل قائمة لهذه الخصائص وتعريفها وتحديد القيم او التكويدات التى تأخذها مثل 1,2,3,4.... الخ. فمثلاً الجنس كمتغير تصنيفى يكود الذكر=1 ، الانثى=2 ، او العمر كمتغير متصل يعبر عنه فى ضوء متوسط عمر العينة بالسنوات او اداة القياس مثل الملاحظة =1، المقابلة=2 ، الاستبيان=3 او غيرها.

مع وجود العديد من خصائص الدراسات التى يجب تكويدها فانها تحتاج الى مساحة كبيرة فى الدراسة ولذلك من المفضل عرض هذه التفاصيل فى جدول فى ملاحق الدراسة، ويجب وصف تكويد احجام التأثير وكذلك تصحيحها من الاخطاء المنهجية. وفى ضوء تكويد احجام التأثير يجب الاجابة عن ثلاثة اسئلة كالتالى (Card, 2012) :

الاول: كيف لاشارة حجم التأثير ان تحدد اتجة النتائج؟، على سبيل المثال ما وراء التحليل للفروق بين الجنسين من الضرورى تحديد ما اذا كانت احجام التأثير الموجبة تشير الى الدرجات المرتفعة للذكور او للاناث.

الثانى: مانوع حجم التأثير المستخدم ولماذا؟، فلو اسخدمت حجم التأثير المعيارى (OR, g, d, r) فيجب الاشارة اليه حتى يكون القارئ على معرفة بنوع حجم التأثير المستخدم.

الثالث: كيف قدرت الطرق المختلفة لتقدير احجام التأثير فى التراث للحصول على حجم التأثير العام؟.

ويجب عرض كيفية حساب ثبات المكودين سواء عبر المكودين مختلفين او نفس المكود عبر نفسه كما تم عرضه فى الفصل الرابع، واذا كان ثبات المكودين منخفض فمن الافضل ادخال تعديلات على بروتوكول التكويد وهذا يمثل نوعاً من الشفافية.

استراتيجية تحليل البيانات Data-Analytic strategy

لان اسلوب ما وراء التحليل غير مألوف لكثير من الباحثين فى كافة التخصصات ونتيجة للاستراتيجيات المختلفة لتنفيذ هذا الاسلوب فمن الضرورى عرض الاستراتيجيات التى يعتمد عليها الباحث فى حسابات هذا الاسلوب ويجب عرضها فى نهاية جزء المنهجية، وفى بعض الاحيان فى بداية عرض جزء النتائج ويوجد على الاقل خمس عناصر اساسية يجب تحديدها فى جزء استراتيجية تحليل البيانات هى كالتالى (Card, 2012):

الاول: كيفية حساب احجام التأثير من الدراسات.

الثانى: تحديد الاوزان المستخدمة للدراسات (مقلوب مربع الخطأ المعيارى).

الثالث: وصف عملية تحليل متوسط احجام التأثير بمعنى تحديد نموذج التحليل سواء نموذج التأثيرات المثبتة او العشوائى فى ضوء نتائج اختبار التجانس او الاختلاف بين احجام التأثير.

الرابع: وصف عملية وطريقة التحليلات المتفاعلة خصوصاً:

(1): قرار اجراء التحليلات المتفاعلة فى ضوء عدم التجانس بين احجام التأثير.

(2): تقييم تفاعلات عديدة فى ضوء متغيرات عديدة فى نفس الوقت او بصورة متتابعة.

(3): لو اجريت تحليل التفاعلات المتسلسل او التتابعى لا بد تحديد متى نتوقف عن التحليل فى ضوء اختبارات الاختلاف البواقى.

(4): تحديد مدخل تحليل التفاعلات فى ضوء تحليل التباين او ما وراء الانحدار.

كذلك يجرى الباحث اختبارات تشخيص تحيز النشر وكذلك تحليل الحساسية وذلك باستبعاد الدراسات غير المنشورة وحساب متوسط حجم التأثير او استبعاد الدراسات منخفضة الجودة وحساب حجم التأثير، وايضاً حساب القوة الاحصائية للاختبارات المتضمنة فى ما وراء التحليل كلما امكن.

النتائج Results

هذا الجزء يتضمن بعض المعلومات المختلفة عن عرض نتائج الدراسات الاولية وفي هذا الجزء يتضمن اربعة اجزاء من المعلومات الضرورية كالتالى (Card, 2012):

الاول: المعلومات الوصفية: وصف عينة الدراسات المتضمنة فى ما وراء التحليل وهذه المعلومات تلخص فى جدول مثل عدد الدراسات K وعدد المشاركين (العينة) N فى كل دراسة وايضاً وصف الدراسات فى ضوء خصائصها التصنيفية مثل (تجريبية المجموعة الواحدة - المجموعتين)، وفى ضوء المتغيرات المتصلة مثل المتوسط والانحراف المعياري لعمر العينة وهذه الخصائص ضرورية اذا كان هدف الباحث اجراء تحليلات المتفاعلة ويمكن وصف المقاييس المستخدم والمنهجيات وكذلك يمكن ان يعلق الباحث على مجمل هذه الخصائص وغيرها ومقارنتها بدراسات وراء التحليل الاخرى فى المجال.

الثانى: مقياس النزعة المركزية واختبار الاختلاف: الخطوة المهمة فى عرض النتائج هى اختيار مقياس النزعة المركزية المناسب الذى غالباً يكون مؤشر المتوسط، والخطأ المعياري لهذا المتوسط و 95% فترات الثقة حول متوسط حجم التأثير وعرض الدلالة الاحصائية حول متوسط حجم التأثير ويجب اعطاء تفسير لمؤشرات المتوسط وفترات الثقة ولا يتم عرضها فى صورة رقمية فقط. واذا كان حجم التأثير فى صورة معامل الارتباط يجب تحديد طبيعة العلاقة سواء موجبة او سالبة وكذلك تفسير حجم التأثير فى ضوء المعايير المتفق عليها.

بالاضافة لمتوسط حجم التأثير فمن الضرورى عرض اختبار التجانس بين احجام التأثير Q ودلالته الاحصائية لمعرفة مدى وجود تجانس او اختلاف بين احجام التأثير وكما نعلم ان نتائج اختبار التجانس يحدد نوع النموذج المستخدم فى ما وراء التحليل سواء مثبت او عشوائى او اجراء التحليلات المتفاعلة.

الثالث: التحليلات المتفاعلة: لو اجرى الباحث بعض التحليلات المتفاعلة لبعض خصائص الدراسات المسببة فى اختلافات بين احجام التأثير فيجب عرض نتائج هذه التحليلات كاملة مثل عرض اختبار Q ودلالته الاحصائية ودرجات الحرية لكل متغير متفاعل وعرض نتائج

ANOVA او تحليل الانحدار ويجب عرض اختبار Q للاختلاف داخل المجموعات او اختلاف البواقي واذا كان المتغير المتفاعل باكثر من مستويين فيجب تحديد اختبارات المقارنات البعدية لتحديد مسبب دلالة Q .

وعند اجراء تحليل التفاعل لمتغير تصنيفي يجب تحديد عدد الدراسات في كل تصنيف من تصنيفات المتغير وكذلك متوسط حجم التأثير لدراسات كل تصنيف.

الرابع: التحليلات التشخيصية Diagnostic Analysis: كما سبق عرضه ان من اهم المهددات لما راء التحليل قضية تحيز النشر التي تؤدي الى تحيز نتائج ما وراء التحليل، فمن الضروري عرض ادلة لتقييم مدى وجود هذه القضية من عدمه من خلال اجراء التحليلات المتفاعلة او عرض شكل القمع او تحليلات الانحدار وذلك لاعطاء صورة صادقة للقارئ عن طبيعة النتائج وعدم تحيزها.

الخامس : تحليل الحساسية Sensitivity analysis: يتضمن اذا وجدت قيم متطرفة لاحجام التأثير فبعد حذفها من التحليل فما هي طبيعة الاستنتاجات هل تتغير ام لا؟. وايضاً يهتم تحليل الحساسية بمدى تأثير الطرق الاحصائية المستخدمة على طبيعة الاستنتاجات من التحليل بمعنى هل تختلف النتائج باختلاف مؤشر حجم التأثير المستخدم مثلاً مؤشر نسبة المخاطرة RR في مقابل نسبة الاحتمال OR او استخدام مؤشر كوهين d في مقابل مؤشر هيدجز g او هل تختلف النتائج باختلاف نموذج التحليل سواء نموذج التأثيرات المثبتة او نموذج التأثيرات العشوائية؟ او هل تختلف النتائج اذا حدث تصحيح لاحجام التأثير من المحددات المنهجية عنها بدون تصحيح؟. ويهتم تحليل الحساسية بالبيانات الغائبة سواء لحجم التأثير او الخصائص المنهجية التي تستخدم كمتغيرات متفاعلة هل تحدث تحيز للنتائج مثل تحيز النشر؟، ايضاً التحليل التراكمي هو الى يهتم بمدى التغير الحادث لحجم التأثير وتباينها باضافة اي حجم تأثير الى الدراسة فهذا يعتبر جزء من تحليل الحساسية.

المناقشة Discussion

فى هذا الجزء يعرض الباحث نتائج ما وراء التحليل فى اطار التراث البحثى للظاهرة التى تناولها الباحث بالتحليل ويجب الابتعاد عن لغة الارقام او الاقلال منها بقدر الامكان وفى هذا الجزء يجب ان يتضمن اربعة مكونات كالتالى (Card, 2012):

الاول: مراجعة النتائج: على الرغم من تجنب تكرار عرض النتائج الا انه فى بعض الاحيان يمكن اعطاء تلخيص للنتائج الرئيسية خاصة اذا كان جزء عرض النتائج طويل وفنى ومعقد ويتضمن حسابات تفصيلية.

الثانى: تفسير وتضمين النتائج Explanations and Implications of Results :

يجب ان تضع فى الاعتبار ان ما وراء التحليل اسلوب يستخدم للاجابة على اسئلة بحثية فمن المفترض ان تنصب المناقشة على تفسير النتائج لاعطاء اجابات عن الاسئلة البحثية وكيف اسهمت هذه النتائج فى زيادة فهم الظاهرة فى المجال، وهل دعمت النتائج النظرية او التراث البحثى ام تعارضت مع النظرية ام ما زالت الحاجة ضرورية لاجراء المزيد من البحث لتتقح النظرية؟، وهل افادت النتائج واضع السياسات وصانعى القرار فى المجال؟.

الثالث: المحددات Limitation: لاتتنسى فى هذا الجزء ان تكون اميناً فى عرضك للمحددات Limits التى واجهتك فى اجراء الدراسة حتى يكون الباحثين الاخرين على وعى بها عند اجراء دراسات مماثلة فى المجال. واذا لم تتمكن من الوصول الى توصيات قاطعة جازمة بسبب اما نقص القوة الاحصائية للاسلوب او نقص حجم عينة الدراسات او اى محدثات اخرى فى الدراسة فلا بد الى ان توصى بالحاجة الى اجراء المزيد من الدراسات للتحقق من صدق النتائج والقدرة على تعميمها.

من الضرورى عرض محدثات او الصعوبات التى واجهتك عند اجراء دراسة ما وراء التحليل بعض هذه المحددات هى قصورات Shortcomings يجب ذكرها بوضوح لان هذا لا يقلل من العمل البحثى انما يذيد الشفافية والموضوعية، ومن اهم هذه المحددات فى ما وراء التحليل صغر حجم العينة والوصول الى استنتاجات عن طبيعة التحليلات المتفاعلة كذلك وجود تحيز النشر فى الدراسة وكذلك نوعية الدراسات ومدى تمثيلها لمجتمع الدراسات فى التراث البحثى للمتغيرين موضع الدراسة ويجب طرح الية ممكنة للدراسات المستقبلية لمحاولة تجنبها اذا كان هذا متاح.

