A Procedure for Combining Sample Correlation Coefficients and Vote Counts to Obtain an Estimate and a Confidence Interval for the Population Correlation Coefficient

Brad J. Bushman
Iowa State University

Morgan C. Wang
University of Central Florida

Missing effect-size estimates pose a particularly difficult problem in meta-analysis. Rather than discarding studies with missing effect-size estimates or setting missing effect-size estimates equal to 0, the meta-analyst can supplement effect-size procedures with vote-counting procedures if the studies report the direction of results or the statistical significance of results. By combining effect-size and vote-counting procedures, the meta-analyst can obtain a less biased estimate of the population effect size and a narrower confidence interval for the population effect size. This article describes 3 vote-counting procedures for estimating the population correlation coefficient in studies with missing sample correlations. Easy-to-use tables, based on equal sample sizes, are presented for the 3 procedures. More complicated vote-counting procedures also are given for unequal sample sizes.

Science is built up with fact, as a house is with stone. But a collection of facts is no more a science than a heap of stones is a house. (Jules Henri Poincare, cited in Olkin, 1990, p. 3)

As the number of primary research studies continues to grow at an exponential rate, it becomes increasingly important to organize and integrate the facts or data from these studies. The task for the reviewer is to integrate the results obtained from a collection of studies that investigate the same phenomenon. There are two general approaches to accomplishing this task: the narrative (or qualitative) approach and the meta-analytic (or quantitative) approach. In the traditional narrative review, the reviewer uses "mental algebra" to integrate the findings from a collection of studies and describes the results in a narrative manner. In the meta-analytic review, the reviewer uses statistical procedures to integrate the findings from a collection of studies and describes the results using numerical effect-size estimates. Traditional narrative reviews are more likely than meta-analytic reviews to depend on the subjective judgments, preferences, and biases of the reviewer (e.g., Cooper & Rosenthal, 1980).

Data are the reviewer's most precious commodity. Unfortunately, missing data represent one of the biggest problems in meta-analysis (Pigott, 1994). Missing effect-size estimates pose a particularly difficult problem in meta-analysis. Often studies do not include enough information to permit the calculation of an effect-size estimate. The two most common "solutions" to the problem of missing effect-size estimates are (a) to set the missing estimates equal to zero and (b) to omit from the review those studies with missing estimates. Both approaches have serious problems because they produce biased results. Setting missing effect-size estimates equal to zero is an overly conservative practice that underestimates the magnitude of the population effect size. Imputing a single value for missing effect-size estimates (e.g., zero or the mean) also produces a sample of estimates that is less variable than what would be expected as a result of sampling error, even if the common population effect size is the same for all studies. The variance of effect-size estimates is used in several meta-analytic procedures (e.g., weighted mean estimates and homogeneity tests).

The omission of studies with missing effect-size estimates is problematic because the proportion of studies omitted is often quite large. For example, we located 12 articles published in *Psychological Bulletin* between the years 1991 and 1993 that used meta-analytic procedures to combine correlation coefficients. In these articles, the proportions of studies with missing correlations ranged from .13 to .49.

Unfortunately, effect-size estimates are likely to be missing for reasons related to the data. Studies with significant results are more likely to report effect-size estimates than are studies with nonsignificant results. The studies that contain effect-size

---

Brad J. Bushman, Department of Psychology, Iowa State University; Morgan C. Wang, Department of Statistics, University of Central Florida.

We would like to thank Larry Hedges for his helpful comments on this article and Kathy Shelley for her assistance with Figure 1. To obtain a free copy of a computer program that can be used to implement all of the procedures described in this article, send a blank floppy disk to Morgan C. Wang, Department of Statistics, University of Central Florida, Orlando, Florida 32828.

Correspondence concerning this article should be addressed to Brad J. Bushman, Department of Psychology, Iowa State University, Ames, Iowa 50011-3180. Electronic mail may be sent via Bitnet to sl.bjb@isumvs or via Internet to s1.bjb@isumvs.iastate.edu.

Psychological Bulletin, 1995, Vol. 117, No. 3, 530–546

estimates may therefore represent a biased subset of the population of studies. Omitting studies with missing effect-size estimates limits the generalizability of the results of a research synthesis.

Rather than discarding studies that do not report enough information to permit the calculation of an effect-size estimate, the meta-analyst can use vote-counting procedures if the studies report the direction of results or the statistical significance of results (Bushman, 1994). Hedges and Olkin (1980) proposed three vote-counting procedures for obtaining an estimate and a confidence interval for the population standardized mean difference. They noted that

> the logic of the three vote-counting procedures can easily be applied to other models. One such case is that of counts of sample correlation coefficients. The main difference in models for correlations is that the correlation coefficient is not normally distributed. (Hedges & Olkin, 1980, p. 367)

In this article, we extend the logic of the vote-counting procedures proposed by Hedges and Olkin (1980) for the standardized mean difference to the correlation coefficient. First, we describe three vote-counting procedures for obtaining an estimate and a confidence interval for the population correlation coefficient, ρ. Easy-to-use tables, based on equal sample sizes, are presented for each type of vote-counting procedure. More complicated vote-counting procedures, based on unequal sample sizes, also are described. Second, we describe effect-size procedures for obtaining an estimate and a confidence interval for ρ. Third, we describe how information from effect-size analyses and vote-counting analyses can be combined to obtain a less biased estimate of ρ and a narrower confidence interval for ρ. We have written a computer program that can be used to implement all of the procedures described in this article. (The computer program is available on request from Morgan C. Wang in the languages of FORTRAN and SAS.)

Level of Significance

Some primary research reports do not contain enough information to calculate effect-size estimates; they may, however, provide information about the magnitude of the treatment effect. Often the information is in the form of a report of the decision yielded by the significance test (e.g., a significant positive correlation) or in the form of a direction of the effect without regard to its statistical significance (e.g., a positive correlation). The first of these corresponds to whether the test statistic exceeds a conventional critical value at a given significance level, such as $\alpha = .05$. The second corresponds to whether the test statistic exceeds the rather unconventional critical value at significance level $\alpha = .5$. For the vote-counting procedures described in this article, we consider both the $\alpha = .5$ and $\alpha = .05$ significance levels.

Estimates and Confidence Intervals Based on Equal Sample Sizes

The vote-counting procedures described in this section assume that each study in a collection of k independent studies has an identical sample size n. This assumption is extremely restrictive because studies frequently have different sample sizes. If the sample sizes are not very different, Hedges and Olkin (1980) recommended treating the studies as if all had the same sample size equal to some "average value." There are several "average values" of the sample size from which to choose, including the arithmetic mean, median, mode, minimum, and maximum values. Hedges and Olkin (1985) offered a general class of means

$$\bar{n}_m = \left(\frac{1}{k}\sum_{i=1}^{k} n_i^{1/m}\right)^m, \quad m \in [0, \infty]. \tag{1}$$

In Equation 1, where m converges to 0, one has the maximum value; when m converges to ∞, one has the minimum value; and, when $m = 1$, one has the arithmetic mean. Gibbons, Olkin, and Sobel (1977) recommended using the square mean root (SMR) for unequal sample sizes, which results when $m = 2$ in Equation 1:

$$\bar{n}_2 = n_{SMR} = \left(\frac{\sqrt{n_1} + \cdots + \sqrt{n_k}}{k}\right)^2. \tag{2}$$

The square mean root is not as influenced by extreme values as the arithmetic mean. The square mean root is more conservative than the arithmetic mean and is less conservative than the minimum value. If the sample sizes of the studies differ substantially, it may not be reasonable to use an average value like the square mean root. In this case, the methods for unequal sample sizes described in the next section should be used.

Vote-counting procedures seem to be quite robust against violations of the equal sample size assumption. Although a proof for the robustness of vote-counting procedures is not available at this time, the illustrative example used in this article grossly violates the equal sample size assumption (sample sizes range from $n = 60$ to $n = 1127$), yet the results obtained from equal and unequal sample size vote-counting procedures are very similar. We are in the process of conducting a simulation study to determine how robust vote-counting procedures are to violations of the equal sample size assumption.

Often a meta-analyst is interested in determining whether a relation exists between the independent variable, X, and the dependent variable, Y, for each study in a collection of k independent studies.[1] That is, the meta-analyst wants to determine whether the effect size is zero for each study. This situation can be stated in terms of null and alternative hypotheses. In the general case, let $T_1, \ldots, T_k$ be independent estimators from k studies of parameters $\theta_1, \ldots, \theta_k$, respectively. If one makes the underlying assumption that the effect sizes are homogeneous, the appropriate null and alternative hypotheses are

$$H_0: \theta_1 = \cdots = \theta_k = \theta = 0$$
$$H_A: \theta_1 = \cdots = \theta_k = \theta > 0. \tag{3}$$

The assumption of homogeneity of effect sizes underlies the va-

[1] We use the term *independent* to refer to manipulated and nonmanipulated variables. Although this usage is not technically correct (i.e., independent variables are manipulated, and participant variables are measured), it simplifies the discussion considerably.

lidity of the meta-analytic procedures described here, and the practicing meta-analyst should test this assumption. Formal statistics for testing the homogeneity assumption have been described in detail elsewhere (e.g., Hedges & Olkin, 1985; Hunter & Schmidt, 1990).

The null hypothesis in Display 3 is rejected if the estimator T of the common effect size θ exceeds the one-sided critical value $C_{\alpha,n}$ from the distribution of T at significance level α.[2] The lower and upper bounds $[\hat{\theta}_L, \hat{\theta}_U]$ of a $100(1 - \alpha)\%$ confidence interval for θ are given by

$$\hat{\theta}_L = T - C_{\alpha/2,n}S_T \leq \theta \leq T + C_{\alpha/2,n}S_T = \hat{\theta}_U, \qquad (4)$$

where $C_{\alpha/2,n}$ is the two-sided critical value from the distribution of T at significance level α and S_T is the standard error of T.

The problem is to obtain an estimate and a confidence interval for θ using the k estimators $T_1, \ldots, T_k$. The essential feature of vote-counting procedures is that the values of $T_1, \ldots, T_k$ are not observed. Even if one does not observe T_i, one can still estimate θ by counting the number of times T_i exceeded the one-sided critical value $C_{\alpha,n}$. If the significance level is $\alpha = .5$, count the number of positive results (i.e., results in the predicted direction). If the treatment has no effect, the proportion of positive results in the population is .5. If the treatment has an effect, the proportion of positive results in the population should be greater than .5. Thus, the appropriate null and alternative hypotheses are

$$H_0: p = .5$$
$$H_A: p > .5, \qquad (5)$$

where p is the proportion of positive results in the population.

