A Meta-Analytic Test of Intergroup Contact Theory

Thomas F. Pettigrew
University of California, Santa Cruz

Linda R. Tropp
Boston College

The present article presents a meta-analytic test of intergroup contact theory. With 713 independent samples from 515 studies, the meta-analysis finds that intergroup contact typically reduces intergroup prejudice. Multiple tests indicate that this finding appears not to result from either participant selection or publication biases, and the more rigorous studies yield larger mean effects. These contact effects typically generalize to the entire outgroup, and they emerge across a broad range of outgroup targets and contact settings. Similar patterns also emerge for samples with racial or ethnic targets and samples with other targets. This result suggests that contact theory, devised originally for racial and ethnic encounters, can be extended to other groups. A global indicator of Allport's optimal contact conditions demonstrates that contact under these conditions typically leads to even greater reduction in prejudice. Closer examination demonstrates that these conditions are best conceptualized as an interrelated bundle rather than as independent factors. Further, the meta-analytic findings indicate that these conditions are not essential for prejudice reduction. Hence, future work should focus on negative factors that prevent intergroup contact from diminishing prejudice as well as the development of a more comprehensive theory of intergroup contact.

Keywords: intergroup prejudice, intergroup contact, meta-analysis

For decades, researchers and practitioners have speculated about the potential for intergroup contact to reduce intergroup prejudice. Some writers thought contact between the races under conditions of equality would only breed "suspicion, fear, resentment, disturbance, and at times open conflict" (Baker, 1934, p. 120). Others proposed that interracial experiences could lead to "mutual understanding and regard" (Lett, 1945, p. 35) and that when groups "are isolated from one another, prejudice and conflict grow like a disease" (Brameld, 1946, p. 245; see also Watson, 1946).

Early studies of intergroup contact provided preliminary tests of these proposals and revealed encouraging trends. After the U.S. Merchant Marine began desegregating, Brophy (1946) found that the more voyages the White seamen took with Blacks, the more positive their racial attitudes became. Likewise, White police officers who worked with Black colleagues later objected less to having Blacks join their police districts, teaming with a Black partner, and taking orders from Black officers (Kephart, 1957).

As studies of intergroup contact grew in number, the Social Science Research Council asked Robin Williams, Jr., a Cornell University sociologist, to review research on group relations. Williams's (1947) monograph, *The Reduction of Intergroup Tensions*, offers 102 testable "propositions" on intergroup relations that include an initial formulation of intergroup contact theory. In particular, he noted that intergroup contact would maximally reduce prejudice when the two

Thomas F. Pettigrew, Department of Psychology, University of California, Santa Cruz; Linda R. Tropp, Department of Psychology, Boston College.

A preliminary report of this research, which included the first 203 studies collected, appeared in Pettigrew and Tropp (2000). The National Science Foundation supported this research (SBR-9709519), with Thomas F. Pettigrew and Stephen Wright serving as coinvestigators. Though listed alphabetically, the co-authors shared equally in the preparation of this manuscript and in their work on this 8-year project. An earlier version of this paper won the Gordon W. Allport Intergroup Research Prize awarded by the Society for the Psychological Study of Social Issues in 2003.

We thank Stephen Wright and the following dedicated research assistants at the University of California, Santa Cruz, and Boston College for their invaluable assistance: Rebecca Boice, Geoffrey Burcaw, Susan Burton, Darcy Cabral, Robert Chang, Daniel Cheron, Vanessa Lee, Kimberly Lincoln, Peter Moore, Danielle Murray, Neal Nakano, Rajinder Samra, Michael Sarette, Christine Schmitt, Amanda Stout, and Gina Vittori. We are grateful to the library staffs of the University of California, Santa Cruz, and Boston College for their help with tracking down hundreds of relevant journal articles in a wide assortment of specialized journals. We thank Makiko Deguchi, Greg Kim, Adele Kohanyi, Daphne Malinsky, and Mark Pettigrew for their help in translating research articles. We also thank Sue Duval, Blair Johnson, David Kenny, and Jack Vevea for their statistical advice and assistance and Jack Dovidio, Samuel Gaertner, Miles Hewstone, and Brian Mullen for their comments on earlier versions of this article. In addition, we wish to express our deep appreciation for the many researchers who dug into their files to find the detailed data needed to compute effect sizes.

Correspondence concerning this article should be addressed to Thomas F. Pettigrew, Department of Psychology, Social Sciences II, University of California, Santa Cruz, California 95064 or to Linda R. Tropp, Department of Psychology, McGuinn Hall, Boston College, Chestnut Hill, Massachusetts 02467. E-mail: pettigr@ucsc.edu or tropp@bc.edu

Journal of Personality and Social Psychology, 2006, Vol. 90, No. 5, 751–783

0022-3514/06/$12.00 DOI: 10.1037/0022-3514.90.5.751

groups share similar status, interests, and tasks and when the situation fosters personal, intimate intergroup contact.

Researchers then began to test the theory more rigorously. Field studies of public housing provided the strongest evidence and marked the introduction of large-scale field research into North American social psychology. In a notable example of this work, Deutsch and Collins (1951) interviewed White housewives across different public housing projects with a design that Campbell and Stanley (1963) later labeled "quasi-experimental." Two housing projects in Newark assigned Black and White residents to separate buildings. Two comparable housing projects in New York City desegregated residents by making apartment assignments irrespective of race or personal preference. The authors found that White women in the desegregated projects had far more optimal contact with their Black neighbors. Moreover, they held their Black neighbors in higher esteem and expressed greater support for interracial housing. Further research extended this work, showing that equal-status interracial contact in public housing related to more positive feelings and intergroup attitudes for both Blacks and Whites (Wilner, Walkley, & Cook, 1952; Works, 1961).

Armed with Williams's initial effort and the rich findings of the housing studies, Allport (1954) introduced the most influential statement of intergroup contact theory in *The Nature of Prejudice*. His formulation of intergroup contact theory maintained that contact between groups under optimal conditions could effectively reduce intergroup prejudice. In particular, Allport held that reduced prejudice will result when four features of the contact situation are present: equal status between the groups in the situation; common goals; intergroup cooperation; and the support of authorities, law, or custom.

Allport's formulation of intergroup contact theory has inspired extensive research over the past half century (Pettigrew, 1998; Pettigrew & Tropp, 2000). These investigations range across a variety of groups, situations, and societies. Going beyond a focus on racial and ethnic groups, investigators have tested the theory with participants of varying ages and with target groups as diverse as elderly, physically disabled, and mentally ill participants. Contact studies also have used a wide variety of research methods and procedures, including archival research (e.g., Fine, 1979), field studies (e.g., Deutsch & Collins, 1951), laboratory experiments (e.g., Cook, 1969, 1978), and surveys (e.g., Pettigrew, 1997; Sigelman & Welch, 1993). In addition to spanning many disciplines, contact theory has been usefully applied to a host of pressing social issues ranging from the racial desegregation of schools (Pettigrew, 1971) and the resolution of ethnopolitical conflicts (Chirot & Seligman, 2001) to explaining regional differences in prejudice (Wagner, van Dick, Pettigrew, & Christ, 2003) and the educational mainstreaming of physically and mentally disabled children (Harper & Wacker, 1985; Naor & Milgram, 1980).

Past Reviews of the Intergroup Contact Literature

Past reviews of this vast literature have often reached conflicting conclusions regarding the likely effects of intergroup contact. Numerous reviews show general support for contact theory, suggesting that intergroup contact typically reduces intergroup prejudice (Cook, 1984; Harrington & Miller, 1992; Jackson, 1993; Patchen, 1999; Pettigrew, 1971, 1986, 1998). However, other reviews reach more mixed conclusions. Amir (1969, 1976) conceded that contact under optimal conditions tends to reduce prejudice among participants, but he stressed that these reductions in prejudice may not generalize to entire outgroups. Moreover, Amir (1976) noted that contact under unfavorable conditions "may increase prejudice and intergroup tension" (p. 308). Likewise, Forbes (1997), a political scientist, concluded that intergroup contact often lowers prejudice at the individual level of analysis but fails to do so at the group level of analysis. Hence, he argued that contact can cure individual prejudice but not group conflict.[1]

Stephan (1987) acknowledged that intergroup contact has the potential to reduce prejudice, but he emphasized the complexity involved in the link between intergroup contact and prejudice. For example, characteristics of the contact setting, the groups under study, and the individuals involved may all contribute to enhancing or inhibiting contact's effects (see also Patchen, 1999; Pettigrew, 1998; Riordan, 1978).

Additional reviews have been more critical regarding the potential for contact to promote positive intergroup outcomes. Ford (1986) examined 53 papers (from six journals) on contact. He found support for the contact hypothesis to be, at best, "premature" and that the research presented in these papers was "grossly insufficient in representing the various settings of daily life" (Ford, 1986, p. 256). McClendon (1974) suggested that "contact research has been rather unsophisticated and lacking in rigor" (p. 47) and concluded that this body of work "would not lead [one] to expect a widespread reduction in prejudice" (p. 52).

Such conflicting views regarding the effects of contact have led some social psychologists to discard contact theory. Indeed, as Hopkins, Reicher, and Levine (1997) asserted, some believe that "the initial hopes of contact theorists have failed to materialize" (p. 306). However, three major shortcomings of these past reviews may account for their divergent conclusions.

Incomplete Samples of Relevant Papers

Whereas there have been literally dozens of partial reviews of the contact literature, we are not aware of any review of this vast literature that attempts to encompass the entire relevant research base. Instead, past reviews have typically been restricted to a particular set of groups, such as racial or ethnic groups (e.g., Patchen, 1999), or a particular contact setting, such as interracial housing (e.g., Cagle, 1973) or schools (e.g., Carithers, 1970). Further examination also reveals that these reviews covered a mean of less than 60 research articles each, compared with the hundreds of studies that comprise the contact research literature. Thus, past reviews offer only limited views regarding the effects of intergroup contact.

Absence of Strict Inclusion Rules

With no inclusion rules, these previous reviews used sharply contrasting definitions of intergroup contact. For example, some reviews included studies that used intergroup proximity, rather than established contact, as the independent variable. This procedure may contribute considerable error to the analysis and obscure the test of contact's influence on prejudice.

[1] Many social psychologists would take issue with Forbes's distinction. If reductions in prejudice generalize broadly from contact, the group level of analysis is necessarily involved (Brown & Hewstone, 2005).

Nonquantitative Assessments of Contact Effects

Moreover, none of the previous reviews used fully quantitative assessments of contact effects. Instead, authors of past reviews have tended to offer subjective judgments of the contact–prejudice relationship that are based on their own readings of a subset of the research literature. Although this approach can be useful, selection biases and differing interpretations of the literature limit the ability to reach definitive conclusions about contact's effects. Thus, quantitative approaches to research synthesis are preferred, as they provide a means for examining replicable patterns of effects across the full accumulation of relevant studies (Johnson & Eagly, 2000; Rosenthal, 1991).

Given these limitations of past reviews and the diverse nature of contact research, the evaluation of this literature requires a meta-analytic approach. Yet, to our knowledge, no investigators have conducted such an analysis of this vast and rich research literature. Thus, a central goal of the present research was to assess the overall effect between intergroup contact and prejudice on the basis of the population of empirical studies that constitute the research literature of the 20th century. With a quantitative means of analyzing a far greater number of relevant studies chosen by strict inclusion rules, we aim to determine more conclusively than past reviews the overall relationship between intergroup contact and prejudice.

The present article reports on such an effort, utilizing 515 individual studies with 713 independent samples and 1,383 non-independent tests. Combined, 250,089 individuals from 38 nations participated in the research. Along with including more than 300 additional studies, this work extends an earlier preliminary analysis, presented in Pettigrew and Tropp (2000), in several important theoretical and empirical directions.

Testing and Reinterpreting Allport's Optimal Conditions

Whereas intergroup contact theory has traditionally held that Allport's optimal conditions are essential, we propose that Allport's conditions facilitate contact's reduction of intergroup prejudice. Social psychology has shown repeatedly that greater exposure to targets can, in and of itself, significantly enhance liking for those targets (Bornstein, 1989; Harmon-Jones & Allen, 2001; Lee, 2001; Zajonc, 1968; see also Homans, 1950). Moreover, studies with social targets have shown that the enhanced liking that results from exposure can generalize to greater liking for other related, yet previously unknown, social targets (Rhodes, Halberstadt, & Brajkovich, 2001). Applying this work to intergroup contact research, the mere exposure perspective suggests that, all things being equal, greater contact and familiarity with members of other groups should enhance liking for those groups. Thus, in the present analysis, we test whether intergroup contact is associated with less prejudice even when Allport's conditions are not established in the contact situation, as well as whether these conditions significantly enhance the degree to which contact promotes positive intergroup outcomes.

Extending our earlier, preliminary analysis (Pettigrew & Tropp, 2000), we examine these issues in three ways. Initially, we use a global indicator of Allport's conditions: structured programs designed to achieve optimal conditions. We test whether including this indicator in the contact situation is necessary to produce positive intergroup outcomes and whether it typically enhances the positive effects of contact. To check for the consistency of these effects, we then compare the results for our full sample with those subsets of cases that either do or do not involve racial and ethnic contact. In addition, for the subset of 134 samples that feature a structured situation boasting many of Allport's proposed conditions, we rate each condition separately and examine the average effect sizes associated with each condition.

Ruling Out Alternative Explanations for Contact–Prejudice Effects

The present analysis also includes additional tests that allow us to test more effectively four alternative explanations for contact–prejudice effects.

Participant Selection and the Causal Sequence Problem

A potential participant selection bias could limit the interpretation of many studies of intergroup contact (Pettigrew, 1998). Instead of optimal contact reducing prejudice, the opposite causal sequence could be at work. Prejudiced people may avoid, and tolerant people may seek, contact with outgroups.

Statistical methods borrowed from econometrics allow researchers to compare roughly the reciprocal paths (contact lowers prejudice vs. prejudice decreases contact) with cross-sectional data. As shown directly in other research (e.g., Herek & Capitanio, 1996), these methods reveal that prejudiced people do indeed avoid intergroup contact. But the path from contact to reduced prejudice is generally much stronger (Butler & Wilson, 1978; Pettigrew, 1997; Powers & Ellison, 1995; Van Dick et al., 2004). Longitudinal studies also have revealed that optimal contact reduces prejudice over time (e.g., Eller & Abrams, 2003, 2004; Levin, van Laar, & Sidanius, 2003), even when researchers have eliminated the possibility of participant selection (e.g., Sherif, 1966). Thus, various methods suggest that, although both sequences operate, the more important effect is that of intergroup contact reducing prejudice.[2]

In the present analysis, we respond to this concern by concentrating on intergroup situations that severely limit choice (see Link & Cullen, 1986). By eliminating the possibility of initial attitudes leading to differential contact, such research provides a clearer indication of the causal relationship between intergroup contact and prejudice. We make use of this method in our analyses by coding samples for the extent to which participants could choose to engage in the contact. Of course, experiments limit choice through randomization of subjects to condition. But our choice rating is not simply a surrogate variable for experimental designs. Almost half of our quasi-experimental samples, and even 31 samples with weaker designs, allowed no participant choice.

The File Drawer: Publication Bias Problem

Another major potential threat to the interpretation of our results pertains to the file drawer problem that besets all literature reviews (Begg, 1994; Rosenthal, 1991). Published studies may form a

[2] See Irish (1952) and Wilson (1996) for other methods that reach the same conclusion.

biased subset of the relevant studies actually conducted, as the statistical significance of a study's results may influence the probability of it being submitted and published. Indeed, investigators have demonstrated this bias in both the social science and medical research literatures (Coursol & Wagner, 1986; Dickersin, 1997; Dickersin, Min, & Meinert, 1992; Easterbrook, Berlin, Gopalan, & Mathews, 1991; Glass, McCaw, & Smith, 1981; Lipsey & Wilson, 1993; Rotton, Foos, Van Meek, & Levitt, 1995; Shadish, Doherty, & Montgomery, 1989; Smith, 1980; Sommer, 1987). Researchers may be reluctant to send in studies with modest or countertheory findings. And journals may publish studies with large effects and reject studies with small or no effects. Thus, reviews may systematically overestimate effect sizes, as they rely heavily on published work.

Estimating publication bias is a difficult task, but numerous investigators have directed considerable attention to the problem in recent years. The many tests for bias are based on particular assumptions and consequently have their unique strengths and shortcomings (Sutton, Song, Gilbody, & Abrams, 2000). In the present research, we investigated this potentially serious problem from multiple directions by using a variety of tests and thereby extend our analysis beyond our initial findings (Pettigrew & Tropp, 2000).

First, following Rosenthal (1991), we calculated a fail-safe index. Though often criticized, this technique is one of the most widely used methods for crudely gauging publication bias (Sutton, Song, et al., 2000). It reveals how many missing studies that average no relationship between contact and prejudice ($Z = .00$) would be required to raise the significance levels above the .05 or .01 level of confidence. But this focus is the index's basic weakness. It assumes that the missing studies will average to no effect. Thus, the fail-safe index underestimates publication bias to the extent that the average of the missing studies runs counter to the hypothesis being tested.

Other rough, initial tests involve the relationship between sample size and effect size. We can test for the possibility of publication bias in several ways. Two preliminary methods are simply to correlate sample sizes with effect sizes and to develop a funnel graph consisting of a scatter diagram with the two variables (Light & Pillemer, 1984). A nonsignificant correlation and a symmetrical funnel graph each suggest minimal publication bias.

The funnel graph in turn relates to two additional methods of testing for publication bias. The "trim-and-fill" technique detects potentially missing studies by adjusting for funnel plot asymmetry (Duval & Tweedie, 2000a, 2000b; Sutton, Duval, Tweedie, Abrams, & Jones, 2000). The general linear model approach of Vevea and Hedges (1995) focuses on the absence of small studies. It assumes that random effects are distributed normally and that the survival probability of a given study can be described by a stop function around such critical probability points as .05 and .01. It should be noted, however, that such funnel graph methods tend to overestimate publication bias (Sterne & Egger, 2000).

Finally, we also tested directly whether this particular literature suffers from the file drawer problem. We compared the effect sizes between intergroup contact and prejudice from published sources (journals and books) and unpublished sources (graduate dissertations, conference papers, and other unpublished manuscripts). This method also makes a critical assumption, namely, that the unpublished studies that we uncovered constitute a random sample of unpublished research on the topic. For this and other reasons, the power of this direct approach, as Begg (1994) noted, is "directly proportional to the assiduousness of the search" (p. 405). Thus, we expended great effort in obtaining as many unpublished manuscripts as possible. We uncovered 88 unpublished contact studies, 70% of which are graduate dissertations.

The Generalization of Effects Problem

A third issue concerns the generalization of contact effects—an issue not fully addressed by Allport (1954; see Pettigrew, 1998). Critics generally concede that intergroup contact often leads to improved attitudes among the participants. But the critical question is whether these altered attitudes generalize beyond the immediate situation to new situations, to the entire outgroup, or even to outgroups not involved in the contact (Pettigrew, 1997, 1998). Such generalization is crucial to the useful application of contact theory. If the changes wrought by contact are limited to the particular situation and the immediate outgroup participants, the practical value of the theory is obviously severely restricted. Hence, we examine whether each test of the link between contact and prejudice involves some generalization beyond the immediate contact situation and participants.

Rigor of Research Studies

A final test of validity involves the relationship between indices of research rigor and the magnitude of the contact–prejudice effect sizes. If less rigorous research was largely responsible for the average effect size between contact and prejudice, we would hesitate to accept it as established. But if the more rigorous studies produce stronger contact effects, it would lend credibility to the results. Meta-analysts have checked on this issue with a variety of generally accepted indices of rigor. We used five rated variables: type of study, type of contact measure, type of control group, quality of the contact measure, and quality of the prejudice measure.

Study and Participant Characteristics as Moderators of Contact–Prejudice Effects

In addition to examining variables of theoretical and methodological interest, we also identify and analyze other possible sources of variability in the contact–prejudice relationship. With our extensive set of contact studies, we consider a range of study and participant characteristics that may moderate the relationship between contact and prejudice. In addition to the aforementioned indicators of research rigor, we examine contact–prejudice effects in relation to the date of publication, the study setting, and the geographical area in which the research was conducted. We also inspect contact–prejudice effects in relation to participant characteristics such as age, gender, and the types of groups involved in the contact.

Method

Inclusion Criteria

We define intergroup contact as actual face-to-face interaction between members of clearly defined groups. From this definition flow several inclusion criteria.

Criterion 1. Because our focus is on the relationship between intergroup contact and prejudice, we consider only those empirical studies in which intergroup contact acts as the independent variable and intergroup prejudice as the dependent variable. Eligible studies include both experimental studies that test for the effects of contact on prejudice and correlational studies that use contact as a correlate or predictor of prejudice.

Criterion 2. Only research that involves contact between members of discrete groups is included. This rule ensures that we examine intergroup outcomes of contact rather than interpersonal outcomes.

Criterion 3. The research must report on some degree of direct intergroup interaction. For inclusion, the intergroup interaction must be observed directly, reported by participants, or occur in focused, long-term situations where direct contact is unavoidable (e.g., small classrooms).

This third rule eliminates a variety of studies that are often cited in previous summaries of contact research. In particular, it excludes research that uses rough proximity or group proportions to infer intergroup interaction. As Festinger and Kelley (1951) made clear a half century ago, proximity is a necessary but not sufficient condition for social contact. One cannot assume contact from the opportunity for contact, such as living in an intergroup neighborhood with no report of actual interaction (e.g., Hamilton & Bishop, 1976; Hood & Morris, 2000). Our rare exceptions carefully demonstrated that the intergroup proximity correlated highly with actual contact, as it did in Deutsch and Collins's (1951) famous interracial housing study.

This rule also omits investigations that attempt to gauge contact with indirect measures such as information about an outgroup (e.g., Taft, 1959). We also excluded studies that asked about attitudes toward contact, unless the researchers directly linked such indicators to prior intergroup experience (e.g., Ford, 1941). In addition, this rule eliminates research that categorizes participants into groups that do not directly interact, as is the case in many minimal group studies (e.g., Otten, Mummendey, & Blanz, 1996; Tajfel, Billig, Bundy, & Flament, 1971).

This inclusion rule is important and differs from many prior reviews of this literature. The extensive research by Catlin (1977) on the impact of interracial dormitories at the University of Michigan illustrates the rule's operation. Catlin found only small relationships between interracial living and prejudice, but we did not enter these aggregate results into our files. However, Catlin also checked on the racial attitudes of White students who reported on whether they had Black acquaintances. We used these results because cross-racial "acquaintance" directly specifies face-to-face contact.

Criterion 4. The prejudice dependent variables must be collected on individuals rather than simply as a total aggregate outcome. Thus, studies concerning the relationships between contact and prejudice were included only if they used individuals as the unit of analysis such that prejudice scores could be examined in relation to individuals' contact experiences.

Locating Relevant Studies

To locate studies that meet our inclusion rules, we used a wide variety of search procedures. First, we conducted computer searches of the psychological (PsychLIT and PsycINFO), sociological (SocAbs and SocioFile), political science (GOV), education (ERIC), dissertation (UMI Dissertation Abstracts), and general research periodical (Current Contents) abstracts published through December 2000. These searches used 54 different search terms ranging from single words (e.g., *contact*, *interracial*) to combined terms (e.g., *age + intergroup contact*, *disabled + contact*). Across the various databases, we conducted three types of searches: by title words, by keyword, and by subject. Following the "descendancy approach" described by Johnson and Eagly (2000), we used the *Social Sciences Citation Index* to check on later citations of seminal contact studies.

To reach the "invisible college" of intergroup contact researchers, we wrote personal letters to researchers who published relevant research and sent for pertinent unpublished conference papers. Reference lists from located studies and previous reviews proved a rich source. We also repeatedly requested relevant materials through e-mail networks of social psychologists in Australia, Europe, and North America. These methods yielded 526 papers (reporting 515 studies), written between 1940 and the end of 2000, that meet our inclusion criteria. These studies yield 713 independent samples and 1,383 individual tests. Most of the studies comprise journal articles published during the past 3 decades (median year of publication = 1986). Slightly more than half of the samples (51%) focused on racial or ethnic target groups. In addition to conducting analyses for the full set of samples, we conducted analyses with this subset exclusively as well as with the remaining subset that used nonracial and nonethnic targets. Whereas 38 different countries contribute data, the United States provides 71% of the studies. Survey and field research constitute 71% of the studies, quasi-experiments 24%, and true experiments 5%. The research typically used college students or adult participants of both sexes who reported on their intergroup contact.

Several prototypical studies predominate. The principal prototype is a questionnaire or survey study. This research uses retrospective reports of personal contact with the outgroup. A quite different prototype uses experiments that involve an intergroup contact intervention and checks to see whether the treatment influences the participants' prejudice. A final prototype uses various quasi-experimental designs. These studies mirror the experimental research, but they lack randomization of participants to condition in between-groups designs. These three basic prototypes shaped the research moderators and other techniques that we use in the meta-analysis.

Computation and Analysis of Effect Sizes

Whereas most past meta-analytic studies in social psychology have used fixed effects models (Johnson & Eagly, 2000), the nature of the contact research literature makes a random effects model more appropriate for our analysis.[3] This model holds that part of the differences in effects across studies is essentially random and involves unidentifiable sources (Hedges, 1994; Hedges & Olkin, 1985; Lipsey & Wilson, 2001; Mosteller & Colditz, 1996; Raudenbush, 1994; Rosenthal, 1995). As Cook et al. (1992) concluded, this approach is "particularly attractive when considering (1) studies that are quite heterogeneous, (2) treatments that are ill-specified, and/or (3) effects that are complex and multidetermined" (p. 310). Because all three of these conditions hold for much of the intergroup contact literature, we used a random effects model. An additional advantage of using the random effects approach is that it allows our findings to be generalized to other contact studies beyond those used in our analysis. This choice is a conservative procedure that markedly reduces the probability levels obtained.

We report Pearson's r as our indicator of effect size throughout our analysis. If the studies do not report this measure, we derived it from other statistics by using the conversion formulas provided by Johnson (1993). A negative mean effect size indicates that greater intergroup contact is associated with lower prejudice.

If researchers reported a relationship as nonsignificant, we conservatively assigned a value of .00 for the effect size. Rosenthal (1995) has questioned this procedure as too conservative and likely to yield misleadingly low effect size estimates. He has recommended that meta-analysts conduct principal analyses both with and without these studies—a procedure that we follow in our summary analyses, although only 17 (2.4%) of our samples are involved. We also follow the suggestion of Johnson and Eagly (2000) and omit these samples when we fit our moderator models to the effect sizes. The small number of samples involved means that they affect only slightly the total effect size.

[3] Although a fixed effects approach was used in our preliminary analyses (Pettigrew & Tropp, 2000), we have now revised our analytic strategy to use a random effects model.

