

## EVALUATION OF AN INTERVIEWEE SKILLS TRAINING PROGRAM IN A NATURAL FIELD EXPERIMENT

MICHAEL A. CAMPION  
Krannert School of Management  
Purdue University

JAMES E. CAMPION  
University of Houston

The purpose of this study was to provide an evaluation of an interviewee skills training program in a field setting using a broad sample of interviewers, jobs, and candidates. The training program ( $n = 158$ ) derived its content from previous literature and used multiple instructional techniques. An experimental design was employed with a self-study (placebo control) group ( $n = 140$ ), random assignment, and many levels of evaluation criteria. Measures exhibited acceptable reliability, and statistical power was high for all analyses. Results indicated that class members responded positively to the program and demonstrated substantial learning. Interviewer evaluations of behavior failed to distinguish between experimental groups, however, and no differences were observed in job offers. Furthermore, no differences were observed between study participants and nonparticipants ( $n = 174$ ) in terms of interview behavior and job offers. This study illustrates the danger of relying solely on reaction or learning criteria in the evaluation of training programs. Several explanations are offered for the findings.

The number of persons seeking jobs or job changes at any given time is substantial (Current Labor Statistics, 1986; Rosenfeld, 1977). Recent trends include an increase in job search and career change services (Solomon, 1986). These services employ a variety of assessment and training strategies to help their clients define career interests and enhance job-hunting skills. Although considerable research is available to guide interest assessment (e.g., Westbrook, 1985), very little work has been done to examine the effectiveness of programs to improve job-hunting skills (Schwab, Rynes, & Aldag, 1987). The purpose of this study was to assess the effectiveness

---

Special thanks to Dr. Paul Thayer and his graduate students in the Training Research Seminar (Psychology 561) at North Carolina State University for their comments and suggestions on the design of this study. Thanks also Drs. Kenneth Wexley of Michigan State University, Michael Harris and Chris Berger of Purdue University, Milton Hakel and Allan Jones of the University of Houston, and Paul Thayer for their comments on an earlier version of this manuscript.

Correspondence and requests for reprints should be addressed to Michael A. Campion, Krannert School of Management, Purdue University, West Lafayette, IN 47907.

of a principal component of many of these programs—interviewee skills training.

While there have been numerous studies exploring interviewee training efforts (Arvey & J. Campion, 1982), the focus has been on inexperienced job candidates. The bulk of the prior research falls within two areas. The first and largest area focuses on improving the interview-taking skills of members of special populations that are disadvantaged in employment. Examples include rehabilitation clients (Keith, Engelkes, & Winborn, 1977; Pinto, 1979; Stevens & Tornatzky, 1976; Venardos & Harris, 1973), psychiatric patients (Furman, Geller, Simon, & Kelly, 1979; Kelly, Laughlin, Claiborne, & Patterson, 1979; Kelly, Urey, & Patterson, 1981; Logue, Zenger, & Gohman, 1968), culturally disadvantaged (Barbee & Keil, 1973; Jackson, 1972; Keil & Barbee, 1973), adolescents with behavioral problems (Braukmann, Fixsen, Phillips, Wolf, & Maloney, 1974), and prison inmates (Speas, 1979).

The second area of research focuses on students, particularly college students. Although some of this research is oriented toward students with communication problems (Hollandsworth, Glazeski, & Dressel, 1978; Hood, Lindsay, & Brooks, 1982) or poor employment outlook (Austin & Grant, 1981; Heimberg, Stanley, & Blankenberg, 1982), most of the studies are oriented toward improving the interview-taking skills of normal students (Harrison et al., 1983; Hollandsworth, Dressel, & Stevens, 1977; Seigfried & Wood, 1983; Wild & Kerr, 1984).

Research findings have generally been promising up to this point, with most studies reporting some success in improving interviewee knowledge or skills. While many studies tried to enhance generalizability by using personnel managers as evaluators of post-training performance, nearly all studies have been conducted in laboratory settings. Only two studies included a field follow-up on employment success (Keith, et al., 1977; Stevens & Tornatzky, 1976), but both had very small sample sizes.

The purpose of this study was to evaluate an interviewee skills training program in a field setting with a large industry sample.

### *Setting and Research Approach*

The present study took advantage of an opportunity to conduct a natural field experiment. The organization was embarking on a voluntary transfer program to address a human resource imbalance. Large numbers of employees in manufacturing and development operations were being interviewed for promotions and lateral transfers to marketing divisions of the organization. Development of an interviewee skills training program was commissioned to enhance the number of employees receiving offers.

However, time and staff resource limitations dictated that not all candidates could receive significant training. Thus, a natural field experiment was justified wherein candidates could be randomly assigned to the limited training slots or to a self-study (placebo control) condition.