If the significance level is $\alpha = .05$, count the number of significant positive results (i.e., results that are statistically significant and in the predicted direction). If the treatment has no effect, the proportion of significant positive results in the population is .05. If the treatment has an effect, the proportion of significant positive results in the population should be greater than .05. Thus, the appropriate null and alternative hypotheses are

$$H_0: p = .05$$
$$H_A: p > .05. \qquad (6)$$

To test the hypotheses given in Equations 5 and 6, one needs an estimator of p. The method of maximum likelihood can be used to obtain an estimator of p. For each study, there are two possible outcomes: a "success" if $T_i > C_{\alpha,n}$, or a "failure" if $T_i \leq C_{\alpha,n}$. If $\alpha = .5$, a success is defined as a positive result and a failure is defined as a negative or null result. If $\alpha = .05$, a success is defined as a significant positive result and a failure is defined as a nonsignificant positive result, a negative result, or a null result. It is convenient to define an indicator variable W_i that takes on the value 1 if the outcome is a success or the value 0 if the outcome is a failure. That is,

$$W_i = \begin{cases} 1 & \text{if } T_i > C_{\alpha,n} \\ 0 & \text{if } T_i \leq C_{\alpha,n} \end{cases}. \qquad (7)$$

The probability of a success is given by

$$\Pr(W_i = 1) = \Pr(T_i > C_{\alpha,n}) = p, \qquad (8)$$

and the probability of a failure is given by

$$\Pr(W_i = 0) = \Pr(T_i \leq C_{\alpha,n}) = 1 - p. \qquad (9)$$

Each W_i has a Bernoulli distribution with parameter p. The maximum likelihood estimator of p is

$$\hat{p} = \sum_{i=1}^{k} w_i/k = U/k, \qquad (10)$$

where w_i is the observed value of W_i (i.e., 1 or 0). Thus, $\hat{p}$ is the proportion of positive or significant positive results obtained from the k studies.

One also can construct a confidence interval for p. It is well known that if the random indicator variables $W_1, \ldots, W_k$ form k Bernoulli trials with parameter p, and if $W = W_1 + \cdots + W_k$, then W has a binomial distribution with parameters k and p (e.g., DeGroot, 1986). Clopper and Pearson (1934) and Hedges and Olkin (1985) have published nomographs for obtaining exact confidence intervals for binomial proportions based on the sample size n and the proportion $\hat{p}$. Figure 1 depicts a nomograph for obtaining exact 95% confidence intervals for binomial proportions.

Example 1. Suppose that a reviewer has a sample of 10 studies with common sample size $n = 50$ and that 7 of the 10 studies found positive correlations. Thus, $\hat{p} = U/k = 7/10 = .70$. Figure 1 can be used to find an exact 95% confidence interval for p. Find the value of $\hat{p} = .70$ on the abscissa, and then move vertically until the bands with the appropriate common sample size $n = 50$ are intersected. At the intersected points, move horizontally to the ordinate to obtain the lower and upper bounds $[\hat{p}_L, \hat{p}_U]$ of the p confidence interval. The exact 95% confidence interval for p is [.575, .821].

Methods also exist for obtaining confidence intervals for p based on large-sample normal approximations to the binomial distribution. The normal distribution provides a good approximation to the binomial distribution if $kp \geq 5$ and if $k(1 - p) \geq 5$. When both of these conditions are met, $\hat{p}$ is approximately normally distributed with mean p and variance $p(1 - p)/k$. That is,

$$z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{k}}} \qquad (11)$$

has a standard normal distribution.

Using $\hat{p}(1 - \hat{p})/k$ to estimate the variance $p(1 - p)/k$, one obtains the lower and upper bounds $[\hat{p}_L, \hat{p}_U]$ of a $100(1 - \alpha)\%$ confidence interval for p,

[2] The alternative hypothesis in Display 3 assumes that the researcher has a priori predictions about the direction of the treatment effect and is therefore using a one-sided test. For a two-sided test, let $\hat{\theta}_i^2$ be an estimator of θ_i^2. In general, it is not possible to use a two-sided test statistic to estimate θ if θ can be either positive or negative (see Hedges & Olkin, 1985, p. 52).

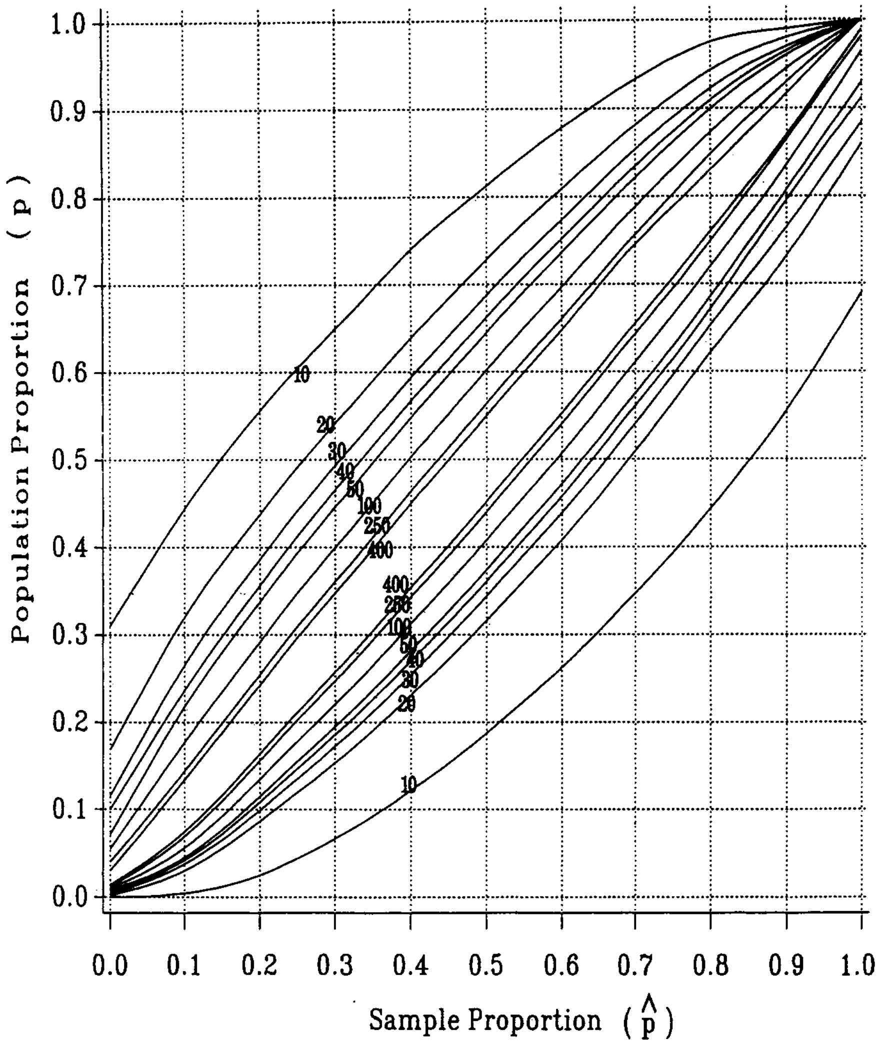

Figure 1. Exact 95% confidence intervals of the parameter *p* of the binomial distribution. The numbers on the curves indicate sample size.

$$\hat{p}_L = \hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{k}} \le p \le \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{k}} = \hat{p}_U, \quad (12)$$

where $z_{\alpha/2}$ is the two-sided critical value of the standard normal distribution. The critical value for a 95% confidence interval is $z_{.025} = 1.96$.

It is well known that the square of a standard normal variate has a chi-square distribution with 1 degree of freedom (e.g., DeGroot, 1986). If Equation 11 is squared, one has

$$z^2 = \left[\frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{k}}}\right]^2 = \frac{k(\hat{p}-p)^2}{p(1-p)}. \quad (13)$$

Consequently, z^2 has a chi-square distribution with 1 degree of freedom. Equation 13 provides an alternative method for obtaining a confidence interval for p. The lower and upper bounds $[\hat{p}_L, \hat{p}_U]$ of a $100(1-\alpha)\%$ confidence interval for p are obtained by solving the quadratic equation

$$\frac{(2\hat{p} + b) \pm \sqrt{b^2 + 4b\hat{p}(1-\hat{p})}}{2(1+b)}, \quad (14)$$

where $b = \chi^2(1)/k$ (see Hedges & Olkin, 1985, p. 56). For a 95% confidence interval, the upper critical value of the chi-square distribution with 1 degree of freedom is $\chi^2_{.05}(1) = 3.841$.

Therefore, two methods are available for constructing confidence intervals for the proportion of positive or significant positive results. Both methods are based on large-sample normal approximations to the binomial distribution. The probability that either confidence interval contains the estimate of the true parameter p is the same for both methods (e.g., .95). Thus, the reviewer should use the method that gives the narrower confidence interval. In general, the method based on the chi-square distribution produces narrower confidence intervals than does the method based on the normal distribution (Hedges & Olkin, 1980). The two methods give similar results when k is large.

Once a confidence interval $[\hat{p}_L, \hat{p}_U]$ has been calculated for p, it can be transformed to a confidence interval $[\hat{\theta}_L, \hat{\theta}_U]$ for θ by finding the value of θ such that $\Pr(T > C_{\alpha,n}) = p$. If T is approximately normally distributed with mean θ and variance S_θ^2, then the probability p can be expressed as an explicit function of θ,

$$p = \Pr(T > C_{\alpha,n}) = \Pr\{(T - \theta)/S_\theta > (C_{\alpha,n} - \theta)/S_\theta\} = 1 - \Phi[(C_{\alpha,n} - \theta)/S_\theta], \quad (15)$$

where $\Phi(.)$ is the standard normal cumulative distribution function. Solving Equation 15 for θ, one obtains

$$\theta = C_{\alpha,n} - S_\theta \Phi^{-1}(1 - p). \quad (16)$$

Equation 16 provides a general large-sample approximate relation between a proportion and an effect size. It is applicable to effect-size indexes that are normally distributed in large samples. More accurate (but more complicated) results can be obtained for specific effect-size indexes by using their exact distributions. This article includes three tables based on the exact distribution of the correlation coefficient.

The population coefficient of correlation between X and Y for the ith study is defined as

$$\rho_i = \frac{\mathrm{Cov}(X_i, Y_i)}{\sigma_{X_i}\sigma_{Y_i}}, \quad i = 1, \ldots, k, \quad (17)$$

where $\mathrm{Cov}(X_i, Y_i)$ is the population covariance of X and Y for the ith study, and σ_{X_i} and σ_{Y_i} are the respective population standard deviations of X and Y for the ith study. The sample estimator of ρ_i is the Pearson product–moment correlation coefficient

$$r_i = \frac{\sum_{j=1}^{n_i} (X_j - \bar{X})(Y_j - \bar{Y})}{\sqrt{\left[\sum_{j=1}^{n_i} (X_j - \bar{X})^2\right]\left[\sum_{j=1}^{n_i} (Y_j - \bar{Y})^2\right]}}, \quad i = 1, \ldots, k, \quad (18)$$

where $\bar{X}$ and $\bar{Y}$ are the respective mean values for the independent and dependent variables.

If one assumes that the correlations are homogeneous, then the hypotheses in Display 3 can be written as

$$H_0: \rho_1 = \cdots = \rho_k = \rho = 0$$
$$H_A: \rho_1 = \cdots = \rho_k = \rho > 0. \quad (19)$$

To estimate ρ, count the number of times r_i exceeds the critical value $r_{\alpha,n}$ from the r distribution at significance level α for common sample size n. For $\alpha = .5$, count the number of studies with positive correlations; for $\alpha = .05$, count the number of studies with significant positive correlations.

