

## فرا-تحلیل، روشی برای مطالعه مطالعات

دکتر محمدتقی ایمان\*

دکتر بیژن خواجه نوری\*\*

### چکیده

از دیرباز اندیشمندان انباشت معرفت علمی را یکی از خصایص مهم علوم دانسته و معتقد بوده‌اند که ما با تکیه به یافته‌های علمی دیگران تلاش می‌کنیم تا بر حل مسایل نظری و عملی خود فایق آییم. از دیرباز مشاهده شده که هریک از پژوهشگران براساس دانسته‌ها و تجربیات خود و دیگران سعی نموده تا در عرصه خاصی از علم به تحقیق پرداخته و نتایج کار خود را به طرق مختلف به اطلاع دیگرانی که در موضوعی خاص دست به تحقیق زده و به نتایج مشابه و گاه به نتایج کاملاً مخالف با یکدیگر رسیده‌اند، برسانند. آنچه در این میان بسیار مهم می‌باشد، این است که چگونه از ورای این همگونی‌ها یا ناهمگونی‌های یافته‌های علمی، با استفاده از روش‌های علمی معتبر، به نتایج و یافته‌های مشترک دست یابیم. روش فرا-تحلیل، به دلیل یاری رسانیدن به پژوهشگران برای ترکیب نتایج تحقیق در یک چارچوب کمی فراهم گردیده است. این مقاله سعی در شناسایی روش فوق و رویه‌هایی که منجر به ترکیب نتایج از تحقیقات گوناگون می‌گردد، دارد.

واژه‌های کلیدی: روش‌شناسی، فراتحلیل، اندازه اثر، مطالعات مستقل، تحلیل همگونی اطلاعات.

## مقدمه

در زندگی روزمره، ما اطلاعات را از منابع گوناگون کسب می‌نماییم و بر اساس آنها تصمیمات خود را شکل می‌دهیم. ما از ابتدای روز تا پایان روز، هزاران تصمیم می‌گیریم، بدین لحاظ ما در زندگی روزانه خود همچون یک فرا-تحلیل‌گر عمل می‌کنیم. به نظر تساو برخی مسایل روزمره زندگی، تحلیل‌هایی را بر مبنای نوعی تجربه‌های دانش پیشین یا حدس‌های ناپروورده ایجاب می‌کند. وی معتقد است که ما فرا - تحلیل‌گر زاده می‌شویم.<sup>(۱)</sup> فرا - تحلیل، روشی برای تحلیل نتایج از چند مطالعه تجربی است. گسترش روزافزون کمی مطالعات و تنوع موضوعات آنها در حوزه‌های گوناگون علمی، نیاز به ترکیب تحقیقات تجربی را به‌طور فزاینده‌ای افزایش داده است. به دلیل رشد این نیاز در بسیاری از حوزه‌های معرفت علمی، اهمیت بازنگری مناسب روش‌شناسی در حال افزایش است. به دلیل نواقص روش‌شناسی‌های سنتی مرور ادبیات همچون روش شرح جزئیات (بررسی گزارشی)<sup>۱</sup> استفاده از روش فرا - تحلیل به عنوان یکی از روش‌شناسی‌های مرور کمی تحقیق که در سال ۱۹۷۶ توسط گلاس و همکارانش شهرتی خاص یافت.<sup>(۲)</sup> در علوم بهداشت، پزشکی، علوم تربیتی و روانشناسی بسیار بیش‌تر از اقتصاد، جامعه‌شناسی، از این روش استفاده شده است.<sup>(۳)</sup> هدف این مقاله آشناساختن خوانندگان با رویه‌ها و مفروضات مرتبط با فرا - تحلیل و نشان دادن شیوه تفکر درباره ترکیب تحقیقات با استفاده از روش‌های آماری است. با کمک این رویه‌ها و شیوه تفکر می‌توان تفاوت‌های موجود در تحقیقات انجام شده را استنتاج کرد و در رسیدن به نتایج کلی و کاربردی از آن بهره جست. فرا - تحلیل راهی برای خارج شدن از نتایج تحقیقات پراکنده و رسیدن به یک نتیجه‌گیری مشترک می‌باشد.

## ۱- تعریف فرا-تحلیل

در خصوص مفهوم فرا - تحلیل، تعاریف کمابیش مشابهی از سوی پژوهشگران بیان گردیده است که ترکیب نتایج، هسته مرکزی تعریف آنها را تشکیل می‌دهد. در این

---

1. narrative

بخش به بررسی چند دیدگاه صاحب نظران در خصوص تعریف مفهوم فوق خواهیم پرداخت. گلاس که خود واضع واژه و مفهوم فرا - تحلیل در سال ۱۹۷۶ می باشد،<sup>(۴)</sup> می نویسد: «فرا - تحلیل اشاره به تحلیلِ تحلیل‌ها دارد... تحلیل آماری مجموعه وسیعی از نتایج تحلیل از مطالعات منفرد با هدف ادغام کردن یافته. این روش بدیل جدی برای بحث‌های علی، شرح جزئیات مطالعات تحقیقی است که کوشش‌های ما را برای معنی دار کردن ادبیات تحقیق را که به سرعت گسترش می‌یابد، راهنمایی می‌کند.<sup>(۵)</sup> دومینس و دیگران در مقاله‌ای فرا - تحلیل را یک مجموعه از رویه‌های آماری مطرح شده می‌دانند که برای انباشتن نتایج تجربی و همبستگی از مطالعات مستقل تهیه شده است. در این انباشتن به مجموعه سؤالات تحقیق توجه می‌شود. به نظر آنها فرا - تحلیل بی‌شبهت به روش‌های تحقیق سنتی، خلاصه‌های آماری را از مطالعات منفرد به عنوان داده به کار می‌برد. آنها معتقدند که فرض کلیدی فرا - تحلیل این است که هر مطالعه برآورد متفاوتی از رابطه اساسی درون جمعیت را فراهم می‌سازد. با انباشتن نتایج از مطالعات مستقل می‌توان ارائه صحیح‌تری از رابطه جمعیت را به دست آورد تا اینکه با برآوردهای مطالعات منفرد.<sup>(۶)</sup> در تعریف دیگر بلند می‌نویسد: روش فرا - تحلیل، ترکیب داده‌ها از چندین مطالعه برای تولید یک برآورد می‌باشد. از نقطه نظر آماری فرا - تحلیل کاربرد صریح<sup>(۷)</sup> روش‌های چند عاملی می‌باشد. به نظر وی ما چند مطالعه از یک چیز مشابه داریم که ممکن است، آزمایشهای کلینیکی یا مطالعات واگیرشناسی انجام شده در کشورهای متفاوت باشد. هر آزمایش به ما برآورد یک اثر را می‌دهد. ما فرض می‌کنیم که این برآوردها از ارزش کلی جمعیت مشابهی<sup>(۸)</sup> به دست آمده‌اند. لذا این فرضیات تحلیل را چک می‌کنیم و اگر قانع شدیم، برآوردهای مطالعات مجزا را ترکیب می‌کنیم تا به برآورد مشترکی برسیم.<sup>(۹)</sup> به نظر چپو و برینو فرا - تحلیل بر این اصل متکی است که نتایج تحقیقی که بر اساس متغیرهای مستقل هستند، می‌توانند ترکیب شوند تا یک مرور کلی موضوع مورد مطالعه و یک چشم‌انداز وسیع‌تر درباره اثرات عناصر طرح مورد آزمایش را ارائه دهند.<sup>(۱۰)</sup> بنابر تعاریف گلاس و دیگران می‌توان چنین بیان نمود که فرا - تحلیل روش تحلیلِ تحلیل‌های آماری مجموعه وسیع از نتایج مطالعات منفرد برای ادغام کردن آنها است. در این روش، نتایج که هم شامل نتایج درونی و هم بین نتایج می‌باشند، ترکیب می‌شوند. بدون یک پشتوانه قوی معرفتی

ترکیب نتایج ناهمگن آزمایش‌ها می‌تواند گمراه‌کننده باشد. به دلیل این ناهمگونی نوعی مناقشه‌آمیز<sup>۱</sup> جمعیت (آزمایشات) اغلب اعتبار فرا - تحلیل مورد نقد قرار گرفته است. فقدان چارچوب نظری مناسب از فهم عمیق فرا - تحلیل جلوگیری می‌کند.<sup>(۱۱)</sup>

## ۲- مقایسه روش فرا - تحلیل با روش‌های سنتی مرور تحقیقات (شرح جزئیات یا بررسی گزارشی)

جکسون در سال ۱۹۸۰ کیفیت روش‌های سنتی (شرح جزئیات) را برای ادغام تحقیقات بررسی کرده است. او نتوانست که ادبیات به خوبی توسعه‌یافته‌ای را درباره روش مرور کردن شرح جزئیات پیدا کند. آزمایش وی بر روی ۳۶ مرور مقالات از مجلات علوم اجتماعی مشهور دوره‌ای قصور و کاستی‌هایی را در حوزه‌های زیر را نشان داد:

۱- انتخاب سؤالات و فرضیات؛ ۲- نمونه‌گیری یا مشخص کردن و شامل نمودن مطالعات؛ ۳- ارزیابی مشخصات مطالعات اولیه؛ ۴- تحلیل مطالعات اولیه؛ ۵- تفسیر نتایج و ۶- گزارش مرور انجام شده.<sup>(۱۲)</sup>

در مقابل، عمده‌ترین ویژگی قابل شناسایی در روش فرا - تحلیل تبدیل نتایج مطالعات منفرد به اندازه‌ای مشترک همچون تفاوت استاندارد شده میانگین (اندازه اثر)<sup>۲</sup> یا ضریب همبستگی (بزرگی اثر) است که به وسیله تحلیل آماری مجموعه نتایج دنبال می‌شود. تحلیل اندازه اثر، حداقل از دو تا از نواقص شرح جزئیات سنتی که به عنوان روش شمارش آرای<sup>۳</sup> مشخص شده‌اند، اجتناب می‌کند. در روش شمارش آرای مطالعات یا به طور معنی‌داری مثبت، منفی و بدون معنی‌داری طبقه‌بندی می‌شوند.<sup>(۱۳)</sup> اولاً روش شمارش آرای می‌تواند خطای نتیجه‌گیری درباره مجموعه مطالعات را به هنگامی که اندازه‌های نمونه مطالعات منفرد و کوچک باشند، به دست دهد. به خصوص هنگامی که تعداد مطالعات بزرگ باشند.<sup>(۱۴)</sup> ثانیاً، روش شمارش آرا، بزرگی رابطه یافت‌شده در مجموعه مطالعات را برآورد نمی‌کند. کوتاهی در برآورد بزرگی‌های اندازه

1. problematic typical heterogeneity.

2. effect size.

3. vote-count.

به خصوص برای پژوهش سؤالات توصیفی پُردردسر است.<sup>(۱۵)</sup> قواعد تصمیم‌گیری غیررسمی که ویژگی اکثر مرورهای شرح جزئیات است، بر آرا منفرد متکی نیست، بلکه بر شمارش آرا استوار است، در جایی که هر رأی ممکن است، به واسطه مسایل اندازه نمونه دچار نقیصه و کمبود باشد. اما به طور کلی، محدودیت اصلی روش شرح جزئیات این است که در این روش طرز رفتار اندازه اثر فقط هنگامی کشف می‌شود که مطالعات مورد بررسی دارای اندازه نمونه برابر و توزیع تک‌نمایی<sup>۱</sup> نشان‌دهنده یک جمعیت باشند، موضوعی که معمولاً پدیده نسبتاً نادری است.<sup>(۱۶)</sup> اغلب اندازه نمونه‌ها از یکدیگر متفاوت و از توزیع‌های چندنمایی کشیده می‌شوند. به علاوه معنی‌داری آماری تابعی از طرز کار اندازه اثر و اندازه نمونه می‌باشد. لذا اگر اندازه نمونه‌ها متفاوت باشند، یافته‌های که از نمونه‌های وسیع به دست می‌آیند، حتی اگر طرز کار واقعی اندازه آنها کوچک و یا مساوی با اندازه سایر مطالعات باشد، به احتمال بیش‌تر معنی‌دار خواهند بود.<sup>(۱۷)</sup> قواعد تصمیم‌گیری فرا - تحلیل بیش‌تر رسمی هستند و از تحلیل آماری اندازه‌های اثر پیروی می‌کنند، در جایی که هر اندازه اثر تحت‌تأثیر مسایل اندازه نمونه قرار نگرفته‌اند.<sup>(۱۸)</sup>

### ۳- تعریف اندازه اثر

با توجه به اینکه مفهوم اندازه اثر مهم‌ترین واژه کلیدی در فرا - تحلیل می‌باشد و در این مقاله بارها تکرار خواهد شد، در اینجا با استفاده از تعریف کوهن سعی در روشن نمودن معنای این مفهوم خواهیم نمود. از آنجا که می‌باید به پرسش کمیّت به نحوی معنی‌دار پاسخ داده شود، می‌باید در خصوص شدت رابطه یا آنچه که اصطلاحاً اندازه اثر (ES) گفته می‌شود، به توافق برسیم و از سوی دیگر، برای تبیین بزرگی یک رابطه نیز لازم است که آنها را به شکل کمی بیان نماییم. لذا اندازه اثر می‌تواند ما را در این امر بسیار یاری نماید. در کتب تحلیل توان آماری برای علوم رفتاری، کوهن اندازه اثر را چنین تعریف می‌نماید:

بدون آنکه قصد داشته باشیم، مفهوم ضمنی علیّت را مراد کنیم، راحت‌تر آن است که عبارت «اندازه اثر» را به یکی از دو معنای زیر به کار ببریم: «حدی که

1. unimodal distribution.

