

ACTION TRAINING FOR CHARISMATIC LEADERSHIP: TWO EVALUATIONS OF STUDIES OF A COMMERCIAL TRAINING MODULE ON INSPIRATIONAL COMMUNICATION OF A VISION

MICHAEL FRESE

Department of Psychology
University of Giessen and London Business School

SUSANNE BEIMEL

Bahn-BKK

SANDRA SCHOENBORN

Department of Psychology
University of Giessen

An action training program that teaches inspirational communication of a vision as part of a training of charismatic leadership for managers is presented (1½ days) and evaluated in 2 studies ($N = 25$ and $N = 22$). We used the research design "nonequivalent dependent variable design" (Cook & Campbell, 1979, p. 118) or "internal referencing strategy" (Haccoun & Hamtieux, 1994), which compared the trained behaviors (charismatic inspirational communication) with behaviors that were not trained (public speech) to control for testing and Hawthorne effects. The training had specific positive effects on those behaviors that were trained but not on those variables that were not trained. Good to excellent effect sizes appeared as a result of the training. We suggest that this research design is useful for evaluation of training effects within the constraints of commercial settings and, moreover, we argue that this design is in many ways superior to a nonequivalent nontraining control group design because it controls for testing effects and for effects that otherwise would need a pseudo-training control group.

This paper is based on two studies and contributes to the management training literature in an area of great importance—charismatic and visionary leadership. Further, in following the call by the recent special issue of *Personnel Psychology* (Autumn, 2002, Vol. 55) to utilize more quasi-experimental designs in work and organizational psychology, we would like to suggest more frequent use of a specific quasi-experimental evaluation design, which requires a minimum of effort, expense, and in-

We thank E-Plus and Der Ruetli who helped to make Study I and Hochtief AG (particularly Alexandra Kamps) for making Study II possible. We also thank the editor and the two anonymous reviewers for their constructive criticism.

Correspondence and requests for reprints should be addressed to Prof. Dr. Michael Frese, Otto Behaghel-Str. 10, 35394 Giessen, Germany; Michael.Frese@psychol.uni-giessen.de.

terruption to the participating firms. Commercial settings are often not prepared to indulge in complicated training evaluations because they are seen as disruptive and expensive, and, therefore, it is useful to develop adequate and minimally invasive evaluation procedures. Finally, we would like to show empirically that one form of training, action training—an often-used tool in Germany but little known outside of Germany—is useful for management training and proves to be useful to improve the quality of trained behaviors in two studies.

Before we present the studies, three issues need to be discussed in some depth: (a) how a vision is communicated to others, (b) design issues related to the approach that we suggest and have used in our studies, and (c) action training as a management training concept. The two studies are then presented and we attempt to draw conclusions on the usability of the training concept, the design idea, and charismatic training.

Charismatic Communication of a Vision

Different concepts of charisma converge on the idea that charismatic leaders communicate a vision that inspires others (Conger & Kanungo, 1998). Thus, it makes sense that a training program for charismatic leadership should emphasize inspirational communication of a vision. Bass (1990) described idealized influence and inspiration as two aspects of transformational leadership (which, in addition, comprises intellectual stimulation and individual consideration). Kirkpatrick and Locke (1996) argued that there are three core aspects of charismatic leadership: vision, vision implementation, and charismatic communication style (cf. Baum, Locke, & Kirkpatrick, 1998). Vision implies that the charismatic leader has a positive, general, and emotionally laden ideal related to strong values. Organizational vision implies that the leader knows what the core values and core tasks of the organization are and what the organization should achieve; a vision should inspire and motivate the subordinates to perform exceptionally well. Managers have to make sure that people understand the implications of the vision. The vision has to be communicated so as to generate enthusiasm and to “inspire” subordinates, peers, or customers.

Evidence suggests that charismatic and inspirational communication is related to performance. Several studies showed that leaders' charisma is related to unit performance (e.g., Geyer & Steyrer, 1997; House, Spangler, & Woycke, 1991; Howell & Frost, 1989; Kirkpatrick & Locke, 1996; Lowe, Kroeck, & Sivasubramaniam, 1996). The most direct test of the effects of charismatic leadership was performed by Howell and Frost (1989). In this study, actors played supervisors with different styles and the charismatic style turned out to be effective for satisfaction and per-

formance, particularly in low productivity norm groups. In addition, their study showed that communication is of particular importance for charisma (cf. also Baum et al., 1998).