Estimates and Confidence Intervals Based on the Proportion of Positive Correlations

Olkin (1973) produced a table for obtaining an estimate and a confidence interval for ρ from the proportion of positive correlations (i.e., $\alpha = .5$). Olkin's table also was included in the book by Hedges and Olkin (1985, pp. 64–65). Several of the values in Olkin's table, however, have absolute errors greater than .01. For this article, we used self-validating numerical methods to ensure that each entry in Table 1 had an absolute error of less than .001 (see Appendix A). Note that Table 1 contains only proportions greater than or equal to .500. Proportions less than .500, which correspond to negative values of ρ, can be obtained by using $1 - \hat{p}$ instead of $\hat{p}$.

Example 2. Suppose that one wishes to obtain an estimate and a confidence interval for ρ from the proportion of positive results in Example 1. Recall that $n = 50$ and $\hat{p} = .70$ in Example 1. In Table 1 for $n = 50$, it can be seen that $\hat{p} = .70$ falls between .687 and .712. The corresponding values of ρ are .07 and .08, respectively. The estimate of ρ must therefore fall between .07 and .08. One can estimate ρ by linear interpolation using the two-point equation of the line. The equation of the line passing through the points (x_1, y_1) and (x_2, y_2) is given by

$$(y - y_1) = \left(\frac{y_2 - y_1}{x_2 - x_1}\right)(x - x_1), \quad (20)$$

where $x_1 \neq x_2$. Using the two-point equation of the line (Equation 20), where the first point is given by $(p_1, \rho_1) = (.687, .07)$ and the second point is given by $(p_2, \rho_2) = (.712, .08)$, one has

$$(\hat{\rho} - \rho_1) = \left(\frac{\rho_2 - \rho_1}{p_2 - p_1}\right)(\hat{p} - p_1)$$
$$= (\hat{\rho} - .07) = \left(\frac{.08 - .07}{.712 - .687}\right)(.70 - .687).$$

Solving for $\hat{\rho}$, one obtains the estimate .075 of ρ.

Table 1 also can be used to convert the p confidence interval to a ρ confidence interval. Recall that the exact 95% p confidence interval $[\hat{p}_L, \hat{p}_U]$ in Example 1 was [.575, .821]. First one obtains the lower bound $\hat{p}_L$. In Table 1 for $n = 50$, it can be seen that $\hat{p}_L = .575$ falls between .555 and .583. The corresponding values of ρ are .02 and .03, respectively. Using the two-point equation of the line (Equation 20), one has

$$(\hat{\rho}_L - .02) = \left(\frac{.03 - .02}{.583 - .555}\right)(.575 - .555).$$

Solving for $\hat{\rho}_L$, one obtains .026. In Table 1 for $n = 50$, it can be seen that $\hat{p}_U = .821$ falls between .757 and .920. The corresponding values of ρ are .10 and .20, respectively. Using the two-point equation of the line (Equation 20), one has

$$(\hat{\rho}_U - .10) = \left(\frac{.20 - .10}{.920 - .757}\right)(.821 - .757).$$

Solving for $\hat{\rho}_U$, one obtains .139. Because the confidence in-

Table 1
Probability That a Correlation Coefficient From a Sample of Size n Will Be Positive When the Population Correlation Coefficient Is ρ

	Population correlation coefficient ρ																		
n	0	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10	.20	.30	.40	.50	.60	.70	.80	.90
3	.500	.505	.510	.515	.520	.525	.530	.535	.540	.545	.550	.600	.650	.700	.750	.800	.850	.900	.950
4	.500	.506	.513	.519	.525	.532	.538	.545	.551	.557	.564	.626	.688	.748	.805	.858	.906	.948	.981
5	.500	.508	.515	.522	.530	.537	.545	.552	.560	.567	.575	.648	.718	.784	.844	.896	.939	.972	.993
6	.500	.508	.517	.525	.534	.542	.551	.559	.568	.576	.584	.666	.743	.813	.873	.923	.960	.985	.997
7	.500	.509	.519	.528	.537	.547	.556	.565	.575	.584	.593	.683	.765	.837	.896	.942	.973	.991	.998
8	.500	.510	.520	.531	.541	.551	.561	.571	.581	.591	.601	.697	.784	.857	.915	.956	.982	.995	1.000
9	.500	.511	.522	.533	.544	.555	.565	.576	.587	.598	.608	.710	.800	.874	.929	.967	.988	.997	1.000
10	.500	.512	.523	.535	.546	.558	.570	.581	.592	.604	.615	.722	.815	.889	.941	.975	.992	.998	1.000
11	.500	.512	.525	.537	.549	.561	.573	.586	.598	.610	.621	.733	.828	.901	.951	.980	.994	.999	1.000
12	.500	.513	.526	.539	.552	.564	.577	.590	.602	.615	.627	.744	.840	.912	.959	.985	.996	.999	1.000
13	.500	.514	.527	.541	.554	.567	.581	.594	.607	.620	.633	.754	.851	.922	.966	.988	.997	1.000	1.000
14	.500	.514	.528	.542	.556	.570	.584	.598	.612	.625	.639	.763	.861	.930	.971	.991	.998	1.000	1.000
15	.500	.515	.529	.544	.558	.573	.587	.602	.616	.630	.644	.771	.871	.938	.976	.993	.999	1.000	1.000
16	.500	.515	.530	.546	.561	.576	.590	.605	.620	.634	.649	.779	.879	.944	.980	.995	.999	1.000	1.000
17	.500	.516	.531	.547	.563	.578	.593	.609	.624	.639	.654	.787	.887	.950	.983	.996	.999	1.000	1.000
18	.500	.516	.532	.549	.565	.581	.596	.612	.628	.643	.658	.794	.894	.955	.985	.997	1.000	1.000	1.000
19	.500	.517	.533	.550	.566	.583	.599	.615	.631	.647	.663	.801	.901	.960	.988	.997	1.000	1.000	1.000
20	.500	.517	.534	.551	.568	.585	.602	.618	.635	.651	.667	.808	.907	.964	.990	.998	1.000	1.000	1.000
21	.500	.518	.535	.553	.570	.587	.605	.622	.638	.655	.671	.814	.913	.967	.991	.998	1.000	1.000	1.000
22	.500	.518	.536	.554	.572	.590	.607	.625	.642	.659	.675	.820	.918	.971	.992	.999	1.000	1.000	1.000
23	.500	.518	.537	.555	.574	.592	.610	.627	.645	.662	.679	.826	.923	.974	.994	.999	1.000	1.000	1.000
24	.500	.519	.538	.557	.575	.594	.612	.630	.648	.666	.683	.831	.927	.976	.995	.999	1.000	1.000	1.000
25	.500	.519	.539	.558	.577	.596	.615	.633	.651	.669	.687	.836	.932	.979	.995	.999	1.000	1.000	1.000
50	.500	.528	.555	.583	.610	.636	.662	.687	.712	.735	.757	.920	.984	.998	1.000	1.000	1.000	1.000	1.000
100	.500	.540	.579	.617	.654	.690	.724	.757	.787	.815	.840	.978	.999	1.000	1.000	1.000	1.000	1.000	1.000
200	.500	.556	.611	.664	.714	.760	.801	.838	.871	.898	.921	.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000
300	.500	.569	.635	.698	.755	.806	.850	.887	.917	.940	.958	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
400	.500	.579	.655	.725	.788	.841	.885	.919	.945	.964	.977	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
500	.500	.588	.672	.749	.814	.868	.910	.941	.963	.978	.987	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
600	.500	.597	.688	.769	.836	.890	.929	.957	.975	.986	.993	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
700	.500	.604	.701	.786	.855	.907	.944	.968	.983	.991	.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
800	.500	.611	.714	.802	.871	.921	.955	.976	.988	.995	.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
900	.500	.618	.726	.816	.885	.933	.964	.982	.992	.997	.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
1,000	.500	.624	.736	.828	.897	.943	.971	.987	.994	.998	.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 2
Studies Correlating Perceived Vulnerability to Human Immunodeficiency Virus Infection With Precautionary Sexual Behavior

Sample	Sample characteristics	n	r	p	Direction
1	Homosexual men	329	.033	*ns*	+
2	Homosexual men	105		*ns*	−
3	Homosexual men	687	.045	*ns*	+
4	Homosexual men	60	.240	*ns*	+
5	Homosexual men	99	−.346	<.05	−
6	Homosexual men	637	.044	*ns*	+
7	Homosexual men	380	−.161	<.05	−
8	Homosexual men	197	−.120	*ns*	−
9	Homosexual men	578	.123	<.05	+
10	Homosexual men	266	−.080	*ns*	−
11	Heterosexual men	544		*ns*	−
12	Heterosexual women	580		*ns*	−
13	College students	513	−.115	<.05	−
14	College students	109	.400	<.05	+
15	College students	1,127	.085	<.05	+
16	College students	459	.069	*ns*	+
17	College students	1,008	−.300	<.05	−
18	College students	294	.236	<.05	+
19	Female college students	205	−.129	*ns*	−
20	Adolescent girls	75	−.180	*ns*	−
21	Adolescent girls	99		*ns*	+
22	Adolescent girls	402	−.073	*ns*	−
23	Female Marines	432	.025	*ns*	+
24	Female drug users	131	−.049	*ns*	−

Note. Nonsignificant results were assumed for samples that did not report p values. *ns* = statistically nonsignificant at the .05 level; + = positive result (i.e., in predicted direction); − = negative result (i.e., in opposite direction).

terval [.026, .139] does not contain the value zero, the null hypothesis that $\rho = 0$ is rejected.

Example 3. The data set used for Example 3 was taken from a meta-analytic review by Gerrard, Gibbons, and Bushman (1994). Table 2 contains the results from 24 independent samples of participants.[3] Correlations could be calculated for 20 of the 24 samples. The studies were conducted to test the hypothesis that perceived risk of human immunodeficiency virus (HIV) infection motivates preventive sexual behaviors (e.g., using condoms). To measure perceived risk to HIV infection, all researchers used some variation of the question "What is the likelihood that you will contract HIV?" or "What is the likelihood that you will develop acquired immunodeficiency syndrome (AIDS)?"

Although the correlations in Table 2 are not homogeneous, they suffice for purposes of illustration. In practice, the reviewer should test whether participant or study characteristics moderate the effects of the independent variable on the dependent variable. Meta-analytic procedures can then be applied to homogeneous subsets of results (see Gerrard et al., 1994, for a discussion of variables that were found to moderate the relation between perceived vulnerability to HIV infection and precautionary sexual behavior).

