RECIPROCAL PEER MANAGEMENT: IMPROVING STAFF INSTRUCTION IN A VOCATIONAL TRAINING PROGRAM

RICHARD FLEMING
AUBURN UNIVERSITY

AND

BETH SULZER-AZAROFF
UNIVERSITY OF MASSACHUSETTS

To test the feasibility and utility of involving peers as sources of feedback, 6 subjects, instructors in a vocational program for adults with mental retardation, participated in a staff training and management program. Subjects' teaching interactions were assessed during baseline, in-service training (on effective teaching), return-to-baseline, peer management, and follow-up phases. Peer management was introduced in multiple baseline fashion across pairs of subjects. Members of each pair were trained to monitor peer teaching, to record and graph data, to provide feedback, and to set goals with the peer. Each pair then performed these procedures on the job for several weeks, during which time 4 of the 6 subjects increased their use of effective teaching methods (over baseline). However, inconsistencies in the magnitude and durability of these increases require that the study be viewed as inconclusive, although it has heuristic value as a promising model for involving co-workers in staff management programs.

DESCRIPTORS: staff management, peer support, vocational instruction, mental retardation, feedback sources

In human services, feedback, typically conveyed by researchers or supervisors in top-down fashion (e.g., Ivancic, Reid, Iwata, Faw, & Page, 1981; Page, Iwata, & Reid, 1982), has been used extensively to improve staff performance (see Reid, Parsons, & Green, 1989, and Repp, Felce, & de Kock, 1987, for reviews). Recently, procedures that increase staff participation, primarily through self-managed monitoring and feedback (e.g., Burgio, Whitman, & Reid, 1983; Richman, Riordan, Reiss, Pyles, & Bailey, 1988), have been studied. However, to date, enabling co-workers to participate in managing contingencies has received scant attention (Egan, Luce, & Hall, 1988; van den Pol, Reid, & Fuqua, 1983).

Peer groups may represent a natural "community" of social reinforcement (Kohler & Greenwood, 1986) that could be harnessed to support the generality of newly acquired work skills. Potential practical advantages also exist. Peers often work in close physical proximity with one another, enabling them to give frequent and immediate feedback and to model desirable behaviors (Sulzer-Azaroff, Fox, Moss, & Davis, 1987). Peer interventions might also be less costly to implement and perhaps more acceptable to staff (Greller, 1980) than those mediated by supervisors. And finally, involving peers in the design of work contingencies that affect them directly may in itself enhance performance (Cotton, Vollrath, Froggatt, Lengnick-Hall, & Jennings, 1988).

Prior research has demonstrated the viability of peer support systems. Paraprofessional staff members have been taught successfully to use instructions, practice, and feedback to train their peers to perform skills related to client safety (van den Pol et al., 1983). Peers have also participated in an internal review system in which written feedback from overnight staff members functioned to improve the quality of client records kept by daytime staff (Egan et al., 1988). Although these two stud-

We thank the administrators and staff of the host facility for their cooperation and support and acknowledge Lorraine Fleury and Cindy Kirshen Fleming for their assistance with data collection and analysis.

Correspondence and requests for reprints should be sent to Richard K. Fleming, Department of Psychology, Auburn University, Auburn, Alabama 36849-5214.

1992, 25, 611-620

ies supported the potential for peer training and management, considerably more research is needed on programs designed to improve and maintain staff's use of more routine forms of service delivery to clients, such as teaching vocational or other functional skills.

We evaluated the effectiveness and acceptability of a peer contingency management program with vocational instructors working with mentally retarded adults. Comprised of observing and monitoring, graphing, delivering praise and constructive feedback, and setting goals, the program was designed to increase and maintain the subjects' rates of high-quality instructional interactions with their client trainees.

METHOD

Subjects and Setting

Six instructors affiliated with a vocational training program based at a large state residential facility for mentally retarded adults participated as subjects. The mean age of the instructors was 32 years (range, 26 to 35 years), and their average length of employment as instructors was 2 years 8 months (range, 2 years to 3 years 8 months). Four held college degrees in human service related areas; 2 held high school degrees. Subjects participated in pairs, with two sets working in one program area and the third in a second area (hereafter referred to as Programs 1 and 2, respectively).

