Chapitre 2

La recherche des variables modératrices à l'aide de la méta-analyse*

Herman Aguinis & Charles A. Pierce***

Sommaire

1. L'approche méta-analytique de Hedges et Olkin (HO) 60
2. Mettre en œuvre des corrections au niveau des études dans la méta-analyse 62
3. Introduire une correction intra-étude dans la procédure de Hedges et Olkin : une procédure en trois étapes 65
4. Une illustration de l'utilisation de la procédure en trois étapes 70
5. Synthèse et conclusion 73

* Ce chapitre est la traduction d'un article publié par les auteurs dans le *Journal of Management* en 1998 dont le titre original est *« Testing Moderator Variable Hypotheses Meta-Analytically »*.

** Kelley School of Business, Indiana University.

*** Fogelman College of Business and Economics, University of Memphis.

Les revues quantitatives de la littérature (i.e. méta-analyses, MA) sont aujourd'hui adoptées de manière consensuelle dans de nombreuses sous-disciplines du management ainsi que dans les autres sciences sociales (Aguinis, Pierce & Quigley, 1993 ; Cooper & Hedges, 1994 ; Cotton & Turttle, 1986 ; Dobbins & Platz, 1986 ; Johnson, 1989). Un des principaux avantages de la méta-analyse, comparée aux revues narratives traditionnelles, est qu'elle permet de tester des hypothèses concernant les effets *de variables modératrices*. Une variable *Z* est considérée comme modératrice de la relation entre les variables *X* et *Y* si l'intensité de cette relation dépend des valeurs ou des niveaux de la variable *Z* (Aguinis, 1995 ; Aguinis, Bommer & Pierce, 1996 ; Aguinis & Pierce, 1998 ; Zedeck, 1971). À l'instar des autres chercheurs (e.g. Aguinis et al., 1994 ; Aguinis & Stone-Romero, 1997 ; Stone-Romero et al., 1994), les méta-analystes testent souvent des hypothèses formulées sur des effets modérateurs (Cooper & Lemke, 1991 ; Mullen et al., 1991). Les méta-analystes en comportement organisationnel se sont, par exemple, intéressés à l'étude de l'effet modérateur de la pression sur la production (*Z*) sur la relation entre la satisfaction au travail (*X*) et la performance au travail (*Y*) (Iaffaldano & Muchinsky, 1985 ; Petty, McGee & Cavender, 1984).

Ce chapitre a pour objectif de présenter les problèmes méthodologiques soulevés par la recherche méta-analytique des effets modérateurs. Plus précisément, nous proposons et décrivons une procédure en trois étapes basée sur l'approche de Hedges et Olkin (1985) qui introduit des corrections d'artefacts méthodologiques et statistiques souvent préconisées et utilisées dans les approches méta-analytiques en psychométrie (e.g. Hunter & Schmidt, 1990).

Dans les sections qui suivent, nous présentons brièvement (1) l'approche de Hedges et Olkin et les avantages liés à l'utilisation du test d'homogénéité *Q*, (2) exposons l'intérêt et les apports d'une correction des artefacts méthodologique et statistique dans le cadre d'une méta-analyse, (3) proposons une procédure en trois étapes afin d'introduire une correction des études dans l'approche de Hedges et Olkin et enfin (4) présentons une illustration de cette procédure.

1. L'approche méta-analytique de Hedges et Olkin (HO)

Hedges et Olkin (Hedges, 1982a, 1982b ; Hedges & Olkin, 1985) proposent une approche méta-analytique qui vise à examiner : (a) l'intensité de la relation entre deux variables, (b) la fluctuation (ou variabilité) des résultats entre les études et (c) les variables modératrices susceptibles d'expliquer cette variabilité des résultats. L'intensité de la relation entre les deux variables est appréciée sur la base d'un estimateur non biaisé de la taille de l'effet (*d*) issu de chaque étude, à partir duquel est calculée une moyenne des *d*.

Le degré de variabilité des estimateurs de l'effet *ds* entre les études est évalué à l'aide de la statistique d'homogénéité *Q*. Un test *Q* statistiquement significatif indique que les grandeurs d'effet *ds* issues de chaque étude ne partagent pas une grandeur d'effet commune et, par conséquent, qu'il est judicieux de rechercher les modérateurs de la relation. Il est en effet recommandé de

poursuivre la recherche des effets modérateurs lorsque la statistique Q est statistiquement significative (Hedges & Olkin, 1985 : Johnson & Turco, 1992). Cette démarche se justifie s'il existe des arguments théoriques en faveur de certains effets supposés modérateurs.

Deux types de procédures statistiques sont employées pour tester les effets modérateurs selon la nature de la variable modératrice, en distinguant notamment les variables *catégorielles* (par exemple, type d'emploi : ouvrier/cadres ; Iaffaldano & Muchinsky, 1985) et les variables *continues* (ex. âge ; Oliver & Hyde, 1993). Dans le cas des variables modératrices *catégorielles* (Hedges, 1982a ; Hedges et Olkin, 1985), chaque étude se voit assigner une valeur numérique correspondant aux modalités de la variable modératrice (e.g. genre, 1 = féminin, 2 = masculin), et se voit affecter dans un sous-groupe en fonction de cette valeur. L'homogénéité des grandeurs d'effets au sein de chaque sous-groupe est ensuite analysée à partir d'un test statistique d'homogénéité intragroupe Q_{wi} et la différence entre les moyennes des grandeurs d'effet au sein de chaque sous-groupe est estimée par un test d'homogénéité inter-groupe Q_B. La présence d'un modérateur potentiel est évaluée par une procédure méta-analytique « complète » (Johnson & Turco, 1992). Ainsi, un test $Q_{wi}s$ non significatif (suggérant que toutes les études au sein de chaque sous-groupe partagent une grandeur d'effet commune) et un test Q_B significatif (suggérant une différence entre la moyenne des grandeurs d'effets de chaque sous-groupe) indiquent la *présence d'une variable modératrice catégorielle*. La présence simultanée d'un Q_B significatif et d'un $Q_{wi}s$ significatif suggère que d'autres modérateurs peuvent exister.

S'agissant des variables modératrices *continues*, la régression par les moindres carrés pondérés (*Weighted Least Squares – WLS*) est utilisée (Hedges, 1982b ; Hedges & Olkin, 1985). Le recours à la régression *WLS* évite la perte d'informations qui pourrait résulter d'une polychotomisation erronée des variables continues (Cohen, 1983). En utilisant la régression *WLS*, l'estimation de la grandeur d'effet d est régressée sur la variable modératrice continue Z (i.e., $\mathbf{d} = \mathbf{Z}\beta + \varepsilon$, où **d** représente un vecteur de valeurs de d, **Z** représente un vecteur de valeurs du modérateur Z, β représente le vecteur des coefficients de régression et ε représente un vecteur des résidus). Les statistiques Q_R et Q_E sont ensuite calculées, avec Q_R testant l'hypothèse nulle selon laquelle le vecteur des coefficients de régression pour le modérateur est égal à zéro ($H_0 : \beta = 0$) et Q_E évaluant la qualité globale du modèle de régression ($H_0 : \varepsilon = 0$). Un Q_R statistiquement significatif indique que la variable Z est un modérateur significatif et un Q_E non significatif indique que le modèle de régression est robuste.