پدیده موردنظر در جمعیت موجود است» یا «حدی که فرضیه صفر، نادرست است». با طی این مسیر، اینک روشن می‌شود که وقتی فرضیه صفر، نادرست باشد، در واقع، تا حد معینی نادرست است، یعنی اندازه اثر به میزان معین غیرصفر در جمعیت موجود است. هرچه این مقدار بیش تر باشد، میزان بروز پدیده مورد مطالعه در جمعیت هم بیش تر خواهد بود.<sup>(۱۹)</sup>

در حال اندازه اثر نوعی از نمره معیار است که ترکیب نتایج از مطالعات را اجازه می‌دهد و معمولاً به وسیله تقسیم تفاوت بین میانگین گروه تجربی و گروه کنترل بر انحراف معیار گروه کنترل تعیین می‌شود.

#### ۴- کاربرد روش فرا - تحلیل

در اجرای روش فرا - تحلیل انواع مختلفی از رویه‌ها به کار رفته است. روزنتال کاربردهای اولیه رویه‌های فرا-تحلیلی را به سه نوع به شرح زیر تقسیم کرده است: (۱) در اولین نوع، هدف خلاصه کردن مجموعه مطالعاتی بود که روابط کلی بین دو متغیر را در هر مطالعه جست‌وجو می‌کرد. اغلب این هدف با تلاش برای برآورد یا برآورد میانگین روابط بین دو متغیر یافت شده در مجموعه مطالعات صورت می‌گرفت. گاهی این هدف با آزمون معنی‌داری، یعنی با تلاش برای تعیین احتمال رابطه به دست آمده بررسی می‌شد. (۲) دومین نوع کاربرد اولیه روش‌های فرا-تحلیل چندان به خلاصه نمودن رابطه بین دو متغیر مربوط نمی‌شد، بلکه با تعیین عواملی سروکار داشت که با تغییرات در بزرگی روابط بین دو متغیر مرتبط بودند (یعنی عواملی که به عنوان متغیرهای تعدیل‌کننده استفاده می‌شوند). (۳) سومین کاربرد، هیچگونه رابطه‌ای را در درون هیچ مطالعه‌ای آزمایش نمی‌نمود. در عوض هر مطالعه فقط داده‌های انبوه شده<sup>۱</sup> برای هر متغیر تهیه می‌کرد. برای مثال، متوسط نگرش مشارکت‌کنندگان در مطالعه یا متوسط عملکرد شناختی آنها. سپس این داده‌های انبوه شده یا متوسط با یکدیگر و یا با دیگر خصیصه‌های مطالعه برای آزمون فرضیه‌ها یا برای پیشنهاد فرضیه‌هایی که می‌بایست در مطالعات به طور ویژه طرح‌ریزی شده بعدی آزمون شوند، مورد آزمون همبستگی قرار می‌گیرند.<sup>(۲۰)</sup>

1. aggregated data.

در بین این کاربردهای اولیه تفاوت‌هایی وجود دارد که به‌طور خلاصه می‌توان آنها را در سه مورد، کلی خلاصه نمود: ۱) نتیجه کاری نوع اول، بیش‌تر برآوردی از متوسط همبستگی (یا سطح  $p$  ترکیب شده مرتبط با آن همبستگی) بود که در همه مطالعات خلاصه شده، یافت می‌شود. ۲) نتیجه نوع دوم، عموماً همبستگی بین برخی خصیصه‌های مطالعات و همبستگی (یا دیگر شاخص‌های اندازه اثر) یافت شده در مطالعات بود. ۳) نوع سوم، به‌طور ساده میانگین داده‌های اخذ شده از هر مطالعه با دیگر میانگین داده‌ها یا با دیگر خصیصه‌های اخذ شده از هر مطالعه را مورد آزمون همبستگی قرار می‌داد.<sup>(۲۱)</sup>

#### ۵ - مراحل انجام فرا - تحلیل

برای انجام فرا - تحلیل مراحل مختلفی را باید پیمود که برخی از مراحل مختص به خود آن می‌باشد، در اینجا مراحل مختلف اجرای این روش تحقیق را از دیدگاه برخی از پژوهشگران مورد بررسی قرار داده و سپس کاربردها و تکنیک‌های آن را به‌طور مبسوط دنبال خواهیم کرد. بلند معتقد است که در فرا - تحلیل دو مرحله وجود دارد: اول ما کنترل می‌کنیم که مطالعات مورد بررسی برآوردها یک چیز مشابهی را فراهم ساخته‌اند. دوم، برآورد مشترک و فاصله اطمینانش را محاسبه می‌کنیم. برای این کار ممکن است که داده‌های اصلی همه مطالعات را داشته باشیم که می‌توان همه آنها را در یک فایل بزرگ داده‌ها ترکیب کنیم یا فقط ممکن است، آمارهای منشر شده آنها را خلاصه کنیم.<sup>(۲۲)</sup>

در مقاله تحقیقی دیگر دوماینس و دیگران در تشریح مراحل مختلف فرا - تحلیل به شکل دستورالعملی چنین ذکر می‌کنند:

۱. تعریف حوزه تحقیق: متغیرهای مستقل، متغیرهای تحقیق شده به‌طور مشترک، علل و نتایج متغیرهای مهم.
۲. استقرار ملاکی برای مطالعات شامل در مرور: مطالعات چاپ شده در مقابل مطالعات چاپ نشده، دوره زمانی که مرور را می‌پوشاند، تعاریف عملیاتی متغیرها، کیفیت مطالعه، سایر.
۳. تعریف نوع اندازه اثر برای استفاده: کوهن، ضریب گشتاوری پیرسون.

۴. جستجو برای مطالعات مربوطه: جستجوی رایانه‌ای، جستجوی دستی، کنفرانس و سمپوزیوم تکنیکی ارایه شده، نامه‌ها برای پژوهشگران در حوزه مورد مطالعه.
۵. انتخاب مجموعه مطالعات نهایی: انجام انتخاب به‌طور فردی، به‌وسیله بیش از یک نفر.
۶. استخراج داده‌ها در مورد متغیرهای مورد علاقه، اندازه‌های نمونه، اندازه‌های اثر، پایایی اندازه‌گیری و سایر خصیصه‌های برجسته هر مطالعه: وقتی که اندازه‌های چندگانه گزارش شده‌اند، همه داده‌ها را استفاده کنید، زیرمجموعه‌های داده‌ها را به‌کار گیرید، میانگین اندازه‌های مطالعات چندگانه را بگیرید تا یک اندازه نتیجه شود.
۷. خصیصه‌های هر یک از مطالعات که ممکن است مرتبط با اندازه اثر گزارش شده در مطالعه باشد، کدگذاری کنید: عوامل طرح تحقیق، خصیصه‌های نمونه، نوع متغیر وابسته و غیره.
۸. کنترل‌های پایایی را بر رویه‌های کدگذاری اجرا کنید: با یک زیرمجموعه داده‌ها، ۱ تا ۴ کدگذار دیگر را استفاده کنید، با همه داده‌ها، ۱ تا ۴ کدگذار دیگر را استفاده کنید.
۹. وقتی اندازه‌های چندگانه‌ای از متغیرهای مستقل و با وابسته وجود دارد، تصمیم بگیرید که آیا آنها را بر اساس تقدم گروه‌بندی شوند یا خیر: گوناگونی نظری<sup>۱</sup> میان متغیرها، اندازه‌گیری عملیاتی گوناگونی میان متغیرها.
۱۰. میانگین و واریانس اندازه‌های اثر از روی مطالعات را تعیین کنید: میانگین اندازه اثر تعدیل شده<sup>۲</sup> به وسیله اندازه نمونه، آزمون مربع خبی برای همگنی محاسبه کنید، N جبرانی<sup>۳</sup> را محاسبه کنید. واریانس بین مطالعات در اندازه اثر برای تعیین متغیرهای تعدیل‌کننده، برآورد منابع مصنوعی<sup>۴</sup> واریانس بین

---

1. theoretical diversity.

2. mean effect size weighted

3. fail safe.

4. artifactual sources.



مطالعات (خطای نمونه‌گیری، تضعیف<sup>۱</sup> به دلیل خطای اندازه‌گیری، و/یا دامنه محدودیت).<sup>۲</sup> برآورد حقیقی<sup>۳</sup> بین واریانس مطالعات. برآورد حقیقی میانگین اندازه اثر تصحیح‌شده، برای اندازه‌گیری و خطای نمونه‌گیری، و دامنه محدودیت.

۱۱. تصمیم بگیرید که آیا برای متغیرهای تعدیل‌کننده جست‌وجو می‌کنید یا خیر: آزمون معنی‌داری (آزمون خی دو)، مقدار واریانس بین مطالعاتی که مصنوعی است. قاعده تجربی<sup>۴</sup>: اگر واریانس محاسبه شده به وسیله واریانس خطا کمتر از ۷۵٪ واریانس همبستگی‌های تعدیل‌شده نمونه باشد، بنابراین ممکن است متغیر تعدیل‌کننده باشد در غیر این صورت تغییر بیش‌تر به دلیل خطای تصادفی می‌باشد (برای مثال، دامنه محدودیت، خطای نمونه‌گیری، یا خطای اندازه‌گیری).

۱۲. تعدیل‌کننده‌های بالقوه را انتخاب کنید: توجهات تئوریک، توجهات اندازه‌گیری عملیاتی میانگین و واریانس اندازه اثر درون زیرمجموعه‌های تعدیل‌کننده را تعیین کنید. رویه مشابه مرحله ۱۰ است.<sup>(۲۳)</sup> کولیک در مقاله‌ای چهار مرحله اساسی در انجام فراتحلیل را به شرح زیر توصیف می‌کند:

(۱) مرورکنندگان می‌بایست، اول مطالعات یک موضوع<sup>۵</sup> را بررسی کرده و سپس رویه‌های ویژه را به‌طور روشن به‌کار می‌بندند.

(۲) خصایص و ویژگی‌های نتایج مطالعات با اصطلاحات کمی تعیین می‌شوند.

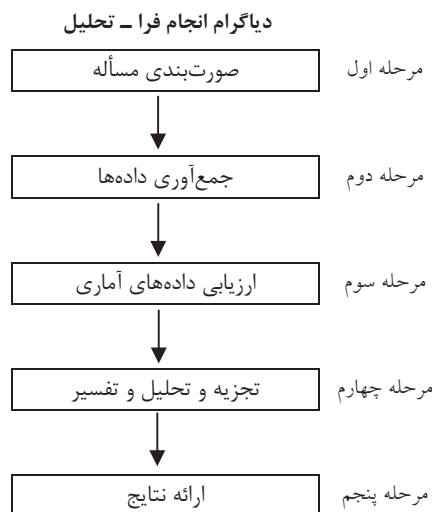
(۳) تا آنجا که ممکن است ویژگی‌های مطالعات کدگذاری می‌شوند.

(۴) رویه‌های آماری برای خلاصه کردن دامنه‌ها و ویژگی‌های مطالعات به جهت مطالعه نتایج به کار بسته می‌شوند.<sup>(۲۴)</sup>

در این مقاله فرا - تحلیل در پنج مرحله عمده به ترتیب خلاصه می‌گردد:

- 
1. attenuation.
  2. range restriction.
  3. true.
  4. rule of thumb.
  5. issue.

