Journal of Applied Psychology
1992, Vol. 77, No. 4, 554-561

0021-9010/92/$3.00

A Meta-Analytic Comparison of the Effectiveness of Smoking Cessation Methods

Chockalingam Viswesvaran and Frank L. Schmidt
Department of Management and Organizations
University of Iowa

Meta-analysis was used to cumulate the results from 633 studies of smoking cessation, involving 71,806 subjects, that reported the proportion of successful quits. Self-care methods do not appear to be as effective as formal intervention methods. Instructional programs involving physicians were not more effective than other instructional programs. Conditioning-based techniques such as aversive methods had success rates similar to those of instructional methods, and among the instructional methods, those incorporating social norms and values were more successful than those relying solely on didactic approaches. Cumulation of quit rates from all available control groups indicated that, on average, 6.4% of the smokers could be expected to quit smoking without any intervention. This figure must be subtracted from the raw success rate to obtain the net success rate for each program. Directions for future research are discussed.

Smoking is a significant source of death and disease (U.S. Department of Health and Human Services, 1989), with more than one out of every six American deaths resulting from cigarette smoking. With increasing evidence confirming the link between smoking and health hazards (Fielding, 1985), programs and policies to combat smoking have proliferated. Many reviews have examined the effectiveness of the methods used in these programs (e.g., Bernstein, 1969; Bruvold & Rundall, 1988; Callner, 1975; Cohen et al., 1989; Keutzer, Lichtenstein, & Mees, 1968; Kinder, Pape, & Walfish, 1980; Kottke, Battista, DeFriese, & Brekke, 1988; Leventhal & Cleary, 1980; Moskowitz, 1983; Schaps, DeBartolo, Moskowitz, Palley, & Churgin, 1981; Schwartz, 1969, 1987; Schwartz & Rider, 1978; Thompson, 1978). Except for Bruvold and Rundall (1988) and Kottke et al. (1988), these reviews are of the traditional narrative type that has been criticized for its subjectivity as well as for its failure to quantitatively cumulate the effects reviewed (Bangert-Drowns, 1986; Cooper, 1984; Cooper & Rosenthal, 1980; Glass, 1976, 1977; Glass, McGaw, & Smith, 1981; Hedges & Olkin, 1985; Hunter & Schmidt, 1990).

In addition, these reviews have analyzed only a very small subset of all available studies. Even the meta-analytic reviews (Bruvold & Rundall, 1988; Kottke et al., 1988) have confined themselves to intervention programs that used control groups. Bruvold and Rundall's (1988) meta-analytic review focused exclusively on school-based intervention programs designed for children and adolescents, and thus restricted itself to 30 empirical studies on this population that had both experimental and control groups. Though these 30 studies are comprehensive of all school-based intervention programs since 1970, this meta-analytic review did not address the effectiveness of the various intervention methods in general. The scope of this limitation is apparent when one considers the fact that there are over 633 studies addressing the effectiveness of the various intervention methods (Schwartz, 1987).

The reason Bruvold and Rundall (1988) restricted their review to designs using control groups was methodological: To meta-analyze d values, one must be able to obtain a d value from each study. This computation is not possible if studies report only the proportion of subjects who quit smoking after going through a program. The same limitation constrained Kottke et al. (1988) to 39 studies that used both experimental and control groups. In the present study we circumvented this methodological limitation and were therefore able to cumulate results across a much larger set of studies.

Furthermore, like Glassian meta-analytic techniques (Glass et al., 1981), the study-effects (Bangert-Drowns, 1986) meta-analytic methods used by Bruvold and Rundall (1988) and Kottke et al. (1988) ignore the impact of statistical artifacts. Before hypothesizing that the effectiveness of a method depends on moderator variables such as individual personality characteristics and situational factors, one must examine whether the observed variance in the effectiveness of a method is due to statistical artifacts. In this article we use Hunter and Schmidt's (1990) meta-analytic methods to cumulate the results of each intervention method across studies, and we correct the observed variance for variance due to statistical artifacts before testing for the presence of any moderating influence.

The practical importance of analyzing the effectiveness of the different intervention methods lies in its policy implications. With increasing societal awareness of the hazards of smoking and government-funded programs to realize a smoke-free society, knowledge of the relative and absolute effectiveness of different intervention methods becomes important. Knowledge of program effectiveness could provide a foundation for policy decisions concerning the funding of programs (e.g., Green, Rimer, & Bertera, 1978; Kristein, 1982, 1983). There are

Correspondence concerning this article should be addressed to Frank L. Schmidt, Department of Management and Organizations, College of Business Administration, University of Iowa, Iowa City, Iowa 52242.

also potential theoretical implications inasmuch as the relative success rates of different intervention methods may reveal theoretically significant facts about human behavior.

Surveys indicate that many employers sponsor programs to help employees quit smoking (Bennett & Levy, 1980; Chovil, Alexander, Gibson, & Altekruse, 1983; Davis, Rosenberg, Iverson, Vernon, & Bauer, 1984; Fielding & Breslow, 1983). The legal conflict between the rights of smokers and nonsmokers and the evolution of legislation to control workplace smoking (Roemer, 1983; Swingle, 1980) indicate that prohibitory policies alone may not withstand legal challenges in court. Thus, employers are increasingly turning to cessation programs (Pearson, 1980) to reduce smoking in the workplace. In this context, a meta-analytic review of the effectiveness of the different intervention programs would be of value to personnel managers, occupational physicians, and chief executive officers.

