WEIGHTING BY INVERSE VARIANCE OR BY SAMPLE SIZE IN META-ANALYSIS: A SIMULATION STUDY

JULIO SÁNCHEZ-MECA AND FULGENCIO MARÍN-MARTÍNEZ
University of Murcia

In meta-analysis, a weighted average effect size is usually obtained to summarize the global magnitude through a set of primary studies. The optimal weight to obtain the unbiased and minimum variance estimator is the inverse variance of each effect-size estimate. In practice, it is not possible to compute the optimal inverse variance because the population effect size is unknown. Hedges and Olkin and Hunter and Schmidt proposed two alternative estimators of optimal weights. In this article, the bias and relative efficiency of both estimators are assessed via Monte Carlo simulation, defining the standardized mean difference as the effect-size index. The number of studies, sample size, magnitude of population effect size, and discrepancy between two population effect sizes were manipulated. Hedges and Olkin's estimator was more efficient, although more biased, than Hunter and Schmidt's estimator. The consequences of applying both alternatives in meta-analyses are discussed.

Meta-analysis aims to quantitatively integrate the results of several studies on a research topic (Cooper & Hedges, 1994; Glass, McGaw, & Smith, 1981). To this aim, an effect-size index (Kirk, 1996) is chosen as the standard measure to express the results, enabling them to be compared between the studies. To obtain a global index of the magnitude of the relation, the effect sizes are averaged, their homogeneity is tested, and if homogeneity is not assumed, possible variables or characteristics influencing the heterogeneity of effect sizes are explored (Bangert-Drowns, 1986; Hedges & Olkin, 1985; Hunter & Schmidt, 1990; Johnson, Mullen, & Salas, 1995; Rosenthal, 1991; Sánchez-Meca and Marín-Martínez, in press).

We would like to thank two anonymous reviewers for their helpful suggestions. We would like to thank two anonymous reviewers and the editor of the journal for their helpful suggestions.
Correspondance concerning this article should be addresssed to Julio Sánchez-Meca, Dpto. Psicología Básica y Metodología, Facultad de Psicología, Apdo. 4021, Universidad de Murcia, 30100 Murcia, Spain, e-mail jsmeca@fcu.um.es

Educational and Psychological Measurement, Vol. 58 No. 2, April 1998 211-220

One of the most widely used effect-size indexes in meta-analysis is the standardized mean difference, d, defined as the difference between two group means (usually experimental vs. control) divided by the within-group standard deviation. The use of this index is indicated especially when the studies to be integrated are experimental or quasi-experimental. Hedges and Olkin (1985) have shown that the best procedure to average a set of independent ds is a weighted average, with the inverse variance of each d as the optimal weight factor.

However, the variance of d depends on the population standardized mean difference, δ, a parameter unknown in practice. Therefore, an estimate of the optimal weight is required. Hedges and Olkin proposed to substitute the sample standardized mean difference, d_i, for the δ parameter in each single study. On the other hand, Hunter and Schmidt (1990) proposed a simpler procedure consisting of weighting by the sample size of each study. Hunter and Schmidt argued that their procedure is less biased than Hedges and Olkin's (1985), even in the case of nonhomogeneous effect sizes.

The statistical properties of these estimators have not been compared. In the present study, a Monte Carlo simulation is carried out to assess the bias and efficiency of the two estimators for averaging independent ds in conditions similar to those of real meta-analyses.

Two Procedures for Averaging *ds*

Conceptually, we assume that a set of k independent studies estimates the same population effect size. Let us also assume that for each study, a standardized mean difference is obtained by

$$g_i = \frac{\overline{y}_i^E - \overline{y}_i^C}{S_i}, \tag{1}$$

where g_i is a positively biased estimator of the population standardized mean difference, δ; $\overline{y}_i^E$ and $\overline{y}_i^C$ are the experimental and control group means, respectively, of the ith study; and S_i is the ith within-group standard deviation computed through

$$S_i = \sqrt{\frac{(n_i^E - 1)(s_i^E)^2 + (n_i^C - 1)(s_i^C)^2}{n_i^E + n_i^C - 2}}, \tag{2}$$

n_i^E and n_i^C being the sample sizes and $(s_i^E)^2$ and $(s_i^C)^2$ the unbiased variances of the two groups in the ith study.