(۱) صورت‌بندی مسأله، (۲) جمع‌آوری داده‌ها، (۳) ارزیابی داده‌های آماری، (۴) تجزیه و تحلیل و تفسیر، و (۵) ارائه نتایج.



#### ۶ - مرحله نخست: صورت‌بندی مسأله

در این مرحله، محقق می‌باید برای متغیرهای موجود تعاریف مفهومی و یا نظری و هم‌چنین تعاریف عملیاتی را ارائه نماید. سؤالاتی که در این مرحله برای پژوهش‌گر مطرح می‌باشد، از این قرار است: (۱) چه عملیاتی برای مفاهیم مربوط به این بررسی مناسب می‌باشد؟ مهم‌تر از این، (۲) کدامیک از منابع برای این بررسی مربوط و کدامیک نامربوطند؟ در این مرحله، هم‌چنین می‌باید فرضیه‌هایی که به تعامل میان دو یا چند متغیر مربوط هستند، ارائه گردد.

در این مرحله متغیرها می‌باید، در ابتدا از بعد مفهومی یا نظری تعریف گردند و ثانیاً از بعد عملیاتی و یا عینی. در این مرحله باید به نکات زیر توجه نمود: (۱) تعاریف مفهومی به لحاظ درجه انتزاع تعداد مواردی که دربر می‌گیرند از یکدیگر متفاوتند. لذا پژوهشگر می‌باید به این امر توجه داشته باشد که هر چه سطح انتزاع کلی‌تر باشد، تعداد منابعی که مورد بررسی فرا می‌گیرد، بیش‌تر خواهد شد و از سوی دیگر، هرچه سطح انتزاع کم‌تر شود تعداد منابع کمتر خواهد شد. بنابراین، بر حسب اهداف و نیات پژوهش‌گر، وی می‌باید دست به انتخاب و گزینش بزند. (۲) برای آنکه مفاهیم را به

سطح عینی مرتبط سازیم، می‌باید آنها را تعریف عملیاتی نماییم. در اینجا این نکته حایز اهمیت است که پژوهشگر می‌باید تعیین کند که تا چه حد احتمال آن می‌رود که یک رویداد نشان‌دهنده نمونه‌های از متغیر مورد نظر وی باشد.

هنگامی که سعی در مشخص نمودن حدود و ثغور مطالعه می‌نماییم تا آنها را در فراتحلیل وارد کنیم، می‌باید تصمیم بگیریم که چه اطلاعاتی از هر سند و مدرکی باید خلاصه شود.

آشکار است که همیشه می‌خواهیم سطح اطمینان و اندازه اثر را ضبط و ثبت کنیم. اگر هر یک از این دو درست نباشد، می‌توانیم آن را، اگر اندازه مطالعه را بدانیم برآورد کنیم. اما چه چیز دیگری می‌باید از هر مطالعه ثبت و ضبط کنیم. پاسخ به این سؤال مسلماً بستگی به اهداف فراتحلیل مورد دارد. لذا در این مرحله کار ما تهیه برگه رمزگذاری است. برگ رمزگذاری برای جمع‌آوری اطلاعات از گزارش‌های مربوط به پژوهش‌های قبلی مورد استفاده قرار می‌گیرد. در اینجا لازم است که بررسی‌کننده دقیقاً در نظر داشته باشد که از هر گزارش چگونه داده‌هایی به دست خواهد آمد. مهم این است که برگ رمزگذاری به شکلی تغییر یابد که در یک بار قرائت هر گزارش پژوهشی، بررسی استاندارد و دقیق آن گزارش ممکن باشد. قواعد مربوط به درست کردن یک برگه رمزگذاری به همان قواعدی شباهت دارد که در چارچوب‌های رمزگذاری و جداول عددی اطلاعات در انجام پژوهش‌های اصلی مورد استفاده قرار می‌گیرد. در هنگام تهیه برگه رمزگذاری به این نکته می‌باید توجه نمود که هر گونه اطلاعات، حتی با کمترین احتمال در داشتن ارتباط با بررسی مورد نظر، می‌باید در مطالعات انجام شده قبلی در نظر گرفته شود.<sup>(۲۵)</sup> در اینجا سه نمونه از اشکال مفید برای خلاصه کردن اطلاعات از مطالعات ارایه می‌شود:

در مقاله تحقیقی که توسط هانسون و دیگران تحت عنوان پیش‌بینی رجعت خطاکاران جنسی<sup>۱</sup> یک فراتحلیل، جهت انجام تحقیق خود رویه کدگذاری زیر را ارائه نمودند:

مثال دوم، از تحقیقی که هریس کوپر به سال ۱۹۸۴ بر روی گره‌های قومی و تفاوت‌های طبقات اجتماعی در نیاز به موفقیت انجام داده است، آورده می‌شود: هریس کوپر (۱۹۸۴) در مقایسه گروه‌های قومی و سطح طبقات اجتماعی بر نیاز به موفقیت

1. predictor of sexual offender recidivism

کار خود را متمرکز کرده بود. کدگذاری وی در تحقیق فوق به شرح زیر می‌باشد:

(۱) استناد کامل؛<sup>۱</sup> (۲) منبع مرجع؛<sup>۲</sup> (۳) جنس آزمودنی، با تعداد هر یک برای هر دو گروه مورد مقایسه؛<sup>۴</sup> (۴) متوسط سن آزمودنی‌ها در هر گروه؛<sup>۵</sup> (۵) موقعیت جغرافیایی هر گروه؛<sup>۶</sup> (۶) سایر قیود؛<sup>۷</sup> (۷) قومیت هر گروه؛<sup>۸</sup> (۸) میانگین (M) و انحراف معیار (s) هر گروه قومی بر اساس نیاز به موفقیت؛<sup>۹</sup> (۹) نوع آزمون معنی‌داری و خطای درجه آزادی (df error) به‌کار رفته؛<sup>۱۰</sup> (۱۰) ارزش (مقدار) آزمون آماری به دست آمده و اثر (df؛ ۱۱) سطح p و اندازه اثر به دست آمده؛<sup>۱۲</sup> (۱۲) نتایج مستقیم؛<sup>۱۳</sup> (۱۳) طبقه اجتماعی هر گروه؛<sup>۱۴</sup> (۱۴) اندازه‌های استاندارد شده؛<sup>۱۵</sup> (۱۵) اساس طبقه‌بندی هر یک از گروه‌ها شامل شغل، دستمزد، پایگاه اجتماعی و سایر؛<sup>۱۶</sup> (۱۶) میانگین و s هر گروه طبقه اجتماعی بر اساس نیاز به موفقیت؛<sup>۱۷-۲۰</sup> (۱۷-۲۰) فقره‌های ۹ تا ۱۲ (بالا) تکرار شده، برای مقایسه طبقه اجتماعی؛<sup>۲۱</sup> (۲۱) اندازه‌های وابسته شامل (n-ach)، TAT، آزمون فرانسوی بینش<sup>۵</sup> یا سایر؛ متغیرهایی که با قومیت یا طبق اجتماعی تعامل دارند.<sup>(۲۶)</sup>

همان‌طور که روزنتال نیز یادآوری می‌کند، ساختن یک شکل برای چکیده کردن نتایج تحقیق با استفاده از فرمت‌های انجام شده پیشین بهتر صورت می‌گیرد. فرمت‌های جدید را می‌توان با هم‌کلاسی‌ها و یا مشاورانی که می‌توانند سایر متغیرها مطرح کنند، مورد مطالعه قرار داد. در نهایت فرم تجدیدنظر شده را می‌توان برای کسانی که در حوزه فرا - تحلیل کار می‌کنند فرستاد به این امید که آنها سایر متغیرهایی که می‌باید کدگذاری شوند را پیشنهاد کنند.<sup>(۲۷)</sup>

## ۷- مرحله دوم: جمع‌آوری داده‌ها

این مرحله با انجام یک انتخاب در مورد تعداد عناصری که در بررسی می‌باید مورد توجه خاص قرار گیرد، سروکار دارد. تشخیص تعداد این عناصر کار مشکلی است، زیرا، اولاً بررسی‌کننده می‌خواهد که انبوه نتایج بررسی، بر تمام پژوهش‌های انجام شده

- 
1. complete citation.
  2. source of reference.
  3. restrictions.
  4. standardized measure.
  5. insight.

قبلی در مورد این موضوع مبتنی باشد و ثانیاً امیدوار است که مطالعات یادشده امکان تعمیم را در میان افراد (یا عناصر) که موضوع این بررسی هستند فراهم سازد.<sup>(۲۸)</sup> در این مرحله به منابع مورد استفاده توجه می‌شود و اهمیت هر یک از آنها مدنظر قرار می‌گیرد.

در روش فرا - تحلیل مسائل عمده و اصلی قبل از این که ما شروع به تحلیل داده‌ها بکنیم برمی‌خیزد. اول، ما باید تعریف روشنی از سؤال داشته باشیم. بنابراین ما فقط مطالعاتی را که به این امر توجه کرده باشند را دربرمی‌گیریم. چنین مطالعاتی که ما از آنها استفاده می‌کنیم، ممکن است تأثیر عمیقی بر نتیجه‌گیری داشته باشند.<sup>(۲۹)</sup> معمولاً در فرا - تحلیل می‌باید همه مطالعات مربوطه را داشته باشیم و به یک جستجوی ساده ادبیات اکتفا نکنیم. همه مطالعات ممکن است در دسترس نباشند. آنهایی که تفاوت معنی‌داری را تولید کرده‌اند ممکن است چاپ شده باشند ولی آنهایی که تفاوت معنی‌داری را به دست نیاورده‌اند، گزارش و چاپ نشده باشند.<sup>(۳۰)</sup> همچنین در یک مطالعه نتایجی که معنی‌دار هستند ممکن است مورد تأکید قرار گیرند و آن بخش‌هایی از داده‌ها که تفاوتی را تولید نکرده‌اند ممکن است که فراموش شوند و توسط محقق و یا محققان بعنوان موضوعاتی که جالب نیستند تلقی گردیده و چاپ نشوند و یا اینکه چاپ آنها توسط متولیان مورد مخالفت و ممانعت شود.<sup>(۳۱)</sup> هم‌چنین گه‌گاه اتفاق می‌افتد که پژوهش‌گری برخی از نتایج تحقیقات بعدی خود را که با نتایج تحقیقات قبلی وی مخالف است ارائه ننماید. بطور کلی منابع مورد استفاده در فرا - تحلیل عبارتند از کتب، مقالات، رساله‌ها، کارهای چاپ نشده. هنگامی که ما همه مطالعاتی را که با تعریف ما جور درمی‌آیند داشتیم، ما آنها را برای رسیدن به برآورد مشترکی از تأثیر طرز عمل یا عامل خطر ترکیب می‌کنیم.

لذا این مطالعات برای تحلیل بعدی، بعنوان داده مبنا<sup>۱</sup> مورد استفاده قرار می‌گیرند. مطالعاتی که پایایی سنجش‌های مورد استفاده را در مطالعات گوناگون و انحراف معیار آنها (برای دامنه انحراف تعديلات)<sup>(۳۲)</sup> در هر یک یا هر دو متغیرها را گزارش می‌دهند، در داده‌های مبنا گنجانده می‌شوند.<sup>(۳۳)</sup>

هنگامی که داده‌ها را به آمار مشترک تبدیل نمودیم، پایایی و دامنه اطلاعات انباشته می‌شوند. اگر همه مطالعات شامل برآوردهای پایایی یا دامنه داده‌ها باشند، سپس اندازه اثر در سطح مطالعه انفرادی می‌تواند تصحیح شود. به هر حال، اکثر مطالعات این اطلاعات را فراهم نمی‌سازد. تحقیق اطلاعات روزنتال ضمن طبقه‌بندی اسناد مورد مطالعه به کتب، مجلات، رساله‌ها و کارهای چاپ نشده پایایی آنها را به صورت شش جفت ممکن منابع اطلاعات به همراه تعداد برآوردهای اندازه اثر فرا - تحلیل را ارائه نموده است. هم‌چنین وی پایایی به دست آمده بین هر دو منبع از هر جفت و سطح  $p$  پایایی به دست آمده نیز نشان می‌دهد.