Schachter (1982) concluded that formal intervention methods are not as effective as self-quit attempts. This controversial conclusion could be due to a real difference in the effectiveness of the intervention methods, or may be the result of differences in the characteristics of smokers utilizing the different methods, or may be the result of statistical artifacts. Schoenbach, Orleans, and Wagner (1988) provided evidence which indicated that the demographic and smoking characteristics of participants in self-quit attempts are similar to those of participants in formal intervention methods. Schachter's conclusions have been criticized for limited generalizability resulting from the use of a small and unrepresentative sample of studies (Jeffery & Wing, 1983; Prochaska, 1983). Finally, on the basis of 10 longitudinal studies of persons who attempted to quit smoking by themselves, Cohen et al. (1989) concluded that the success rates reported for self-quit methods are no higher than those reported for formal intervention methods. The present study sheds some light on this question.

To assess quantitatively the effectiveness of the different intervention methods, a meta-analytic review that cumulates across studies the success rates of each intervention method and accounts for statistical artifacts is appropriate. Using the meta-analytic methods of Hunter and Schmidt (1990), we conducted a meta-analytic review of the different intervention methods to effect smoking cessation with data from a comprehensive qualitative review by Schwartz (1987). The data set comprised 633 empirical studies that report the percentage of smokers who participated in an intervention program and were successful in quitting the habit. Though the Hunter–Schmidt meta-analytic procedures have been applied mostly to correlations or effect sizes, those procedures can be extended to cumulate other statistics. For example, Coggin and Hunter (1987) applied these methods to regression slopes.

This meta-analytic review of the effectiveness of intervention methods to achieve smoking cessation is unique in three ways. First, no review has summarized quantitatively such a comprehensive data base of smoking intervention methods. Second, no existing review has examined whether the observed variance in the effectiveness of intervention methods is due to statistical artifacts. Only if the observed variance is greater than the variance due to statistical artifacts should one hypothesize that moderator variables such as personal and situational factors affect the methods' effectiveness. Finally, this review is the first to cumulate proportions quantitatively and to correct the observed variability in proportions for variance due to the artifact of sampling error.

Method

Hunter and Schmidt's (1990) methods cumulate the effects found by different studies to arrive at the best estimate of the mean effect of the intervention method. Furthermore, by correcting the observed effect sizes for statistical artifacts, either individually by each study or by correcting the mean effect size for the mean effect of artifacts, these methods yield the best possible estimate of the mean true effect size of the intervention methods. Knowledge of the mean true effect size is necessary to compare meaningfully the effectiveness of any two intervention methods. In this meta-analytic review, because of the limited information available in the individual studies, only the artifact of sampling error could be addressed. The observed mean proportion across studies is usually based on a large sample size, which mitigates the impact of sampling error and thus provides a better estimate of the intervention effect than the individual studies. Hence, we compare the different intervention methods on the basis of their sample-size-weighted means (mean proportions). Furthermore, in many areas of scientific enquiry (e.g., validity generalization), sampling error has been found to account for most of the observed variance. Hence, it is important to examine whether sampling error alone is sufficient to explain the observed variation across situations in the effectiveness of the various intervention methods.

Database

This meta-analysis cumulates the results of 633 studies of smoking cessation methods. The criteria for inclusion (Schwartz, 1987, p. 11) were as follows: (a) The studies examined smoking cessation or abstinence, not merely reduction in smoking levels; (b) only cigarette smoking (and not smoking of pipes or cigars or the chewing of smokeless tobacco) was examined; (c) the studies were based on at least 6 persons —case studies were excluded; and (d) the studies reported follow-up results for at least 3 months. Though the focus was on the United States and Canada, some studies from England, Sweden, France, and Scandinavian countries were also included. The studies all used adults (college students, employees, patients, etc.) as subjects.

The studies were classified into 15 categories on the basis of the predominant method (i.e., the major component of the intervention program) used to effect smoking cessation. Four dimensions were used to identify the predominant method used: (a) the general approach (e.g., educational, behavior modification, medical), (b) the organization (e.g., proprietary, medical group, university, religious group, voluntary), (c) the investigator or leader (e.g., psychologist, physician, educator, lay person), and (d) the specific approach (e.g., groups, hypnosis, aversive conditioning). The 15 categories were as follows: self-care, educational, 5-day plans, group withdrawal clinics, medication, nicotine chewing gums, hypnosis, acupuncture, smoke aversion techniques, other aversive techniques, miscellaneous, physician advice programs, physician intervention programs more than counseling, programs for patients with pulmonary disease, and programs for patients with cardiac disease. Success was assessed after 3 months in 93 intervention studies, after 6 months in 214 studies, and after 12 months in 226 studies. Most of the studies (614 out of 633) assessed the effectiveness of intervention within 12 months. The citations for these 633 studies (along with their sample sizes and success rates) are available on request.

Intercoder agreement in summarizing or extracting information from the primary studies is a concern in meta-analysis. This was not an issue in this meta-analysis because the data were available in a coded and tabulated form. The question of interest here is the intercoder agreement in the original reviews (Schwartz, 1969, 1987; Schwartz & Rider, 1978). Schwartz (1969, pp. 484–485) and Schwartz and Rider (1978, p. 3) stated that ambiguities and confusions were resolved by contacting the authors of the primary studies and requesting additional information and clarification.

Lack of attention to the methodological issues surrounding the reliability of measures is rampant in the medical literature (Bailar & Mosteller, 1986; Baum et al., 1981; Hedges, 1987; Hunter & Schmidt, 1990, p. 38). Of the 11 published reviews of smoking programs (4 of them in the *Psychological Bulletin*), only 2 (Bruvold & Rundall, 1988; Kottke et al., 1988) addressed the issue of intercoder agreement. Bruvold and Rundall (1988, p. 57) merely stated that discrepancies in coding were resolved in a meeting of the raters (with no information on the percentage of decisions thus resolved), whereas Kottke et al. (1988, p. 2884) stated that when there was disagreement in the coding, the principal author resolved it by reading the article himself (again no information was given as to the percentage of decisions thus resolved). It is therefore not surprising that Schwartz also merely indicated that disagreements were resolved by contacting the authors of the primary studies. He did not provide information on the percentage of decisions in coding and tabulating the results that were resolved through consultations with the authors of the primary studies.