To remove the bias in the g_i index, the accurate correction factor includes the gamma function (see Hedges, 1981, p. 111), although Hedges derived a satisfactory approximation for most applications, $c(m)_i$, enabling the computation of an approximately unbiased d_i index, which is even more efficient than the g_i index:

$$d_i = c(m)_i g_i, \tag{3}$$

where

$$c(m)_i + 1 - \frac{3}{4m_i - 9}, \tag{4}$$

being $m_i = n_i^E + n_i^C$.

The variance of d_i is given by

$$\sigma_{d_i}^2 = \frac{n_i^E + n_i^C}{n_i^E n_i^C} + \frac{\delta^2}{2\,(n_i^E + n_i^C)}. \tag{5}$$

The accurate variance includes the factor $(N_i - 1) / (N - 3)$, with $N_i = n_i^E = n_i^C$. That is,

$$\sigma_{d_i}^2 = \frac{N_i - 1}{N_i - 3}\left[\frac{n_i^E + n_i^C}{n_i^E n_i^C} + \frac{\delta^2}{2\,(n_i^E + n_i^C)}\right]. \tag{6}$$

Hedges and Olkin (1985) and Hunter and Schmidt (1990) obviate this factor because its magnitude is generally very close to 1.00. Only with low sample sizes will this omission slightly affect the variance (Hunter & Schmidt, 1990).

The optimal weights, w_i, to average a set of independent ds are given by the inverse variance of each d_i (Hedges & Olkin, 1985, p. 110):

$$w_i = \left[\frac{n_i^E + n_i^C}{n_i^E n_i^C} + \frac{\delta^2}{2\,(n_i^E + n_i^C)}\right]^{-1}. \tag{7}$$

As shown in Equation 7, the optimal weights depend on the sample sizes and the unknown population effect size, δ. Hedges and Olkin proposed to substitute the sample d_i for δ in Equation 7:

$$w_i^{HO} = \left[\frac{n_i^E + n_i^C}{n_i^E n_i^C} + \frac{d_i^2}{2\,(n_i^E + n_i^C)}\right]^{-1}, \tag{8}$$

where w_i^{HO} is the weight factor proposed in Hedges and Olkin's procedure. Therefore, they compute the weighted average, $\overline{d}_{HO}$, through

$$\overline{d}_{HO} = \frac{\sum w_i^{HO} d_i}{\sum w_i^{HO}}. \tag{9}$$

Hunter and Schmidt (1990) contended that it is more accurate to substitute the average d weighted by sample size for δ in Equation 7, than the individual ds. Because the resulting weights reduce to sample sizes (see appendix), the average standardized mean difference, $\bar{d}_{HS}$, proposed by Hunter and Schmidt is given by

$$\bar{d}_{HS} = \frac{\sum N_i d_i}{\sum N_i}, \quad (10)$$

where $N_i = n_i^E + n_i^C$ is the sample size of the ith study. Although the equations presented assume a set of homogeneous effect sizes, they are also applicable to effect sizes obtained from different populations.

Method

The simulation study was programmed in GAUSS (1992). Two normally distributed populations with homogeneous variances were defined, $[N(\mu^E, \sigma^2), N(\mu^C, \sigma^2)]$, where μ^E and μ^C are the experimental and control population means, respectively, and σ^2 is the common population variance. From these populations, pairs of independent random samples were generated with n^E and n^C as sample sizes. The simulated studies were accomplished with an experimental and a control group. The parameter effect size, δ, was defined as (Hedges & Olkin, 1985, p. 76):

$$\delta = \frac{\mu^E - \mu^C}{\sigma}. \quad (11)$$

Each pair of generated samples simulated the data in a primary research, where the d index (Equation 3) was computed. A total of k studies simulating the data of a meta-analysis were generated. The studies within the same meta-analysis could estimate a single population effect size (homogeneous case) or more than one population effect size (heterogeneous case). To simulate the heterogeneous case, the simplest situation has been assumed: a dichotomous moderator variable with two different parameter effect sizes for half of the studies in each simulation.