الخامس: الاستنتاجات Conclusions: فى نهاية تقرير البحث يجب اعطاء التأثير العالى للبحث فى المجال ويجب صياغة عبارة واضحة لكيفية احداث دراسة ما وراء التحليل نمواً فى فهمنا للظاهرة.

المراجع References

كأى بحث، فان دراسة ما وراء التحليل تتضمن قائمة من المراجع. وفى هذا الشأن يجب عرض الدراسات المتضمنة فى ما وراء التحليل ويجب ان تذكر عبارة قبل عرض المراجع هى ان المراجع التى تأخذ رمز نجمة(*) هى الدراسات السابقة التى تضمنت فى عينة دراسة ما وراء التحليل.

الملاحق Appendices

تضع المجلات العلمية معايير مختلفة لمحتوى الدراسات التى بصدد نشرها اما ان توضع فى متن الدراسة او توضع فى قائمة الملاحق فى نهاية المقال او الدراسة او يمكن ان تكون متاحة الكترونياً عبر موقع على الانترنت. ونظراً للمساحة المحددة للنشر من قبل المجلة فيفضل ان توضع الجداول التلخيصية لتكويد الدراسات المتضمنة فى التحليل فى الملاحق وهذا يتيح للقارئ الحكم على مدى شمولية وموثوقية عملية التكويد كوسيلة لاستخلاص المعلومات من الدراسات.

استخدام الجداول والاشكال فى تقرير ما وراء التحليل

تمدنا الجداول والاشكال بقدر كبير من البيانات التى يمكن استخلاص منها المعلومات المفيدة، ويعرض Card (2012) بعض المداخل لعرض نتائج ما وراء التحليل فى الجداول والاشكال:

الجداول Tables: يوجد نوعين من الجداول لتلخيص نتائج ما وراء التحليل: جداول لعرض المعلومات التلخيصية مثل متوسط حجم التأثير، وجداول تلخص المظاهر المكودة للدراسات المفردة المتضمنة فى ما وراء التحليل وهذين النوعين من الجداول هما الاكثر شيوعاً فى عرض نتائج ما وراء التحليل، وهذا ما اتضح من خلال مسح (Borman & Grigg 2009) حيث ان هذه الجداول متضمنة فى من 74% الى 89% من دراسات ما وراء التحليل المنشورة.

الاول: الجداول التلخيصية **Summary Tables**: تستخدم لعرض المعلومات المتجمعة المتحصل عليها جراء دمج او مقارنة الدراسات المتعددة، وهذه المعلومات تتضمن مؤشرات النزعة المركزية مثل المتوسط والوسيط، وتوزيع احجام التأثير مثل المدى، الانحراف المعياري، اختبار الاختلاف Q ، مؤشرات الاختلاف I^2 ، ونتائج التحليلات المتفاعلة. ولو ان الدراسة تضمنت تحليل وحيد لما وراء التحليل فيمكن عرض هذه المعلومات في عبارة تقريرية لان عرضها في جدول يكون ضيق.

مثال: فيما يلي الجدول التلخيصي لعرض نتائج مثال البيانات المعروض في الكتاب لنتائج 22 دراسة للعلاقة بين العدوان ورفض الاقران (في: Card, 2012):

الجدول (1.14): النتائج التلخيصية لدراسة ما وراء التحليل لـ 22 دراسة.

Heterogeneity (Q ; $df = 21$)	291.17***
I^2	92.8%
Random-effects mean r	.33***
95% confidence interval	.24 to .40
Moderator effects	
Reporter ($Q_{b(3)}$)	65.14***
Observation ($k = 3$)	.10 _a
Parent ($k = 1$)	.52*** _c
Peer ($k = 3$)	.38*** _b
Self ($k = 0$)	—
Teacher ($k = 6$)	.19*** _a
Age ($Q_{\text{regression}(1)}$)	9.31**
β_{Zr}	-.011
\hat{r} at 5, 10, and 15 years	.41, .37, .22

* $p < .05$; ** $p < .01$; *** $p < .001$.

الصفين الاولين يعرضا اختبار التجانس ودلالته الاحصائية ومؤشر I^2 لتكميم مقدار الاختلاف بين احجام التأثير اما الصفين الثالث والرابع يعرضا متوسط حجم التأثير للتأثيرات العشوائية مع دلالتها الاحصائية وفترات الثقة حول المتوسط. اما الاعمدة التالية يعرضوا نتائج التحليلات المتفاعلة لمتغيرين من خصائص الدراسة احدهما تصنيفي هي طبيعة قياس العدوان والاخر متصل هو عمر عينة للدراسات. وبعد عرض اختبار الاختلاف بين التصنيفات الاربعة

$Q_{Between}$ للمتغير التصنيفي والعلاقة الارتباطية بين المتغيرين في دراسات كل تصنيف من التصنيفات الاربعة ودلالاتها الاحصائية وعرض $Q_{Regression}$ للمتغير المتصل العمر ودلالته الاحصائية ومعامل الانحدار اللامعياري (BzF) لانحدار حجم التأثير على متغير العمر والارتباطات المتنبأ بها على قيم مختلفة للمتغير المتفاعل (العمر). ويمكن ان يتسع هذا الجدول اذا تضمن التحليل متغيرات متفاعلة اخرى مثل مقارنة بين الدراسات المنشورة وغير المنشورة لتقييم تحيز النشر.

الثاني: جداول الدراسات المفردة Tables of Individual Studies: من الافضل ان يمد الباحث القارئ بقائمة مفصلة من التكويدات لكل دراسة متضمنة في ما وراء التحليل ويتضمن هذا الجدول اسم المؤلف، تاريخ النشر، حجم العينة، الخصائص المنهجية (نوعية المنهج، نوعية التصميم، ثبات المقاييس، وغيرها)، وحجم التأثير في الدراسة وتعديل او تصحيح حجم التأثير من الاخطاء المنهجية. وتعرض الدراسات اما حسب الترتيب الابددي او الترتيب الزمني بتاريخ النشر للمقال او الدراسة.

مثال: فيما يلي جدول الدراسات المفردة لـ 22 دراسة للعلاقة بين العدوان ورفض الاقران (Card, 2012):

الجدول (2.14): البيانات التفصيلية لدراسة ما وراء التحليل لـ 22 دراسة.

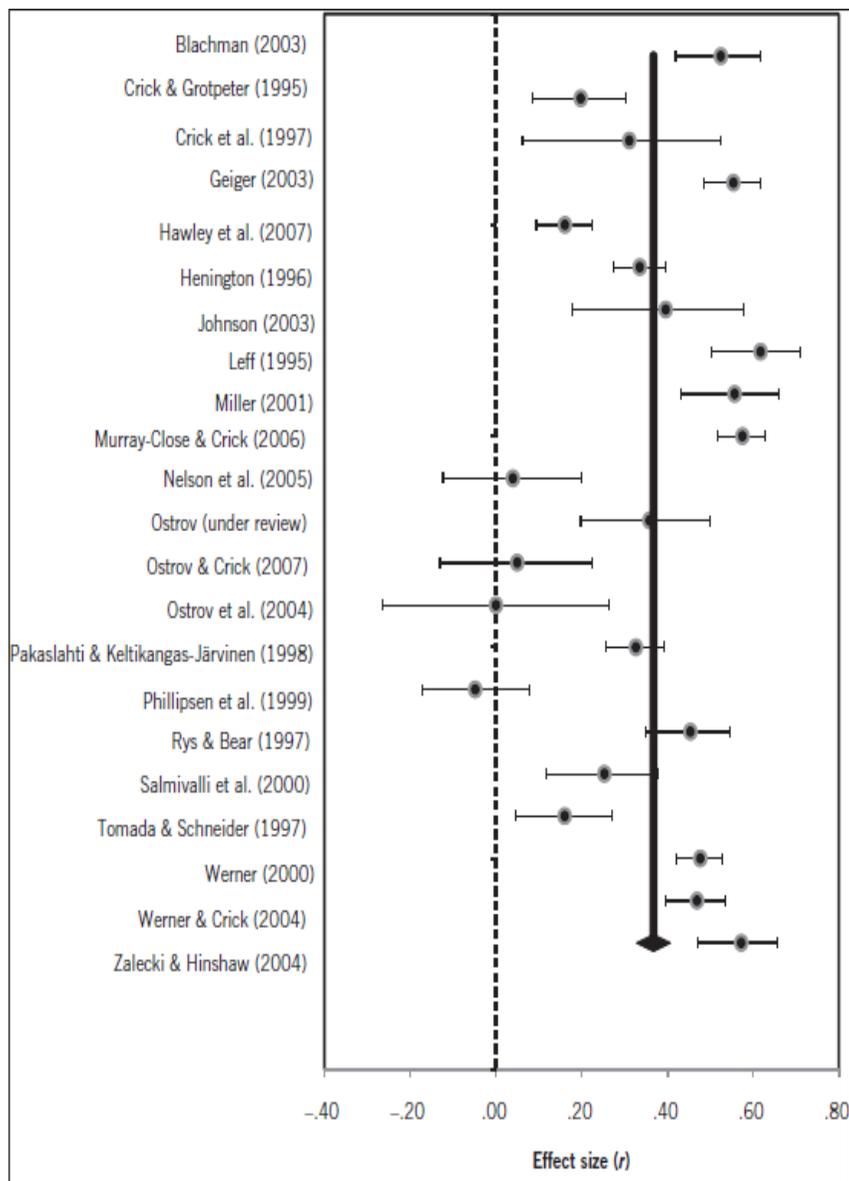
Study	Sample size (N)	Method	Age (years)	Effect size (r)
Ostrov (under review) ^a	139	Obs.	3.6	.181
Ostrov & Crick (2007)	132	Obs.	4.1	.049
Ostrov et al. (2004) ^b	60	Obs.	4.6	.000
Blachman (2003)	228	Parent	9.2	.525
Crick et al. (1997)	65	Peer	4.5	.443
Nelson et al. (2005)	180	Peer	4.8	.090
Johnson (2003)	74	Peer	6.0	.531
Henington (1996)	904	Peer	7.5	.336
Geiger (2003)	458	Peer	8.0	.554
Werner (2000)	881	Peer	8.0	.477
Werner & Crick (2004)	517	Peer	8.0	.469
Phillipsen et al. (1999)	262	Peer	8.7	-.048
Murray-Close & Crick (2006)	590	Peer	9.0	.575
Tomada & Schneider (1997)	314	Peer	9.0	.313
Crick & Grotmeter (1995)	491	Peer	9.4	.198
Leff (1995)	151	Peer	9.5	.617
Rys & Bear (1997)	266	Peer	9.5	.556
Pakaslahti & Keltikangas-Järvinen (1998)	839	Peer	14.5	.326
Hawley et al. (2007)	929	Peer	14.7	.161
Salmivalli et al. (2000)	209	Peer	15.5	.253
Miller (2001)	150	Peer	16.0	.557
Ostrov (under review) ^a	139	Teacher	3.6	.513
Crick et al. (1997)	65	Teacher	4.5	.167
Nelson et al. (2005)	180	Teacher	4.8	-.011
Johnson (2003)	74	Teacher	6.0	.074
Tomada & Schneider (1997)	314	Teacher	9.0	.000 ^b
Rys & Bear (1997)	266	Teacher	9.5	.338

والصف في هذا الجدول يمثل الدراسة بينما العمود يمثل خاصية الدراسة وفي هذا الجدول يوجد متغيرين متفاعلين أحدهما تصنيفي طريقة Method قياس العدوان أما ملاحظة Observation أو الاقران Peer أو المعلم Teacher أو الاباء Parent والآخر متصل عمر العينة.