Haring et al. (1981) presented empirical data indicating that intercoder agreement in meta-analyses is a function of the judgmental nature of the items coded. Their review of meta-analyses indicated that eight of the nine items lowest in coder agreement were judgments (e.g., the quality of the study) as opposed to calculation-based variables (e.g., effect sizes, number of subjects). Jackson (1980) and Hattie and Hansford (1982, 1984) also provided data which indicate that problems of intercoder agreement in meta-analyses are negligible for coding computation-based numerical variables.

In this review we used four items of information extracted from the primary studies: (a) the number of subjects in the treatment, (b) the number of subjects who were successful in quitting the habit, (c) the time period of follow-up, and (d) the predominant treatment method used. The first three of these four pieces of information are numerical values recorded from the primary studies. In addition, because most studies were conducted to assess the effectiveness of a particular intervention method, we expected there to be no ambiguity in classifying a study into 1 of the 15 categories (in fact, only 10 studies were categorized as miscellaneous). This was confirmed by J. L. Schwartz (personal communication, December 18, 1991), who indicated that he had no problem in classifying the studies into categories and did not have to seek clarification from the authors of the primary studies to code the four items of information used in this meta-analysis. Finally we coded the four items used in this meta-analysis for 20 of the 633 studies. We had perfect agreement with the information reported by Schwartz (1987) in his review.

Control groups were used only sporadically; only 41 of the 633 studies used a control group. Of the 41 control groups, 10 were used in studies evaluating the effectiveness of aversive techniques other than smoke aversion, and 5 were used in evaluations of physician advice interventions. Evaluations of smoke aversive techniques used 6 control groups, and 4 control groups were used in evaluating the effectiveness of hypnosis. Three control groups were used in each of the following three categories: nicotine chewing gum, acupuncture, and physician interventions more than counselling. There were 2 control groups from studies of self-care evaluations and 2 from studies evaluating the 5-day plans. Finally, of the studies categorized as educational, miscellaneous, and cardiac patients, 1 study in each category used a control group. Thus, the control groups were used to a similar extent across all intervention methods, though similarity in usage is not a necessary requirement for comparison, as we explain further on.

In a natural setting involving health care, it is sometimes difficult to obtain a control group (i.e., a group in which no intervention is attempted). This limitation in individual studies can be circumvented in a meta-analysis because the control group for one intervention study can serve as the control group for other intervention studies that do not have control groups (Hunter, 1990). Control groups (which by definition receive no intervention) are used to estimate the smoking quit rate without any intervention. Thus, the "treatment" in the control group is the same across all studies (i.e., no treatment), and findings for control groups in studies that have control groups can be generalized to studies that do not have control groups. The control group figure used should represent the average across all studies with control groups. A meta-analysis of the results from all available control groups provides a better benchmark to evaluate any intervention than does the control group from that particular study because the idiosyncrasies of particular control samples are averaged out. Hence, in our meta-analytic cumulation, we meta-analyzed the 41 control group samples together to obtain a more robust and reliable estimate of the average quit rate of smoking without any intervention. This resulting mean quit rate was then compared to the mean quit rates for different interventions, thus allowing estimates of "net" quit rates.

Of the categories analyzed, *self-care methods* involved persons quitting by themselves with the aid of self-help books and via mailing campaigns (in which information encouraging the smokers to quit, along with steps that facilitate quitting, are provided). Though different manuals had been used in self-quit programs, Cummings, Emont, Jaen, and Sciandra (1988) presented data which indicated that differences in manuals have no influence on study outcomes. Though self-care is implicated in all methods (i.e., even success in formal treatment methods depends on self-care), we use the term *self-care methods* as a descriptor for programs in which the essential feature is the absence of contractualized treatment. Self-care methods differ from control groups in that participants in self-care methods are known to have received literature regarding smoking cessation and to have been provided with self-quit kits, whereas subjects in control groups were not provided with similar material by the experimenters (though some subjects in control groups may possibly have been exposed to self-quit material from other sources). Studies using Lichtenstein and Danaher's (1977) self-help book, the *American Cancer Society Quit Kit*, graduated filters, the Stanford 5-city quit kit, American Lung Association (ALA) leaflets, and various freedom-from-smoking manuals are included in this category.

Educational methods, on the other hand, involved a classroom setting and a concentrated imparting of knowledge by an informed authority in the smoking health area. Lectures, films, and group discussions are used. Lectures provide information on relaxation techniques, insight into smoking habits, and health risk appraisals. An example is the Freshstart programs conducted by the American Cancer Society, which involves 4 hr of instruction.

Five-day plans usually involve the screening of a film showing surgery on a cancerous lung, along with displays of actual lung specimens. Evaluation studies of programs offered by the Seventh Day Adventist church, the Atlanta Lung Association, and the St. Helena Health Center are included in this category. In 1976, the Seventh Day Adventist church changed the name of their program from the Five-Day plan to the Breathe Free plan and copyrighted the program (other commercial organizations cannot duplicate it). But because the basic underlying format remained unchanged, we included both program evaluations in the same category.

Group withdrawal clinics emphasized self-appraisal and practicing abstinence under controlled conditions, followed by maintenance.

These groups used peer and group norms to effect smoking cessation. Evaluations of quit-smoking clinics conducted by the American Cancer Society (eight 2-hr sessions over a 4-week period) and various clinics offered by the local chapters of the American Lung Association (e.g., the 8-week group program called UNsmoke) are included in this category.

Medication involved the prescription of pharmaceutical agents to help the subjects quit smoking; these were classified as substitutes (which serve as replacements to nicotine; e.g., lobeline), deterrents (which diminish smoking desires by irritating the oral and nasal mucosa; e.g., benzocaine lozenges), and vegetable-base products (though the mode of action of these products is not clear, Schwartz & Rider, 1978, p. 11; e.g., *Anena sativa*, an oat product).