We used two primary units of analysis: independent samples and individual tests. The use of independent samples is helpful because they are more numerous than are studies and allow detailed comparisons. Tests are especially numerous and allow us to compare effect sizes for even more detailed factors. But multiple tests reported for a single sample violate assumptions of independence, and we have an average of almost two tests per sample. Thus, we used tests as our unit of analysis for just one potential moderator (type of generalization) that can be measured at this level only.

For each category of effect sizes, we calculated the mean effect size, weighted for sample size. Thus, the larger and more reliable samples contribute proportionately more to the mean. The homogeneity of each set of effect sizes was examined by calculating the homogeneity statistic Q that has an approximate chi-square distribution with $k-1$ degrees of freedom, where k is the number of effect sizes (Hedges & Olkin, 1985).

Weighting by sample size, however, posed a problem. Five especially large studies constitute only 1% of our study file but 28% of the total number of participants. To keep these few studies from having such enormous influence on the results, we capped their study sizes at 5,000, sample sizes at 3,000, and tests at 2,000 participants. Whereas these caps are arbitrary, only seven samples and 17 tests are involved. As Table 1 shows, omitting the studies that report only "nonsignificant" effects and applying these sample size caps results in only trivial differences of the average effect size across studies, samples, and tests for both fixed effects models and random effects models.

Ratings for Studies and Samples

We rated and recorded 16 separate variables at their appropriate levels of analysis. Tables 2–5 and 11–12 provide the particular categories used for the ratings of each of these variables, and the Appendix supplies the ratings for each sample.

Ratings of study characteristics include the date and source of publication and whether it used a within- or between-subjects design. In addition, we rated research quality in several ways. We established ratings for the type of study (e.g., survey, quasi-experiment, or experiment). We also rated the type of control group used in studies with between-subjects designs (e.g., no, some, or considerable earlier contact with the outgroup). We found that many of these control groups actually had prior contact with the target group; such "leakage" obviously renders these groups as less adequate controls. Another indicator of research rigor checked on the type of contact measure used (e.g., whether the contact was assumed from long-term situations in which contact was unavoidable, reported by the participants, or directly observed by the researchers). The quality of the measures for contact and prejudice was rated according to whether they consisted of either a single item, a multi-item scale with unreported or low reliability ($\alpha < .70$), a multi-item scale with high reliability ($\alpha \geq .70$), an experimental manipulation (for contact indicators only), or other forms.

The ratings of participant characteristics included age, sex, and geographical area of the study and the kind of target group involved (e.g., racial or ethnic, elderly, mentally ill). Additional ratings focused on the contact situation, including the setting of the contact (e.g., educational, residential, laboratory) and whether participants had any choice in participating in the contact. Another rating concerned the type of generalization involved (e.g., to outgroup members within the situation, to the whole outgroup, across situations, or to other outgroups).

Two independent judges achieved kappas above .80 for all ratings of these variables (after we collapsed error-prone adjacent rating categories). Thus, if a four-category variable had a disproportionately large number of

Table 1
Summary of Effect Sizes for Contact and Prejudice

Level of analysis	*r*	95% CL	*Z*	*k*	*N*
Studies					
All studies					
Fixed	−.225	−.23/−.22	−113.96***	515	250,089
Random[a]	−.205	−.22/−.19	−27.12***	515	250,089
With data corrections[b]					
Fixed	−.209	−.21/−.20	−94.92***	500	202,742
Random[a]	−.210	−.22/−.20	−28.93***	500	202,742
Samples					
All samples					
Fixed	−.225	−.23/−.22	−114.15***	713	250,089
Random[a]	−.210	−.22/−.20	−31.22***	713	250,089
With data corrections[b]					
Fixed	−.210	−.21/−.21	−94.96***	696	199,830
Random[a]	−.215	−.23/−.20	−32.24***	696	199,830
Tests					
All tests					
Fixed	−.218	−.22/−.22	−154.96***	1,383	494,912
Random[a]	−.214	−.22/−.20	−39.83***	1,383	494,912
With data corrections[b]					
Fixed	−.204	−.21/−.20	−127.15***	1,365	381,723
Random[a]	−.217	−.23/−.21	−38.31***	1,365	381,723

Note. These analyses were conducted with Fisher's z-transformed r values. Mean effects and confidence limits listed in this table have been transformed back to the r-metric from the z-transformed estimates obtained in these analyses. r = correlation coefficient representing the mean effect size; 95% CL = the 95% confidence limits of r; $Z = z$ test for the mean effect sizes; p = probability of z test; k = number of samples associated with the mean effect size; N = total number of participants.
[a] Random effects variance components (based on Fisher's z-transformed r values) ranged from .019 to .024 for studies and samples and from .030 to .036 for tests.
[b] Data corrections involved capping especially large numbers of participants (5,000 for studies, 3,000 for samples, 2,000 for tests) and excluding 15 nonsignificant studies from the analysis.
*** $p < .001$.

errors at Categories 2 and 3, we combined Categories 2 and 3 to form a three-category rating. The median kappa for all ratings was .86. All discrepancies between the raters were resolved through further discussion.

We also conducted ratings to examine whether the contact situation approached the optimal context specified by Allport's key conditions. We began by attempting to rate each of Allport's four conditions individually for each study. However, this procedure proved impossible given the lack of information about situational characteristics provided by the vast majority of our 515 studies. Reliable ratings were largely unattainable for all but a subset of the studies that directly addressed the characteristics of the contact.

Consequently, we conducted comparisons by using a global measure of Allport's contentions. This procedure actually offers a more direct test of the original theory, as Allport advanced his four conditions as a necessary package for positive contact effects rather than as a listing of variables that must be considered individually. In particular, we rated for all samples whether the contact situations involved structured programs designed to approximate Allport's optimal conditions.[4]

Next, for the 134 samples with contact in the context of structured programs, we attempted to conduct more fine-grained ratings for each of Allport's conditions. Though these research studies often implemented the conditions together, we used "yes" and "no" ratings to discern whether the program clearly and explicitly (a) focused participants on common goals, (b) emphasized a cooperative environment, (c) presented the groups with equal status, and (d) demonstrated authority sanction for the contact. Ratings of these variables by two independent judges yielded kappas between .76 and .97, with a median kappa of .84.

Results

Examining the Overall Pattern of Effects: Does Intergroup Contact Reduce Prejudice?

Table 1 reveals the inverse association between intergroup contact and prejudice for all studies, samples, and tests for both fixed effects analyses and random effects analyses. The mean estimates for the contact–prejudice effect size are consistent across units of analysis, data corrections, and types of analysis. With random effects analysis, the 515 studies, 713 samples, and 1,383 tests yield mean *r*s that range from −.205 to −.214. Our data corrections of eliminating the studies, samples, and tests that did not provide precise effect sizes and capping the largest studies, samples, and tests made little difference. It is these files—with the 17 "nonsignificant" samples removed and the largest samples and tests capped—that we use in the following analyses. The fixed effects analyses and random effects analyses show similar mean effect sizes, although, as expected, the *Z*s of the random effects analyses are sharply reduced. We use the random effects model for all subsequent analyses. In sum, the initial answer to our query is that intergroup contact generally relates negatively and significantly to prejudice.[5]

Though in most empirical contexts, psychologists would consider this effect size to be "small" to "medium" in magnitude (Cohen, 1988), we should emphasize several points. First, given the large number of samples, the effect is highly significant ($p <$ .0001). Second, 94% of the samples show an inverse relationship between contact and prejudice. A scatter plot of the effect sizes by sample size reveals that the effects center around an average *r* of approximately −.21, which corresponds closely to the overall mean effect size (see Figure 1). Finally, we later note markedly higher mean effect sizes for subsets of samples from rigorous studies.

At the same time, these effect sizes are highly heterogeneous across the samples ($Q_w(695) = 4{,}990.44$, $p < .0001$). Indeed, even when we remove the results of one fifth of the samples that are the largest outliers, highly significant heterogeneity remains. Such vast heterogeneity is, of course, precisely what intergroup contact theory predicts. These studies are highly diverse as to research methods, participants, situations, and targets, all of which are potential moderators of the link between contact and prejudice. As with most meta-analyses, the ultimate thrust of our analysis is not so much the gross effect sizes but more so the moderating variables that suggest the conditions under which intergroup contact reduces prejudice. Before turning to this task, however, we first test for four threats to validity that provide alternative explanations for the findings shown in Table 1.

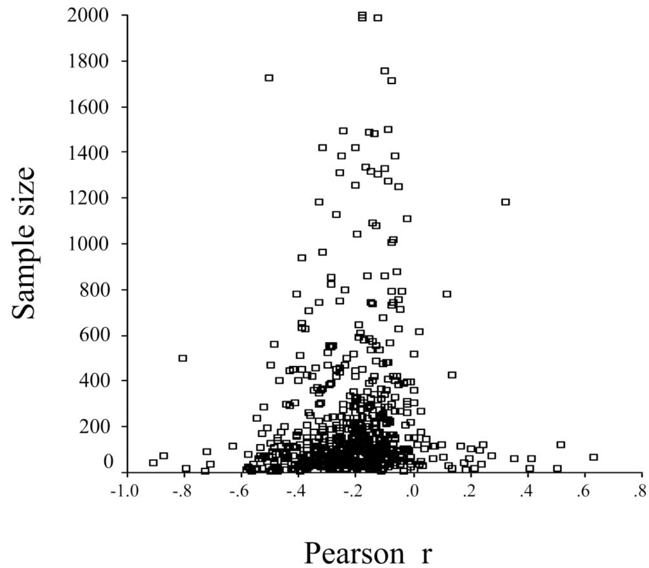

Figure 1. Scatter plot of mean contact–prejudice effect sizes (*r*) in relation to sample size.

Tests for Threats to Validity

The causal sequence problem: Examining choice to engage in contact. The negative link between contact and prejudice may largely reflect the avoidance of contact by prejudiced people. If this is so, the effect sizes for those studies that provided participants with choice as to whether to engage in the intergroup contact

[4] As a secondary, indirect indicator of Allport's conditions, we recorded whether cross-group friendship served as the measure of contact. Friendship typically involves cooperation and common goals as well as repeated equal-status contact over an extended period and across varied settings (Pettigrew, 1997). Many researchers have pointed to the role intimacy can play in reducing prejudice (Amir, 1976; Patchen, 1999; Williams, 1947), such that close, cross-group relationships may be especially likely to promote positive intergroup outcomes. Thus, together with a focus on Allport's optimal conditions for contact, we also examined the effects of cross-group friendships.

[5] With a more stringent criterion for examining contact effects, we conducted supplementary analyses that excluded the 39 cases in which some degree of intergroup contact was assumed. Mean effect sizes for the remaining cases were −.209 for studies (−.214 with data corrections) and −.221 for samples (−.220 with data corrections).

Table 2
Testing Threats to Validity for Contact–Prejudice Effect Sizes

Variable	r	95% CL	Z	k	N	Q_B
Participant choice (samples)						
No choice	−.280	−.31/−.25	−16.13***	116	15,133	
Some choice	−.190	−.21/−.17	−18.45***	279	95,267	
Full choice	−.218	−.24/−.20	−21.51***	301	89,430	
Between-classes effect						21.52***
Publication source (samples)						
Published	−.211	−.23/−.20	−28.38***	577	162,085	
Unpublished	−.237	−.27/−.21	−14.51***	119	37,745	
Between-classes effect						2.17
Type of generalization (tests)						
Within situation	−.231	−.26/−.20	−13.03***	152	31,554	
Across situations	−.244	−.33/−.15	−5.20***	17	7,553	
Whole outgroup	−.213	−.22/−.20	−36.08***	1,164	333,608	
To other outgroups[a]	−.190	−.28/−.10	−3.89***	18	3,396	
Between-classes effect						1.61

Note. These analyses were conducted with Fisher's *z*-transformed *r* values. Mean effects and confidence limits listed in this table have been transformed back to the *r*-metric from the *z*-transformed estimates obtained in these analyses. Random effects variance components (based on Fisher's *z*-transformed *r* values) ranged from .022 to .023 for analyses based on samples and was .032 for the analysis based on tests. As in Table 1, r = correlation coefficient representing the mean effect size; 95% CL = the 95% confidence limits of *r*, Z = z test for the mean effect sizes; p = probability of z test; k = number of samples associated with the mean effect size; N = total number of participants; Q_B = between-classes test of homogeneity.
[a] Homogeneity can be obtained with less than 20% of the cases trimmed.
*** $p < .001$.

should reveal larger effect sizes than do those that provided no such choice. In other words, only the studies with choice risk having a participant selection bias. Table 2 provides the results relevant to this issue.

Note that the no-choice samples provide a significantly larger mean effect size (mean $r = -.280$) than do those samples in which participants had some choice (mean $r = -.190$), $Q_B(1) = 20.58$, $p < .0001$, or full choice (mean $r = -.218$), $Q_B(1) = 8.98$, $p < .01$. The fact that the no-choice studies were, in general, of higher quality magnifies this difference between these three types of studies. The basic correlation between choice and effect size, $r = -.086$, $p < .05$, becomes nonsignificant, $r = .005$, $p = .89$, when we partial out four indicators of research quality.[6] But the key finding is that the studies that allow the participant selection bias to operate do not typically yield the larger effect sizes that would be predicted by participant selection bias.

The file drawer problem: The application of multiple tests. First, we apply Rosenthal's (1991) fail-safe index. According to our uncorrected effect size estimate for all samples based on the random effects model (see Table 1), it would require more than 1,200 missing samples that average no effect to erase the significance of the intergroup contact and prejudice association at the 5% level of confidence. This figure is considerably larger than the 713 samples uncovered by our intensive 6-year search. Next, we check on publication bias by determining that the relationship between sample sizes and effect sizes is not significant for either the original set of 713 samples, $r = -.02$, $p = .67$, or the 696 samples included in our analysis, $r = .04$, $p = .33$. Large samples provide more reliable results, and this lack of a relationship between sample size and effect size is a crude indicator of limited publication bias.

Figure 1 provides a scatter diagram using the two variables. The graph roughly resembles a funnel, as is suggested by Light and Pillemer (1984). Most important, the funnel is not sharply skewed, and the mean effect remains approximately the same regardless of sample size. Hence, the mean ($r = -.216$), median ($r = -.205$), and mode ($r = -.210$) of the distribution of samples are similar. The more symmetrical the funnel, the more it suggests that publication bias is not a major problem with this dataset.

Duval and Tweedie's (2000a, 2000b) "trim-and-fill" method was used to adjust for missing studies by focusing on funnel plot asymmetry. With Z as the effect size and with the random effects model, Duval used her technique to estimate that about 72 (10.3%) samples were missing. When she filled in for these missing data, the effect size estimate increased to a Z of $-.245$, with 95% confidence limits of $-.258$ to $-.231$. This result suggests a mean effect that is comparable to those reported in Table 1.

Contradicting these results in part is an analysis that uses Vevea and Hedges's (1995) general linear model approach. Vevea kindly conducted this analysis for us and found that the sample file was missing numerous small studies with small effects. After adjustment for these cases, his method did not find a significant overall relationship between contact and prejudice ($Z = -.02$, *ns*). However, for those 118 samples with between-groups designs and strong controls, the adjusted effect size did reach statistical significance (mean $Z = -.109$, one-tailed, $p < .02$).

Finally, as a direct test for publication bias, we compare (in Table 2) the negative mean effect sizes between intergroup contact and prejudice from published sources (journals and books) and unpublished sources (dissertations, conference papers, and other unpublished manuscripts). Note that the unpublished work has a

[6] The four rated variables for research quality included ratings of study type, quality of the contact and prejudice measures, and the quality of the control groups, as detailed later in the text.

Table 3
Type of Study, Control Group, and Contact Measure as Moderators for Contact–Prejudice Effect Sizes

Variable	r	95% CL	Z	k	N	Q_B
Type of study (samples)						
Surveys and field studies	−.204	−.22/−.19	−26.53***	492	180,386	
Quasi-experiments[a]	−.237	−.27/−.21	−15.64***	168	16,497	
Experiments[a]	−.336	−.40/−.27	−9.94***	36	2,947	
Between-classes effect						18.51***
Type of control group (samples)						
Within design	−.221	−.24/−.20	−23.58***	365	116,091	
No contact control	−.244	−.27/−.21	−15.60***	119	33,817	
Some contact control[a]	−.209	−.24/−.18	−14.69***	156	35,155	
Extensive contact control[a]	−.138	−.18/−.09	−5.96***	56	14,767	
Between-classes effect						15.78***
Type of contact measure (samples)						
Observed contact[a]	−.246	−.27/−.22	−19.50***	249	25,247	
Self-reported contact	−.210	−.23/−.19	−25.29***	408	162,292	
Assumed contact	−.132	−.18/−.08	−4.75***	39	12,291	
Between-classes effect						16.44***

Note. These analyses were conducted with Fisher's z-transformed r values. Mean effects and confidence limits listed in this table have been transformed back to the r-metric from the z-transformed estimates obtained in these analyses. Random effects variance components (based on Fisher's z-transformed r values) were .022 for each analysis. r = correlation coefficient representing the mean effect size; 95% CL = the 95% confidence limits of r, Z = z test for the mean effect sizes; p = probability of z test; k = number of samples associated with the mean effect size; N = total number of participants; Q_B = between-classes test of homogeneity.
[a] Homogeneity can be obtained with less than 20% of the cases trimmed.
*** $p < .001$.

slightly larger mean effect size between contact and prejudice (mean $r = -.237$) than does published work (mean $r = -.211$) although this difference is not significant, $Q_B(1) = 2.17$, $p = .14$.

Thus, all but one of our indicators suggest that a file drawer publication bias does not pose a major threat to the results of Table 1. However, the one notable exception—the results of Vevea and Hedges's (1995) test—lends caution in interpreting the following findings. But even this test uncovers a significant relationship between intergroup contact and diminished prejudice in studies that use between-groups designs with strong controls.

The generalization of effects problem. Do contact effects extend beyond the immediate situation? The summary results shown in Table 2 provide an affirmative answer. A total of 152 tests examined effects within the contact situation and focused exclusively on outgroup members directly involved in the contact. As shown in Table 2, their average effects correspond closely with the mean effects of our full analysis (mean $r = -.231$).

Most of the tests, however, concerned generalized effects of contact on prejudice toward the entire outgroup. These 1,164 tests provide an average effect that is not significantly weaker than the effects obtained for individual outgroup members within the contact situation (mean $r = -.213$), $Q_B(1) = .94$, $p = .33$. In addition, only 17 of the tests, drawn from nine samples, checked on contact's effects on prejudice across situations (but see Gathing, 1991; Nesdale & Todd, 1998, 2000). These few tests rendered considerable generalization (mean $r = -.244$). Finally, 18 additional tests checked on contact effects on prejudice toward outgroups not involved in the contact. This rarely considered form of generalization also operates (mean $r = -.190$).[7] Taken together, these results suggest a far wider generalization net of contact effects than is commonly thought.

Research rigor: Examining multiple tests. An additional test of validity involves the relationship between indices of research rigor and the magnitude of the contact–prejudice effect sizes. Results from five rated variables reveal that greater research rigor is routinely associated with larger effect sizes. Put differently, the less rigorous studies sharply reduce the overall relationships observed between contact and prejudice.

Study type. One measure of research rigor involves the type of study. Table 3 shows that samples tested with true experiments (mean $r = -.336$) yield significantly larger effects than do those tested with either quasi-experiments (mean $r = -.237$), $Q_B(1) = 6.72$, $p < .01$, or surveys and field studies (mean $r = -.204$), $Q_B(1) = 15.99$, $p < .001$. Note that contact's effects on prejudice in experiments ($r = -.336$) approach what Cohen (1988) described as a "large" effect size for psychological data ($d = -.713$).

In addition to demonstrating differences in effect sizes associated with research rigor, this result is relevant to the causal sequence problem discussed previously. True experiments, with their random assignment of participants to condition, remove the possibility of a selection bias operating in those who participate in intergroup contact.

Quality of control groups used. Another indicator of research rigor concerns the quality of control groups used in the research with between-subjects designs. Table 3 shows that for the samples with between-subjects designs, the less contact the control group had with the target outgroup prior to the study, the larger the mean effect sizes. Thus, the samples with control groups that had no

[7] We excluded 14 other tests from one study that attained even larger effects because it used "intergroup friends" as its contact measure (Pettigrew, 1997).

Table 4
Quality of Contact and Prejudice Indicators as Moderators for Contact–Prejudice Effect Sizes

Variable	r	95% CL	Z	k	N	Q_B
Quality of contact measure (samples)						
Single item	−.195	−.22/−.17	−14.95***	151	64,927	
Multiple items ($\alpha < .70$)	−.195	−.22/−.17	−16.31***	182	72,187	
Multiple items ($\alpha \geq .70$)	−.298	−.33/−.26	−14.59***	60	22,289	
Experimental manipulation	−.295	−.33/−.26	−17.08***	129	10,168	
Other	−.175	−.20/−.15	−12.64***	174	30,259	
Between-classes effect						54.94***
Quality of prejudice measure (samples)						
Single item[a]	−.233	−.28/−.18	−8.83***	44	11,508	
Multiple items ($\alpha < .70$)	−.190	−.21/−.17	−21.05***	384	110,407	
Multiple items ($\alpha \geq .70$)	−.246	−.27/−.22	−22.13***	241	76,469	
Other[a]	−.293	−.37/−.22	−7.23***	27	1,446	
Between-classes effect						20.86***

Note. These analyses were conducted with Fisher's z-transformed r values. Mean effects and confidence limits listed in this table have been transformed back to the r-metric from the z-transformed estimates obtained in these analyses. Random effects variance components (based on Fisher's z-transformed r values) ranged from .020 to .022 for each analysis. r = correlation coefficient representing the mean effect size; 95% CL = the 95% confidence limits of r, Z = z test for the mean effect sizes; p = probability of z test; k = number of samples associated with the mean effect size; N = total number of participants; Q_B = between-classes test of homogeneity.
[a] Homogeneity can be obtained with less than 20% of the cases trimmed.
*** $p < .001$.

prior outgroup contact (mean $r = -.244$) had a higher mean effect than did samples with controls that had either some prior outgroup contact (mean $r = -.209$), $Q_B(1) = 2.77$, $p = .09$, or extensive prior outgroup contact (mean $r = -.138$), $Q_B(1) = 10.71$, $p = .001$. In addition, samples with within-subject designs had an average effect size (mean $r = -.221$) that did not differ significantly from that of all between-subjects samples combined (mean $r = -.217$), $Q_B(1) = 0.07$, $p = .79$.

Type of contact measure. Table 3 shows differences in mean effects between samples with contrasting contact measures. Samples with directly observed contact yield the highest mean effect (mean $r = -.246$). Significantly smaller effects were obtained from samples that used self-report measures of contact (mean $r = -.210$), $Q_B(1) = 6.39$, $p = .01$, or assumed contact from a close, ongoing situation in which some degree of contact was unavoidable (mean $r = -.132$), $Q_B(1) = 11.82$, $p < .001$.[8]

Quality of contact and prejudice measures. The quality of the contact and prejudice indicators is highly influential. Multiple-item measures with low or unreported reliabilities render weaker effects than do other measures. This finding is important because, as shown in Table 4, contact researchers have often used these measures. Moreover, for contact indicators, the samples tested with reliable multiple-item measures or experimentally manipulated contact (mean $r = -.296$) provide significantly larger effect sizes than do those with other measures combined (mean $r = -.189$), $Q_B(1) = 53.22$, $p < .0001$. For prejudice indicators, the samples tested with unreliable multiple-item measures provide smaller effects (mean $r = -.190$) than does each of the other types of measures: single items (mean $r = -.233$), $Q_B(1) = 2.89$, $p < .09$, reliable multi-item scales (mean $r = -.246$), $Q_B(1) = 16.38$, $p < .0001$, and other reliable measures of the dependent variables (mostly high interrater reliability; mean $r = -.293$), $Q_B(1) = 6.60$, $p < .01$. Note also in Table 4 that the quality of the contact measures is more closely related to the effect sizes than is the quality of the prejudice measures.

Evaluating the Role of Allport's Conditions

Having addressed the major threats to validity, we can proceed with an investigation of more specific questions relevant to our research goals. Of particular interest are tests of whether Allport's stated conditions contribute to positive contact outcomes and whether such conditions are necessary for positive outcomes to occur.

Global test: Structured optimal contact. Our global predictor involves the issue of whether the contact consisted of a structured program that the researchers designed to establish Allport's optimal conditions in the contact situation. Table 5 shows that the 134 samples with optimal contact conditions yield significantly greater reductions of prejudice (mean $r = -.287$) than do the other samples (mean $r = -.204$), $Q_B(1) = 20.19$, $p < .0001$.[9]

Is this result largely a function of Allport's optimal conditions, or does it merely reflect other aspects of this subset of contact research? We addressed this question by conducting a regression analysis that includes as predictors the structured program test of Allport's conditions and our strongest methodological moderators: the type of study, the quality of the contact and prejudice measures,

[8] It should be noted that ratings for type of contact measure are strongly associated with ratings for the type of study, $r = .66$, $p < .001$.

[9] In addition, a less direct test of Allport's conditions involves tests for intergroup friendship. Only 4 of the 134 samples that experienced optimal structured contact used friends as the measure of contact. Yet, paralleling the findings for optimally structured contact, the 154 tests that used intergroup friendship as the measure of contact (mean $r = -.246$) showed a significantly stronger effect than did the remaining 1,211 tests (mean $r = -.212$), $Q_B(1) = 4.42$, $p < .05$.

Table 5
Structured Programs as a Moderator for Contact–Prejudice Effect Sizes

Variable	r	95% CL	Z	k	N	Q_B
Structured programs (samples)						
Program[a]	−.287	−.32/−.25	−16.09***	134	10,400	
No program	−.204	−.22/−.19	−28.11***	562	189,430	
Between-classes effect						20.19***

Note. These analyses were conducted with Fisher's z-transformed r values. Mean effects and confidence limits listed in this table have been transformed back to the r-metric from the z-transformed estimates obtained in these analyses. The random effects variance components (based on Fisher's z-transformed r values) was .022 for this analysis. r = correlation coefficient representing the mean effect size; 95% CL = the 95% confidence limits of r, Z = z test for the mean effect sizes; p = probability of z test; k = number of samples associated with the mean effect size; N = total number of participants; Q_B = between-classes test of homogeneity.
[a] Homogeneity can be obtained with less than 20% of the cases trimmed.
*** $p < .001$.

and the adequacy of the control group.[10] Table 6 reveals significant relationships between ratings on the structured program variable and the methodological moderators. Samples with structured programs tended to use more rigorous procedures, more reliable measures, and better controls. An inverse variance weighted regression analysis was then conducted with SPSS macros, developed by Wilson (2002), which provide the appropriate parameters and probability values for meta-analytic data (see also Lipsey & Wilson, 2001).

Table 7 displays the regression results. Ratings of the quality of the contact and prejudice measures and the adequacy of the control groups all relate significantly to the magnitude of the contact–prejudice effect sizes. To further demonstrate the combined importance of these three methodological predictors, we formed a subset of 77 samples that boasted the most rigorous category for each of these variables. The mean effect for this rigorous subset ($r = -.323$) proved far stronger than did that of the remaining, less rigorous samples ($r = -.202$), $Q_B(1) = 35.96$, $p < .0001$. Thus, when properly tested with rigorous measures and research procedures, studies of contact–prejudice relationships typically yield larger effects.