The following parameters were manipulated: (a) the sample size of each study, $N = n^E + n^C$ (with $n^E = n^C$), 30, 50, 80, and 100 being the mean sample sizes in the k studies; (b) the number of studies, k, with values 10, 20, 40, and 100; (c) following Cohen (1988), the value of the parameter effect sizes in the homogeneous case meta-analyses was manipulated with values of δ = 0.2, 0.5, and 0.8; (d) to study the heterogeneous case, half of the k studies were generated so that the parameter effect size was δ_1, and the other half shared the parameter effect size δ_2. The discrepancy between δ_1/δ_2 was manipulated across the following conditions: 0.5/0.4, 0.6/0.5, 0.5/0.2, 0.8/0.5, 0.5/0.0, and 1/0.5.

To simulate the sample sizes, Ns, of the k studies in a meta-analysis, some properties of the sample-sizes distribution in 30 real meta-analyses in different areas of educational and behavioral sciences were assessed[1]. The Pearson skewness index of the sample-sizes distribution was computed through all the meta-analyses, obtaining a value of +1.464. In accord with this value, four vectors of five Ns each were selected: [12, 16, 18, 20, 84], [32, 36, 38, 40, 104], [62, 66, 68, 70, 134], and [82, 86, 88, 90, 154], all with the skewness of +1.464 and averaging 30, 50, 80, and 100, respectively. Each vector was replicated 2, 4, 8, and 20 times, to get meta-analyses of 10, 20, 40, and 100 studies, respectively.

For each one of the 4 (sample size) × 4 (number of studies) × 9 (parameter effect sizes) = 144 conditions defined, 1,000 replications were generated. Thus, 144,000 meta-analyses were simulated. The averages $\bar{d}_{HO}$ (Equation 9) and $\bar{d}_{HS}$ (Equation 10) were computed in each one of these replications (or meta-analyses).

In the homogeneous case, the bias of each of the two procedures, $\bar{d}_{HO}$ and $\bar{d}_{HS}$, was assessed through the 1,000 replications of the same meta-analysis as the mean difference between the 1,000 observed averages and the constant parameter effect size. In the heterogeneous case, the bias was computed as the mean difference between the 1,000 averages observed and the mean of the two parameter effect sizes. In both the homogeneous and heterogeneous cases, the variability of the procedures was assessed by the standard deviations of $\bar{d}_{HO}$ and $\bar{d}_{HS}$ estimators, respectively, in the 1,000 replications of the same meta-analysis.

Results and Discussion

Tables 1 and 2 show the bias and standard deviations of the two estimators, $\bar{d}_{HO}$ (Hedges and Olkin procedure) and $\bar{d}_{HS}$ (Hunter and Schmidt procedure), through all the manipulated conditions. The bias in $\bar{d}_{HO}$ is negative in all of the conditions, with an average value of −0.0109, whereas the bias in $\bar{d}_{HS}$ is always lower than $\bar{d}_{HO}$, with a practically null average value of 0.0002. In the homogeneous case (Table 1) the bias in $\bar{d}_{HO}$ increases the greater the parameter effect size; however, $\bar{d}_{HS}$ remains approximately constant. In the heterogeneous case (Table 2), $\bar{d}_{HO}$ also increases the bias with the discrepancy between δ_1 and δ_2, whereas $\bar{d}_{HS}$ is scarcely affected. On the other hand, in both the homogeneous and heterogeneous cases, the sample size influences the bias in $\bar{d}_{HO}$: the lower the sample size, the higher the bias; $\bar{d}_{HS}$ is not affected. The number of studies, k, does not seem to influence neither $\bar{d}_{HO}$ nor $\bar{d}_{HS}$.