جدول شماره ۱: پایایی منابع اطلاعات برای نمونه فراتحلیل<sup>(۳۴)</sup>

سطح $p$ پایایی	پایایی $r$ منابع	تعداد ( $n$ ) فرا - تحلیل‌ها	جفت‌های منابع
۰/۰۰۵	۰/۸۹	۱۰	مجله؛ رساله
۰/۰۰۶	۰/۶۵	۷	مجله؛ چاپ‌نشده
۰/۰۰۸	۰/۸۵	۷	رساله؛ چاپ‌نشده
۰/۰۲۵	۰/۸۲	۶	کتاب؛ مجله
۰/۰۲	۰/۹۶	۴	کتاب؛ رساله
۰/۰۰۵	۱/۰۰	۳	کتاب؛ چاپ‌نشده
۰/۰۱۴	۰/۸۷	۶/۵	میانه
۰/۰۰۸	۰/۸۵	۷	میانه وزنی (تعدیل‌شده)
-	۰/۸۳	-	میانگین ( $n-2$ )

میانه (غیروزنی و وزنی) با پایایی ۰/۸۷ و ۰/۸۵ و میانگین وزنی  $r$ ، ۰/۸۳ نشان‌دهنده آن است که درجه بالایی از پایایی بطور متوسط بین جفت‌های گوناگون منابع اطلاعات وجود دارد.<sup>(۳۵)</sup> همان‌طور که روزنتال تذکر می‌دهد، پایایی بالای منابع اطلاعات بدین معنی نیست که لزوماً برآوردهای اندازه اثر آنها در یک مرور فراتحلیلی معین مشابه باشد. دو منبع می‌تواند پایایی کاملی داشته باشند ( $r=1/00$ ) اما تفاوت بالایی در برآوردهای اندازه اثر آنها وجود داشته باشد.<sup>(۳۶)</sup>

#### مرحله سوم: ارزشیابی داده

پس از آن که اطلاعات جمع‌آوری شدند می‌باید کیفیت هر یک از داده‌های آماری ارزیابی نمود. هر یک از داده‌های آماری را با مدارک و شواهد پیرامونی بررسی کرد تا

مبادا عواملی غیرمرتبط با موضوع پژوهش آن را به انحراف کشانده باشند.<sup>(۳۷)</sup> در این مرحله پژوهشگر در تلاشی علمی سعی در قضاوت در مورد ورود داده‌های آماری مورد بحث در تحقیق خود می‌نماید. ارزشیابی داده‌ها مستلزم ایجاد معیارهایی برای قضاوت در مورد کارایی روش شناختی چگونگی تولید اطلاعات می‌باشد. پژوهشگر در این مرحله می‌باید تمامی عوامل بالقوه مؤثر را برای هر یک از داده‌های آماری مورد توجه قرار دهد؛ عواملی که می‌توانند رابطه میان یک داده آماری را با مسأله مورد نظر از بین ببرند. سپس محقق باید تشخیص دهد که آیا این عوامل آنقدر مهم هستند که لازم باشد آن داده آماری از تحقیق کنار گذاشته شود یا نه.<sup>(۳۸)</sup>

در پژوهش‌های اولیه تک‌تک داده‌های آماری یا توزیع نمونه‌ای مورد مقایسه قرار می‌گیرند تا معلوم گردد که آیا تفاوت آنقدر زیاد هست که اعتبار تحقیق را زیر سؤال ببرد یا نه، در بررسی‌های پژوهش‌ها نیز ردیابی فاصله‌های آماری با مطالعه دامنه رابطه سروکار دارد تا تفاوت میان شدت رابطه در یک بررسی. هم‌چنین در هر نوع بررسی می‌توان با محاسبه مجدد نتایج یا مطالعه ارزش‌گذاری‌های خام، خطای در ثبت آمار و ارقام را تشخیص داد. بررسی‌کننده ممکن است جهت کنترل محاسبات محقق اولیه به محاسبه مجدد ارزش‌های آزمون‌های آماری بپردازد. در بررسی پژوهش‌های انجام شده علاوه بر ارزیابی ناهمگنی و خطا تنها یک معیار بالقوه، برای کنار گذاشتن اطلاعات وجود دارد و آن اعتبار شیوه‌های مورد استفاده در پژوهش است. بررسی‌کننده باید ببیند که آیا مطالعات مورد بررسی با دقت انجام شده است تا بتوان به نتایج آنها اعتماد کرد یا خیر.<sup>(۳۹)</sup>

در مرحله ارزیابی باید توجه داشت که علی‌رغم اینکه در تحقیقات اولیه نظام‌های دقیق نظارتی وجود دارند تا از نتایج مصنوعی که معلول اثر انتظار آزمایشگر است، جلوگیری نمایند. در بررسی پژوهش‌های انجام شده چنین نظام نظارتی وجود ندارد. عبارت دیگر بررسی‌کنندگان در مورد موضع‌گیری‌های خود و جهت‌گیری نتایج در مرحله ارزشیابی آگاهی کامل دارند. لذا احتمال این که ارزشیابی تحقیق تحت تأثیر نتایج حاصل از آن قرار گیرد، وجود دارد.<sup>(۴۰)</sup> و همانطور که گلاس می‌گوید کاری که در مرحله بعدی می‌باید انجام شود، پذیرفتن یکی دو پژوهش «قابل قبول» به‌عنوان حقیقت موضوع است.<sup>(۴۱)</sup> به نظر کوپر می‌باید پژوهش‌ها را براساس بررسی استقرایی تفاوت در پژوهش‌های گوناگون حذف نمود تا براساس شیوه اشتتاجی زیرا در شیوه استتجاجی

ذهنیت بررسی کننده می تواند بر ارزیابی وی از کیفیت متون پژوهشی اثر بگذارد.<sup>(۴۲)</sup> گلاس و اسمیت معتقدند که چنانچه نتایج حاصل از پژوهش های مختلف همخوانی داشته باشند، ویژگی های ضعیف مربوط به طراحی هر پژوهش، پژوهش های دیگر را نفی می کند. ثانیاً داوری کیفی استنتاجی لازم برای حذف بعضی از پژوهش ها، مورد به مورد فرق می کند و از موضع گیری های شخصی بررسی کننده تاثیر می پذیرد. آنها نهایتاً مدعی می شوند که از نادیده گرفتن معیارهای کیفی دفاع نکرده اند، بلکه تأثیر کیفیت طراحی را بر نتایج پژوهش یک مسأله استنتاجی تجربی، و نه یک نظرگاه استنتاجی تلقی کرده اند.<sup>(۴۳)</sup> به نظر کوپر اگر به شکل تجربی نشان داده شود که پژوهش های انجام شده با روشهای «خوب»، نتایجی متفاوت از پژوهش های انجام شده با روشهای «بد» دارند، آنگاه می توان نتایج حاصل از پژوهش های «خوب» را باور کرد. اما اگر تفاوتی وجود نداشته باشد، معقول این است که پژوهش های «بد» را هم در نظر بگیریم چرا که متضمن تنوع هایی در روش هستند (از قبیل تنوع در نمونه ها و مکان ها) که استفاده از آنها به پاسخ گویی پرسش های موجود در پیرامون حوزه موضوعی کمک می کند.<sup>(۴۴)</sup>

#### مرحله چهارم: تجزیه و تحلیل و تفسیر

در این مرحله داده های آماری جداگانه ای که به وسیله محقق گردآوری شده با یکدیگر تلفیق می شوند. گزاره واحدی را درباره مسأله به وجود می آورند. هم چنین تفسیر مستلزم آن است که محقق، الگوهای اطلاعاتی روشمند را از عوامل اختلال زا یا نوسانات اتفاقی بازشناسی کند.<sup>(۴۵)</sup>

در این بخش روشهای تبدیل آمارهای مطالعه به اندازه اثر، برآورد اندازه اثرها، برآورد سطوح معنی داری، مقایسه اندازه اثرها، مقایسه سطوح معنی داری، ترکیب نتایج، ترکیب سطوح معنی داری، ترکیب احتمالات را بررسی خواهیم کرد.

#### تبدیل آمارهای مطالعه به اندازه اثر

هنگامی که داده های مبنا یا پایگاه اطلاعاتی فراهم گردید، لازم است که آمارهای مطالعات منفرد و جدا را به اندازه مشترک برای انباشتگی<sup>۱</sup> بعدی تبدیل کرد (خواه  $r$  یا  $d$ ).<sup>(۴۶)</sup>

1. accumulation.



سنجش اندازه اثر بستگی به مسأله و حتی عادات حوزه‌های متفاوت تحقیق دارد.<sup>(۴۷)</sup> مطالعات شبه تکراری ممکن است مقیاسهای متفاوتی از سنجش را به کار می‌برند. ما نیاز به شاخصی از اندازه اثر داریم که بستگی به درجه بندی دلخواهی متغیر وابسته نداشته باشد.<sup>(۴۸)</sup> گلاس در سال ۱۹۷۶، ۲ مقیاس شاخص آزاد<sup>(۴۹)</sup> را پیشنهاد می‌کند:

(۱) انحراف تفاوت میانگین کوهن این شاخص اندازه اثر کاربرد بیش تری دارد.

(۲) ضریب همبستگی گشتاوری  $r$

جداول ۲ و ۳ فرمول‌ها و رویه‌های تبدیل آماره‌های مطالعات را به آماره‌های  $r$  و  $d$

مرور می‌کند:<sup>(۵۰)</sup>

جدول ۲: فرمول‌ها و رویه‌ها برای تبدیل آماره‌های مطالعه به  $r$

معادله	آماره‌ای که می‌باید تبدیل شود	فرمول برای تبدیل به $r$	توضیح
۱	t	$r = \sqrt{\frac{2}{t^2 + df}}$	می‌توان هم برای آزمون‌های t جفت‌شده یا غیرجفت‌شده به کار برد.
۲	F	$r = \sqrt{\frac{F}{F + df(e)}}$	فقط برای تحلیل واریانس یکطرفه به کار می‌رود
۳	تحلیل واریانس دو طرفه	$r = \sqrt{\frac{(fa * dfa)}{(fa * dfa) + (fb * dfb) + (fab * dfab) + df(e)}}$	اثر اصلی مورد نظر fa = dfab = dfb = fab = dfab = df(e) = درجه آزادی برای A دومین اثر اصلی درجه آزادی برای B تعامل اثرها تعامل درجه آزادی خطای درجه آزادی
۴	$\chi^2$	$r = \sqrt{\frac{2}{\chi^2 + N}}$	اندازه نمونه N = فقط برای وقتی که درجه آزادی مساوی ۱ با df=۱ باشد.
۵	d	$r = \frac{d}{d^2 + \frac{4(N-2)}{N}}$	d کوهن = اندازه‌های نمونه ترکیب شده = N
۶	p	(۱) تبدیل ارزش p دو دنباله به p یک دنباله (برای مثال $p/2$ ) (۲) جستجوی Z مرتبط شده در جدول احتمال طبیعی	خواه ارزشهای p دقیق یا وقتی که نویسنده p تقریبی (برای مثال $p < 0.05$ ) گزارش می‌دهد، می‌توان به کار برد.

جدول شماره ۳: فرمول‌ها و رویه‌ها برای تبدیل آماره‌های مطالعه به  $d$ 

توضیحات	فرمول برای تبدیل به $d$	آماره‌ها برای تبدیل	مورد
$\bar{X}_e =$ میانگین گروه تجربی $\bar{X}_c =$ میانگین گروه کنترل $sp =$ انحراف معیار ادغام شده (درون موضوعات مورد مطالعه)	$d = \frac{\bar{X}_e - \bar{X}_c}{sp}$	میانگین‌ها و انحراف معیارها	۷
$N_e =$ تعداد گروه تجربی (آزمایشی) $N_c =$ تعداد گروه کنترل $S_e^2 =$ واریانس گروه تجربی $S_c^2 =$ واریانس گروه کنترل	$S_p^2 = \frac{(N_e - 1) S_e^2 + (N_c - 1) S_c^2}{(N_e + N_c - 2)}$	واریانس درون آزمودنی‌های ادغام شده	۸
می‌توان با آزمونهای $t$ جفتی یا غیر جفتی به کار برد.	$d = \frac{2t}{\sqrt{df}}$	$t$	۹
فقط با تحلیل واریانس یکطرفه به کار می‌رود.	$d = \frac{2\sqrt{F}}{\sqrt{df(\text{error})}}$	$F$	۱۰
—	$d = \frac{2r}{\sqrt{1-r^2}}$	$r$	۱۱

## اندازه اثر و معنی‌داری آماری

هنگامی که ما در نتیجه یا نتایج تحقیق صحبتی به میان می‌آوریم، منظورمان نتایجی که پژوهشگر اولیه به دست آورده نیست زیرا اغلب بطوری مبهم به نتایج واقعی مرتبط هستند. اغلب اتفاق می‌افتد که بین بخش نتایج<sup>۲</sup> و بحث و نتیجه‌گیری<sup>۳</sup> تغییر و دگرگونی<sup>۴</sup> رخ می‌دهد. همچنین منظور از نتیجه آزمون  $F$  مختلط<sup>۵</sup> با  $df > 1$  در صورت کسر<sup>۶</sup> یا آزمون<sup>۲</sup> مختلط با  $df > 1$  نیز نیست زیرا در هر دو مورد ما جواب‌های کمی برای پاسخگویی که اغلب مبهم<sup>۷</sup> هستند، به دست می‌آوریم. لذا منظور از نتیجه تحقیق پاسخ به سؤالات زیر می‌باشد. چه رابطه‌ای بین هر متغیر  $x$  و با هر متغیر  $y$  وجود دارد؟ متغیرهای  $x$  و  $y$  فقط با این الزام که روابطشان مورد علاقه ما می‌باشند، انتخاب خواهند

1. subject.
2. results.
3. discussion.
4. metamorphosis.
5. omnibus.
6. numerator.
7. imprecise.