Nonetheless, the structured program indicator of Allport's conditions remains a significant predictor of contact–prejudice effects ($\beta = -.099$, $p < .03$) even when entered with these methodological moderators. As such, this multivariate model provides a stronger test for Allport's theory of intergroup contact than do the univariate comparisons for structured programs. Still, mean comparisons reported in Table 5 indicate that the inverse relationship between contact and prejudice persists—though not as strongly—even when the contact situation is not structured to match Allport's conditions.

Specific tests of individual conditions. We conducted a series of tests with ratings of individual contact conditions for the 134 samples with structured programs. These cases were rated as having authority sanction, an unsurprising finding that was virtually assured by the implementation of programs designed to promote Allport's conditions.

As a first step, we conducted mean comparisons between samples that were rated as with or without each of the three remaining conditions (i.e., common goals, cooperation, and equal status). These tests showed no significant differences in mean contact–prejudice effects for samples rated with and without common goals, $Q_B(1) = 1.89$, $p = .17$, cooperation, $Q_B(1) = 0.03$, $p = .86$, or equal status, $Q_B(1) = 0.70$, $p = .40$. We also compared samples that included all four of Allport's conditions with those that did not include all four conditions, and we found no significant differences in mean contact–prejudice effects, $Q_B(1) = 1.48$, $p = .22$. Additional analyses indicated that ratings of common goals, cooperation, and equal status were highly correlated with each other (rs ranging from .51 to .63, $p < .001$), with 72% of the samples rated as having at least three of Allport's four optimal conditions. We then conducted inverse weighted regression analyses (see Lipsey & Wilson, 2001; Wilson, 2002) to test common goals, cooperation, and equal status as predictors for contact–prejudice effect sizes. The models revealed no significant effects for these three conditions when either entered simultaneously as predictors (βs ranging from .02 to .18, $p > .15$) or when entered separately alongside our methodological moderators (βs ranging from .05 to .06, $p > .50$).

Given that none of the three conditions emerged as a significant, independent predictor, additional analyses were conducted to examine whether authority sanction might play a special role in predicting the contact–prejudice effect sizes. For this analysis, samples rated as having only authority sanction ($k = 31$) were compared with samples rated as having two or more of Allport's conditions ($k = 103$) as well as with the remaining samples in our analysis ($k = 564$). Results show that the mean effect for samples with only authority sanction (mean $r = -.286$) did not differ significantly from the mean effect for samples with two or more of Allport's conditions (mean $r = -.290$), $Q_B(1) = 0.01$, $p = .93$, whereas both of these groups showed significantly stronger effects than did the remaining samples in our analysis (mean $r = -.204$), $Q_B(1) = 6.10$, $p < .05$, and $Q_B(1) = 16.18$, $p < .001$.

Subset Analyses for Racial or Ethnic Samples and Other Samples

To check for consistency in general patterns of effects, we conducted additional analyses examining contact–prejudice rela-

[10] For the regression analyses, ratings of the quality of the contact and prejudice measures were dichotomized such that ratings would indicate either high reliability (e.g., multi-item scale with high reliability, experimental manipulation, high interrater reliability) or low reliability (e.g., single-item measure, multi-item measure with low or unknown reliability). Ratings of the control measure were trichotomized: (a) the control group had no prior contact or the sample used a within-subject design, (b) the control group had some prior contact, or (c) the control group had extensive prior contact with the outgroup.

Table 6
Correlation Matrix and Descriptive Statistics of Predictor Variables

Predictor variable	1	2	3	4	5
1. Type of study (3)	—				
2. IV quality (2)	.539***	—			
3. DV quality (2)	.016	.219***	—		
4. Type of control (3)	.058	−.009	−.106**	—	
5. Program (2)	.570***	.390***	.102**	.095*	—
M	1.34	1.27	1.39	1.39	1.19
SD	.57	.45	.49	.63	.40

Note. Numbers in parentheses represent the number of levels for each variable. For type of study, 1 = survey or field study, 2 = quasi-experiment, 3 = experiment; for independent variable (IV) and dependent variable (DV) quality, 1 = other, 2 = reliable indicator; for type of control, 1 = within-subjects design or between-subjects design with no prior contact, 2 = some prior contact, 3 = considerable prior contact; for program, 1 = no structured program, 2 = structured program.
* $p < .05$. ** $p < .01$. *** $p < .001$.

tionships across two subsets of cases. As approximately half of the samples in our analysis (51%) involved contact between racial and ethnic groups, we analyze these cases and the remaining cases as two separate subsets. Contact theory was originally developed to address racial and ethnic prejudices, but recent decades have witnessed a massive use of the theory for a range of different target groups. Is this expansion of contact theory justified? And do these nonracial and nonethnic samples yield meta-analytic patterns that are similar to those for racial and ethnic samples?

Comparisons across the racial and ethnic subsets and the nonracial and nonethnic subsets yield virtually identical mean estimates of contact–prejudice effect sizes (mean $r = -.218$ and $-.220$, respectively), $Q_B(1) = 0.027$, $p = .87$. Table 8 presents results for each subset in relation to our four strongest methodological moderators. Higher quality of the contact and prejudice measures tend to show larger average effect sizes for samples in both subsets. At the same time, study type and type of control group proved especially important for the nonracial and nonethnic samples, whereas quality of the prejudice measures proved particularly important for the racial and ethnic samples. Overall, however, the patterns of results observed for these subsets largely reflect those obtained in the full analysis.

We then examined contact–prejudice effects for each subset in relation to the global indicator of Allport's conditions (see Table 9). Structured programs developed in line with Allport's conditions enhanced contact–prejudice effects for both subsets of cases, though the effects tended to be stronger among the nonracial and nonethnic samples, $Q_B(1) = 19.67$, $p < .001$, than among the racial and ethnic samples, $Q_B(1) = 2.62$, $p = .11$. At the same time, no significant differences in mean contact–prejudice effects emerged between structured program samples with racial and ethnic targets and nonracial and nonethnic targets, $Q_B(1) = 1.23$, $p = .27$.

Paralleling our analysis of the full sample, regression analyses then examined the structured program variable and four methodological moderators as predictors for contact–prejudice effects in each subset (see Table 10). These analyses reveal some variability in the degree to which the different methodological indicators predict the contact–prejudice effects. In addition, the structured program variable testing Allport's contentions consistently emerges as a marginally significant predictor of contact–prejudice effects for both the racial and ethnic samples, $\beta = -.112$, $p = .069$, and the remaining samples, $\beta = -.105$, $p = .094$.

Overall, then, results from both subsets closely resemble the findings from the full analysis. Moreover, although there are some slight differences associated with methodological factors, the preponderance of the evidence indicates similar patterns of effects across the two subsets of samples.

Table 7
Summary of Inverse Variance Weighted Regression Model Predicting Contact–Prejudice Effect Sizes

Predictor variable	*B*	*SE*	β	*Z*	*p*	Model summary
Type of study	.001	.017	.002	.035	.972	
IV quality	−.088	.018	−.206	−4.775	.000	
DV quality	−.031	.014	−.084	−2.231	.026	
Type of control	.034	.010	.121	3.303	.001	
Program	−.053	.024	−.099	−2.219	.027	
R^2						.10***
Q_{Model}						77.29***
k						696

Note. This analysis was conducted with Fisher's *z*-transformed *r* values. The random effects variance component for this analysis (based on Fisher's *z*-transformed *r* values) was .020. *B* = raw regression coefficient; *SE* = standard error for the regression coefficient; β = standardized regression coefficient; *Z* = *z* test for the regression coefficient; *p* = probability of *z* test; R^2 = proportion of variance accounted for; Q_{Model} = test of whether the regression model explains a significant portion of variability across effect sizes (see Wilson, 2002); *k* = number of samples included in the analysis.
*** $p < .001$.

Supplementary Analyses of Participant and Study Moderators

A final set of analyses examines several additional participant and study variables as potential moderators for contact–prejudice effects.[11]

Target group. Extending our analysis of the intergroup contexts under study, Table 11 presents mean effect sizes for the many types of target groups studied in the contact literature. We consistently find significant relationships between intergroup contact

[11] Comparisons of samples with and without Allport's conditions were not conducted in relation to these variables because they would have involved tests with extremely small numbers of cases.

Table 8
Indicators of Research Rigor as Moderators for Contact–Prejudice Effect Sizes Among Racial and Ethnic Samples and Nonracial and Nonethnic Samples

	Racial and ethnic samples					Nonracial and nonethnic samples				
Variable	*r*	95% CL	*Z*	*k*	Q_B	*r*	95% CL	*Z*	*k*	Q_B
Type of study										
Surveys and field studies	−.215	−.23/−.20	−22.05***	299		−.186	−.21/−.16	−15.26***	193	
Quasi-experiments	−.211	−.26/−.16	−8.25***	54		−.251	−.29/−.22	−13.49***	114	
Experiments	−.221	−.34/−.09	−3.37***	9		−.377	−.44/−.31	−9.67***	27	
Between-classes effect					0.03					27.89***
Quality of contact measure										
Single item	−.210	−.25/−.17	−10.98***	65		−.184	−.22/−.15	−10.40***	86	
Multiple items ($\alpha < .70$)	−.201	−.23/−.17	−14.34***	128		−.181	−.22/−.14	−8.29***	54	
Multiple items ($\alpha \geq .70$)	−.323	−.36/−.28	−14.00***	44		−.226	−.30/−.15	−5.76***	16	
Experimental manipulation	−.236	−.30/−.17	−6.77***	32		−.314	−.35/−.28	−15.24***	97	
Other	−.170	−.20/−.14	−9.52***	93		−.181	−.22/−.14	−8.52***	81	
Between-classes effect					31.85***					34.98***
Quality of prejudice measure										
Single item	−.235	−.29/−.18	−8.24***	34		−.225	−.34/−.11	−3.63***	10	
Multiple items ($\alpha < .70$)	−.182	−.20/−.16	−15.77***	210		−.200	−.23/−.17	−14.09***	174	
Multiple items ($\alpha \geq .70$)	−.259	−.29/−.23	−16.37***	105		−.235	−.26/−.21	−15.03***	136	
Other	−.344	−.44/−.24	−6.31***	13		−.235	−.35/−.12	−3.91***	14	
Between-classes effect					23.73***					2.97
Type of control group										
Within design	−.228	−.25/−.21	−19.59***	217		−.209	−.24/−.18	−13.68***	148	
No contact control	−.166	−.22/−.11	−6.25***	39		−.284	−.32/−.25	−14.96***	80	
Some contact control	−.220	−.26/−.18	−11.16***	73		−.197	−.23/−.16	−9.86***	83	
Extensive contact control	−.179	−.24/−.12	−5.99***	33		−.081	−.15/−.01	−2.28*	23	
Between-classes effect					6.56					29.93***

Note. These analyses were conducted with Fisher's *z*-transformed *r* values. Mean effects and confidence limits listed in this table have been transformed back to the *r*-metric from the *z*-transformed estimates obtained in these analyses. Random effects variance components (based on Fisher's *z*-transformed *r* values) ranged from .019 to .024. *r* = correlation coefficient representing the mean effect size; 95% CL = the 95% confidence limits of *r, Z* = *z* test for the mean effect sizes; *p* = probability of *z* test; *k* = number of samples associated with the mean effect size; Q_B = between-classes test of homogeneity.
* $p < .05$. *** $p < .001$.

andprejudice across contexts, though the magnitudes of the contact–prejudice effect sizes vary in relation to different target groups. The largest effects emerge for samples involving contact between heterosexuals and gay men and lesbians (mean $r = -.271$). These effects are significantly larger than are those for the other samples combined (mean $r = -.211$), $Q_B(1) = 5.34$, $p = .02$. Research focused on contact with the physically disabled (mean $r = -.243$) also provides a larger-than-average effect size. The most studied target groups, racial and ethnic groups (mean $r = -.214$), and research on contact with the mentally disabled (mean $r = -.202$) yield average effects. But research with other target groups generally produces smaller effects. In particular, samples concerning contact with the mentally ill and the elderly combined (mean $r = -.183$) render significantly lower mean effects than do the other target groups combined (mean $r = -.221$), $Q_B(1) = 4.51$, $p = .03$.

Table 9
Structured Programs as a Moderator for Contact–Prejudice Effect Sizes Among Racial and Ethnic Samples and Nonracial and Nonethnic Samples

	Racial and ethnic samples					Nonracial and nonethnic samples				
Variable	*r*	95% CL	*Z*	*k*	Q_B	*r*	95% CL	*Z*	*k*	Q_B
Program	−.262	−.32/−.20	−8.30***	40		−.299	−.34/−.26	−13.80***	94	
No program	−.210	−.23/−.19	−22.37***	322		−.194	−.22/−.17	−17.21***	240	
Between-classes effect					$Q_B(1) = 2.62$					19.67***

Note. These analyses were conducted with Fisher's *z*-transformed *r* values. Mean effects and confidence limits listed in this table have been transformed back to the *r*-metric from the *z*-transformed estimates obtained in these analyses. Random effects variance components for these analyses (based on Fisher's *z*-transformed *r* values) were .022. *r* = correlation coefficient representing the mean effect size; 95% CL = the 95% confidence limits of *r, Z* = *z* test for the mean effect sizes; *p* = probability of *z* test; *k* = number of samples associated with the mean effect size; Q_B = between-classes test of homogeneity.
*** $p < .001$.

Table 10

Summary of Inverse Variance Weighted Regression Model Predicting Contact–Prejudice Effect Sizes Among Racial and Ethnic Samples and Nonracial and Nonethnic Samples

	Racial and ethnic samples						Nonracial and nonethnic samples					
Predictor variable	*B*	*SE*	β	*Z*	*p*	Statistic	*B*	*SE*	β	*Z*	*p*	Model summary
Type of study	.062	.027	.147	2.29	.022		−.036	.022	−.112	−1.61	.108	
IV quality	−.095	.025	−.216	−3.79	.000		−.068	.027	−.162	−2.53	.011	
DV quality	−.059	.020	−.161	−2.93	.003		−.023	.020	−.059	−1.14	.256	
Type of control	−.003	.014	−.010	−.18	.857		.064	.015	.213	4.16	.000	
Program	−.073	.040	−.112	−1.82	.069		−.049	.029	−.105	−1.68	.094	
R^2						.10***						.15***
Q_{Model}						40.45***						59.74***
k						362						334

Note. These analyses were conducted with Fisher's *z*-transformed *r* values. Random effects variance components (based on Fisher's *z*-transformed *r* values) ranged from .019 to .020. *B* = raw regression coefficient; *SE* = standard error for the regression coefficient; β = standardized regression coefficient; *Z* = *z* test for the regression coefficient; *p* = probability of *z* test; R^2 = proportion of variance accounted for; Q_{Model} = test of whether the regression model explains a significant portion of variability across effect sizes (see Wilson, 2002); *k* = number of samples included in the analysis; IV = independent variable; DV = dependent variable.
*** $p < .001$.

Age. Table 11 also shows that the effects obtained with children (mean $r = -.239$) and college students (mean $r = -.231$) do not significantly differ from those obtained with adolescents (mean $r = -.208$), $Q_B(1) = 1.20$ and 1.37, respectively, $p > .20$. At the same time, effects for children are marginally stronger, $Q_B(1) = 3.59$, $p = .06$, and effects for college students are significantly stronger, $Q_B(1) = 5.49$, $p < .05$, than are those obtained for adults (mean $r = -.197$). That college students yield significantly stronger average effects than do adults is consistent with Sears's (1986) contentions that college students are generally more open to change than are older adults.

Sex. Participants' sex proves to be a minor factor in interpreting contact–prejudice effects (see Table 11). The difference between all-male and all-female samples is not significant, $Q_B(1) = 0.70$, $p = .40$.

Table 11

Participant Predictors of Contact–Prejudice Effect Sizes Across Samples

Variable	*r*	95% CL	*Z*	*k*	*N*	Q_B
Target groups						
Sexual orientation	−.271	−.32/−.22	−10.49***	42	12,059	
Physically disabled	−.243	−.28/−.21	−12.91***	93	15,584	
Race, ethnicity	−.214	−.23/−.20	−23.62***	362	133,249	
Mentally disabled[a]	−.207	−.26/−.15	−7.16***	40	6,116	
Mentally ill[a]	−.184	−.23/−.14	−8.41***	66	17,218	
Elderly	−.181	−.23/−.13	−6.73***	54	6,424	
Other[a]	−.192	−.25/−.13	−6.27***	39	9,180	
Between-classes effect						11.95
Age of participants						
Children (1–12 years)	−.239	−.28/−.20	−11.30***	82	10,207	
Adolescents	−.208	−.24/−.18	−12.68***	114	45,602	
College students	−.231	−.25/−.21	−20.50***	262	46,553	
Adults	−.197	−.22/−.18	−17.81***	238	97,468	
Between-classes effect						6.68
Sex of participants						
Females[a]	−.214	−.26/−.17	−9.06***	63	13,183	
Males[a]	−.185	−.23/−.14	−7.56***	59	15,598	
Both or undetermined	−.218	−.23/−.20	−29.58***	574	171,049	
Between-classes effect						1.83

Note. These analyses were conducted with Fisher's *z*-transformed *r* values. Mean effects and confidence limits listed in this table have been transformed back to the *r*-metric from the *z*-transformed estimates obtained in these analyses. Random effects variance components (based on Fisher's *z*-transformed *r* values) were 0.23 for each analysis. *r* = correlation coefficient representing the mean effect size; 95% CL = the 95% confidence limits of *r*; *Z* = *z* test for the mean effect sizes; *p* = probability of *z* test; *k* = number of samples associated with the mean effect size; *N* = total number of participants. Q_B = between-classes test of homogeneity.
[a] Homogeneity can be obtained with less than 20% of the cases trimmed.
*** $p < .001$.

Geographic area. With 72% of our samples conducted in the United States, it is important to determine whether there are significant differences in effect sizes in contact research conducted elsewhere. A general test across the six geographical areas revealed no significant differences in effects, $Q_B(5) = 1.88$, $p = .87$ (see Table 12). And a focused test shows that there is virtually no difference in effect sizes between U.S. (mean $r = -.215$) and non–U.S. samples (mean $r = -.217$), $Q_B(1) = .01$, $p = .90$.

Contact setting. Various research settings relate significantly to the size of the effects (see Table 12). Although there likely are differences in intensity and duration of contact among these settings, their discrepant mean effects are suggestive. The smallest mean effect results from intergroup contact through tourism and travel. Though based on only 13 samples from nine studies, this tourism effect size (mean $r = -.113$) is significantly smaller than is that of the other samples combined (mean $r = -.217$), $Q_B(1) = 3.84$, $p < .05$. By contrast, the largest mean effects emerge from contact that occurs in recreational and laboratory settings. The 48 samples studied in these settings provide a mean effect (mean $r = -.287$) that is significantly larger than that of the other settings combined (mean $r = -.211$), $Q_B(1) = 6.86$, $p < .01$.

Date of study. Though early samples studied prior to 1960 uncovered slightly larger average effects (mean $r = -.228$), the dominant trend is for recent research to reveal greater mean effects than does earlier work. Thus, the 415 samples tested after 1979 yield a significantly larger average effect (mean $r = -.236$) than do the 281 samples tested prior to 1980 (mean $r = -.184$), $Q_B(1) = 15.59$, $p < .0001$. It is tempting to speculate how major events, such as American racial conflict in the 1960s, might have shaped this difference. However, the difference across the two time periods is explained largely by the increased rigor of modern research. Relative to earlier work, contact research since 1979 has used more rigorous measures and procedures, as indicated by the quality of the contact measure, $\chi^2(1) = 13.70$, $p < .001$, the quality of the prejudice measure, $\chi^2(1) = 52.62$, $p < .001$, and the quality of the controls used in the research, $\chi^2(2) = 12.14$, $p < .01$. When these indicators of research rigor are controlled, the difference in effect sizes between the early and late intergroup contact samples is sharply reduced but remains statistically significant, $\beta = -.08$, $p < .05$.

Discussion

These meta-analytic findings shed important light on long-standing debates in the contact literature concerning the central questions of whether contact reduces prejudice and the role that Allport's conditions play in promoting positive intergroup outcomes.

Table 12
Study Predictors of Contact–Prejudice Effect Sizes Across Samples

Variable	r	95% CL	Z	k	N	Q_B
Geographic area of research						
United States	−.215	−.23/−.20	−26.81***	501	133,598	
Europe	−.217	−.25/−.18	−10.96***	80	36,799	
Israel[a]	−.196	−.26/−.13	−5.42***	24	6,808	
Canada	−.232	−.30/−.16	−6.19***	21	4,732	
Australia and New Zealand[a]	−.259	−.34/−.18	−6.11***	16	3,704	
Africa, Asia, Latin America	−.205	−.25/−.16	−8.45***	54	14,189	
Between-classes effect						1.88
Research setting						
Laboratory[a]	−.273	−.35/−.19	−6.25***	22	1,754	
Recreational[a]	−.299	−.37/−.23	−7.60***	26	2,168	
Work, organizational	−.224	−.27/−.18	−10.20***	73	16,608	
Educational	−.213	−.24/−.19	−16.72***	209	52,980	
Residential[a]	−.202	−.25/−.16	−8.48***	57	8,778	
Tourism, travel	−.113	−.22/−.01	−2.08*	13	2,211	
Mixed and other	−.213	−.23/−.19	−21.82***	296	115,331	
Between-classes effect						11.14
Date of publication						
Prior to 1960	−.228	−.27/−.19	−10.12***	57	19,667	
1960–1969[a]	−.176	−.21/−.14	−9.18***	83	16,350	
1970–1979	−.169	−.20/−.14	−11.24***	141	44,297	
1980–1989	−.233	−.26/−.21	−16.81***	165	37,217	
1990–2000	−.238	−.26/−.22	−21.82***	250	82,299	
Between-classes effect						21.15***

Note. These analyses were conducted with Fisher's z-transformed r values. Mean effects and confidence limits listed in this table have been transformed back to the r-metric from the z-transformed estimates obtained in these analyses. Random effects variance components (based on Fisher's z-transformed r values) ranged from .022 to .023 for each analysis. r = correlation coefficient representing the mean effect size; 95% CL = the 95% confidence limits of r, Z = z test for the mean effect sizes; p = probability of z test; k = number of samples associated with the mean effect size; N = total number of participants; Q_B = between-classes test of homogeneity.
[a] Homogeneity can be obtained with less than 20% of the cases trimmed.
* $p < .05$. *** $p < .001$.

Does Intergroup Contact Reduce Prejudice?

The meta-analytic results clearly indicate that intergroup contact typically reduces intergroup prejudice. Synthesizing effects from 696 samples, the meta-analysis reveals that greater intergroup contact is generally associated with lower levels of prejudice (mean $r = -.215$). Moreover, the mean effect rises sharply for experiments and other rigorously conducted studies. In addition, 94% of the samples in our analysis show an inverse relationship between intergroup contact and prejudice.

Additional findings suggest that these relationships between contact and prejudice are not artifacts of either participant selection or publication bias. Consistent with past research, the particularly strong effects observed for experimental studies confirm that contact can cause meaningful reductions in prejudice. Moreover, the investigations that allowed no choice for their participants to avoid the intergroup contact yield a slightly larger mean effect size in reducing prejudice than do studies that allowed choice. In addition, of the six tests we conducted to test for publication bias, all but one indicate that this bias is not a serious threat to the validity of our results, and the one exception still revealed a significant contact–prejudice effect among the most rigorous research studies.

Results from our analysis also show that intergroup contact effects typically generalize beyond participants in the immediate contact situation. Indeed, the generalization of contact's effects appears to be far broader than what many past commentators have thought. Not only do attitudes toward the immediate participants usually become more favorable, but so do attitudes toward the entire outgroup, outgroup members in other situations, and even outgroups not involved in the contact. This result enhances the potential of intergroup contact to be a practical, applied means of improving intergroup relations.

The findings also reveal that intergroup contact may be useful for reducing prejudice in a variety of intergroup situations and contexts. The patterns of contact–prejudice effects observed for racial and ethnic samples closely resemble those observed for the remaining samples in our analysis. Moreover, although we observe variability in the magnitude of contact–prejudice effects across different intergroup contexts, the relationships between contact and prejudice remain significant across samples involving different target groups, age groups, geographical areas, and contact settings. These results support the recent extension of intergroup contact theory to a variety of intergroup contexts, beyond its original focus on racial and ethnic groups. In sum, our meta-analytic results provide substantial evidence that intergroup contact can contribute meaningfully to reductions in prejudice across a broad range of groups and contexts.

What Role Do Allport's Conditions Play in Helping Contact to Reduce Prejudice?

Results from the meta-analysis also offer important insights regarding the role of Allport's conditions in reducing prejudice through intergroup contact. Consistent with much of the intergroup contact literature (see Allport, 1954; Pettigrew, 1998), those samples that experienced carefully structured contact situations designed to meet Allport's optimal conditions achieved a markedly higher mean effect size than did other samples. Moreover, a multivariate model shows that structured contact predicted stronger contact–prejudice effects, beyond that explained by multiple indices of research rigor. This trend emerged for racial and ethnic samples as well as for the remaining samples in our analysis. Taken together, these results show that establishing Allport's optimal conditions in the contact situation generally enhances the positive effects of intergroup contact.

At the same time, Allport's conditions are not essential for intergroup contact to achieve positive outcomes. In particular, we found that samples with no claim to these key conditions still show significant relationships between contact and prejudice. Thus, Allport's conditions should not be regarded as necessary for producing positive contact outcomes, as researchers have often assumed in the past. Rather, they act as facilitating conditions that enhance the tendency for positive contact outcomes to emerge.

Moreover, further examination of Allport's conditions suggests that institutional support may be an especially important condition for facilitating positive contact effects. Although the present analysis offers a relatively crude test, samples with structured programs showed significantly stronger contact–prejudice effects than the remaining samples, irrespective of whether they were rated as having conditions beyond authority support. At the same time, it is important to note that our ratings of authority support were conducted in the context of structured programs designed to approximate optimal conditions for positive intergroup contact. Hence, although authority support appears to play an important role, this condition should not be conceived of or implemented in isolation. Institutional support for contact under conditions of competition or unequal status can often enhance animosity between groups, thereby diminishing the potential for achieving positive outcomes from contact (see Sherif, 1966). Thus, consistent with Allport's original contentions, we believe that optimal conditions for contact are best conceptualized as functioning together to facilitate positive intergroup outcomes rather than as entirely separate factors.

