SOCIAL SKILLS TRAINING IN ADULT PSYCHIATRIC POPULATIONS: A META-ANALYSIS

PATRICK W. CORRIGAN
University of Chicago

Summary — A meta-analysis was conducted on 73 studies of social skills training in four adult psychiatric populations: developmentally disabled, psychotic, nonpsychotic, and legal offenders. Findings from this analysis showed that patients participating in social skills training programs broadened their repertoire of skills, continued to demonstrate these skills several months after treatment, and showed diminished psychiatric symptoms related to social dysfunctions. Although results from an ANCOVA comparing effect sizes across the four populations (with design quality as a covariate) were nonsignificant, consistent trends suggested that social skills training had the greatest effect on developmentally disabled groups and the least effect on offender groups. In addition, social skills training was found to be relatively more effective in outpatient than inpatient settings.

In the past 25 years, there has been a veritable explosion of research on social skills training that has produced numerous comprehensive reviews of the topic (Bellack & Hersen, 1979; Corrigan, Schade, & Liberman, in press; Curran & Monti, 1982; Marzillier, 1978; McFall, 1982; Morrison & Bellack, 1984; Wallace et al., 1980). Social skills training programs have been developed for children, adolescents, and adults in school, psychiatric, and medical settings. The content areas of social skills training have broadened from a narrow focus on conversational behavior to include assertiveness and problem-solving skills. The components of skills training — instructions, modeling, rehearsal, feedback, and homework — have been studied alone and in varying combinations to apportion their relative influence on outcome. Skills training packages have been compared to other treatment modalities including cognitive-behavioral techniques, psychotropic medication administration, and behavior family management to determine their independent and synergistic influence on patient populations.

Assertions regarding the utility of social skills training have varied from initial excitement that it offered a panacea for psychopathologies involving social problems to a more sober look at its limitations (Curran, 1979; Liberman, Nuechterlein, & Wallace, 1982). Despite enormous amounts of literature demonstrating the effectiveness of social skills training to enhance the acquisition, maintenance, and generalization of skills, as well as to diminish symptoms related to social dysfunction, several questions remain. Are acquired skills maintained several months after treatment? Do skills generalize to settings outside the training milieu? Do some psychiatric populations respond better to social skills training than others? Is social skills training more effective when conducted in community settings? As a first step toward answering these questions, a meta-analysis was conducted on adults with psychiatric diagnoses who had

The author completed this project as a post-doctoral scholar with the UCLA Clinical Research Center for Schizophrenia and Psychiatric Rehabilitation.

Requests for reprints should be addressed to Dr. Patrick W. Corrigan, Department of Psychiatry, University of Chicago, 5841 S. Maryland Avenue, Chicago, Ill. 60637, U.S.A.
participated in social skills training investigations.

Method

The Psychological Abstracts of the American Psychological Association were searched for all social skills training studies on adult psychiatric samples published between January 1970 and August 1988. Only studies with subjects between the ages of 18 and 62 were included. Studies on normal groups or analogue populations, in the area of health psychology, or with head trauma victims were excluded. Findings from the Psychological Abstracts were augmented by a search of Blampied’s (1982) comprehensive bibliography of social skills training. A total of 464 publications conforming to these criteria were identified of which 150 were empirical studies. The remainder of the publications were single case experiments \((N = 56)\), literature reviews \((N = 98)\), foreign language publications \((N = 64)\), or unpublished dissertations \((N = 96)\).

Of the 150 empirical articles, 73 studies (49%) reported statistics that could be transformed into an estimate of effect size. The list of studies included in this meta-analysis may be obtained by writing to the author. Effect size was determined from the ratio of the differences between group means and an estimate of the pooled variance (Hedges & Olkin, 1985) or from the estimate of effect size calculated from \(t\)- and \(F\)-tests (Wolf, 1986). Studies were excluded if they reported nonparametric tests or ANOVAs across more than two groups; these studies yield more biased estimates of effect size (Wolf, 1986). Several of the 73 studies included “sub-studies” comprised of independent samples which tested more than one social skills training condition. As a result, the 73 studies yielded 90 index comparisons.