Charismatic and inspirational communication is characterized by content and stylistic components. In terms of content, charisma was characterized by stressing the importance of the project, by sharing a vision related to the project, by increasing confidence of the subordinates, and by stressing a common goal in Howell and Frost's (1989) study. Shamir, Arthur, and House (1994) argued similarly and found charismatic speeches to contain elements of collective identity, positive references to collective and individual self-worth and efficacy of followers, references to similarity between leader and followers, moral justifications and values, long-term goal orientation and visions, and references to hope and faith (cf. Den Hertog & Verburg, 1997).

In terms of stylistic components, Howell and Frost (1989) used certain paralinguistic concepts to portray a powerful, confident, and dynamic presence of the leader. The charismatic leaders used a captivating and engaging voice tone, pacing and sitting, leaning forward, direct eye contact, and animated facial expressions. Although there are conflicting results on paralinguistic characteristics (e.g., it was not effective in Kirkpatrick & Locke's [1996] study), paralinguistic communication criteria are important for bringing across the emotional side of charismatic leadership (Holladay & Coombs, 1994) as well as a strong delivery style (Awamleh & Gardner, 1999).

The literature consistently suggests that communicating a vision and inspiring people to work for this vision is central to charismatic behavior (Baum, Locke, & Kirkpatrick, 1998). With the exception of paralinguistic communication, the literature is also consistent in its description of which aspects of communication are effective. The literature does not, however, provide evidence that communication of a vision can be changed through training.

Training for enhancing charismatic behavior has been called for by Barling, Weber, and Kelloway (1996) and can be based on Howell and Frost's instructions to the leaders in their experiment (1989; Kirkpatrick & Locke, 1996; Shamir, Arthur, & House, 1994). In Germany—where this training was done—improving the visionary leadership of managers is particularly called for because most managers do not have a long-term vision (Schwertfeger, 1997).

Design Issues for Training Evaluation in Commercial Settings

A particular problem of training evaluation is to convince business leaders of the need for a control group. A pretest/posttest design has

several disadvantages, particularly its inability to control for maturation, history, testing, and instrumentation effects (Campbell, 1963), as well as for the regression effect (Cook, Campbell, & Peracchio, 1990). However, commercial firms and personnel departments usually do not support training evaluation with a randomized assignment to experimental and control group. Moreover, managers are usually not willing to participate in research from which they do not receive any direct benefits. Consequently, much of training evaluation research—even in the U.S.—uses a simple pretest/posttest design (Sackett & Mullen, 1993).¹ This was also the case in our studies. Management that paid a high sum for the training would not have been prepared to accept a waiting control, or any other form of control groups (let alone random assignment). An additional reason that makes a control group awkward to evaluate in our training program is the fact that inspirationally communicating a vision presupposes that a vision must be developed in the first place. Thus, the control group would have had to develop a vision but would not have been trained in how to communicate the vision; it is unlikely that a company would have paid for such a training of their managers.

However, there is an even more important effect that is, surprisingly, rarely discussed in training research: the Hawthorne or placebo effect. Simply reflecting upon one's behavior as a manager may already produce changes in behavior. Such an effect may appear as a general result of participating in any training and it can be likened to a placebo effect in the fields of medicine and psychotherapy (cf. Hollon & DeRubeis, 1981). Goldstein (1974) subsumed this effect under the umbrella term of Hawthorne effect. Unfortunately, using a random control group design with a nontreatment control cannot rule out such a placebo or Hawthorne effect because differences between training and nontreatment control groups may be due either to the specific training or to self-reflection that happens as part of any training (and is, therefore, not specific to a particular training procedure). Thus, training and self-reflection are confounded within such a design. For this reason, good training evaluation research would also need to employ a pseudotraining group (similarly to a pseudotherapy group or placebo group in psychotherapy research). A pseudotraining control group would assemble a group of participants outside the workplace, would involve them in thinking about their own leadership behavior, and would allow people

¹Niemiec, Sikorski, Clark, and Walberg (1992, p. 300) concluded from their meta-analysis of commercial management training programs in the U.S.: "Research standards were routinely ignored or forgotten. . . Much of what is claimed as research amounts to cheerleading, hawking of a specific service, or justification for budget increase. . . corporate training evaluation is in a dismal state."

to talk and reflect about what they think is important. (It is obvious that such a control group is hardly justifiable in a commercial setting.)