The use of *nicotine chewing gum* is based on the assumption that, because nicotine is the addicting agent, a substitute for nicotine will provide the solution to the smoking habit. Studies evaluating the effectiveness of nicorette and other carbonate buffers are included in this category.

The evaluation studies included in the *hypnosis* category vary in the time and effort spent on the individual participants. The therapeutic techniques used range from simple suggestions to a complex sequence of imagery, relaxation, and counseling. Many studies fail to report in sufficient detail the techniques used. The presence of a hypnotherapist was used as the basis for including a study in this category.

Under *acupuncture* are included the studies evaluating the effectiveness of acupuncture techniques. Acupuncture is based on the Eastern theory of connections in the body. Under the category of *smoke aversion* are included the studies that involved blowing warm, stale smoke on subjects' faces while they smoked. The studies included in the category of *other aversive techniques* used electric shocks, unpleasant tastes, smells, or noises, or covert sensitization (requiring the subjects to imagine unpleasant outcomes while they smoked). These studies are based on the behaviorist assumptions that the reinforcing aspects of any stimulus can be reduced and that the stimulus may be made aversive if it is paired with a noxious response.

The studies evaluating the effectiveness of interventions by physicians are grouped into two categories: those in which the *physicians provide only advice* and those in which there was *more active participation from the physicians*. The studies based on patient samples were classified into two categories: *pulmonary patients* versus *cardiac patients*. The *miscellaneous* category included 10 studies with a total sample of 1,400 subjects. These studies used a combination of emotional role playing, transcendental meditation, cruises, hypnosis, breathing therapy, and exercise. This category included methods that used fear films such as *Dying for a Fag?* along with biofeedback and peer pressures.

The effectiveness of the intervention methods was examined in a workplace setting in 51 of the 633 studies. Of these studies, self-care methods were used in 1 study, instructional methods in 13 studies, physician counseling in 6 studies, and nicotine chewing gum and hypnosis in 2 and 3 studies, respectively. The remaining 26 studies examined the effectiveness of aversive techniques for effecting smoking cessation in the workplace.

Procedure

Each of the 15 categories just mentioned (involving one predominant method) was meta-analyzed separately. We had available for each study the sample sizes (n_i) used and the proportion who quit smoking (P_i). For each category, the mean quit rate ($\bar{P}$) was computed as $\Sigma n_i P_i / \Sigma n_i$. The observed variance is given as $\Sigma n_i(P_i - \bar{P})^2/\Sigma n_i$. Thus, we obtained the sample-size-weighted mean and the sample-size-weighted variance of observed quit rates for each intervention method.

The formula for the sampling-error variance of proportions was obtained from standard statistical textbooks as $P_i Q_i / n_i$, where $Q_i = 1 - P_i$. A sample-weighted mean of the sampling-error variance of proportions to be cumulated was then obtained as $(\Sigma n_i P_i Q_i / n_i)/\Sigma n_i = \Sigma P_i Q_i / \Sigma n_i$. This sample-weighted mean sampling-error variance was then subtracted from the observed variance, yielding an estimate of the true variance plus variance due to other artifacts, such as unreliability (Hunter & Schmidt, 1990), that could not be corrected for. Lacking any information about these other artifacts, we can assume our true variance estimate to be an upper bound value on the real variability of the effectiveness of an intervention method.

After analyzing the 15 categories separately, we ran 5 more meta-analyses. First, we combined all educational programs, 5-day plans, and group withdrawal clinics and meta-analyzed them together (K = number of studies = 98), obtaining an estimate of the effectiveness of the *instructional* approaches. The second analysis combined medication and nicotine chewing gum studies (K = 69) and provided an estimate of the effectiveness of *drug-based* programs. Then studies using smoke aversion methods and studies examining other aversive techniques (K = 281) were meta-analyzed together to estimate the effectiveness of *aversive* techniques in effecting smoking cessation. Finally, two more meta-analyses, one with programs involving physicians (K = 33) and another with studies using only patients (K = 51), were conducted. The meta-analysis of *medically sponsored* programs estimated the effectiveness of programs employing physicians in their capacity as health experts. Comparing this estimate with the effectiveness of instructional approaches allowed evaluation of the claim that the source of information affects effectiveness. The meta-analysis of *patient-based* studies enabled us to examine whether the medical condition of the subject acts as a catalyst in effecting smoking cessation.

Given the small number of studies (51 out of 633) conducted in a workplace setting, we were unable to examine the effectiveness of each of the various (15) intervention methods in the workplace. We were able to compare only the effectiveness of *instructional methods in the workplace* (13 studies) and *aversive techniques in the workplace* (26 studies). Finally, we conducted a meta-analysis of the 41 *control* groups.

Results

The results of these 23 meta-analyses are presented in Table 1. Column 1 indicates the predominant intervention method used by the studies meta-analyzed, and the number of studies included in that meta-analysis is shown in column 2. The total sample sizes involved, the mean quit rate, the observed variance, the sampling-error variance, and the percentage of variance accounted for by sampling error are shown in the subsequent columns. The last two columns contain the corrected standard deviation and the 80% credibility intervals. The corrected standard deviation can be taken as an upper bound value, and the 80% credibility values are computed under the assumption that the proportions are approximately normally distributed. The 80% credibility values provide the range of values after the most extreme 10% of the outcomes on each end are eliminated.