Program 1 was a prevocational program in which clients with profound or severe mental retardation completed sorting and assembly tasks. Pair 1 worked in one of two large, adjacent rooms, and Pair 3 worked in the other. Each subject taught a group of approximately four clients in a work area (5 m by 5 m). Among the 16 clients, one was blind, four had ambulation problems, four exhibited aggression, four often injured themselves, and one was dually diagnosed with a bipolar affective disorder.

Program 2 was a vocational program located in a gymnasium-sized room. Each staff member in this program, including the subjects in Pair 2, trained groups of six moderately to severely mentally retarded clients who processed contract work involving two- to four-piece assemblies and sorting. Of the 12 clients taught by subjects in Pair 2, one was hearing impaired, three were nonambulatory, and four exhibited aggression.

In each program, the trainer (first author) met with the instructors and their immediate supervisor, described the project, and asked for volunteers. All 4 instructors in Program 1 agreed to participate. Because of practical limitations (e.g., availability of equipment and research assistants), the trainer asked the supervisor of Program 2 to suggest only 4 of the 7 instructors in the program who might benefit most from participating. All 4 who were identified agreed to participate, but the members of one pair subsequently withdrew at the start of the peer management phase, explaining that they felt "threatened" by the prospect of giving evaluative feedback to one another. All subjects read and signed an informed consent form.

Prior to beginning observations, the 4 subjects in Program 1 chose partners, basing their selection on the physical proximity of their work areas, existing supportive relationships, and common work schedules. The subjects in Program 2 maintained pairings to which their supervisor had assigned them.

Equipment and Materials

A videocamera (Panasonic Model AG-100), videotapes, a wireless microphone (Realistic Model 32-1226), a television set (Magnavox Model CF-4137WA02), a videocassette recorder (General Electric Model VG-7510), a cassette tape recorder/player (Panasonic Model RQ-356), written instructions on effective client teaching and peer management methods, and forms for recording and graphing observational data and writing feedback and goal statements were used.

Dependent Variables

Teaching Interactions. To address the wide range of instructional needs of clients in the subjects' groups (from intensive shaping of subskills with some clients to intermittent reinforcement of independent task completion with others), two types of instructional interactions were selected for train-
ing and measurement purposes. The first, a complete teaching interaction, was defined as a sequence consisting of prompting (verbal instruction, gesturing, modeling, and partial or full physical guidance) to a level necessary to produce a complete client response, or series of responses, and then providing a rewarding consequence (see below). The second, a rewarding consequence, was defined as providing praise, task-related feedback, a pat, hug, tangible or edible item, or any combination of these consequences contingent upon a client’s active engagement with work materials or within 3 s of completion of a task-related response. Although by definition the latter measure was a subset of the former, the two types of interactions were functionally distinct in that complete teaching interactions were used to train new skills, whereas rewarding consequences were also delivered independent of teaching interactions to maintain clients’ use of previously acquired skills. It should be noted that in scoring these two response classes, a rewarding consequence delivered as part of a complete teaching interaction was never counted again as a rewarding consequence.

Client behavior. In Program 2, clients of varying abilities were paid on a piece-rate basis for stuffing boxes with tubes of glue and spare caps. On an hourly basis, trainers counted and recorded the number of pieces their clients completed. These data were used to calculate the average hourly production rates per week for each of the groups. Because clients were frequently assigned to different work groups and thus were taught by different subjects, client outcome data were not assessed in Program 1.

Measurement and Interrater Agreement

Subjects were videotaped for 5 min within the same 30-min period each day by one of two research assistants. Prior to filming, the assistant handed the subject a clip-on wireless microphone. If demands for client care required subjects to leave their groups during filming, recording stopped; it then resumed when the subject returned. Filming was canceled on several occasions because of a variety of unforeseen circumstances, including extreme disruption of a client, equipment failure, or illness of a subject.