The second major problem for training evaluation is the testing effect. Measurement and participating in a study inadvertently leads to some feedback, and measurement provides hints of required behaviors. Thus, measurement produces changes that are not due to training but that are confounded with a training effect within a pretest/posttest design. The testing effect may be one of the reasons why pretest/posttest designs lead to higher effect sizes than the more adequate pretest/posttest control group design (Carlson & Schmidt, 1999).

In the spirit of Sackett and Mullen (1993) and Salas and Cannon-Bowers (2001), who argued for the development of alternative and pragmatically useful evaluation designs, we decided to base our training evaluation on a design concept called "nonequivalent dependent variable design" (Cook & Campbell, 1979, p. 118) or "internal referencing strategy" (Haccoun & Hamtieux, 1994). Briefly, the design is based on a single treatment group and compares two sets of dependent variables—one set that is hypothesized to be affected by the training and another set that is hypothesized not to be affected by the training. We think that this training design is advantageous for training evaluation in commercial settings. The design does not require a control group or a pseudo-training group because it effectively controls for testing and Hawthorne effects. We measured not only those behaviors that we intended to train (we call them experimental variables) but also those behaviors that were not the focus of the training (we call them control variables) and compared them with each other. Because all trainees participate in a real training program, we can effectively rule out nonspecific placebo effects if the training only leads to changes in the experimental variables but not in the control variables. Further, testing effects can be ruled out as well because both types of variables are tested.

This type of evaluation design should be used much more frequently: To our knowledge there has been no other published evaluation using this design, bar Haccoun & Hamtieux (1994), and it has not been used at all for leadership training evaluation. We think that this design is particularly useful for leadership training. The design is inexpensive, easy to use, unobtrusive, and, therefore, does not lead to the same resistance from management as randomized control group designs. Moreover, it may actually be better than a nontraining random control group because this design controls for the abovementioned placebo effect. It is superior to a nonequivalent (i.e., nonrandom) control group because there is no danger that an unmeasured variable that differentiates between the nonequivalent control and the training groups may interact with the

training in question (e.g., self-efficacy may be higher in the nonequivalent control group because those people volunteer for such a control group who perceive that they do not need this particular training).

One problem of the nonequivalent dependent variable design (Cook & Campbell, 1979) or internal referencing strategy (Haccoun & Hamtieaux, 1994) is that the experimenter is able to control how hard or how easy it is to get significant differences between the experimental and control variables. An experimenter can either take very distant variables as "control variables" or variables that have a high degree of similarity to what is trained (Cook & Campbell, 1979). If very distant variables are taken, it is easy to achieve significant differences. Therefore, it is necessary that control variables are included that are conceptually similar but distinct from those variables that are trained. Thus, experimenters must stay in the middle ground between choosing control variables that are obviously different and control variables that are too similar (and, therefore, co-change as a result of the training). The best control variables are those that allow potential Hawthorne or placebo effects (e.g., as a result of having time to self-reflect on one's performance in a certain area) to be tested and that are, at the same time, not directly affected by the experimental manipulation. Because we trained inspirational communication and because this communication was presented in the form of a speech, we measured traditional "public speaking" ("rhetoric") effects as untrained "control variables." There are obvious similarities between trainings in public speaking and inspirational communication because both imply that a one-way communication is taking place between a speaker and a group of people. Indeed, participants in our training program often spontaneously thought of principles of good public speaking during their training. However, there are also differences: An inspirational speech is primarily related to the content—the vision—and is oriented to be inspirational and, therefore, should increase emotional and motivational consequences in contrast to a public speech. By measuring untrained public speaking variables (control variables), we could evaluate whether our training program had a significant effect over and above testing effects.