Suppose that one wishes to obtain an estimate and a confidence interval for ρ based on the proportion of positive results in Table 2. This example severely violates the equal sample size assumption (i.e., sample sizes range from $n = 60$ to $n = 1{,}127$), but it suffices for purposes of illustration. Using Equation 4, one obtains the square mean root for the 24 results,

$$n_{SMR} = \left(\frac{\sqrt{329} + \cdots + \sqrt{131}}{24}\right)^2 = 337.65 \text{ or } 338.$$

Because the correlation between perceived vulnerability and preventive sexual behavior was positive for 11 of the 24 results, the estimate of p is $\hat{p} = U/k = 11/24 = .458$. Because Table 1 contains only values of p greater than .500, the value $1 - \hat{p} = 1 - .458 = .542$ is used instead of the value $\hat{p} = .458$, and the sign of $\hat{\rho}$ is changed from positive to negative. The estimate of ρ is negative because $\hat{p} = .458$ is less than .5. Linear interpolation in Table 1 yields the estimate $\hat{\rho} = -.006$.

If $\rho = 0$, the normal distribution should provide a good approximation to the binomial distribution in this example because $kp = k(1 - p) = 24(.5) = 12 > 5$. The approximate 95% confidence interval based on the standard normal distribution given in Equation 12 is

$$.458 - 1.96\sqrt{\frac{.458(1 - .458)}{24}} \le p \le .458 + 1.96\sqrt{\frac{.458(1 - .458)}{24}}$$

or, simplifying, [.259, .658]. The approximate 95% confidence

[3] A list of the studies given in Table 2 can be obtained from Brad J. Bushman.

interval based on the chi-square distribution given in Equation 14 is

$$\frac{[2(.458)+0.160] \pm \sqrt{0.160^2 - 4(0.480)(.458)(1-.458)}}{2(1+0.160)},$$

where $b = \chi^2_{.05}(1)/k = 3.841/24 = 0.160$. Simplifying, one has [.279, .649]. Because the confidence interval based on chi-square distribution theory is narrower than is the confidence interval based on normal distribution theory (i.e., .370 < .399), it is used to obtain a confidence interval for ρ. Linear interpolation in Table 1 yields the ρ confidence interval [−.032, .021]. Because the ρ confidence interval contains the value zero, the null hypothesis that $\rho = 0$ is not rejected.

Estimates and Confidence Intervals Based on the Proportion of Significant Positive Correlations

Table 3 can be used to obtain an estimate and a confidence interval for ρ from the proportion of positive results significant at the $\alpha = .05$ significance level. Self-validating numerical methods were used to ensure that each entry in Table 3 had an absolute error of less than .001 (see Appendix A).

Example 4. Suppose that one wishes to obtain an estimate and a confidence interval for ρ based on the proportion of the total number of results in Table 2 that were positive and statistically significant at the $\alpha = .05$ level. In Table 2, 4 of the 24 results are positive and significant. Thus, the estimate of the proportion of significant positive results is $\hat{p} = U/k = 4/24 = .167$. The estimate of ρ is positive because the proportion of significant positive results is greater than .05. Linear interpolation in Table 3 yields the estimate $\hat{\rho} = .037$.

Checking the assumptions for the normal approximation to the binomial distribution, one has $kp = 24(.05) = 1.2$ and $k(1-p) = 24(.95) = 22.8$. Thus, if $\rho = 0$, the normal distribution probably will not provide a good approximation to the binomial distribution because $kp < 5$. The confidence interval for p based on chi-square distribution theory is [.067, .358], and the confidence interval for p based on standard normal distribution theory is [.018, .316]. Because the former p confidence interval is narrower than is the latter p confidence interval, it is used to obtain a confidence interval for ρ. Linear interpolation in Table 3 yields the ρ confidence interval [.008, .070]. Because the ρ confidence interval does not include the value zero, the null hypothesis that $\rho = 0$ is rejected. Recall that the confidence interval based on the proportion of positive results, [−.032, .021], included the value zero.

Estimates and Confidence Intervals Based on the Proportion of Positive and Negative Significant Correlations

Vote-counting procedures also can be used to examine the extent of publication bias. If studies with significant results are more likely to be published than are studies with nonsignificant results, then published studies may reflect only a biased subsample of the total collection of studies in the population. This problem can be minimized if the reviewer counts both positive and negative significant results, as Hedges and Olkin (1980) stated:

> When both positive and negative significant results are counted, it is possible to dispense with the requirement that the sample available is representative of all studies conducted. Instead, the requirement is that the sample of positive and negative *significant* results is representative of the population of positive or negative *significant* results. If only statistically significant findings tend to be published, this requirement is probably more realistic. (p. 366)

Suppose that one wishes to obtain an estimate and a confidence interval for θ using the k estimators $T_1, \ldots, T_k$. In this case, because the values $T_1, \ldots, T_k$ are not observed, one counts the number of times $T_i > C_{\alpha/2,n}$ and the number of times $T_i < -C_{\alpha/2,n}$. That is, one counts the number of two-sided tests that found significant positive or significant negative results. To obtain a confidence interval for θ, find the values $[\hat{\theta}_L, \hat{\theta}_U]$ that correspond to the upper and lower confidence limits $[\hat{p}_L, \hat{p}_U]$ for p. In this case, however, p denotes the proportion of the total number of significant results that were positive. That is,

$$p = \Pr\{\text{positive significant result} \mid \text{a significant result}\} = \frac{p^+}{p^+ + p^-}, \quad (21)$$

where p^+ is the probability of a significant positive result, defined as $\Pr(T_i > C_{\alpha/2,n})$, and p^- is the probability of a significant negative result, defined as $\Pr(T_i < -C_{\alpha/2,n})$. The maximum likelihood estimator of p is

$$\hat{p} = \frac{\hat{p}^+}{\hat{p}^+ + \hat{p}^-}, \quad (22)$$

where $\hat{p}^+$ and $\hat{p}^-$ are the respective proportions of positive and negative significant results out of the total number of significant results in the k independent studies.

Table 4 can be used to obtain a confidence interval for ρ from the proportion of significant correlations that were positive. Table 4, like Table 1, contains only proportions greater than .500. Proportions less than .500, which correspond to negative values of ρ, can be obtained by using $1 - \hat{p}$ instead of $\hat{p}$. Self-validating numerical methods were used to ensure that each entry in Table 4 had an absolute error of less than .001 (see Appendix A).

Example 5. Eight of the 24 results in Table 2 are statistically significant, 4 results are positive, and 4 are negative. Thus, $\hat{p}^+ = 4/8$ and $\hat{p}^- = 4/8$. The estimate of p is

$$\hat{p} = \frac{\hat{p}^+}{\hat{p}^+ + \hat{p}^-} = \frac{4/8}{4/8 + 4/8} = \frac{4}{4+4} = .5.$$

Table 4 shows that the estimate of ρ is $\hat{\rho} = 0$.

If $\rho = 0$, the normal distribution might not provide a good approximation to the binomial distribution in this example because $kp = k(1-p) = 8(.5) = 4 < 5$. The confidence interval for p based on chi-square distribution theory is [.215, .785], and the confidence interval for p based on standard normal distribution theory is [.154, .846]. Because the former p confidence interval is narrower than the latter p confidence interval, it is used to obtain a confidence interval for ρ. Linear interpola-

Table 3
Probability That a Correlation Coefficient From a Sample of Size n Will Be Significant at the .05 Significance Level When the Population Correlation Coefficient Is ρ

	Population correlation coefficient ρ																		
n	0	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10	.20	.30	.40	.50	.60	.70	.80	.90
3	.050	.051	.052	.052	.053	.054	.055	.056	.057	.057	.058	.068	.079	.093	.110	.132	.163	.211	.311
4	.050	.051	.052	.054	.055	.056	.058	.059	.061	.062	.064	.080	.102	.130	.167	.219	.295	.416	.629
5	.050	.052	.053	.055	.057	.058	.060	.062	.064	.066	.068	.092	.123	.166	.224	.305	.419	.583	.809
6	.050	.052	.054	.056	.058	.060	.062	.065	.067	.069	.072	.102	.143	.200	.277	.383	.524	.705	.900
7	.050	.052	.054	.057	.059	.062	.064	.067	.070	.073	.076	.112	.162	.232	.327	.453	.612	.791	.947
8	.050	.052	.055	.058	.060	.063	.066	.069	.072	.075	.079	.121	.181	.263	.375	.517	.684	.853	.970
9	.050	.053	.055	.058	.061	.064	.068	.071	.074	.078	.082	.130	.198	.293	.419	.573	.744	.897	.981
10	.050	.053	.056	.059	.062	.066	.069	.073	.077	.081	.085	.138	.215	.322	.460	.624	.793	.928	.984
11	.050	.053	.056	.060	.063	.067	.071	.075	.079	.083	.088	.146	.232	.350	.499	.669	.833	.950	1.000
12	.050	.053	.057	.060	.064	.068	.072	.076	.081	.086	.091	.154	.248	.377	.536	.710	.866	.965	1.000
13	.050	.053	.057	.061	.065	.069	.073	.078	.083	.088	.093	.162	.264	.403	.570	.746	.892	.976	1.000
14	.050	.054	.057	.061	.065	.070	.075	.079	.085	.090	.096	.170	.280	.428	.603	.778	.914	.984	1.000
15	.050	.054	.058	.062	.066	.071	.076	.081	.086	.092	.098	.178	.295	.452	.633	.806	.931	.989	1.000
16	.050	.054	.058	.062	.067	.072	.077	.082	.088	.094	.101	.185	.311	.475	.661	.831	.945	.992	1.000
17	.050	.054	.058	.063	.068	.073	.078	.084	.090	.096	.103	.192	.325	.498	.687	.853	.957	.995	1.000
18	.050	.054	.059	.063	.068	.074	.079	.085	.092	.098	.106	.200	.340	.520	.712	.872	.966	.996	1.000
19	.050	.054	.059	.064	.069	.075	.080	.087	.093	.100	.108	.207	.354	.541	.735	.889	.973	.997	1.000
20	.050	.054	.059	.064	.070	.075	.082	.088	.095	.102	.110	.214	.369	.561	.756	.904	.978	.998	1.000
21	.050	.055	.059	.065	.070	.076	.083	.089	.097	.104	.112	.221	.382	.581	.775	.916	.983	.999	1.000
22	.050	.055	.060	.065	.071	.077	.084	.091	.098	.106	.115	.228	.396	.600	.793	.928	.987	.999	1.000
23	.050	.055	.060	.066	.072	.078	.085	.092	.100	.108	.117	.235	.410	.618	.810	.938	.989	.999	1.000
24	.050	.055	.060	.066	.072	.079	.086	.093	.101	.110	.119	.242	.423	.635	.826	.946	.992	.999	1.000
25	.050	.055	.061	.066	.073	.080	.087	.095	.103	.112	.121	.249	.436	.652	.840	.953	.994	1.000	1.000
50	.050	.058	.066	.075	.085	.097	.109	.123	.137	.153	.170	.405	.692	.902	.984	.999	.999	1.000	1.000
100	.050	.061	.074	.089	.106	.125	.147	.171	.197	.226	.257	.642	.922	.994	1.000	1.000	1.000	1.000	1.000
200	.050	.066	.086	.111	.140	.173	.212	.255	.303	.354	.408	.887	.997	1.000	1.000	1.000	1.000	1.000	1.000
300	.050	.070	.097	.130	.170	.217	.272	.332	.397	.465	.535	.968	1.000	1.000	1.000	1.000	1.000	1.000	1.000
400	.050	.074	.106	.148	.199	.259	.328	.403	.482	.562	.640	.992	1.000	1.000	1.000	1.000	1.000	1.000	1.000
500	.050	.078	.115	.165	.226	.299	.380	.468	.557	.644	.724	.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000
600	.050	.081	.124	.181	.253	.337	.430	.528	.624	.713	.791	.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
700	.050	.084	.132	.197	.278	.373	.477	.582	.682	.770	.843	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
800	.050	.087	.140	.213	.303	.408	.521	.632	.732	.817	.882	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
900	.050	.089	.148	.228	.328	.442	.562	.676	.776	.855	.913	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
1,000	.050	.092	.156	.243	.353	.475	.600	.716	.813	.886	.936	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 4