Average Success Rates

The mean quit rates in column 4 of Table 1 provide a basis for comparison across the intervention methods. We find values ranging from 7% for physician advice to 42% for programs involving patients with cardiac disease. The medical condition of

Table 1
Meta-Analysis Results for Smoking Intervention Studies

Type of program	No. of studies	Total sample	Mean *P* (quit rate)	Observed variance	Sampling-error variance	% variance accounted for	Corrected *SD*[a]	80% credibility interval[b]
1. Self-care	24	3,585	.15	.0091	.00097	10.7	.09	.03–.27
2. Educational	27	3,352	.24	.0107	.00149	13.9	.10	.11–.37
3. Five-day plans	25	7,828	.26	.0082	.00056	6.8	.09	.14–.38
4. Group withdrawal clinics	46	11,580	.30	.0150	.00073	4.9	.12	.15–.45
5. Medication	29	6,810	.18	.0135	.00059	4.4	.11	.04–.32
6. Nicotine chewing gum	40	4,866	.16	.0094	.00134	14.3	.09	.04–.28
7. Hypnosis	48	6,020	.36	.0378	.00143	3.8	.19	.12–.60
8. Acupuncture	19	2,992	.30	.0362	.00094	2.6	.19	.06–.54
9. Smoke aversion	103	2,557	.31	.0240	.00733	30.5	.13	.14–.48
10. Other aversive techniques	178	3,926	.27	.0262	.00732	27.9	.14	.09–.45
11. Miscellaneous	10	1,400	.35	.0103	.00108	10.5	.10	.22–.48
12. Physician advice	17	7,190	.07	.0015	.00014	9.3	.04	.02–.12
13. Physician interventions more than mere advice	16	3,486	.18	.0021	.00085	40.5	.04	.13–.23
14. Pulmonary patients	17	1,661	.34	.0416	.00189	4.5	.20	.08–.60
15. Cardiac patients	34	4,553	.42	.0271	.00163	6.0	.16	.22–.62
16. Instructional (2 + 3 + 4)[c]	98	22,760	.28	.0126	.00078	6.2	.11	.14–.42
17. Drug based (5 + 6)[c]	69	11,676	.17	.0119	.00088	7.4	.10	.04–.30
18. Aversive (9 + 10)[c]	281	6,483	.29	.0257	.00736	28.6	.14	.11–.47
19. Medically sponsored (12 + 13)[c]	33	10,676	.11	.0041	.00036	8.8	.06	.03–.19
20. Patient based (14 + 15)[c]	51	6,214	.40	.0322	.00169	5.2	.17	.18–.62
21. Instructional methods in work sites	13	976	.30	.0142	.00257	18.1	.11	.16–.44
22. Aversive methods in work sites	26	1,041	.25	.0235	.00409	17.4	.14	.07–.43
23. Control groups	41	3,295	.06	.0062	.00080	12.9	.07	–.02–.16

[a] Standard deviation of success rates after the effects of sampling error have been eliminated (see Method section for explanation). [b] The 80% credibility interval values indicate the range of proportions after the extreme 10% of values on either side of the mean have been eliminated from the distribution. [c] Meta-analysis results when studies in those categories are combined.

the patients appears to promote successful quitting, as evidenced by the 42% success rate of cardiac patients followed by pulmonary patients with a 34% success rate.

A surprising finding is the low effectiveness rate of programs based on physician counseling when compared with other instruction-based programs. Medically sponsored programs involving health professionals (i.e., physicians) regarded as authorities in health matters might have been expected to be more effective than other educational and instructional programs. The fact that this is not so suggests that the content and not the source of information is important in determining success. Conversely, this finding may merely reflect the fact that physicians typically do far less extensive counseling than other health care professionals. Instruction-based programs (28%) are comparable to aversive techniques in general (29%) and are more effective than drug-based programs (17%).

Conditioning-based techniques, such as the aversive methods, have a success rate (29%) similar to those of instructional methods. Among the instructional methods, group withdrawal clinics (30%) are more effective than 5-day plans (26%), which are more effective than educational methods (24%) that passively transmit factual knowledge. This supports the conclusion that traditional didactic approaches that ignore social norms, *self-esteem* enhancement, and environmental contingencies are not as effective as instructional methods incorporating those factors.

Contrary to the conclusions of Schachter (1982), self-care methods resulted in a success rate of only 15%, as against 36% for hypnosis, 31% for smoke aversion techniques, 30% for group withdrawal clinics, 27% for other aversive techniques, 26% for 5-day plans, 30% for acupuncture, and 24% for educational methods. Self-care methods were not better than physician intervention programs involving more than mere counseling or drug-based programs (15% vs. 18% and 17%, respectively).

In the workplace setting, instructional methods had an average success rate of 30% (as against 28% across all settings) and were found to be more effective than aversive techniques, which had an average success rate of 25% in the work setting (compared to 29% across all settings). It appears that in nonwork settings (e.g., school-based programs), aversive techniques (e.g., use of fear videos or electric shocks) are more effective than instructional methods, but in workplaces they perform more poorly than the instructional methods. This finding is also consistent with previous surveys that indicated a worker preference for self-care methods (Hallett, 1986; Walsh, 1984). However, these differences are all small.

The analysis indicates a high success rate (35%) for the miscellaneous category. This high value probably reflects the fact that programs incorporating multiple methods and components have increased chances of meeting the needs of more individuals (Mumford, Schlesinger, & Glass, 1982), thus producing higher success rates than programs focusing on single components or fewer methods of achieving cessation.

The quit rate in the control groups indicates the percentage of smokers who would have quit without any intervention. When analyzing the effect of an intervention across situations, the control group figure used should also represent the average across the situations considered. This estimated average effect size for the control groups is more robust (in terms of sampling error) and is a better benchmark against which to evaluate any intervention than is the control group for that particular study because the idiosyncrasies of particular control samples are averaged out. Our analysis across 41 control groups, with a total sample of 3,295, indicated that 6.4% of smokers can be expected to quit without any intervention.