With respect to the variability of the two estimators, $\bar{d}_{HO}$ is systematically more efficient than $\bar{d}_{HS}$, although the discrepancies are not very pronounced.

Table 1
Bias and Standard Deviations (in parentheses) of Hedges and Olkin (1985) and Hunter and Schmidt (1990) Estimators: Homogeneous Case

		δ = 0.2		δ = 0.5		δ = 0.8	
k	*N*	Hedges & Olkin	Hunter & Schmidt	Hedges & Olkin	Hunter & Schmidt	Hedges & Olkin	Hunter & Schmidt
10	30	–.0093	–.0035	–.0139	.0007	–.0206	.0038
		(.1094)	(.1126)	(.1121)	(.1158)	(.1179)	(.1225)
	50	–.0044	–.0009	–.0093	–.0002	–.0158	–.0016
		(.0867)	(.0883)	(.0896)	(.0914)	(.0904)	(.0924)
	80	–.0043	–.0021	–.0048	.0009	–.0097	–.0006
		(.0687)	(.0695)	(.0697)	(.0706)	(.0737)	(.0745)
	100	–.0015	.0002	–.0042	.0003	–.0081	–.0008
		(.0634)	(.0640)	(.0650)	(.0657)	(.0670)	(.0678)
20	30	–.0078	–.0018	–.0173	–.0016	–.0162	.0091
		(.0790)	(.0817)	(.0791)	(.0814)	(.0812)	(.0845)
	50	–.0039	–.0002	–.0063	.0032	–.0161	–.0010
		(.0636)	(.0649)	(.0639)	(.0654)	(.0654)	(.0664)
	80	–.0025	–.0001	–.0072	–.0013	–.0114	–.0021
		(.0499)	(.0505)	(.0488)	(.0495)	(.0533)	(.0541)
	100	–.0006	.0013	–.0049	–.0003	–.0096	–.0020
		(.0437)	(.0442)	(.0436)	(.0440)	(.0454)	(.0461)
40	30	–.0074	–.0010	–.0154	.0005	–.0272	–.0016
		(.0571)	(.0589)	(.0577)	(.0598)	(.0591)	(.0612)
	50	–.0016	.0023	–.0100	–.0004	–.0159	–.0004
		(.0433)	(.0441)	(.0442)	(.0452)	(.0448)	(.0459)
	80	–.0024	–.0000	–.0084	–.0023	–.0080	.0016
		(.0354)	(.0358)	(.0342)	(.0347)	(.0359)	(.0365)
	100	–.0024	–.0005	–.0067	–.0019	–.0085	–.0007
		(.0324)	(.0327)	(.0309)	(.0311)	(.0337)	(.0341)
100	30	–.0070	–.0005	–.0165	–.0004	–.0252	.0011
		(.0350)	(.0361)	(.0351)	(.0364)	(.0376)	(.0392)
	50	–.0060	–.0021	–.0109	–.0011	–.0149	.0007
		(.0274)	(.0280)	(.0284)	(.0291)	(.0298)	(.0304)
	80	–.0037	–.0013	–.0058	.0004	–.0084	.0014
		(.0222)	(.0224)	(.0228)	(.0231)	(.0237)	(.0240)
	100	–.0008	.0011	–.0053	–.0004	–.0067	.0013
		(.0197)	(.0199)	(.0200)	(.0202)	(.0203)	(.0205)

Note. *k* = number of studies; *N* = sample size; δ = parameter effect size.