استخدام الأشكال لعرض نتائج ما وراء التحليل مفيد جداً لعرض قدر كبير من المعلومات مثل المؤشرات التلخيصية مثل المتوسط أو الانحراف المعياري والتأثيرات المتفاعلة وتحيز النشر. ويطرح (Card 2012) ثلاثة أنواع من الأشكال لعرض نتائج ما وراء التحليل كالآتي:

إرسومات أو اشكال الغابة Forest Plots او شكل عرض فترات الثقة

يسمح للباحث ان يرى الغابة او الخشب والاشجار هي نادراً ما تستخدم في العلوم الاجتماعية ولكنها شائعة الاستخدام في دراسات ما وراء التحليل في العلوم الطبية (Borman & Grigg, 2009)، ومثال لهذا العرض كالتالي:



الشكل (1.14): مثال للعرض البياني Forest لـ 22 دراسة.

وفي هذا العرض توضع الدراسات الـ 22 على يسار الشكل بينما على يمين الشكل تعرض المعلومات حول المتوسط (الدوائر السوداء)، و 95% فترات الثقة (الخطوط الأفقية) لكل دراسة في ما وراء التحليل حيث ان فترات الثقة الضيقة تعنى تقدير دقيق لحجم التأثير في الدراسة، اما الخط الرأسى السميك | يمثل المتوسط الموزون لاحجام التأثير، اما الخط الرأسى المنقوط ||| يشير الى النتيجة الصفرية بمعنى ان معامل الارتباط $r = 0.0$ لتوضيح الدراسات التي لها تاثيرات دالة احصائية.

وهذا التمثيل البياني يقدم قدراً من المعلومات تتضمن:

الأول: عرض المعلومات الخاصة بتقديرات احجام التأثير لكل دراسة فى ما وراء التحليل وهذا مفيد ومتشابهه مع الجداول التى سبق عرضها.

الثانى: عرض الخط الراسى لمتوسط حجم التأثير يظهر مدى ابتعاد او اقتراب احجام التأثير للدراسات من المتوسط الموزون ويظهر احجام التأثير المتطرفة ومقدار الدقة ففترات الثقة الضيقة او القصيرة تعنى دقة افضل، اى ان الدقة للدراسات تقاس بمدى طول او قصر الخط الافقى الدال على فترة الثقة.

الثالث: يقدم معلومات بصرية فيما يخص مدى اختلاف او تجانس الدراسات واذا كانت فترات الثقة لا تتضمن متوسط حجم التأثير(الخط الراسى) فهذا يقدم دليل بصرى لدلالة اختبار الاختلاف ومدى احجام التأثير للدراسات حول الخط الراسى يقدم دليل على عدم تجانس فى احجام التأثير، وهذا الشكل مفيد فى توضيح الدراسات التى لها احجام تأثير متطرفة(البعيدة على يسار او يمين الدراسات الاخرى مع فترات ثقة لا تتناسب مع بقية الدراسات).

رسومات الساق والورقة **Steam and Leaf Plots**

هذا الشكل يستخدم بصورة متزايدة فى الدراسات والبحوث لعرض معلومات عن النزعة المركزية والتشتت وخصائص توزيع احجام التأثير مثل التفرطح والالتواء وكذلك يعتبر وسيلة هامة لتشخيص القيم المتطرفة لاحجام التأثير، ويتكون هذا الشكل من جزئين هما الساق على يسار الخط الراسى ويتضمن مدى احجام التأثير كما فى الشكل الاتى لـ 22 دراسة متضمنة فى دراسة ما وراء التحليل (Card (2012):

.7	
.6	2
.5	35677
.4	0578
.3	1346
.2	05
.1	66
.0	450
-0	5
-1	

الشكل(2.14): شكل الورقة والساق لـ22 دراسة للعلاقة بين العدوان ورفض الاقران.

والواضح ان الساق يتضمن كل احجام التأثير الممكنة حيث تبدأ بالرقم العشري 1.- اسفل الساق انتهاء بـ 70. حيث امام 5. يوجد القيم 35677 وهذا يعنى وجود خمسة احجام تاثير لخمسة دراسات هي 53، 55، 56، 57، 57. ، بينما يوجد حجم تاثير وحيد هو 05.- وهكذا. وهذا يعنى ان الاعداد على يسار الخط الراسى تمثل الساق بينما الاعداد على يمين الخط الراسى تمثل الاوراق حيث كل رقم عشرى يعرض دراسة وحيدة وان اعلى معامل ارتباط بين العدوان ورفض الاقران هو 62.. لاحظ ان عرض الاوراق يتم من الاصغر الى الاكبر مثل 35667.

وهذا الشكل يعطى معلومات مفيدة مثل عدد الدراسات المتضمنة فى التحليل حيث يعرض على يسار الشكل التكرارات للدراسات او احجام التأثير وان احجام التأثير تكون فى المدى من 0.50 الى 0.59 اكبر من عدد الدراسات فى المدى من 0.60 الى 69. ومن الشكل يمكن الاستنتاج بمدى وجود تجانس او اختلاف بين احجام التأثير وكذلك يمدنا بمعلومات عن شكل التوزيع لاحجام التأثير سواء اعتدالى او ملتوى او متفرطح، ففى الشكل السابق يبدو ان التوزيع ملتوى ناحية القيم السالبة او ناحية اليسار.

ويعطى شكل الساق والورقة تشخيص للقيم المتطرفة ولم يعطى فى هذا المثال قيم متطرفة مما يدل على عدم وجود قيم متطرفة. وأشار (Borman & Grigg (2009 الى ان 30% من دراسات ما وراء التحليل تعرض شكل الورقة والساق فى نتائجها.

شكل القمع

كما سبق عرضه فان هذا الشكل يستخدم لتشخيص تحيز عينة الدراسات او النشر وهو تمثيل بيانى يوضع حجم العينة على احد المحاور السينى او الصادى بينما حجم التأثير على المحور الاخر. وفقاً للمسح التقييمى الذى قام به (Borman & Grigg (2009 فان 12.5% من الدراسات المنشورة تعرض هذا الشكل فى عرضها للنتائج.

وقدم (Higgins & Green (2006 اطاراً يتضمن العناصر المتضمنة فى تقرير دراسة ما وراء التحليل كالاتى:

1. الغلاف: العنوان، تفاصيل اخرى، وعنوان التواصل (E-mail).

2. المستخلص يتضمن: الخلفية النظرية، الاهداف، استراتيجية البحث، محكات الانتقاء او التضمين، جمع البيانات والتحليل، النتائج الرئيسية، والاستنتاجات.

3. المقدمة او الخلفية النظرية وتتضمن:

1.3. وصف المشكلة واهميتها.

2.3. وصف البرنامج المراد تقييم فعاليته ودوره فى الممارسة الحالية.

3.3. كيفية عمل البرنامج اوالنموذج المنطقى له.

3.3. اعطاء مبررات لاجراء دراسة ما وراء التحليل.

4. اهداف الدراسة.

5. الطريقة او المنهجية وتتضمن:

1.5. معايير تضمين الدراسات فى التحليل (نوع الدراسات، نوع العينة وعمرها، نوع التدخلات(المجموعة الضابطة والتجريبية، المجموعة الواحدة)، وغيرها.

2.5. استراتيجية البحث فى التراث البحثى (قواعد البيانات، التقارير الحكومية والمؤسسية، المؤتمرات، البحث اليدوى، المحركات البحثية العالمية مثل Google ، الاتصال الشخصى، ومستخلصات الدراسات، وغيرها).

3.5. جمع البيانات وطرق التحليل: وتتضمن:

- الدراسات المختارة فى ضوء معايير الانتقاء وعدد المقدرين او المكودين وثبات المقدرين.

- واستخلاص البيانات من الدراسات وادارتها وكيفية انشاء قاعدة البيانات وشكل جدول استخلاص البيانات وعدد المقدرين وطريقة معالجة البيانات.

- تقدير الجودة المنهجية للدراسات المتضمنة فى التحليل مثل المنهجية المستخدمة.

- مقاييس حجم التأثير المناسبة لطبيعة المتغيرات سواء بيانات تصنيفية (RD, RR, OR) او بيانات متصلة.

- وحدة التحليل: هل التصميم يتضمن مجموعات عديدة او المجموعة الضابطة والتجريبية.

- كيفية التعامل مع البيانات الغائبة.

- تقدير الاختلاف بين احجام التأثير سواء الاختلاف الاكلنيكي او الاختلاف الاحصائي.

- تقدير تحيز التقرير او تحيز النشر من خلال الاختبارات او شكل القمع.

- اختيار نموذج التحليل سواء المثبت او العشوائي.

- اجراء التحليلات المتفاعلة فى ضوء بعض الخصائص للدراسات وذلك اذا وجد اختلاف بين احجام التأثير.

- تحليل الحساسية.

6. النتائج: وتتضمن:

- الدراسات المستبعدة واسباب الاستبعاد.

- عرض ووصف للدراسات المتضمنة مثل حجم العينة وتاريخ النشر والتصميم المستخدم وغيرها.

- عرض متوسط حجم التأثير ودلالته الاحصائية وفترات الثقة حول المتوسط.

- عرض اختبارات لتشخيص الاختلاف بين احجام التأثير.

- التحليلات المتفاعلة سواء باستخدام ANOVA او الانحدار.

- عرض الرسومات البيانية خاصة شكل القمع لتشخيص التحيز.

7. المناقشة: يتضمن هذا الجزء:

- عرض ملخص للنتائج.

- التفسيرات والتناقضات والاتفاق مع الدراسات السابقة. -

8. الاستنتاجات: تتضمن توصيات للممارسة وتوصيات بحثية مقترحة.

9. محددات الدراسة ان وجدت.

10. المراجع: تتضمن:

- الدراسات المتضمنة في التحليل.

- الدراسات المستبعدة.

- الدراسات التي استعان بها الباحث في كتابة تقرير البحث.

11. الجداول: خصائص الدراسات المتضمنة والدراسات المستبعدة.

12. الاشكال: مثل Forest وشكل القمع.

ومن اهم مهدهدات الاستدلالات من ما وراء التحليل (Matt & Cook 2009):

- الدراسات ذات احجام العينات الصغيرة حيث تؤدي انخفاض القوة الاحصائية ومن شأنه عدم الحصول على حجم التأثير الحقيقي.
- انخفاض الثبات في الدراسات الاولية.
- المدى الضيق للمتغيرات المتضمنة في التحليل حيث تؤدي الى تقلص حجم التأثير.
- غياب احجام تأثير في التحليل نتيجة عدم تمثيل عينة ما وراء التحليل لمجتمع الدراسات.
- عدم دقة او عدم ثبات عملية التكويد.
- تحيز عينة احجام التأثير.
- تحيز النشر.
- تحيز في حساب احجام التأثير.
- نقص الاستقلال الاحصائي بين احجام التأثير.
- عدم اجراء موازنة دقيقة لاحجام التأثير.

- استخدام غير صحيح لنموذج التحليل سواء نموذج تأثيرات مثبتة او نموذج تأثيرات عشوائية.