To provide a rigorous evaluation, an evaluation plan originally proposed by Kirkpatrick was utilized (1959a, 1959b, 1960a, 1960b). He suggested four levels of criteria: (1) Reaction—How well did the participants like the training program? (2) Learning—What facts, principles, or techniques did the participants understand and absorb? (3) Behavior—Did the program change any of the participants' behavior back on the jobs (or in the interviews)? (4) Results—Did the program have an impact on important outcomes for the organization or participants (like getting job offers)? Although many research texts on training encourage this form of in-depth evaluation (e.g., Goldstein, 1986; McGehee & Thayer, 1961; Wexley & Latham, 1981), published studies using all four levels in an industrial training evaluation are rare (e.g., Latham & Saari, 1979).

### *Method*

#### *Sample*

A broad sample of jobs, interviewers, and interviewees was obtained from a large electronics company. Jobs to be staffed included sales, engineering, programming, computer operations, administration, secretarial, and other support positions. Ninety-nine interviewers participated; the group consisted of first- through third-level marketing and support managers (e.g., sales, marketing, sales engineering, customer service, office administration, etc.).

Interviewees came from a facility that was responsible for developing and manufacturing electronic equipment. The experimental group contained 158 and the control group 140 employees. For some analyses, data were also available on 174 nonparticipants. Nonparticipants consisted of employees who did not volunteer for the program (see *Procedure*).

Specific sample sizes for each measure are described below. In all cases, statistical power is at least 95% to detect a medium effect size at  $p < .05$ . (Note: According to Cohen, 1977, a medium effect size is equal to a .5 *SD* difference between groups, or  $r = .24$ ).

Mean tenure of the interviewees was 2.1 years ( $SD = 2.9$ ) for job and 7.5 years ( $SD = 6.9$ ) for company. Interviewees' current jobs were 43.7% professional, 22.3% technical, 19.4% administrative, and 14.6% manufacturing. Approximately 40% had a four-year college degree or more, and 72.6% had some education past high school. Only nonsignificant or trivial differences existed between the training and self-study groups, or between

study participants and nonparticipants, on these demographic variables and on age, race, and sex.

### *Procedure*

Selection took place in three stages. First, presentations were made to employees describing available positions. Second, managers from the marketing organizations screened interested candidates by reviewing personnel files. Third, marketing managers returned to the facility to conduct the interviews.

These stages were repeated for each marketing organization recruiting at this facility. All employees were available for recruiting, except for small numbers of employees with critical skills (e.g., some programming positions) or in departments with unusually high workloads. Candidates could be interviewed for more than one job, and they could be interviewed by several managers from the same marketing organization or by managers from different organizations.

To schedule interviews, current managers of employees were called and informed of times and places. During the conversation, managers were told: "An Interview-Taking Skills Program is available to help your employee prepare for the upcoming interview. If interested, have your employee call (staff member's name and extension)." When (if) the employee called, he or she was randomly assigned (using a Hewlett Packard random number generator, Model HP97) to either the training or self-study condition as follows. Training: "We have a training program designed to help you prepare for your upcoming interview to enhance your interview-taking skills. It is a half-day program. Are you interested?" Self-study: "We have some reading materials designed to help you prepare for your upcoming interview. If you are interested, we can have them delivered to you."

It was necessary to have employees call back for random assignment (rather than simply telling the manager), because many managers had more than one employee taking interviews. Also, this procedure limited the study to volunteers and avoided the sending of reading materials to uninterested employees, thus allowing separate analyses of "pure" nonparticipants. Approximately 63% of those solicited called back and were assigned to one of the study conditions.

There were 68 (22.8%) exceptions to the random assignment procedure for a variety of reasons (e.g., specific requests for training from employees assigned to self-study, or scheduling conflicts causing employees assigned to training to instead be given only reading materials). All analyses were conducted both with and without these exceptions, and no differences were found. Therefore, only results for the total sample will be presented.

*Training condition.* Most research on interviewee skills training involves some combination of lecture, discussion, role playing or other form of practice, feedback and reinforcement, and occasional videotaping (e.g., Barbee & Keil, 1973; Hollandsworth, et al., 1978; Keil & Barbee, 1973; Kelly, et al., 1979, 1981; Logue et al., 1968). Efforts to clearly differentiate the effectiveness of various techniques have yielded mixed results (Austin & Grant, 1981; Hollandsworth, et al., 1977; Speas, 1979; Venardos & Harris, 1973). Furthermore, recent meta-analyses on managerial training suggest that combinations of techniques, such as lecture and discussion plus role playing or other practice, are effective on a wide array of criteria (Burke & Day, 1986). Thus, this study used a combination of lecture, discussion, role playing, and feedback.

The program had three sections. First, "Advice on How to take Interviews" consisted of lecture and discussion and drew largely from popular books on the subject (e.g., Bolles, 1982; Medley, 1984; Pettus, 1981; Robertson, 1978). It consisted of eight topics: appearance and dress (e.g., Dress conservatively. Check appearance before entering the interview.), interviewing etiquette (e.g., Don't smoke or chew gum. Remember the interviewer's name.), preparation (e.g., Learn about the product or service before the interview. Bring a pen and notebook.), answering questions (e.g., Answer questions specifically and decisively. Remember you are selling yourself.), attitudes (e.g., You must show that you are interested. You must communicate enthusiasm and a positive attitude.), nervousness (e.g., Interviewers expect a little nervousness. Do not retract an answer in the face of silence.), verbal expression (e.g., Be aware of proper grammar. Speak clearly and watch your pronunciation.), and nonverbal behavior (e.g., Maintain good eye contact. Be an active listener.).