شد. بنابراین پاسخ به این سؤال دو بخش خواهد آمد: ۱) برآورد بزرگی روابط<sup>۱</sup> یا اندازه اثر و ۲) نشانه صحت<sup>۲</sup> یا پایایی<sup>۳</sup> اندازه اثر برآورد شده. از بحث بالا به این مسأله می‌رسید که اهمیت آماری اندازه اثر چیست و رابطه بین اندازه اثر و آزمون معنی‌داری چیست فرمول زیر نشان‌دهنده رابطه کلی این دو را نشان می‌دهد:

$$\text{آزمون معنی‌داری} = \text{اندازه اثر} \times \text{اندازه مطالعه}^{(۵۱)}$$

در اینجا محاسبات مربوط به پایایی بازیافت و برخی اصلاحات در فرایند اندازه‌گیری فراتحلیل در زیر بیان می‌شود:

**پایایی بازیافت:**<sup>۴</sup> هنگامی که همه نتایج اطلاعات بر اساس طرح کدگذاری دقیق پیدا شدند لازم است که پایایی مرزهای نتایج تحقیق به دست آید. اگر هر محقق فراتحلیل فقط یک یا دو شاخص تحقیق را به کار برد، مقدار قابل ملاحظه‌ای از بازیافت مطالعات را از دست خواهد داد.<sup>(۵۲)</sup> لذا تمام تلاش در مرحله بازیافت این است که بیش از یک یا دو شاخص مطالعات را شامل شوند.

**تصحیح نتایج تحقیق:** همچون سایر تحقیقات در فرا - تحلیل نیز امکان اشتباه و خطا امری طبیعی است. این اشتباه از دو منبع می‌تواند باشد. یا از سوی فراتحلیل‌گر و یا از سوی نویسنده‌ای که مطالعه‌اش خلاصه می‌گردد. مطالعه دقیق مقالات اصلی از سوی فراتحلیل‌گر اغلب اوقات اشتباهات را آشکار می‌سازد. این اشتباهات را اغلب پیش از آنکه فرایندهای فرا - تحلیل به کار رود، می‌توان تصحیح کرد. یکی از اشتباهات مهم و کم‌تر قابل تصحیح و حتی برای تشخیص در اجرای فرا - تحلیل ثبت و ضبط داده، به‌عنوان داده‌ای که توسط محقق اصلی به‌دست آمده می‌باشد.<sup>(۵۳)</sup>

**تصحیح برای انحراف نمونه:** هنگامی که داده‌ها به آماره‌های مشترک تبدیل شدند، پایایی و دامنه اطلاعات انباشته می‌شوند. اگر همه مطالعات شامل برآوردهای پایایی یا دامنه داده‌ها باشند، سپس اندازه اثر در سطح هر مطالعه انفرادی می‌تواند تصحیح شود. اما اکثر مطالعات این اطلاعات را به‌دست نمی‌دهند. هانترو و اشمیت رویه‌هایی برای

1. magnitude of the relationship.

2. indication accuracy.

3. reliability.

4. reliability of retrieval.

برآورد اصلاحات برای پایایی و دامنه عزیمت<sup>۱</sup> با ساختن توزیع برای متغیرهای مستقل و وابسته فراهم نموده‌اند. هنگامی که پایایی و دامنه اطلاعات جمع‌آوری شدند، مرحله بعدی، برآورد اریب سو به پایین<sup>۲</sup> است که توسط خطای نمونه‌گیری سبب شده است. این اشاره به واریانس تصادفی که به خاطر اندازه نمونه است، دارد. میانگین تعدیل شده همبستگی<sup>۳</sup> نمونه به شرح زیر است:

$$\bar{r} = \frac{\sum [N_i r_i]}{\sum N_i}$$

در اینجا  $N_i$  تعداد آزمودنی‌ها در مطالعه است، و  $r_i$  اندازه اثر برای هر مطالعه منفرد می‌باشد. میانگین تعدیل شده همبستگی به وسیله فرمول واریانس زیر تعریف می‌گردد:

$$S_r^2 = \frac{\sum [N_i (r_i - \bar{r})^2]}{\sum N_i}$$

میانگین تعدیل شده  $d$  نمونه و واریانس مرتبط با آن به همین روش محاسبه می‌شوند.<sup>(۵۴)</sup>

**محاسبه و تصحیح واریانس خطا:** وقتی میانگین تعدیل شده همبستگی نمونه تحت تأثیر خطای نمونه‌گیری قرار گرفته، واریانس آن به شدت بالا می‌رود. یک رویه دومرحله‌ای برای تصحیح واریانس میانگین تعدیل شده همبستگی نمونه به طریق زیر به کار می‌رود. اولین مرحله محاسبه واریانس، خطای نمونه‌گیری است.

$$S_{er}^2 = \frac{K(1 - \bar{r}^2)^2}{\sum N_i}$$

در جایی که  $k$  تعداد مطالعات در تحلیل است. برای برآورد واریانس جمعیت اریبی شده،<sup>۵</sup> تصحیح نشده برای اندازه‌گیری، یا دامنه عزیمت، واریانس خطای نمونه‌گیری از  $S_r^2$  کم می‌شود.<sup>(۵۵)</sup>

$$S_{p_{xy}}^2 = S_r^2 - S_{er}^2$$

### مقایسه و ترکیب نتایج تحقیق

در بررسی و سنجش نتایج تحقیق به طور کلی دو راه اصلی را می‌توان به کار برد. یکی

1. departures.
2. downwards bias.
3. sample weighted mean corrolation.
4. subjects.
5. biased population variance.

براساس اطمینان آماری آنها (برای نمونه سطوح  $p$ ) و دیگری برحسب اندازه اثر آنها (برای مثال  $d$  یا تفاوت میانگین تقسیم بر انحراف معیار مشترک  $\sigma$  یا  $s$  و یا پیرسون  $r$ ).<sup>(۵۶)</sup> از سوی دیگر، دو فرایند تحلیلی اصلی برای مجموعه مطالعات وجود دارند که برای ارزشیابی به کار می‌روند: (۱) مقایسه و (۲) ترکیب. بنابراین با ادغام این چهار نوع می‌توان چنین نتیجه گرفت که در خصوص مقایسه می‌توان در جست‌وجوی دو موضوع بود، یکی بررسی اینکه آیا سطح معنی‌داری یک مطالعه از سطح معنی‌داری مطالعه دیگر متفاوت است؟ هم‌چنین می‌توان بررسی نمود که آیا اندازه اثر (برای مثال  $d$  یا  $r$ ) یک مطالعه از اندازه اثر مطالعه دیگر تفاوت معنی‌داری دارد؟ از سوی دیگر، نیز می‌توان آزمون‌های معنی‌داری مطالعات متفاوت را با یکدیگر و هم می‌توان برآوردهای اندازه اثرها را با یکدیگر ترکیب کرد.

با توجه به مطالب بالا می‌توان چنین گفت که در یک تحقیق فرا - تحلیل با داشتن دو مطالعه مستقل می‌توان رویه‌های زیر را به کار برد:

### مقایسه مطالعات

#### آزمون معنی‌داری

اگرچه در تحقیقات اغلب مقایسه اندازه اثرها از اولویت بیش‌تری برخوردارند، لیکن در برخی از موارد ما تنها می‌توانیم ارزش‌ها و یا مقادیر  $p$  را مقایسه نماییم و به اجبار به آن اکتفا کنیم. هنگامی که دو تحقیق مورد مطالعه قرار گیرند، سطح  $p$  یک دامنه جهت مقایسه مورد نیاز می‌باشد. بنابراین اگر ما به دست آوریم که  $t(30) = ۳/۰۳$  با استفاده از جدول توزیع نرمال ما می‌توانیم  $p$  مربوط به آن را پیدا نماییم که در مورد مثال فوق با یک دامنه برابر با  $۰/۰۰۲۵$  می‌باشد. سپس برای هر  $p$ ،  $Z$  آن را پیدا می‌کنیم که انحراف طبیعی معیار منطبق با مقدار  $p$  می‌باشد. از آنجا که هر یک از دو  $p$  می‌باید یک دامنه باشند،  $Z$ های مرتبط با آنها نیز همان علامت آنها را خواهند داشت. تفاوت بین دو  $Z$  وقتی بر  $\sqrt{2}$  تقسیم شود، یک  $Z$  جدید به دست می‌آید که منطبق با مقدار  $p$  است. در صورتی که دو  $Z$  واقعا متفاوت نباشند، تفاوت بین  $Z$ ها می‌تواند، بزرگ‌تر و خیلی بزرگ باشد. لذا به‌طور مختصر:

$$^{(۵۷)} \frac{Z_1 - Z_2}{\sqrt{2}} \text{ به عنوان } Z \text{ توزیع می‌شود.}$$

مثال: در مثالی فرضی مطالعات A و B نتایجی با جهت‌های مخالف ارائه می‌دهند و هیچ‌کدام معنی‌دار نیستند. یک  $p = 0.06$ ، یک دامنه است و دیگری  $0.12$ ، یک دامنه، اما در دامنه مخالف. Zهای منطبق با این  $p$ ها یافته شده در جدول منحنی نرمال  $1.56 +$  و  $1.18 -$  می‌باشند. با توجه به معادله بالا خواهیم داشت:

$$\frac{Z_1 - Z_2}{\sqrt{2}} = \frac{1.56 - (-1.18)}{1.42} = 1.94$$

مقدار  $p$  مرتبط با  $Z = 1.94$ ،  $0.026$  یک دامنه است یا  $0.052$  دو دامنه لذا دو مقادیر  $p$  ممکن است، به‌طور معنی‌داری متفاوت از یکدیگر به نظر برسند.<sup>(۵۸)</sup>

### برآورد اندازه اثر

زمانی ما از خود می‌پرسیم که آیا دو مطالعه مورد بررسی موضوع یکسانی را بیان می‌کنند یا خیر. به عبارت دیگر، آیا نتایج (برحسب اندازه اثر برآورد اثر) به‌طور معقولی با یکدیگر سازگاری دارند. یا آنها به‌طور معنی‌داری ناهمگن هستند. برای پاسخ به این سؤال ما می‌توانیم از شاخص اندازه اثر  $r$  و یا سایر شاخص‌ها همچون  $d$  استفاده نماییم. در اینجا از  $r$  استفاده می‌کنیم. برای هر دو مطالعه مورد مقایسه ما اندازه اثر  $r$  را محاسبه می‌کنیم و برای هر یک از این  $r$ ها  $Z_r$  فیشر را پیدا کنیم. جداول تبدیل  $r$  به دست آمده به  $Z_r$  فیشر معمولاً در اکثر کتب آمار وجود دارد. لذا وقتی  $N_1$  و  $N_2$  تعداد واحدهای نمونه‌گیری در هر یک از دو مطالعه ما را ارائه می‌دهند، کمیّت

$$\frac{Z_{r1} - Z_{r2}}{\sqrt{\frac{1}{N_1 - 3} + \frac{1}{N_2 - 3}}}$$

<sup>(۵۹)</sup> به عنوان  $Z$  توزیع می‌شود.