Several index comparisons measured more than one dependent variable resulting in the determination of multiple effect sizes in the same study. The median number of effect sizes per index comparison was 2.3 with comparisons ranging from 1 to 12 effect sizes. The 90 comparisons produced a total of 255 effect sizes representing findings based on data from 1836 subjects. Multiple effect sizes within a single comparison are not independent. Therefore, their significance can be exaggerated if averaged within subsequent analyses (Kulik, 1983). The median effect size across variables within a comparison was determined to provide a more conservative index of outcome (Rosenthal & Rubin, 1986).

The 90 index comparisons were evaluated using guidelines established elsewhere (Glass, McGaw, & Smith, 1981; Giblin, Sprengle, & Sheehan, 1985). The criteria included judgements regarding (1) methodological variables, (2) dependent measures, and (3) treatment variables and were discussed more fully in respective sections of this paper.

Results

How big an effect size is big? Strength of effect can also be determined by comparison of effect sizes produced from other meta-analyses on similar topics. An independent meta-analysis investigating the efficacy of generic clinical interventions with chronic, mentally ill individuals found a mean effect size of 0.32 (Videka-Sherman, 1988). The significance of effect sizes may also be determined by calculating whether the mean effect size is significantly different from zero; this statistic is reported as a \(t\)-test throughout the paper where appropriate.

Findings from meta-analyses are confounded by the design of studies from which effects are gleaned (i.e., within-subject investigations tend to over-estimate effect size compared to between-group studies) or by the statistics used to determine the results (i.e., effect sizes determined from descriptive statistics tend to be more conservative than effect sizes derived from inferential statistics). Although effect sizes in this analysis were determined from
studies that included within- and betweensubject designs as well as studies that reported inferential and descriptive statistics, results of a separate analysis showed that these biases did not confound my findings. In addition, results of an analysis of homogeneity showed that effect sizes in this analysis were homogenous. Homogeneity tests whether the range of effect sizes across several studies generated from dependent variables and presumed to load into a common factor actually represents a solitary construct (Wolf, 1986).

(1) Methodological Design Quality

One of the benefits of meta-analysis is a description of the experimental design and dependent variables used to investigate the treatment of interest. A strategy was developed for determining an index of design quality from 10 criteria necessary for good treatment outcome research: sufficient sample size, valid diagnosis, presence of a control group, random assignment, specification of concurrent drug treatment, low treatment mortality, standardized treatment protocols, blind raters, experienced trainers, and nonreactive dependent measures. The criteria were gleaned from seminal work on experimental design and psychotherapy research (Campbell & Stanley, 1963; Luborsky, Singer, & Luborsky, 1975) and summed to provide an overall design quality. In this summation, the criteria were given equal weights, based on findings from a separate study which showed that research psychologists as a group do not give priority to any single criterion. Most of the social skills training studies had at least 10 subjects per group, low treatment mortality, and a control group. Less than half the studies reported the use of trainers that were experienced with social skills training or raters that were blind to treatment conditions. Only 43% of studies reported subjects’ diagnoses. Half of the reported diagnoses were based on interview format while half were made from self-report measures like the MMPI. Only 49% of the studies discussed the use of medications.

Total design quality ranged from 2 to 10 across the 73 studies with a mean of 6.7 (SD = 1.83). The average design quality per year sampled in the meta-analysis revealed a slight increasing trend during the early 1970s suggesting that the quality of design was improving in social skill training investigations. However, overall design quality distribution was relatively flat, showing that researchers have not heeded the recurrent call for better designed investigations of social skills training. Frequently, meta-analyses are criticized for combining outcomes with qualitatively different experimental designs into a summary index, thereby lumping apples and oranges. The problem was diminished in this meta-analysis by adjusting effect size by design quality. Subsequent comparisons included ANCOVAs with effect size adjusted by the design quality covariate.