Second, "Preparing and Practicing Answers to Likely Questions" asked participants to develop answers to 20 questions on three topics: yourself (e.g., What are your strengths? What are your long-term goals?), education and experience (e.g., What future plans do you have for education and training? What do you like most and least about your current job?), and match with the job opening (e.g., Why do you want this job? Why do you think you would be successful?). Participants then divided into groups of three to role play being interviewed and to give feedback to each other. The instructor observed and gave feedback.

Third, "Characteristics of Interviews and Interviewers" addressed the question: What do we know about interviews that might help prepare an interviewee to be more effective? It included a presentation and discussion of a 20-minute company-developed videotape designed to train managers on how to conduct interviews. It gave advice (e.g., put the candidate at ease. Encourage self-evaluation.), and it illustrated the proper behaviors

with actors. Also included was a lecture and discussion on what interviewees might learn from the research on interviews (e.g., importance of first impressions, heavy influence of candidate's verbal skills. For reviews see Arvey & J. Campion, 1982, or Schmitt, 1976.) Finally, evaluation factors thought to be negatively weighted in interviews were reviewed (e.g., lack of planning for career, overemphasis on money).

A four-hour program was developed because it was a typical length for in-house training and was adequate to cover the content available. Time was divided 40:40:20 among the three sections, respectively. Class sizes were limited to 20 students or less to enhance interaction.

*Self-Study (Placebo Control) Condition.* A control condition, which included a placebo, was needed because it would have been inappropriate to turn away volunteers for the program without any perceived assistance.

The self-study package consisted of copies of the overhead transparencies presented in the training and a cover memo that encouraged study and preparation. Each page contained a list of short statements or discussion points on how to take interviews. These materials avoided the potential bias of totally different content between the two conditions, yet they provided an inert or placebo treatment (Cook & Campbell, 1979) because few principles of learning were present. Aside from missing examples and explanations of the instructor and fellow students, the self-study condition also lacked active participation in discussion and role-playing practice and feedback.

The self-study condition was considered an adequate placebo for two additional reasons. First, much of the content consisted of little more than what is available in the popular press. Second, it was expected that candidates would not use the package or spend much time preparing for their interviews.

### *Measures*

Measures were developed for each level of training evaluation criteria (Kirkpatrick, 1959a, 1959b, 1960a, 1960b). Most measures were organized around the three sections of the training program.

1. *Reaction.* After classroom training, participants were asked two questions about each section: "To what extent did the (section) improve your knowledge/skills in the interview taking process?" and "To what extent do you think the (section) will improve your effectiveness in the upcoming interview(s)?" All questions utilized a five-point rating scale ranging from 1 (to a very great extent) to 5 (to no extent). On the basis of previous psychophysical research (Bass, Cascio, & O'Connor, 1974), adjectives that would enhance discriminability were selected to describe

the anchor points. Reaction data were available on 146 class participants (92.0% completion rate).

2. *Learning.* Participants were given a three-question essay test before and after the class: "What advice are you aware of on how to take interviews?", "What are some likely questions you might be asked in the upcoming interview(s)?", and "What do you know about the nature of interviews and interviewers that might be helpful in taking interviews?"

Questions were scored by counting correct answers. An answer was correct in both the before and after tests if it was consistent with some aspect of the training content. This liberal criterion allowed considerable variation in (correct) performance in the test given before the class. Scorers were blind as to whether a test was given before or after the class. Interrater reliabilities between the two scorers on a sample of 30 tests ranged from .90 to .97 ( $p < .05$ ) across the three questions of the test. Data were available on 156 class participants (98.7% completion rate).

In addition, participants were asked: "To what extent have you thought out answers to potential interview questions or otherwise prepared for the upcoming interview?" The five-point rating scale described above was utilized. This question was only asked prior to the class, and it was analyzed separately from the tests.

3. *Behavior.* Behavior measures were collected on both training and self-study groups, and in some cases on nonparticipants. Two measures were developed: interviewer-completed evaluation and candidate-completed evaluation. Because interviewers typically conducted 12 or more interviews per day and evaluations ideally should be completed immediately after each interview, a lengthy behavioral measure (cf. Latham & Wexley, 1981) could not be used. Thus, a four-question, interviewer-completed evaluation was developed. Two questions reflected the course content: "How well did the candidate appear to be prepared for the interview (e.g., good appearance, anticipated questions, well-thought-out views, etc.?)" and "How well did the candidate perform from an interview-taking skills perspective (e.g., good verbal expression, good eye contact, proper interviewing behavior, good attitude, calm, etc.?)" Interviewers were also asked: "Overall, how good was the skills match between the candidate's background and your job opening?" and "Understanding that this is not an official expression of interest, what is the likelihood that the candidate might receive a job offer?" Ratings were made on a five-point scale ranging from 1 (very good) to 5 (very poor). Evaluations were collected immediately after each interview.