برای مثال دو نمونه فرضی A و B به ترتیب نتایجی در جهت مخالف با یکدیگر با اندازه اثرهای  $(N = 15)$ ،  $r = 0.60$  و  $(N = 100)$ ،  $r = -0.20$  را به دست می‌دهند.  $Z_r$  فیشرهای منطبق با این  $r$ ها به ترتیب برابر با  $0.69$  و  $-0.20$  می‌باشند. با توجه به معادله مذکور

$$\frac{Z_{r1} - Z_{r2}}{\sqrt{\frac{1}{N_1 - 3} + \frac{1}{N_2 - 3}}} = \frac{(0.69) - (-0.20)}{\sqrt{\frac{1}{12} + \frac{1}{97}}} = 2.91$$

به عنوان  $Z$  تفاوت بین دو اندازه اثر. ارزش  $p$  مرتبط با  $Z = 2.91$  مساوی با  $0.002$  یک دامنه یا  $0.004$  دو دامنه است. بنابراین، این دو اندازه اثرها به‌طور معنی‌داری با یکدیگر متفاوت هستند.<sup>(۶۰)</sup>

### ترکیب مطالعات

در دو مطالعه مستقل، ما می‌توانیم هم آزمون‌های معنی‌داری را با یکدیگر ترکیب نماییم و هم اندازه اثرها را. حال در ابتدا ما به بررسی ترکیب آزمون‌های معنی‌داری و در دنباله به بررسی ترکیب اندازه اثرها خواهیم پرداخت.

### ترکیب آزمون معنی‌داری

بعد از آنکه ما نتایج دو مطالعه مستقل را مقایسه نمودیم، ترکیب سطوح  $p$  دو مطالعه مورد نظر در عمل کار آسانی خواهد بود. بنابراین برآورد کلی احتمالی را به دست خواهیم آورد که دو سطح  $p$  ممکن است، به دست بیاورند (اگر که فرض صفر، یعنی احتمال نبود رابطه بین  $x$  و  $y$  درست باشد). برای این امر روش‌های زیر وجود دارد که در اینجا روش بسیار ساده اضافه نمودن  $Z$ ها که روش استافر نامیده می‌شود، به کار می‌بریم. این روش همچون روش مقایسه مقادیر  $p$  است. این روش در ابتدا از ما می‌خواهد تا سطوح دقیق  $p$  را برای هر دو مطالعه به دست آورده و سپس  $Z$  منطبق با هر یک را از جدول  $p$  ها پیدا نماییم. هر دو  $p$  می‌باید یک دامنه بوده و  $Z$ های منطبق با آنها نیز علایم یکسانی داشته باشند.<sup>(۱۱)</sup> جمع دو  $Z$  هنگامی که تقسیم بر  $\sqrt{2}$  شود، یک  $Z$  جدید ارایه می‌دهد. این  $Z$  جدید منطبق با مقدار  $p$  است که نتایج دو مطالعه ترکیب شده می‌توانست اتفاق بیفتند، اگر فرض صفر صحیح می‌بود. بنابراین

به عنوان  $Z$  توزیع شده

$$\frac{Z_1 + Z_2}{\sqrt{2}}$$

سپس ما می‌توانیم به هر  $Z$  به وسیله  $df$  آن، کمیت برآورد شده آن یا هر وزن مورد علاقه، وزن دهیم. رویه کلی برای وزن دادن به  $Z$ ها ضرب هر  $Z$  با وزن مورد نظر، جمع نمودن  $Z$ ها و تقسیم جمع  $Z$ های وزن داده شده بر جذر جمع وزن‌های مربع شده به شرح زیر است:

$$Z = \frac{W_1 Z_1 + W_2 Z_2}{\sqrt{2}}$$

وزنی

برای مثال، در دو مطالعه فرضی  $A$  و  $B$  نتایج در جهت‌های مخالف ارایه می‌شوند. و هر دو معنی‌دار هستند. یک  $p$ ،  $0/05$  یک دامنه، و دیگری  $0/000001$  یک دامنه اما در دامنه‌های مخالف.  $Z$ های منطبق با این  $p$ ها در جدول انحراف نرمال به ترتیب  $-1/64$

و ۵/۲۰ می‌باشند. بنابراین از معادله بالا خواهیم داشت:

$$\frac{Z_1 + Z_2}{\sqrt{2}} = \frac{(-1.64) + (5.20)}{1.41} = 2.52$$

به عنوان Z نتایج ترکیب شده مطالعات A و B. مقدار p مرتبط با z ۲/۵۲، ۰/۰۰۶ یک دامنه یا ۰/۰۱۲ دو دامنه می‌باشد. بنابراین، p ترکیب شده نتیجه مطمئن‌تری از دو نتیجه را حمایت می‌کند. حال، در این مثال، اگر ما از هیئتی از خبرگان فهمیدیم که مطالعه A وزنی (W1) برابر با ۳/۴ براساس ارزیابی پایایی درونی کسب می‌کند، درحالی که مطالعه B فقط وزنی W2 برابر با ۰/۹ کسب نموده است. Zها برای مطالعات A و B به ترتیب برابر ۱/۶۴- و ۵/۲۰ می‌باشد. بنابراین با استفاده از معادله پیشین خواهیم داشت:

$$\frac{(3.4)(-1.64) + (0.9)(5.20)}{\sqrt{(3.4)^2 + (0.9)^2}} = \frac{-0.896}{3.517} = 0.25$$

به عنوان Z نتایج ترکیب شده مطالعات A و B. مقدار p مرتبط با این z، ۰/۴۰ یک دامنه یا ۰/۸۰ دو دامنه است. توجه کنید که وزن دادن، به نتیجه نامعنی‌داری در این مثال می‌رسد. در مرحله اول که وزن به معادله نداده بودیم، مقدار p معنی‌دار بود. یعنی p=۰/۰۱۲ دو دنباله. اگر وزن دادن با درجه آزادی (df) انجام دهیم تا کیفیت تحقیق، و اگر (df) برای مطالعات A و B به ترتیب ۳۶ و ۱۴۴ باشد، z وزنی می‌شود:

$$\frac{(36)(-1.64) + (144)(5.20)}{\sqrt{(36)^2 + (144)^2}} = \frac{689.76}{148.43} = 4.65$$

این نتیجه نشان می‌دهد که Z ترکیب شده (p < ۰/۰۰۰۰۰۲) یک دامنه) به‌طور قوی به دلیل تفاوت اساسی در df بین دو مطالعه در جهت Z با df بزرگ‌تر حرکت می‌کند. توجه کنید که وقتی که به Zها به وسیله df وزن می‌دهیم، سعی می‌کنیم که به اندازه مطالعه نقش بزرگ‌تری در تعیین ترکیب شده بدهیم. این نقش بسیار بزرگ‌تر است، زیرا اندازه‌ای که مطالعه تاکنون دارد در هر Z تعیین شده است و بنابراین برای دومین بار در روند وزن‌گذاری وارد شده است.<sup>(۶۲)</sup>

## N جبرانی

از آنجا که همه مقایسه‌ها، برای آنکه توسط بررسی‌کننده ردیابی شوند، شانس مساوی



ندارند، لذا احتمال آنکه نتایج غیرمعنی دار توسط بررسی کننده یافت شود، کم تر از نتایج معنی دار می باشد. این واقعیت به معنی آن است که شیوه Zها ممکن است، سطحی از احتمال را ایجاد کند که نوعی برآورد پایین تر از سطح واقعی باشد (احتمال خطای نوع اول). یکی از مزایای روش Zها آن است که محاسبه N جبرانی را ممکن می کند. سؤالی که N جبرانی به آن پاسخ می گوید این است: «چه تعداد مقایسه منتهی به تأیید فرضیه صفر (یعنی  $Z_{st} = 0$ ) را باید به نتایج مقایسه های ردیابی شده اضافه کرد تا این نتیجه گیری را که رابطه ای وجود دارد، تغییر بدهد؟

یکی از مزایای روش جمع Zها آن است که محاسبه N جبرانی را ممکن می کند. وقتی که سطح معنی داری انتخاب شده  $P < 0.05$  باشد، فرمول محاسبه این مقدار به قرار زیر است:

$$N_{FS0.05} = \left( \frac{\sum_{i=1}^N Z_i}{1.645} \right)^2 - N$$

در فرمول بالا:  $N_{FS} = 0.05$  تعداد مقایسه های اضافی صفر که برای بالا بردن احتمال ترکیب شده فقط تا حد  $P < 0.05$  لازم است؛

$1/645 =$  انحراف نرمال استاندارد همراه با  $P < 0.05$  (یکسویه)؛ سایر کمیت ها نیز قبلاً توضیح داده شده است.

N جبرانی بر این فرض استوار است که مجموعه پژوهش های ردیابی نشده مساوی با نتیجه صفر مطلق است. این فرض باعث محدودیت اعتبار آن می شود. حال این سؤال مطرح می شود که N جبرانی چقدر باید بزرگ باشد که بتوان نتیجه گرفت که یافته های مقابل نتایج صفر ردیابی نشده مقاومت می کند؟ روزنتال می گوید که، ۵ برابر تعداد پژوهش های ردیابی شده به علاوه ۱۰ است. البته هیچ قاعده مشخصی در این باب وجود ندارد.<sup>(۶۳)</sup>

### ترکیب برآورد اندازه اثر

هنگامی که ما می خواهیم نتایج دو مطالعه را ترکیب کنیم، حداقل به همان میزان نیز به برآورد اندازه اثر ترکیب شده علاقمندیم. درست همان طور که در مورد مقایسه دو برآورد اندازه اثر انجام دادیم؛ ما می باید به  $r$  به عنوان اندازه اثر اولیه مان در ترکیب

اندازه اثرها توجه کنیم. بدیهی است که سایر شاخص‌های برآورد اندازه اثر همچون  $d$  نیز در دسترس هستند.

برای هر دو مطالعه جهت ترکیب،  $r$  را محاسبه می‌کنیم و  $Z_r$  فیشر مرتبط با آن را خواهیم داشت:

$$\frac{Z_{r1} + Z_{r2}}{2} = \bar{Z}_r$$

به‌عنوان  $Z_r$  فیشر مرتبط با میانگین  $r$  ما. سپس  $r$  به  $Z_r$  یا  $r$  جدول با جست‌وجوی  $r$  مرتبط با میانگین  $\bar{Z}_r$  مان به کار می‌بریم.

### رویه‌های فرا – تحلیلی برای هر تعداد مطالعات مستقل

در این بخش تلاش می‌شود که روش‌های مقایسه و ترکیب نتایج را برای هر تعداد مطالعه مستقل تعمیم دهیم. در این مورد برای مقایسه مطالعات از دو نوع آزمون یکی پراکندگی و دیگر متمرکز می‌توان سود جست. در مطالعه آزمون‌های پراکنده ما می‌آموزیم که آیا دو مطالعه  $A$  و  $B$  میان خودشان با توجه به سطوح معنی‌داری با اندازه اثرشان متفاوت هستند یا خیر، اما نمی‌آموزیم که چگونه آنها متفاوتند. در آزمون‌های متمرکز وقتی ما مطالعات را مقایسه می‌کنیم به جهت سطوح معنی‌داری آنها و یا اندازه‌های تأثیرشان، ما یاد می‌گیریم که آیا مطالعات به‌طور معنی‌داری میان خودشان به‌نحوی نظری قابل پیش‌بینی یا معنی‌داری متفاوتند یا خیر. لذا آزمون‌های مهم فرضیات می‌توانند به‌وسیله استفاده از آزمون‌های متمرکز ساخته شوند. حال با این توضیح مختصر به مقایسه مطالعات تحت آزمون پراکندگی و سپس آزمون متمرکز خواهیم پرداخت:

### مقایسه مطالعات آزمون پراکندگی

آزمون معنی‌داری با سه سطح  $p$  یا بیش‌تر برای مقایسه، ما اول انحراف طبیعی،  $Z$ ، منطبق با هر  $p$  را پیدا می‌کنیم. همه سطوح  $p$  می‌باید یک دامنه باشند و  $Z$ های منطبق نیز می‌باید همان علامت‌ها را داشته باشند. معنی‌داری آماری ناهمگن  $Z$  می‌تواند از  $\chi^2$  محاسبه شده به شرح زیر به دست می‌آید:

$$\sum (Z_j - \bar{Z})^2$$

توزیع شده است به‌عنوان  $\chi^2$  با  $K-1$ df

در این معادله،  $Z_j$ ،  $Z$  برای هر یک مطالعه است،  $\bar{Z}$  میانگین همه  $Z$ های به دست آمده و  $k$  تعداد مطالعات مورد محاسبه می باشد.