1.1 Arguments en faveur de l'utilisation de la statistique d'homogénéité Q

De récentes simulations Monte-Carlo[1] ont montré que les hypothèses sur les variables modératrices – vérifiées à l'aide de la statistique Q décrite plus haut –

1 NdT. La méthode Monte-Carlo désigne une approche statistique visant à calculer une valeur en utilisant des procédés aléatoires.

affichent des taux d'erreur de type I[1] satisfaisant et une puissance statistique[2] convenable. Ainsi, une étude Monte-Carlo a montré que (a) le taux d'erreur de type I varie de 0,07 à 0,10 et (b) que la puissance statistique s'étend de 0,63 à 0,91 quelles que soient les tailles d'échantillon et le nombre d'études intégrées dans la méta-analyse (Sagie et Koslowsky, 1993). Les résultats d'une autre simulation Monte-Carlo – s'intéressant particulièrement au cas des petits échantillons – confirment l'intérêt de la statistique *Q* : (a) le taux d'erreur de type I s'étend de 0,041 à 0,045 et (b) le taux de puissance statistique est acceptable (e.g. lorsqu'une méta-analyse combine cinq études, la puissance statistique varie de 0,709 à 1,000) (Alliger, 1995).

En définitive, l'approche de Hedges et Olkin permet au méta-analyste de tester l'hétérogénéité des résultats entre les études et d'identifier la présence de variables modératrices. De surcroît, les investigations Monte-Carlo récentes indiquent que les hypothèses examinées par la statistique *Q* semblent pouvoir être testées sans risque d'écarts importants avec les taux d'erreur de type I généralement admis et avec une puissance statistique suffisante. Finalement, un autre avantage de l'approche de Hedges et Olkin est qu'elle est facile à utiliser à la fois pour tester les effets modérateurs de variables quantitatives et qualitatives.

2. Mettre en œuvre des corrections au niveau des études dans la méta-analyse

Les promoteurs des approches méta-analytiques en psychométrie s'appuient sur les arguments de la *théorie de la mesure* pour soutenir l'idée qu'une large part de la variance observée entre les études portant sur une relation d'intérêt *X-Y* est le résultat d'*artefacts* (Hunter et Schmidt, 1990 ; Raju et al. 1991 ; Schmidt, 1992 ; Schmidt et al. 1993). Dit autrement, les différences de résultats observées entre les études peuvent être dues (a) à des *artefacts* statistiques et méthodologiques et/ou (b) à des effets modérateurs. Par conséquent, afin de mieux estimer la relation entre *X* et *Y* dans la population étudiée, les chercheurs doivent (a) contrôler l'effet des *artefacts* en mettant en œuvre des méthodes de recherche robustes et (b) corriger la variabilité *artefactuelle* entre les études, en la retirant de la variance totale observée dans les estimations de la taille de l'effet. L'objectif de cette correction n'est pas d'éliminer toutes les sources de variation, mais seulement la variabilité inter-étude qui résulte des *artefacts* statistiques et méthodologiques (parmi lesquels on retrouve l'erreur d'échantillonnage et l'erreur de mesure de la variable dépendante et l'étendue de mesure limitée ou *"range restriction"* ; Aguinis et Whitehead, 1997).

Les chercheurs qui encouragent l'adoption de telles approches psychométriques soutiennent que les revues quantitatives non expérimentales (notamment fondées sur les coefficients de corrélation *r* comme métrique de la taille de l'effet) et expérimentales (fondées sur les *d* comme estimateur de l'effet) souffrent des

1 NdT. Une erreur de type I désigne une erreur de première espèce qui consiste à rejeter une hypothèse nulle alors qu'elle est fausse (ou encore à trouver une différence inexistante)

2 NdT. La puissance statistique est égale à 1 – β (β est ici l'erreur de seconde espèce, c'est-à-dire celle qui consiste à accepter une hypothèse nulle alors qu'elle est vraie). Une expérience plus puissante est une expérience qui a plus de chance de rejeter à juste titre une hypothèse nulle erronée.

mêmes lacunes : la variabilité de la taille de l'effet n'est pas seulement le résultat de différences réelles dues à des modérateurs, mais serait aussi liée à des *artefacts* méthodologiques et statistiques. Ainsi, en l'absence de correction de la variabilité artefactuelle, la variabilité inter-étude peut être attribuée à de faux effets modérateurs en commettant une erreur statistique de type I (voir Hunter et Schmidt, 1990, pp. 23-29 pour une illustration).

Certains chercheurs ont toutefois critiqué l'utilisation des corrections d'artefacts. Tout d'abord, James et ses collègues (James et al., 1986 ; James et al., 1992) rappellent que plusieurs artefacts peuvent être corrélés à des modérateurs contextuels importants (le climat organisationnel, par exemple). Ainsi, James et ses collègues soutiennent qu'en corrigeant les artefacts méthodologiques, le méta-analyste peut aussi être amené à corriger (partiellement) l'effet modérateur. Dès lors, les chercheurs risquent d'obtenir des résultats fallacieux.

Une deuxième controverse concernant ces corrections a été lancée par Murphy (1993) qui a démontré que la correction des artefacts peut tout aussi bien réduire la variabilité des tailles de l'effet que l'augmenter (voir Johnson, Mullen & Salas, 1995, p. 100, tableau 5 pour une illustration). Il existerait plusieurs situations où la distribution des effets corrigés présentera une variance plus importante qu'une distribution d'effets non corrigés (voir Murphy, 1993 pour un examen détaillé). Toutefois, les conclusions de Murphy ne doivent pas être nécessairement interprétées en opposition avec l'approche psychométrique de la méta-analyse : la variabilité entre les études en termes de taille de l'effet n'est pas indésirable en soi, c'est la variabilité causée par les artefacts que le méta-analyste cherche à minimiser.

Une troisième interrogation, s'agissant de la correction des artefacts, porte sur le fait que l'application de ces corrections peut conduire à une grandeur d'effet associée à une variance négative. En d'autres termes, la variance inter-étude due à des artefacts méthodologiques et statistiques peut être supérieure à 100 % (e.g. Rothstein, Schmidt, Erwin, Owens et Sparks, 1990, p. 179, tableau 5). Ce phénomène peut être interprété comme une preuve que (a) les corrections méthodologiques et statistiques sont intercorrélées entre elles, ou (b) qu'il existe une anomalie statistique. Ces deux interprétations suggèrent finalement une faiblesse des procédures méta-analytiques introduisant la correction des artefacts. Toutefois, plusieurs chercheurs ont montré que les résultats des méta-analyses au sein desquelles les artefacts expliquent plus de 100 % de la variance inter-étude sont un phénomène fréquent et sans conséquence. Si la variance inter-étude est entièrement artefactuelle, alors 100 % de la variance inter-étude observée est prise en compte dans les artefacts. Dans de telles situations, les valeurs de la variance artefactuelle sont supérieures à 100 % en raison de l'erreur d'échantillonnage de second ordre (Burke, 1996 ; Callender et Osburn, 1988 ; Rothstein et al., 1990).