Interrater reliabilities (intraclass correlations) ranged from .39 to .54 across items ( $p < .05$ ). Agreement, defined as one point or less difference on the rating scale, ranged from 67.2% to 88.9% across items ( $p < .05$ ); this method of calculating the significance of agreement is based on Lawlis

& Lu (1972). Although levels of reliability and agreement seem modest, they are considered acceptable because most candidates were interviewed for more than one job, and evaluations correlated strongly with actual job offers (see *Results* measure). Data were available on 473 interviews with study participants and 220 interviews with nonparticipants (92.2% completion rate).

The candidate-completed evaluation was collected from both training and self-study groups and consisted of six questions. Three elicited reactions to the three different sections of the class: "To what extent do you feel the training program section on (name of section) helped make you more effective in this interview?" The five-point scale of extent described above was used. Three additional questions were asked: "Overall, how well did the interview go?"; "In your opinion, what is the likelihood you might get an offer?"; and "If the job were offered, what is the likelihood you might accept?" The five-point scale ranging from very good to very poor described above was used. Interviewee-completed evaluations were mailed out after the interviews and collected before any offers were made.

Interrater reliabilities ranged from .85 to .92 for the items on the training ( $p < .05$ ) and from .23 ( $p < .10$ ) to .58 ( $p < .05$ ) for the items on the overall interview and chances of offer/acceptance. Agreement ranged from 75.7% to 98.0% across items (one point criterion;  $p < .05$ ). Again, reliability and agreement are conservative because many candidates interviewed for a number of different jobs. Data were returned from 289 interviews (71.5% completion rate).

Positive correlations exist between all items of the interviewer-completed evaluation and two items of the candidate-completed evaluation: how well the overall interview went ( $r = .21$  to  $.35$ ,  $p < .05$ ) and likelihood of an offer ( $r = .23$  to  $.42$ ,  $p < .05$ ). No correlations were observed with the items on the value of the training or with the likelihood that the candidate would accept an offer. These findings are consistent with previous research on the relationship between interviewers' and candidates' evaluations (M. Campion, 1980).

In addition to measures developed for the study, the company's official applicant evaluations were also collected when possible. These evaluations were completed by interviewers and required written comments on various factors (e.g., education/training, work experience, job interest) and a summary rating on a five-point scale from 1 (outstanding candidate) to 5 (unacceptable candidate). They were completed within a day of the interviews, but frequently not immediately after each interview. Interrater reliability was .49 ( $p < .05$ ), and agreement was 71.3% (one point criterion;  $p < .05$ ). Data were available on 289 interviews with study participants and 152 interviews with nonparticipants (58.7% completion rate).



The company's official applicant evaluations were not identifiable by interviewer's name, thus only averaged data across interviews per candidate on the company's evaluations ( $M = 1.56$  interviews per candidate,  $SD = 1.01$ ) could be compared with the averaged data on the interviewer-completed evaluations developed for this study ( $M = 2.12$  interviews per candidate,  $SD = 1.53$ ). Correlations ranged from .57 to .72 ( $p < .05$ ), but these aggregate analyses do not insure reliability at the individual level (James, 1982).

4. *Results.* Data were collected on number of jobs interviewed for and offers received, and a ratio of offers to jobs was computed. As an assessment of reliability, correlations with this ratio ranged from -.53 to -.77 with the interviewer-completed evaluation and -.65 with the company's official applicant evaluation ( $p < .05$ ). Data were available on 151 training members, 127 self-study members, and 174 nonparticipants (95.8% completion rate).

Only behavior and results criteria were collected for self-study members to avoid "testing effects" (Campbell & Stanley, 1963). Typically, self-study programs do not include externally administered evaluations or quizzes.

All instruments included statements on confidentiality and assurances that refusing to participate would not adversely affect chances of getting a job offer.

## *Results*

### *1. Reaction*

Table 1 indicates that training participants were positive on all aspects of the program. All means are significantly more favorable than the middle point on the scales. On average, 78% of the participants were favorably impressed with the program and believed that it had improved their interview-taking skills and effectiveness.

### *2. Learning*

Before attending the class, only 33.5% of participants had prepared for upcoming interviews to even a moderate extent ( $M = 2.88$ ,  $SD = .99$ ). Table 2 shows significant before-after differences on all questions of the test, with improvements ranging from 23% to 57%.

TABLE 1  
*Means, Standard Deviations, and Percentage Favorable on  
 Course Evaluations*

Item	<i>M</i>	<i>SD</i>	% Favorable
Advice—			
improved skills	1.94	.77	79
improved effectiveness	1.90	.79	79
Preparing answers—			
improved skills	1.86	.75	79
improved effectiveness	1.77	.77	84
Characteristics of interviews—			
improved skills	2.08	.90	71
improved effectiveness	1.97	.92	76

Note:  $n=146$ . Five-point rating scales, with 1 most favorable. All means are significantly smaller (i.e., more favorable) than the 3 rating or "moderate extent" anchor on the scales ( $t$  tests,  $p < .05$ ). Percentage favorable reflects proportion indicating the training improved their skills/effectiveness "to a very great extent" or "to a considerable extent."