وقدم (Pigott 2012) عدة توصيات لزيادة جودة دراسة ما وراء التحليل كالتالى:

- **الرؤية الواضحة لمشكلة البحث والاسلوب:** يتطلب اسلوب ما وراء التحليل فهم عميق لطبيعة الاسئلة المراد الاجابة عليها لان هذا الاسلوب يتسم بالحدائثة الى حداً ما فى تطبيقه فى الدراسات، وكذلك فهم لطبيعة الاسلوب المستخدم ومواصفات الدراسات المتضمنة وحجم عينة الدراسات المناسب للتحليل واختيار نوعية النموذج المستخدم وكيفية تفسير الفروق بين الدراسات وكيفية تفسير النتائج ومدى الاستفادة منها فى تحسين الممارسات البحثية واستخدامها لصانعى السياسات المختلفة.
- **وجود خطة مسبقة لدراسة ما وراء التحليل :** وجود تصور مسبق عن طبيعة العلاقة بين المتغيرين وتحديد اهم خصائص الدراسات المراد تكويدها واهم العوامل التى يمكن ان تسبب الاختلافات بين الدراسات لتضمينها فى التحليل لاجراء التحليلات المتفاعلة.
- **الدقة فى تفسير نتائج ما وراء التحليل:** الفهم الواضح للمشكلة ووضع الخطة المناسبة للتحليل يقود الى الوصول الى نتائج تتسم بالموثوقية، واثناء التفسير راعى انه يجب تحويل حجم التأثير اللوغاريتمى الى حجم تأثير قابل للتفسير. ويجب ان يكون التفسير مرتبط بالظرية سواء بتأييدها او رفضها او محاولة ادخال تحسينات او تعديلات عليها كذلك ان وجدت اختلافات بين احجام التأثير فيجب البحث عن العوامل التى احدثت هذه الاختلافات من خلال اجراء ANOVA او ما وراء الانحدار.

تجنب اشكاليات عند كتابة عرض نتائج ما وراء التحليل

قدم (Card 2012) عشرة اشكاليات شائعة عند عرض نتائج ما وراء التحليل وفيما يلى عرض هذه الاشكاليات والمقترحات لمحاولة التغلب عليها:

1. **عدم الربط بين المنطقية المفاهيمية وتحليلات البيانات:** احد اهم المشاكل عند كتابة تقرير نتائج ما وراء التحليل عدم الربط بين الخلفية النظري للمراجعة او للدراسة والتحليلات والنتائج المعروضة. اى تحليل يتم اجرائه يكون لسبب وهذا السبب يكون معروضاً فى مقدمة الدراسة،

وبعض الباحثين يعرضوا احياناً بعض التحليلات الاستكشافية ولكن يجب التركيز في المقام الاول على التحليلات التي تجيب على الاسئلة البحثية او تخدم اهداف الدراسة.

2. **عدم اعطاء عرض وتفصيلات كافية عن المنهجية:** في دراسات ما وراء التحليل يجب عرض وصف كافي عن استراتيجيات البحث التي اعتمد عليها الباحث في الحصول على الدراسات، معايير التضمين او الاستبعاد، اجراءات التأكيد لخصائص الدراسات، كيفية حساب حجم التأثير، واستراتيجية التحليل المستخدمة. ويجب عرض الدراسة قبل نشرها على متخصصين في ما وراء التحليل لمراجعة النسخة الاولى من الدراسة.

3. **عدم كتابة تفاصيل النتائج:** عند عرض نتائج ما وراء التحليل يجب عرض المعلومات الخاصة بالدراسات المفردة والاحصائيات التلخيصية مثل متوسط حجم التأثير واختبار الاختلاف بين احجام التأثير ودلالته الاحصائية وتحليل المتفاعلات ان وجدت واختبار تشخيص التحيز وعرض الاشكال البيانية.

4. **عرض التعقيدات الفنية للاسلوب للجاجة عن سؤال البحث:** يعتبر اسلوب ما وراء التحليل من الاساليب التي تتسم بالتعقيد الى حد ما ولكن بقدر الامكان استخدام الفنيات البسيطة كلما كان ممكن لان استخدام الفنيات المعقدة يمكن ان يقع الباحث في اخطاء. ومن الصعب فهم هذه الفنيات المعقدة ولا مانع من استخدام الفنيات المعقدة اذا كانت ضرورية للجاجة عن الاسئلة البحثية المعقدة. ولتجنب عدم فهم تقرير بحث ما وراء التحليل اعرض البحث على زملائك لترى مدى فهمهم للبحث واذا وجدت صعوبة او استنتاج غير دقيق فيجب اخذ ذلك في الاعتبار في الصورة النهائية للنشر.

5. **نسيان لماذا تجرى دراسة ما وراء التحليل:** وهذا واضح عند كتابة جزء المناقشة فينسى الباحث الوصول الى استنتاج واضح عن النتائج ما وراء التحليل للجاجة عن اسئلة البحث ومدى اسهام الدراسة في زيادة الفهم للظاهرة المراد دراستها وكذلك اصدار توصيات احياناً لا ترتبط بمشكلة البحث.

6. **الفشل في تحديد محددات عينة الدراسات:** لاتعتقد ان الدراسات المتضمنة في ما وراء التحليل تمثل مجتمع الدراسات تمثيلاً دقيقاً، فتوجد محددات في الحصول على الدراسات وكذلك

بعد استبعاد العديد من الدراسات فى ضوء معايير التضمن والاستبعاد فاصبحت عينة صغيرة نسبياً فى ضوء مجتمع الدراسات للمتغيرين المراد دراستهما ويجب عرض نتائج تقييم تحيز النشر وتحديد خصائص هذه الدراسات يحد من القدرة التعميمية لنتائج الدراسة.

7. **عدم عرض الملامح الوصفية للدراسات:** الفشل فى اعطاء المعلومات الاساسية والوصفية عن الدراسات المتضمنة فى ما وراء التحليل وهذه المعلومات او خصائص الدراسات على درجة كبيرة من الاهمية سواء فى استخدام بعض الخصائص فى اجراء التحليلات المتفاعلة او للوصول الى استنتاجات من الدراسة ولذلك من الضرورى عرض هذه الخصائص فى بداية جزء النتائج.

8. **استخدم نموذج التأثيرات المثبتة فى التحليل فى حالة وجود تجانس بين احجام التأثير:** يفترض نموذج التأثيرات المثبتة ان الاختلاف بين احجام التأثير يرجع الى خطأ المعاينة فقط، واذا لم يتحقق التجانس بين احجام التأثير من الافضل استخدام نموذج التأثيرات العشوائية او نموذج التأثيرات المثبتة مع تحليل التفاعلات او النموذج المختلط. ويتم تقييم عدم التجانس بين احجام التأثير من خلال اختبار Q واذا كانت له دلالة احصائية فيجب استخدام نموذج التأثيرات العشوائية.

9. **عدم عرض محددات تحليلات ما وراء التحليل المتفاعلة Meta-analysis moderator analyses:** تعتبر التحليلات المتفاعلة من التحليلات الاساسية فى ما وراء التحليل وان نتائج هذه التحليلات مرتبطة بخصائص الدراسات ويجب عدم تفسير نتائجها بصورة سببية لان طبيعة المتغيرات المتضمنة فى هذه التحليلات غير تجريبية.

10. **لا تعتقد وجود طريق صحيح Right way لاجراء وكتابة ما وراء التحليل على الرغم من كل التوصيات الملموسة فى هذا العرض الا انها توصيات مجردة ومطلقة ولكن من الضرورى اتباع الصيغ المتبعة المحددة من قبل الجمعيات العلمية مثل الجمعية النفسية الامريكية APA وكذلك المجالات العلمية فى هذا الشأن.**

وعلى الرغم من مميزات ما وراء التحليل كاسلوب لتحليل وتلخيص النواتج الكمية للتراث البحثى الخاص بموضوع ما ولكنه كاي اسلوب يشوبه العديد من المحددات والمشاكل ويمكن عرض

البعض منها المرتبط بتفسير النتائج للممارسة ولصانعي السياسة وللبحث. وبعض هذه المحددات كالتالي (Lipsey & Wilson 2001):

المناسبة المنهجية للاساس البحثي Methodological adequacy

جودة نتائج ما وراء التحليل تتوقف على طبيعة الدراسات المتضمنة فى التحليل. فاذا كانت الدراسات المتضمنة فى التحليل فقيرة المنهجية فلا نتوقع ان نصل الى نتائج مفيدة وصادقة بالتالى تحد من قدرة الباحث على تعميم النتائج.

الدراسات المتضمنة فى دراسة ما وراء التحليل متنوعة فى الجودة المنهجية فبعضها جيد وبعضها يتضمن محددات منهجية وعلى ذلك فلا بد على الباحث ان يختبر الى مدى تختلف النتائج باختلاف الجودة المنهجية للدراسات وهل نفس النتائج المتحصل عليها من كل الدراسات هى نفسها للدراسات عالية الجودة وللدراسات منخفضة الجودة ولتحقيق ذلك يمكن تصنيف الدراسات الى دراسات جيدة ودراسات فقيرة المنهجية واجراء التحليل المتفاعل. ويمكن تضمين بعض الخصائص المنهجية مثل عمر العينة لاجراء تحليل الانحدار الموزون لمعرفة مدى اسهامها فى التنبؤ بحجم التأثير وهذه النوعية من التحليلات للمظاهر المختلف للمنهجية لتحديد تأثيرها على حجم التأثير يطلق عليها تحليلات الحساسية Sensitivity Analyses.

تداخل الملامح الجوهرية والمنهجية

يمكن تودى نتائج الدراسات الضعيفة منهجياً الى نتائج غير موثوق بها. ويقوم الباحث بالمقارنة بين احجام التأثير فى ضوء متغيرات تصنيفية مختلفة مثل المقارنة بين احجام التأثير لمعالجات مختلفة كالمقارنة بين العلاج السلوكى المعرفى فى مقابل اسلوب الاسترخاء فى احجام التأثير او المقارنة بين الذكور والاناث او المرحلة التعليمية او الفئة العمرية، وهذه التحليلات تهدف الى دراسة العلاقة بين هذه المتغيرات واحجام التأثير.

اهمية التباين

بالتأكيد يوجد فروق فردية بين الافراد فى الدراسات الاولية كذلك يوجد نفس الفروق بين احجام التأثير المتضمنة فى ما وراء التحليل وليس من الحكمة لباحث ما وراء التحليل اعطاء اهتمامه

لقيم متوسط حجم التأثير بدون الوضع فى الاعتبار تباين احجام التأثير. وفى ما وراء التحليل يتم تحديد مجتمع الدراسات التى تقيس نفس المتغيرين او نفس المفاهيم ومن الطبيعى ان يوجد اختلاف بين احجام التأثير للدراسات المختلفة لاسباب عديدة، وفى الوضع الطبيعى اذا كان توزيع احجام التأثير متجانس بمعنى انه من المتوقع ان لا تختلف احجام التأثير اكثر مما هو متوقع نتيجة خطأ المعاينة وفى هذه يستخدم الباحث نموذج التأثيرات المثبتة.

وعندما يوجد اختلاف بين احجام التأثير فليس من الحكمة ان يتم دمج كل احجام التأثير لحساب المتوسط ومن الضرورى تحديد لماذا اختلفت احجام التأثير من خلال تحديد خصائص الدراسات المنهجية والجوهرية وى ملامح اخرى ويتم تكويدها للمساعدة فى تحديد العوامل التى تسبب هذا الاختلاف فرما يكتشف الباحث ان المفاهيم الاجرائية المختلفة لنفس المفهوم، المعالجات المختلفة، العينات المختلفة وغيرها من العوامل يمكن ان تسبب الاختلاف فى احجام التأثير من ثم يصبح تحديد مصادر الاختلاف بين احجام التأثير هدف اسمى من حساب متوسط حجم التأثير. ويعتبر نموذج التأثيرات العشوائية من افضل الخيارات للتعامل مع توزيع غير متجانس لاحجام التأثير.