Research assistants were trained to record accurately the frequencies of each of the two types of staff teaching interactions. To obtain a conservative measure of interrater agreement, the 5-min sample was partitioned into 15 20-s intervals. A prerecorded audiocassette signaled when to begin observing and when to record for each interval. Response frequencies were scored and agreements and disagreements assessed on an interval-by-interval basis by the trainer, or in some cases by another assistant, on 78 (25.2%) of the 309 observations. For each observation, agreement scores were calculated by dividing the number of agreements by the number of agreements plus disagreements and then multiplying by 100. For complete teaching interactions, indices of overall agreement averaged 91.4% (range, 66.7% to 100%), occurrence agreement averaged 80.7% (range, 50.0% to 100%), and nonoccurrence agreement averaged 91.7% (range, 50.0% to 100%). For rewarding consequences, indices of overall agreement averaged 88.7% (range, 73.3% to 100%), occurrence agreement averaged 79.8% (range, 0% to 100%), and nonoccurrence agreement averaged 87.9% (range, 20.0% to 100%). Low scores tended to occur when only one or two instances of a response occurred during a 5-min observation and the raters disagreed. To maintain scoring accuracy, if indices of agreement fell below 80% on two consecutive tapes, the trainer retrained the research assistants; however, he never changed their scores.

Experimental Design and Procedures

The peer management package was introduced in multiple baseline fashion across the three pairs of subjects following nonconcurrent baselines of varying lengths.

Baseline. Subjects were filmed daily until the data were observed to be relatively stable. Stability was assessed by applying the rule that, before ending a phase, the final three data points contained no new high or low scores. Next, to ensure the subjects’ instructional competence, the trainer conducted in-service training on the same days as Ses-
sessions 22, 23, and 24 for Pair 1; Sessions 16, 17, and 18, with a refresher on the day of Session 25, for Pair 2; and Sessions 17, 19, and 20 for Pair 3. During two 1-hr sessions, conducted in a small conference room over the first 2 days, pairs of subjects were provided, orally and in writing, a rationale, definitional criteria, and guidelines for the use of the targeted teaching skills. Subjects were also required to list, for each of their clients, one or more specific behaviors they planned to train or reinforce and prompting sequences and rewarding consequences they knew or suspected would be effective. On the 3rd day, instruction was provided on the job in sessions lasting approximately 45 min each. The trainer (a) demonstrated correct use of the teaching behaviors directly with several of the subjects’ clients; (b) coached the subjects as they practiced the skills themselves, providing rewarding and corrective feedback; and (c) had each subject demonstrate, unassisted, the correct delivery of at least two complete teaching interactions with each of four clients in a 5-min period, and the dispensing of at least eight rewarding consequences distributed across clients during a separate 5-min period. Finally, subjects were asked to “use the two skills we covered—complete teaching and rewarding consequences—as correctly and as frequently and consistently as you can.” In total, training for each pair lasted an average of 2 hr and 45 min.

Peer management. In preparation for their roles as peer managers, the trainer taught the subjects in pairs to observe and record, graph data, provide performance feedback, and set goals based on their previous performance. Training sessions lasting 30 to 60 min were held on the same days as Sessions 37, 40, and 44 for Pair 1; Sessions 41, 43, and 44 for Pair 2; and Sessions 65, 66, and 67 for Pair 3. Written materials, adapted from well-known sources in organizational behavior management (Daniels & Rosen, 1986; Fellner & Sulzer-Azaroff, 1984; Prue & Fairbank, 1981), were reviewed, providing a rationale, definitions, and criteria for the effective use of each skill.

Training exercises were also conducted. Under the trainer’s guidance, subjects were required to write and read aloud at least three feedback comments describing recently observed, exemplary aspects of their partner’s teaching. Subjects evaluated their feedback against a set of quality criteria and reworded their comments, if necessary, until they met the criteria. Subjects were also provided with personalized graphs on which they practiced plotting data and setting goals until they could perform these skills correctly. Four baseline data points, derived from the trainer’s data, represented a weekly average of that subject’s combined use of the two instructional skills from the previous 4 weeks. Although the trainer presented these data to subjects, he subsequently never showed them his own graphs or data summaries and did not deliver any information of this sort again.

Next, subjects were taught to observe and record frequencies of one another’s use of the instructional skills on the job. As subjects watched one another teach, the trainer taught them to (a) label occurrences of each response accurately, (b) time the 5-min observations using a stop-watch, (c) record the total number of occurrences of each type of interaction on a form, (d) write at least two feedback statements and one suggestion on the form, and (e) hand the completed form immediately to their partner. Peer observations and recording were separate from and independent of the intervention observations. Finally, the pairs were taught to conduct weekly review meetings in which they summarized feedback from their earlier observations, used data they had recorded to complete their graphs, noted some aspect of their teaching they planned to improve, and set a quantitative goal for increasing or maintaining their rates of teaching during the coming week.