TABLE 2  
*Means, Standard Deviations, and *t* Tests on the Class Tests*

Item	Before test			After test			<i>t</i>
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	
Advice	135	3.90	2.37	127	4.85	2.59	6.63
Likely questions	156	3.22	1.81	141	5.05	2.87	3.60
Nature of interviews	146	2.62	1.88	136	3.49	2.18	3.09

Note: All  $t$  tests significant at  $p < .05$ .

### 3. Behavior

Strong intercorrelations exist among items of the interviewer-completed evaluation (Table 3), and standard deviations reveal no apparent range restriction (Table 4). Comparisons between training and self-study groups show no significant differences on any of the items, and the multivariate analysis of variance (MANOVA) is also nonsignificant (Wilks Lambda = .99,  $F = .93$ ,  $df = 4, 437$ , n.s.). A significant MANOVA is observed between study participants and nonparticipants (Wilks Lambda = .98,  $F = 2.92$ ,  $df = 4, 646$ ,  $p < .05$ ). Individual  $t$  tests are not significant (Table 5), however, indicating that the ratings, by themselves, did not distinguish between groups.

Intercorrelations among the items of the candidate-completed evaluation range from .10 to .82, with the highest intercorrelations among the items on the training program (Table 6). Standard deviations reveal no apparent range restriction (Table 7). The MANOVA is significant (Wilks Lambda = .89,  $F = 4.57$ ,  $df = 6, 229$ ,  $p < .05$ ), and differences in favor of the training group exist on items reflecting evaluations of the training but not on items evaluating the overall interview and chances of offer/acceptance.

TABLE 3

*Intercorrelations Among the Items of the Interviewer-completed Evaluation*

	Performed in interview	Skills match	Likelihood of offer
Prepared for interview	.72	.56	.54
Performed in interview		.51	.54
Skills Match			.78

Note:  $n$ 's = 653 to 691. All correlations significant at  $p < .05$ .

TABLE 4

*Means, Standard Deviations, and  $t$  Tests on the Interviewer-completed Evaluation*

Item	Training group			Self-Study group			$t$
	$n$	$M$	$SD$	$n$	$M$	$SD$	
Prepared for interview	252	2.14	.89	221	2.04	.86	-1.31
Performed in interview	251	2.14	.94	222	2.01	.85	-1.57
Skills match	244	2.90	1.15	214	2.76	1.34	-1.33
Likelihood of offer	242	3.25	1.45	212	3.00	1.45	-1.84

Note: All  $t$  tests nonsignificant at  $p < .05$ . Five-point rating scales, with 1 most favorable.

TABLE 5

*Means, Standard Deviations, and  $t$  Tests on the Interviewer-completed Evaluation Comparing Study Participants and Nonparticipants*

Item	Participants			Nonparticipants			$t$
	$n$	$M$	$SD$	$n$	$M$	$SD$	
Prepared for interview	473	2.09	.88	220	2.08	.81	-.16
Performed in interview	473	2.08	.90	219	2.16	.85	1.19
Skills match	458	2.84	1.15	215	2.73	1.16	-1.13
Likelihood of offer	454	3.13	1.46	213	3.25	1.47	.98

Note: All  $t$  tests nonsignificant at  $p < .05$ . Five-point rating scales, with 1 most favorable.

Analysis of effect size shows 11% of variance accounted for in the dependent variables (Maxwell, Camp, & Arvey, 1981).

The company's official applicant evaluation also shows no obvious range restriction ( $M = 3.32$ ,  $SD = 1.28$ ). As with the other interviewer-completed evaluation, no significant difference is observed between training and self-study groups ( $M = 3.35$  and  $SD = 1.32$  versus  $M = 3.24$  and  $SD = 1.23$ , respectively;  $t = -1.15$ , n.s.) and between study participants and nonparticipants ( $M = 3.30$  and  $SD = 1.28$  versus  $M = 3.34$  and  $SD = 1.29$ , respectively;  $t = 1.03$ , n.s.).

#### 4. Results

Both the number of jobs interviewed for and offers per candidate range from one to seven ( $M = 1.82$  and  $SD = 1.16$  versus  $M = .77$  and  $SD =$

TABLE 6

*Intercorrelations among the Items of the Candidate-completed Evaluation*

	Preparing answers	Characteristics of interviews	Overall evaluation	Likelihood of offer	Likelihood of accept
Advice	.75	.82	.22	.18	.15
Preparing answers		.73	.21	.22	.10(a)
Characteristics of interviews			.22	.21	.13
Overall evaluation				.67	.25
Likelihood of offer					.39

Note: *n*'s = 252 to 280. All correlations significant at  $p < .05$ , except (a).