(2) Dependent Measures

Several dependent variables have been collected on social skills training investigations and may be classified into four categories: dependent variables which validated that skills were acquired, dependent variables which determined whether the acquisition of skills subsequently diminished symptoms and improved personal adjustment, dependent variables which demonstrated that skills were maintained for a period of time after training had ended, and dependent variables which showed that skills had generalized to settings outside the training situation. Each dependent variable category subsumed several different measurement strategies. The number of studies using each strategy, the mean effect size of all studies including that measure, and its standard deviation are summarized in Table 1. The table also included the median effect size of all measures in the four measurement categories. Significant differences between
Table 1

Summary Statistics of Commonly Measured Dependent Variables

<table>
<thead>
<tr>
<th>Variables</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>No. of Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acquired Skills</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Role play</td>
<td>1.70†</td>
<td>1.56</td>
<td>57</td>
</tr>
<tr>
<td>Nat. obs</td>
<td>1.34†</td>
<td>1.38</td>
<td>17</td>
</tr>
<tr>
<td>Self-report</td>
<td>0.93†</td>
<td>0.76</td>
<td>28</td>
</tr>
<tr>
<td>Median acquisition</td>
<td>1.43†</td>
<td>1.37</td>
<td>82</td>
</tr>
<tr>
<td>Reduced Symptoms and Improved Personal Adjustment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General psychopathology</td>
<td>1.08†</td>
<td>0.78</td>
<td>18</td>
</tr>
<tr>
<td>Social anxiety</td>
<td>0.83†</td>
<td>0.73</td>
<td>25</td>
</tr>
<tr>
<td>Hostility</td>
<td>0.79*</td>
<td>0.58</td>
<td>2</td>
</tr>
<tr>
<td>Self-esteem</td>
<td>1.01†</td>
<td>0.65</td>
<td>7</td>
</tr>
<tr>
<td>Median symptom/adjust</td>
<td>0.99†</td>
<td>0.75</td>
<td>35</td>
</tr>
<tr>
<td>Acquired Skills Generalized to Other Settings</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median generalization</td>
<td>0.92†</td>
<td>0.52</td>
<td>19</td>
</tr>
<tr>
<td>Acquired Skills Maintained After Social Skills Training</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maintenance of acquired skills</td>
<td>1.40†</td>
<td>1.70</td>
<td>31</td>
</tr>
<tr>
<td>Symptom Reduction and Personal Adjustment After Social Skills Training</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maintenance of symptom/adjustment</td>
<td>1.20†</td>
<td>0.64</td>
<td>15</td>
</tr>
</tbody>
</table>

*p < .05; †p < .01.

effect sizes in Table 1 and zero are marked accordingly.

Acquisition. The acquisition of targeted skills immediately after training has been measured using role play, naturalistic observation, and self-report. Role play was the most frequent measure of behavior change used in these studies (N = 57) with the Behavior Assertiveness Test (Eisler, Hersen, Miller, & Blanchard, 1975; Eisler, Miller, & Hersen, 1973) and the Confederate's Role Play Test (Goldsmith & McFall, 1975) as frequent examples. Researchers more commonly constructed role play situations specific to the content of the social skills training module tested in their investigation and subsequently established reliability of the instrument through high inter-rater agreement.

Naturalistic observation has been argued to be a more reliable and valid measure of skill acquisition than role play (Bellack, Hersen, & Turner, 1979; Curran, 1978). However, this measurement strategy was used in only 17 studies. The average effect size for studies using naturalistic observation was smaller than for studies using role play. Self-report measures (N = 28) were relatively easy to collect and were typically interview or paper-and-pencil assessments of the subjects' probable response to social interactions. The Wolpe-Lazarus Assertiveness Scale (Wolpe & Lazarus, 1966) and the Rathus Assertiveness Scale (Rathus, 1973) were the most common measures in this category. Self-report measures have comparatively reduced sensitivity to social skills training treatment effects as well.

Symptom reduction and personal adjustment. The various instruments measuring reduction of psychiatric symptoms were collapsed into one category representing change in symptoms characteristic of that diagnostic group. Improvement in psychotic samples was assessed most frequently by overall changes on the Brief Psychiatric Rating Scale (Overall & Gorham,
1962) while diminution in dysphoria was measured on nonpsychotic depressed groups using instruments like the Beck Depression Inventory (Beck et al., 1961). Social skills training effect size across all psychiatric samples was significantly greater than zero.

Personal adjustment was sampled by coding studies for variables that assessed social anxiety, interpersonal hostility, and self-esteem. These variables were infrequently measured, with hostility and low self-esteem used in only two and seven studies, respectively. Social anxiety was assessed more often (N = 25) primarily with the Social Avoidance and Distress Scale and the Fear of Negative Evaluation Scale (Watson & Friend, 1969). Results showed that effect sizes for social anxiety were significantly greater than zero.