Moving Toward a Reformulation of Intergroup Contact Theory

Combined with other recent empirical advances, these meta-analytic findings suggest new ways of thinking about the likely effects of intergroup contact. We posit that the process underlying contact's ability to reduce prejudice involves the tendency for familiarity to breed liking. Emphasized by Homans (1950) and demonstrated experimentally by Zajonc (1968), this phenomenon leads to the prediction that intergroup contact will induce liking under a wide range of conditions. Research has consistently found evidence for the relationship between exposure and liking with a range of targets (e.g., Bornstein, 1989; Harmon-Jones & Allen, 2001; Lee, 2001) and across varied research settings (e.g., Moreland & Zajonc, 1977; Zajonc & Rajecki, 1969). Moreover, recent work has demonstrated that the increases in liking that derive from exposure can generalize to greater liking for related, yet unknown, targets (Rhodes et al., 2001); this is comparable to the generalization of contact's effects to unknown outgroup members.

These mere exposure findings also help to explain why Allport's optimal conditions prove not to be essential for the positive effects of contact to emerge. Although 94% of the 713 samples in our analysis showed an inverse relationship between intergroup contact and prejudice, only 19% of the samples involved contact situations structured in line with Allport's conditions.[12] Consider two relevant examples: Van Dyk (1990) found that rural Afrikaans-speaking White housewives who had close contact with their African domestic workers had more favorable attitudes toward Africans in general ($r = -.09$). Conducted during the tense final days of South Africa's apartheid policy, this contact situation sharply violates Allport's key conditions. Likewise, Crain and Weisman (1972) found that adult African Americans who reported having played with Whites as children were less anti-White ($r = -.08$), although they had experienced racially segregated neighborhoods and elementary schools. Like these examples, many of the meta-analysis' studies conspicuously lack Allport's key conditions for positive contact outcomes and yet report some reduction in prejudice.

In turn, these trends beg the following question: If Allport's optimal conditions are not essential for achieving positive intergroup outcomes, then what might be necessary? An answer to this central question is forming from the confluence of several new lines of contemporary research.

Work on the relationship between familiarity and liking suggests that uncertainty reduction is an important mechanism underlying these relationships (e.g., Lee, 2001). Complementing this view, emerging perspectives have pointed to the significance of reducing intergroup anxiety to achieve reductions in prejudice from contact (Dijker, 1987; Islam & Hewstone, 1993; Stephan & Stephan, 1985; Stephan et al., 2002). Intergroup anxiety refers to feelings of threat and uncertainty that people experience in intergroup contexts. These feelings grow out of concerns about how they should act, how they might be perceived, and whether they will be accepted (Stephan & Stephan, 1985; see also Berger & Calabrese, 1975; Blascovich, Mendes, Hunter, & Lickel, 2000; Gudykunst, 1985; Mendes, Blascovich, Lickel, & Hunter, 2002). Indeed, Stephan, Stephan, and Gudykunst (1999) have begun the task of combining the uncertainty reduction and threat reduction theories.

A rapidly growing research literature supports this fresh perspective. Studies have shown repeatedly that contact can reduce feelings of threat and anxiety about future cross-group interactions (Blair, Park, & Bachelor, 2003; Blascovich, Mendes, Hunter, Lickel, & Kowai-Bell, 2001; Islam & Hewstone, 1993; Paolini, Hewstone, Cairns, & Voci, 2004; Stephan & Stephan, 1985). Moreover, recent studies have demonstrated that intergroup anxiety mediates the relationships between intergroup contact and prejudice (e.g., Paolini et al., 2004; Stephan et al., 2002; Voci & Hewstone, 2003). Thus, more positive contact outcomes can be achieved to the extent that anxiety is reduced (Brown & Hewstone, 2005). Reducing negative feelings such as anxiety and threat represents an important means by which intergroup contact diminishes prejudice.[13]

Directions for Future Research

These findings, along with recent work on familiarity and liking, suggest a new orientation for future theory and research on intergroup contact. In particular, social psychologists must grant greater attention to the negative factors that deter intergroup contact from diminishing prejudice. When Williams (1947) and Allport (1954) were fashioning intergroup contact theory, they assumed that most contact did not reduce prejudice. Hence, they sought to specify the positive features of those contact situations that could maximize the potential for contact to reduce prejudice and promote positive intergroup outcomes. Ever since, explorations of contact theory have focused largely on positive factors. But the meta-analytic data reveal that the knowledge gained from past contact research is limited by its primary emphasis on positive features of the contact situation. Factors that curb contact's ability to reduce prejudice are now the most problematic theoretically, yet the least understood. These negative factors, ranging from intergroup anxiety (Stephan & Stephan, 1985) to authoritarianism and normative restraints (Pettigrew, Wagner, Stellmacher, & Christ, 2006), deserve to become a major focus of future contact research. Such an emphasis would allow a more comprehensive understanding of conditions that both enhance and inhibit the potentially positive effects of contact.

New developments also suggest that the effects of these factors are likely to be moderated by the degree to which group membership is salient during contact. Voci and Hewstone (2003) have shown that anxiety mediates the relationship between contact and prejudice when group salience is high but that such mediation is less pronounced when group salience is low.

Other studies have demonstrated that contact effects are more likely to generalize when group membership is salient (Brown & Hewstone, 2005). Indeed, this Hewstone and Brown (1986) contention may explain why the meta-analytic results reveal such widespread generalization. It is likely that the demands of the contact research situation (or the need for reflection by those reporting on past contact) led to high group salience in most of the studies.

These advances raise the possibility of the development of a model considerably more complex and complete than Allport's original "contact hypothesis." Contemporary research has examined a range of additional mediators of contact effects, including perspective taking (Craig, Cairns, Hewstone, & Voci, 2002), broadened views of the ingroup (e.g., Gaertner & Dovidio, 2000; Pettigrew, 1998; Sherif, 1966), and the perceived importance of the contact (Van Dick et al., 2004). The search for mediators has also involved an expanded investigation of contact effects. Beyond the influence of contact on prejudice, researchers have tested the effects of intergroup contact on such variables as intergroup differentiation and outgroup variability (Islam & Hewstone, 1993; Oaker & Brown, 1986; Paolini et al., 2004), ingroup pride (London & Linney, 1993), and a willingness to trust and forgive the outgroup (Hewstone et al., 2005).

[12] It is possible that this result could reflect a selection bias involving the intergroup situations researchers choose to study. But this type of situational selection bias appears highly unlikely. Our file contains many studies, such as the examples just described, where the contact situation is far less than optimal. More important, most of the studies in this meta-analysis involve survey and questionnaire research. Here, the subjects report on whatever intergroup contact they have had. Thus, there is limited information regarding the contact conditions, and the researchers had no control over the situations involved.

[13] Not all contact experiences are positive, of course. Although most of the contact studies in our analysis focused on positive contact outcomes, some recent work has shown that negative intergroup experiences can enhance feelings of anxiety and threat and hinder the development of positive orientations toward the outgroup (Plant, 2004; Plant & Devine, 2003; Stephan & Stephan, 1985; Tropp, 2003).

Given the current state of the research literature, there is little need to demonstrate further contact's general ability to lessen prejudice. Results from the meta-analysis conclusively show that intergroup contact can promote reductions in intergroup prejudice. Moreover, the meta-analytic findings reveal that contact theory applies beyond racial and ethnic groups to embrace other types of groups as well. As such, intergroup contact theory now stands as a general social psychological theory and not as a theory designed simply for the special case of racial and ethnic contact.

Still, continued advances in understanding intergroup contact require more extensive longitudinal research. To date, findings from longitudinal studies typically have shown the persistence of the prejudice reduction achieved by contact (e.g., Eller & Abrams, 2003; Levin et al., 2003). But such studies are rare. In addition to learning about the persistence of contact effects, it is necessary to determine the effects of long-term intergroup contact. Similar to mere exposure effects, we predict that, with continued contact, the reduction of prejudice would asymptote at some point and provide few further gains.

In addition, more elaborate models are needed to integrate and account for these varied intergroup contact effects. Some such models have begun to emerge (e.g., Brown & Hewstone, 2005; Gaertner & Dovidio, 2000; Pettigrew, 1998), many of which use complex structural models (e.g., Eller & Abrams, 2003; Paolini et al., 2004; Van Dick et al., 2004; Voci & Hewstone, 2003; Wagner et al., 2003). For the future, multilevel models that consider both positive and negative factors in the contact situation, along with individual, structural, and normative antecedents of the contact, will greatly enhance researchers' understanding of the nature of intergroup contact effects. And as the contact literature continues to expand rapidly with rigorous methods and attention to theory, we anticipate that the future will witness the development of such comprehensive models of intergroup contact.

References

Allport, G. W. (1954). *The nature of prejudice.* Reading, MA: Addison Wesley.

Amir, Y. (1969). Contact hypothesis in ethnic relations. *Psychological Bulletin, 71,* 319–342.

Amir, Y. (1976). The role of intergroup contact in change of prejudice and race relations. In P. Katz & D. A. Taylor (Eds.), *Towards the elimination of racism* (pp. 245–308). New York: Pergamon Press.

Baker, P. E. (1934). *Negro–White adjustment.* New York: Association Press.

Begg, C. B. (1994). Publication bias. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 399–409). New York: Sage.

Berger, C. R., & Calabrese, R. J. (1975). Some explorations in initial interaction and beyond: Toward a developmental theory of interpersonal communication. *Human Communication Research, 1,* 99–112.

Blair, I. V., Park, B., & Bachelor, J. (2003). Understanding intergroup anxiety: Are some people more anxious than others? *Group Processes and Intergroup Relations, 6*(2), 151–169.

Blascovich, J., Mendes, W. B., Hunter, S. B., & Lickel, B. (2000). Stigma, threat, and social interactions. In T. F. Heatherton, R. E. Kleck, M. R. Hebl, & J. G. Hull (Eds.), *The social psychology of stigma* (pp. 307–333). New York: Guilford Press.

Blascovich, J., Mendes, W. B., Hunter, S. B., Lickel, B., & Kowai-Bell, N. (2001). Perceiver threat in social interactions with stigmatized others. *Journal of Personality and Social Psychology, 80,* 253–267.

Bornstein, R. F. (1989). Exposure and affect: Overview and meta-analysis of research, 1968–1987. *Psychological Bulletin, 106,* 263–289.

Brameld, T. (1946). *Minority problems in the public schools.* New York: Harper.

Brophy, I. N. (1946). The luxury of anti-Negro prejudice. *Public Opinion Quarterly, 9,* 456–466.

Brown, R., & Hewstone, M. (2005). An integrative theory of intergroup contact. *Advances in Experimental Social Psychology, 37,* 255–343.

Butler, J. S., & Wilson, K. L. (1978). The American soldier revisited: Race relations and the military. *Social Science Quarterly, 59*(3), 451–467.

Cagle, L. T. (1973). Interracial housing: A reassessment of the equal-status contact hypothesis. *Sociology and Social Research, 57,* 342–355.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research.* Chicago: Rand McNally.

Carithers, M. W. (1970). School desegregation and racial cleavage, 1954–1970: A review of the literature. *Journal of Social Issues, 26,* 25–47.

Catlin, J. B. (1977). *The impact of interracial living on the racial attitudes and interaction patterns of White college students.* Unpublished doctoral dissertation, University of Michigan.

Chirot, D., & Seligman, M. E. P. (2001). *Ethnopolitical warfare: Causes, consequences, and possible solutions.* Washington, DC: American Psychological Association.

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

Cook, S. W. (1969). Motives in a conceptual analysis of attitude-related behavior. In J. Brigham & T. Weissbach (Eds.), *Racial attitudes in America: Analyses and findings of social psychology* (pp. 250–260). New York: Harper and Row.

Cook, S. W. (1978). Interpersonal and attitudinal outcomes in cooperating interracial groups. *Journal of Research and Development in Education, 12,* 97–113.

Cook, S. W. (1984). Cooperative interaction in multiethnic contexts. In N. Miller & M. B. Brewer (Eds.), *Groups in contact: The psychology of desegregation* (pp. 155–185). Orlando, FL: Academic Press.

Cook, T. D., Cooper, H., Cordray, D. S., Hartman, H., Hedges, L. V., Light, R. J., et al. (1992). Some generic issues and problems for meta-analysis. In T. D. Cook, H. Cooper, D. S. Cordray, H. Hartman, L. V. Hedges, R. J., Light, et al. (Eds.), *Meta-analysis for explanation: A casebook* (pp. 283–320). New York: Sage.

Coursol, A., & Wagner, E. E. (1986). Effect of positive findings on submission and acceptance rates: A note on meta-analysis bias. *Professional Psychology, 17,* 136–137.

Craig, J., Cairns, E., Hewstone, M., & Voci, A. (2002). *Young people's attitudes to and contact with members of the religious out-group.* Unpublished manuscript, University of Ulster.

Crain, R. L., & Weisman, C. S. (1972). *Discrimination, personality, and achievement.* New York: Seminar Press.

Deutsch, M., & Collins, M. (1951). *Interracial housing: A psychological evaluation of a social experiment.* Minneapolis: University of Minnesota Press.

Dickersin, K. (1997). How important is publication bias? A synthesis of available data. *AIDS Education and Prevention, 9*(Suppl. A), 15–21.

Dickersin, K., Min, Y. I., & Meinert, C. L. (1992). Factors influencing the publication of research results: Follow up of applications submitted to two institutional review boards. *Journal of the American Medical Association, 267,* 867–872.

Dijker, A. J. M. (1987). Emotional reactions to ethnic minorities. *European Journal of Social Psychology, 17,* 305–325.

Duval, S. J., & Tweedie, R. L. (2000a). A nonparametric "trim and fill" method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association, 95,* 89–98.

Duval, S. J., & Tweedie, R. L. (2000b). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56*(2), 455–463.

Easterbrook, P. J., Berlin, J. A., Gopalan, R., & Mathews, D. R. (1991). Publication bias in clinical research. *Lancet, 337,* 867–872.

Eller, A. L., & Abrams, D. (2003). "Gringos" in Mexico: Cross-sectional and longitudinal effects of language school-promoted contact on intergroup bias. *Group Processes and Intergroup Relations, 6,* 55–75.

Eller, A. L., & Abrams, D. (2004). Come together: Longitudinal comparisons of Pettigrew's reformulated intergroup contact model and the Common Ingroup Model in Anglo-French and Mexican-American contexts. *European Journal of Social Psychology, 34,* 229–256.

Festinger, L., & Kelley, H. (1951). *Changing attitudes through social contact.* Ann Arbor: University of Michigan, Institute for Social Research.

Fine, G. A. (1979). The Pinkston settlement: An historical and social psychological investigation of the contact hypothesis. *Phylon, 40,* 229–242.

Forbes, H. D. (1997). *Ethnic conflict: Commerce, culture, and the contact hypothesis.* New Haven, CT: Yale University Press.

Ford, R. N. (1941). Scaling experience by a multiple-response technique: A study of White–Negro contacts. *American Sociological Review, 6,* 9–23.

Ford, W. S. (1986). Favorable intergroup contact may not reduce prejudice: Inconclusive journal evidence, 1960–1984. *Sociology and Social Research, 70,* 256–258.

Gaertner, S. L., & Dovidio, J. F. (2000). *Reducing intergroup bias: The common ingroup identity model.* Philadelphia: Psychology Press.

Gathing, L. (1991). Generality vs. specificity of attitudes toward people with disabilities. *British Journal of Medical Psychology, 64,* 55–64.

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

Gudykunst, W. B. (1985). A model of uncertainty reduction in intercultural encounters. *Journal of Language and Social Psychology, 4,* 79–97.

Hamilton, D. L., & Bishop, G. D. (1976). Attitudinal and behavioral effects of initial integration of White suburban neighborhoods. *Journal of Social Issues, 32,* 47–67.

Harmon-Jones, E., & Allen, J. J. B. (2001). The role of affect in the mere exposure effect: Evidence from physiological and individual differences approaches. *Personality and Social Psychology Bulletin, 27,* 889–898.

Harper, D. C., & Wacker, D. P. (1985). Children's attitudes toward disabled peers and the effects of mainstreaming. *Academic Psychology Bulletin, 7,* 87–98.

Harrington, H. J., & Miller, N. (1992). Research and theory in intergroup relations: Issues of consensus and controversy. In J. Lynch, C. Modgil, & S. Modgil (Eds.), *Cultural diversity and the schools* (Vol. 2, pp. 159–178). London: Falmer.

Hedges, L. V. (1994). Fixed effects models. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 285–299). New York: Sage.

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

Herek, G. M., & Capitanio, J. P. (1996). "Some of my best friends": Intergroup contact, concealable stigma, and heterosexuals' attitudes toward gay men and lesbians. *Personality and Social Psychology Bulletin, 22,* 412–424.

Hewstone, M., & Brown, R. (1986). Contact is not enough: An intergroup perspective on the "contact hypothesis." In M. Hewstone & R. Brown (Eds.), *Contact and conflict in intergroup encounters* (pp. 1–44). Cambridge, MA: Basil Blackwell.

Hewstone, M., Cairns, E., Voci, A., Paolini, S., McLernon, F., Crisp, R. J., et al. (2005). Intergroup contact in a divided society: Challenging segregation in Northern Ireland. In D. Abrams, J. M. Marques, & M. A. Hogg (Eds.), *The social psychology of inclusion and exclusion.* Philadelphia: Psychology Press.

Homans, G. C. (1950). *The human group.* New York: Harcourt, Brace, & World.

Hood, M. V., III, & Morris, I. L. (2000). Brother, can you spare a dime? Racial/ethnic context and the Anglo vote on Proposition 187. *Social Science Quarterly, 81,* 194–206.

Hopkins, N., Reicher, S., & Levine, M. (1997). On the parallels between social cognition and the "new racism." *British Journal of Social Psychology, 36,* 305–329.

Irish, D. P. (1952). Reactions of Caucasian residents to Japanese-American neighbors. *Journal of Social Issues, 8,* 10–17.

Islam, M. R., & Hewstone, M. (1993). Dimensions of contact as predictors of intergroup anxiety, perceived outgroup variability, and outgroup attitude: An integrative model. *Personality and Social Psychology Bulletin, 19,* 700–710.

Jackson, J. W. (1993). Contact theory of intergroup hostility: A review and evaluation of the theoretical and empirical literature. *International Journal of Group Tensions, 23,* 43–65.

Johnson, B. T. (1993). *DSTAT: Software for the meta-analytic review of research literatures.* Hillsdale, NJ: Erlbaum.

Johnson, B. T., & Eagly, A. H. (2000). Quantitative synthesis of social psychological research. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social psychology* (pp. 496–528). Cambridge, England: Cambridge University Press.

Kephart, W. M. (1957). *Racial factors and urban law enforcement.* Philadelphia: University of Pennsylvania Press.

Lee, A. Y. (2001). The mere exposure effect: An uncertainty reduction explanation revisited. *Personality and Social Psychology Bulletin, 27,* 1255–1266.

Lett, H. A. (1945). Techniques for achieving interracial cooperation. *Proceedings of the Institute on Race Relations and Community Organization.* Chicago: University of Chicago and the American Council on Race Relations.

Levin, S., van Laar, C., & Sidanius, J. (2003). The effects of ingroup and outgroup friendships on ethnic attitudes in college: A longitudinal study. *Group Processes and Intergroup Relations, 6,* 76–92.

Light, R. J., & Pillemer, D. B. (1984). *Summing up: The science of reviewing research.* Cambridge, MA: Harvard University Press.

Link, B. G., & Cullen, F. T. (1986). Contact with the mentally ill and perceptions of how dangerous they are. *Journal of Health and Social Behavior, 27,* 289–303.

Lipsey, M. W., & Wilson, D. B. (1993). The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta-analysis. *American Psychologist, 48,* 1181–1209.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis.* Thousand Oaks, CA: Sage.

London, L. H., & Linney, J. A. (1993). *Kids' College: Enhancing children's understanding and acceptance of cultural diversity.* Unpublished manuscript, Loyola University.

McClendon, M. J. (1974). Interracial contact and the reduction of prejudice. *Sociological Focus, 7,* 47–65.

Mendes, W. B., Blascovich, J., Lickel, B., & Hunter, S. (2002). Challenge and threat during social interaction with White and Black men. *Personality and Social Psychology Bulletin, 28,* 939–952.

Moreland, R. L., & Zajonc, R. B. (1977). Is stimulus recognition a necessary condition for the occurrence of exposure effects? *Journal of Personality and Social Psychology, 35,* 191–199.

Mosteller, F., & Colditz, G. A. (1996). Understanding research synthesis (meta-analysis). *Annual Review of Public Health, 17,* 1–23.

Naor, M., & Milgram, R. M. (1980). Two preservice strategies for preparing regular class teachers for mainstreaming. *Exceptional Children, 47,* 126–129.

Nesdale, D., & Todd, P. (1998). Intergroup ratio and the contact hypothesis. *Journal of Applied Social Psychology, 28,* 1196–1217.

Nesdale, D., & Todd, P. (2000). Effect of contact on intercultural acceptance: A field study. Intergroup ratio and the contact hypothesis. *International Journal of Intercultural Relations, 24,* 341–360.

Oaker, G., & Brown, R. (1986). Intergroup relations in a hospital setting: A further test of social identity theory. *Human Relations, 39,* 767–778.

Otten, S., Mummendey, A., & Blanz, M. (1996). Intergroup discrimination in positive and negative outcome allocations: Impact of stimulus valence, relative group status, and relative group size. *Personality and Social Psychology Bulletin, 22,* 568–581.

Paolini, S., Hewstone, M., Cairns, E., & Voci, A. (2004). Effects of direct and indirect cross-group friendships on judgments of Catholics and Protestants in Northern Ireland: The mediating role of an anxiety-reduction mechanism. *Personality and Social Psychology Bulletin, 30,* 770–786.

Patchen, M. (1999). *Diversity and unity: Relations between racial and ethnic groups*. Chicago: Nelson-Hall.

Pettigrew, T. F. (1971). *Racially separate or together?* New York: McGraw-Hill.

Pettigrew, T. F. (1986). The contact hypothesis revisited. In H. Hewstone & R. Brown (Eds.), *Contact and conflict in intergroup encounters* (pp. 169–195). Oxford, England: Basil Blackwell.

Pettigrew, T. F. (1997). Generalized intergroup contact effects on prejudice. *Personality and Social Psychology Bulletin, 23,* 173–185.

Pettigrew, T. F. (1998). Intergroup contact theory. *Annual Review of Psychology, 49,* 65–85.

Pettigrew, T. F., & Tropp, L. R. (2000). Does intergroup contact reduce prejudice? Recent meta-analytic findings. In S. Oskamp (Ed.), *Reducing prejudice and discrimination: Social psychological perspectives* (pp. 93–114). Mahwah, NJ: Erlbaum.

Pettigrew, T. F., Wagner, U., Stellmacher, J., & Christ, O. (2006). *Why does authoritarianism predict prejudice? The mediators of a global phenomenon.* Manuscript submitted for publication.

Plant, E. A. (2004). Responses to interracial interactions over time. *Personality and Social Psychology Bulletin, 30,* 1458–1471.

Plant, E. A., & Devine, P. G. (2003). The antecedents and implications of interracial anxiety. *Personality and Social Psychology Bulletin, 29,* 790–801.

Powers, D. A., & Ellison, C. G. (1995). Interracial contact and Black racial attitudes: The contact hypothesis and selectivity bias. *Social Forces, 74,* 205–226.

Raudenbush, S. W. (1994). Random effects models. In H. Cooper & L. V. Hedges (Eds.), *Handbook of research synthesis* (pp. 301–321). New York: Sage.

Rhodes, G., Halberstadt, J., & Brajkovich, G. (2001). Generalization of mere exposure effects to averaged composite faces. *Social Cognition, 19,* 57–70.

Riordan, C. (1978). Equal-status interracial contact: A review and revision of the concept. *International Journal of Intercultural Relations, 2,* 161–185.

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

Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin, 118,* 183–192.

Rotton, J., Foos, P. W., Van Meek, L., & Levitt, M. (1995). Publication practices and the file drawer problem: A survey of published authors. *Journal of Social Behavior and Personality, 10,* 1–13.

Sears, D. O. (1986). College sophomores in the laboratory: Influences of a narrow database on social psychology's view of human nature. *Journal of Personality and Social Psychology, 51,* 515–530.

Shadish, W. R., Doherty, M., & Montgomery, L. M. (1989). How many studies are in the file drawer? An estimate from the family/marital psychotherapy literature. *Clinical Psychology Review, 9,* 589–603.

Sherif, M. (1966). *In common predicament.* Boston: Houghton Mifflin.

Sigelman, L., & Welch, S. (1993). The contact hypothesis revisited: Black–White interaction and positive racial attitudes. *Social Forces, 71,* 781–795.

Smith, M. L. (1980). Publication bias and meta-analysis. *Evaluation in Education, 4,* 22–24.

Sommer, B. (1987). The file drawer effect and publication rates in menstrual cycle research. *Psychology of Women Quarterly, 11,* 233–242.

Stephan, W. G. (1987). The contact hypothesis in intergroup relations. In C. Hendrick (Ed.), *Review of personality and social psychology*: *Group processes and intergroup relations* (Vol. 9, pp. 13–40). Newbury Park, CA: Sage.

Stephan, W. G., Boniecki, K. A., Ybarra, O., Bettencourt, A., Ervin, K. S., Jackson, L. A., et al. (2002). The role of threats in the racial attitudes of Blacks and Whites. *Personality and Social Psychology Bulletin, 28,* 1242–1254.

Stephan, W. G., & Stephan, C. W. (1985). Intergroup anxiety. *Journal of Social Issues, 41,* 157–175.

Stephan, W. G., Stephan, C. W., & Gudykunst, W. B. (1999). Anxiety in intercultural relations: A comparison of anxiety/uncertainty management theory and integrated threat theory. *International Journal of Intercultural Relations, 23,* 613–628.

Sterne, J. A. C., & Egger, M. (2000). High false positive rates for trim and fill method. *British Journal of Medicine.* Retrieved February 5, 2004, from http://bmj.com/cgi/eletters/ 320/7249/1574#EL1.

Sutton, A. J., Duval, S. J., Tweedie, R. L., Abrams, K. R., & Jones, D. R. (2000). Empirical assessment of effect of publication bias on meta-analysis. *British Journal of Medicine, 320,* 1574–1577.

Sutton, A. J., Song, F., Gilbody, S. M., & Abrams, K. R. (2000). Modeling publication bias in meta-analysis: A review. *Statistical Methods in Medical Research, 9,* 421–445.

Taft, R. (1959). Ethnic stereotypes, attitudes, and familiarity: Australia. *Journal of Social Psychology, 49,* 177–186.

Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behavior. *European Journal of Social Psychology, 1,* 149–178.

Tropp, L. R. (2003). The psychological impact of prejudice: Implications for intergroup contact. *Group Processes and Intergroup Relations, 6,* 131–149.

Van Dick, R., Wagner, U., Pettigrew, T. F., Christ, O., Wolf, C., Petzel, T., et al. (2004). The role of perceived importance in intergroup contact. *Journal of Personality and Social Psychology, 87*, 211–227.

Van Dyk, A. C. (1990). Voorspellers van etniese houdings in ′n noue kontaksituasie [Determinants of ethnic attitudes in a close contact situation]. *South African Journal of Psychology, 20,* 206–214.