To determine the net effectiveness of the intervention methods in Table 1, 6.4 percentage points must be subtracted from each success rate. This has not been done for the success rates reported in Table 1 because of space considerations. However, readers should make these adjustments in evaluating the net effect of each intervention. The reader should also note that this adjustment produces larger percentage reductions for some interventions than others. For example, for the self-care category, the adjustment is .15 − .064 = .086, a 42.7% reduction, whereas for hypnosis, the adjustment is .36 − .064 = .30, a 16.7% reduction. Thus, the control group adjustment changes the *ratios* of success rates. For example, the net effectiveness of hypnosis is 3.5 times greater than that of self-care (.30/.086 = 3.488). But for unadjusted success rates, this ratio is only 2.4 (.36/.15 = 2.4). Thus, in comparing relative success rates, the control group adjustment is very important. Furthermore, for some intervention methods, this adjustment has important implications for the *absolute* value of the intervention. For example, the net success rate of medically sponsored programs (No. 19 in Table 1) is only 4.6% (i.e., .11 − .064 = .046), and the net success rate of physician advice is nearly zero (.60%).

Variability of Success Rates

The mean success rates are only part of the story. Success rates varied widely within most program types across different studies. The percentage of observed variance accounted for by sampling error ranged from 3% to 41%. These relatively small percentages are due in part to the fact that many studies are based on large samples and hence have little sampling error. The corrected standard deviations are moderate in size, compared with, for example, those from validity generalization studies. The unweighted, average corrected standard deviation in Table 1 is only .12, and the sample-size-weighted average is only .11. If it were possible to correct for other artifacts in addition to sampling error, the corrected standard deviations would undoubtedly be considerably smaller, and many might be zero. Nevertheless, it is possible that there are moderators that affect the effectiveness of the different intervention methods. However, there is strong evidence that mean success rates are different for different methods. Pooling across all 633 studies and applying meta-analysis yields a corrected standard deviation of .15, whereas the mean corrected standard deviation (sample-size weighted) across methods is only .11. Thus, the standard deviation across all methods is 36% larger than the average standard deviation within methods.

Discussion

Our analysis of the effectiveness of the different intervention methods is based on the assumption that there are no systematic differences in important methodological features across the different intervention methods. A potential systematic difference across the intervention methods that would affect our comparisons is differences in the follow-up periods used to assess effectiveness. This potential concern becomes a reality only if there are systematic differences in follow-up periods across the intervention methods and if these differences are correlated with obtained effectiveness rates. For example, if studies of self-help methods used 1-week follow-up periods, whereas studies of educational methods assessed effectiveness after one year, comparisons across these methods must take account of this systematic difference.

We examined the possibility that differences in the follow-up periods used affect study results by correlating the quit rates and the follow-up time periods for the studies within each of the 15 categories. Thus, a correlation coefficient was computed for each category to summarize the association between success rates and follow-up periods used. The mean observed correlation (between the assessed effectiveness and the follow-up period used to assess effectiveness) across the 15 categories was less than .01 (only .006). This low value indicates that assessed effectiveness does not depend on the follow-up periods used. Furthermore, all of the observed variation in the 15 correlations was found to be due to sampling error, indicating that the mean correlation of .006 is the best estimate of the relationship (between the assessed effectiveness and the follow-up period used to assess effectiveness) for all of the smoking intervention methods. That is, it appears that length of follow-up period does not affect the outcome (success rate) for any of the intervention methods.

Another factor that could affect success rates for different intervention programs is the duration of the programs. Other things being equal, one would expect longer programs to be more successful. But lack of information in many studies precluded an analysis of program duration. In short, our findings indicate the relative effectiveness of the different methods as they exist rather than the effectiveness they would have if all were made the same length. (This is similar to comparing the operational validities of tests in personnel selection instead of true score validities.)

The conflicting conclusions advanced by previous researchers (Cohen et al., 1989; Schachter, 1982) about the superiority of self-help methods over formal intervention methods are most likely the result of use of a small sample of unrepresentative studies. For example, we can expect 80% of self-care methods to have success rates between 3% and 27%, whereas the success rate for smoke aversion methods varies from 14% to 48% and that of acupuncture ranges from 6% to 54%. Given these wide variations, conflicting inferences are likely when a small subset of studies is reviewed. Our findings lead us to conclude that, on average, formal intervention methods have been more effective than self-care methods.

One individual personality characteristic hypothesized as a potential moderator in the literature on the effectiveness of various intervention methods is the degree of addiction of the subjects. The nicotine addiction model (Schachter, 1977; U.S. Department of Health & Human Services, 1987) postulates that heavy smokers will have less success in quitting. However, Schachter (1982) and Rzewnicki and Forgays (1987) found empirical data that do not support this prediction. Cohen et al. (1989), using seven studies, found some evidence supporting the nicotine addiction model. On the basis of the information given in their article, we computed a correlation coefficient (computed from the reported chi-square values indicating the degree of difference in quit rates between light and heavy smokers) for each study and meta-analyzed the resulting seven correlations. The mean was .09 (d = .18) for long-term effects and .07 (d = .14) for short-term effects. Although the effect was in the predicted direction, the difference in the success rates of light and heavy smokers does not appear to be striking.

The results of our analyses are consistent with the conclusions inferred in related domains of research, such as compliance with prescribed medical regimens, patient education, and treatment of addictive behaviors (e.g., Mullen, Green, & Persinger, 1985). Though the procedures used in the intervention methods as well as the criteria for success vary considerably, these domains have in common the underlying concept of attempts to modify individual behavior to enhance the effectiveness of health care. This common core reflecting the psychoeducational component of patient care (Bartlett, 1980; Devine & Cook, 1983; Levy, Iverson, & Walberg, 1980; Mumford et al., 1982) can be used as the basis for general conclusions about educational or behavioral interventions to improve the overall health care of the population. The findings applicable in one area can be generalized and extended to other areas, enabling scientific cumulativeness, replicability, and theory-driven research.