### مقایسه برآورد اثر

در اینجا ما می خواهیم ناهمگنی آماری سه اندازه اثر یا بیش تر را ارزیابی نماییم. در اینجا  $r$  را به عنوان شاخص برآورد کننده یا برآورد کننده اندازه اثر به کار می بریم. برای هر سه یا بیش تر مطالعه مورد مقایسه ما،  $r$  اندازه اثرها را محاسبه می کنیم،  $Z_r$  فیشر منطبق با آنها و  $N-3$  (برای مثال سندیکور و کوکران، ۱۹۶۷، ۱۹۸۰)، را محاسبه می کنیم.  $N$  تعداد واحدهای نمونه گیری است که بر روی آنها هر  $r$  بنا شده است. سپس معنی داری آماری ناهمگنی  $r$ ها می تواند از یک  $\chi^2$  به دست آید زیرا  $\sum (N_j-3)(z_{rj}-\bar{z})^2$  به عنوان  $\chi^2$  توزیع می شود با  $K-1$ df. در این معادله  $Z_r$ ،  $Z_{rj}$  فیشر منطبق با هر  $r$  و  $\bar{Z}_r$  میانگین وزنی  $Z_r$  است، یعنی

$$\bar{Z}_r = \frac{\sum (N_j-3) Z_{rj}}{\sum (N_j-3)}$$

### مقایسه مطالعه آزمون متمرکز

همچون آزمون های پراکندگی، در اینجا ما با یافتن انحراف طبیعی استاندارد  $Z$ ، منطبق با هر سطح  $p$  شروع می کنیم. همه سطوح  $p$  می باید یک دامنه باشند و  $Z$ های منطبق می باید علامت یکسانی داشته باشند. معنی داری آماری آزمون تباین<sup>۱</sup> هر فرضیه ویژه ای درباره مجموعه سطوح  $p$  می تواند از  $Z$  محاسبه شده به شرح زیر به دست می آید:

$$Z = \frac{\sum \lambda_j z_j}{\sqrt{\sum \lambda_j^2}}$$

در این معادله  $\lambda_j$  به لحاظ نظری پیش بینی مشتق شده یا وزن تباین برای هریک از مطالعات است. چینی فرض می شود که جمع  $\lambda_j$ ها صفر خواهد شد. و  $Z_j$ ،  $Z$  برای هریک مطالعه خواهد بود.

### ترکیب مطالعات برای هر اندازه مطالعات

ترکیب مطالعات در دو بخش آزمون معنی داری و اندازه اثر مورد مطالعه قرار می گیرد.

1. contrast testing.

در اینجا همچون گذشته، ابتدا ترکیب آزمون معنی داری را بررسی می‌کنیم و سپس به ترکیب اندازه اثر خواهیم پرداخت.

**آزمون معنی داری:** پس از آنکه هر مجموعه‌ای از سه یا چند مطالعه را به لحاظ نتایج مقایسه نمودیم، ترکیب سطوح  $p$  مجموعه مطالعات لازم است. در روشی که در اینجا جهت ترکیب آزمون معنی داری استفاده می‌کنیم، فقط نیاز به محاسبه  $Z$  برای هر یک از سطوح  $p$  داریم. همه  $p$ ها می‌باید یک دامنه باشند.  $Z$ های مخالف علامت منفی را کسب می‌نمایند. سپس، جمع  $Z$ ها را تقسیم بر مجذور تعداد ( $K$ ) مطالعات می‌کنیم و آمارهای جدید توزیع شده‌ای تحت عنوان  $Z$  به ما می‌دهد. به‌طور خلاصه:

$$Z = \frac{\sum Z_j}{\sqrt{K}}$$

توزیع شده به عنوان  $Z$

در اینجا ممکن است که ما به هر یک از  $Z$ ها به‌وسیله درجه آزادی‌شان یا کیفیت برآورد شده آنها و یا هر وزن دیگری وزن بدهیم. بدین لحاظ فرمول ما به شرح زیر می‌باشد.

$$Z = \frac{\sum w_j Z_j}{\sum w_j^2} \quad (64)$$

وزنی  $Z$

#### ترکیب اندازه اثرها

هنگامی که ما نتایج سه یا چند مطالعه را ترکیب نمودیم به ترکیب برآورد اندازه اثر علاقمند خواهیم بود. در اینجا  $r$  را به‌عنوان برآوردکننده اندازه اثر اولیه در نظر خواهیم گرفت. برای هر سه یا چند مطالعه که می‌باید ترکیب شوند، ما  $r$  و  $Zr$  فیشر مرتبط با آن را محاسبه می‌کنیم و خواهیم داشت:

$$\frac{\sum Zr}{K} = \bar{Z}_r$$

در اینجا  $K$  به تعداد مطالعات ترکیب‌شده اشاره دارد. هم‌چنین  $\bar{Z}_r$  فیشر نیز منطبق با میانگین  $r$  ما می‌باشند. برای یافتن  $r$  مرتبط با میانگین  $Z_r$  موردنظرمان از جدول  $Z_r$  فیشر استفاده می‌کنیم. می‌توانیم به هر  $Z_r$ ،  $df$  آن را یعنی  $N-3$  و با کیفیت تحقیق برآورد شده و یا هر وزن تعیین‌شده‌ای را برای وزن داده تخصیص دهیم. میانگین وزنی  $Z_r$  به شرح زیر به‌دست می‌آید.

$$\bar{Z}_r = \frac{\sum w_j Z_{rj}}{\sum w_j}$$

## تجزیه و تحلیل همگونی مطالعات

به طور کلی همگونی یا ناهمگونی درجه تغییرپذیری<sup>۱</sup> بین مطالعات را در گروهی از مطالعات توصیف می‌کنند. احتمالاً ترکیب نتایج حاصل از مجموعه مطالعات همگن مناسب است، اما خیلی از روش‌شناسان معتقدند که نتایج حاصل از مطالعات ناهمگن نباید ترکیب شوند. از سوی دیگر، شاخص‌های همگونی این سؤال را مطرح می‌کنند که آیا واریانس اندازه اثر، به طور معنی‌داری با آنچه از خطای نمونه‌گیری انتظار می‌رود، فرق می‌کند یا نه؟ اگر پاسخ به این منفی باشد، برخی از روش‌شناسان معتقدند که بررسی‌کننده باید تجزیه و تحلیل را در همین نقطه متوقف کند. اما اگر پاسخ به سؤال مذکور مثبت باشد، یعنی اندازه اثر واریانس بزرگ‌تر از آنچه که به طور تصادفی انتظار می‌رود نشان بدهند، بررسی‌کننده منابع دیگر واریانس را هم که بالقوه وجود دارند مورد بررسی قرار می‌دهد.<sup>(۶۵)</sup> برای آزمون همگنی استاندارد از فرمول زیر استفاده می‌شود:

$$Q = \sum_{i=1}^N \frac{(d_i - d)^2}{\sigma(d_i)}$$

تحت فرض صفر  $\sigma = \dots = \sigma_{n_i}$  و  $Q \sim \chi_{n-1}^2$

و یا با استفاده از ضریب همبستگی  $r$ ، آزمون همگونی ضریب همبستگی به شرح زیر می‌باشد:

$$Q = \sum_{i=1}^N (n_i - 3)(z_i - z)^2$$

تحت فرض صفر  $P_i = \dots = P_{n_i}$  و  $Q \sim \chi_{n-1}^2$

همان‌طور که ملاحظه می‌شود، شاخص  $Q$  دارای توزیع مجذور  $k$  با درجه آزادی  $N-1$ ، یعنی کمتر از تعداد مقایسه‌هاست. برای تفسیر توزیع مجذور  $k$  در جایی که درجه آزادی  $N-1 = N$  (تعداد مطالعات است)، می‌باشد، اگر  $P < 0.05$  باشد، ناهمگونی معنی‌داری وجود دارد. راه دیگر برای این کار مقایسه ساده  $Q$  با  $N-1$  است، یعنی اگر  $Q > N-1$  باشد، ناهمگونی معنی‌داری وجود دارد.<sup>(۶۶)</sup> برای ارزیابی همگنی به شاخص تفاوت نیز نیازمندیم. برای اینکه مشخص شود که مجموعه شاخص‌های تفاوت، همگن هستند یا ناهمگن، بررسی‌کننده باید  $Q_t$  را که توسط هجز و اوکلین در ۱۹۸۵ ارائه

1. variability.

گردیده، استفاده نمایند. فرمول محاسبه شاخص  $Q_t$  به شرح زیر می‌باشد:

$$Q_t = \frac{\sum_{i=1}^N w_i d_i^2 - \left( \sum_{i=1}^N w_i d_i \right)^2}{\sum_{i=1}^N w_i}$$

تفسیر  $Q_t$  نیز مانند  $Q$  است. اگر مقدار حاصله از مقدار دامنه بالای مجذور  $k$  در سطح معنی‌داری خاصی بیش‌تر باشد، بررسی‌کننده این فرضیه را که واریانس اندازه اثرها، تنها معدل خطای نمونه‌گیری بوده است، رد می‌کند. در جدول شماره ۷ مقدار  $Q_t$  برابر با ۴/۵ است مقدار شاخص برای مجذور  $k$  در  $P < 0/05$  بر مبنای ۶ درجه آزادی، معادل ۱۲/۵۶ است. از اینرو فرضیه‌ای که مدعی است، خطای نمونه‌گیری باعث تفاوت شاخص‌های تفاوت می‌شود، نمی‌توان رد کرد. به عبارت دیگر، مجموعه مطالعات همگن هستند.<sup>(۶۷)</sup>

به نظر کوپر، روش‌هایی که به آزمایش این موضوع می‌پردازند که آیا تفاوت‌های روش‌شناختی یا مفهومی میان مقایسه‌ها، توجیه‌کننده واریانس در اندازه‌های اثر هستند یا خیر، شامل سه مرحله‌اند. در ابتدا برای هر یک از زیر گروه‌های موجود در پژوهش‌ها، شاخص  $Q$  به‌طور جداگانه محاسبه می‌شود. به عنوان مثال برای مقایسه چهار شاخص تفاوت اول در جدول ۷، با سه شاخص آخر برای هر یک از گروه‌بندی‌ها، شاخص  $Q$  به‌طور جداگانه محاسبه می‌شود. سپس مقادیر مربوط به این شاخص‌های  $Q$  با هم جمع می‌شوند تا شاخصی موسوم به  $Q_w$  به دست آید. این مقدار، سپس از  $Q_t$  کسر می‌شود تا  $Q_b$  به دست آید. فرمول:

$$Q_b = Q_t - Q_w$$

شاخص  $Q_b$  برای آزمایش این موضوع مورد استفاده قرار می‌گیرد که آیا میانگین تأثیرات گروه‌بندی‌ها همگن هستند یا نه. سپس با استفاده از جدول مجذور  $k$  ارجاع همچون موارد بالا تفسیر می‌شود. اگر میانگین شاخص‌های تفاوت همگن باشند، عامل گروه‌بندی، واریانس تأثیرات را به فراتر از آنچه که با خطای نمونه‌گیری همراه است توجیه نمی‌کند. اگر  $Q_b$  از مقدار شاخص بیش‌تر باشد، قابل گروه‌بندی در واریانس اندازه‌های تأثیر نقش مهمی خواهد داشت در جدول ۷، در مقایسه شاخص‌های تفاوت چهارگانه اولیه و سه شاخص آخر، معادل ۰/۹۸ است. این نتیجه در درجه آزادی معنی‌دار تلقی نمی‌شود.

شاخص ضریب همبستگی (r) برای تجزیه و تحلیل همگنی در مورد شاخص ضریب همبستگی هم روش مشابهی وجود دارد که در آن از فرمول زیر استفاده می‌شود

$$Q_t = \sum_{i=1}^N (n_i - 3) z_i^2 - \left[ \frac{\sum_{i=1}^N (n_i - 3) z_i}{\sum_{i=1}^N (n_i - 3)} \right]^2$$

نتایج تحلیل همگنی با استفاده از شاخص ضریب همبستگی در جدول ۸ ارائه شده است.

### ترکیب احتمالات

برای ترکیب احتمالات از روشهای زیر می‌توان استفاده نمود.