Vevea, J. L., & Hedges, L. V. (1995). A general linear model for estimating effect size in the presence of publication bias. *Psychometrika, 60,* 419–435.

Voci, A., & Hewstone, M. (2003). Intergroup contact and prejudice toward immigrants in Italy: The mediational role of anxiety and the moderational role of group salience. *Group Processes and Intergroup Relations, 6,* 37–54.

Wagner, U., van Dick, R., Pettigrew, T. F., & Christ, O. (2003). Ethnic prejudice in East and West Germany: The explanatory power of intergroup contact. *Group Processes and Intergroup Relations, 6,* 22–36.

Watson, G. (1946). Unity for action. *Jewish Affairs, 1,* 3–22.

Williams, R. M., Jr. (1947). *The reduction of intergroup tensions.* New York: Social Science Research Council.

Wilner, D. M., Walkley, R. P., & Cook, S. W. (1952). *Human relations in interracial housing: A study of the contact hypothesis.* Minneapolis: University of Minnesota Press.

Wilson, D. B. (2002). *SPSS macros for meta-analytic data*. Retrieved January 14, 2002, from http://mason.gmu.edu/~dwilsonb/ma.html

Wilson, T. C. (1996). Prejudice reduction or self-selection? A test of the contact hypothesis. *Sociological Spectrum, 16,* 43–60.

Works, E. (1961). The prejudice-interaction hypothesis from the point of view of the Negro minority group. *American Journal of Sociology, 67,* 47–52.

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology, 9*(Monograph Suppl. 2), 1–27.

Zajonc, R. B., & Rajecki, D. W. (1969). Exposure and affect: A field experiment. *Psychonomic Science, 17,* 216–217.

Appendix: Ratings of Samples Included in the Meta-Analysis

Reference	Sample	*r*	Test	Statistic	Choice	Pub	Type	B/W	Control	IV	IVQ	DVQ	Prog	Target	Age	Sex	Geo	Set	*N*
Abu-Hilal (1986)	1	−.210	*r*	−.210	full	2	1	W	1	1	3	3	1	1	coll	b/u	6	m/o	353
Adams (1992)	1	−.345	*r*	−.345*	full	2	1	W	1	1	2	3	1	1	coll	b/u	1	m/o	42
	2	−.323	*r*	−.323*	full	2	1	W	1	1	2	3	1	1	coll	b/u	1	m/o	26
	3	−.508	*r*	−.508*	full	2	1	W	1	1	2	3	1	1	coll	b/u	1	m/o	51
	4	−.346	*r*	−.346*	full	2	1	W	1	1	2	3	1	1	coll	b/u	1	m/o	67
	5	−.300	*r*	−.300*	full	2	1	W	1	1	2	3	1	1	coll	b/u	1	m/o	58
Aday et al. (1991)	1	−.311	*t*	−2.29*	none	1	2	B	2	2	4	3	2	2	child	b/u	1	m/o	49
Aday et al. (1993)	1	−.487	*t*	−3.48	none	1	2	B	2	2	4	3	2	2	adol	b/u	1	rec	39
Alderfer et al. (1992)	1	−.105	*F*	5.29*	some	1	2	B	4	2	99	2	2	1	adult	b/u	1	org	477
Aljeaid (1986)	1	−.330	*M/SD*	*p* = .000	some	2	1	B	2	1	1	3	1	1	coll	b/u	1	edu	296
Allport & Kramer (1946)	1	−.154	Prop	51/66*	full	1	1	B	4	1	2	2	1	1	coll	b/u	1	m/o	393
Alreshoud & Koeske (1997)	1	−.290	*r*	−.290	full	1	2	W	1	1	3	3	1	1	coll	b/u	1	m/o	74
Altrocchi & Eisdorfer (1961)	1	−.200	Prop	82/93	some	1	2	W	1	2	99	2	1	5	coll	f	1	edu	49
	2	−.181	Prop	93/100	some	1	2	W	1	2	99	2	1	5	coll	f	1	edu	192
Amir & Ben-Ari (1985)	1	−.088	*t*	−.088	full	1	2	W	1	2	99	3	1	1	adult	b/u	3	trav	483
Amir & Garti (1977)	1	−.158	*t*	−2.80	some	1	1	W	1	2	99	2	1	1	adol	f	3	rec	78
	2	−.123	*t*	−1.16	some	1	1	W	1	2	99	2	1	1	adol	f	3	rec	22
Amir et al. (1978)	1	−.067	*t*	−2.75*	some	1	1	W	1	3	99	2	1	1	adol	b/u	3	edu	419
	2	+.023	*t*	+1.15*	some	1	1	W	1	3	99	2	1	1	adol	b/u	3	edu	614
Amsel & Fichten (1988)	1	−.419	*t*	−4.95*	full	1	1	B	2	1	1	2	1	4	coll	b/u	4	m/o	117
Angermeyer & Matshinger (1997)	1	−.134	Prop	36/50*	some	1	1	B	3	1	1	2	1	5	adult	b/u	2	m/o	1,484
Anthony (1969)	1	−.361	*t*	−2.44*	full	1	1	B	4	2	4	3	2	4	coll	b/u	1	rec	42
Antonak (1981)	1	−.150	*r*	−.150*	full	1	1	W	1	1	1	3	1	4	coll	b/u	1	m/o	326
Antonak et al. (1989)	1	−.132	*p*	.000	full	1	1	W	1	1	2	2	1	6	adult	b/u	1	m/o	557
Archie & Sherrill (1989)	1	−.096	Prop	54/67*	some	1	1	B	2	3	99	3	1	4	child	b/u	1	edu	229
Arguc (1995)	1	−.015	*r*	−.015*	some	2	1	W	1	1	1	1	1	1	adult	m	2	m/o	96
Arikan & Uysal (1999)	1	−.054	*t*	−2.74*	some	1	1	W	3	1	1	3	1	5	coll	b/u	6	edu	630
Arkar & Eker (1992)	1	−.092	*p*	.400	none	1	1	B	3	1	1	3	1	5	adult	b/u	6	org	84
Aronson & Page (1980)	1	−.140	*p*	.364*	full	1	2	B	2	2	99	2	2	5	coll	b/u	1	org	42
Auerbach & Levinson (1977)	1	+.519	Prop	88/38	none	1	2	B	3	3	99	2	1	2	coll	b/u	1	edu	120
Bagget (1981)	1	−.090	*t*	−0.68*	none	1	2	B	3	2	4	2	2	2	child	b/u	1	edu	56
Ballard et al. (1977)	1	−.339	*M/SD*	*p* = .051*	some	1	2	B	3	2	99	4	2	6	child	b/u	1	edu	33
Barnard & Benn (1987)	1	−.197	*F*	11.34	some	1	3	W	1	2	4	2	2	1	coll	m	1	lab	48
Barnea & Amir (1981)	1	.000	*p*	*ns*	some	1	1	B	4	3	99	2	1	1	coll	b/u	3	m/o	209
	2	.000	*p*	*ns*	some	1	1	B	4	3	99	2	1	1	coll	b/u	3	m/o	209
Basu & Ames (1970)	1	−.484	*r*	−.484*	full	1	1	W	1	1	2	2	1	1	coll	b/u	1	m/o	562
Beh-Pajooh (1991)	1	−.341	*M/SD*	*p* = .000*	some	1	1	B	2	1	1	2	1	6	coll	b/u	2	edu	132
Bekker & Taylor (1966)	1	−.258	*M/SD*	*p* = .01	none	1	1	B	2	3	1	2	1	2	coll	b/u	1	m/o	100
Belan (1996)	1	−.077	*F*	1.75*	some	2	1	B	4	1	2	3	1	5	adult	b/u	1	m/o	296
Bell (1962)	1	−.216	*t*	−2.07*	full	1	1	B	3	1	99	2	1	4	adult	b/u	1	m/o	110
Benedict et al. (1988)	1	−.193	*r*	−.193*	full	1	1	W	1	1	1	2	1	99	adult	b/u	1	m/o	112
	2	−.143	*r*	−.143*	full	1	1	W	1	1	1	2	1	99	adult	b/u	1	m/o	112
Benedict et al. (1992)	1	–.205	*r*	−.205*	full	1	1	W	1	1	2	2	1	99	adult	b/u	1	m/o	314
Berg & Wolleat (1973)	1	+.235	*M/SD*	*p* = .02	some	1	1	B	2	1	2	3	1	1	child	b/u	1	m/o	100
Bergmann & Erb (1997)	1	−.205	Prop	20/40*	full	1	1	B	2	1	1	2	1	1	adult	b/u	2	m/o	2,102
Bicknese (1974)	1	−.195	*p*	.25*	full	1	2	W	1	2	99	2	1	1	coll	b/u	2	edu	19
	2	−.108	*p*	.52*	full	1	2	W	1	2	99	2	1	1	coll	b/u	2	edu	18
	3	−.173	*p*	.25*	full	1	2	W	1	2	99	2	1	1	coll	b/u	2	edu	22
Biernat (1990)	1	−.118	*r*	−.118*	full	1	1	W	1	1	1	2	1	1	coll	b/u	1	res	78
	2	−.275	*r*	−.275*	full	1	1	W	1	1	1	2	1	1	coll	b/u	1	res	90
Biernat & Crandall (1994)	1	−.349	*r*	−.349*	full	1	1	W	1	1	3	2	1	99	coll	b/u	1	m/o	116
Borus et al. (1973)	1	−.250	*r*	−.250	some	1	1	W	1	1	1	3	1	1	adult	m	1	m/o	1,385
Bowman (1979)	1	−.201	Prop	25/44	full	1	1	B	2	1	2	1	1	3	adult	b/u	5	m/o	322
Bradnum et al. (1993)	1	−.163	*M/SD*	*p* = .006	none	1	1	B	3	3	99	2	1	1	adol	b/u	6	edu	294
	2	−.221	*M/SD*	*p* = .000	none	1	1	B	3	3	99	2	1	1	adol	b/u	6	edu	336
Brewer & Campbell (1976)	1	−.089	*r*	−.089*	some	1	1	W	1	1	2	2	1	1	adult	b/u	6	m/o	1,500
Brigham (1993)	1	−.158	*r*	−.158*	full	1	1	W	1	1	2	3	1	1	coll	b/u	1	m/o	280
	2	−.361	*r*	−.361*	full	1	1	W	1	1	2	3	1	1	coll	b/u	1	m/o	81
Brigham & Barkowitz (1978)	1	−.420	*r*	−.420	full	1	1	W	1	1	2	2	1	1	coll	b/u	1	m/o	76

(*Appendix continues*)

Appendix (*continued*)

Reference	Sample	*r*	Test	Statistic	Choice	Pub	Type	B/W	Control	IV	IVQ	DVQ	Prog	Target	Age	Sex	Geo	Set	*N*
	2	−.130	*r*	−.130	full	1	1	W	1	1	2	2	1	1	coll	b/u	1	m/o	86
Brigham & Malpass (1985)	1	−.580	*r*	−.580	full	1	1	W	1	1	2	3	1	1	coll	b/u	1	m/o	78
Brigham & Ready (1985)	2	−.210	*r*	−.210	full	1	1	W	1	1	2	3	1	1	coll	b/u	1	m/o	90
Brink & Harris (1964)	1	−.202	Prop	18/36	full	1	1	B	3	1	99	2	1	1	adult	b/u	1	m/o	1,257
Britt et al. (1996)	1	−.070	*r*	−.070*	full	1	1	W	1	1	2	2	1	1	coll	b/u	1	m/o	131
Brockington et al. (1993)	1	−.180	*r*	−.180*	some	1	1	W	1	1	2	2	1	5	adult	b/u	2	m/o	1,987
Brockman & D'Arcy (1978)	1	−.131	Prop	51/64	some	1	1	B	3	1	1	2	1	5	adult	b/u	4	m/o	221
Brooks & Fricdrich (1970)	1	−.222	Prop	36/58	some	1	1	B	3	1	1	2	1	99	adult	b/u	1	m/o	85
	2	−.143	Prop	23/36	some	1	1	B	3	1	1	2	1	99	adult	b/u	1	m/o	146
Brooks et al. (1973)	1	−.125	*r*	−.125	full	1	1	W	1	1	1	1	1	1	coll	b/u	1	m/o	54
	2	−.300	*r*	−.300	full	1	1	W	1	1	1	1	1	1	coll	b/u	1	m/o	56
Brophy (1945)	1	−.433	Prop	28/72	none	1	1	B	4	3	99	2	1	1	adult	m	1	org	447
Brown (1997)	1	−.290	*r*	−.290*	some	2	1	W	1	1	2	3	1	1	coll	b/u	1	m/o	190
Brown & Albee (1966)	1	+.245	*t*	+2.73*	none	1	1	B	2	3	99	2	1	1	adult	m	1	m/o	120
Brown et al. (1986)	1	−.143	*r*	−.143*	full	1	1	W	1	1	1	4	1	99	adult	f	2	org	29
	2	−.238	*r*	−.238*	full	1	1	W	1	1	1	4	1	99	adult	m	2	org	16
	3	−.168	*r*	−.168*	full	1	1	W	1	1	1	4	1	99	adult	m	2	org	30
	4	−.010	*r*	−.010*	full	1	1	W	1	1	1	4	1	99	adult	m	2	org	39
	5	−.010	*r*	−.010*	full	1	1	W	1	1	1	4	1	99	adult	m	2	org	33
Brown et al. (1999)	1	−.450	*r*	−.450	some	1	1	W	1	1	3	1	1	1	coll	b/u	2	m/o	85
	2	−.230	*r*	−.230	some	1	1	W	1	1	3	1	1	1	coll	b/u	2	m/o	217
Brown et al. (1999/2001)	1	−.205	*M/SD*	$p = .001$*	full	2	2	B	3	1	2	3	1	1	adult	b/u	2	m/o	262
Bucich-Naylor (1978)	1	−.041	*M/SD*	$p = .73$*	none	2	2	B	4	2	4	3	2	4	child	b/u	1	edu	69
Bullock (1976a/1976b/1978)	1	−.298	*r*	−.298*	full	1	1	W	1	1	2	2	1	1	adol	b/u	1	m/o	2,076
	2	−.101	*r*	−.101*	full	1	1	W	1	1	2	2	1	1	adol	b/u	1	m/o	1,755
Buono (1981)	1	−.175	*r*	−.175*	full	2	2	W	1	2	99	2	1	1	adult	b/u	1	res	121
	2	+.029	*r*	+.029*	full	2	2	W	1	2	99	2	1	1	adult	b/u	1	res	50
Burgin & Walker (2000)	1	−.373	Mult	$p = .000$*	some	2	2	B	3	1	3	2	1	1	adol	f	5	rec	137
Butler & Wilson (1978)	1	−.156	*r*	−.156*	some	1	1	W	1	1	2	3	1	1	adult	b/u	1	org	1,490
	2	−.131	*r*	−.131*	some	1	1	W	1	1	2	3	1	1	adult	b/u	1	org	3,000
Caditz (1976)	1	−.131	MW	$p = .069$*	full	1	1	B	3	1	2	2	1	1	adult	b/u	1	m/o	196
Campbell (1958)	1	−.067	*p*	$p = .01$*	some	1	1	W	1	3	99	3	1	1	adol	b/u	1	edu	746
Canter & Shoemaker (1960)	1	−.263	*M/SD*	$p = .042$*	some	1	2	W	1	2	99	2	1	5	coll	f	1	org	30
Carlson & Widaman (1988)	1	−.290	*F*	70.4*	full	1	1	B	2	2	99	2	1	1	coll	b/u	3	trav	823
Carstensen et al. (1982)	1	−.382	*F*	4.35	none	1	2	B	3	2	4	3	2	2	child	b/u	1	edu	26
Carter & Mitchell (1956)	1	−.218	*t*	−2.18	some	1	1	B	3	1	1	2	1	1	adol	b/u	1	m/o	124
Casey (1978)	1	−.210	*r*	−.210*	some	1	1	W	1	1	99	2	1	4	adult	b/u	1	edu	100
Caspi (1984)	1	−.488	Mult	$p = .001$	some	1	2	B	2	2	99	2	2	2	child	b/u	1	edu	38
Catlin (1977)	1	−.084	*r*	−.084*	some	2	1	W	1	1	1	2	1	1	coll	b/u	1	edu	570
Chadwick et al. (1971)	1	−.243	*r*	−.243*	some	1	1	W	1	1	2	2	1	1	adol	b/u	1	edu	300
	2	−.155	*r*	−.155*	some	1	1	W	1	1	2	2	1	1	adol	b/u	1	edu	35
Chang (1973)	1	−.197	χ^2	8.87	full	1	1	W	1	1	2	2	1	1	coll	b/u	1	m/o	238
Chang (1998)	1	−.021	*r*	−.021*	some	1	1	W	1	1	1	3	1	1	adult	b/u	1	org	260
	2	−.113	*r*	−.113*	some	1	1	W	1	1	1	3	1	1	adult	b/u	1	org	244
Chen et al. (1970)	1	−.254	Prop	31/56	some	1	1	B	3	1	1	1	1	1	coll	b/u	3	m/o	99
Chinsky & Rappaport (1970)	1	−.125	*p*	.230	full	1	2	B	2	2	4	2	1	5	coll	b/u	1	org	90
	2	−.213	χ^2	10.8*	none	1	2	W	1	2	4	2	1	99	adult	b/u	1	org	119
Chou & Mak (1998)	1	−.085	*r*	−.085*	some	1	1	W	1	1	2	2	1	5	adult	b/u	6	m/o	1,273
Cleland & Cochran (1961)	1	−.023	*p*	.750*	some	1	2	W	1	2	99	2	1	6	adol	b/u	1	res	98
Clément et al. (1977)	1	−.140	*F*	4.68*	full	1	1	B	2	2	99	1	1	1	adol	b/u	4	trav	253
Clore et al. (1978)	1	−.151	*F*	2.56*	none	1	2	B	3	2	4	2	2	1	child	b/u	1	rec	112
Clunies-Ross & O'Meara (1989)	1	−.337	*M/SD*	$p = .009$*	none	1	3	B	3	2	4	3	2	4	child	b/u	5	rec	60
Colca et al. (1982)	1	−.263	*F*	4.75*	none	1	2	B	4	2	4	2	2	1	child	b/u	1	edu	64

Appendix (*continued*)

Reference	Sample	*r*	Test	Statistic	Choice	Pub	Type	B/W	Control	IV	IVQ	DVQ	Prog	Target	Age	Sex	Geo	Set	*N*
Cook & Wollersheim (1976)	1	−.013	Mult	*p* = .87*	some	1	2	B	2	3	99	2	1	6	adol	b/u	1	edu	150
Cook (1969)	1	−.316	Prop	65/91	none	1	3	B	2	2	4	3	2	1	coll	f	1	lab	46
Cookston (1973)	1	−.372	*F*	5.10*	some	2	2	B	3	2	99	2	2	1	coll	b/u	1	edu	47
	2	−.272	*F*	4.92*	some	2	2	B	3	2	99	2	2	1	coll	b/u	1	edu	62
Cotten-Huston & Waite (2000)	1	−.267	Mult	*p* = .000*	some	1	1	W	1	1	1	3	1	3	coll	b/u	1	m/o	150
Couper et al. (1991)	1	−.156	*t*	−1.68*	none	1	2	B	2	2	4	4	2	2	adol	b/u	1	lab	114
Cousens & Crawford (1988)	1	−.309	*M/SD*	*p* = .000*	some	1	1	B	3	1	2	3	1	5	adult	b/u	5	m/o	158
Cowen et al. (1958)	1	+.039	*t*	+0.39	full	1	1	B	3	1	1	3	1	4	adult	b/u	1	edu	101
Crain & Weisman (1972)	1	−.074	*r*	−.074	full	1	1	W	1	1	1	2	1	1	adult	m	1	m/o	1,715
	2	−.110	*r*	−.110	full	1	1	W	1	1	1	2	1	1	adult	f	1	m/o	2,043
Creech (1977)	1	−.213	*F*	18.2*	full	1	2	W	1	2	4	3	1	5	coll	m	1	org	95
Crull & Bruton (1979)	1	−.196	*M/SD*	*p* = .000*	full	1	1	B	2	1	1	2	1	99	coll	b/u	1	m/o	1,043
D'Augelli (1989)	1	−.143	*r*	−.143*	full	1	1	W	1	1	1	3	1	3	coll	b/u	1	m/o	101
D'Augelli & Rose (1990)	1	−.200	*r*	−.200*	full	1	1	W	1	1	1	2	1	3	coll	b/u	1	m/o	218
Davidson et al. (1983)	1	−.330	*r*	−.330*	full	1	1	W	1	1	2	2	1	1	adult	b/u	5	m/o	150
Dellmann-Jenkins et al. (1986)	1	−.188	Prop	26/44*	none	1	2	B	2	2	4	2	2	2	child	b/u	1	m/o	30
Dellmann-Jenkins et al. (1991)	1	−.235	Prop	85/98*	some	1	2	B	3	2	4	4	2	2	child	b/u	1	m/o	31
Desforges et al. (1991)	1	−.150	*M/SD*	*p* = .21*	none	1	2	W	1	2	4	2	2	5	coll	b/u	1	lab	35
	2	−.291	*M/SD*	*p* = .01*	none	1	2	W	1	2	4	2	2	5	coll	b/u	1	lab	29
Deutsch & Collins (1951)	1	−.288	Prop	40/69*	none	1	1	B	3	1	2	2	1	1	adult	f	1	res	390
Deutsche Shell (2000)	1	−.250	*r*	−.250	some	1	1	W	1	1	3	3	1	1	adol	b/u	2	m/o	3,000
Di Tullio (1982)	1	−.873	*M/SD*	*p* = .000*	none	2	3	B	2	2	4	3	1	6	adult	m	1	org	76
Diamond & Lobitz (1973)	1	−.335	*t*	−3.00*	full	1	2	B	1	2	4	2	2	99	coll	b/u	1	m/o	73
	2	−.447	*t*	−3.46*	full	1	2	W	1	2	4	2	2	99	adult	m	1	m/o	12
Dijker (1987)	1	−.159	*r*	−.159*	full	1	1	W	1	1	2	3	1	1	adult	b/u	2	m/o	95
Distefano & Pryer (1970)	1	−.157	*p*	.06*	full	1	2	W	1	2	4	2	1	5	adult	b/u	1	res	71
Dodson (1970)	1	−.494	*t*	−4.40*	full	2	2	W	1	2	4	3	1	1	coll	b/u	1	edu	15
	2	−.034	*t*	−0.13*	full	2	2	W	1	2	4	3	1	1	coll	b/u	1	edu	8
Doka (1985–1986)	1	−.013	Prop	51/52*	full	1	2	B	4	2	4	2	1	2	adol	b/u	1	rec	48
Donaldson & Martinson (1977)	1	−.253	*p*	.05*	none	1	2	B	3	2	4	3	1	4	coll	b/u	1	edu	120
Dooley & Frankel (1990)	1	−.528	*p*	.000*	full	1	2	W	1	2	99	3	2	2	adol	b/u	4	res	21
Drake (1957)	1	−.007	Prop	33/34*	none	1	1	B	3	1	1	2	1	2	coll	b/u	1	m/o	397
Dubey (1979)	1	+.134	χ^2	7.72	full	1	1	B	4	1	2	2	1	99	adult	b/u	6	res	428
	2	−.173	χ^2	3.25	full	1	1	B	4	1	2	2	1	99	adult	b/u	6	res	109
Duckitt (1984)	1	−.201	*r*	−.201*	full	1	1	W	1	1	1	2	1	3	adult	b/u	6	m/o	1,420
Dunbar (2000)	1	−.443	*r*	−.443	some	2	1	W	1	1	2	2	1	1	coll	b/u	2	m/o	125
Eaton & Clore (1975)	1	−.185	*t*	−1.96	some	1	2	B	4	2	4	4	2	1	child	b/u	1	rec	112
Eberhardt & Mayberry (1995)	1	−.120	*r*	−.120*	full	1	1	W	1	1	3	3	1	4	adult	b/u	5	org	172
Eddy (1986)	1	−.121	Prop	50/62*	some	1	2	W	1	2	99	2	1	2	coll	b/u	1	org	56
Eller (2000)	1	−.084	*r*	−.084*	some	2	1	W	1	1	3	3	1	1	coll	b/u	2	m/o	104
	2	−.363	*r*	−.363*	full	2	1	W	1	1	3	3	1	99	coll	b/u	2	trav	102
	3	−.210	*r*	−.210*	none	2	1	W	1	1	3	3	1	99	adol	b/u	2	edu	708
Eller & Abrams (1999)	1	−.275	*r*	−.275*	some	2	1	W	1	1	3	3	1	1	coll	b/u	6	m/o	67
Eller et al. (1999/2000)	1	−.297	*r*	−.297*	some	2	1	W	1	1	2	3	1	1	coll	b/u	6/1	edu	239
	2	−.300	*r*	−.300*	some	2	1	W	1	1	2	3	1	1	coll	b/u	6/1	edu	90
Eller et al. (2000)	1	−.272	*r*	−.272*	some	2	1	W	1	1	3	3	1	1	adult	b/u	6/1	m/o	207
Ellis & Vasseur (1993)	1	−.437	*r*	−.437*	none	1	3	W	1	1	2	2	1	3	coll	b/u	1	m/o	108
Emerton & Rothman (1978)	1	+.124	*t*	+1.37	full	1	1	W	1	1	2	2	1	4	coll	b/u	1	edu	30
Ervin (1993)	1	−.025	*r*	−.025*	some	2	1	W	1	1	2	1	1	1	coll	b/u	1	m/o	100
	2	−.239	*r*	−.239*	some	2	1	W	1	1	2	1	1	1	coll	b/u	1	m/o	130
Eshel & Dicker (1995)	1	−.260	*M/SD*	*p* = .001*	some	1	1	B	4	2	99	1	1	1	adol	b/u	3	edu	160
Esposito & Peach (1983)	1	−.728	*p*	.001	some	1	1	W	1	2	99	2	2	4	child	b/u	1	edu	9

(*Appendix continues*)

Appendix (*continued*)