Furthermore, these short discrepancies are very similar through all of the conditions, without any interactive effect of interest. In the homogeneous case, the greater the parameter effect size, the greater the variability of the two estimators. In the heterogeneous case, the discrepancy between the two parameter effect sizes scarcely affected the variability of the estimators. As

Table 2
Bias and Standard Deviations (in parentheses) of Hedges and Olkin (1985) and Hunter and Schmidt (1990) Estimators: Heterogeneous Case

		$\delta_1 = 0.5, \delta_2 = 0.4$		$\delta_1 = 0.6, \delta_2 = 0.5$		$\delta_1 = 0.5, \delta_2 = 0.2$	
k	*N*	Hedges & Olkin	Hunter & Schmidt	Hedges & Olkin	Hunter & Schmidt	Hedges & Olkin	Hunter & Schmidt
10	30	–.0127	.0004	–.0196	–.0036	–.0113	.0011
		(.1130)	(.1163)	(.1174)	(.1220)	(.1139)	(.1182)
	50	–.0044	.0042	–.0031	.0073	–.0077	.0005
		(.0884)	(.0901)	(.0912)	(.0933)	(.0899)	(.0922)
	80	–.0051	.0003	–.0053	.0011	–.0035	.0024
		(.0711)	(.0722)	(.0749)	(.0758)	(.0713)	(.0728)
	100	–.0026	.0016	–.0062	–.0010	–.0023	.0029
		(.0627)	(.0632)	(.0673)	(.0679)	(.0618)	(.0628)
20	30	–.0142	.0001	–.0158	.0017	–.0114	.0009
		(.0832)	(.0858)	(.0828)	(.0857)	(.0781)	(.0813)
	50	–.0093	–.0005	–.0134	–.0030	–.0076	.0011
		(.0636)	(.0651)	(.0624)	(.0637)	(.0639)	(.0657)
	80	–.0037	.0020	–.0049	.0021	–.0073	–.0013
		(.0496)	(.0503)	(.0490)	(.0497)	(.0491)	(.0499)
	100	–.0037	.0009	–.0042	.0014	–.0055	–.0003
		(.0455)	(.0460)	(.0465)	(.0469)	(.0456)	(.0463)
40	30	–.0142	.0002	–.0168	.0014	–.0166	–.0037
		(.0582)	(.0604)	(.0574)	(.0596)	(.0578)	(.0600)
	50	–.0100	–.0010	–.0107	.0002	–.0051	.0036
		(.0446)	(.0456)	(.0453)	(.0464)	(.0436)	(.0448)
	80	–.0055	.0002	–.0058	.0012	–.0064	–.0003
		(.0339)	(.0343)	(.0357)	(.0361)	(.0341)	(.0348)
	100	–.0030	.0016	–.0052	.0005	–.0052	.0000
		(.0314)	(.0317)	(.0307)	(.0311)	(.0314)	(.0320)
100	30	–.0137	.0011	–.0179	.0001	–.0147	–.0016
		(.0342)	(.0356)	(.0354)	(.0367)	(.0358)	(.0372)
	50	–.0088	.0003	–.0113	–.0002	–.0089	–.0003
		(.0289)	(.0295)	(.0283)	(.0290)	(.0284)	(.0291)
	80	–.0058	–.0000	–.0072	–.0001	–.0066	–.0005
		(.0228)	(.0231)	(.0231)	(.0234)	(.0222)	(.0226)
	100	–.0050	–.0003	–.0053	.0005	–.0062	–.0009
		(.0197)	(.0199)	(.0200)	(.0202)	(.0197)	(.0200)

(continued)

expected, in both homogeneous and heterogeneous cases, the greater the number of studies, *k*, and sample size, *N*, the lower the variability of the two estimators.