الفجوات البحثية والقدرة التعميمية Research Gaps and Generalizability

من النواتج القيمة لما وراء التحليل هو اعطاء وصف ملخص لطبيعة حال الدراسات فى تراث الظاهرة فمن الضرورى فحص التوزيع التكرارى لتوصيفات او خصائص الدراسات المنهجية او الجوهرية مثل نوعية التصميم والقياسات المستخدمة والعينات والمفاهيم الاجرائية والمعالجات غيرها، وهذه الخصائص تكشف عن القصورات والمحددات فى التراث البحثى للظاهرة واذا لم تمثل هذه الخصائص فى الدراسات المتضمنة فى ما وراء التحليل فان تعميم النتائج يكون موضع شك وتساؤل. واحد التضمينات هو ان التأثيرات فى ما وراء التحليل الخاصة بمدى كفاءة معالجة معينة لفئة معينة ليس من الضرورى ان تكون ممثلة للتأثيرات التى تحدث فى الممارسة الاكلينيكية فى المواقف غير البحثية.

واحد محددات تعميم النتائج ان الدراسات تتم فى مواقف معملية وليس من الضرورى ان تعميم نتائجها فى الممارسات الطبيعية فى المدرسة او المستشفيات او العيادات هذة تختلف عن طبيعة الابحاث التجريبية الحقيقية المعملية الصارمة.

تطبيقات ما وراء التحليل للممارسة ولصانعى السياسات

اجراء دراسة ما وراء التحليل بصورة متقنة وجيدة يمكن ان يكون لها اسهامات فعالة فى الممارسة العملية وصناعة السياسات والقرارات، فالدراسات المفردة تختلف فيما بينها فى النتائج والجودة ولذلك فان ما وراء التحليل يعطى اطار او رؤية شاملة للدراسات المفردة. وهذا التجميع يقلل خطأ المعاينة المسئول عن تضارب نتائج الدراسات. يرى (Hunter & Schmidt 2004) ان ما وراء التحليل يمدنا باجراء اميريقى كمي لتحديد النواقص المنهجية اذا تحقق التى:

الاول: ان يجمع الباحث مجموعة من الدراسات الشاملة التى تمثل مجتمع الدراسات تمثيلاً جيداً.
الثانى: تحديد ماهية العيوب المنهجية فى الدراسات.

الثالث: تطبيق ما وراء التحليل لكل الدراسات الجيدة والضعيفة منهجياً وليس كما هو سائد فى دراسات ما وراء التحليل استبعاد الدراسات الضعيفة منهجياً فى ضوء معايير الاستبعاد التى يضعها الباحث فتعامل دراسات ما وراء التحليل مع الجانب المضى ويستبعد الجانب المظلم وعليه ما الفائدة اذا كان الهدف الاسمى من ما وراء التحليل تقييم التراث البحثى ومحاولة تحديد المحددات المنهجية واخبار الباحثين فى المجال باهمية تلاشى هذه المحددات.

الرابع: اذا وجدت اختلافات بين الدراسات فى التحليل فان هذا الاختلاف ربما او لا يفسر عن طريق دراسات ما وراء التحليل المنفصلة الضعيفة منهجياً و الجيدة منهجياً. ومن الافضل تحديد هذة المشكلات المنهجية وقياسها وتصحيحها افضل من استبعادها وهذا المدخل الى تبناه (Hunter & Schmidit 2004) فى كتابهما: طرق ما وراء التحليل: تصحيح الخطأ والتحيز فى نتائج البحوث.

عرض (2009) Borenstien et al. العديد من القضايا الخلافية والجوهرية اثناء اجراء دراسة ما وراء التحليل كالتالى:

هل توجد منطقية لخلط دراسات بمنهجيات وعينات ومقاييس مختلفة في تحليل واحد؟. شبهه البعض تحليل ما وراء التحليل كالخلط بين التفاح والبرتقال، ولكن Rosenthal من المطورين الاوائل لاسلوب ما وراء التحليل قدم لهم اجابة مفيدة وهو انه يمكن جعل هذا الخلط له معنى اذا كنت مهتم بعمل سلطة فواكهة فهذ اكثر فائدة من اكل البرتقال فقط او التفاح فقط.

الهدف من ما وراء التحليل ليس فقط تجميع بيانات من مجموعة متماثلة من الدراسات وتوليها لتقدير حجم التأثير العام، ولكن من الممكن الذهاب ابعد من ذلك من خلال تحديد مدى تشتت او ابتعاد احجام التأثير عن المتوسط العام. فاذا كان الهدف تحديد اثر او فعالية التدخلات التدريسية على التحصيل فهل من الافضل تضمين الدراسات التي تتضمن الدراسات التي تتضمن قياس المهارات اللغوية والمهارات الرياضية؟، اذا كان الهدف تقدير الاثر على الاداء بصفة عامة فالاجابة نعم ولكن اذا كان الهدف تقدير الاثر على المهارة اللغوية فقط فالاجابة لا. وهل من الافضل تضمين الدراسات فى المرحلة الاعدادية والثانوية معاً؟، الاجابة تتوقف على طبيعة وهدف البحث.

ايضاً هل من المنطقى تضمين دراسات تجريبية مع دراسات ارتباطية او ملاحظة فى نفس التحليل؟، وماذا عن الدراسات شبه تجريبية؟، وهل من المقبول تضمين دراسات من تصميم مجموعات مستقلة مع دراسات من تصميم مجموعات مرتبطة او متزاوجة؟ الاجابة على هذه الاسئلة تتوقف على طبيعة السؤال البحثى.

هل يجب ان تكون الدراسات المتضمنة فى التحليل متشابهة؟.

من وجهة النظر الاحصائية لا يوجد قيود على التماثل او التشابه للدراسات المتضمنة فى التحليل سواء تشابه فى ضوء العينات او التدخلات او المنهجيات، يكون التحليل مفيد وله معنى اذا تناولنا هذه التنوعات فى التحليل فهذا مفيد فى تفسير الظاهرة موضع الدراسة ودراسة مدى استقرار وتماسل النظرية. ومدى اتساع معايير التضمين ومحدوديتها يتوقف على طبيعة الدراسة

فمثلاً فى الدراسات الدوائية والعلاج بادوية كلما كانت الدراسات المتضمن فى التحليل محكمة ومتجانسة ومعايير التضمين محددة تماماً لفئة معينة فان هذا يعطى تقديرات دقيقة. ولكن فى معظم تحليلات ما وراء التحليل معايير التضمين تكون واسعة وليست مقيدة وهذا يعطى قدراً كبيراً من التنوع بين الدراسات وهذا امر مرغوب فى ما وراء التحليل Borestein et al., (2009).

ودراسة ما وراء التحليل الجيدة هى التى تتضمن تنوع فى متغيراتها المتفاعلة التى تفسر الاختلاف فى احجام التأثير وكذلك يمكن تفسير الاختلاف بين احجام التأثير بالاعتماد على نموذج التأثيرات العشوائية، وايضاً بتضمين المتغيرات التصنيفية مثل الجنس او المرحلة التعليمية فهذا يفيد فى اجراء ما وراء تحليلات منفصلة فى الدراسة الواحدة وهذا يشار اليه ما وراء التحليل للمجموعات الفرعية Subgroup meta-analysis ولو كان التأثير متماثل عبر المجموعات المختلفة عندئذ يقال ان للتاثير مناعة او ضلابة Robust وهذا لن يحدث بتبنى معايير تضمين محدودة، بالتالى اذا كان الهدف تقدير متوسط حجم التأثير فلا مانع من معايير محدودة ضيقة ولكن اذا كان الهدف الكشف عن التنوع والاختلاف بين احجام التأثير فلا مفر من اتساع معايير التضمين.

هل من الممكن تضمين دراسات من تصميمات بحثية مختلفة؟.

النوعية المناسبة من الدراسات المتضمنة فى ما وراء التحليل يتوقف على نوعية السؤال البحثى الذى يطرحه الباحث فى دراسته، فمثلاً ما وراء التحليل الذى يتضمن تقييم فعالية برنامج او تدخل معين سيعتمد على الدراسات التجريبية وايضاً يعتمد على نوعية التجارب هل التصميمات التى تتضمن مجموعتين واى انتقاء للمجموعتين للعينة عشوائية او مقصودة، ام تصميم المجموعة الواحدة ويعتمد الباحث على الدراسات السببية الارتباطية كل هذا يتوقف على الهدف من الدراسة.

ويرى البعض تضمين كل الدراسات التجريبية سواء كانت حقيقية او شبه تجريبية وكذلك تضمين كل التصميمات ثم دراسة الفروق فى احجام التأثير بين التصميمات المختلفة من خلال تحليل المتفاعل. بينما يرى البعض ان ما وراء التحليل الذى يهتم بدراسة تاثيرات التدخلات تكون

مقصورة فقط على الدراسات التجريبية الحقيقية التي تتمتع بصدق داخلي مرتفع وبكل تأكيد يترتب على ذلك استبعاد الدراسات الملاحظة الارتباطية التي تتناول العلاقة بين المتغيرين على مجموعة واحدة بدون معالجة، ولكن Borenstein et al.(2009) يرى ان الافضلية في الدراسات المتضمنة ليس بطبيعة التصميم ولكن بمدى جودة الدراسة في اعطاء تقدير غير متحيز لحجم التأثير ولكن يجب تحليل الدراسات التجريبية منفصلة عن دراسات الارتباطية وربما يتم دمجها في تحليل واحد لو لم يختلفوا في حجم التأثير ويتناولوا نفس السؤال.

حتى في دمج دراسات المقارنة بين متوسطات المجموعة التجريبية في مقابل المجموعة الضابطة يستخدم المجموعات المستقلة ام المجموعات المرتبطة، وفي هذا الشأن لا يوجد مانع او اى مشكلة فنية في دمج البيانات من تصميم المجموعات المستقلة مع تصميم المجموعات المرتبطة.

هل تدمج دراسات باستخدام تحليلات مختلفة او عرضت النتائج بصورة مختلفة؟

يتعامل دراسة ما وراء التحليل مع نتائج مختلفة لنفس المتغيرين. فبعض الدراسات تقيس المتغير التابع في ضوء المتوسط وبعضها في ضوء اعداد مثل قياس الفترة الزمنية في ضوء عدد الايام او الاسابيع وهذا شائع في تخصصات مثل الزراعة والطب اونسبة الافراد او النباتات التي حدث لها تحسن، حتى في نفس الدراسة يتم عرض النتائج في صور مختلفة نتيجة لاستخدام مداخل مختلفة للتحليل البيانات. ومهما اختلفت طرق عرض النتائج لنفس المتغيرين عبر الدراسات المختلفة فان العبرة هي بكيفية حساب حجم التأثير حيث يمكن تحويل الصيغ المختلفة لحجم التأثير الى مؤشر موحد مثل معامل الارتباط r .

كم عدد الدراسات اللازمة لاجراء ما وراء التحليل؟

السؤال الشائع المطروح في ما وراء التحليل هو كم عدد الدراسات اللازمة لاجراء ما وراء التحليل؟، يرى Borenstein et al.(2009) باستخدام نموذج التأثيرات المثبتة لاجراء ما وراء التحليل يمكن الاعتماد على دراستين لان الحصول على ملخص لحجم التأثير لدراستين افضل واكثر دقة من دراسة واحدة، المهم ان اهتمام الباحث ليس دراسة الاختلافات بين احجام التأثير لان هذا يهتم بشئ اخر ابعد من خطأ المعاينة، وملخص حجم التأثير يتم توصيفة في فترات الثقة التي تصف عدم التاكيد لتقدير حجم التأثير.