Reference	Sample	*r*	Test	Statistic	Choice	Pub	Type	B/W	Control	IV	IVQ	DVQ	Prog	Target	Age	Sex	Geo	Set	*N*
Esposito & Reed (1986)	1	−.490	*M/SD*	*p* = .000*	some	1	1	B	3	2	99	4	2	4	child	b/u	1	edu	92
Evans (1976)	1	−.539	*M/SD*	*p* = .001*	none	1	3	B	1	2	4	3	2	4	coll	b/u	1	lab	60
Felton (1975)	1	−.473	*t*	−2.84	full	1	2	W	1	2	99	3	2	4	adult	f	1	org	7
Fenrick & Petersen (1984)	1	−.451	Mult	*p* = .000*	none	1	2	B	3	2	4	2	2	99	child	b/u	1	edu	63
Fichten & Amsel (1986)	1	.000	*p*	*ns*	some	1	1	B	3	1	1	2	1	4	coll	b/u	4	m/o	115
Fichten et al. (1988)	1	−.206	*p*	.05	full	1	1	B	2	1	2	1	1	4	adult	b/u	4	edu	91
Fichten et al. (1989)	1	−.158	*t*	−1.73	full	1	1	B	2	1	1	2	1	4	coll	b/u	4	m/o	125
Finchilescu (1988)	1	−.335	*F*	14.3*	some	1	1	B	2	2	99	2	1	1	coll	b/u	6	org	113
Florian & Kehat (1987)	1	−.079	*M/SD*	*p* = .46	none	1	1	B	3	2	4	3	2	4	adol	b/u	1	m/o	88
Floyd (1970)	1	−.170	Prop	44/61	some	2	1	B	2	1	2	2	1	5	adult	f	1	m/o	131
Foley (1977)	1	−.070	*F*	−0.070	full	1	1	W	1	3	99	2	1	1	adult	m	1	m/o	40
	2	−.223	*r*	−.223	some	1	1	W	1	3	99	2	1	1	adult	m	1	org	30
Ford (1973)	1	−.503	Prop	24/74*	some	1	1	B	3	1	2	2	1	1	adult	f	1	res	72
	2	−.148	Prop	43/58*	none	1	1	B	3	1	2	2	1	1	adult	f	1	res	73
Friedman (1975)	1	−.266	*M/SD*	*p* = .05	none	2	2	B	2	2	4	3	2	4	child	b/u	1	edu	55
Friesen (1966)	1	−.250	*r*	−.250	some	2	1	W	1	1	2	3	1	4	adult	b/u	6	org	241
	2	−.310	*r*	−.310	some	2	1	W	1	1	2	3	1	4	adult	b/u	6	org	135
Furnham & Gibbs (1984)	1	−.178	*F*	8.10*	full	1	1	B	3	1	1	2	1	4	adol	b/u	2	m/o	135
Furnham & Pendred (1983)	1	.000	*p*	*ns*	some	1	1	W	1	1	2	2	1	4	adult	b/u	2	m/o	96
Furuto & Furuto (1983)	1	−.232	*p*	.01*	none	1	3	B	3	2	4	2	2	1	coll	b/u	1	lab	124
Gaertner et al. (1994)	1	−.328	*r*	−.328*	some	1	1	W	1	1	2	2	1	1	adol	b/u	1	edu	1,181
Gaertner et al. (1999)	1	−.143	*F*	12.11*	none	1	3	B	2	2	4	2	2	99	coll	b/u	1	lab	576
Gardner et al. (1969)	1	.106	O	*p* = .38*	full	1	1	B	2	2	99	2	1	1	adult	b/u	6	m/o	68
Gardner et al. (1973)	1	−.115	O	*p* = .01*	full	1	1	W	1	1	1	2	1	1	coll	b/u	6	edu	250
Gardner et al. (1974)	1	−.134	*t*	−3.92*	full	1	1	W	1	1	99	2	1	1	adol	b/u	2	trav	211
Gelber (1993)	1	−.299	*M/SD*	*p* = .013*	none	2	3	W	1	2	4	3	2	4	coll	b/u	1	lab	37
	2	−.378	*M/SD*	*p* = .001*	none	2	3	W	1	2	4	3	2	4	coll	b/u	1	lab	37
	3	−.415	*M/SD*	*p* = .000*	none	2	3	W	1	2	4	3	2	4	coll	b/u	1	lab	38
Gelfand & Ullmann (1961)	1	−.295	*t*	−2.31*	full	1	2	B	3	2	4	3	2	5	coll	m	1	org	59
Gentry (1987)	1	−.096	*p*	.18	full	1	1	W	1	1	1	3	1	3	coll	m	1	m/o	96
	2	−.191	*p*	.006	full	1	1	W	1	1	1	3	1	3	coll	f	1	m/o	105
Gerbert et al. (1991)	1	−.147	*p*	.040*	some	1	1	B	4	1	1	3	1	99	adult	b/u	1	m/o	1,320
Gething (1991)	1	−.256	*M/SD*	*p* = .000	some	1	1	B	4	1	2	3	1	4	adult	b/u	5	m/o	460
Glass & Meckler (1972)	1	−.483	*M/SD*	*p* = .004	full	1	2	W	1	2	4	1	2	6	adult	f	1	edu	18
Glass & Trent (1980)	1	−.034	*F*	1.74	none	1	1	W	1	1	1	2	1	2	adol	b/u	1	res	388
Glassner & Owen (1976)	1	−.236	*r*	−.236*	full	1	1	W	1	1	1	2	1	3	coll	b/u	1	edu	61
Glock et al. (1975)	1	−.259	Prop	53/66	full	1	1	B	2	1	1	3	1	1	adol	b/u	1	edu	750
	2	−.186	Prop	41/60	full	1	1	B	2	1	1	3	1	1	adol	b/u	1	edu	608
	3	−.101	Prop	31/41	full	1	1	B	2	1	1	3	1	1	adol	b/u	1	edu	1,328
Glover & Smith (1997)	1	−.083	*t*	−0.53	none	1	1	B	3	2	99	2	1	1	child	b/u	1	edu	41
	2	−.503	*t*	−2.45	none	1	1	B	3	2	99	2	1	1	child	b/u	1	edu	19
Goldstein & Simpkins (1973)	1	−.471	*t*	−4.12*	full	1	1	W	1	2	99	3	2	99	coll	b/u	1	org	15
Gordon & Hallauer (1976)	1	−.330	χ^2	6.71*	full	1	2	W	1	2	99	2	2	2	coll	b/u	1	res	40
Gosse & Sheppard (1979)	1	−.244	*M/SD*	*p* = .000	full	1	1	B	3	1	1	2	1	4	adol	b/u	4	m/o	273
	2	.027	*M/SD*	*p* = .662	full	1	1	B	3	1	1	2	1	4	adol	b/u	4	m/o	268
	3	−.477	*M/SD*	*p* = .000	full	1	1	B	3	1	1	2	1	4	coll	b/u	4	m/o	155
Goto (2000)	1	−.395	*r*	−.395*	some	2	1	W	1	1	2	2	1	1	adol	b/u	1	m/o	511
	2	−.309	*r*	−.309*	some	2	1	W	1	1	2	2	1	1	adol	b/u	1	m/o	135
Gottlieb & Corman (1975)	1	−.020	*F*	0.63*	some	1	1	W	1	1	1	2	1	6	adult	b/u	1	m/o	394
Grack & Richman (1996)	1	−.713	*F*	35.15	none	1	3	B	2	2	4	3	2	3	coll	b/u	1	lab	34
Graffi & Minnes (1988)	1	.000	*p*	*ns*	full	1	1	W	1	1	1	3	1	6	child	b/u	4	edu	120
Grantham & Block (1983)	1	−.089	*p*	.15*	some	1	1	W	1	1	1	2	1	5	coll	b/u	1	edu	289
	2	−.095	*p*	.042	some	1	1	W	1	1	1	2	1	5	coll	b/u	1	edu	229
Gray & Thompson (1953)	1	−.141	O	*p* = .000	full	1	1	W	1	1	1	2	1	1	coll	b/u	1	m/o	400
	2	−.447	O	*p* = .000	full	1	1	W	1	1	1	2	1	1	coll	b/u	1	m/o	300

Appendix (*continued*)

Reference	Sample	*r*	Test	Statistic	Choice	Pub	Type	B/W	Control	IV	IVQ	DVQ	Prog	Target	Age	Sex	Geo	Set	*N*
Green & Stoneman (1989)	1	.000	*p*	*ns*	some	1	1	W	1	1	1	3	1	6	adult	b/u	1	m/o	117
Greenland & Brown (1999)	1	−.549	*r*	−.549*	some	1	1	W	1	1	3	2	1	1	coll	b/u	2	m/o	236
	2	+.177	*r*	+.177*	some	1	1	W	1	1	3	2	1	1	coll	b/u	2	m/o	40
Gregory (1997)	1	−.189	*r*	−.189	full	1	1	W	1	1	2	2	1	4	coll	b/u	1	m/o	140
Gronberg (1982)	1	−.341	*t*	−3.58*	none	2	2	B	2	2	99	2	2	4	child	b/u	1	edu	97
	2	−.346	*t*	−3.62*	none	2	2	B	2	2	99	2	2	4	child	b/u	1	edu	96
	3	−.336	*t*	−3.57*	none	2	2	B	2	2	99	2	2	4	child	b/u	1	edu	100
Gruesser (1950)	1	−.142	*t*	−3.73*	full	2	1	B	3	1	2	2	1	1	adol	b/u	1	*r*es	737
Gundlach (1950)	1	−.317	Prop	12/44*	some	1	1	B	3	3	99	1	1	1	adult	b/u	1	org	1,418
Haddock et al. (1993)	1	−.170	*r*	−.170*	full	1	1	W	1	1	2	2	1	3	coll	b/u	4	m/o	151
Hale (1998)	1	−.384	*t*	−2.94	some	1	1	B	3	1	2	2	1	2	adult	b/u	1	m/o	50
Hall (1969)	1	.000	*p*	*ns*	some	2	3	B	3	2	4	2	2	6	coll	b/u	1	org	264
Hall (1998)	1	−.186	*p*	.000*	full	2	2	B	2	2	4	3	2	99	child	b/u	1	*r*ec	303
Hamblin (1962)	1	−.230	*r*	−.230	some	1	1	W	1	1	2	3	1	1	adult	b/u	1	m/o	100
	2	−.170	*r*	−.170	some	1	1	W	1	1	2	3	1	1	adult	b/u	1	m/o	100
Hansen (1982)	1	−.484	*t*	−5.44	full	1	1	B	2	1	1	3	1	3	coll	b/u	1	m/o	107
Harding & Hogrefe (1952)	1	−.030	Prop	55/58*	full	1	1	B	2	1	1	3	1	1	adult	b/u	1	org	210
Haring et al. (1958)	1	−.250	*t*	−.266*	full	1	2	B	1	2	4	2	2	4	adult	b/u	1	edu	106
Haring et al. (1987)	1	−.248	*p*	.05*	full	1	3	B	2	2	4	4	2	1	adol	b/u	1	m/o	59
Harlan (1942)	1	−.804	*M/SD*	*p* = .000	full	1	1	B	1	1	1	3	1	1	coll	b/u	1	m/o	502
Harper & Wacker (1985)	1	−.196	O	*p* = .05	some	1	1	B	3	1	2	2	1	4	child	b/u	1	edu	100
Harris & Fiedler (1988)	1	.000	*p*	*ns*	some	1	1	W	1	1	2	2	1	2	child	b/u	1	edu	157
Hastings & Graham (1995)	1	−.107	*F*	5.97*	full	1	1	W	1	1	1	2	1	6	adol	b/u	2	edu	128
Hastings et al. (1998)	1	−.256	*F*	6.06*	some	1	1	B	3	1	1	3	1	4	coll	f	2	m/o	87
Hatanaka (1982)	1	−.225	*r*	−.225*	none	2	2	W	1	2	99	2	2	1	adult	b/u	1	lab	128
Hazzard (1983)	1	−.111	*r*	−.111*	full	1	1	W	1	1	2	3	1	4	child	b/u	1	m/o	367
Hébert et al. (2000)	1	−.200	*t*	−.285*	some	1	1	B	3	1	1	2	1	5	adol	b/u	4	m/o	284
Helmstetter et al. (1994)	1	−.211	*M/SD*	*p* = .006*	some	1	1	B	3	1	2	3	1	4	adol	b/u	1	m/o	161
Herek (1988)	1	−.064	*r*	−.064*	full	1	1	W	1	1	1	3	1	3	coll	f	1	m/o	73
	2	−.048	*r*	−.048*	full	1	1	W	1	1	1	3	1	3	coll	m	1	m/o	37
	3	−.135	*r*	−.135*	full	1	1	W	1	1	1	3	1	3	coll	f	1	m/o	220
	4	−.124	*r*	−.124*	full	1	1	W	1	1	1	3	1	3	coll	m	1	m/o	169
Herek (1999)	1	−.389	*M/SD*	*p* = .000*	full	2	1	B	2	1	2	3	1	3	adult	f	1	m/o	652
	2	−.302	*M/SD*	*p* = .000*	full	2	1	B	2	1	2	3	1	3	adult	m	1	m/o	524
Herek & Capitanio (1996)	1	−.354	*F*	52.3*	full	1	1	B	2	1	1	3	1	3	adult	b/u	1	m/o	422
Herek & Capitanio (1997)	1	−.189	*M/SD*	*p* = .000*	full	1	1	B	2	1	1	2	1	3	adult	b/u	1	m/o	594
Herek & Glunt (1993)	1	−.392	*F*	152.9	full	1	1	B	2	1	1	3	1	3	adult	b/u	1	m/o	937
Herman (1970)	1	−.160	Prop	37/53	full	1	1	W	1	2	99	2	1	1	coll	b/u	1	m/o	56
Hicks & Spaner (1962)	1	−.479	*t*	−4.69*	none	1	1	B	2	2	99	2	1	5	coll	b/u	1	org	78
	2	−.201	Mult	*p* = .01*	full	1	1	B	2	2	99	2	1	5	coll	b/u	1	org	330
Hill (1984)	1	−.316	*r*	−.316*	full	1	1	W	1	1	2	2	1	1	adult	b/u	2	m/o	200
Hillis (1986)	1	+.075	*t*	+0.77*	none	2	2	B	2	2	99	2	2	4	child	b/u	1	edu	117
Hillman & Stricker (1996)	1	−.280	*r*	−.280	some	1	1	W	1	1	2	3	1	2	coll	b/u	1	m/o	241
Hoeh & Spuck (1975)	1	−.283	*M/SD*	*p* = .121*	full	1	2	W	1	2	4	3	2	1	adol	b/u	1	*t*rav	15
Hofman & Zak (1969)	1	−.208	*p*	.046*	full	1	1	W	1	2	2	2	1	1	adol	b/u	6	m/o	46
Holmes et al. (1999)	1	−.157	*r*	−.157*	some	1	1	W	1	1	3	3	1	5	adult	b/u	1	m/o	83
Holtzman (1956)	1	−.148	χ^2	23.6*	some	1	1	W	1	1	2	2	1	1	coll	b/u	1	m/o	539
Holzberg & Gewirtz (1963)	1	−.526	*t*	−4.50	full	1	2	B	2	2	4	2	2	5	coll	b/u	1	org	59
Horenczyk & Bekerman (1997)	1	−.207	*M/SD*	*p* = .000	full	1	2	W	1	2	99	2	2	1	adol	b/u	6	*r*ec	148
	2	−.148	*M/SD*	*p* = .076	full	1	2	W	1	2	99	2	1	1	adol	b/u	6	*r*ec	72
Hortacsu (2000)	1	−.155	*r*	−.155	some	1	1	W	1	1	3	2	1	1	coll	m	6	edu	47
	2	−.034	*r*	−.034	some	1	1	W	1	1	3	2	1	1	coll	m	6	edu	49
Hraba et al. (1996)	1	−.208	*r*	−.208*	some	1	1	W	1	1	2	2	1	1	coll	b/u	1	m/o	208
	2	−.181	*r*	−.181*	some	1	1	W	1	1	2	2	1	1	coll	b/u	1	m/o	193
Hughey (1988)	1	−.166	*r*	−.166*	some	2	1	W	1	1	2	3	1	4	adult	b/u	1	edu	162
Hunt (1960)	1	−.158	*p*	.01	full	1	1	W	1	1	1	1	1	1	adult	b/u	1	m/o	133
Hunt & Hunt (2000)	1	−.186	*r*	−.186*	some	1	1	W	1	1	3	3	1	4	coll	b/u	1	m/o	274

(*Appendix continues*)

Appendix (*continued*)

Reference	Sample	r	Test	Statistic	Choice	Pub	Type	B/W	Control	IV	IVQ	DVQ	Prog	Target	Age	Sex	Geo	Set	N
Ibrahim (1970)	1	−.176	Prop	54/71	full	1	1	B	3	1	3	3	1	1	coll	b/u	1	m/o	402
Ichildov & Even-Dar (1984)	1	−.312	M/SD	p = .002*	full	1	1	W	1	2	2	3	2	1	adol	b/u	3	m/o	49
Iguchi & Johnson (1966)	1	−.222	Prop	32/55*	full	1	2	B	2	2	4	3	2	5	coll	b/u	1	res	98
Ijaz (1980)	1	−.108	F	1.92*	some	2	1	B	4	1	2	2	1	1	adol	b/u	4	edu	164
Ingamells et al. (1996)	1	−.199	t	−2.30	some	1	1	B	3	1	2	2	1	5	adult	b/u	2	m/o	133
Irish (1952)	1	−.063	p	.30*	some	1	1	B	4	1	99	2	1	1	adult	b/u	1	res	267
Islam & Hewstone (1993)	1	−.490	r	−.490*	full	1	1	W	1	1	2	1	1	1	coll	b/u	6	m/o	65
	2	−.240	r	−.240*	full	1	1	W	1	1	2	1	1	1	coll	b/u	6	m/o	66
Ivester & King (1977)	1	.000	p	ns	some	1	1	W	1	1	1	3	1	2	adol	b/u	1	m/o	413
Jackman & Crane (1986)	1	−.268	Prop	38/64*	full	1	1	B	3	1	2	2	1	1	adult	b/u	1	m/o	1,131
Jaffe (1966/1967)	1	−.115	M/SD	p = .21	full	1	1	B	2	1	1	2	1	6	adol	b/u	1	m/o	119
James (1955)	1	−.907	Prop	5/95	none	1	1	W	1	2	99	4	2	1	adol	b/u	2	edu	43
James-Valutis (1993)	1	−.188	r	−.188	some	2	1	W	1	1	2	3	1	1	coll	b/u	1	edu	213
Jaques et al. (1970)	1	.000	p	ns	some	1	1	W	1	1	1	3	1	4	coll	b/u	2	m/o	360
	2	.000	p	ns	some	1	1	W	1	1	1	3	1	4	coll	b/u	2	m/o	307
	3	−.111	p	.005	some	1	1	W	1	1	1	3	1	4	coll	b/u	2	m/o	322
Jeffries & Ransford (1969)	1	−.307	Prop	16/46*	full	1	1	B	2	1	2	1	1	1	adult	b/u	1	m/o	99
Johannsen et al. (1964)	1	−.153	p	.075*	full	1	2	B	3	2	99	3	1	5	coll	b/u	1	org	135
Johnson & Johnson (1981)	1	−.240	t	−1.56	none	1	3	B	2	2	4	4	2	5	child	b/u	1	edu	40
Johnson & Johnson (1985)	1	−.377	t	−1.82*	none	1	3	B	3	2	4	2	2	4	child	b/u	1	edu	20
Johnson & Marini (1998)	1	−.469	r	−.469	full	1	1	W	1	1	3	3	1	1	adol	b/u	1	edu	3,000
	2	−.557	r	−.557	full	1	1	W	1	1	3	3	1	1	adol	b/u	1	edu	2,648
Johnstone (1992)	1	−.258	r	−.258*	full	2	1	W	1	1	3	3	1	4	coll	b/u	1	m/o	185
	2	−.229	r	−.229*	full	2	1	W	1	1	3	3	1	4	coll	b/u	1	m/o	189
Jones (1960)	1	−.515	Prop	23/74*	some	2	1	B	4	2	99	4	1	1	adult	b/u	1	org	76
Jones et al. (1981)	1	−.364	M/SD	p = .006*	none	1	3	W	3	2	4	3	2	4	child	b/u	1	edu	25
	2	−.537	Mult	p = .000*	none	1	3	B	3	2	4	3	2	1	child	b/u	1	edu	74
Kalson (1976)	1	−.414	χ^2	5.14	full	1	1	W	1	2	99	2	2	6	adult	b/u	1	rec	15
Kamal & Maruyama (1990)	1	−.320	r	−.320	full	1	1	W	1	1	2	2	1	1	coll	b/u	1	m/o	187
Kanouse-Roberts (1977)	1	−.197	t	−1.17*	full	2	2	B	3	2	99	2	1	2	adol	f	1	org	34
Katz & Yochanan (1988)	1	−.533	M/SD	p = .000*	none	1	1	B	2	2	99	3	2	1	child	b/u	3	edu	108
Kelly et al. (1958)	1	−.286	r	−.286*	full	1	1	W	1	1	2	2	1	1	coll	b/u	1	m/o	547
Kephart (1957)	1	−.128	Prop	39/51*	some	1	1	B	3	1	1	2	1	1	adult	m	1	org	1,081
Kidwell & Booth (1977)	1	−.086	p	.08*	full	1	1	B	4	1	1	2	1	2	adult	b/u	1	edu	409
Kierscht & DuHoux (1980)	1	−.474	F	40.46	none	1	2	B	2	2	4	2	1	4	child	b/u	1	edu	140
Kisabeth & Richardson (1985)	1	−.175	M/SD	p = .277	none	1	2	B	3	2	4	2	2	4	coll	b/u	1	edu	41
Kish & Hood (1974)	1	−.334	p	.013	none	1	2	W	1	2	99	2	1	5	coll	b/u	1	org	28
	2	−.232	p	.31	none	1	2	W	1	2	99	2	1	5	coll	b/u	1	org	10
	3	−.228	p	.05	none	1	2	W	1	2	99	2	1	5	coll	b/u	1	org	37
Kishi & Meyer (1994)	1	−.270	M/SD	p = .049*	some	1	2	B	2	2	4	2	1	4	adol	m	1	edu	53
	2	−.317	M/SD	p = .006*	some	1	2	B	2	2	4	2	1	4	adol	f	1	edu	74
Kleinman (1983)	1	−.203	t	−2.63*	full	2	2	W	1	2	99	3	2	5	coll	f	1	org	40
Knox et al. (1986)	1	−.414	r	−.414*	full	1	1	W	1	1	3	2	1	2	coll	b/u	4	m/o	110
Knussen & Niven (1999)	1	−.130	r	−.130	some	1	1	W	4	1	2	2	1	99	adult	b/u	2	org	174
Kobe & Mulick (1995)	1	+.019	r	+.019*	full	1	1	W	1	1	2	3	1	6	coll	b/u	1	m/o	37
Kocarnik & Ponzetti (1986)	1	−.163	t	−.87*	some	1	1	B	3	3	99	2	2	2	child	b/u	1	edu	30
Koslin et al. (1969)	1	−.339	M/SD	p = .000*	some	1	2	B	2	2	4	2	1	1	child	b/u	1	edu	64
	2	−.344	M/SD	p = .000	some	1	2	B	2	2	4	2	1	1	child	b/u	1	Edu	65
Kosmitzki (1996)	1	−.106	t	−1.24	full	1	1	B	2	1	2	3	1	1	adult	b/u	1	m/o	254
	2	−.187	t	−1.98	full	1	1	B	2	1	2	3	1	1	adult	b/u	1	m/o	137
Krajewski & Flaherty (2000)	1	−.206	M/SD	p = .085*	some	1	1	B	3	1	1	2	1	6	adol	m	1	m/o	70
	2	−.110	M/SD	p = .177*	some	1	1	B	3	1	1	2	1	6	adol	f	1	m/o	74

Appendix (*continued*)