### روش‌های اصلی

روش جمع کردن احتمالات (قابلیت کاربرد وقتی که  $\sum P$  نزدیک به واحد یا کم‌تر باشد):

$$p = \frac{(p)N}{N!}$$

روش جمع کردن tها

$$z = \frac{t}{\sqrt{df / (df - 2)}}$$

روش اضافه کردن zها:

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

روش جمع کردن zهای وزنی

$$Z = \frac{T}{\sigma T} = \frac{df_1 Z_1 + df_2 Z_2 + \dots + df_n Z_n}{\sqrt{df_1^2 + df_2^2 + \dots + df_n^2}}$$

روش آزمون میانگین p

$$Z = (0.50 - \bar{p})(\sqrt{12N})$$

روش آزمون میانگین z

$$t = \frac{z/N}{\sqrt{(S_z^2)/N}}$$

### مرحله پنجم: ارائه نتایج

در مرحله نهایی گزارشی از بررسی انجام شده ارائه گردیده و به وسیله آن بررسی به پایان می‌رسد. انجمن روان‌شناسی آمریکا (۱۹۸۳) پیشنهاد نمود که در بررسی‌ها،

چارچوبی شبیه به آنچه که در گزارش‌های پژوهشی معمول است، اختیار شود؛ چارچوبی که شامل بخش‌های مقدمه، روش‌ها، نتایج و مباحث باشد. به نظر کوپر مقدمه باید عرصه مفهومی مسأله پژوهشی و بین اهمیت آن مسأله باشد. برخلاف مقدمه گزارش‌های پژوهشی، مقدمه گزارش بررسی‌های پژوهشی می‌باید جامع‌تر باشد. بررسی‌کننده باید سعی نماید تا در مورد مسائل نظری و روش‌شناختی مربوط به پرسش پژوهشی، مروری تاریخی و کامل نماید. مقدمه می‌باید مسأله مورد مطالعه را به دقت بافت‌مند<sup>۱</sup> نماید. در بخش روشها فراتحلیل‌گر می‌باید به نحوی بررسی متون را ارایه دهد. موضوع بعدی که می‌باید به آن پرداخت معیارهای ارتباط موضوعی است؛ معیارهایی که در متون ردیابی شده می‌باید مورد توجه قرار گیرد. در مرحله سوم می‌باید روش‌شناسی‌های نوعی<sup>۲</sup> را تشریح کرد. در چهارمین مرحله فرا - تحلیل‌گر چگونگی شناسایی آزمون‌های مستقل فرضیه را مشخص نماید. در مرحله پنجم باید ویژگی‌های مطالعات مربوط به پژوهش‌های اولیه را توصیف کرد. و در آخرین مرحله می‌باید به روش‌های مورد استفاده در فرایند تجزیه و تحلیل کمی فرا - تحلیل پرداخت. در بخش نتایج فرا - تحلیل‌گر می‌باید در مورد متون پژوهشی و مبانی آماری یافته‌های ادغامی توصیفی اجمالی ارایه نماید.<sup>(۲۸)</sup> در پایان می‌باید یک بخش را به بحث و نتیجه‌گیری اختصاص داد. این بخش همان کارکردهایی را که در پژوهش‌های اولیه انجام می‌دهند، انجام خواهد داد.

### بحث و نتیجه‌گیری

در پایان به مسایل و مشکلاتی که در انجام روش فرا - تحلیل فراروی پژوهشگران می‌باشد، پرداخته و جهت انجام بهتر این روش پیشنهاداتی ارایه می‌گردد:

در مراحل انجام فرا - تحلیل برخی از موانع وجود دارد که پاره‌ای به تمامی روش‌های تحقیق است و پاره دیگر نیز مختص این روش می‌باشد. موانع ویژه برای فراتحلیل شامل موارد زیر می‌گردد:

۱- به ندرت نتایج مطالعات متفاوت به طور دقیق مورد موافقت قرار می‌گیرند.

1. contextualize

2. prototypical methodologies



۲- اگر نویسندگان علاقمند به حمایت نتیجه‌گیری ویژه‌ای باشند، آنها می‌توانند مطالعاتی که حمایت از آن نتیجه‌گیری می‌کند را دربرگیرد و مطالعاتی که حمایت نمی‌کنند، حذف نمایند. آیا نویسندگان در مقاله‌شان به‌طور دقیق توضیح می‌دهند که مطالعه آن شامل چه اساسی می‌شود؟ و آیا دلایلیشان معنی‌دار است؟

۳- تمایلی وجود دارد که بیش‌تر مطالعاتی که نوعی اثر مثبت را نشان می‌دهند، چاپ شوند تا آنهایی که این را نشان نمی‌دهند. لذا این اثر به‌عنوان اریبی چاپ‌شده<sup>۱</sup> شناخته می‌شود. آبرامی و دیگران برای بهبود در مراحل انجام فرا - تحلیل موارد زیر را پیشنهاد می‌کنند:

۱- برای بهبود شاخص شمول ویژه که با ملاک انتخاب مدرک مورد مطالعه سروکار دارد، معتقدند که برخلاف پیشنهاد گلاس در خصوص معیار شمول وسیع - که در آن می‌توان بسیاری از مطالعات را حتی با وجود پایین بودن سطح کیفیت روش‌شناختی آنها وارد نمود - و برخلاف معیار بهترین مدرک که توسط سالیوان ارایه گردیده است - که در آن مطالعات با سطح کیفیت پایین خارج می‌شود و تنها مطالعاتی که هم به لحاظ اعتبار درونی و هم به لحاظ اعتبار بیرونی در سطح بالایی هستند، مورد بررسی قرار گیرند - می‌باید موقعیتی میانه را اخذ نمود. مرورکنندگان می‌باید از اتکا شدید به آزمون‌های آماری برای ادغام مطالعاتی که کیفیت‌های گوناگونی دارند، خودداری کنند. هم‌چنین برای ورود اطلاعات می‌باید به آنها وزن داده شود.

۲- برای بهبود کدگذاری خصیصه‌های تحقیق، پیشنهاد می‌شود که مرورکنندگان اول از همه به خلاصه نظام‌یافته خصیصه‌های تحقیق شده، محاسبه شده و هم‌چنین ذکر شده درون مطالعات متعهد شوند.

۳- برای بهبود تحلیل داده‌ها، پیشنهاد می‌شود که مرورکنندگان در آزمون متوسط نتایج و در آزمون‌های خصیصه‌ها توجه بیش‌تری به پرهیز از خطای نوع دوم مطالعه نشان بدهند.<sup>(۶۹)</sup>

## پی‌نوشت‌ها

1. Tsao, AndyC. (1998). What is the Meta - Analysis, <Http://Server.am.ndhu.edu.tw/~chtsao/ma.html>.
2. Abrami, Philip C. , Cohen, Peter A., Apollonia, Sylvia d' ( 1988). Implementation Problems in Meta-Analysis, Review of Educational Research, Vol. 58, No. 2, pp. 151-179.
3. Philips, Jodeph M. (1994). Farmer Education and Farmer Efficiency: a meta-analysis, Economic Development and Cultural Change,
4. Barroman, Nicholas (1998). A Survey of Meta - Analysis, [http:// www.mscs.dal.ca/~barrowma/ma.html](http://www.mscs.dal.ca/~barrowma/ma.html).
5. Glass, G.V(1977). Integrating findings: The meta - analysis of research. Review of Research in education, 5, 351-379, in Lyons, Larry C. & Manassas Va, (1998). Meta-Analysis: Methods of Accumulating Results across Research Domains, <http://www.nsinc.com /solomon/metaA/Mapage 3.html>.
6. Lyons, Larry c, Manassa Va. & Virginia, Portions (1998). Meta-Analysis: Methods of Accumulating Results across Research Domains, <http://www.mnsinc.com/Solomon/Meta A/Ma page 3.html>.
7. Straight forward.
8. Same Global Population Value.
9. Balan, Martin (2000). An Introduction to Medical Statistics, <http://www.Sghms.ac.uk/phs/staff/jmb/meta.htm>.
10. Chiu, Irene & Brennau (1990). The Effectiveness of some Techniques for Improving Mail Survey Response Rates: A meta-analysis, Marketing Bulletin, Vol 1, PP.13-18.

11. Tsao, 1998.
12. Jackson, G.B (1980). Methods for integrative reviews. Review of Educational Research, Vol. 50, pp. 438-460.
13. Franklin, Cynthia, Darlene Grant, Jacquelin Corcoran, Pamela O'Miller, Linda Bultman, (1997). Effectiveness of Prevention Programs for Adolescent Pregnancy: A Meta - Analysis, Journal of Marriage and the Family, Vol. 59, pp.551-567.
14. Abrami, 1988.
15. Ibid, 1988.
16. Franklin, et al, 1997.
17. Ibid, 1997.
18. Abrami, 1988.
۱۹. کوپر، هریس ام. (۱۳۷۹)، پژوهش ترکیبی: راهنمای بررسی متون برای پژوهش، ترجمه محمد علی حمید رفیعی، تهران: دفتر پژوهشهای فرهنگی، ص ۱۴۲-۱۴۳.
20. Rosenthal, Robert (1984). Meta - Analytic Procedures for Social Research, London: Sage Publication., p.11.
21. Ibid, 1984, pp.11-12.
22. Balan, 2000.
23. Lyons, et al, 1998.
24. Kulik, James A. How can chemists Use Educational - Technology Effectively, Journal of Chemical Education, 60(November 1983): 957-959.
۲۵. کوپر، (۱۳۷۹). ص ۴۰-۴۳.
26. Rosenthal, 1984, pp.47-48.
27. Rosenthal, 1984, p.48.
۲۸. کوپر، هریس ام. (۱۳۷۹). ص ۱۶-۱۷.
29. Thompson, S.G. (1993) Controversies in Meta-analysis: The case of

- Trials of Serum Cholesterol Reduction, Statistical Methods in Medical research, 173-93. in Bland, Martin, (2000), Meta-Analysis: data from several studies, <http://www.sghms.ac.uk/phs/staff/jmb/meta.htm>.
30. Pocock, S.J. and Hughes, M.D.(1990) Estimation Issues in Clinical trial and overviews. Statistical in medicine, 657-71 in Bland, Martin, (2000), Meta-Analysis: data from several studies, <http://www.sghms.ac.uk/phs/staff/jmb/meta.htm>.
31. Bland, 2000.
32. Deviation Adjustment
33. Lyons et al, 1998, p.3.
34. Rosenthal, 1984, p.41.
35. Ibid, 1984, p.42.
36. Ibid, 1984, p.42.
۳۷. کوپر، هریس ام. (۱۳۷۹)، ص ۱۷.
۳۸. همانجا، ۱۳۷۹، ص ۸۹.
۳۹. همانجا، ۱۳۷۹، ص ۹۰.
۴۰. همانجا، ۱۳۷۹، ص ۹۱.
41. Glass, G. (1976). Primary, Secondary, and Meta - Analysis of Research, Educational Research, Vol. 5, p.3-8.
۴۲. کوپر، هریس ام. (۱۳۷۹)، ص ۹۶.
۴۳. همانجا، ۱۳۷۹، ص ۹۶.
۴۴. همانجا، ۱۳۷۹، ص ۹۶.
۴۵. همانجا، ۱۳۷۹، ص ۱۷.
46. Lyons et al, 1998, p.4.
47. Hall, J.A, L. Tickle-Degnen, R. Rosenthal, and F. rosteller. 1994, Hypotheses and problems in research synthesis. In Cooper, H.and L.V. Hedges, The Handbook of Research synthesis. NewYork: Russell sage foundation.

48. Hedges, L.V., and I. Olkin. 1985. *Statistical Methods for Meta-Analysis*. San Diego: Academic press. In Barrowman, Nicholas, *A Survey of Meta-Analysis*, 1998.
49. Scale - Free Indices.
50. Lyons et al, 1998, p.4
51. Rosenthal, 1984, p.19-20.
52. Glass, G.V, McGaw, B., & Smith, M.L.(1981). *Meta-analysis in social research*. Beverly Hills, CA: Sage, p.63-65. in Rosenthal, Robert (1984). *Meta - Analytic Procedures for Social Research*, London: Sage Publication.
53. Rosenthal, 1984, p.50.
54. Lyons et al, 1998.
55. Ibid, 1988.
56. Rosenthal, 1984, p.63.
57. Ibid, 1984, p.66.
58. Ibid, 1984, p.66.
59. Ibid, 1984, p.67.
60. Ibid, 1984, p.67.
61. Ibid, 1984, p.74.
62. Ibid, 1984, pp.73-74.
63. کوپر، هریس ام. (۱۳۷۹)، ص ۱۳۷-۱۳۹.
64. Ibid, 1984, pp. 88-89.
65. کوپر، هریس ام. (۱۳۷۹)، ص ۱۶۰.
66. Fless, JL, Gross. AJ (1991). *Meta - Analysis in Epidemiology, with Special Referrnce to Studies of the Association between Exposure to Environment Tabacco Smoke and Lung Cancer: A critice*. J. Clin Epidemio, Vol. 44, pp.127-139.

۶۷. کوپر، هریس ام. (۱۳۷۹)، ص ۱۶۰-۱۶۱.

۶۸. همانجا، ۱۳۷۹، ص ۱۷۱-۱۷۹.

69. Abrami, 1988.