TABLE 7

*Means, Standard Deviations, and t Tests on the Candidate-completed Evaluation*

Item	Training group			Self-Study group			<i>t</i>
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	
Advice	165	2.30	.87	114	2.76	.99	4.12*
Preparing answers	168	2.13	.94	115	2.79	1.12	5.35*
Characteristics of interviews	163	2.39	.85	114	2.79	1.01	3.51*
Overall evaluation	164	1.85	.77	121	1.84	.89	-.05
Likelihood of offer	148	2.32	1.07	116	2.19	.99	-.99
Likelihood of accept	145	2.08	1.07	111	2.15	1.11	.56

Note: Five-point rating scales, with 1 most favorable.

\* $p < .05$

1.02, respectively). The ratio of offers to jobs averages approximately one-third ( $M = .37$ ), and there is wide variation among candidates ( $SD = .43$ ). The comparison between training and self-study groups is nonsignificant ( $M = .37$  and  $SD = .42$  versus  $M = .38$  and  $SD = .42$ , respectively;  $t = .30$ , n.s.). The same occurs when study participants are compared to nonparticipants ( $M = .37$  and  $SD = .42$  versus  $M = .37$  and  $SD = .44$ , respectively;  $t = -.02$ , n.s.).

### Discussion

The purpose of this study was to evaluate the effectiveness of an interviewee skills training program. An experimental design was employed with a self-study (placebo control) group and random assignment. A field setting was used with a large and broad sample of interviewers, jobs, and candidates. Multiple measures were utilized, which exhibited acceptable interrater reliability and agreement, as well as range and variation. Statistical power was high for all analyses. The training program derived its content from previous literature and used multiple instructional techniques. Finally, the program was evaluated on four levels of training criteria ranging from participant reactions to job offers.

Results indicated that class members responded positively to the content of the training. They felt the program improved their interview-taking skills and would enhance their effectiveness. Class members also demonstrated substantial learning of the training content, and they were more positive than the self-study group on a follow-up evaluation of interviewing behavior.

Actual interviewer evaluations failed to distinguish between training and self-study groups, however, and no differences were observed in job offers. In addition, no differences existed between study participants and nonparticipants on these criteria. In different terms, this research replicates earlier work in the management training domain by finding positive training effects on subjective and objective learning measures, as well as what Burke and Day (1986) refer to as subjective behavior measures. But these positive effects were not found for more objective behavior measures or for organizational results. This study further illustrates the often noted danger of relying solely on reaction or learning measures in the evaluation of training programs (e.g., Goldstein, 1986; Kirkpatrick, 1959a; McGehee & Thayer, 1961; Wexley & Latham, 1981).

Several potential explanations can be offered for the lack of significance on some of the criteria. For example, one difference between this research and earlier work using behavioral measures is the context of evaluation. Most previous work employed passive observers evaluating videotaped interviews in laboratory settings. In this study, evaluations were collected from the interviewers themselves following actual interviews. The laboratory setting, by enhancing experimental control, may be able to detect training differences that are imperceptible to more active interviewers in face-to-face interviews. The interview-conducting (e.g., rapport building, questioning) and information-collecting (e.g., listening, note taking) responsibilities of real interviewers may detract from their ability to detect subtle differences in interview-taking behavior. This potential for method variance needs to be examined, for it has implications for the generalizability of laboratory findings to actual interview settings.

A potential explanation for the nonsignificant findings on job offers concerns the availability of candidate information prior to the interview. Because candidates were employees of the same organization, considerable information was available before the actual interviews (e.g., work history, performance appraisals, some limited aptitude test data). It has long been known that first impressions are potent predictors of final interview outcomes (Springbett, 1954, cited in Webster, 1982). More recent research suggests that this sort of previewing influences both data collection and evaluation in the interview (Dipboye, Fontenelle, & Garner, 1984; Tucker & Rowe, 1977), and it has been argued that pre-interview evaluations are

self-fulfilling (Dipboye, 1982). In a direct recent comparison, resumé credentials had a greater impact on evaluations than either verbal or nonverbal behavior in a laboratory setting (Rasmussen, 1984).

Concern with availability of prior information is enhanced when candidates in this study are contrasted with those of most other studies in the area. In this study, candidates were experienced employees, and many were applying for professional or technical jobs where skills match may be much more important than performance in the interview. In studies with students, interview performance may be more important because the students are so similar in experience and background. With special populations (e.g., psychiatric patients), interpersonal behavior problems obvious in the interview may far outweigh any work experience for the usually unskilled jobs being filled.

But students and other new entrants into the work force compose only an eighth of all the unemployed (Current Labor Statistics, 1986) and none of the equally large group of employed persons looking for new jobs (Rosenfeld, 1977). Thus, subjects in this study may be more similar to candidates in the majority of interview situations; that is, experienced workers pursuing jobs or job changes. Furthermore, the setting is similar to many interviewing situations. Prescreening is based on credentials, followed by multiple interviews, and offer decisions are made by the management team on the basis all information (both credentials and interviews). That decisions were not predetermined by credentials alone is evidenced by the fact that candidates received offers from only one-third of the positions for which they were interviewed.