Maintenance and generalization of effects. Maintenance was assessed by determining the effect size of dependent variable categories measured after several months passed since the completion of social skills training. Of the 90 index comparisons, 31 provided information on the maintenance of skill acquisition and 15 assessed maintenance of symptom reduction and personal adjustment. The average follow-up period was 5.4 months (SD = 5.2) with a range from 1 month to 2 years. Mean effect size of measures taken after the completion of treatment paralleled post-treatment findings; measures of maintenance of acquisition effects were associated with greater effect sizes than follow-up measures of symptom reduction. These findings suggested that specific treatment effects, particularly skills acquisition, can be observed several months after treatment.

Studies measuring generalization of trained skills to situations outside of training were coded for the presence of instruments assessing change in leisure, employment, and dating skills. Despite repeated calls for measurement of this construct (Robertson, Richardson, & Youngson, 1984; Wallace et al., 1980), only 19 studies measured any aspect of generalization. The mean effect size of all generalization measures collapsed under one construct was nevertheless fairly high and significant.

(3) Variables Relevant to Treatment

Effect size by diagnostic group. A subset of the variables in Table 1 were used to determine effect across diagnosis. Role play was selected because of the frequency with which it had been used in the 90 index comparisons. In addition, the median effect size in each of the four dependent variable categories—the acquisition of trained skills, its subsequent effect on symptom reduction and personal adjustment, the maintenance of acquired skills over time, and the generalization of skills to other settings—and the median overall effect size were analyzed across four diagnostic groups.

A developmentally disabled group consisted of mentally retarded individuals alone; no studies were found on autistic subjects. A psychotic group consisted of 11 schizophrenic samples and six samples which had mixed psychotic diagnoses. A nonpsychotic group consisted of seven samples with mild depression and 13 samples requiring psychiatric help for life problems, but not receiving a specific DSM-III diagnosis. An offender group comprised 11 samples of drug and alcohol abusing individuals, two sex offender samples, and two other criminal offender samples. Twenty-three studies were excluded because they examined mixed psychotic and nonpsychotic groups.

The effect sizes for the four psychiatric groups are summarized in Table 2. ANCOVAs were conducted on group means for the variables using study design quality as a covariate and are also summarized in Table 2. Although nonsignificant, trends suggested that the greatest treatment effects were obtained with developmentally disabled individuals across most variables. Developmentally disabled samples may have shown large effect sizes because they started closer to the floor on most dependent measures and progressed through a greater performance range to reach "normal"
Table 2

Effect Size for Six Dependent Variable Categories Across Four Diagnostic Classes Adjusted by Design Quality

<table>
<thead>
<tr>
<th>Variables</th>
<th>Developmentally Disabled</th>
<th>Psychotic</th>
<th>Nonpsychotic</th>
<th>Offenders</th>
<th>ANCOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N = 11</td>
<td>N = 12</td>
<td>N = 8</td>
<td>N = 7</td>
<td>F(3, 33) = 2.04</td>
</tr>
<tr>
<td>Roleplay</td>
<td>3.25†</td>
<td>1.58†</td>
<td>1.47†</td>
<td>1.66†</td>
<td>p = .13</td>
</tr>
<tr>
<td>Median acquisition</td>
<td>2.58†</td>
<td>1.37†</td>
<td>1.42†</td>
<td>1.15†</td>
<td>F(3, 55) = 2.07</td>
</tr>
<tr>
<td>Median symptom and adjustment</td>
<td>3.22</td>
<td>0.96*</td>
<td>1.00†</td>
<td>0.54</td>
<td>p = .11</td>
</tr>
<tr>
<td>Median generalization</td>
<td>0.49</td>
<td>1.74†</td>
<td>1.25†</td>
<td>1.19</td>
<td>†</td>
</tr>
<tr>
<td>Maintenance of acquired skill</td>
<td>2.89†</td>
<td>1.22*</td>
<td>1.05*</td>
<td>0.87</td>
<td>†</td>
</tr>
<tr>
<td>Maintenance of symptom/adjust</td>
<td>1.09</td>
<td>1.74*</td>
<td></td>
<td></td>
<td>†</td>
</tr>
<tr>
<td>Overall ES</td>
<td>2.07†</td>
<td>1.31†</td>
<td>1.33†</td>
<td>1.06†</td>
<td>F(3, 62) = 1.37</td>
</tr>
<tr>
<td></td>
<td>N = 15</td>
<td>N = 17</td>
<td>N = 20</td>
<td>N = 15</td>
<td>p = .26</td>
</tr>
</tbody>
</table>

*p < .05; †p < .01; †The number of studies per cell was too small for an ANCOVA.