2.1 Arguments en faveur de l'introduction de corrections

Malgré les controverses susmentionnées, les corrections des artefacts méthodologiques et statistiques telles que l'erreur de mesure et l'étendue de mesure limitée (range restriction) sont ancrées dans une vieille tradition de la méthodologie en sciences sociales qui soutient que les tailles de l'effet corrigées (des artefacts)

obtiennent de meilleures estimations que les tailles de l'effet observées (non corrigées) (Spearman, 1904 ; Thorndike, 1949, pp. 104-105).

Tout d'abord, s'agissant de l'erreur de mesure, il est bien connu que les artefacts réduisent la relation observée entre deux variables d'intérêt par rapport à la relation entre les deux construits sous-tendant ces mesures. L'objectif de la méta-analyse est de mieux comprendre la relation entre des construits et non pas seulement la relation entre des mesures discutables de ces construits (Schmidt et Hunter, 1996). Si les tailles de l'effet ne sont pas corrigées de l'erreur de mesure, les tailles de l'effet méta-analysées sont systématiquement biaisées à la baisse. De plus, les écarts d'erreur de mesure entre les études augmentent artificiellement la variance inter-étude des estimateurs de l'effet. Cette variabilité, causée par le différentiel des erreurs de mesure et non pas par des variables théoriquement modératrices, peut conduire le chercheur à conclure à tort que les tailles de l'effet varient entre les études et qu'il existerait donc des variables modératrices. Hunter et Schmidt (1990, pp. 117-125) fournissent une discussion détaillée et une synthèse des avantages d'introduire une correction des erreurs de mesure dans les investigations méta-analytiques (voir aussi Muchinsky, 1996 pour une synthèse sur la correction de l'erreur de mesure).

Il existe également une longue tradition dans la littérature en sciences sociales sur la nécessité de corriger l'étendue de mesure limitée (« *range restriction* »). Dès 1903, Pearson estime que les estimations de la taille de l'effet calculée à partir de données censurées ou d'étendue limitée sous-estiment les effets dans la population étudiée. Les méta-analystes sont souvent confrontés à des tailles de l'effet calculées à partir de scores qui couvrent seulement le niveau le plus faible des scores de la population (e.g. dans la recherche sur la sélection du personnel). Les effets des *range restrictions* sont de deux ordres : (a) les *range restrictions* produisent un biais à la baisse des tailles de l'effet et (b) les écarts de niveau des *range restrictions* entre les études augmentent la variabilité inter-étude liée aux estimateurs de l'effet. Par conséquent, les chercheurs sont invités à introduire des corrections des « *range restrictions* » (Thorndike, 1949, pp. 169-176 ; voir aussi Ree, Carrettan, Earles et Albert, 1994). Hunter et Schmidt (1990, pp. 125-133) décrivent les corrections des « *range restrictions* » en détail dans leur ouvrage et proposent quelques exemples.

En définitive, un des principaux avantages d'introduire des corrections est de permettre le contrôle de la variabilité artefactuelle dans un ensemble d'estimateurs de l'effet issus d'étude individuelle. Plus précisément, les artefacts qui accroissent la variabilité inter-étude peuvent être neutralisés par des corrections statistiques. Par conséquent, la probabilité de commettre une erreur de type I en attribuant la variance artefactuelle inter-étude à de « faux » effets modérateurs est maintenue à un niveau résiduel.

L'approche méta-analytique de Hedges et Olkin permet de corriger un certain type d'artefacts méthodologiques (i.e. erreur d'échantillonnage par le biais de la statistique *Q* et l'erreur de mesure de la variable dépendante). Plus précisément, à la page 135 de l'ouvrage d'Hedges et Olkin (1985), l'équation 39 montre que les estimateurs *ds* peuvent être individuellement corrigés de l'erreur de mesure sur la variable dépendante. De même, l'équation 40 (p. 136) indique comment

calculer une estimation pondérée de la moyenne corrigée *d* (i.e. d^R) et l'équation 43 (p. 137) montre comment calculer la statistique d'homogénéité Q^R. Cependant, les formules de Hedges et Olkin (1985) permettent uniquement de corriger les *ds* fondés sur l'erreur de mesure. Il existe d'autres artefacts tels que la *range restriction* et la dichotomisation des variables continues qui sont fréquents dans les recherches en management.

Même si Hedges et Olkin (1985) suggèrent de corriger les erreurs de mesure des études de façon individuelle et que l'approche proposée est compatible avec l'utilisation de corrections, les méta-analystes n'ont manifestement pas conscience des possibilités offertes par cette approche. Nous avons mené une synthèse de toutes les méta-analyses publiées dans le *Psychological Bulletin* et le *Journal of Applied Psychology* entre janvier 1991 et janvier 1996. 30 des 55 articles au sein desquels une méta-analyse a été menée utilisent l'approche de HO. Toutefois, aucun de ces articles ne propose une correction de l'erreur de mesure. Ainsi, il semble que – malgré la possibilité offerte par l'approche de HO de corriger l'erreur de mesure – les utilisateurs de l'approche de HO choisissent de ne pas introduire cette correction de façon délibérée ou tout simplement parce qu'ils ne la connaissent pas.

En somme, l'approche de HO permet la correction des artefacts au niveau de chaque étude. Toutefois, cette approche ne permet que la correction de l'erreur de mesure, au-delà de la correction de l'erreur d'échantillonnage qui est prise en compte à travers la statistique *Q*. Finalement, bien que le cadre d'analyse de HO n'est pas incompatible avec la correction des tailles de l'effet, les méta-analystes utilisant cette approche en tiennent rarement compte.

3. Introduire une correction intra-étude[1] dans la procédure de Hedges et Olkin : une procédure en trois étapes

L'approche de Hedges et Olkin (HO) propose d'utiliser la statistique *Q* afin d'identifier l'existence de variables modératrices, mais cette statistique ne tient compte que de l'erreur de mesure. Dès lors, les utilisateurs de l'approche de Hedges et Olkin n'intègrent pas la correction de l'ensemble des artefacts méthodologique et statistique. Nous proposons une procédure en trois étapes afin (a) de corriger la variabilité intra-étude due aux artefacts méthodologiques et statistiques, (b) de tester l'homogénéité totale des tailles de l'effet calculées après avoir contrôlé les artefacts méthodologiques et (c) de tester l'influence de variables modératrices potentielles. Chaque étape est décrite dans le détail ci-après.

1 NdT. Dans le modèle à effet fixe (cf. chapitre 1), les vraies grandeurs d'effet sont considérées comme identiques, quelle que soit l'étude. Cette hypothèse ne permet pas de tenir compte de l'influence des caractéristiques de l'étude sur la taille de l'effet. Le modèle aléatoire qui est retenu le plus fréquemment par les méta-analystes consiste à considérer que la vraie grandeur de l'effet est une variable aléatoire distribuée normalement autour d'une grandeur d'effet constante ρ de variance v^2. L'objectif de la méta-analyse est d'estimer le paramètre ρ. Dans ce cadre d'analyse, la variance totale d'une observation ρ_i au niveau d'une étude est la somme de la variance inter-étude et de la variance intra-étude.