Several methodological problems need to be addressed in primary studies of smoking cessation programs. One such concern addresses the problem of nonresponse during the follow-up. Most of the primary studies are based on participants who reply to follow-ups. These people may disproportionately represent the successful quitters, which would result in a response bias (Berglund et al., 1974). But this would be expected to be true for control group respondents also, and thus the net success rates should not reflect any such response bias. As indicated earlier, the net success rates are the more accurate indicators in any event. Another consideration is the fact that the operational definition of follow-up time period varies. Follow-up time period is operationalized as the time elapsed either from the start of the program or from the time the program ends. Furthermore, abstinence is measured either as abstinence at just the follow-up point in time (e.g., Do you smoke now?) or as abstinence over the entire follow-up period (e.g., Have you ever smoked since the end of the intervention program?). Uniformity in these procedures would be very desirable. Finally, the validity of some of the abstinence measures used (e.g., self-reports, carboxyhemoglobin levels in blood) can be questioned. However, these methodological problems are expected to affect all programs (including control groups) about equally (Schwartz, 1987) and therefore should not alter the adjusted relative success rates. Future research should address the concerns identified in this meta-analytic review and provide a stronger evidential basis for evaluating the effectiveness of the different intervention methods.

Conclusions

On the basis of 633 studies and a sample size of 71,806, we can conclude at present that (a) the average observed success rate across all programs is 25% but is only 18.6% after adjustment for the control group success rate; (b) the most successful methods appear to be instructional and conditioning-based methods (net success rates of 22% and 23%, respectively); and (c) drug-based and medically sponsored programs appear to be the least effective (net success rates of 11% and 5%, respectively) of the methods.

References

Bailar, J. C., III, & Mosteller, F. (Eds.). (1986). *Medical uses of statistics.* Waltham, MA: NEJM Books.

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

Bartlett, E. E. (1980). The contributions of consumer health education to primary care practice: A review. *Medical Care, 18,* 862–871.

Baum, M. L., Anish, D. S., Chalmers, T. C., Sacks, H. S., Smith, H., Jr., & Fagerstrom, R. M. (1981). A survey of clinical trials of antibiotic prophyloxis in colon surgery: Evidence against further use of no-treatment controls. *New England Journal of Medicine, 305,* 795–799.

Bennett, D., & Levy, B. S. (1980). Smoking policies and smoking cessation programs of large employers in Massachusetts. *American Journal of Public Health, 70,* 629–631.

Berglund, F., Bernstein, D. A., Fisinger, R. A., Hochbaum, G. M., Lichtenstein, F., Schwartz, J. L., & Straits, B. C. (1974). *Guidelines for research on the effectiveness of smoking cessation programs.* Chicago: American Dental Association, National Interagency Council on Smoking & Health.

Bernstein, D. A. (1969). Modification of smoking behavior: An evaluative review. *Psychological Bulletin, 71,* 418–440.

Bruvold, W. H., & Rundall, T. G. (1988). A meta-analysis and theoretical review of school based tobacco and alcohol intervention programs. *Psychology and Health, 2,* 53–78.

Callner, D. A. (1975). Behavioral treatment approaches to drug abuse: A critical review of research. *Psychological Bulletin, 82,* 143–164.

Chovil, A. C., Alexander, G. R., Gibson, J. J., & Altekruse, J. M. (1983). Occupational health services in South Carolina manufacturing plants: Results of a survey. *Public Health Reports, 98,* 597–603.

Coggin, T. D., & Hunter, J. E. (1987). A meta-analysis of pricing risk factors in APT. *The Journal of Portfolio Management, 14,* 35–38.

Cohen, S., Lichtenstein, E., Prochaska, J. O., Rossi, J. S., Gritz, E. R., Carr, C. R., Orleans, C. T., Schoenbach, V. J., Biener, L., Abrams, D., DiClemente, C. C., Curry, S., Marlatt, G. A., Cummings, K. M., Emont, S. L., Giovino, G., & Ossip-Klein, D. (1989). Debunking myths about self quitting: Evidence from 10 prospective studies of persons who attempt to quit smoking by themselves. *American Psychologist, 44,* 1355–1365.

Cooper, H. M. (1984). *The integrative research review: A systematic approach.* Beverly Hills, CA: Sage.

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

Cummings, K. M., Emont, S. L., Jaen, C., & Sciandra, R. (1988). Form and quitting instructions as factors influencing the impact of a self-administered quit smoking program. *Health Education Quarterly, 15,* 199–216.

Davis, M. F., Rosenberg, K., Iverson, D. C., Vernon, T. M., & Bauer, J.