ولكن باستخدام نموذج التأثيرات العشوائية حيث يكون الاختلافات بين احجام التأثير ضرورة قصوى ومصدر اخر للخطأ فان عدد الدراسات الصغيرة يجعل تقدير الخطأ غير صادق او ثابت لان تقدير التباين بين الدراسات τ^2 يعتبر مصدر اخر للخطأ وليس فقط خطأ المعاينة كما في نموذج التأثيرات المثبتة من ثم عددالدراسات الصغيرة يسبب تقدير تحيز للخطأ المعيارى المرتبط بمتوسط حجم التأثير ويسبب تضخم او تقدير موجب متحيز لمتوسط حجم التأثير ويعطى تقديرات غير دقيقة لفترات الثقة. ولكن (Pigott 2012) يرى ان الاجابة الكاملة والحقيقية لهذا السؤال تتوقف على القوة الاحصائية للاختبارات المتضمنة فى ما وراء التحليل سوا متوسط حجم التأثير او الاختلاف او تحليل المتفاعل فى ضوء نموذج التأثيرات المثبتة ونموذج التأثيرات العشوائية، وعليه فان القاعدة فى عينة الدراسات فى ما وراء التحليل هى "الاكثر افضل" وهى قاعدة فى الدراسات الاولية لان تضمين عدد قليل من الدراسات فى التحليل فمن الافضل عدم تلخيصها احصائياً ومن الافضل مراجعتها كيفياً ولكن هذا لا يعنى عدم اجراء ما وراء التحليل بل اجرائه وفى هذه الحالة يجب عدم الوثوق فى النتائج وتعميمها.

حقيقة فى ظل تراكم دراسات ما وراء التحليل ظهر الجيل الثانى من ما وراء التحليل وهو ما وراء التحليل من الرتبة الثانية Second-Order Meta-analysis وهى اجراء ما وراء التحليل لدراسات ما وراء التحليل نفسها بالتالى فان عينة الدراسات ليست دراسات اولية انما دراسات ما وراء التحليل ولكن المشكلة التحليلية لاجراء هذا النوع من الدراسات هى خطأ المعاينة(للمزيد يمكن الرجوع الى (Hunter & Schmidt, 2004).

معايير الجودة فى دراسة ما وراء التحليل

قدم (Cooper 2009) قائمة تتضمن عدد من الخطوات الارشادية او المعايير التى يمكن ان تساعد على جودة الدراسة وتقييمها ومدى مقبوليتها للنشر وهى كالاتى:

تحديد المشكلة

- هل توجد منطقية ومبررات لدراسة المشكلة؟.
- هل التعريفات النظرية واضحة للمتغيرات الاساسية؟
- هل متغيرات الدراسات الاولية مناسبة لتمثل مفاهيم دراسة ما وراء التحليل؟

- هل يوجد عدد كافي من الدراسات الاولية تناولت متغيرات دراسة ما وراء؟
- هل التصميمات البحثية واضح ومناسبة للاجابة عن الاسئلة؟
- هل قواعد او معايير التضمنين او الاستبعاد واضحة؟

جمع البيانات

- هل توجد استراتيجيات بحثية عديدة وقواعد بيانات للبحث عن الدراسات المناسبة؟
- هل استخدم الباحث مصطلحات مترادفة للمتغيرات الرئيسية اثناء البحث فى قواعد البيانات؟
- هل استخلاص البيانات من الدراسات تميز بالموضوعية والثبات؟
- هل استخلص الباحث معلومات تكويدية ملائمة من الدراسات المتعددة؟
- هل الطرق المستخدمة فى استخلاص البيانات مرضية، موثقة، واضحة، وقابلة للاعادة من قبل الاخرين؟
- هل تعامل مع البيانات الغائبة(الدراسات الغائبة) بطرق مناسبة؟

حساب وتوليف احجام التأثير

- هل استخدم مؤشر حجم التأثير المناسب؟
- هل قدر متوسط حجم التأثير؟
- هل قدر فترات الثقة؟
- هل قدر الدلالة الاحصائية لمتوسط حجم التأثير؟
- هل اختار النموذج الاحصائى المناسب؟
- هل حسب المتوسط الموزون فى ضوء مقلوب تباين الخطأ المعيارى؟
- هل حسب درجة الاختلاف بين احجام التأثير؟
- هل اجرى تحليل التفاعلات اذا حدث اختلاف بين احجام التأثير؟

اختبار ضلاعة النتائج

- هل حسب تحليل الحساسية واستخدمه فى تفسير النتائج؟
- هل تناول قضية البيانات الغائبة من خلال اختبارات تحيز النشر؟

عرض النتائج

- هل عرض النتائج فى ضوء جداول واشكال بشكل جيد؟
- هل الاستخلاصات من الدراسة واضحة وكاملة ومرضية؟

تفسير النتائج

- هل اصدر تعميمات و محددات الدراسة؟
- هل تفسير النتائج مرتبطة بما انتجته الدراسة من نتائج ومرتبطة بالقيمة العملية بمجال الدراسة ؟
- هل حدد نواحي معينة بحاجة الى دراسات اخرى؟

ومن خلال العرض السابق لفصول الكتاب يمكن لمعد هذا الكتاب من طرح قائمة لكيفية قراءة وتقييم جودة دراسة ما وراء التحليل كالاتى:

- هل العنوان مصاغ بطريق توحى انها دراسة ما وراء التحليل؟
- هل الاسئلة مصاغة بطريقة مناسبة؟
- هل استراتيجية البحث واضحة وشاملة (قواعد البيانات او اى مصادر اخرى)؟
- هل معايير التضمين والاستبعاد واضحة وكافية؟
- هل قيم الجودة المنهجية للدراسات بقدر الامكان؟
- هل بروتكول التكويد شامل وواضح؟
- هل تحقق من موضوعية عملية التكويد وثباتها؟
- هل عينة الدراسات ممثلة لمجتمع الدراسات فى التراث البحثى؟
- هل اختار حجم التأثير المناسب لطبيعة المتغيرات؟
- هل اعتمد على اكثر من حجم تأثير من نفس الدراسة؟
- هل قدر حجم التأثير للعينات الفرعية فى الدراسة؟
- هل اجرى تصحيح لحجم التأثير من المحددات المنهجية كلما امكن؟
- هل قدر الدلالة الاحصائية
- هل قدر فترات الثقة حول متوسط حجم التأثير؟

- هل حدد البرنامج المستخدم لاجراء التحليل ومبررات ذلك؟
- هل حدد مدخل التحليل المستخدم كطريقة Hedges او Hunter & Schmidt ؟
- هل حدد نموذج التحليل المستخدم سوا مثبت او عشوائى او مختلط ومبررات اختياره؟
- هل قدر اختبار الاختلاف بين الدراسات؟
- واذا وجد اختلاف هل اجرى التحليل المتفاعل لتحديد سبب الاختلاف؟
- هل تحقق من تحيز النشر والبيانات الغائبة؟
- هل اجرى تحليل الحساسية فى الدراسة؟
- هل نتائج الدراسة منطقية وصحيحة؟
- هل عرض النتائج فى جداول واضحة؟
- هل عرض النتائج فى رسومات بيانية خاصة شكل Forest plot؟
- هل اعطى تفسيرات واضحة للنتائج؟
- هل طرح الاستفادة والاهمية من نتائج الدراسة؟
- هل عرض الصعوبات ومحددات الدراسة؟

المراجع

- Abelson, R. P. (1995). *Statistics as principled argument*. Hillsdale, NJ: Erlbaum.
- Alhija, F. N., & levey, A. (2009). Effect Size reporting practices in published articles. *Educational and Psychological Measurement*, 69, 245- 265.
- American psychological Association (APA). (2001). *Publication manual of the American Psychological Association* (5th Ed.). Washington, DC: Author.
- American psychological Association (APA). (2010). *Publication manual of the American Psychological Association* (6th Ed.). Washington, DC: Author.
- Amstrong, J. S. (2007). Significance tests harm progress in forecasting. *International Journal of Forecasting*, 23, 321-327.
- Bangert-Drowns, R. L. (1986). Review of developments in meta-analysis method. *Psychological Bulletin*, 99, 388-399.
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173-1182.
- Begg, C. B., & Mazumdar, M. (1994). Operating characteristics for publication bias. *Biometrics*, 5, 1088-1101.
- Bloom, B. S. (1964). *Stability and change in human characteristics*. New York: Wiley.
- Borenstein, M. (2009). Effect size for continuous data. IN H. Cooper, L. V. Hedge, & J. C. Valentine(Eds.), *The handbook of research synthesis and meta analysis* (2nd.ed) (pp. 222-235). New York: Russell Sage Foundation.

- Borenstein, M., Hedges, L. V., Higgins, J. P. T., Rothstein, H. R. (2009). *Introduction to meta-analysis*. West Sussex: John Wiley & Sons, Ltd.
- Borman, G. D. & Grigg, J. A. (2009). Visual and narrative interpretation. In H. Cooper, L. V. Hedges, & J. C. Valentine(Eds.), *The handbook of research synthesis and meta-analysis (2nd ed., pp. 497-519)*. New York: Russell Sage.
- Breaugh, J. A. (2003). Effect Size Estimation: factors to consider and Mistakes to Avoid. *Journal of Management*, 29, 79-97.
- Card, N. A. (2012). *Applied meta-analysis for social sciences research*. New York: Guilford Press, Inc.
- Card, N. A., Stucky, B. D., Sawalani, G. M., & Little, T. D. (2008). Direct and indirect aggression during childhood and adolescence: A meta-analytic review of gender differences, inter-correlations, and relations to maladjustment. *Child Development*, 79, 1185-1229.
- Cheung, M. W-L. (2008). A model for integrating fixed-, random-, and mixed-effects meta analyses in structural equation modeling. *Psychological Methods*, 13, 182-202.
- Cleophas, T. J., & Zwinderman, A. H. (2017). *Modern meta-analysis: review and update of methodologies*. Switzerland: Springer International Publishing.
- Cohen, J. (1962). The statistical power of abnormal–social psychological research: A review. *Journal of Abnormal and Social Psychology*, 65, 145- 153.
- Cohen, J. (1969). *Statistical power analysis for the Behavioral sciences*. London: Academic Press.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd Ed.)*. Hillsdale, NJ: Erlbaum.
- Cohen, J. (1990). Things I have learned (So far). *American Psychologist*, 45, 1304 – 1312.

Cohen, J. (1994). The Earth is round ($P < 0.05$). *American Psychologist*, 49, 997 – 1003.

Cohen, J., & Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/ correlation analysis for behavioral sciences*(3rd.ed). Mahwah, NJ: Erlbaum.

Cohen, L. D., & Becker, B. J. (2003). How meta-analysis increase statistical power. *Psychological Methods*, 8, 243-253.

Cooper, H. & Hedges, L. V. (Eds.). (1994). *The handbook of research synthesis*. New York: Russell Sage.

Cooper, H. (1998). *Synthesizing research: A guide for literature reviews*. Thousand Oaks, CA: Sage.

Cooper, H. M. (1988). Organizing knowledge syntheses: A taxonomy of literature reviews. *Knowledge in society*, 1, 104-126.

Cooper, H. M., & Rosenthal, R. (1980). Statistical versus traditional procedures for summarizing research findings. *Psychological Bulletin*, 87, 442-449.