Our results confirm the statement of Hunter and Schmidt (1990) contending the practical absence of bias in the $\bar{d}_{HS}$ estimator, even with heterogeneous

Table 2 Continued

		$\delta_1 = 0.8, \delta_2 = 0.5$		$\delta_1 = 0.5, \delta_2 = 0.0$		$\delta_1 = 1, \delta_2 = 0.5$	
k	N	Hedges & Olkin	Hunter & Schmidt	Hedges & Olkin	Hunter & Schmidt	Hedges & Olkin	Hunter & Schmidt
10	30	–.0148	.0079	–.0139	–.0033	–.0369	–.0044
		(.1147)	(.1196)	(.1121)	(.1179)	(.1176)	(.1229)
	50	–.0144	.0010	–.0109	–.0028	–.0240	.0006
		(.0907)	(.0933)	(.0924)	(.0956)	(.0892)	(.0929)
	80	–.0066	.0040	–.0045	.0022	–.0191	–.0000
		(.0727)	(.0739)	(.0713)	(.0733)	(.0700)	(.0717)
	100	–.0095	–.0003	–.0067	–.0008	–.0142	.0034
		(.0658)	(.0668)	(.0632)	(.0648)	(.0669)	(.0687)
20	30	–.0253	–.0019	–.0125	–.0016	–.0324	.0012
		(.0819)	(.0856)	(.0787)	(.0826)	(.0850)	(.0892)
	50	–.0132	.0023	–.0084	–.0001	–.0212	.0036
		(.0621)	(.0637)	(.0611)	(.0634)	(.0656)	(.0681)
	80	–.0099	.0013	–.0073	–.0007	–.0214	–.0019
		(.0522)	(.0534)	(.0484)	(.0498)	(.0512)	(.0524)
	100	–.0085	.0010	–.0090	–.0031	–.0190	–.0013
		(.0444)	(.0451)	(.0441)	(.0453)	(.0460)	(.0469)
40	30	–.0225	.0011	–.0112	.0004	–.0332	.0007
		(.0576)	(.0602)	(.0556)	(.0585)	(.0587)	(.0622)
	50	–.0157	.0001	–.0078	.0007	–.0255	–.0008
		(.0444)	(.0455)	(.0425)	(.0440)	(.0446)	(.0461)
	80	–.0118	–.0006	–.0064	.0003	–.0198	–.0003
		(.0352)	(.0358)	(.0355)	(.0365)	(.0362)	(.0373)
	100	–.0112	–.0014	–.0057	.0004	–.0201	–.0022
		(.0324)	(.0329)	(.0306)	(.0314)	(.0328)	(.0335)
100	30	–.0232	.0013	–.0118	–.0001	–.0336	.0009
		(.0367)	(.0383)	(.0369)	(.0386)	(.0354)	(.0372)
	50	–.0140	.0020	–.0077	.0009	–.0246	.0002
		(.0286)	(.0294)	(.0273)	(.0283)	(.0296)	(.0308)
	80	–.0110	.0003	–.0055	.0012	–.0197	.0000
		(.0230)	(.0234)	(.0212)	(.0218)	(.0234)	(.0241)
	100	–.0096	.0001	–.0075	–.0014	–.0181	–.0001
		(.0202)	(.0205)	(.0201)	(.0206)	(.0208)	(.0214)

Note. k = number of studies; N = sample size; δ_1 and δ_2 = parameter effect sizes.

effect sizes. The same is not true of the $\bar{d}_{HO}$ estimator, which shows a slight negative bias, more pronounced with small sample sizes, large population effect sizes, and high heterogeneity among population effect sizes. The bias in $\bar{d}_{HO}$ is produced substituting d_i for δ in estimating the inverse variance of each d_i (Equation 8); thus, other factors equal, there is a negative relationship

between effect size, d_i, and the estimated weighting factor, w_i^{HO}. Although the negative bias in the $\bar{d}_{HO}$ estimator does not seem to be of practical importance, it must be applied with caution when sample sizes in a meta-analysis are about 30 or less, population effect size is more than 0.8, and the influence of some moderator variable is clear. In these circumstances, we advise computing both $\bar{d}_{HO}$ and $\bar{d}_{HS}$ estimates to compare possible discrepancies. If the differences are relevant, $\bar{d}_{HS}$ will be the estimator of election.