During each succeeding week, each pair conducted two 5-min peer observations, after which they exchanged written feedback and suggestions and held review meetings approximately 15 min long. Out of concern that the trainer’s presence could function as a discriminative stimulus for improved teaching, peer observations and review meetings were not monitored. To help ensure that subjects implemented the procedures correctly, the
trainer collected and reviewed their completed forms and graphs at the end of each week and sent them written feedback on the morning of the next working day. However, to ensure that this feedback did not contaminate the effects of the peer contingencies on teaching performance, the trainer limited his comments to those that presumably would affect only the subjects’ correct and consistent use of the peer management procedures. Representative examples included: “Your feedback to (name of partner) was constructive, specific, and rewarding. For example, you commented specifically on the quality of her teaching interactions with several clients. “ “Your selection of a goal followed the criteria we discussed for effective goal-setting, being challenging but within reach.”

In an effort to reduce the time subjects spent performing the peer management procedures, 5 weeks after the onset of the peer management phase the trainer instructed each pair to conduct only one observation per week, provide fewer feedback statements (but at least one) and suggestions, and shorten review meetings, if they so desired. Graphing and goal setting continued as before. The trainer also withdrew his weekly feedback. These instructions and changes were presented on the same days as Session 60 for Pair 1, Session 68 for Pair 2, and Session 89 for Pair 3.

Follow-up measurement and return to baseline. One follow-up observation was made after 1 month, and three consecutive observations were made after 2 months. Subjects were contacted a day or two prior to these observations to reconfirm their permission to be filmed.

Consumer satisfaction survey. Following termination of the peer management phase, subjects completed privately and independently a survey on which they rated their satisfaction with the various components of the program. Several items concerning in-service training and the peer management package were rated on a 5-point Likert-type scale on which scores of 1 indicated “strong disagreement,” 3 indicated “no difference,” and 5 indicated “strong agreement.” Subjects were also asked to select which, if any, components of the peer management package they would like to continue to use. Finally, subjects were asked to rate their preferences for three sources of feedback—supervisor, peer, and a combination of peer and supervisor. A rating of 1 indicated the “least preferred” source, 2 indicated “acceptable,” and 3 indicated “most preferred.”

RESULTS

Teaching Interactions

It was anticipated that, over the course of the study, the subjects’ delivery of complete teaching interactions would decrease and rewarding consequences would increase as clients acquired skills and came to use them more independently. However, this pattern failed to emerge. Except for Julie, who primarily used rewarding consequences because her clients performed more independently than those in the other groups, all subjects continued to distribute their use of both types of instructional interactions over the course of the study. Thus, for simplicity, the two measures were combined to yield an overall index of teaching.

Figure 1 presents the frequency of all teaching interactions for the 6 subjects across baseline, peer management, and follow-up phases. Means across the respective phases (excluding the 1-month follow-up observation) were 14.8, 23.5, and 20.0 for Deborah; 17.5, 20.3, and 19.7 for Susan; 2.9, 9.0, and 9.7 for Julie; 8.2, 12.6, and 17.7 for Jason; 8.6, 14.2, 18.7 for Barbara; and 9.4, 16.2, and 12.0 for Ashley.

Although most subjects’ baseline performance was variable, all subjects except Susan and Jason established a consistent level and range of variability by the end of baseline. During the peer management phase, rates of teaching increased substantially for Deborah and Julie, but to a lesser extent for Barbara and Ashley. Rates for Susan and Jason were maintained approximately at the levels observed following in-service training. Follow-up measurement indicated that Susan’s, Jason’s, and Barbara’s rates of teaching approximated their rates during the peer management phase, whereas Debo-
Figure 1. Frequency of teaching interactions for all subjects and pairs. Days on which training of teaching and peer management skills occurred are marked by arrows. A 4-week break in data collection occurred during baseline in Pair 3 when Ashley sustained an injury. Upon her return, a 1-hr refresher course (arrow) was given. Two 2-week vacation breaks also occurred in Pair 3. Weekly probes are presented for each subject in that pair during her partner’s absence. (FU, follow-up.)
Work Productivity of Students in Pair 2

Figure 2 presents work productivity rates for the groups of clients instructed by Julie and Jason. Corresponding rates of the subjects' teaching are presented as weekly averages. After the subjects had exchanged feedback for several weeks, their clients began to increase their rates of production, a trend that occurred up until the time subjects were instructed to reduce their peer management efforts.