The Action Training Approach

We used action theory based training or, for short, action training, to improve communication skills in managers. Action training is based on action theory and has been shown to be useful in learning computer and other cognitive-behavioral skills (Carroll & Mack, 1984; Carroll, Mack, Lewis, Grischkowsky, & Robertson, 1985; Frese et al., 1988; Martocchio, 1994). Action training approach has some overlap with other training

concepts, such as exploratory training (Bruner, 1966; Greif & Keller, 1990) and behavioral modeling (Bandura, 1986; Latham & Saari, 1979). Our two studies were not designed to contrast different training procedures; however, we need to explain our approach and we shall, therefore, use behavioral modeling (Bandura, 1986; Latham & Saari, 1979) as a template with which to compare action training. The components of action training (Frese & Zapf, 1994; Hacker & Skell, 1993; Semmer & Pfäfflin, 1978) are the following: action-oriented mental model, learning by doing, motivation by experiencing the difference between present state and future goals, feedback in training, supporting transfer, and the necessity to routinize behavior.²

First, people need to develop an *action-oriented mental model* of what constitutes effective actions within certain situations (Hacker, 1992). An action-oriented mental model is a cognitive representation of the starting situation, the goal state, and how the present situation can be transferred into a future state. Action-oriented mental models help employees to work more efficiently and effectively (Frese & Zapf, 1994; Hacker, 1992). The action-oriented mental model is not exhaustive and not complete; rather, it consists of a set of "principles" (rules of thumb). The use of such principles proved to work well in training (Parry & Reich, 1984; Volpert, Frommann, & Munzert, 1984) and is well known since the study by Judd (1908). Providing principles is similar to teaching learning points, as suggested by Latham and Saari (1979) within a behavior modeling approach. However, the learning points refer to the behaviors of the model to be emulated (Latham & Saari, 1979), but in the case of action training, principles are more generally taught and are not tied to a behavioral model displaying the required behaviors. In our training program, we described the empirical research and theories of charismatic leadership and derived 13 principles from this literature to describe charismatic and inspirational communication. These principles were presented in a handout entitled: "Principles of good visionary communication."

Second, action theory suggests that people *learn by doing*; therefore, participants are supposed to take an active approach and to learn via role-play procedures. Theoretically, the following arguments speak for learning by doing: (a) The cognitive apparatus is built for action. Action theory assumes that people can only survive and procreate if they show an active orientation toward the environment. Cognitions (Gibson, 1979; Hofsten, 1985; Neisser, 1985) and emotions (Frijda, 1986) are

² Action theory based training or action training is not the same as action learning (Conger & Benjamin, 1999; Revans, 1982). Although action theory agrees that people learn by doing things, action training acknowledges the need to develop formal training seminars outside the work activities.

at the service of actions and the ability to keep up actions against most odds. (b) Cognitions develop their regulatory function over actions during action. The connection between thought and action is developed via actions. There can be cognitions that do not regulate actions (Semmer & Frese, 1985). One problem of training is that people learn abstract principles but may not be able to *connect* these principles to specific actions (Semmer & Frese, 1985). People often develop attitudes and lofty goals as a result of training but they are not really translated into actions and, therefore, have no regulatory power over actions (e.g., in the case of procrastination). Action theory assumes that when people learn knowledge through actions, there is a tighter relationship between cognitions and actions and the regulatory function of cognitions is acquired while learning skills and knowledge (Semmer & Frese, 1985). This is similar to Anderson's (1983) concept of compilation, taken to be a necessary step to get from declarative to procedural knowledge. (c) Actions help people to explore a system and situational parameters that decide when to use and when not to use a skill. Exploration has been shown to lead to better performance (Dormann & Frese, 1994) even when people explored against the suggestion provided by the trainer. Therefore, in our training program, participants were asked to role-play communicating a vision inspirationally to their employees, and every participant had to perform several times so that the general principles were connected to their actions.

Third, participants can be motivated in training by experiencing that they are not yet good enough in the required skills. Action theory suggests that motivation increases through negative feedback (at least as long as it is a safe environment in which negative feedback is not threatening the ego) because it tells the person that a certain goal (e.g., demonstrating some skill) is not yet achieved (difference between *present state and future goal state*). The wider the gap between present state and future goal state, the more motivated the individual is to attempt to learn a skill (given enough self-efficacy and high value for the skill). Such a proposal is consistent with goal setting theory's insistence that high goals produce high motivation (Locke & Latham, 1990). Participants in action training are sometimes asked to role-play the desired behavior before they are trained: Thus, participants get to know their weaknesses and this increases their motivation to participate in the training. Such a recommendation is in contrast to the social cognitive approach to training (Latham & Saari, 1979): Based on behavior theory, this approach assumes that one should be careful not to frustrate the participants. A related issue is the issue of learning from errors. Action training suggests learning from errors (Frese, 1995; Frese et al., 1991). The training environment must be structured in such a way that people feel safe to