Hunter and Schmidt (1990) do not mention, however, the variability of the two estimators. Our findings suggest that $\bar{d}_{HO}$ is systematically more efficient than $\bar{d}_{HS}$. Nevertheless, averaging all the manipulated conditions, $\bar{d}_{HO}$ is only 2.8% more efficient than $\bar{d}_{HS}$. Although the differences in efficiency between $\bar{d}_{HS}$ and $\bar{d}_{HO}$ are small, the lower variability in $\bar{d}_{HO}$ will provide more accurate estimates of population effect size than those of $\bar{d}_{HS}$.

Finally, we must note that in our simulation study, we have assumed null correlation between sample size and effect size. It is not unusual to find real meta-analyses with negative correlations between sample size and effect size (Light, Singer, & Willett, 1994). In this case, the possible differences between the two estimators of population effect size can be greater than those evidenced in the present report.

APPENDIX

Hunter and Schmidt (1990) proposed to express the variance of d index given in Equation 5 as a function of N, being that $N = n^E + n^C$. Supposing that $n^E = n^C$, Equation 5 is equivalent to

$$\sigma^2_{d_i} = \frac{4}{N_i}\left(1 + \frac{\delta^2}{8}\right). \quad (12)$$

To estimate the variance of d_i, Hunter and Schmidt (1990) proposed to substitute $\bar{d}_{HS}$, defined in Equation 10, for the unknown δ parameter. So, the inverse variance of each d_i is given by

$$w_i^{HS} = \frac{N_i}{4}\left(1 + \frac{\bar{d}^2_{HS}}{8}\right)^{-1}. \quad (13)$$

To simplify, let

$$a_i = \frac{1}{4}\left(1 + \frac{\bar{d}^2_{HS}}{8}\right)^{-1}, \quad (14)$$

where a_i is a constant in the k studies, then $a_i = a$. Substituting Equation 14 for Equation 13:

$$w_i^{HS} = N_i a. \quad (15)$$

The $\bar{d}_{HS}$ estimator is given by

$$\bar{d}_{HS} = \frac{\sum w_i^{HS} d_i}{\sum w_i^{HS}}. \tag{16}$$

Finally, substituting Equation 15 for Equation 16, we obtain the Hunter and Schmidt (1990) estimator defined in Equation 10:

$$\bar{d}_{HS} = \frac{\sum N_i\, a\, d_i}{\sum N_i\, a} = \frac{\sum N_i d_i}{\sum N_i}. \tag{17}$$

Thus, following Hunter and Schmidt (1990), substituting $\bar{d}_{HS}$ for δ in Equation 7 to estimate the optimal weights implies simply weighting each sample-effect size, d_i, by its sample size, N_i.

Note

1. The meta-analyses were published in 18 journals (e.g., *Clinical Psychology Review*, *Journal of Applied Psychology*, *Journal of Consulting and Clinical Psychology*, *Journal of Educational Psychology*, *Journal of Personality and Social Psychology*, and *Psychological Bulletin*). The list of the references is available from the authors.

References

Bangert-Drowns, R. L. (1986). Review of developments in meta-analytic method. *Psychological Bulletin, 99*, 388-399.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

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

GAUSS. (1992). *The GAUSS system* (Version 3.0). Maple Valley, WA: Aptech Systems.

Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research*. Beverly Hills, CA: Sage.

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics, 6*, 107-128.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. New York: Academic Press.

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

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

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

Light, R. J., Singer, J. D., & Willett, J. B. (1994). The visual presentation and interpretation of meta-analyses. In H. M. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 439-453). New York: Russell Sage.

Rosenthal, R. (1991). *Meta-analytic procedures for social research* (Rev. ed.). Newbury Park, CA: Sage.

Sánchez-Meca, J., & Marín Martínez, F. (in press) Testing continuous moderators in meta-analysis: A comparison of procedures. *British Journal of Mathmatical and Statistical Psychology*