(1984). Work site health promotion in Colorado. *Public Health Reports, 99,* 538–543.
Devine, E. C., & Cook, T. D. (1983). A meta-analytic analysis of effects of psychoeducational interventions on length of post surgical hospital stay. *Nursing Research, 32,* 267–274.
Fielding, J. E. (1985). Smoking: Health effects and control. *New England Journal of Medicine, 313,* 491–498.
Fielding, J. E., & Breslow, L. (1983). Health promotion programs sponsored by California employers. *American Journal of Public Health, 73,* 538–542.
Glass, G. V. (1976). Primary, secondary and meta-analysis of research. *Educational Researcher, 5*(10), 3–8.
Glass, G. V. (1977). Integrating findings: The meta-analysis of research. *Review of Research in Education, 5,* 351–379.
Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research.* Beverly Hills, CA: Sage.
Green, L. W., Rimer, B., & Bertera, R. (1978). How cost effective are smoking cessation methods? *World Smoking and Health, 3,* 33–40.
Hallett, R. (1986). Smoking intervention in the work place: Review and recommendations. *Preventive Medicine, 15,* 213–231.
Haring, M. J., Okun, M. A., Stock, W. A., Miller, W., Kinney, C., & Ceurvorst, W. R. (1981, April). *Reliability issues in meta-analysis.* Paper presented at the annual meeting of the American Educational Research Association, Los Angeles.
Hattie, J. A., & Hansford, B. C. (1982). Self-measures and achievement: Comparing a traditional review of literature with a meta-analysis. *Australian Journal of Education, 26,* 71–75.
Hattie, J. A., & Hansford, B. C. (1984). Meta-analysis: A reflection on problems. *Australian Journal of Psychology, 36,* 239–254.
Hedges, L. V. (1987). How hard is hard science, how soft is soft science: The empirical cumulativeness of research. *American Psychologist, 42,* 443–455.
Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analyses.* San Diego, CA: Academic Press.
Hunter, J. E. (1990, May). *The control group and meta-analysis.* Unpublished manuscript, Department of Psychology, Michigan State University.
Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings.* Newbury Park, CA: Sage.
Jackson, G. B. (1980). Methods for integrative reviews. *Review of Educational Research, 50,* 438–460.
Jeffery, R. W., & Wing, R. A. (1983). Recidivism and self cure of smoking and obesity: Data from population studies. *American Psychologist, 38,* 852.
Keutzer, C. S., Lichtenstein, E., & Mees, H. L. (1968). Modification of smoking behavior: A review. *Psychological Bulletin, 70,* 520–533.
Kinder, B. N., Pape, N. E., & Walfish, S. (1980). Drug and alcohol education programs: A review of outcome studies. *The International Journal of Addictions, 15,* 1035–1054.
Kottke, T. E., Battista, R. N., DeFriese, G. H., & Brekke, M. L. (1988). Attributes of successful smoking cessation interventions in medical practice: A meta-analysis of 39 controlled trials. *Journal of American Medical Association, 259,* 2883–2889.
Kristein, M. M. (1982). The economics of health promotion at the work site. *Health Education Quarterly, 9,* 27–36.
Kristein, M. M. (1983). How much can business expect to profit from smoking cessation? *Preventive Medicine, 12,* 358–381.
Leventhal, H., & Cleary, P. D. (1980). The smoking problem: A review of the research and theory in behavioral risk modification. *Psychological Bulletin, 88,* 370–405.
Levy, S. R., Iverson, B. K., & Walberg, H. J. (1980). Nutrition education research: An interdisciplinary evaluation and review. *Health Education Quarterly, 7,* 107–126.
Lichtenstein, E., & Danaher, B. G. (1977). Behavior therapy and the control of cigarette smoking. In J. Steinfeld, W. Griffiths, K. Ball, R. M. Taylor (Eds.), *Health consequences, education, cessation activities, and social action: Vol. II. Proceedings of the third world conference on smoking and health* (HEW Publication No. 77-1413). Washington, DC: U.S. Government Printing Office.
Moskowitz, J. M. (1983). Preventing adolescent substance abuse through drug education. In T. Glynn, C. Leukefeld, & J. Ludford (Eds.), *Preventing adolescent drug abuse: Intervention strategies* (pp. 233–249). Rockville, MD: National Institute on Drug Abuse.
Mullen, P. D., Green, L. W., & Persinger, G. S. (1985). Clinical trials of patient education for chronic conditions: A comparative meta-analysis of intervention types. *Preventive Medicine, 14,* 753–781.
Mumford, E., Schlesinger, H. J., & Glass, G. V. (1982). The effects of psychological intervention on recovery from surgery and heart attacks: An analysis of the literature. *American Journal of Public Health, 72,* 141–151.
Pearson, C. E. (1980). The emerging role of the occupational physician in preventive medicine, health promotion, and health education. *Journal of Occupational Medicine, 22,* 104–106.
Prochaska, J. O. (1983). Self changers versus therapy changers versus Schachter. *American Psychologist, 38,* 853–854.
Roemer, R. (1983). The role and evolution of legislation to control smoking. *World Smoking and Health, 8,* 4–14.
Rzewnicki, R., & Forgays, D. G. (1987). Recidivism and self-cure of smoking and obesity: An attempt to replicate. *American Psychologist, 42,* 97–100.
Schachter, S. (1977). Nicotine regulation in heavy and light smokers. *Journal of Experimental Psychology: General, 106,* 5–12.
Schachter, S. (1982). Recidivism and self-cure of smoking and obesity. *American Psychologist, 37,* 436–444.
Schaps, E., DeBartolo, R., Moskowitz, J., Palley, C. S., & Churgin, S. (1981). A review of 127 drug abuse prevention program evaluations. *Journal of Drug Issues, 11,* 17–43.
Schoenbach, V. J., Orleans, C. T., & Wagner, E. H. (1988). *Recruiting smokers to a self-help cessations study in an HMO: The free and clear experience.* Unpublished manuscript, University of North Carolina at Chapel Hill, School of Public Health.
Schwartz, J. L. (1969). A critical review and evaluation of smoking control methods. *Public Health Reports, 84,* 483–506.
Schwartz, J. L. (1987). *Review and evaluation of smoking cessation methods: The United States and Canada* (NIH Publication No. 87-2940). Washington, DC: U.S. Department of Health and Human Services.
Schwartz, J. L., & Rider, G. (1978). *Review and evaluation of smoking control methods: The United States and Canada, 1969–1977* (HEW Publication No. 79-8369). Atlanta, GA: U.S. Bureau of Health Education.
Swingle, M. (1980). The legal conflict between smokers and nonsmokers: The majestic vice versus the right to clean air. *Missouri Law Review, 45,* 444–475.
Thompson, E. L. (1978). Smoking education programs 1960–1976. *American Journal of Public Health, 68,* 250–257.
U.S. Department of Health and Human Services. (1987). *Nicotine addiction: A report of the Surgeon General* (Publication No. HHS/PHS/CDC-88-8406). Washington, DC: U.S. Government Printing Office.
U.S. Department of Health and Human Services. (1989). *Reducing health consequences of smoking: 25 years of progress. A report of the Surgeon General* (Publication No. HHS/PHS/CDC/89-8411). Washington, DC: U.S. Government Printing Office.
Walsh, D. C. (1984). Corporate smoking policies: A review and an analysis. *Journal of Occupational Medicine, 26,* 17–22.

Received October 7, 1991
Revision received February 3, 1992
Accepted February 5, 1992 ■