make errors and that the frustration that can appear as a result of making errors is reduced (Heimbeck, Frese, Sonnentag, & Keith, 2003). Action theory's emphasis on making errors and using errors as feedback is not shared by sociocognitive theory, which argues that errors are wasteful and that training programs should teach the right skills throughout (Debowski, Wood, & Bandura, 2001). In our training, the participants experienced firsthand that they were not yet good enough to inspire people for their vision. In addition, they were told not to "play it safe" but were encouraged to experiment, to make errors, and to learn from their own errors and those of other participants.

Fourth, participants should get *feedback* on their behavior. Action training suggests that feedback should not only be provided by the trainer and other participants, but also by one's own critical look at one's performance. Action theory uses action feedback extensively because it is assumed that feedback is the primary mechanism by which action gets corrected and learning occurs (Frese & Zapf, 1994; Miller, Galanter, & Pribram, 1960). Positive and negative feedback is provided with a functional task perspective ("if you mention your vision just once, people will forget; you noticed that your audience did not really remember your vision after your presentation"). In the beginning of the training process, feedback needs to be more detailed and the trainer is more involved in providing feedback; later, feedback can be provided more in the sense of a summary label ("provide more emotional expression"). Action training provides positive and negative feedback and uses the principles as a guide to provide feedback (e.g., "Your use of gestures has been very good but you did not vary your speed of delivery enough. Varying the speed of speech increases the emotional flavor and underlines what is important to you. You should speak faster, for example, when you want to say that..."). Action training does not attempt to phrase negative feedback positively, because negative feedback helps the person to get to know which aspects of the skills he or she has not yet mastered (on the other hand, feedback should not crush self-confidence). It has been a concern of action theory to provide task feedback rather than trainer feedback. Trainer feedback is artificial feedback (Holding, 1965) or feedback intervention (Kluger & DeNisi, 1996). Artificial feedback may make trainees dependent on this feedback so that they are at a loss when they do not receive this feedback any longer in their normal work situation. Our training program took three lessons from the feedback discussion: (a) the trainer provided feedback in the beginning but phased out his or her feedback with time and then encouraged the participants to self-generate feedback; (b) feedback was given with a functional task perspective, thus relating the feedback to the task ("if you do not look at the audience, they will be less motivated to listen to you") to reduce

the artificial nature of feedback interventions; and (c) the trainer emphasized feedback from the task. For example, he or she might ask the other participants whether they felt inspired by the inspirational speech and would use this feedback as task feedback for the participant to improve his or her performance. To our knowledge, this approach to feedback by action training is not a systematic part of behavioral modeling approaches to leadership skill training.

Fifth, action training attempts to maximize transfer and emphasizes this issue. One way to enhance transfer is to develop a mental model with principles; these principles can be transferred to the work situation (Judd, 1908; Skell, 1972; cf. Kozlowski & Salas, 1997). Further, people are asked to think of examples from their work so that they develop knowledge on applying the principles to their work (enhancing transfer knowledge as suggested by von Papstein & Frese, 1988). Therefore, our training program provided principles of actions, an operative mental model, and attempted to increase transfer knowledge. Moreover, participants were asked to think of specific dates when they would use their newly developed skills (and to write these dates into their agendas).

Sixth, action theory argues that new skills developed during the training program compete with old, well rehearsed routines (Frese & Zapf, 1994). Therefore, we found it necessary to repeat performance during training so that a certain amount of routinization occurs and to emphasize and discuss practice issues (e.g., the possibility of falling back into the trap of using their old routines) for the time directly after training (cf. Driskell, Willis, & Cooper, 1992).

Thus, action theory has helped us to develop such a training program. There are certain principles (e.g., on feedback, on learning by doing, on operative mental models, and the importance of establishing new routines and warning of the danger of falling into past routines) that action theory as a cognitive-behavioral theory emphasizes more strongly than other approaches. However, in order not to produce misunderstandings, we would like to emphasize that action training also acknowledges the importance of processes, for example, of learning from others during role-play, and that there is an appreciable degree of overlap to the behavioral modeling approach.