Conditional Probability $Pr(\rho, n) = Pr\{r > 0 \mid |r| > c_{r,n}\}$ That a Correlation Coefficient r from a Sample of Size n Is Positive Given That the Absolute Value of r Exceeds the $\alpha = .05$ Critical Value for the Population Correlation Coefficient ρ

	Population correlation coefficient ρ																		
n	0	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10	.20	.30	.40	.50	.60	.70	.80	.90
3	.500	.508	.516	.524	.531	.539	.547	.555	.563	.570	.578	.654	.725	.790	.848	.897	.937	.968	.989
4	.500	.512	.525	.537	.550	.562	.574	.586	.598	.610	.622	.732	.823	.891	.938	.968	.986	.995	.999
5	.500	.516	.533	.549	.565	.581	.597	.613	.628	.644	.659	.790	.883	.940	.972	.988	.996	.999	1.000
6	.500	.520	.540	.559	.578	.598	.617	.635	.654	.671	.689	.832	.920	.965	.986	.995	.999	1.000	1.000
7	.500	.523	.546	.568	.590	.612	.634	.655	.675	.695	.714	.864	.943	.978	.993	.998	.999	1.000	1.000
8	.500	.526	.551	.576	.601	.625	.649	.672	.694	.715	.736	.887	.958	.986	.996	.999	1.000	1.000	1.000
9	.500	.528	.556	.583	.610	.637	.662	.687	.711	.733	.755	.906	.969	.991	.997	.999	1.000	1.000	1.000
10	.500	.530	.560	.590	.619	.647	.674	.701	.725	.749	.771	.920	.976	.994	.998	1.000	1.000	1.000	1.000
11	.500	.532	.565	.596	.627	.657	.686	.713	.739	.763	.786	.932	.981	.995	.999	1.000	1.000	1.000	1.000
12	.500	.534	.569	.602	.635	.666	.696	.724	.751	.776	.799	.942	.985	.997	.999	1.000	1.000	1.000	1.000
13	.500	.536	.572	.607	.642	.674	.706	.735	.762	.788	.811	.950	.988	.998	1.000	1.000	1.000	1.000	1.000
14	.500	.538	.576	.613	.648	.682	.714	.745	.773	.799	.822	.956	.991	.998	1.000	1.000	1.000	1.000	1.000
15	.500	.540	.579	.618	.655	.690	.723	.754	.782	.808	.832	.962	.992	.999	1.000	1.000	1.000	1.000	1.000
16	.500	.542	.582	.622	.661	.697	.731	.762	.791	.818	.841	.966	.994	.999	1.000	1.000	1.000	1.000	1.000
17	.500	.543	.586	.627	.666	.704	.738	.770	.800	.826	.850	.970	.995	.999	1.000	1.000	1.000	1.000	1.000
18	.500	.545	.589	.631	.672	.710	.745	.778	.807	.834	.857	.974	.996	.999	1.000	1.000	1.000	1.000	1.000
19	.500	.546	.591	.635	.677	.716	.752	.785	.815	.841	.865	.977	.997	1.000	1.000	1.000	1.000	1.000	1.000
20	.500	.548	.594	.639	.682	.722	.759	.792	.822	.848	.871	.979	.997	1.000	1.000	1.000	1.000	1.000	1.000
21	.500	.549	.597	.643	.687	.728	.765	.798	.828	.854	.877	.981	.998	1.000	1.000	1.000	1.000	1.000	1.000
22	.500	.550	.600	.647	.692	.733	.771	.804	.834	.861	.883	.983	.998	1.000	1.000	1.000	1.000	1.000	1.000
23	.500	.552	.602	.651	.696	.738	.776	.810	.840	.866	.889	.985	.998	1.000	1.000	1.000	1.000	1.000	1.000
24	.500	.553	.605	.654	.701	.743	.782	.816	.846	.872	.894	.986	.999	1.000	1.000	1.000	1.000	1.000	1.000
25	.500	.554	.607	.658	.705	.748	.787	.821	.851	.877	.898	.988	.999	1.000	1.000	1.000	1.000	1.000	1.000
50	.500	.579	.654	.723	.782	.832	.872	.904	.928	.947	.961	.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000
100	.500	.613	.715	.799	.863	.909	.941	.962	.976	.985	.990	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
200	.500	.658	.788	.877	.933	.964	.981	.990	.995	.997	.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
300	.500	.691	.834	.918	.962	.983	.992	.997	.999	.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
400	.500	.717	.866	.943	.977	.991	.996	.999	.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
500	.500	.739	.890	.958	.985	.995	.998	.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
600	.500	.758	.908	.969	.990	.997	.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
700	.500	.775	.922	.976	.993	.998	.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
800	.500	.789	.934	.982	.995	.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
900	.500	.802	.943	.985	.996	.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
1,000	.500	.814	.951	.988	.997	.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

tion in Table 4 yields the ρ confidence interval [−.016, .016]. Because the ρ confidence interval contains the value zero, the null hypothesis that $\rho = 0$ is not rejected. Recall that the confidence interval based on the proportion of positive results also included the value zero.

Estimates and Confidence Intervals Based on Unequal Sample Sizes

The vote-counting methods described in the previous section can be extended to handle unequal sample sizes (Hedges & Olkin, 1985). For each study, observe whether T_i exceeds some critical value C_{α,n_i}. Note that the sample size associated with the critical value has a subscript i. Because the sample sizes for the studies may differ, the critical values also may differ. For each study, there are two possible outcomes: a "success" if $T_i > C_{\alpha,n_i}$, or a "failure" if $T_i \le C_{\alpha,n_i}$. If $\alpha = .5$, a success is defined as a positive result and a failure is defined as a negative or null result. If $\alpha = .05$, a success is defined as a significant positive result and a failure is defined as a nonsignificant positive result, a negative result, or a null result. Let W_i be an indicator variable that takes on the value 1 if $T_i > C_{\alpha,n_i}$ or the value 0 if $T_i \le C_{\alpha,n_i}$. That is,

$$W_i = \begin{cases} 1 & \text{if } T_i > C_{\alpha,n_i} \\ 0 & \text{if } T_i \le C_{\alpha,n_i}. \end{cases} \quad (23)$$

The probability of a success is given by

$$\Pr(W_i = 1) = \Pr(T_i > C_{\alpha,n_i}) = p_i, \quad (24)$$

and the probability of a failure is given by

$$\Pr(W_i = 0) = \Pr(T_i \le C_{\alpha,n_i}) = 1 - p_i. \quad (25)$$

If T_i is approximately normally distributed with mean θ and variance S_θ^2, then the probability p_i can be expressed as an explicit function of θ,

$$p_i = \Pr(T_i > C_{\alpha,n_i}) = \Pr\{(T_i - \theta)/S_\theta > (C_{\alpha,n_i} - \theta)/S_\theta\} = 1 - \Phi[(C_{\alpha,n_i} - \theta)/S_\theta]. \quad (26)$$

The log-likelihood function is given by

$$L(\theta) = \sum_{i=1}^{k} [w_i \ln(p_i) + (1 - w_i)\ln(1 - p_i)], \quad (27)$$

where ln(.) is the natural logarithmic function. Because $n_1, \ldots, n_k$ are known and the data $w_1, \ldots, w_k$ are observed, the log-likelihood function $L(\theta)$ is a function of θ alone. Thus, $L(\theta)$ can be maximized over θ to obtain the maximum likelihood estimator $\hat{\theta}$. Unfortunately, there is generally no closed form expression for the maximum likelihood estimator $\hat{\theta}$, and $\hat{\theta}$ must be obtained numerically. The simplest way to obtain $\hat{\theta}$ is to calculate $L(\theta)$ for a coarse grid of θ values and use the maximum value of $L(\theta)$ to determine the region where $\hat{\theta}$ lies. Finer grids of θ can then be used to determine a more precise estimate of $\hat{\theta}$.

When k is large, $\hat{\theta}$ is approximately normally distributed, and the large-sample variance of $\hat{\theta}$ is given by

$$\mathrm{Var}(\hat{\theta}) = \left[\sum_{i=1}^{k} \frac{[D_i^{(1)}]^2}{p_i(1 - p_i)}\right]^{-1}, \quad (28)$$

where $D_i^{(1)} = \partial p_i/\partial\theta$ and the derivative is evaluated at $\theta = \hat{\theta}$ (see Hedges & Olkin, 1985, p. 70). The $100(1 - \alpha)\%$ confidence interval $[\hat{\theta}_L, \hat{\theta}_U]$ is given by

$$\hat{\theta}_L = \hat{\theta} - C_{\alpha/2}\sqrt{\mathrm{Var}(\hat{\theta})} \le \theta \le \hat{\theta} + C_{\alpha/2}\sqrt{\mathrm{Var}(\hat{\theta})} = \hat{\theta}_U. \quad (29)$$

Suppose that one observes whether the correlation r_i is positive for each study in a collection of k independent studies. If one assumes that the effect sizes are homogeneous and $n_1, \ldots, n_k$ become large at the same rate (i.e., $n_1/N, \ldots, n_k/N$ are fixed, where $N = n_1 + \cdots + n_k$), then the maximum likelihood estimator $\hat{\rho}$ of ρ is approximately normally distributed with mean ρ and variance $(1 - \rho^2)^2/N$ (see Hedges & Olkin, 1985, p. 234). Consequently, the maximum likelihood estimate for the probability of a positive result is given by

$$\hat{p}_i = \Pr(T_i > C_{\alpha,n_i}) = \Pr(r_i > 0) = \Pr\left[\frac{\sqrt{n_i}(r_i - \hat{\rho})}{(1 - \hat{\rho}^2)} > \frac{-\sqrt{n_i}\hat{\rho}}{(1 - \hat{\rho}^2)}\right] = 1 - \Phi\left[\frac{-\sqrt{n_i}\hat{\rho}}{(1 - \hat{\rho}^2)}\right], \quad (30)$$

and the derivative $D_i^{(1)}$ is given by (see Appendix B)

$$D_i^{(1)} = \frac{\partial p_i}{\partial \hat{\rho}} = \sqrt{\frac{n_i}{2\pi}}\left[\frac{1 + \hat{\rho}^2}{(1 - \hat{\rho}^2)^2}\right]\exp\left\{-\frac{1}{2}\left[\frac{n_i\hat{\rho}^2}{(1 - \hat{\rho}^2)^2}\right]\right\}. \quad (31)$$