Consumer Satisfaction

Subjects' mean ratings indicated that (a) both in-service training ($M = 4.75$) and peer management ($M = 4.33$) helped them improve their performance, and (b) observing ($M = 4.33$) and receiving feedback ($M = 4.83$) were more helpful than goal setting ($M = 4.17$) and graphing ($M = 3.83$). When asked what components of the peer
man­age­ment pro­gram they might like to con­tin­ue to use, all 6 sub­jects noted observ­ing and ex­chang­ing feed­back, only 1 sub­ject noted goal set­ting, and 4 sub­jects noted the end-of-the-week re­view ses­sions. Fi­nally, sub­jects’ mean ra­tions of pre­ference for source of feed­back were (a) im­me­di­ate super­visor ($M = 1.83$), (b) peer ($M = 2.83$), and (c) com­bi­na­tion of super­visor and peer ($M = 1.60$).

DISCUSSION

Al­though our re­sults are es­sen­tially in­con­clu­sive, the cur­rent study holds heu­ristic value as a prom­ising model for in­volv­ing para­profes­sional peers as re­ci­procal man­agers of one an­other’s teach­ing per­for­mance. Sev­er­al sub­jects in­creased their in­struc­tional in­ter­ac­tions over ba­line during the pe­er man­age­ment phase. How­ever, lim­i­ta­tions of the pre­sent study, such as in­con­sis­tencies in the mag­nitude and du­ra­bil­ity of change across sub­jects, and sev­er­al method­o­log­i­cal shortcom­ings sug­gest the need for fur­ther re­search in this area.

De­borah and Julie clear­ly in­creased their teach­ing dur­ing pe­er man­age­ment, but other sub­jects’ im­pro­vem­ents were more vari­able. Al­though the study was not de­signed to as­sess sys­tem­at­i­cally the ef­fects of in-service train­ing on teach­ing per­for­mance, Su­san and Jason ap­peared to have bene­fited from this in­i­tial train­ing. Per­haps in-service train­ing, shown con­sis­tently to be in­ad­e­quate with di­rect-service staff (Reid et al., 1989), may be more ef­fective with in­struc­tional staff (e.g., Lang­one, Ko­or­land, & Oso­roff, 1987). This may be es­pe­cial­ly true in the pre­sent case, where staff mem­bers had been hired based on their pri­or ed­u­ca­tion, teach­ing ex­pe­ri­ence, and ex­pressed de­sire to teach. Fi­nally, in­creases in Bar­bara’s and Ash­ley’s per­for­mances were of rel­a­tively low mag­nitude. If sus­tained through­out the day, such changes would re­sult, of course, in sub­stan­tially more teach­ing. How­ever, the gen­er­al­i­ty of staff be­ha­vi­or change across the day was not as­ses­sed in this study and thus re­mains an im­port­ant di­rec­tion of fu­ture re­search.

Sev­er­al method­o­log­i­cal lim­i­ta­tions also war­rant com­ment. One lim­i­ta­tion con­cerns the pos­si­ble re­ac­tiv­i­ty of the ob­ser­va­tion sys­tem. Ex­tended ba­se­lines, cou­pled with the sub­jects’ pre­vious ex­pos­ure to video­taping on the job, prob­a­bly ser­ved to limit re­ac­tiv­i­ty. How­ever, the use of a wire­less mi­cropho­ne and the pre­dict­able ob­ser­va­tion sched­ule may have be­come dis­crimin­a­tive stim­uli for in­creased teach­ing in­ter­ac­tions. In sub­se­quent re­search, the po­ten­tial for re­ac­tiv­i­ty could be re­duced fur­ther by con­duct­ing ob­ser­va­tions of lon­ger du­ra­tion (e.g., 30 min), ran­domly sched­ul­ing ob­ser­va­tions to oc­cur through­out the day, or us­ing fre­quent, unan­nounced mo­ment­ary time sam­ples (e.g., Burgio et al., 1983).