Another explanation for the lack of differences on the behavior and results criteria is that the four-hour training program, although typical for industry and adequate to cover the material, may have had an impact that was not sufficient to influence actual interviewing behavior. Alternatively, the interview-taking skills of the candidates may have been high to begin with, so making improvements was more difficult. Either way, it is possible that a longer and more individualized training program may be more effective.

A related concern is that the reading materials in the self-study condition provided too much information. But this concern is negated by the likelihood that the materials received little study from recipients, most of this information is available in the popular press, and the nonparticipants in the study performed at the same level as the participants on the behavior and results criteria.

A final potential explanation relates to the literature itself. Interviewee skills training may not be as useful for industry populations as with special populations (e.g., rehabilitation clients, psychiatric patients) or students. The advice may be so commonsense and so publicly available that a special

program to teach it is simply not useful to an industry population. The findings of the present study do not contradict any existing strong evidence that this training has a positive effect with samples from industry.

In conclusion, researchers and practitioners may need to consider setting and subject variables more carefully. The present findings are most apt to generalize to field settings where experienced job candidates are being recruited. Under these conditions, a four-hour training program is unlikely to give the job seeker an advantage. Replications with other jobs and job candidates, as well as extensions utilizing more intensive training, are necessary before broader conclusions can be drawn.

#### REFERENCES

- Arvey RD, Campion JE. (1982). The employment interview: A summary and review of recent research. *PERSONNEL PSYCHOLOGY*, 35, 281-322.
- Austin MF, Grant TN. (1981). Interview training for college students disadvantaged in the labor market: Comparison of five instructional techniques. *Journal of Counseling Psychology*, 28, 72-75.
- Barbee JR, Keil EC. (1973). Experimental techniques of job interview training for the disadvantaged: Videotape feedback, behavior modification, and microcounseling. *Journal of Applied Psychology*, 58, 209-213.
- Bass BM, Cascio WF, O'Connor EJ. (1974). Magnitude estimations of expressions of frequency and amount. *Journal of Applied Psychology*, 59, 313-320.
- Bolles RN. (1982). *What color is your parachute? A practical manual for job-hunters and career-changers*. Berkeley, CA: Ten Speed Press.
- Braukmann CJ, Fixsen DL, Phillips EL, Wolf MM, Maloney DM. (1974). An analysis of a selection interview training package for predelinquents at Achievement Place. *Criminal Justice and Behavior*, 1, 30-42.
- Burke MJ, Day RR. (1986). A cumulative study of the effectiveness of managerial training. *Journal of Applied Psychology*, 71, 232-245.
- Campbell DT, Stanley JC. (1963). *Experimental and quasi-experimental designs for research*. Chicago: Rand-McNally.
- Campion MA. (1980). Relationship between interviewers' and applicants' reciprocal evaluations. *Psychological Reports*, 47, 1335-1338.
- Cohen J. (1977). *Statistical power analysis for the behavioral sciences* (rev. ed). New York: Academic Press.
- Cook TD, Campbell DT. (1979). *Quasi-experimentation: Design & analysis issues for field settings*. Chicago: Rand-McNally.
- Current labor statistics. (1986, November). *Monthly Labor Review*, 109(11), 74.
- Dipboye RL. (1982). Self-fulfilling prophecies in the selection-recruitment interview. *Academy of Management Review*, 7, 579-586.
- Dipboye RL, Fontenelle GA, Garner K. (1984). Effects of previewing the application on interview process and outcomes. *Journal of Applied Psychology*, 69, 118-128.
- Furman W, Geller M, Simon SJ, Kelly JA. (1979). The use of a behavior rehearsal procedure for teaching job-interviewing skills to psychiatric patients. *Behavior Therapy*, 10, 157-167.
- Goldstein IL. (1986). *Training in organizations: Needs assessment, development, and evaluation* (2nd ed.). Monterey, CA: Brooks/Cole.