3.1 Étape 1

La variance inter-étude liée à des sources artefactuelles est retranchée de la variabilité totale inter-étude observée. La variabilité artefactuelle peut être neutralisée en corrigeant individuellement chaque estimation de la taille de l'effet issue d'une étude singulière (Hunter et Schmidt, 1990, chapitre 3). Les informations nécessaires à la correction des artefacts peuvent être (a) obtenues à partir des études primaires elles-mêmes, (b) estimées à partir des études précédentes, (c) calculées à partir d'un sous-échantillon d'études au sein desquelles l'information est reportée ou (d) estimées à partir de régression ou des techniques de maximum de vraisemblance. Roth (1994) propose une discussion détaillée et des recommandations techniques afin de résoudre le problème des données manquantes susceptibles d'être utiles pour corriger les artefacts statistiques.

La première étape de la procédure consiste à corriger individuellement chaque estimation de la taille de l'effet des artefacts identifiés dans la littérature (Hunter et Schmidt, 1990)[1]. Prenons comme exemple le cas de la correction *d'un seul type* d'artefact, l'estimation de la taille de l'effet r issue d'une étude individuelle peut être corrigée de « *l'étendue de mesure limitée* » (*range restriction*)[2] à partir du coefficient de corrélation dans la population $\hat{\rho}$ qui est obtenu en divisant r/a avec :

$$a = \frac{u}{\sqrt{(u^2 - 1)r^2 + 1}} \qquad (1)$$

où r représente l'estimation de la taille de l'effet issue d'une étude, et u le rapport entre l'écart-type de l'échantillon et celui de la population (Hunter et Schmidt, 1990, p. 48)

Si, en revanche, le méta-analyste s'intéresse à des effets expérimentaux (ou dichomotiques) et cumule des statistiques *d*, la variabilité inter-étude « artefactuelle » peut être évaluée directement en corrigeant les *ds*. En effet, les *ds* peuvent être facilement convertis en coefficient de corrélation de Pearson (*rs*). Ainsi, la correction des études individuelles (intra-étude) s'applique aux *rs*. Quelques méta-analystes (Hunter et Schmidt, 1990 ; Rosenthal,1991) recommandent de transformer les *ds* en *rs* pour faciliter leur interprétation et permettre l'utilisation de techniques multivariées telles que l'analyse de corrélation partielle, l'analyse des chemins causaux (*path analysis*), ou la régression multiple.

1 NdT. Hunter et Schmidt (1990, p. 45, tableau 2.1.) proposent une liste de onze artefacts méthodologiques et statistiques à laquelle nous renvoyons le lecteur.

2 La corrélation est un indicateur standardisé (variant de 0 à 1 en valeur absolue), dont la valeur va dépendre du degré de variation de la variable indépendante. Au sein d'une population parfaitement homogène (sans variance) sur la variable indépendante, la corrélation avec la variable dépendante est égale à zéro. Si l'on fixe le niveau d'erreur autour de la droite de régression et que l'on considère les individus s'éloignant de plus en plus de la variable indépendante, les corrélations peuvent s'approcher de 1. Si les corrélations issues de différentes études doivent être comparées, alors les différences de corrélations liées à des différences de répartition (variance) sur la variable indépendante doivent être contrôlées. C'est le problème de l'étendue de la variation sur la variable indépendante (*range variation* ou *range restriction in the independent variable*) (Hunter et Schmidt, 1990, p. 47).

Convertir les *ds* en *rs* nécessite que les *ds* soient d'abord eux-mêmes convertis en coefficients de corrélation bisériale ($r_{pb}s$)[1] et que ces coefficients soient ensuite transformés en *rs*, car les $r_{pb}s$ sous-estiment les *rs* lorsque la taille des sous-groupes est inégale. L'équation 2 (Wolf, 1986, p. 35) peut être utilisée pour transformer les *ds* en $r_{pb}s$:

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

Glass et Stanley (1970, p. 171) ont proposé la formule suivante pour passer de $r_{pb}s$ à *rs* :

$$r = \frac{r_{pb}\sqrt{n_1 n_2}}{uN} \qquad (3)$$

Où n_1 représente la taille de l'échantillon du sous-groupe 1, n_2 la taille de l'échantillon du sous-groupe 2, $N = n_1 + n_2$ et u représente l'ordonnée (i.e. hauteur) d'une unité de distribution normale qui représente au moins $100 \times (n_1/N)$ % de l'aire sous la courbe.

Enfin, chaque *r* intra-étude corrigé doit être converti en statistique *d* en utilisant l'équation 4 (Wolf, 1986, p. 35) :

$$d = \frac{2r}{\sqrt{1 - r^2}} \qquad (4)$$

À ce stade, toutes les grandeurs d'effet (*rs* ou *ds*) sont maintenant exprimées dans la même métrique *d* et ont été individuellement corrigées des artefacts méthodologique et statistique. Ce sont ces *ds* corrigés qui vont être considérés dans la deuxième étape.

3.2 Étape 2

Les *ds* intra-études corrigés sont soumis à un test d'homogénéité qui s'appuie sur une statistique *Q* modifiée (la formule initiale est présentée par Hedges et Olkin, 1985, p. 123). La statistique *Q* se distribue suivant une loi du khi-deux à $k - 1$ degré de liberté avec *k* correspondant au nombre d'études :

$$Q = \sum \frac{(d_i - d_+)^2}{\hat{\sigma}'^2(d_i)} \qquad (5)$$

où d_i représente l'estimation de la grandeur d'effet de l'étude *i* corrigée (des artefacts) et ajustée (du biais lié à la taille de l'échantillon en multipliant l'estimation par $1 - [3/(4N - 9)]$, Hedges et Olkin, 1985, p. 81), d_+ est la grandeur d'effet moyenne corrigée et ajustée et $\hat{\sigma}'^2(d_i)$ est la variance corrigée de d_i (cf. chapitre 1).

1 Un coefficient de corrélation bisériale (ou bisérielle) est un indicateur d'association entre une variable dichotomique et une variable quantitative dont les valeurs peuvent être considérées comme des points séparés (par exemple, une échelle de Likert à 7 points demandant à un individu d'évaluer sa satisfaction au travail).

Les grandeurs d'effet ajustées et non biaisées (Johnson, 1989) sont utilisées parce qu'il s'agit d'estimations plus précises de la grandeur d'effet dans la population. De surcroît, les estimations ajustées minimisent la variance inter-étude des grandeurs d'effet (Hedges et Olkin, 1985).

Soulignons que la variance $\hat{\sigma}'^2(d_i)$ de l'équation (5) diffère de $\hat{\sigma}^2(d_i)$ proposée généralement par Hedges et Olkin. La variance non corrigée de l'erreur d'échantillonnage se calcule comme suit (cf. Hedges et Olkin, 1985, p. 86) :

$$\hat{\sigma}^2(d_i) = \frac{n_1 + n_2}{n_1 n_2} + \frac{d_i^2}{2(n_1 + n_2)} \qquad (6)$$

Avec n_1 correspondant à la taille de l'échantillon du groupe expérimental et n_2 à la taille de l'échantillon du groupe de contrôle. (Lorsque la statistique *d* est calculée à partir d'un coefficient de corrélation de Pearson *r* et par conséquent que l'on dispose d'un échantillon de taille *N*, on considère que $N = n_1 + n_2$ et $n_1 = n_2$; Johnson, 1989).