Sec­ond, al­though our data sug­gest that in three cases im­proved per­for­mance was main­tained at fol­low-up, these ob­ser­va­tions were lim­ited in num­ber, and it is pos­si­ble that re­new­ing (and an­nounc­ing) ob­ser­va­tions af­ter 1-mon­th breaks may have oc­casioned re­ac­tiv­i­ty. More­over, in the ab­sence of a with­draw­ing phase, pres­um­ing that ex­pe­rience with pe­er man­age­ment func­tioned to main­tain the sub­jects’ per­for­mance re­mains con­ject­ural. Fu­ture in­ves­ti­gators should at­tempt to ex­tend rou­tine ob­ser­va­tions into fol­low-up (e.g., Par­sons, Schepis, Reid, McCarn, & Green, 1987) and use ex­per­i­men­tal de­signs that per­mit a func­tion­al anal­y­sis of the vari­ables main­taining per­for­mance, such as se­quent­i­al, par­tial, or par­tial-sequen­tial with­draw­ing de­signs (Rusch & Kaz­din, 1981).

Fi­nally, al­though the train­er’s in­form­al as­sess­ments of the sub­jects’ forms and graphs in­di­cat­ed that they rap­id­ly came to per­form the pe­er man­age­ment skills com­pe­tent­ly on the job, a more sys­tem­at­i­cal eva­lu­a­tion of the in­teg­rity of the in­de­pen­dent vari­able is war­ranted. For ex­ample, sub­jects might be re­quired to vi­deo- or audiotape some of their pe­er man­age­ment ses­sions, pro­vid­ing an ad­di­tion­al and per­haps more valid source of data.

Pe­er man­age­ment may of­fer a more cost­ef­fective al­ter­na­tive to su­per­vis­ory sys­tems. For each pair, the train­er spent a total of about 6 hr pro­vid­ing in-service and pe­er man­age­ment train­ing and 1 hr per week over 5 weeks ex­tend­ing sup­por­tive feed­back. In prac­tice, train­ing time could be dra­sti­cally re­duced by train­ing lar­ger groups. Al­so, pend­ing the re­sults of fu­ture re­search, the pack­age could be stream­lined with­out loss of effec­tive­ness by elimi­
nating one or more of its components, such as the review meetings or goal setting.

Supportive feedback probably was necessary to ensure the subjects’ correct and continued use of the peer management procedures, as evidenced by two observations: (a) Subjects improved the quality of their feedback and goal selection when the trainer provided feedback initially, and (b) subjects discontinued their formal use of the procedures following the peer management phase. External prompting and feedback have been used similarly in other participative management studies (Burgio et al., 1983; van den Pol et al., 1983), and their necessity is further supported by reports of incomplete or incorrect use of supervisory procedures by subjects in the absence of such support (Babcock, Sulzer-Azaroff, Sanderson, & Scibak, 1992; Page et al., 1982).

In light of the positive consumer satisfaction ratings, it is important to recall that 2 subjects withdrew from the study, expressing their apprehension about giving and receiving peer feedback. Although emphasizing the positive and constructive characteristics of feedback and gradually shaping the requisite skills from the outset could help mitigate an individual’s reluctance to give peer feedback, some staff members might still prefer or respond more favorably to alternative sources of feedback (e.g., supervisor, self).

Demonstrating collateral benefits to clients enhances the external validity of staff training and management programs (Greene, Willis, Levy, & Bailey, 1978). Although observed increases in the work productivity of two groups of clients in the current study were promising, a functional relationship between these increases and the peer management program was not demonstrated. Furthermore, although the clients’ production was monitored frequently and carefully, those data were not checked systematically for interobserver agreement. Staff management researchers should continue to refine their methods for assessing accompanying changes in client performance by selecting valid outcome measures and analyzing the functional effects of their interventions on those measures.

REFERENCES


Received October 20, 1991
Initial editorial decision December 30, 1991
Revision received March 19, 1992
Final acceptance May 29, 1992
Action Editor, John M. Parrish