- Harrison RP, Horan JJ, Torretti W, Gamble K, Terzella J, Weir E. (1983). Separate and combined effects of a cognitive map and a symbolic code in the learning of a modeled social skill (job interviewing). *Journal of Counseling Psychology, 30*, 499-505.
- Heimberg RG, Stanley J, Blankenberg R. (1982). Preparing unemployed youth for job interviews: A controlled evaluation of social skills training. *Behavior Modification, 6*, 299-322.
- Hollandsworth JG, Dressel ME, Stevens J. (1977). Use of behavioral versus traditional procedures for increasing job interview skills. *Journal of Counseling Psychology, 24*, 503-510.
- Hollandsworth JG, Glazeski RC, Dressel ME. (1978). Use of social skills training in the treatment of extreme anxiety and deficient verbal skills in the job interview setting. *Journal of Applied Behavior Analysis, 11*, 259-269.
- Hood EM, Lindsay WR, Brooks N. (1982). Interview training with adolescents: A controlled group study incorporating generalization and social validation. *Behaviour Research and Therapy, 20*, 581-592.
- Jackson GG. (1972). The use of role-playing job interviews with Job Corps females. *Journal of Employment Counseling, 9*(3), 130-139.
- James LR. (1982). Aggregation bias in estimates of perceptual agreement. *Journal of Applied Psychology, 67*, 219-229.
- Keil EC, Barbee JR. (1973). Behavior modification and training the disadvantaged job interviewee. *Vocational Guidance Quarterly, 22*, 50-56.
- Keith RD, Engelkes JR, Winborn BB. (1977). Employment-seeking preparation and activity: An experimental job-placement training model for rehabilitation clients. *Rehabilitation Counseling Bulletin, 21*, 159-165.
- Kelly JA, Laughlin C, Claiborne M, Patterson J. (1979). A group procedure for teaching job interviewing skills to formerly hospitalized psychiatric patients. *Behavior Therapy, 10*, 299-310.
- Kelly JA, Urey JR, Patterson J. (1981). Small group job interview skills training in the mental health care setting. *Behavioral Counseling Quarterly, 1*, 202-212.
- Kirkpatrick DT. (1959a). Techniques for evaluating training programs. *American Society of Training Directors Journal, 13*(11), 3-9.
- Kirkpatrick DT. (1959b). Techniques for evaluating training programs. Part 2—learning. *American Society of Training Directors Journal, 13*(12), 21-26.
- Kirkpatrick DT. (1960a). Techniques for evaluating training programs. Part 3—behavior. *American Society of Training Directors Journal, 14*(1), 13-18.
- Kirkpatrick DT. (1960b). Techniques for evaluating training programs. Part 4—results. *American Society of Training Directors Journal, 14*(2), 28-32.
- Latham GP, Saari LM. (1979). Applications of social learning theory to training supervisors through behavioral modeling. *Journal of Applied Psychology, 64*, 239-246.
- Latham GP, Wexley KN. (1981). *Increasing productivity through performance appraisal*. Reading MA: Addison-Wesley.
- Lawlis GF, Lu E. (1972). Judgment of counseling process: Reliability, agreement, and error. *Psychological Bulletin, 78*, 17-20.
- Logue PE, Zenner M, Gohman G. (1968). Video-tape role playing in the job interview. *Journal of Counseling Psychology, 15*, 436-438.
- Maxwell SE, Camp CJ, Arvey RD. (1981). Measures of strength of association: A comparative examination. *Journal of Applied Psychology, 66*, 525-534.
- McGehee W, Thayer PW. (1961). *Training in business and industry*. New York: Wiley.
- Medley HA. (1984). *Sweaty palms: The neglected art of being interviewed*. Berkeley, CA: Ten Speed Press.
- Pettus T. (1981). *One on one: Win the interview, win the job*. New York: Random House.



- Pinto RP. (1979). An evaluation of job interview training in a rehabilitation setting. *Journal of Rehabilitation*, 45, 71-76.
- Rasmussen KG. (1984). Nonverbal behavior, verbal behavior, resumé credentials, and selection interview outcomes. *Journal of Applied Psychology*, 69, 551-556.
- Robertson J. (1978). *How to win in a job interview*. Englewood Cliffs, NJ: Prentice-Hall.
- Rosenfeld C. (1977). The extent of job search by employed workers. *Monthly Labor Review*, 100(3), 58-62.
- Schwab DP, Rynes SL, Aldag RJ. (1987). Theories and research on job search and choice. In Roland KM, Ferris GR (Eds.), *Research in personnel and human resources management* (Vol. 5, pp. 129-166). Greenwich, CT: JAI Press.
- Seigfried WD, Wood K. (1983). Reducing college student's compliance with inappropriate interviewer requests: An educational approach. *Journal of College Student Personnel*, 24, 66-71.
- Schmitt N. (1976). Social and situational determinants of interview decisions: Implications for the employment interview. *PERSONNEL PSYCHOLOGY*, 29, 79-101.
- Solomon J. (1986, May 22). Career-change services spread, offering advice on where and how to job hunt. *The Wall Street Journal*, p. 31.
- Speas CM. (1979). Job-seeking interview skills training: A comparison of four instructional techniques. *Journal of Counseling Psychology*, 26, 405-412.
- Stevens W, Tornatzky L. (1976). The effects of job interview skills workshop on drug-abuse clients. *Journal of Employment Counseling*, 13, 156-163.
- Tucker DH, Rowe PM. (1977). Consulting the application form prior to the interview: An essential step in the selection process. *Journal of Applied Psychology*, 62, 283-287.
- Venardos MG, Harris MB. (1973). Job interview training with rehabilitation clients: A comparison of videotape and role-playing procedures. *Journal of Applied Psychology*, 58, 365-367.
- Webster EC. (1982). *The employment interview: A social judgment process*. Ontario, Canada: S.I.P. Publications.
- Westbrook BW. (1985). Review of Strong-Campbell Interest Inventory. In Mitchell JV (Ed.), *The ninth mental measurements yearbook* (Vol. 2). Lincoln, NE: University of Nebraska Press.
- Wexley KN, Latham GP. (1981). *Developing and training human resources in organizations*. Glenview, IL: Scott, Foresman.
- Wild BK, Kerr BA. (1984). Training adolescent job-seekers in persuasion skills. *Vocational Guidance Quarterly*, 33, 63-69.