Cooper, H., DeNeve, K., & Charlton, K. (1997). Finding the missing science: The fate of studies submitted for review by a human subjects committee. *Psychological Methods*, 2, 447-452.

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

Cumming, G., & Finch, S. (2001). A primer on the understanding, use, and calculation of confidence intervals that are based on central and noncentral distribution. *Educational and Psychological Measurement*, 61, 530-572.

De Smidt, G. A., & Gorey, K. M. (1997). Unpublished social research: Systematic replication of a recent meta-analysis of published intervention effectiveness research. *Social Work Research*, 21, 58-62.

Dickersin, K., Chan, S., Chalmers, T. C., Sacks, H. S. & Smith, H. Jr. (1987). Publication bias and clinical trials. *Controlled Clinical Trials*, 8, 343-353.

Dickerson, K. (2005). Publication bias: Recognizing the problem, understanding its origins and scope, and preventing harm. In H. R. Rothstein, A. J. Sutton, & M. Borenstein (Eds.). *Publication bias in meta-analysis: Preventing, assessment, and adjustments* (pp.11-34). Chichester: John Wiley & Sons Ltd.

Dulark, J. A. (1995). *Reading and understanding multivariate statistics*. Washington, DC: American Psychological Association.

Dulark, J. A. (2009). How to select, calculate, and interpret Effect Size. *Journal of Pediatric Psychology*, 3, 40, 917- 928.

Dunlap, W. P. (1994). Generalizing the common language effect size indicators to bivariate normal correlations. *Psychological Bulletin*, 116, 509- 611.

Duval, S. I., & Tweedie, R. L. (2000). Trim and fill: A simple funnel-plot based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, 56, 455-463.

Egger, M., Smith. D. G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, 315, 752-754.

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

Eysenck, H. J. (1952). The effects of psychotherapy: An evaluation. *Journal of Consulting Psychology*, 16, 319-324.

Eysenck, H. J. (1978). An exercise in mega-silliness. *American Psychologist*, 33, 517.

Ferguson, C. J. (2009). An effect Size primer: A Guide for clinicians and Researchers. *Professional Psychology: Research and Practice*, 2, 1- 7.

Fern, E. F., & Monroe, K. B. (1996). Effect Size Estimates: Issues and problems in interpretation. *Journal of Consumer Research*, 23, 89- 105.

Fidler, F., Cumming, G., Thomason, N., Pannuzzo D., Smith, J., Fyffe, P., Edmonds, H., Harrington, C., & Schmitt, R. (2005). Toward improved statistical reporting in the journal of consulting and clinical psychology. *American psychological Association*, 73, 136-143.

Field, A. P. (2003). The problems using fixed-effects models of meta-analysis on real-world data. *Understanding Statistics*, 2, 77-96.

Field, A. P. (2005). "Is meta-analysis of correlation coefficients accurate when population correlations vary?" *Psychological Methods*, 10, 444-467.

Frohlich, M., Emrich, E., Pieter, A., & Stark, R. (2009). Outcome Effect Sizes in Sport sciences. *International Journal of Sports Sciences*, 3, 175-179.

Furr, R. M. (2008). Summary for effect size and their links to inferential statistic. Available from: <http://www.wfu.edu/ufurrrm/sizeformulapdf>.

Ghiselli, E. E. (1955). The measurement of occupational aptitude. *University of California Publications in Psychology*, 8, 101-216.

Glass, G. V. (1976). Primary, secondary and meta-analysis of research. *Educational Researcher*, 5, 3-8.

Glass, McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research*. Beverly Hills, CA: Sage.

Gotzsche, P. C., Hroegjartsson, A., Maric, K., & Trendal, B. (2007). Data extraction errors in meta-analysis that use standardized mean differences. *Journal of the American Medical Association*, 298, 430-437.

Greenhouse, J. B., & Iyengar, S. (1994). Sensitivity analysis and diagnostics. In H. Cooper & L. V. Hedges (Ed.), *Handbook of research synthesis* (pp. 383-398). New York: Russell Sage.

Greenwald, A. G., Gonzalez, R., Harris, R. J., & Guthrie, D. (1996). Effect sizes and p values: what should be reported and

- what should be replicated? *Psychological Physiology*, 33, 175-183.
- Grissom, R. J., & Kim, J. J. (2011). *Effect sizes for research: Univariate and multivariate applications (2nd.ed)*. New York, NY: Routledge.
- Hagger, M. S. (2006). Meta-analysis in sport and exercise research: Review, recent, development, and recommendations. *European Journal of Sport*, 6, 103-115.
- Hall, J. A., & Brannick, M. T. (2002). Comparison of two random effects methods of met-analysis. *Journal of Applied Psychology*, 87, 377-389.
- Hayes, W. L. (1963). *Statistics for psychology*. New York: Holt, Rinehart & Winston.
- Hayes, W. L. (1988). *Statistics (4th Ed.)*. New York: Holt, Rinehart, & Winston.
- Hedges, L. (1992). Modeling publication selection in meta-analysis. *Statistical Science*, 7, 227-236.
- Hedges, L. V. & Piggot, T. D. (2001). The power of statistical tests in meta-analysis. *Psychological Methods*, 6, 203-217.
- Hedges, L. V. (1981). Distribution Theory of Glass's estimator of effect Size and related estimators. *Journal of Educational Behavioral Statistics*, 6, 107- 128.
- Hedges, L. V. (1982). Fitting categorical models to effect sizes from a series of experiments. *Journal of Educational Statistics*, 7, 119-137.
- Hedges, L. V. (1987). How hard is hard science, how soft is soft science: The empirical cumulativeness of research. *American Psychologist*, 42, 443-445.
- Hedges, L. V. (2009). Statistical considerations. IN H. Cooper, L. V. Hedge, & J. C. Valentine(Eds.), *The handbook of research synthesis and meta analysis(2nd.ed)* (pp. 37-47). New York: Russell Sage Foundation.

- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.
- Hedges, L. V. & Vevea, J. L. (1998). Fixed- and random-effects models in meta-analysis. *Psychological Methods*, 3, 486-504.
- Higgins, J. P. T., & Green, S. (Eds.) (2006). *Cochrane handbook for systematic review of interventions*. Chichester, UK: John Wiley & Sons.
- Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, 21, 1539-1558.
- Hill, C. J., Bloom, H. S., Black, A. R., & Lipsey, M. W. (2007). *Empirical Benchmarks for Interpreting Effect Sizes in Research*. MDRC working papers on Research Methodology, Available from: <http://www.Mdrc.Org>.
- Hinkle, D. E., & Wiersma, W., & Jurs, S. G. (1994). *Applied statistics for the behavioral sciences (3rd.ed)*. Boston: Houghton.
- Hopewell, S., Clark, M., & Mallett, S. (2005). Grey literature and systematic reviews. In H. R. Rothstein, A. J. Sutton, & M. Borenstein (Eds.). *Publication bias in meta-analysis: Preventing, assessment, and adjustments* (pp. 49-72). Chichester: John Wiley & Sons Ltd.
- Howell, D. C. (2013). *Statistics methods for psychology (8th.ed)*. Belmont: Wadsworth, Cengage Learning.
- Huedo-Medina, T. B., Sanchez-Meca, J., Marin-Martinez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistics or I^2 index? *Psychological Methods*, 11, 193-206.
- Hunt, M. (1997). *How science takes stock*. New York: Russell Sage.
- Hunter, J. E. (1997). Needed: A ban on the significance test. *Psychological Science*, 8, 3-7.

Hunter, J. E., & Schmidt, F. L. (1977). A critical analysis of the statistical and ethical implications of various definitions of tests fairness. *Psychological Bulletin*, 83, 1053-1071.

Hunter, J.E., & Schmidt, F. L. (1990). *Meta-analysis: Correcting error and bias in research findings*. Newbury Park, CA: Sage Publication, Inc.

Hunter, J.E., & Schmidt, F. L. (2000). Fixed effects vs. random effects meta-analysis models: Implications for cumulative research knowledge in psychology. *International Journal of Selection and Assessment*, 8, 275-292.

Hunter, J.E., & Schmidt, F. L. (2004). *Meta-analysis: Correcting error and bias in research findings (2nd Ed)*. Newbury Park, CA: Sage Publication, Inc.

Hunter, J.E., Schmidt, F. L., & Jackson, G. B. (1982). *Meta-analysis: Cumulating research findings across studies*. Beverly Hills, CA: Sage.

Iversson, A., Anderson, M. B., Johnson, V., & Lindwall, M. (2013). To adjust or not adjust: Nonparametric effect size, Confidence intervals, and Real – world meaning. *Psychology of Sport and Exercise*, 14, 97– 102.

Iyengar, S., & Greenhouse, J. (1988). Selection models and the file drawer problem. *Statistical Science*, 3,109-135.

Kampense, V. B., Dyba, T., Hannay, J. E., & Sjoberg, D. L. (2007). A Systematic Review of effect Size in Software engineering experiments. *Information and Software Technology*, 49, 1073- 1086

Kelley, K. & Preacher, K. J. (2012). On effect size. *Psychological Methods*, 17, 137-152.

Keselman , H. J., Huberty, C. J., Lix, L. M., Olejnik, S., Cribbie, R. A., Donahue, B., Kowalchuk, R. K., Lowman, L. L., Petoskey, M. D., Keselnon, J. C., & Levin. J. R. (1998). Statistical practices of educational Researchers: on analysis of their ANOVA, MANOVA, and ANCOVA analysis. *Review of Educational Research*, 68, 350- 386.

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

Kirk, R. E. (2001). Promoting good Statistical practice: Some Suggestions. *Educational and Psychological Measurement*, 61, 213 – 218.

Kirk, R. E. (2001). Promoting good Statistical practice: Some Suggestions. *Educational and Psychological Measurement*, 61, 213 – 218.

Kisamore, J. L., & Brannick, M. T. (2008). An illustration of the consequences of meta-analysis model choice. *Organizational Research Methods*, 11, 35-53.

Kline, R. B. (2013): *Beyond significance testing: Statistics reform in the behavioral science* (2nd Ed). Washington: American Psychological Association.

Koricheva, J., & Gurevitch, J. (2013). Place of meta-analysis among other methods of research synthesis. In J. Koricheva, J. Gurevitch, & K. Mengersen (Eds.), *Handbook of meta-analysis in ecology and evolution*(pp. 3-13). Oxford: Princeton University Press.

Leandro, G. (2005). *Meta-analysis in medical research: The handbook for the understanding and practice of meta- analysis*. Oxford: Blackwell Publishing Ltd.

Ledesma, R. D., Macbeth, G., & Dekohan, N. G. (2009). Computing Measures with vista: The visual statistics system. Tutorials, in *Quantitative Methods for Psychology*, 5, 25-34.

Ledesma, R. D., Macbeth, G., & Dekohan, N. G. (2009). Computing Measures with vista: The visual statistics system. *Tutorials in Quantitative Methods for Psychology*, 5, 25-34.

Levine , T . & Hullett, C. (2002). Eta squared, partial Eta squared, and misreporting of effect Size in communication research . *Human Communication Research*, 28, 612- 625.

Light, R. J., & Pillemer, D. B. (1984). *Summing up: The science of reviewing research*. Cambridge, MA: Harvard University Press.

Light, R. J., Singer, J. D., & Willett, J. B. (1994). The visual presentation and interpretation of meta-analyses. In H. V. Hedges (eds.), *The handbook of research synthesis*. New York: Russell Sage.