Example 6. The proportion of positive results in Table 2 is used to illustrate how to obtain an estimate and a confidence interval for ρ using vote-counting procedures based on unequal sample sizes. Inserting the values of W_i for each of the 24 results into the log-likelihood function $L(\rho)$ given in Equation 27, for values of ρ ranging from −.3 to .3 in steps of .1, gives the values in the left-hand column of Table 5. This coarse grid of ρ values reveals that the maximum value of $L(\rho)$ lies between $\rho = -.1$ and $\rho = .1$. To obtain an estimate of ρ to two decimal places, one calculates the log-likelihood function values between $\rho = -.10$ and $\rho = .10$ in steps of .01. This fine grid of ρ values, given in the middle column of Table 5, reveals that the maximum value of $L(\rho)$ lies between $\rho = -.01$ and $\rho = .01$. To obtain an estimate of ρ to three decimal places, one calculates the log-likelihood function values between $\rho = -.010$ and $\rho = .010$ in steps of .001. The extra fine grid of ρ values is given in the right-hand column of Table 5. Because the log-likelihood function value is largest at $\rho = -.002$, the maximum likelihood estimate of ρ is $\hat{\rho} = -.002$. Recall that the estimate based on equal sample size vote-counting procedures was −.006. Thus, even though the sample sizes in Table 2 vary greatly (i.e., from $n = 60$ to $n = 1{,}127$), the estimates of ρ based on equal and unequal sample size vote-counting procedures are quite similar.

Table 6 gives the estimates of p_i, $D_i^{(1)}$, and $[D_i^{(1)}]^2/[p_i(1 - p_i)]$ for each sample in Table 2. Substituting the values given in Table 6 into Equation 28, one obtains the large-sample variance $\mathrm{Var}(\hat{\rho}) = 1/58{,}485.25 = .000017$. The 95% confidence interval for ρ given in Equation 29 is

$$-.002 - 1.96\sqrt{.000017} \le \rho \le -.002 + 1.96\sqrt{.000017}$$

Table 5
Log-Likelihood Function Values for ρ Based on Data From Studies of the Relation Between Perceived Vulnerability to Human Immunodeficiency Virus Infection and Preventive Sexual Behavior

Coarse grid		Fine grid		Extra fine grid	
ρ	$L(\rho)$	ρ	$L(\rho)$	ρ	$L(\rho)$
−.3	−185.470	−.10	−44.624	−.010	−16.840
−.2	129.946	−.09	−39.316	−.009	−16.793
−.1	−44.624	−.08	−34.553	−.008	−16.751
0	−16.636	−.07	−30.335	−.007	−16.716
.1	−45.285	−.06	−26.665	−.006	−16.687
.2	−126.748	−.05	−23.551	−.005	−16.663
.3	−221.005	−.04	−21.002	−.004	−16.646
		−.03	−19.028	−.003	−16.634
		−.02	−17.638	−.002	−16.628
		−.01	−16.840	−.001	−16.629
		0	−16.636	0	−16.636
		.01	−17.024	.001	−16.648
		.02	−17.999	.002	−16.667
		.03	−19.549	.003	−16.690
		.04	−21.660	.004	−16.720
		.05	−24.316	.005	−16.756
		.06	−27.500	.006	−16.800
		.07	−31.198	.007	−16.845
		.08	−35.398	.008	−16.899
		.09	−40.094	.009	−16.959
		.10	−45.285	.010	−17.024

or, simplifying, [−.010, .006]. Recall that the confidence interval for ρ based on the equal sample size vote-counting procedure was [−.032, .021]. The confidence interval based on the unequal sample size procedure is narrower than is the confidence interval based on the equal sample size procedure. Both confidence intervals contain the value zero.

Estimates and Confidence Intervals Based on Effect-Size Procedures

In this section, we describe effect-size procedures for obtaining an estimate and a confidence interval for ρ. In the next section, we use effect-size procedures to combine correlation coefficients and vote counts.

When correlation coefficients can be obtained from primary research reports, one estimator of the population correlation coefficient is the average of the sample correlation coefficients. Combining sample correlation coefficients, however, is complicated by the fact that the distribution of the correlation coefficient is not normal when $\rho \neq 0$. To remedy this problem, Fisher (1921) proposed a transformation from r to a quantity z that is approximately normally distributed with approximate standard error $1/\sqrt{n-3}$. The relation between r and z is given by

$$z = \frac{1}{2}\ln\frac{1+r}{1-r}. \tag{32}$$

To obtain a weighted average of the correlation coefficients, r_+, one first obtains a weighted average of the z values,

$$z_+ = \frac{\sum_{i=1}^{k}(n_i-3)z_i}{\sum_{i=1}^{k}(n_i-3)}, \quad i = 1, \ldots, k \tag{33}$$

(Cooper, 1989, p. 108). The value r_+ can then be obtained from the value z_+ by means of the z to r transformation

$$r = \frac{\exp\{2z\}-1}{\exp\{2z\}+1}. \tag{34}$$

There are two reasons for the choice of $(n_i - 3)$ as the weight in Equation 33. First, $(n-3)$ is the inverse of the variance for Fisher's z transformation. Second, z_+ is an unbiased estimator of $\zeta = \frac{1}{2}\ln(1+\rho/1-\rho)$ (i.e., r_+ is an unbiased estimator of ρ).

To obtain a $100(1-\alpha)\%$ confidence interval for the population correlation coefficient ρ, one first obtains the upper and lower bounds $[\hat{\zeta}_L, \hat{\zeta}_U]$ of a $100(1-\alpha)\%$ confidence interval for the population z transformation parameter ζ,

$$\zeta_L = z_+ - z_{\alpha/2}\frac{1}{\sqrt{N-3k}} \leq \zeta \leq z_+ + z_{\alpha/2}\frac{1}{\sqrt{N-3k}} = \zeta_U, \tag{35}$$

where $N = n_1 + \cdots + n_k$ and $z_{\alpha/2}$ is the two-sided critical value of the standard normal distribution. The z to r transformation (Equation 34) can then be used to obtain the upper and lower confidence interval bounds $[\hat{\rho}_L, \hat{\rho}_U]$ for ρ from the upper and lower confidence interval bounds $[\hat{\zeta}_L, \hat{\zeta}_U]$ for ζ.

Example 7. Suppose that one wants to obtain an estimate

Table 6
Computations for Obtaining the Large-Sample Variance of $\hat{\rho}$ Based on Data From Studies of the Relation Between Perceived Vulnerability to Human Immunodeficiency Virus Infection and Preventive Sexual Behavior

Sample	p_i	$D^{(1)}{}_i$	$[D^{(1)}{}_i]^2/[p_i(1-p_i)]$
1	.485	−22.718	2,066.23
2	.492	−12.840	659.65
3	.479	−32.806	4,312.34
4	.494	−9.707	376.97
5	.492	−12.468	621.96
6	.480	−31.592	3,998.78
7	.485	−24.413	2,386.35
8	.489	−17.584	1,237.46
9	.481	−30.097	3,628.72
10	.487	−20.430	1,670.72
11	.481	−29.201	3,415.43
12	.481	−30.149	3,641.26
13	.482	−28.358	3,220.95
14	.492	−13.082	684.78
15	.473	−41.981	7,069.73
16	.483	−26.827	2,882.13
17	.475	−39.712	6,324.33
18	.486	−21.478	1,846.51
19	.489	−17.938	1,287.70
20	.493	−10.853	471.20
21	.492	−12.468	621.96
22	.484	−25.109	2,524.43
23	.483	−26.028	2,712.70
24	.491	−14.341	822.96
Total			58,485.25

Table 7
Computations for Effect-Size Analysis

Sample	n	r	z	$n-3$	$(n-3)z$
1	329	.033	.033	326	10.758
3	687	.045	.045	684	30.780
4	60	.240	.245	57	13.965
5	99	−.346	−.361	96	−34.656
6	637	.044	.044	634	27.896
7	380	−.161	−.162	377	−61.074
8	197	−.120	−.121	194	−23.474
9	578	.123	.124	575	71.300
10	266	−.080	−.080	263	−21.040
13	513	−.115	−.116	510	−59.160
14	109	.400	.424	106	44.944
15	1,127	.085	.085	1,124	95.540
16	459	.069	.069	456	31.464
17	1,008	−.300	−.310	1,005	−311.550
18	294	.236	.241	291	70.131
19	205	−.129	−.130	202	−26.260
20	75	−.180	−.182	72	−13.104
22	402	−.073	−.073	399	−29.127
23	432	.025	.025	429	10.725
24	131	−.049	−.049	128	−6.272
Total	7,988			7,928	−178.214

Note. Samples for which correlation coefficients could not be estimated were omitted from the effect-size analysis (i.e., Samples 2, 11, 12, and 21).

and a confidence interval for ρ using the 20 sample correlation coefficients in Table 2. Using the sample sizes, z-transformation values, and $(n_i - 3)z_i$ values in Table 7, one obtains a weighted average of the z-transformation values (Equation 33):

$$z_+ = \sum_{i=1}^{20} (n_i - 3)z_i \Big/ \sum_{i=1}^{20} (n_i - 3) = -178.214/7{,}928 = -.023.$$

Applying the z to r transformation (Equation 34), one obtains the estimate

$$r_+ = \frac{\exp\{2z_+\} - 1}{\exp\{2z_+\} + 1} = \frac{\exp\{2(-.023)\} - 1}{\exp\{2(-.023)\} + 1} = -.023.$$

The estimate based on effect-size procedures is quite similar to the estimates based on vote-counting procedures.

To obtain a 95% confidence interval for ρ, one first obtains a 95% confidence interval for ζ using Equation 35:

$$-.023 - 1.96\frac{1}{\sqrt{7988 - 3(20)}} \le \zeta \le -.023 + 1.96\frac{1}{\sqrt{7988 - 3(20)}}.$$

Simplifying, one has [−.045, .001]. Applying the z to r transformation (Equation 34), one obtains the ρ confidence interval [−.045, −.001]. Contrary to expectation, the estimate of ρ is significantly less than zero.

Combining Information From Effect-Size and Vote-Counting Procedures

The meta-analyst generally has access to at least one of three types of data from primary research reports: (1) information that can be used to calculate effect-size estimates, (2) information about whether the hypothesis tests found statistically significant relations, and (3) information about the direction of outcomes. These data are rank ordered, from most to least, in terms of the amount of information they contain (Hedges, 1986). Effect-size procedures can be used to integrate the first type of data, whereas vote-counting procedures can be used to integrate the second and third types of data. Effect-size procedures are, of course, better than vote-counting procedures because they use all of the information from the studies that do report correlations. The problem, however, is that some studies do not report correlations. Furthermore, the correlations are likely to be missing for reasons related to the data (i.e., significant correlations are more likely to be reported than are nonsignificant correlations). If studies with missing correlations do report the direction of results or the statistical significance of results, then the meta-analyst can use vote-counting procedures to estimate the missing correlations. Effect-size procedures can then be used to integrate the potentially larger group of studies that includes those for which correlations cannot be calculated.