Status. Social skills training produced the smallest effect size for legal offenders. Effect sizes were relatively similar for psychotic and nonpsychotic samples. Effect sizes within a diagnostic class were compared to determine if effects varied across dependent variable categories. The developmentally disabled sample showed high effect sizes on all dependent variable categories except the median of generalization measures. The limited generalization effects may be explained by their cognitive deficits. Perhaps the reduced intellectual abilities of developmentally disabled subjects prevented transfer of learned skills to new settings.

Both psychotic and nonpsychotic samples acquired, maintained, and generalized targeted skills. However, the effect of these new skills on symptom reduction was muted, suggesting that acquired skills do not directly replace maladaptive behaviors. Offenders demonstrated relatively large skill acquisition but did not retain the skills well. The large increase in the offenders' repertoire had less an effect on symptoms and social adjustment measures. In part, these findings may represent the syntonic character of offenders' maladaptive behaviors.

**Effect size by treatment setting.** Past research has suggested that skills training conducted in outpatient settings produced better outcome than training completed in hospitals (Test & Stein, 1978). Of the 90 indexed comparisons in this meta-analysis, 38% examined programs that were carried out within inpatient settings while 62% examined outpatient social skills training programs. Outpatient treatment occurred in a variety of settings, with mental health clinics being the most frequent (N = 43), followed by sheltered workshops (N = 7), and community residential settings (N = 3). The effect sizes for inpatient and outpatient social skills training programs across the four psychiatric populations are presented in Table 3. The skewed distributions of treatment settings across diagnostic groups may represent the therapeutic zeitgeist in which these studies were conducted. For example, developmentally disabled patients had only been tested at outpatient settings; this finding paralleled the movement to return these patients to the
social skills training strongly and consistently enhanced acquisition of skills that were maintained for several months after treatment. Social skills training seemed to be better facilitated when training occurred in outpatient rather than inpatient settings. Although large effects were demonstrated in all four adult psychiatric categories, specific variables within each group seemed to interact with effect size. For example, developmentally disabled patients may have shown the greatest effects because they had the lowest skill functioning measures, leaving a broader possible range of acquisition. Offender groups may have shown the smallest effect sizes because they are coerced into treatment. Studies must continue to emphasize experimental designs which adequately test social skills training hypotheses. Specifically, future social skills training protocols should incorporate nonreactive behavioral measures, independent and blind raters, treatment fidelity measures, and better evidence of patient diagnosis.

This meta-analysis is also interesting for the questions it is unable to address. Studies were coded for content (conversation skills versus assertiveness versus problem solving) and for component techniques (instructions, modeling, role play, feedback, homework, reinforcement). Few studies, however, were completed on these topics so that the relative efficacy of each could not be analyzed.

Findings from the meta-analysis suggest processes that may mediate social skills training effects. Research has shown that skill learning is related to cognitive processes (Bowen et al., 1980) so that models which link cognitive functioning with skills training provide a useful model for future treatment and research programs (Corrigan, Schade, & Liberman, in press). Similarly, our findings suggest that social skills training effects interact with patients’ motivation level, especially in offender groups. Future research should determine the relative size of treatment outcomes generated by the interaction of skills training and incentive strategies. Theoretical founda-
tions like these provide fertile ground for explaining current phenomena as well as for guiding future research.

Acknowledgements — I would like to thank Mark Schade, Geralyn Freeland-Soper, Ruth Davies-Farmer, and Melinda Stolley for their work on the meta-analysis. This project was made possible in part by a grant from First Step, Evanston, Illinois.

References

Luborsky, L., Singer, B., & Luborsky, L. (1975). Comparative studies of psychotherapies: Is it that "everyone has won and all must have prizes?" Archives of General Psychiatry, 32, 985-1002.