Une fois les tailles de l'effet corrigées des artefacts, leurs variances liées aux erreurs d'échantillonnage sont affectées. Dans le prolongement de l'exemple présenté dans l'étape 1, considérons $u_1 \ldots u_k$ les facteurs de correction de chaque étude *k* (Hunter et Schmidt, 1990, p. 254, voir aussi l'équation 1). Les variances liées aux erreurs d'échantillonnage *corrigées* sont (Hedges, 1995) :

$$\sigma'^2_1 = u_1^2\hat{\sigma}_1^2,\ \hat{\sigma}'^2_2 = u_2^2\hat{\sigma}_2^2,\ \ldots,\ \hat{\sigma}'^2_k = u_k^2\hat{\sigma}_k^2 \qquad (7)$$

Au besoin, l'équation 7 peut être développée afin de corriger les variances d'autres artefacts comme l'erreur de mesure sur la variable dépendante (Hedges, 1995).

$$\hat{\sigma}'^2_1 = u_1^2 w_1^2\hat{\sigma}_1^2,\ \hat{\sigma}'^2_2 = u_2^2 w_2^2\hat{\sigma}_2^2,\ \ldots,\ \hat{\sigma}'^2_k = u_k^2 w_k^2\hat{\sigma}_k^2 \qquad (8)$$

où *w* représente le facteur de correction de l'erreur de mesure ($w = 1/[r_{yy}]^{1/2}$).

Une fois les *ds* corrigés au cours de la première étape, et que les équations 7 et 8 ont permis de contrôler les erreurs d'échantillonnage lié aux artefacts, l'ensemble de la procédure de Hedges et Olkin est acceptable (Hedges, 1995).

La valeur et la significativité statistique du test *Q* (équation 5) permettent ensuite d'évaluer et de déterminer si l'estimation moyenne des grandeurs d'effet dans l'échantillon d'études est homogène (si elles partagent bien une grandeur d'effet commune). Un test *Q* significatif indique la présence d'une variance inexpliquée et, par conséquent, conduit à chercher les causes de cette hétérogénéité. De surcroît, si la théorie prédit un effet modérateur particulier, il est conseillé de mener un test des effets modérateurs (étape 3) quand bien même le test *Q* serait non significatif (Hedges et Olkin, 1985 ; Johnson et Turco, 1992).

3.3 ÉTAPE 3

Le modèle à effets fixes d'Hedges et Olkin est mis en œuvre pour tester la présence de variables modératrices à partir des estimateurs *ds* corrigés et ajustés au niveau des études et de la variance corrigée des erreurs d'échantillonnage. Comme nous l'avons vu, la nature des variables modératrices, catégoriques ou continues, conduira à choisir une approche plutôt qu'une autre.

Afin de tester des modèles catégoriels, les équations 9 et 11 proposent des formules adaptées pour calculer Q_B et Q_{wi} (les formules d'origine sont présentées dans Hedges et Olkin, 1985, pp. 154-155). Q_B suit une loi de distribution du khi-deux à p-1 degrés de liberté, avec p correspondant au nombre de classes ou de modalités du modérateur potentiel :

$$Q_B = \sum \frac{(d_{i+} - d_{++})^2}{\hat{\sigma}'^2(d_{i+})}, \qquad (9)$$

où d_{i+} représente l'estimation de la taille de l'effet moyenne corrigée et ajustée de la $ni^{ème}$ modalité de la variable modératrice, d_{++} représente la moyenne ajustée des k estimations des tailles de l'effet corrigées au niveau individuel, et $\hat{\sigma}'^2(d_{i+})$ représente la variance de la taille de l'effet corrigé d_{i+}. Soulignons que cette variance intraclasse corrigée est différente de la variance non corrigée (proposée par Hedges et Olkin, 1985, p. 152) parce qu'elle combine les variances inter et intraclasses corrigées :

$$\hat{\sigma}'^2(d_{i+}) = \left(\sum \frac{1}{\hat{\sigma}'^2(d_{ij})} \right)^{-1} \qquad (10)$$

Q_w suit une distribution du khi-deux à m – 1 degrés de liberté où m représente le nombre d'estimateurs de l'effet pour une modalité donnée du modérateur :

$$Qw_i = \sum \frac{(d_{ij} - d_{i+})^2}{\hat{\sigma}'^2(d_{ij})} \qquad (11)$$

où d_{ij} représente l'estimation de la taille de l'effet corrigée de la $ni^{ème}$ classe du modérateur et la $ni^{ème}$ étude, d_{i+} correspond à l'estimation de la taille de l'effet corrigée et ajustée pour la $ni^{ème}$ classe du modérateur et $\hat{\sigma}'^2(d_{ij})$ est la variance de la taille de l'effet corrigée de d_{ij}. L'équation 11 est utilisée pour chaque classe, niveau ou catégorie de la variable modératrice.

Afin de tester des modèles continus, les équations 12 et 13 présentent les formules pour calculer Q_R et Q_E (voir Hedges et Olkin, 1985 pp. 168-172). Q_R suit une distribution du khi-deux avec l degré de liberté avec l correspondant au nombre de coefficients de régression au sein du vecteur β :

$$Q_R = \hat{\beta}' \hat{\sum}_{\beta}^{-1} \hat{\beta} \qquad (12)$$

où $\hat{\beta}'$ représente la transposée des vecteurs des coefficients de régression estimés par les moindres carrés généralisés (MCG) et $\hat{\Sigma}_{\beta}$ est une estimation de la matrice de covariance Σ_{β}.

Q_E suit une distribution du khi-deux à $k - p - 1$ degrés de liberté avec k égal au nombre d'études et p au nombre de variables prédictives ou explicatives :

$$Q_E = d' \hat{\sum}_d^{-1} d - Q_R \qquad (13)$$

où d' représente la transposée d'un vecteur de grandeurs d'effet d corrigées, $\hat{\Sigma}_d$ représente une estimation de la matrice diagonale de covariance Σ_d, et Q_R correspond à la valeur obtenue à partir de l'équation 12.

4. Une illustration de l'utilisation de la procédure en trois étapes

Afin d'illustrer l'utilisation de la procédure méta-analytique en trois étapes que nous venons de décrire, nous avons sélectionné arbitrairement un ensemble de données tiré de l'ouvrage de Hunter et Schmidt (1990, p. 24, tableau 1.1, études 1 à 10). Le tableau 1 présente un ensemble fictif de 10 études qui examinent indépendamment la relation entre l'implication dans l'organisation et la satisfaction au travail. Pour chacune des 10 études, on dispose de (a) la taille de l'échantillon (N), (b) du coefficient de corrélation de Pearson (r) entre l'implication et la satisfaction, et (c) du sexe du répondant (masculin *vs* féminin). Considérons que l'objectif de l'étude est de tester l'hypothèse selon laquelle le sexe du répondant est une variable modératrice de la relation entre l'implication organisationnelle et la satisfaction au travail et que le fait d'être une femme renforce cette relation. Passons maintenant à la procédure méta-analytique en trois étapes. Les calculs peuvent aisément être réalisés à l'aide d'un tableur voire d'une simple calculatrice si le nombre d'études n'est pas trop important.