Reference	Sample	*r*	Test	Statistic	Choice	Pub	Type	B/W	Control	IV	IVQ	DVQ	Prog	Target	Age	Sex	Geo	Set	*N*
Kuelker (1996)	1	−.128	*r*	−.128*	some	2	1	W	1	1	2	3	1	5	adult	b/u	4	m/o	489
Kulik et al. (1969)	1	−.123	*t*	−2.20	full	1	1	B	2	2	99	2	1	5	coll	b/u	1	org	318
Kurtzweil (1995)	1	−.162	*r*	−.162*	some	2	1	W	1	1	2	3	1	1	coll	b/u	1	m/o	240
Ladd et al. (1984)	1	−.544	*F*	26.92	some	1	1	W	1	2	99	4	2	4	adol	b/u	1	edu	16
Lambert et al. (1990)	1	−.225	Prop	82/96*	some	1	2	B	3	2	4	2	2	2	child	b/u	1	edu	31
Lance (1987)	1	−.450	Prop	18/61	some	1	2	B	2	1	99	2	1	3	coll	b/u	1	edu	46
Lance (1992)	1	−.332	Prop	35/68	none	1	1	B	2	2	99	3	1	3	coll	b/u	1	edu	228
Lance (1994)	1	−.294	Prop	11/35*	full	1	1	B	3	1	1	3	2	3	coll	b/u	1	m/o	140
Landis et al. (1985)	1	−.570	*M/SD*	*p* = .015*	none	1	3	B	4	2	4	4	2	1	coll	m	1	lab	18
Larsen (1997)	1	−.261	*r*	−.261*	some	2	1	W	1	1	2	2	1	1	adult	b/u	1	m/o	11
	2	−.110	*r*	−.110*	some	2	1	W	1	1	2	2	1	1	adult	b/u	1	m/o	6
Lazar et al. (1971)	1	−.332	*M/SD*	−2.33	none	1	2	B	2	2	4	3	2	4	child	b/u	1	edu	44
Leach (1990)	1	−.206	Mult	*p* = .041*	some	2	1	B	3	1	1	2	1	4	coll	b/u	1	m/o	98
Lebhart & Munz (1999)	1	−.179	Prop	39/57*	full	2	1	B	3	1	1	2	1	1	adult	b/u	2	m/o	1,999
Leonard (1964)	1	−.219	*t*	−4.15	full	1	1	W	1	2	99	2	1	1	coll	b/u	2	*t*rav	85
Lepore & Brown (1997)	1	−.410	*r*	−.410	full	1	1	W	1	1	2	3	1	1	coll	b/u	2	edu	162
Lessing et al. (1976)	1	−.047	*F*	2.35	some	1	1	W	1	1	1	2	1	1	adult	f	1	edu	269
LeUnes et al. (1975)	1	−.399	*F*	32.4*	full	1	2	B	2	2	4	2	1	6	coll	b/u	1	*r*es	179
Levine et al. (1969)	1	−.203	Prop	45/65*	some	2	1	W	1	1	1	1	1	1	adol	b/u	1	m/o	419
	2	−.232	Prop	40/63*	some	2	1	W	1	1	1	1	1	1	adol	b/u	1	m/o	500
Levinson (1954)	1	−.220	*p*	.05	full	1	1	W	1	2	99	2	2	1	adult	b/u	1	edu	28
	2	−.310	*p*	.01	full	1	1	W	1	2	99	2	2	1	adult	b/u	1	edu	28
Levinson & Schermerhorn (1951)	1	−.192	*M/SD*	*p* = .125	full	1	1	W	1	2	99	2	2	1	adult	b/u	1	edu	32
Levy et al. (1993)	1	−.199	*F*	13.4*	some	1	1	B	2	1	2	3	1	4	adult	b/u	1	org	324
Lewis & Cleveland (1966)	1	−.187	*p*	.031*	some	1	2	B	3	2	99	3	2	5	coll	b/u	1	org	134
Lewis & Frey (1988)	1	−.439	*F*	15.2*	some	1	2	B	2	2	99	3	2	99	coll	b/u	1	edu	66
Leyser & Abrams (1983)	1	−.210	*M/SD*	*p* = .000	some	1	2	B	3	2	99	3	2	4	coll	b/u	1	edu	289
Leyser & Price (1985)	1	−.111	*M/SD*	*p* = .39	none	1	2	B	2	2	99	2	1	4	child	b/u	1	edu	60
Leyser et al. (1986)	1	−.176	*M/SD*	*p* = .005*	none	1	2	B	3	2	99	3	2	4	child	b/u	1	edu	244
Li & Yu (1974)	1	−.007	*M/SD*	*p* = .92*	full	1	1	B	2	1	1	2	1	1	coll	b/u	1	m/o	220
	2	+.043	*M/SD*	*p* = .60	full	1	1	B	2	1	1	2	1	1	coll	b/u	1	m/o	145
Liebkind et al. (2000)	1	−.325	*r*	−.325	some	1	1	W	1	1	2	3	1	1	adult	b/u	2	m/o	104
	2	−.185	*r*	−.185	some	1	1	W	1	1	2	3	1	1	adult	b/u	2	m/o	185
	3	−.152	*r*	−.152	some	1	1	W	1	1	2	3	1	1	adult	b/u	2	m/o	86
Link & Cullen (1986)	1	−.266	*M/SD*	*p* = .001*	some	1	1	B	3	1	3	3	1	5	adult	b/u	1	m/o	153
	2	−.249	*M/SD*	*p* = .003*	some	1	1	B	3	1	3	3	1	5	adult	b/u	1	m/o	151
Lombardi (1963)	1	.000	*p*	*ns*	some	1	1	B	3	1	2	3	2	1	adol	b/u	1	edu	344
Lombroso et al. (1976)	1	−.096	O	*p* = .01	some	1	1	W	1	1	1	2	1	5	adol	b/u	3	m/o	360
London & Linney (1993)	1	−.064	*M/SD*	*p* = .634	full	2	1	W	1	2	99	2	2	1	child	b/u	1	*r*ec	28
	2	−.141	*M/SD*	*p* = .362	full	2	1	W	1	2	99	2	2	1	child	b/u	1	*r*ec	21
Loomis & Schuler (1948)	1	+.195	O	*p* = .03*	full	1	1	W	1	3	99	2	1	1	adult	b/u	6	*t*rav	62
Lopes & Vala (2000)	1	−.207	*F*	23.2*	full	2	1	B	3	1	1	2	1	1	adult	b/u	2	m/o	520
Lopez (1993)	1	−.086	*r*	−.086*	none	2	1	W	1	2	99	2	1	1	coll	b/u	1	edu	480
	2	−.072	*r*	−.072*	none	2	1	W	1	2	99	2	1	1	coll	b/u	1	edu	165
	3	−.045	*r*	−.045*	none	2	1	W	1	2	99	2	1	1	coll	b/u	1	edu	92
Luiz & Krige (1981/1985)	1	−.313	*t*	−2.09*	some	1	2	W	1	2	4	3	2	1	adol	f	6	edu	10
	2	−.567	*t*	−4.13	some	1	2	W	1	2	4	3	2	1	adol	f	6	edu	9
MacKenzie (1948)	1	−.276	Prop	16/42*	full	1	1	B	3	1	1	2	1	1	adult	m	1	org	36
	2	−.169	Prop	14/28*	full	1	1	B	3	1	1	2	1	1	adult	f	1	org	88
	3	−.206	Prop	21/41*	full	1	1	B	3	1	1	2	1	1	adult	m	1	org	40
	4	−.180	Prop	29/46*	full	1	1	B	3	1	1	2	1	1	adult	f	1	org	26
	5	−.265	Prop	41/67*	full	1	1	B	3	1	1	2	1	1	adult	m	1	org	142
	6	−.290	Prop	24/58*	full	1	1	B	3	1	1	2	1	1	adult	f	1	org	85
	7	−.347	Prop	20/53*	full	1	1	B	3	1	1	2	1	1	adult	b/u	1	org	118
MacLean & Gannon (1995)	1	.000	*p*	*ns*	full	1	1	W	1	1	3	3	1	4	coll	b/u	5	m/o	341
Malla & Shaw (1987)	1	−.127	*p*	.280*	full	1	2	B	3	2	4	2	1	5	coll	f	4	edu	71
Maluso (1992)	1	−.136	*M/SD*	*p* = .013*	none	2	2	B	4	2	99	2	2	1	coll	b/u	1	edu	339
Mann (1959/1960)	1	−.207	*p*	.01	some	1	1	W	1	2	4	4	2	1	adult	b/u	1	edu	78
Maoz (2000)	1	−.265	*t*	−3.89*	some	1	2	W	1	2	99	2	2	1	adol	b/u	3	edu	50
	2	−.137	*t*	−1.87*	some	1	2	W	1	2	99	2	2	1	adol	b/u	3	edu	46

(*Appendix continues*)

Appendix (*continued*)

Reference	Sample	*r*	Test	Statistic	Choice	Pub	Type	B/W	Control	IV	IVQ	DVQ	Prog	Target	Age	Sex	Geo	Set	*N*
Maras & Brown (1996)	1	−.329	*F*	5.40*	full	1	2	B	2	2	4	2	2	4	child	b/u	2	edu	44
Marin (1984)	1	−.244	*r*	−.244*	full	1	1	W	1	1	1	2	1	1	coll	b/u	1	m/o	100
Marin & Salazar (1985)	1	+.325	*r*	+.325*	full	1	1	W	1	1	2	2	1	1	coll	b/u	6	m/o	1,184
Marion (1980)	1	−.192	*p*	.01	some	1	1	W	1	2	99	3	1	1	coll	b/u	1	m/o	90
Marks (1992)	1	−.150	*r*	−.150*	none	2	1	W	1	1	3	3	1	99	adult	b/u	1	org	293
Martin (2000)	1	−.270	Mult	*p* = .000*	some	2	1	B	2	1	2	2	1	1	adult	b/u	2	m/o	420
Marx (1967)	1	+.117	Prop	30/20	some	1	1	B	4	1	2	2	1	1	adult	b/u	1	m/o	782
	2	−.014	Prop	11/12	some	1	1	B	4	1	2	2	1	1	adult	b/u	1	m/o	154
Masson & Verkuyten (1993)	1	−.406	*r*	−.406*	full	1	1	W	1	1	2	3	1	1	adol	b/u	2	m/o	160
Mathisen (2000)	1	−.099	*r*	−.099*	none	1	2	W	1	1	2	2	1	5	adol	b/u	1	edu	132
Maurice et al. (1975)	1	−.161	*t*	−1.82	full	2	1	W	1	2	99	1	1	5	adult	b/u	4	org	31
Maxmen (1979)	1	−.098	*t*	−2.08*	full	1	2	W	1	2	99	2	1	5	coll	b/u	1	edu	111
McClenahan et al. (1996)	1	−.073	*p*	.17*	some	1	1	W	1	3	99	2	1	1	adol	b/u	2	edu	167
	2	−.074	*p*	.30*	some	1	1	W	1	3	99	2	1	1	adol	b/u	2	edu	96
McConkey et al. (1983)	1	−.163	Prop	36/52*	full	1	1	B	2	1	2	2	1	5	adol	f	2	m/o	858
	2	−.094	Prop	62/71*	full	1	1	B	2	1	2	2	1	5	adol	m	2	m/o	482
McCrady & McCrady (1976)	1	−.064	*p*	.43*	full	1	1	W	1	3	99	2	1	1	coll	b/u	1	trav	77
McDonald et al. (1987)	1	−.509	*t*	−8.33*	none	1	1	B	3	3	99	2	1	6	child	b/u	5	edu	198
	2	−.279	*t*	−2.80	none	1	1	B	3	3	99	2	1	6	adult	b/u	5	edu	104
McGuigan (1959)	1	−.318	O	*p* = .000*	full	1	2	B	2	2	99	2	1	1	coll	f	1	m/o	179
McKirnan & Hamayan (1984)	1	−.237	*F*	11.5*	full	1	1	W	1	1	1	3	1	1	adol	b/u	1	m/o	48
McRainey (1981)	1	−.209	*p*	.31*	some	2	2	B	2	2	99	3	1	4	adult	f	1	edu	24
Meer & Freedman (1966)	1	−.096	*t*	−0.68	some	1	1	B	2	3	99	3	1	1	adult	b/u	1	res	100
Merkwan & Smith (1999)	1	.000	*p*	*ns*	some	1	2	W	1	3	99	2	2	1	adult	b/u	1	m/o	27
Meshel (1997)	1	−.222	*M/SD*	*p* = .152*	some	2	3	B	3	2	4	3	2	2	adol	b/u	1	edu	42
	2	−.429	*M/SD*	*p* = .012	some	2	3	W	1	2	4	3	2	2	adult	b/u	1	edu	17
Meyer et al. (1980)	1	+.411	*p*	.001	full	1	2	B	3	2	99	3	1	2	adult	m	1	org	64
Milem (1992)	1	−.203	*r*	−.203*	full	2	1	W	1	1	1	3	1	1	coll	m	1	edu	3,000
	2	−.221	*r*	−.221*	full	2	1	W	1	1	1	3	1	1	coll	f	1	edu	3,000
Miller et al. (1998)	1	−.185	*t*	−3.63	full	2	1	W	1	1	2	2	1	1	coll	b/u	1	m/o	93
Millham et al. (1976)	1	−.037	*F*	4.25	full	1	1	W	1	1	1	2	1	3	coll	b/u	1	m/o	795
Mills et al. (1998)	1	−.019	*M/SD*	*p* = .844*	some	1	1	B	4	1	1	2	1	2	coll	b/u	1	m/o	104
Moeschl (1978)	1	−.066	*r*	−.066	some	2	1	W	1	1	2	2	1	2	coll	b/u	1	m/o	144
Mohr & Rochlen (1999)	1	−.325	*F*	38.7*	full	1	1	W	1	1	2	3	1	3	adult	b/u	1	m/o	305
	2	−.186	*r*	−.186*	full	1	1	W	1	1	2	3	1	3	adult	b/u	1	m/o	315
Monroe & Howe (1971)	1	+.632	*M/SD*	*p* = .000	some	1	1	B	4	3	99	2	1	6	adol	m	1	edu	70
Morin (1974)	1	−.263	Mult	*p* = .110*	full	1	1	W	1	2	99	2	1	3	coll	b/u	1	edu	18
Morris (1964)	1	−.168	*t*	−2.43*	some	1	1	W	1	2	4	3	2	5	coll	f	1	org	51
Morris & Jeffries (1968)	1	−.155	Mult	*p* = .000*	some	1	1	B	3	1	2	2	1	1	adult	b/u	1	m/o	583
Mosher-Ashley & Ball (1999)	1	−.135	MW	*p* = .07	none	1	1	B	3	1	1	2	1	2	coll	b/u	1	res	123
Most et al. (1999)	1	−.070	*M/SD*	*p* = .55*	some	1	1	B	3	2	99	3	1	4	adol	m	3	edu	70
	2	−.010	*M/SD*	*p* = .93*	some	1	1	B	3	2	99	3	1	4	adol	f	3	edu	70
Murphy et al. (1993)	1	.000	*p*	*ns*	some	1	1	W	1	1	1	3	1	6	adult	b/u	2	m/o	155
Murphy-Russell et al. (1986)	1	−.302	*M/SD*	*p* = .006*	none	1	2	B	3	2	4	3	1	2	coll	b/u	1	lab	84
Mussen (1950)	1	−.033	*M/SD*	*p* = .76	full	1	1	W	1	2	99	3	1	1	child	m	1	rec	106
Nabuzoka & Renning (1997)	1	−.796	*M/SD*	*p* = .000	some	1	2	B	3	2	99	2	1	6	child	b/u	6	edu	20
	2	+.183	*M/SD*	*p* = .41	some	1	2	B	3	2	99	2	1	6	child	b/u	6	edu	10
Naor & Milgram (1980)	1	−.383	Mult	*p* = .001*	full	1	2	B	2	2	99	2	2	6	adult	f	3	edu	80
Narukawa (1995)	1	−.083	*M/SD*	*p* = .21*	some	1	1	B	3	1	1	3	1	6	coll	m	6	m/o	228
	2	−.085	*M/SD*	*p* = .20*	some	1	1	B	3	1	1	3	1	6	coll	f	6	m/o	230
Nash (1976)	1	+.276	*r*	+.276*	full	1	2	B	2	2	99	2	1	1	coll	b/u	1	m/o	73
Naus (1973)	1	−.126	*r*	−.126*	some	1	1	W	1	1	2	3	1	2	coll	b/u	1	m/o	43
NCCJ (2000)	1	−.143	Prop	47/61*	some	1	1	B	3	1	2	2	1	1	adult	b/u	1	m/o	2,586
Neprash (1953)	1	−.312	Prop	11/27*	full	1	1	B	4	1	2	2	1	1	child	m	1	res	61
Nesdale & Todd (1998)	1	−.224	*p*	.011*	some	1	1	B	3	3	99	2	1	1	coll	b/u	5	edu	127
	2	−.210	*p*	.02*	some	1	1	B	3	3	99	2	1	1	coll	b/u	5	edu	119
Nesdale & Todd (2000)	1	−.203	*p*	.09*	some	1	2	B	4	1	2	2	1	1	coll	b/u	1	edu	69
Neto (2000)	1	−.460	*r*	−.460	none	1	1	W	1	1	3	2	1	1	adol	b/u	1	edu	118

Appendix (*continued*)

Reference	Sample	*r*	Test	Statistic	Choice	Pub	Type	B/W	Control	IV	IVQ	DVQ	Prog	Target	Age	Sex	Geo	Set	*N*
Newberry & Parish (1987)	1	−.630	*F*	43.57	none	1	2	B	2	2	4	2	1	4	child	b/u	1	rec	114
	2	−.722	*F*	67.76	none	1	2	B	2	2	4	2	1	4	child	b/u	1	rec	90
	3	−.525	*F*	22.44	none	1	2	B	2	2	4	2	1	4	child	b/u	1	rec	76
	4	−.517	*F*	20.84	none	1	2	B	2	2	4	2	1	4	child	b/u	1	rec	70
	5	−.100	*F*	0.66	none	1	2	B	2	2	4	2	1	4	child	b/u	1	rec	105
Newswanger (1996)	1	−.158	Prop	58/73*	none	1	2	B	4	2	99	1	1	1	coll	b/u	1	res	144
Ng et al. (1997)	1	−.235	*r*	−.235*	full	1	1	W	1	1	3	1	1	2	adult	b/u	5	m/o	100
Nieuwoudt & Thom (1980)	1	−.430	*M/SD*	*p* = .000	none	1	2	W	1	2	4	3	2	2	coll	b/u	6	m/o	290
	2	−.497	*M/SD*	*p* = .000	none	1	2	W	1	2	4	3	2	2	coll	b/u	6	m/o	467
	3	−.576	*M/SD*	*p* = .000	none	1	1	W	1	2	4	3	2	2	coll	b/u	6	m/o	27
	4	−.375	*M/SD*	*p* = .001	none	1	2	W	1	2	4	3	2	2	coll	b/u	6	m/o	42
Nishi-Strattner & Myers (1983)	1	.000	*p*	*ns*	some	1	1	W	1	1	2	3	1	3	child	b/u	1	m/o	52
	2	.000	*p*	*ns*	some	1	1	W	1	1	2	3	1	3	adult	b/u	1	m/o	52
Noels & Clément (1996)	1	−.496	*r*	−.496*	full	1	2	W	1	1	3	3	1	1	coll	b/u	4	m/o	125
	2	−.458	*r*	−.458*	full	1	2	W	1	1	3	3	1	1	coll	b/u	4	m/o	94
Nosse (1993)	1	−.335	O	*p* = .002	full	1	2	W	1	2	99	3	2	4	coll	b/u	1	m/o	37
Nosse & Gavin (1991)	1	−.249	*p*	.05*	full	1	1	W	1	2	99	3	2	4	coll	b/u	1	m/o	31
Oaker & Brown (1986)	1	−.400	*r*	−.400*	some	1	1	W	1	1	1	1	1	99	coll	f	2	org	23
	2	+.135	*r*	+.135*	some	1	1	W	1	1	1	1	1	99	coll	f	2	org	17
Ogendengbe (1993)	1	−.464	Prop	23/76	full	1	1	B	2	1	1	2	1	5	adult	b/u	6	m/o	174
Olejnik & LaRue (1981)	1	−.388	Prop	56/93	some	1	2	B	3	2	4	2	2	2	adol	b/u	1	edu	634
Pagtolun-an & Clair (1986)	1	−.243	*F*	4.47*	none	1	3	B	3	2	4	3	1	3	coll	b/u	1	edu	71
Palmerton & Frumkin (1969)	1	−.329	*r*	−.329	some	1	1	W	1	1	1	2	1	4	adult	b/u	1	m/o	81
Paris (1991)	1	−.115	*t*	−1.78	some	2	1	B	4	1	2	3	1	4	adult	b/u	1	edu	237
Parker et al. (1998)	1	−.234	*M/SD*	*p* = .06*	full	1	2	B	3	2	3	2	2	1	adult	b/u	1	edu	54
Patchen (1982/1983)/ Patchen et al. (1977)	1	−.189	*r*	−.189*	some	1	1	W	1	1	2	2	1	1	adol	b/u	1	edu	2,146
	2	−.125	*r*	−.125*	some	1	1	W	1	1	2	2	1	1	adol	b/u	1	edu	1,986
Penn et al. (1994)	1	−.064	*p*	.24	full	1	1	B	2	1	1	3	1	5	coll	b/u	1	m/o	329
Peterson (1974)	1	−.059	χ^2	1.42*	some	1	1	B	2	3	99	3	1	6	child	b/u	1	edu	420
Petrangelo (1976)	1	−.209	*t*	−1.79*	full	2	1	B	4	2	99	3	1	4	coll	b/u	1	edu	172
Pettigrew (1997)	1	−.259	*r*	−.259*	full	1	1	W	1	1	3	3	1	1	adult	b/u	1	m/o	437
	2	−.272	*r*	−.272*	full	1	1	W	1	1	3	3	1	1	adult	b/u	1	m/o	453
	3	−.319	*r*	−.319*	full	1	1	W	1	1	3	3	1	1	adult	b/u	1	m/o	962
	4	−.388	*r*	−.388*	full	1	1	W	1	1	3	3	1	1	adult	b/u	1	m/o	449
	5	−.423	*r*	−.423*	full	1	1	W	1	1	3	3	1	1	adult	b/u	1	m/o	452
	6	−.298	*r*	−.298*	full	1	1	W	1	1	3	3	1	1	adult	b/u	1	m/o	470
	7	−.341	*r*	−.341*	full	1	1	W	1	1	3	3	1	1	adult	b/u	1	m/o	455
Petzel (2000)	1	−.281	*r*	−.281*	some	2	1	W	1	1	1	3	1	1	adult	b/u	2	res	553
Philips (1963)	1	−.020	*p*	.73*	some	1	1	B	2	1	1	3	1	5	adult	f	1	m/o	300
Phinney ey al. (1997)	1	−.364	*r*	−.364	some	1	1	W	1	1	3	3	1	1	adol	b/u	1	edu	261
	2	−.523	*r*	−.523	some	1	1	W	1	1	3	3	1	1	adol	b/u	1	edu	286
Pinquart et al. (2000)	1	+.030	*M/SD*	*p* = .88	full	1	2	B	3	2	4	2	2	2	child	b/u	2	org	32
	2	−.153	*M/SD*	*p* = .49	full	1	2	B	3	2	4	2	2	2	adult	b/u	2	org	20
Pleck et al. (1988)	1	−.230	*r*	−.230	some	1	1	W	1	1	3	3	1	3	adult	b/u	1	org	237
Porter & O'Connor (1978)	1	−.188	*t*	−2.10	some	1	2	W	1	2	4	2	2	2	coll	b/u	1	edu	30
Prather & Chovan (1984)	1	−.010	*p*	.94	full	1	1	W	1	2	99	2	1	5	child	b/u	1	edu	25
Preston & Robinson (1974)/Robinson & Preston (1976)	1	−.062	*p*	.44*	some	1	2	B	3	2	4	2	2	1	adult	b/u	1	org	154
	2	−.435	*p*	.0000*	some	1	2	B	3	2	4	2	2	1	adult	b/u	1	org	116
Price (2000)	1	−.078	*t*	−1.25*	full	2	2	W	1	2	99	3	2	1	adol	b/u	5	m/o	64
Proller (1989)	1	−.267	*M/SD*	*p* = .044	full	1	2	B	3	2	99	2	2	2	child	b/u	1	res	57
	2	+.067	*M/SD*	*p* = .60	full	1	2	B	3	2	99	2	2	2	adult	b/u	1	res	61
	3	−.240	*M/SD*	*p* = .078	full	1	2	B	3	2	99	2	2	2	adult	b/u	1	res	54
	4	−.136	Mult	*p* = .326*	full	1	2	W	1	2	99	2	2	2	child	b/u	1	res	26
Pryer et al. (1969)	1	−.108	*p*	.34*	full	1	2	W	1	2	4	2	1	5	adult	b/u	1	res	39
	2	−.202	*p*	.153*	full	1	2	W	1	2	4	2	1	5	adult	b/u	1	res	25
Rabushka (1969)	1	−.156	Prop	59/73*	full	1	1	B	3	1	1	1	1	1	coll	b/u	6	edu	86
	2	−.364	Prop	33/60*	full	1	1	B	3	1	1	1	1	1	coll	b/u	6	edu	51

(*Appendix continues*)

Appendix (*continued*)

Reference	Sample	*r*	Test	Statistic	Choice	Pub	Type	B/W	Control	IV	IVQ	DVQ	Prog	Target	Age	Sex	Geo	Set	*N*
Radcliffe (1972)	1	−.096	*F*	1.22*	full	2	1	W	1	1	2	3	1	1	adult	b/u	1	edu	219
Ralph (1968)	1	.000	*p*	*ns*	full	1	2	W	1	2	99	3	2	5	coll	f	1	m/o	40
Rapier et al. (1972)	1	−.114	O	*p* = .05*	some	1	1	W	1	3	99	2	1	4	child	b/u	1	edu	148
Read & Law (1999)	1	−.247	*M/SD*	*p* = .006*	some	1	1	B	2	1	1	2	1	5	coll	b/u	5	m/o	126
Reed (1980)	1	−.044	Prop	56/60*	some	1	1	B	4	1	1	2	1	99	adult	b/u	1	m/o	716
Reigrotski & Anderson (1959)	1	−.194	Prop	49/68	full	1	1	B	2	1	2	2	1	1	adult	b/u	2	m/o	2,006
	2	−.168	Prop	45/62	full	1	1	B	2	1	2	2	1	1	adult	b/u	2	m/o	2,041
Reinsch & Tobis (1991)	1	−.175	*M/SD*	*p* = .338	full	1	1	B	4	2	4	2	2	2	coll	b/u	1	org	30
Rich et al. (1983)	1	−.155	*t*	−1.56	none	1	2	B	2	2	4	2	2	2	child	b/u	1	lab	99
Rich et al. (1995)	1	−.100	*p*	.27*	some	1	2	W	1	2	4	3	2	1	adol	b/u	3	edu	60
Rimmerman (1998)	1	−.097	*p*	.433*	some	1	1	W	1	1	2	3	1	6	adult	b/u	3	m/o	120
Rimmerman et al. (2000)	1	−.268	*F*	6.92	some	1	1	B	3	1	1	2	1	4	coll	b/u	3	edu	102
Riordan (1987)	1	−.189	*t*	−1.97	full	1	1	B	4	1	2	2	1	1	adult	b/u	1	m/o	102
Robbins et al. (1992)	1	−.110	*r*	−.110	some	1	1	W	1	1	2	3	1	3	coll	b/u	2	edu	112
	2	−.280	*r*	−.280	some	1	1	W	1	1	2	3	1	3	coll	b/u	2	edu	63
	3	−.270	*r*	−.270	some	1	1	W	1	1	2	3	1	3	coll	b/u	2	edu	28
Roberts (1988)	1	−.146	*r*	−.146*	some	1	1	W	1	1	1	3	1	1	adult	b/u	1	m/o	745
Robinson (1987)	1	−.410	*r*	−.410*	some	2	1	W	1	1	1	3	1	99	coll	b/u	1	edu	781
Renning & Nabuzoka (1993)	1	−.260	O	*p* = .142*	none	1	2	W	1	2	99	4	2	6	child	b/u	6	edu	16
Rooney & Jason (1986)	1	−.562	*F*	16.6*	some	1	2	W	4	2	4	4	2	1	child	b/u	1	edu	36
	2	+.217	*F*	0.99*	some	1	2	B	4	2	4	4	2	1	child	b/u	1	edu	20
Roper (1990)	1	−.171	*M/SD*	*p* = .002*	full	1	1	B	3	1	2	2	1	6	adult	b/u	1	rec	331
Rose (1961)	1	−.120	Prop	55/67*	full	1	1	B	4	1	99	2	1	1	adult	b/u	1	m/o	175
Rose et al. (1953)	1	−.239	Prop	33/57*	full	1	1	B	3	1	1	2	1	1	adult	b/u	1	res	471
Rosenblith (1949)	1	−.098	Prop	46/56	full	1	1	B	3	1	2	2	1	1	coll	b/u	1	m/o	859
Rosencranz & McNevin (1969)	1	−.103	*p*	.08*	some	1	1	B	3	1	2	2	1	2	coll	b/u	1	m/o	287
Rowlett (1981)	1	−.475	*M/SD*	*p* = .016	full	2	2	B	3	2	4	3	1	4	coll	b/u	1	edu	26
	2	−.507	*M/SD*	*p* = .011	full	2	2	B	3	2	4	3	1	4	coll	b/u	1	edu	25
Rusalem (1967)	1	−.370	*p*	.05	none	1	2	W	1	2	99	2	2	4	adol	b/u	1	edu	14
	2	.000	*p*	*ns*	none	1	2	W	1	2	99	2	2	4	adol	b/u	1	edu	14
Rusinko et al. (1978)	1	−.068	*M/SD*	*p* = .002*	some	1	1	W	1	1	3	2	1	99	adol	b/u	1	m/o	1,020
Sadler & Blair (1999)	1	−.357	*r*	−.357*	full	2	1	W	1	1	2	2	1	3	coll	b/u	1	edu	251
Sakaris (2000)	1	−.120	*r*	−.120*	none	2	1	W	1	1	2	3	1	99	adult	b/u	1	edu	77
Salter & Teger (1975)	1	−.231	*p*	.238*	full	1	1	W	1	2	99	2	1	1	coll	b/u	1	org	13
	2	−.393	*p*	.01*	full	1	1	W	1	2	99	2	1	1	coll	b/u	1	trav	22
San Miguel & Millham (1976)	1	−.145	*t*	−1.44*	none	1	3	B	2	2	4	4	1	3	coll	m	1	lab	96
Sandberg (1982)	1	−.062	*M/SD*	*p* = .217	some	1	1	B	2	3	99	2	1	6	child	b/u	1	edu	400
Sayler (1969)	1	−.154	*F*	1.38*	none	2	3	B	3	2	4	3	1	1	coll	b/u	1	edu	140
	2	−.074	*F*	0.35	none	2	3	B	3	2	4	3	1	1	coll	b/u	1	edu	64
Scarberry et al. (1996)	1	−.426	*M/SD*	*p* = .016*	none	2	3	W	1	2	4	2	2	99	coll	b/u	1	lab	16
Schcibe (1965)	1	−.152	*t*	−3.07*	full	1	1	W	1	2	4	3	1	5	coll	b/u	1	org	99
Schneider (1994)	1	−.302	*r*	−.302*	some	2	1	W	1	1	1	1	1	1	adol	b/u	2	m/o	237
	2	−.291	*r*	−.291	some	2	1	W	1	1	1	1	1	1	adol	b/u	2	m/o	557
Schneider & Lewis (1984)	1	−.240	Prop	37/61	full	1	1	B	2	1	1	1	1	3	adult	m	1	m/o	800
	2	−.289	Prop	33/62	full	1	1	B	2	1	1	1	1	3	adult	f	1	m/o	853
Schwarzwald et al. (1985)	1	−.305	*F*	2.60	some	1	1	W	1	3	99	2	1	1	adol	b/u	3	edu	2,530
Seefeldt (1987)	1	+.355	*p*	.006*	none	1	1	B	4	2	99	2	1	2	child	b/u	1	res	60
Segal (1965)	1	−.258	*p*	.01	full	1	1	B	4	1	1	2	1	1	coll	m	1	edu	100
	2	−.288	*p*	.01	full	1	1	B	4	1	1	2	1	1	coll	m	1	edu	80
Sellin & Mulchahay (1965)	1	−.101	Prop	56/66*	some	1	2	W	1	2	99	2	1	6	adol	b/u	1	res	144
Selltiz et al. (1963)	1	−.102	*p*	.058*	full	1	1	B	4	1	2	2	1	1	coll	m	1	edu	343
	2	−.134	*p*	.062	full	1	1	B	4	1	2	2	1	1	coll	m	1	edu	194
Selznick & Steinberg (1969)	1	−.125	Prop	24/36	some	1	1	W	1	1	2	1	1	1	adult	b/u	1	m/o	1,302
	2	+.024	Prop	42/39	some	1	1	W	1	1	2	1	1	1	adult	b/u	1	m/o	175
Semmel & Dickson (1966)	1	−.374	*M/SD*	*p* = .000	full	1	1	B	2	1	1	2	1	4	coll	b/u	1	m/o	426
Sewell & Davidsen (1956)	1	−.219	*p*	.05*	full	1	1	W	1	1	99	4	1	1	coll	b/u	2	m/o	40