The combined procedure that we are proposing here has at least two advantages over the effect-size procedure. First, the combined procedure yields a less biased estimate of the population correlation coefficient than does the effect-size procedure. If significant correlations are more likely to be reported than are nonsignificant correlations, then the effect-size procedure will tend to overestimate the population correlation coefficient. That is, the estimate will be too large if ρ is positive and too small if ρ is negative. Second, the combined procedure yields a narrower confidence interval for the population correlation coefficient than does the effect-size procedure. The width of the 95% confidence interval for ζ is $(2)(1.96)(1/\sqrt{N - 3k})$. Clearly, the standard error $(1/\sqrt{N - 3k})$ will be smaller when the studies with missing correlations are included in the analysis because k and N will be larger. Although the interest here is in obtaining a confidence interval for ρ rather than a confidence interval for ζ, a narrower confidence interval for ζ implies a narrower confidence interval for ρ. We are in the process of conducting a simulation study to test the accuracy of the combined procedure.

To implement the combined procedure, the meta-analyst first uses vote-counting procedures to estimate the population correlation coefficient from the proportion of positive and significant positive results found in all studies. Let $\hat{\rho}_1$ denote the estimator based on the sample correlations alone, $\hat{\rho}_2$ denote the estimator based on the proportion of significant positive results, and $\hat{\rho}_3$ denote the estimator based on the proportion of positive results. The estimator $\hat{\rho}_2$ is used for studies that do not report correlations but do report the statistical significance of results. The estimator $\hat{\rho}_3$ is used for studies that do not report correlations or the statistical significance of results but do report the direction of results. Finally, effect-size procedures are used to obtain an estimate and a confidence interval for ρ using all three types of estimators. We illustrate the combined procedure in Example 8.

Example 8. The effect-size analysis used in Example 7 omitted the 4 samples that did not include enough information to estimate correlation coefficients (i.e., Samples 2, 11, 12, and

Table 8
Computations for Combining Effect Sizes and Vote Counts

Sample	n	r	z	$n-3$	$(n-3)z$
1	329	.033	.033	326	10.758
2	105	−.002	−.002	102	−0.204
3	687	.045	.045	684	30.780
4	60	.240	.245	57	13.965
5	99	−.346	−.361	96	−34.656
6	637	.044	.044	634	27.896
7	380	−.161	−.162	377	−61.074
8	197	−.120	−.121	194	−23.474
9	578	.123	.124	575	71.300
10	266	−.080	−.080	263	−21.040
11	544	−.002	−.002	541	−1.082
12	580	−.002	−.002	577	−1.154
13	513	−.115	−.116	510	−59.160
14	109	.400	.424	106	44.944
15	1,127	.085	.085	1,124	95.540
16	459	.069	.069	456	31.464
17	1,008	−.300	−.310	1,005	−311.550
18	294	.236	.241	291	70.131
19	205	−.129	−.130	202	−26.260
20	75	−.180	−.182	72	−13.104
21	99	.002	.002	96	0.192
22	402	−.073	−.073	399	−29.127
23	432	.025	.025	429	10.725
24	131	−.049	−.049	128	−6.272
Total	9,316			9,244	−180.462

Note. The correlations for Samples 2, 11, 12, and 21 were estimated with vote-counting procedures. The original sample sizes for Samples 2, 11, 12, and 21 were used in the calculations.

21). In Example 8, we use vote-counting procedures to estimate correlations for these 4 samples. We then use effect-size procedures to obtain a confidence interval for ρ using the information from all 24 samples.

We use the estimator $\hat{\rho}_3$ in this example because all four samples did not report correlations or the statistical significance of results but did report the direction of results. The equal and unequal sample size estimates of ρ based on the proportion of positive results were −.006 and −.002, respectively. We use the unequal sample size estimate because the number of participants in Samples 2, 11, 12, and 21 are not equal (i.e., sample sizes range from $n = 99$ to $n = 580$). Applying the r to z transformation (Equation 32) to the estimate −.002, one has

$$z_2 = z_{11} = z_{12} = \frac{1}{2} ln \frac{1+r}{1-r} = \frac{1}{2} \ln \frac{1-.002}{1+.002} = -.002.$$

$z_{21} = .002$ because the direction for this study is positive. Using the sample sizes, z-transformation values, and $(n_i - 3)z_i$ values in Table 8, one obtains a weighted average of the z-transformation values (Equation 33):

$$z_+ = \sum_{i=1}^{24} (n_i - 3)z_i \Big/ \sum_{i=1}^{24} (n_i - 3) = -180.462/9{,}244 = -.020.$$

Applying the z to r transformation (Equation 34), one obtains the estimate

$$r_+ = \frac{\exp\{2(-.020)\} - 1}{\exp\{2(-.020)\} + 1} = -.020.$$

To obtain a 95% confidence interval for ρ, one first obtains a 95% confidence interval for ζ:

$$-.020 - 1.96 \frac{1}{\sqrt{9316 - 3(24)}} \le \zeta$$

$$\le -.020 + 1.96 \frac{1}{\sqrt{9316 - 3(24)}}.$$

Simplifying, one has [−.040, .0003]. Applying the z to r transformation (Equation 34), one obtains the confidence interval [−.040, .0003] for ρ. Because the confidence interval includes the value zero, the null hypothesis that $\rho = 0$ is not rejected.

Note that the confidence interval based on the results from all 24 samples included in the combined effect-size and vote-counting analysis is narrower than is the confidence interval based on the results from the 20 samples included in the effect-size analysis (.040 < .044). The information provided by the 4 extra studies reduced the width of the confidence interval by .004.

Results All in the Same Direction

One limitation of vote-counting procedures is that they cannot be used if all of the results are in the same direction. The method of maximum likelihood cannot be used if $\hat{p}$ is unity or zero because there is not a unique corresponding value of $\hat{\rho}$. For example, $\hat{p}$ could be unity when sample sizes and correlations are large. If all of the results are in the same direction, one can obtain a Bayes estimate of p (e.g., Hedges & Olkin, 1985). For example, one can assume that p has a truncated uniform prior to distribution. That is, one assumes that there is a value p_0 such that any value of p in the interval $[p_0, 1]$ is equally likely. The Bayes estimator of p is given by

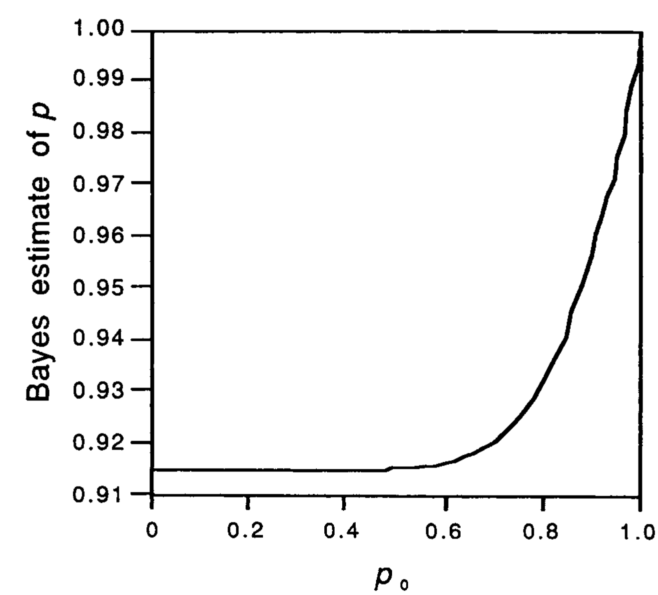

Figure 2. Bayes estimates for the results from $k = 10$ studies for values of p_0 ranging from 0 to .9999.

$$\hat{p} = \frac{(k+1)(1-p_0^{k+2})}{(k+2)(1-p_0^{k+1})}. \quad (36)$$

Example 9. Suppose that a meta-analyst has a sample of 10 studies and that all 10 studies found positive results. Also suppose that the meta-analyst believes that the proportion of positive results is at least .5 and that any value .5 or greater is equally likely. That is, the meta-analyst believes that any value of p in the interval [.5, 1] is equally likely. The Bayes estimate of p given in Equation 36 is

$$\hat{p} = \frac{(10+1)(1-.5^{(10+2)})}{(10+2)(1-.5^{(10+1)})} = .9169.$$

For purposes of comparison, Figure 2 gives Bayes estimates for values of p_0 ranging from 0 to .9999 for $k = 10$. In practice, of course, the meta-analyst selects the value of p_0 on a priori grounds.

Conclusions

Missing effect-size estimates pose a particularly difficult problem in meta-analysis. Often studies do not include enough information to permit the calculation of a correlation coefficient. Even if primary research reports do not contain enough information to calculate correlations, they may still provide information about the magnitude of the linear relation between the independent and dependent variables. Often the information is in the form of a report of the decision yielded by the significance test (e.g., a significant positive correlation) or in the form of a direction of the relation without regard to its statistical significance (e.g., a positive correlation). In this article, we have described three vote-counting procedures that use information about the direction and statistical significance of study results to obtain an estimate and a confidence interval for the population correlation coefficient. Easy-to-use tables, based on equal sample sizes, were presented for all three vote-counting procedures. More complicated vote-counting procedures, based on unequal sample sizes, also were described. We view vote-counting procedures as supplements rather than as alternatives to effect-size procedures. Vote-counting procedures should be used to estimate missing sample correlations, but they should not be used to obtain an estimate and a confidence interval for the population correlation coefficient unless none of the studies contain enough information to calculate sample correlation coefficients. We recommend that meta-analysts use the procedure we have proposed for combining information from effect-size and vote-counting procedures because it uses as much information as possible to obtain an estimate and a confidence interval for the population correlation coefficient. The combined procedure yields a less biased estimate of the population correlation coefficient, and a narrower confidence interval for the population correlation coefficient, than does the effect-size procedure. Whenever possible, meta-analysts should use all available data to estimate population effect sizes and confidence intervals. Data are the meta-analyst's most precious commodity; they are the stones that are used to build the structure called "science."

References

Bushman, B. J. (1994). Vote-counting procedures in meta-analysis. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 193–213). New York: Russell Sage Foundation.

Clopper, C. J., & Pearson, E. S. (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika, 26,* 404–413.

Cooper, H. M. (1989). *Integrating research: A guide for literature reviews* (2nd ed.). Newbury Park, CA: Sage.

Cooper, H., & Rosenthal, R. (1980). Statistical versus traditional procedures for summarizing research findings. *Psychological Bulletin, 87,* 442–449.

DeGroot, M. H. (1986). *Probability and statistics* (2nd ed.). Reading, MA: Addison-Wesley.

Fisher, R. A. (1921). On the 'probable error' of a coefficient of correlation deduced from a small sample. *Metron, 1,* 1–32.