4.1 Étape 1

Tableau 2.1 — Une illustration de la procédure méta-analytique en trois étapes à partir des données de Hunter et Schmidt (1990, p. 24, Tableau 1.1.)

Étude	N	r	Sexe	a	$\hat{\rho}$	d	d_i	$(d_i - d_+)^2$	$\hat{\sigma}^2(d_i)$	$\hat{\sigma}'^2(d_i)$	$(d_{ij} - d_{i+})^2$
1	20	.46	F	.65	.71	2.03	1.95	.0529	.29478	.10612	.0036
2	72	.32	M	.62	.52	1.20	1.19	.2809	.06540	.02354	.0064
3	29	.10	M	.60	.17	.34	.33	1.9321	.13978	.05032	.8836
4	30	.45	M	.64	.70	1.96	1.91	.0361	.19387	.06979	.4096
5	71	.18	F	.61	.30	.62	.61	1.2321	.05900	.02124	1.9600
6	62	.45	F	.64	.70	1.96	1.93	.0441	.09469	.03409	.0064
7	25	.56	F	.67	.83	3.03	2.93	1.4641	.33150	.11934	.8464

Étude	N	r	Sexe	a	$\hat{\rho}$	d	d_i	$(d_i - d_+)^2$	$\hat{\sigma}^2(d_i)$	$\hat{\sigma}'^2(d_i)$	$(d_{ij} - d_{i+})^2$
8	46	.41	M	.64	.65	1.69	1.66	.0036	.11696	.04211	.1521
9	22	.55	F	.67	.82	2.90	2.79	1.1449	.35865	.12911	.6084
10	69	.44	F	.64	.69	1.89	1.87	.0225	.08322	.02996	.0196

Notes : N = taille de l'échantillon

r = coefficient de corrélation de Pearson entre l'implication organisationnelle et la satisfaction au travail

F = Féminin

M = Masculin

a = facteur de correction de la "range restriction" calculé à partir de l'équation 1

$\hat{\rho} = r/a$, coefficient de corrélation corrigé de la « range restriction »

d = transformation de $\hat{\rho}$ à d en utilisant l'équation 4

$d_i = d$ corrigé (des artefacts) et ajusté $d_i = d \times 1 - [3/(4N - 9)]$

d_+ = moyenne générale = moyenne de tous les d_is = 1,72 (r = .65) dans cet exemple

$(d_i - d_+)^2$ = écart au carré entre la grandeur d'effet de chaque étude et la moyenne générale

$\hat{\sigma}^2(d_i)$ = variance non corrigée de l'erreur d'échantillonnage calculée à partir de l'équation 6

$\hat{\sigma}'^2(d_i)$ = variance corrigée de l'erreur d'échantillonnage calculée à partir de l'équation 7

$(d_{ij} - d_{i+})^2$ = écart au carré entre la grandeur d'effet de chaque étude et la moyenne du groupe

La première étape consiste à corriger les estimations de la grandeur d'effet au niveau de chaque étude des sources de variation artefactuelle. Afin d'être cohérents avec l'exemple proposé précédemment, nous corrigeons les *r*s de chaque étude afin d'obtenir pour chaque étude une corrélation $\hat{\rho}$. ($\hat{\rho} = r/a$, cf. équation 1). Pour simplifier, considérons que u (le rapport entre les écarts-types de l'échantillon et celui de la population) est égal à 0.60 pour les 10 études. Le tableau 1 présente les valeurs de a pour chaque étude (calculées à partir de l'équation 1) et les valeurs de chaque $\hat{\rho}$, c'est-à-dire le coefficient de corrélation corrigé de la *"range restriction"*. Ensuite, il convient de convertir chaque coefficient de corrélation corrigé en d en utilisant l'équation 4. Soulignons ici que la transformation en métrique d est nécessaire afin de calculer la statistique Q dans les étapes 2 et 3. Cependant, afin d'être cohérents avec la métrique d'origine et afin de faciliter l'interprétation des résultats, les *d*s peuvent être convertis en r à n'importe quel moment de la procédure. À la fin de la première étape, les grandeurs d'effet sont exprimées dans la métrique d et ont été individuellement corrigées des artefacts méthodologiques et statistiques.

4.2 Étape 2

Comme il est suggéré dans l'approche de Hedges et Olkin, l'homogénéité des grandeurs d'effet *d*s corrigées est testée à l'aide de la statistique modifiée Q (cf. équation 5). Cette étape impose que chaque grandeur d'effet d soit convertie en d_i en multipliant chaque d par $1 - [3/(4N - 9)]$. Comme prévu, on observe de plus grandes différences entre d (corrigé et non ajusté) et d_i (corrigé et ajusté) lorsque les études portent sur de petits échantillons (cf. tableau 1, colonnes 7 et 8).

Après avoir calculé les *d*s, il convient de calculer l'estimateur moyen corrigé et ajusté d_+ qui est simplement la moyenne de tous les d_is (d_+ = 1,72 ou converti

en métrique r égal à 0,65). Ensuite, nous calculons le numérateur de l'équation 5, c'est-à-dire la différence au carré entre chaque grandeur d'effet individuelle d_i et la moyenne d_+ ($[d_i - d_+]^2$). Nous calculons derrière les valeurs du dénominateur de l'équation 5, c'est-à-dire les variances corrigées de l'erreur d'échantillonnage ($\hat{\sigma}'^2(d_i)$). On obtient ces valeurs en calculant tout d'abord les variances non corrigées (équation 6) pour chaque d_i puis ensuite en utilisant l'équation 7 pour convertir chaque variance non corrigée $\hat{\sigma}^2(d_i)$ en variance corrigée $\hat{\sigma}'^2(d_i)$. Dans cet exemple, les estimateurs de la grandeur d'effet sont les coefficients *rs* et par conséquent on ne dispose que d'un échantillon de taille *N*. Ainsi, en accord avec l'approche de Hedges et Olkin, dans l'équation 6, nous considérons que $N = n_1 + n_2$ et $n_1 = n_2$ (Johnson, 1989 ; Johnson et Eagly, sous presse).

Finalement, la statistique Q est calculée, comme indiqué dans l'équation (5), $Q(9) = 132{,}15$ avec $p < 0{,}01$. Dès lors, il est possible de conclure que les grandeurs d'effets corrigées évaluant le lien entre l'implication organisationnelle et la satisfaction au travail sont hétérogènes. Ainsi, il convient de passer à la troisième étape de la procédure afin de tester l'hypothèse selon laquelle l'hétérogénéité constatée entre les résultats des études existantes peut être expliquée par l'effet *modérateur* de la variable « sexe du répondant ».

4.3 Étape 3

La troisième étape consiste à calculer les statistiques Q_B et Q_{wi} à l'aide des équations 9 et 11 respectivement. Toutes les informations nécessaires à ces calculs sont présentées dans le tableau 1.