Lipsey, M. W. (1998). Design sensitivity: statistical power for Applied experimental research. In L. Beckman & D. J. Rog (Eds.), *Handbook of Applied Social Research Methods* (PP. 39- 68). Thousand Oaks, CA: Sage.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks: Sage Publications.

Littell, J. H., Corcoran, J., & Pillai, V. (2008). *Systematic review and meta-analysis*. OXFORD: University Press.

Matt, G. E., & Cook, T. D. (2009). Threats of the validity of generalized inferences. IN H. Cooper, L. V. Hedge, & J. C. Valentine(Eds.), *The handbook of research synthesis and meta-analysis(2nd.ed)* (pp. 538-557). New York: Russell Sage Foundation.

Maxwell, S. E. (2004). The persistence of underpowered studies in psychological research: Causes, consequences, and remedies. *Psychological Methods*, 9, 147-163.

McGrath, R. E., & Meyer, G. J. (2006). When effect Size disagree: the case of r and d. *Psychological Methods*, 11, 386 – 401.

McGraw, k. O., & Wong, S. P. (1992). A Common language effect size Statistic. *Psychological Bulletin*, 111, 361- 365.

Mcleod, B. D., & Weis, J. R. (2004). Using dissertation to examine potential bias in child and adolescent clinical trials. *Journal of Consulting and Clinical Psychology*, 72, 235-251.

McMillan, J. H. (2011). Reporting and discussing Effect Size: Still the road less traveled. *Practical Assessment Research & Evaluation*, 16, 1- 11.

Mengersen, K., Schmid, C. H., Jennions, & Gurevitch, J. (2013). Statistical models and approaches to inference. In J. Koricheva, J. Gurevitch, & K. Mengersen (Eds.), *Handbook of meta-analysis in ecology and evolution*(pp. 89-107). Oxford: Princeton University Press.

Olejnik, S., & Algina, J. (2003). Generalized Eta and omega squared statistics: Measures of effect Size for some common Research design. *Psychological Methods*, 434 – 44 .

Olejnik. S., & Algina, J. (2000). Measures of effect Size for comparative studies: Applications, interpretations, and limitations. *Contemporary Educational Psychology*, 25, 241 – 286.

Onwuegbuzie, A. J. (2002). Common analytical and interpretational errors in educational research: An analysis of the 1998 volume of the *British Journal of Educational Psychology*. *Educational Research Quarterly*, 26, 11-12.

Pedhazur, W. J., & Schmelkin, L. p. (1991). *Measurement, design, and analysis: an integrated approach*. Hillsdale, NJ: Erlbaum

Pigott, T. D. (2012). *Advances in meta-analysis*. New York: Springer Science + Business Media, LLC.

Pigott, T. D., & Wu, M. J. (2008). Methodological issues in meta-analysis standard deviations: Comments on Bond and Depaulo (2008). *Psychological Bulletin*, 134, 498-500.

Prentice, D. A., & Miller, D. T. (1992). When small effect are impressive. *Psychological Bulletin*, 112, 160- 164.

Reed, J. G. & Baxter, P. M. (2009). Using reference database. IN H. Cooper, L. V. Hedge, & J. C. Valentine(Eds.), *The handbook of research synthesis and meta analysis*(2nd.ed) (pp. 73-101). New York: Russell Sage Foundation.

Richardson, J. T. E. (1996). Measures of effect Size. *Behavior Research Methods, Instruments, and Computers*, 28, 12- 22.

Richardson, J. T. E. (2011). Eta squared and partial Eta squared as measures of effect Size in educational Research. *Educational Research Review*, 6, 135- 147.

Rodriguez, M. C., & Maeda, Y. (2006). Meta-analysis of coefficient alpha. *Psychological Methods*, 11, 306-322,.

Rosenthal, R. (1961, September). On the social psychology of the psychological experiment: With particular reference to experimenter bias. In H. W. Riecken(Chair), *On the social psychology of the psychological experiment*. Meeting of the American of the American Psychological Association, New York.

Rosenthal, R. (1984). *Meta-analysis procedures for social research*. Beverly Hills, CA: Sage.

Rosenthal, R. (1991). *Meta-analysis procedures for social research*(2nd ed). Newbury Park, CA: Sage.

Rosenthal, R. (1994). Statistically describing and combining studies. In H. Cooper & L. V. Hedges (Eds.),*The handbook of research synthesis*(pp.231-244). New York: Russell Sage Foundation.

Rosenthal, R., & DiMatteo, M. R. (2001). Meta-analysis: Recent development in quantitative methods for literature reviews. *Annual Review of Psychology*, 52, 59-82.

Rosenthal, R., & Rubin, D. B. (1978). Interpersonal expectancy effects: The first 345 studies. *The Behavioral and Brain Sciences*, 3, 377-415.

Rosenthal, R., Rosnow, R. L., & Rubin, D. B (2000). *Contrasts and effect sizes in behavioral Research: Acorretational approach*. New York: Cambridge university press.

Rosenthal,R., & Rubin, D. B. (1982). A simple, general purpose display of magnitude of experimental effects. *Journal of Educational Psychology*, 74, 166-169.

Rosnow, R. L., & Rosenthal, R. (1989). Statistical Procedures and The Justification of Knowledge in Psychological science. *American Psychologist*, 44, 1276 – 1286.

Rothstein, H. R., Sutton, A. J., & Borenstein M. (2005 Eds.). *Publication bias in meta-analysis: Preventing, assessment, and adjustments*. Chichester: John Wiley & Sons Ltd.

Rubin, D. B. (1990). A new prespective on meta-analysis. In K. W. Wachter & M. L. Straf (Eds.), *the future of meta-analysis*. New York: Russell Sage.

Schmid, C. H., Stewart, G. B., Rothstein, H. R., Lajeunesse, M. J., & Gurevitch, J. (2013). Software for statistical meta-analysis. In J. Koricheva, J. Gurevitch, & K. Mengersen (Eds.), *Handbook of meta-analysis in ecology and evolution (pp. 174-191)*. Oxford: Princeton University Press.

Schmidt, F. L. (1992). What do data really mean? Research findings, meta-analysis, and cumulative knowledge in psychology. *American Psychologist*, 47, 1173-1181.

Schmidt, F. L. (1996). Statistical significance testing and cumulative knowledge in psychology: Implications for the training of researchers. *Pschological Methods*, 1, 115-129.

Schmidt, F. L., & Hunter, J. E. (1997). Eight comman but false objections to the discontinuation of significance testing in the analysis of research data. In L. L. Harlow, S. A. Mulaik, & J. H. Steiger (Eds.), *What if there were no significance tests? (pp.37-64)*. Mahwah, NJ: Erlbaum.

Schmidt, F.L., & Hunter, J. E. (1977). Development of a general solution to the problem of validity generalization. *Journal of Applied Psychology*, 62, 529-540.

Schulze, R. (2004). *Meta-analysis: A comparison of approaches*. Cambridge, MA: Hogrefe and Huber.

Sedlmeier, p., & Gigerenzer, G. (1989). Do studies of statistical power have an effect on the power of studies?. *psychological Bulletin*, 105, 309- 316.

Shadish, W. R. (1995). The Logic of generalization: Five principles common to experiments and ethnographies. *American Journal of Community Psychology*, 23, 419-247.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.

Shadish, W. R., Robinson, L., & Lu, C. (1999). *ES: A computer program and manual for effect size calculation*. St. Paul, MN: Assessment Systems Cooperation.

Shadish, W., & Myers, D. (2004). Campbell Collaboration research design policy brief. Retrieved from <http://www.campbellcollaboration.org/MG/ResDesPolicyBrief.pdf>.

Shourt, P.E. (1997). Should significance tests be banned? Introduction to a special section exploring the pros and cons. *Psychological Science*, 8, 1-2.

Simmons, J. P., Nelson, L. D. & Simonsohn, U. (2011). False-positive psychological: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22, 1359-1366.

Smith, M., & Glass, G. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist*, 32, 752-760.

Snyder, P., & Lawson, S. (1993). Evaluating results using corrected and uncorrected effect size estimates. *Journal of Experimental Education*, 16, 334- 349.

Stern, J. A. C., Becker, B. J., Egger, M. (2005). The funnel plot. In H. R. Rothstein, A. J. Sutton, & M. Borenstein (Eds.). *Publication bias in meta-analysis: Preventing, assessment, and adjustments* (pp. 75-98). Chichester: John Wiley & Sons Ltd.

Sterne, J. A. C., & Egger, M. (2005). Regression methods to detect publication and other bias in meta-analysis. In H. R. Rothstein, A. J. Sutton, & M. Borenstein (Eds.), *Publication bias in meta-analysis: Preventing, assessment, and adjustments* (pp. 99-110). Chichester: John Wiley & Sons Ltd.

- Steven, J. P. (1980). Power of the multivariate analysis of variance tests. *Psychological Bulletin*, 88, 728- 737.
- Stout, D. E., & Ruble, T. L. (1995). Assessing the practical Significance of empirical Results in accounting education Research: The use of Effect Size information. *Journal of Accounting Education*, 13, 281- 298.
- Sun, S., Pan, W., & Wang, L. L. (2010). A comprehensive review of effect size reporting and interpreting practices in academic Journals in education and psychology. *Journal of Educational Psychology*, 102, 989- 1004.
- Tatsuoka, M. (1973). *An examination of the statistical properties of a multivariate measure of strength of readerships*. Urbana: university of Illinois. (ERIC Document. Service No. ED 099406).
- Tatsuoka, M. (1993). Effect Size. In G. Keren & C. Lewis (Eds.), *A Handbook for Data analyses in the behavioral sciences: Methodological Issues*. New Jersey: Lawrence Erlbaum Associates, publishers.
- Thompson, B. (1998). Statistical Significance and effect Size Reporting: portrait of a possible future. *Research in the Schools*, 5, 33 -38.
- Thompson, B. (Ed.). (1993). Statistical significance testing in contemporary practice (special issue). *Journal of Experimental Education*, 61(4).
- Thorndike, R. I. (1933). The effect of the interval between test and retest on the constancy of the IQ. *Journal of Educational Psychology*, 25, 543-549.
- Trusty, J., Thompson, B., & petrocelli , J. V. (2004). Practical guide for reporting effect Size in Quantative Research in the Journal of Counseling & Development. *Journal of Counseling & Development*, 82, 106- 110.

Turner, H., Boruch, R., Laveberg, J., Schoeneberger, J., & de Moya, D. (2004). *Electronic registers of trials*. Paper presented at the Fourth Annual Campbell Collaboration Colloquium: A First look at the Evidence; Washington, DC, 18-20 February.

Vacha-Haase, T., & Thompson, B. (2004). How to estimate and interpret various effect Sizes. *Journal of Counseling and Psychology*, 51, 473-481.

Wang, Z., & Thompson, B. (2007). In The Pearson r^2 biased and if so, what on the best correction formula. *Journal of Experimental Education*, 75, 109- 125.

Wilson, D. V. (2009). Systematic coding. *IN H. Cooper, L. V. Hedge, & J. C. Valentine (Eds.), The handbook of research synthesis and meta analysis(2nd.ed) (pp. 160-175)*. New York: Russell Sage Foundation.

Wolf, F. M. (1986). *Meta-analysis: Quantitative methods for research synthesis*. Beverly Hills: Sage Publications.

Wolf, I. M. (1990). Methodological observations on bias. *In K. W. Wachter & M. L. Straf (Eds.). The future of meta-analysis (pp. 139-151)*.New York: Russell: Sage.

Zientek, L. R, Capraro, M. M., & Capraro, R. M. (2008). Reporting practices in quantative teacher education research: On look at the evidence cited in AERA panel report. *Educational Researcher*, 37, 208 - 216.