Gerrard, M., Gibbons, F. X., & Bushman, B. J. (1994). *The relation between perceived vulnerability to HIV and precautionary sexual behavior.* Manuscript submitted for publication.

Gibbons, J. D., Olkin, I., & Sobel, M. (1977). *Selecting and ordering populations: A new statistical methodology.* New York: Wiley.

Hedges, L. V. (1986). *Estimating effect sizes from vote counts or box score data.* Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Hedges, L. V., & Olkin, I. (1980). Vote-counting methods in research synthesis. *Psychological Bulletin, 88,* 359–369.

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.

Kennedy, W. J. (1990). Special purpose numerical tools for approximation functions. *Statistical Computing and Statistical Graphics Newsletter, 1,* 3–6.

Olkin, I. (1973). *Do positive population correlation coefficients yield positive sample correlation coefficients?* (Technical Report No. 73). Stanford, CA: Stanford University, Department of Statistics.

Olkin, I. (1990). History and goals. In K. W. Wachter & M. L. Straf (Eds.), *The future of meta-analysis* (pp. 3–10). New York: Russell Sage Foundation.

Pigott, T. D. (1994). Methods for handling missing data in research synthesis. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 163–175). New York: Russell Sage Foundation.

Wang, M., & Kennedy, W. J. (1990). Comparison of algorithms for bivariate normal probability over a rectangle based on self-validating results from interval analysis. *Journal of Statistical Computation and Simulation, 37,* 13–25.

Wang, M., & Kennedy, W. J. (1992). A numerical method for accurately approximating multivariate normal probabilities. *Computational Statistics and Data Analysis, 13,* 197–210.

Wang, M., & Kennedy, W. J. (in press). A self-validating numerical method for computation of central and noncentral *F* probabilities and percentiles. *Statistics and Computing.*

Appendix A

Computational Details for Obtaining the Values for Tables 1, 3, and 4

Under the alternative hypothesis given in Display 19, the probability density function of the sample Pearson product-moment correlation coefficient r (Equation 18) is

$$f_{\rho,n}(r) = \sum_{j=0}^{\infty} d_j \rho^j (1-\rho^2)^{(n-1)/2} \frac{r^j(1-r^2)^{n/2-2}}{B[(j+1)/2, n/2-1]}, \quad \text{(A1)}$$

where

$$d_j = \frac{\Gamma\left[\frac{(n+j-1)}{2}\right]}{\Gamma\left(\frac{j}{2}+1\right)\Gamma\left[\frac{(n-1)}{2}\right]},$$

$B(.,.)$ is the complete beta function, and $\Gamma(.)$ is the gamma function (Olkin, 1973). Under the null hypothesis given in Display 19, the probability density function given in Equation A1 reduces to

$$f_{\rho=0,n}(r) = \frac{(1-r^2)^{n/2-2}}{B(1/2, n/2-1)}. \quad \text{(A2)}$$

The values in Table 1 were obtained by evaluating the probability

$$\Pr_{\rho,n}(r > 0) = \int_0^1 f_{\rho,n}(r)dr. \quad \text{(A3)}$$

It is easy to verify that

$$\Pr_{\rho,n}(r > 0) = \frac{(1-\rho)^{(n-1)/2}}{2} \sum_{j=0}^{\infty} d_j \rho^j. \quad \text{(A4)}$$

Although Equation A4 can easily be evaluated on a computer, numerical difficulties such as rounding and cancellation can cause serious precision problems. To ensure that each entry in Table 1 had an absolute error of less than .001, we used self-validating numerical methods (see Kennedy, 1990; Wang & Kennedy, 1990, 1992).

The values in Table 3 were obtained by evaluating the probability

$$\Pr_{\rho,n}(r > C_{\alpha,n}) = \int_{C_{\alpha,n}}^1 f_{\rho,n}(r)dr, \quad \text{(A5)}$$

where $C_{\alpha,n}$ is the one-sided critical value, such that $\alpha = \int_{C_{\alpha,n}}^1 f_{\rho=0,n}(r)dr$, and $0 < \alpha < .5$. To evaluate Equation A5, one must first obtain $C_{\alpha,n}$ by solving the equation

$$\alpha = \int_{C_{\alpha,n}}^1 f_{\rho=0,n}(r)dr = \int_{C_{\alpha,n}}^1 \frac{(1-r^2)^{n/2-2}}{B(1/2, n/2-1)}dr = \left(\frac{1}{2}\right)\frac{\int_{C^2_{\alpha,n}}^1 t^{1/2-1}(1-t)^{n/2-2}dt}{B(1/2, n/2-1)} = \left(\frac{1}{2}\right)I(n/2-1, 1/2, 1-C^2_{\alpha,n}), \quad \text{(A6)}$$

where $I(.,.,.)$ is the incomplete beta function. The beta percentile subroutine can be used to solve Equation A6 for $C_{\alpha,n}$ (Wang & Kennedy, in press). After $C_{\alpha,n}$ has been obtained, the probability given in Equation A5 can be evaluated as follows:

$$\Pr_{\rho,n}(r > C_{\alpha,n}) = \int_{C_{\alpha,n}}^1 f_{\rho,n}(r)dr = \sum_{j=0}^{\infty} d_j\rho^j(1-\rho^2)^{(n-1)/2} \times \int_{C_{\alpha,n}}^1 \frac{r^j(1-r^2)^{n/2-2}}{B[(j+1)/2, n/2-1]}dr = \frac{(1-\rho^2)^{(n-1)/2}}{2} \times \sum_{j=0}^{\infty} d_j\rho^j I(n/2-1, (j+1)/2, 1-C^2_{\alpha,n}). \quad \text{(A7)}$$

Equation A7 contains a sequence of incomplete beta functions that can be solved with the beta probability subroutine (Wang & Kennedy, in press).

The values in Table 4 were obtained by computing the probability

$$\Pr_{\rho,n}(r > C_{\alpha/2,n} | r > C_{\alpha/2,n} \text{ or } r < -C_{\alpha/2,n}), \quad \text{(A8)}$$

where $C_{\alpha/2,n}$ is the two-sided critical value, such that

$$\alpha = \int_{C_{\alpha/2,n}}^1 f_{\rho=0,n}(r)dr + \int_{-1}^{-C_{\alpha/2,n}} f_{\rho=0,n}(r)dr,$$

and $0 < \alpha < .5$. Because $f_{\rho=0,n}(r)$ is symmetric, $\alpha = I(n/2-1, 1/2, 1-C^2_{\alpha,n})$. The beta percentile subroutine can be used to find $C^2_{\alpha,n}$. After $C^2_{\alpha,n}$ has been obtained, the probability given in Equation A8 can be written as follows:

$$\Pr_{\rho,n}(r > C_{\alpha/2,n} | r > C_{\alpha/2,n} \text{ or } r < -C_{\alpha/2,n}) = \frac{\Pr_{\rho,n}(r > C_{\alpha/2,n})}{\Pr_{\rho,n}(r > C_{\alpha/2,n}) + \Pr_{\rho,n}(r < -C_{\alpha/2,n})}. \quad \text{(A9)}$$

Methods similar to those used for Table 3 were used to calculate $\Pr_{\rho,n}(r > C_{\alpha/2,n})$ for Table 4. The $\Pr_{\rho,n}(r < -C_{\alpha/2,n})$ is given by

$$\Pr_{\rho,n}(r < -C_{\alpha/2,n}) = \int_{-1}^{-C_{\alpha/2,n}} f_{\rho,n}(r)dr = \sum_{j=0}^{\infty} d_j\rho^j(1-\rho^2)^{(n-1)/2} \times \int_{-1}^{-C_{\alpha/2,n}} \frac{r^j(1-r^2)^{n/2-2}}{B[(j+1)/2, n/2-1]}dr = \frac{(1-\rho^2)^{(n-1)/2}}{2} \times \sum_{j=0}^{\infty} d_j(-\rho)^j I(n/2-1, (j+1)/2, 1-C^2_{\alpha,n}). \quad \text{(A10)}$$

Equation A10 contains a sequence of incomplete beta functions that can be solved with the beta probability subroutine.

(*Appendix B follows on next page*)

Appendix B

Computational Details for Obtaining the Derivative $D_i^{(1)}$

The derivative $D_i^{(1)}$ in Equation 31 was obtained as follows:

$$D_i^{(1)} = \frac{\partial p_i}{\partial \hat{\rho}} = \frac{\partial \int_{-\infty}^{\frac{+\sqrt{n_i}\hat{\rho}}{(1-\hat{\rho}^2)}} \frac{1}{2\pi} \exp\left\{\frac{-t^2}{2}\right\} dt}{\partial \hat{\rho}}$$

$$= \frac{1}{2\pi} \exp\left\{-\frac{1}{2}\left[\frac{-\sqrt{n_i}\hat{\rho}}{(1-\hat{\rho}^2)}\right]^2\right\} \frac{\partial\left[\frac{+\sqrt{n_i}\hat{\rho}}{(1-\hat{\rho}^2)}\right]}{\partial \hat{\rho}}$$

$$= \frac{1}{2\pi} \exp\left\{-\frac{1}{2}\left[\frac{n_i\hat{\rho}^2}{(1-\hat{\rho}^2)^2}\right]\right\} \frac{+\sqrt{n_i}(1+\hat{\rho}^2)}{(1-\hat{\rho}^2)^2}$$

$$= \sqrt{\frac{n_i}{2\pi}}\left[\frac{1+\hat{\rho}^2}{(1-\hat{\rho}^2)^2}\right]\exp\left\{-\frac{1}{2}\left[\frac{n_i\hat{\rho}^2}{(1-\hat{\rho}^2)^2}\right]\right\}.$$

Note, first, that by symmetry, $1 - \Phi(a) = \Phi(-a)$. Thus,

$$1 - \Phi\left[\frac{-\sqrt{n_i}\hat{\rho}}{(1-\hat{\rho}^2)}\right] = \Phi\left[\frac{\sqrt{n_i}\hat{\rho}}{(1-\hat{\rho}^2)}\right].$$

Second,

$$\frac{\partial f(x)g(x)}{\partial x} = f'(x)g(x) + f(x)g'(x).$$

Thus,

$$\frac{\partial\left[\frac{\sqrt{n_i}\hat{\rho}}{(1-\hat{\rho}^2)}\right]}{\partial \hat{\rho}} = \frac{\partial}{\partial \hat{\rho}}(\sqrt{n_i}\hat{\rho})(1-\hat{\rho}^2)^{-1}$$

$$= \sqrt{n_i}(1-\hat{\rho}^2)^{-1} + \sqrt{n_i}(1-\hat{\rho}^2)^{-2}(-1)(-2\hat{\rho})$$

$$= \sqrt{n_i}(1-\hat{\rho}^2)^{-2}[(1-\hat{\rho}^2) + 2\hat{\rho}^2]$$

$$= \frac{\sqrt{n_i}(1+\hat{\rho}^2)}{(1-\hat{\rho}^2)^2}.$$

Received July 26, 1993
Revision received October 14, 1994
Accepted October 17, 1994 ■