Calcul de Q_B. On utilise l'équation 9 pour calculer Q_B et ainsi obtenir une grandeur d'effet moyenne corrigée et ajustée (d_{i+}) à partir des d_i*s* de chaque sous-groupe de modérateur (ici. masculin/féminin. Il y a six études dont les répondants sont des femmes et 4 dont les répondants sont des hommes). Pour les hommes, $d_{1+} = 1{,}27$ ($r = 0{,}54$, cf. eq. 2) et pour les femmes $d_{2+} = 2{,}01$ ($r = 0{,}71$, cf. eq. 2). On obtient ensuite les écarts au carré entre la grandeur d'effet moyenne dans chaque groupe et la moyenne de tous les k estimations de la grandeur d'effet corrigée au niveau individuel (($[d_{i+} - d_{++}]^2$, avec $d_{++} = 1{,}72$). Cela conduit au calcul d'un écart au carré de 0,2025 pour les hommes et de 0,0841 pour les femmes qui sera utilisé comme numérateur dans l'équation 9.

S'agissant du dénominateur de l'équation 9, on calcule la variance des grandeurs d'effet corrigées pour le groupe des hommes $\hat{\sigma}'^2(d_{1+})$ et celui des femmes $\hat{\sigma}'^2(d_{2+})$ en utilisant l'équation 10. Il est relativement aisé d'obtenir ces valeurs, car chaque variance d'estimateur d'effet corrigé intraclasse est ni plus ni moins que l'agrégation des variances des estimateurs de l'effet au sein de chaque classe ($\hat{\sigma}'^2(d_i)$) qu'on obtient au cours de la première étape de la procédure. La variance pour les hommes est de 0,00996 et la variance pour les femmes est de 0,00739. Ensuite, étant donné les valeurs de $[d_{i+} - d_{++}]^2$ et de $\hat{\sigma}'^2(d_{i+})$, $Q_B(1) = 32{,}54$, $p < 0{,}01$. Par conséquent, on peut conclure qu'il existe un effet modérateur statistiquement significatif du sexe sur le lien entre l'implication organisationnelle et la satisfaction au travail et que cette relation est plus forte pour les femmes ($r = 0{,}71$) que pour les hommes ($r = 0{,}54$).

Calcul de Q_{wi}. Afin de calculer Q_{wi} pour chacun des deux sous-groupes, nous devons tout d'abord obtenir l'écart au carré entre les estimateurs de l'effet de chaque étude au sein de chaque groupe et la moyenne du groupe ($[d_{ij} - d_{i+}]^2$). La taille de l'effet issue de chaque étude est présentée dans le tableau 1 (i.e. d_is au sein de chaque classe) et les moyennes des deux classes sont présentées sous le tableau (pour les hommes d_{1+}= 1,27 et pour les femmes d_{2+} = 2,01). L'information nécessaire pour le dénominateur est également présentée dans le tableau 1 : $\hat{\sigma}'^2(d_{ij})$ correspond à la variance de la taille de l'effet corrigée pour les études issues de la même classe (i.e. $\hat{\sigma}'^2(d_i)s$). Ainsi, il ne reste plus qu'à calculer Q_{wi} à partir de $[d_{ij} - d_{i+}]^2$ et $\hat{\sigma}'^2(d_{ij})$ pour chaque étude. Pour la classe des hommes, $Q_{wi}(3) = 59{,}82$, $p < 0{,}01$ et pour celle des femmes $Q_{wi}(5) = 80{,}58$, $p < 0{,}01$.

En définitive, à partir de ces résultats, nous pouvons conclure qu'il existe une forte hétérogénéité des résultats obtenus par les 10 études existantes dans la littérature consacrée au lien entre l'implication organisationnelle et la satisfaction au travail. Ainsi, des effets modérateurs semblent affecter cette relation. Par ailleurs, nous avons identifié une hétérogénéité *entre* les classes ainsi qu'*au sein* de chaque classe lorsque nous avons testé l'effet modérateur de la variable « sexe du répondant ». Ces résultats suggèrent (a) qu'il existe des modérateurs qui affectent la relation entre l'implication et la satisfaction, (b) que le sexe du répondant est l'un de ces modérateurs et (c) que la présence d'une hétérogénéité *intraclasse* indique qu'il existe certainement d'autres modérateurs, c'est-à-dire d'autres variables qui peuvent affecter l'intensité de la relation entre l'implication organisationnelle et la satisfaction au travail.

5. Synthèse et conclusion

Comme l'ont indiqué Hall et Rosenthal (1991, p. 447), « *la recherche de variables modératrices n'est pas seulement une entreprise intellectuellement stimulante, mais aussi ...au cœur de la recherche scientifique* »[1]. Au regard de l'avancée des théories en management au cours de ces dernières décennies, et du nombre toujours croissant de modèles théoriques prenant en considération des relations complexes entre les phénomènes étudiés, il n'est pas surprenant que les méta-analyses des effets modérateurs suscitent de plus en plus d'intérêts.

Ce chapitre insiste sur l'utilisation de la correction des artefacts méthodologique et statistique dans le cadre de l'approche méta-analytique de Hedges et Olkin. Il serait judicieux de mener à l'avenir des études Monte-Carlo pour examiner les effets de ce type de correction dans l'approche de Hedges et Olkin et d'étudier plus particulièrement les propriétés statistiques des tests Q ajustés à l'aide de ces corrections. Cela étant, il existe déjà quelques preuves empiriques (Pearson, 1903 ; Alliger, 1995 ; Sagie et Koslowsky, 1993) concernant (a) les avantages d'introduire ce type de corrections des effets et (b) la précision des tests d'homogénéité Q fondés sur des tailles de l'effet corrigées. Par conséquent, nous recommandons vivement l'utilisation de ce type de corrections de la variabilité entre études liée

1 *"The search for moderator variables is not only an exciting intellectual enterprise but, indeed... it is at the very heart of the scientific enterprise."*

aux artefacts. Nous conseillons également le recours aux statistiques *Q* qui incorporent une correction de la variance liée aux erreurs d'échantillonnage. Le recours aux statistiques *Q* permet aux chercheurs d'expliquer la variabilité des grandeurs d'effet entre les études existantes en testant ouvertement des hypothèses sur les variables modératrices.

BIBLIOGRAPHIE

Aguinis, H. (1995). Statistical power problems with moderated multiple regression in management research. *Journal of Management*, 21: 1141-1158

Aguinis, H., Bommer, W. H. & Pierce, C. A. (1996). Improving the estimation of moderating effects by using computer-administered questionnaires. *Educational and Psychological Measurment*, 56: 1043-1047.

Aguinis, H. & Pierce, C. A. (1998). Heterogeneity of error variance and the assessment of moderating effects of categorical variables: A conceptual review. *Organizational Research Methods*, 56: 296-314.

Aguinis, H., Pierce, C. A. & Quigley, B. M. (1993). Conditions under which a bogus pipeline procedure enhances the validity of self-reported cigarette smoking: A meta-analytic review. *Journal of Applied Social Psychology*, 23: 352-373.

Aguinis, H. Pierce, C. A. & Quigley, B. M. (1995). Enhancing the validity of self-reported alcohol and marijuana consumption using a bogus pipeline procedure: A meta-analytic review. *Basic and Applied Social Psychology*, 16: 515-527.