Appendix (*continued*)

Reference	Sample	*r*	Test	Statistic	Choice	Pub	Type	B/W	Control	IV	IVQ	DVQ	Prog	Target	Age	Sex	Geo	Set	*N*
Shafer et al. (1989)	1	−.091	*F*	1.00*	some	1	1	B	3	1	2	2	1	6	adult	b/u	1	org	212
Sheare (1974)	1	−.468	*F*	112.2	none	1	3	B	2	2	4	2	1	6	adol	m	1	res	400
Sheehan (1980)	1	−.052	*t*	−2.89*	some	1	2	W	1	2	99	2	1	1	child	b/u	1	edu	759
	2	−.104	*t*	−5.43*	some	1	2	W	1	2	99	2	1	1	child	b/u	1	edu	674
	3	−.116	*t*	−3.84*	some	1	2	W	1	2	99	2	1	1	child	b/u	1	edu	272
Shera & Delva-Tauiliili (1996)	1	−.414	*M/SD*	$p = .007$	none	1	2	B	2	2	4	3	2	5	coll	b/u	1	edu	33
	2	−.427	*M/SD*	$p = .01$	none	1	2	B	2	2	4	3	2	5	coll	b/u	1	edu	39
Sherif et al. (1961)	1	−.581	χ^2	7.43*	none	1	3	W	1	2	4	2	2	99	adol	m	1	rec	11
	2	−.499	χ^2	4.50*	none	1	3	W	1	2	4	2	2	99	adol	m	1	rec	9
Shibuya (2000)	1	−.190	*r*	−.190	some	1	1	W	1	1	1	3	1	1	adult	m	6	m/o	137
Shoemake & Rowland (1993)	1	−.130	*t*	−2.03*	full	1	2	W	1	2	4	2	1	2	coll	b/u	1	m/o	60
Sigelman & Welch (1991)	1	−.053	Prop	58/63*	full	1	1	B	4	1	1	2	1	1	adult	b/u	1	m/o	1,250
Siller & Chipman (1964)	1	−.020	*r*	−.020*	some	1	1	W	1	1	2	3	1	4	coll	b/u	1	m/o	1,108
Simon (1995)	1	−.323	*r*	−.323*	full	1	1	W	1	1	1	2	1	3	coll	f	1	m/o	360
	2	−.453	*r*	−.453*	full	1	1	W	1	1	1	2	1	3	coll	m	1	m/o	204
Simoni (1996)	1	−.527	*r*	−.527*	full	1	1	W	1	1	1	3	1	3	coll	b/u	1	m/o	170
Simpson et al. (1976)	1	−.265	*t*	−1.70*	none	1	2	B	3	2	4	3	2	4	child	b/u	1	edu	38
Singer (1966)	1	−.235	*F*	7.03	some	2	1	B	3	3	99	3	1	1	child	b/u	1	edu	120
	2	−.271	*F*	7.60	some	2	1	B	3	3	99	3	1	1	child	b/u	1	edu	96
Slavin (1979)	1	−.223	*F*	15.38*	none	1	2	B	4	2	4	1	2	1	adol	b/u	1	edu	294
Slavin & Madden (1979)	1	−.062	*p*	.001*	some	1	1	W	1	1	2	2	1	1	adol	b/u	1	edu	1,387
	2	−.073	*p*	.001*	some	1	1	W	1	1	2	2	1	1	adol	b/u	1	edu	1,004
Slininger et al. (2000)	1	−.226	*M/SD*	$p = .008$*	none	1	2	B	3	2	4	2	1	6	child	m	1	edu	69
	2	−.114	*M/SD*	$p = .21$*	none	1	2	B	3	2	4	2	1	6	child	f	1	edu	62
Smith (1955)	1	−.475	*p*	.001	full	1	1	W	1	1	99	2	1	1	coll	b/u	2	trav	24
	2	+.407	*p*	.01	full	1	1	W	1	1	99	2	1	1	coll	b/u	2	trav	20
	3	+.238	*p*	.05	full	1	1	W	1	1	99	2	1	1	coll	b/u	2	trav	34
Smith (1969)	1	−.083	*M/SD*	$p = .282$	full	1	1	B	3	2	99	2	1	5	coll	b/u	1	res	167
Smith (1994)	1	−.408	Prop	27/68	full	1	1	B	2	1	1	2	1	1	adult	f	1	res	110
	2	−.152	Prop	44/59	full	1	1	B	2	1	1	2	1	1	adult	f	1	res	75
	3	−.394	Prop	31/70	full	1	1	B	2	1	1	2	1	1	adult	f	1	res	122
	4	+.068	Prop	53/47	full	1	1	B	2	1	1	2	1	1	adult	f	1	res	118
Smith-Castro (2000)	1	−.329	*r*	−.329	some	2	1	W	1	1	1	2	1	1	adol	b/u	6	m/o	742
	2	−.294	*r*	−.294	some	2	1	W	1	1	1	2	1	1	adol	b/u	6	m/o	383
Spangenberg & Nel (1985)	1	−.198	*t*	−2.66*	full	1	1	B	2	1	99	3	1	1	adult	b/u	6	org	195
Sparling & Rogers (1985)	1	−.397	O	$p = .17$*	full	1	2	W	3	2	4	3	2	2	adol	b/u	1	rec	6
Stager & Young (1981)	1	−.053	*M/SD*	$p = .14$*	some	1	1	W	1	2	99	2	1	6	adol	b/u	1	edu	382
Stainback et al. (1984)	1	−.326	*r*	−.326	full	1	1	W	1	1	1	3	1	6	adult	b/u	1	edu	91
Stangor et al. (1996)	1	−.244	*r*	−.244*	full	1	1	W	1	1	2	3	1	1	coll	b/u	1	m/o	83
Starr & Roberts (1982)	1	−.078	*p*	.034*	some	1	1	B	2	1	2	3	1	1	adult	b/u	1	m/o	734
Stephan & Rosenfield (1978a)	1	−.390	*r*	−.390	some	1	2	W	1	1	2	2	2	1	child	b/u	1	edu	65
Stephan & Rosenfield (1978b)	1	−.370	*r*	−.370	some	1	2	W	1	1	3	3	1	1	child	b/u	1	edu	151
	2	−.370	*r*	−.370	some	1	2	W	1	1	3	3	1	1	child	b/u	1	edu	96
	3	−.390	*r*	−.390	some	1	2	W	1	1	3	3	1	1	child	b/u	1	edu	64
Stephan & Stephan (1985)	1	−.159	*r*	−.159*	some	1	1	W	1	1	3	2	1	1	coll	b/u	1	m/o	83
Stephan & Stephan (1989)	1	−.177	*r*	−.177*	full	1	1	W	1	1	3	1	1	1	coll	b/u	1	m/o	123
	2	−.205	*r*	−.205*	full	1	1	W	1	1	3	1	1	1	coll	b/u	1	m/o	133
Stephan et al. (2000)	1	−.291	*r*	−.291*	some	1	1	W	1	1	3	3	1	1	coll	b/u	6	m/o	126
	2	−.243	*r*	−.243*	some	1	1	W	1	1	3	2	1	1	coll	b/u	6	m/o	130
Stewart (1988)	1	−.413	*M/SD*	$p = .024$	some	1	2	W	1	2	99	3	1	4	coll	b/u	1	edu	15
Stohl (1985)	1	−.335	*F*	6.15*	none	1	2	B	2	2	99	3	2	1	coll	b/u	1	m/o	49
Stouffer et al. (1949)	1	−.442	Prop	23/67	none	1	1	W	1	3	99	1	1	1	adult	m	1	org	60
	2	−.502	Prop	20/71*	none	1	1	B	2	3	99	1	1	1	adult	m	1	org	1,725
Strauch (1970)	1	−.180	*F*	4.15	some	1	2	B	2	3	99	2	1	4	adol	b/u	1	edu	124
Strauch et al. (1970)	1	−.217	*F*	1.98	full	1	2	W	1	2	4	2	2	6	coll	b/u	1	res	10
Strohmer et al. (1984)	1	−.372	*r*	−.372*	full	1	1	W	1	1	2	3	1	4	adult	b/u	1	m/o	210

(*Appendix continues*)

Appendix (*continued*)

Reference	Sample	*r*	Test	Statistic	Choice	Pub	Type	B/W	Control	IV	IVQ	DVQ	Prog	Target	Age	Sex	Geo	Set	*N*
Surace & Seeman (1967)	1	−.173	*r*	−.173*	full	1	1	W	1	1	2	2	1	1	adult	b/u	1	m/o	159
	2	+.014	*r*	+.014*	full	1	1	W	1	1	2	2	1	1	adult	b/u	1	m/o	124
Tait & Purdie (2000)	1	−.169	*M/SD*	*p* = .000*	some	1	1	B	4	1	1	3	1	4	adult	b/u	5	m/o	1,338
Taylor & Dear (1981)	1	−.140	*t*	−9.33*	some	1	1	W	1	1	2	3	1	5	adult	b/u	4	m/o	1,090
Thomas et al. (1985)	1	.000	*p*	*ns*	some	1	1	B	4	1	2	2	1	4	child	b/u	5	m/o	88
Thompson (1993)	1	−.292	*F*	6.25	none	1	3	W	1	2	4	2	1	99	coll	b/u	1	lab	17
	2	−.269	*F*	10.65	none	1	3	W	1	2	4	2	1	99	coll	b/u	1	lab	34
	3	−.086	*F*	0.71	none	1	3	W	1	2	4	2	1	99	coll	b/u	1	lab	24
Togonu & Odebiyi (1985)	1	.000	Prop	55/55*	some	1	1	B	2	1	1	2	1	4	adult	b/u	6	m/o	519
Towles-Schwen & Fazio (1999)	1	−.200	*r*	−.200	some	2	1	W	1	1	2	4	1	1	coll	b/u	1	edu	190
Trent et al. (1979)	1	−.219	*p*	.05	none	1	2	W	4	2	4	3	2	2	adol	b/u	1	m/o	80
	2	−.219	*p*	.05	none	1	2	W	4	2	4	3	2	2	adol	b/u	1	m/o	80
Triandis & Vassiliou (1967)	1	+.200	Prop	78/59*	full	1	1	B	4	1	99	2	1	1	adult	m	2	m/o	103
	2	−.155	Prop	62/77*	full	1	1	B	4	1	99	2	1	1	adult	m	2	m/o	80
Tropp (2000)	1	−.325	*r*	−.325*	some	2	1	W	1	1	4	3	1	1	coll	b/u	1	m/o	80
Tropp & Stout (1999)	1	−.233	*r*	−.233*	full	2	1	W	1	1	4	4	1	1	coll	b/u	1	m/o	74
	2	−.194	*r*	−.194*	full	2	1	W	1	1	4	4	1	1	coll	b/u	1	m/o	50
	3	−.079	*r*	−.079*	full	2	1	W	1	1	4	4	1	1	coll	b/u	1	m/o	39
Trubowitz (1969)	1	−.132	*M/SD*	*p* = .15*	some	1	2	B	2	2	99	2	1	1	child	b/u	1	edu	121
	2	+.101	*M/SD*	*p* = .263*	some	1	2	B	2	2	99	2	1	1	child	b/u	1	edu	122
Trute & Loewen (1978)	1	−.257	*M/SD*	*p* = .044*	some	1	1	B	4	1	3	3	1	5	adult	b/u	4	m/o	62
Trute et al. (1989)	1	−.121	*r*	−.121*	some	1	1	W	1	1	2	2	1	5	adult	b/u	4	m/o	537
Tsukashima & Montero (1976)	1	−.020	Prop	52/54*	some	1	1	B	3	1	1	2	1	1	adult	b/u	1	m/o	308
Tuckman & Lorge (1958)	1	−.074	Prop	34/41*	full	1	1	B	4	1	1	2	1	2	adult	b/u	1	m/o	792
Turman & Holtzman (1955)	1	−.061	Prop	54/60	some	1	1	W	1	1	2	2	1	1	adult	b/u	1	edu	144
	2	.000	*p*	*ns*	some	1	1	W	1	1	2	2	1	1	adult	b/u	1	edu	150
Van Den Berghe (1962)	1	−.329	Prop	29/62	full	1	1	B	4	1	99	2	1	1	coll	b/u	6	m/o	345
Van Dick & Wagner (1995)	1	−.179	*r*	−.179*	some	2	1	W	1	1	1	3	1	1	adult	m	2	m/o	134
Van Dick et al. (2000)	1	−.210	*r*	−.210*	some	2	1	W	1	1	1	2	1	1	adult	b/u	2	m/o	3,000
Van Dyk (1990)	1	−.091	*t*	−1.75*	some	1	1	W	1	1	2	3	1	1	adult	f	6	res	91
Van Ossenbruggen (1999)	1	−.290	*r*	−.290*	some	2	1	W	1	1	2	2	1	1	adult	b/u	2	m/o	3,000
Van Weerden-Dijkstra (1972)	1	.000	*p*	*ns*	some	1	1	W	1	1	2	2	1	6	adult	b/u	2	m/o	418
Verkuyten & Masson (1995)	1	−.333	*r*	−.333*	some	1	1	W	1	1	3	3	1	1	adol	b/u	2	m/o	372
	2	−.180	*r*	−.180*	some	1	1	W	1	1	2	3	1	1	adol	b/u	2	m/o	158
Voeltz (1980)	1	−.259	*M/SD*	*p* = .000*	full	1	1	B	3	2	99	3	1	4	child	b/u	1	edu	1,310
Vornberg & Grant (1976)	1	−.144	*F*	3.71*	some	1	2	W	1	2	4	1	1	1	adol	b/u	6	edu	44
Wagner et al. (1989)	1	−.229	*r*	−.229*	some	1	1	W	1	1	2	1	1	1	adol	b/u	2	m/o	60
	2	−.048	*r*	−.048*	some	1	1	W	1	1	2	1	1	1	adol	b/u	2	m/o	50
Walsh (1971)	1	−.205	*t*	−5.09*	some	1	2	W	1	2	4	3	2	5	coll	f	1	org	147
Ward & Rana-Deuba (2000)	1	−.350	*r*	−.350	some	1	1	W	1	1	2	2	1	1	adult	b/u	1	m/o	104
Ward et al. (1998)	1	−.271	*t*	−2.03*	some	1	2	W	1	2	4	2	1	2	coll	b/u	6	rec	13
Webster (1961)	1	+.163	χ^2	3.05*	some	1	1	B	2	3	99	3	1	1	adol	b/u	1	edu	115
	2	−.049	χ^2	0.23*	some	1	1	B	2	3	99	3	1	1	adol	b/u	1	edu	97
Weerbach & DePoy (1993)	1	−.305	*r*	−.305*	full	1	1	B	2	1	99	2	1	5	coll	b/u	1	org	90
Weigert (1976)	1	−.180	*r*	−.180	some	1	1	W	1	1	2	2	1	1	adult	m	1	org	454
Weinberg (1978)	1	−.167	*t*	−2.06*	some	1	2	B	4	2	2	2	1	4	coll	b/u	1	edu	202
Weis & Dain (1979)	1	−.233	*p*	.001	full	1	1	W	1	1	1	3	1	3	adult	b/u	1	m/o	100
Weisel (1988)	1	−.147	*M/SD*	*p* = .143*	some	1	1	B	2	3	99	3	1	4	adol	b/u	3	edu	99
Weiss (1987)	1	−.190	*r*	−.190	full	1	1	W	1	1	1	2	1	1	adult	b/u	2	m/o	648
Weller & Grunes (1988)	1	−.104	*p*	.31	full	1	1	B	3	2	99	3	1	5	coll	m	3	org	95
Whaley (1997)	1	−.245	*r*	−.245*	some	1	1	W	1	1	2	3	1	5	adult	b/u	2	m/o	1492
Whitley (1990)	1	−.315	*r*	−.315*	full	1	1	W	1	1	1	3	1	3	coll	b/u	1	m/o	366
Wilder (1984)	1	−.099	*t*	−0.53*	none	1	3	B	3	2	4	1	1	99	coll	f	1	lab	62

Appendix (*continued*)

Reference	Sample	r	Test	Statistic	Choice	Pub	Type	B/W	Control	IV	IVQ	DVQ	Prog	Target	Age	Sex	Geo	Set	N
Williams (1964)	1	−.134	Prop	45/58*	full	1	1	B	3	1	2	2	1	1	adult	b/u	1	m/o	422
	2	−.209	Prop	26/47*	full	1	1	B	3	1	2	2	1	1	adult	b/u	1	m/o	287
	3	−.415	Prop	33/74*	full	1	1	B	3	1	2	2	1	1	adult	b/u	1	m/o	302
	4	−.260	Prop	55/82*	full	1	1	B	3	1	2	2	1	1	adult	b/u	1	m/o	195
	5	−.202	Prop	33/53*	full	1	1	B	3	1	2	2	1	1	adult	b/u	1	m/o	445
	6	−.259	Prop	37/63*	full	1	1	B	3	1	2	2	1	1	adult	b/u	1	m/o	150
	7	−.162	Prop	34/50*	full	1	1	B	3	1	2	2	1	1	adult	b/u	1	m/o	147
	8	−.173	Prop	38/55*	full	1	1	B	3	1	2	2	1	1	adult	b/u	1	m/o	150
	9	−.153	Prop	40/56*	full	1	1	B	3	1	2	2	1	1	adult	b/u	1	m/o	283
	10	−.183	Prop	35/56*	full	1	1	B	3	1	2	2	1	1	adult	b/u	1	m/o	207
	11	−.079	Prop	44/52*	full	1	1	B	3	1	2	2	1	1	adult	b/u	1	m/o	219
Williams (1972)	1	−.156	*M/SD*	$p = .031$	full	2	3	B	4	2	4	2	2	1	adol	b/u	1	edu	192
Wilner et al. (1955)	1	−.309	Prop	33/63*	some	1	1	B	3	1	2	2	1	1	adult	b/u	1	res	120
	2	−.292	Prop	23/50*	some	1	1	B	3	1	2	2	1	1	adult	f	1	res	66
	3	−.212	Prop	58/69*	some	1	1	B	3	1	2	2	1	1	adult	f	1	res	154
	4	−.209	Prop	35/54*	some	1	1	B	3	1	2	2	1	1	adult	f	1	res	98
	5	−.156	Prop	26/41*	some	1	1	B	3	1	2	2	1	1	adult	f	1	res	68
	6	−.099	Prop	18/26*	some	1	1	B	3	1	2	2	1	1	adult	f	1	res	67
	7	−.258	Prop	35/61*	some	1	1	B	3	1	2	2	1	1	adult	f	1	res	125
	8	−.194	Prop	27/46*	some	1	1	B	3	1	2	2	1	1	adult	f	1	res	108
Wilson (1984)	1	−.380	*r*	−.380	some	1	1	W	1	1	3	3	1	1	adult	b/u	1	m/o	628
	2	−.350	*r*	−.350	some	1	1	W	1	1	3	3	1	1	adult	b/u	1	m/o	362
Wilson (1996)	1	−.173	*r*	−.173*	full	1	1	W	1	1	2	2	1	1	adult	b/u	1	res	580
Wilson & Lavelle (1990)	1	−.205	*p*	.005	some	1	1	W	1	2	99	2	1	1	child	b/u	6	edu	94
	2	−.174	*p*	.01	some	1	1	W	1	2	99	2	1	1	child	b/u	6	edu	110
With & Rabbie (1985)	1	−.272	*F*	14.63	full	1	1	B	3	1	2	3	1	1	adult	b/u	2	res	196
Wolsko et al. (2000)	1	−.157	*r*	−.157*	some	2	1	W	1	1	3	2	1	1	coll	b/u	1	m/o	148
	2	−.121	*r*	−.121*	some	2	1	W	1	1	3	2	1	1	coll	b/u	1	m/o	122
Wood (1990)	1	−.176	*r*	−.176*	some	2	1	W	1	1	2	2	1	1	coll	f	1	edu	105
	2	−.048	*r*	−.048*	some	2	1	W	1	1	2	2	1	1	coll	m	1	edu	80
Wood & Sonleitner (1996)	1	−.210	*r*	−.210*	some	1	1	W	1	1	3	3	1	1	adult	b/u	1	m/o	292
Works (1961)	1	−.218	Prop	31/53	some	1	1	B	3	3	99	1	1	1	adult	m	1	res	68
	2	−.287	Prop	21/49	some	1	1	B	3	3	99	1	1	1	adult	f	1	res	76
Wright & Klein (1966)	1	−.345	χ^2	19.8*	full	1	1	B	3	2	99	2	1	5	adult	b/u	1	res	179
Wright & Tropp (2000)	1	−.162	*p*	.002*	some	2	2	B	4	2	99	2	1	1	child	b/u	1	edu	356
Yeakley (1998)	1	−.301	Prop	29/59	full	2	2	B	3	2	4	4	2	99	coll	b/u	1	edu	26
Yinon (1975)	1	−.447	*F*	23.6*	full	1	1	B	3	2	99	2	1	1	adult	b/u	3	res	80
Young (1998)	1	−.196	*t*	−1.30*	none	2	3	B	4	2	4	2	2	1	coll	m	1	edu	43
	2	−.360	*t*	−1.97*	none	2	3	B	4	2	4	2	2	1	coll	m	1	edu	29
Yum & Wang (1983)	1	−.055	*p*	.10*	full	1	1	B	3	1	2	3	1	1	adult	b/u	1	m/o	876
Zakay (1985)	1	−.214	Prop	54/74	some	1	2	B	2	2	99	3	1	4	adol	b/u	3	edu	191
Zani & Kirchler (1995)	1	−.244	*r*	−.244*	full	1	1	W	1	1	3	3	1	1	adult	b/u	2	res	222
Zentralarchiv… (1996)	1	−.270	*r*	−.270*	some	2	1	W	1	1	1	3	1	1	adult	b/u	2	m/o	2,945
Zeul & Humphrey (1971)	1	−.486	Prop	23/71	some	1	1	B	3	1	2	2	1	1	adult	f	1	res	50

Note. r = effect size (correlation) for each sample; Test = test from which the effect size was derived (r = correlation, F = F test from analysis of variance, t = t test, MW = Mann–Whitney U test, χ^2 = chi-square, Prop = proportions or frequencies, p = p value, *M/SD* = means and standard deviations, O = other, Mult = multiple tests); Statistic = the original statistic before conversion to r (P = probability level; the proportional ratios represent the percentages of the contact and no contact groups in the high prejudice category; asterisks indicate that the statistic shown is a composite formed from multiple tests); Choice = the degree of choice available to the subject about whether to participate in the contact (none = no choice, some = some choice, full = full choice); Pub = the publication status of the study (1 = published, 2 = unpublished); Type = the type of study conducted (1 = surveys and field studies, 2 = quasi-experiments, 3 = experiments); B/W = the study design (B = between subjects, W = within subjects); Control = the control group (1 = within design, 2 = no contact control, 3 = some contact control, 4 = extensive contact control); IV = the type of contact indicator (1 = reported contact, 2 = observed contact, 3 = assumed contact); IVQ = the independent variable quality (1 = single item, 2 = multiple items, low reliability [$\alpha < .70$ or unreported], 3 = multiple items, high reliability [$\alpha > .69$], 4 = experimental manipulation, 99 = other); DVQ = the dependent variable quality (1 = single item, 2 = multiple items, low reliability [$\alpha < .70$ or unreported], 3 = multiple items, high reliability [$\alpha > .69$], 4 = other reliable indicator); Prog = the global rating of structured programs designed to approximate Allport's optimal conditions (1 = no program, 2 = structured program); Target = the target group for contact in the study (1 = race, ethnicity, or religion, 2 = elderly, 3 = sexual orientation, 4 = physical disability, 5 = mental illness, 6 = mental disability, 99 = other); Age = the age of the participants (child = children, adol = adolescents, coll = college students, adult = adults); Sex = the sex of the participants (m = all males, f = all females, b/u = both or unspecified); Geo = the geographical area of the participants (1 = United States, 2 = Europe, 3 = Israel, 4 = Canada, 5 = Australia and New Zealand, 6 = Africa, Asia, and Latin America); Set = the setting of the study (lab = laboratory, edu = educational, org = organizational, res = residential, rec = recreational, trav = travel/tourism, m/o = mixed and other); N = the sample size.

Received November 13, 2003
Revision received June 28, 2005
Accepted August 11, 2005 ■