Aguinis, H., Pierce, C. A. & Stone-Romero, E. F. (1994). Estimating the power to delect dichotomous moderators with moderated multiple regression. *Educational and Psychological Measurement*, 54: 690-692.

Aguinis, H. & Stone-Romero, E. F. (1997). Methodological artifacts in moderated multiple regression and their effects on statistical power. *Journal of Applied Psychology*, 82: 192-206.

Aguinis, H. & Whitehead, R. (1997). Sampling variance in the correlation coefficient under indirect range restriction: Implications for validity generalization. *Journal of Applied Psychology*, 82: 528-538.

Alliger, G. M. (1995). The small sample performance of four tests of the difference between pairs of meta-analytically derived effects sizes. *Journal of Management*, 21: 789-799.

Burke, M. J. (1996). *An impirical examination of the effects of second-order sampling error on selection test validity coefficients.* In M. J. Burke (Chair), Quantitative issues in personnel selection research and practice. Symposium conducted at the meeting of the Society for Industrial and Organizational Psychology, San Diego, CA.

Callender, J. C. & Osburn, H. G. (1988). Unbiased estimation of sampling variance of correlations. *Journal of Applied Psychology*, 73: 312-315.

Cohen, J. (1983). The cost of dichotomization. *Applied Psychology Measurement*, 7: 249-253.

Cooper, H. M. & Hedges, L. V. (Eds.), (1994). *The handbook of research synthesis*. New York: Russell Sage Foundation.

Cooper, H. M. & Lemke, K. M. (1991). On the role of meta-analysis in personality and social psychology. *Personality and Social Psychology Bulletin*, 17: 245-251.

Cotton, J. L. & Tuttle, J. M. (1986).Employee turnover: A meta-analysis and review with implications for research. *Academy of Management Review*, 11: 55-70.

Dobbins, G. H., & Platz, S. J. (1986). Sex differences in leadership: How real are they? *Academy of Management Review*, 11: 118-127.

Glass, G. V. & Stanley, J. C. (1970). *Statistical methods in education and psychology*. Englewood Cliffs, NJ: Prentice-Hall.

Hall, J. A. & Rosenthal, R. (1991). *Testing for moderator variables in meta-analysis: Issues and methods*. Communication Monographs, 58: 437-448.

Hedges, L. V. (1982a). Fitting categorical models to effect sizes from a series of experiments. *Journal of Educational Statistics*, 7: 119-137.

Hedges, J. V. (1982b). Fitting continuous models to effect size data. *Journal of Educational Statisics*, 7: 245-270.

Hedges, L. V. & Olkin, I. (1985). *Statistical Methods for Meta-Analysis*. Orlando, FL: Academic Press.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings*. Newburry Park, CA: Sage.

Iaffaldano, M. T. & Muchinsky, P. M. (1985). Job satisfaction and job performance: A meta-analysis. *Psychological Bulletin*, 97: 251-273.

James, L. R., Demaree, R. G. & Mulaik, S. A. (1986). A note on validity generalization procedures. *Journal of Applied Psychology*, 71: 440-450.

James, L. R., Demaree, R. G., Mulaik, S. A. & Ladd, R. T. (1992). Validity generalization in the context of situational models, *Journal of Applied Psychology*, 77: 3-14.

Johnson, B. T. (1989). *DSTAT: Software for the meta-analytic review of research literatures*. Hillsdate, NJ: Erlbaum.

Johnson, B. T., Mullen, B. & Salas, E. (1995). Composition of three major meta-analytic approaches, *Journal of Applied Psychology*, 80. 94-106.

Johnson, B. T., Mullen. B. & Salas, E. 1995, Comparison of three major meta-analytic approaches. *Journal of Applied Psychology*, 80: 94-106.

Johnson, B. T. & Turco, R. M., 1992. *The value of goodness-of-fit indices in meta-analysis; A comment on Hall and Rosenthal*, Communication Monographs, 59: 388-396.

Muchinsky, P. M, 1996. The correction for attenuation. *Educational and Psychological Measurement*. 56: 63-75.

Mullen, B., Salas, E. & Miller, N, 1991. Using meta-analysis to test theoretical hypotheses in social psychology. *Personality and Social Psychology Bulletin*. 17: 258-264.

Murphy, K. R., 1993. *The situational specificity of validities: Correcting for statistical artifacts does not always reduce the trans-situational variability of correlation coefficients*.

Oliver, M. B. & Hyde, J. S., 1993. Gender differences in sexuality: A Meta-Analysis. *Psychological Bulletin*. 114.-29-51.

Pearson, K. 1903. Mathematical contributions to the theory of evolution: II. On the influenceof natural selection on the variability and correlation of organs. *Royal Society Philosophical Transactions*. 200 (Series A) 1-66.

Petty, M. M., McGee, G. W. & Cavender, J. W. 1984. A meta-analysis of the relationships between individual job satisfaction and individual performance. *Academy of Management Review*, 9: 712-721.

Raju, N. S., Burke, M. J., Norinand. J. & Langlcis, G. M. 1991. A new meta-analytic approach. *Journal of Applied Psychology*, 16: 432-446.

Ree, M. J., Carretta, T. R., Earles, J. A. & Albert, W. 1994. Sign changes when correcting for range restriction: A note on Pearson's and Lawley's selection formulas. *Journal of Applied Psychology*. 79: 298-301.

Rosenthal, R. 1991. *Meta-analytic procedures for social research* (rev., ed.,). Beverly Hills. CA; Sage.

Roth, P. L. 1994. Missing data: A conceptual review for applied psychologists. *Personnel Psychology*. 47: 537-560.

Rothstein, H. R., Schmidt, F. L., Erwin, R. W., Owens, W. A. & Sparks, C. P. 1990. Biographical data in employment selection: Can validities be made generalizable? *Journal of Applied Psychology*. 75: 175-184.

Sagie, A. & Koslowsky, M. 1993. Detecting moderators with meta-analysis: An evaluation and comparison of techniques. *Personnel Psychology*. 46: 629-640.

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. & Hunter, J. E. 1996. Measurement error in psychologica! research: Lessons from 26 research scenarios. *Psychological Methods*. 1: 199-223.

Schmidt, F. L., Law. K., Hunter, J. E., Rothstein, H. R., Peariman. K. & McDaniel, M. 1993. Refinements in validity generalization meihods; Implications for the situational specificity hypothesis. *Journal of Applied Psychology*. 7-12.

Spearman, C. 1904. The proof and measurement of association between two things. *American Journal of Psychology* 15:72-101.

Stone-Romero, E. F., Aihger, G. M. & Aguinis. H. 1994. Type II error problems in the use of moderated multiple regression for the detection of moderating effects of dichotomous variables. *Journal of Management*, 20: 167-178.

Thomdike, R. L. 1949. *Personnel Selection*. New York: Wiley.

Wolf, F. M. 1986. *Meta-analysis: Quantitative methods for research synthesis*. Newbury Park, CA: Sage.

Zedeck, S. 1971. Problems with the use of "moderator" variables. *Psychological Bulletin*. 76: 295-310.