Study I

Study I provides an evaluation of an action training program designed to teach managers to improve their skills to charismatically communicate a vision. For the evaluation, we utilized a nonequivalent dependent variable design that has not been applied to leadership training

before. The most important method for evaluating the training is the assessment of changes in communication acts before and after the training through expert ratings.

Methods

Sample

The sample consisted of midlevel managers (in all, $N = 25$) who participated in one of three courses of a management training program conducted by the first author at a commercial training institute.³ The managers were from a modern, fast-growing, start-up mobile phone provider in Germany, aged between 25 and 55 years, and mainly men (two were women). All of the participants had an academic background, mostly in the technical and business areas. They took part in the training as a routine part of their education to become a first and second line supervisor (group or department manager). The participants were informed that we evaluated the training.

Training Procedure

The training consisted of two components. First, the managers were supposed to develop an appropriate vision for their group or department. Second, both paralinguistic and content issues of charisma were trained to increase managers' skills to inspire their group—this part was called inspirational speech. The training module evaluated in this paper refers to the second part, but the development of a vision was also important and needs to be described as well.

The appendix presents an overview of the training. After a brief introduction to the function of a vision, the participants were asked to prepare and then to enthusiastically role-play a speech about their vision in order to inspire their subordinates to work harder toward this vision. Because participants have implicit standards of performance and are able to recognize whether or not a speech was inspirational, they would notice the discrepancy between what an inspirational speech ought to be and how their speech actually was at this time (this is Time 1 = T1). At this point, no further trainer feedback was given. However, all participants filled out a questionnaire on how much their colleagues had inspired them directly after each role-play (role-plays were videotaped to be coded later).

³The typical group size of a training in good commercial schools in Germany is usually between 8 and 16—this was also the case here.

Then the trainer lectured on the importance of visions for managers at all levels. Although charismatic leadership and visions have often been described as more important for higher level managers, there is evidence that charismatic leadership is related to higher performance for lower level managers as well (summarized by Barling et al., 1996). Obviously, the vision must be appropriate for this level and should not contradict the higher level visions of the company. Visions were described to be long-term, future-oriented goals with emotional appeal that are embedded in a set of values. A vision should be short (at best, a few words), and should make it possible to develop some kind of image of what needs to be achieved in a group or department (Baum et al., 1998). The power of visions was presented in this lecture by providing examples of visions of other firms and departments and by presenting the empirical literature.

Principles of good visions were presented so that a good operative mental model could be developed. The trainer asked the managers to develop an improved vision and this vision was discussed and modifications were suggested. An example of such a vision might be that a particular department functions as a competence and action center for a technical tool frequently used by the company. This would imply continuous development of knowledge, active approaches to convince intrafirm clients and upper level management of the usefulness of these tools, and to be easily accessible for questions.

After the vision was developed, the concept of charismatic leadership was introduced by presenting empirical studies. Inspirational speech content and formal principles were presented and discussed. These principles were related to Howell and Frost's (1989) and to Shamir et al.'s (1994) presentation of charismatic leadership and include the paralinguistic aspects of using gestures, variations of speed and loudness of speech, orientation (e.g., leaning) toward the audience, eye contact, as well as the content issues of repeated description of the vision, use of metaphors and images, increase of group self-efficacy, positive emotional appeal, value orientation, and the use of "we" instead of "I-you" sentences.

The participants were then asked to prepare an improved version of their inspirational speech in a second step and to role-play it in front of the group while pretending to be in front of their subordinates. The speeches were videotaped and questionnaires (the same as before) were filled out (at T2). After each speech, the videotape was shown and the trainer and the participants provided feedback by discussing how far each participant had followed the principles of charismatic communication. Feedback was mostly informative and presented in a matter-of-fact style. No attempt was made to be particularly positive or negative. After feedback, the trainer asked the participants to think of situations in

their work life in which they could use these principles and in which they could communicate their vision for their department.

On the next day, the inspirational speech was presented a third time (and videotaped) after a preparation time of 10 minutes (at T3). Participants were asked to observe those principles that they found difficult to adhere to before. No specific feedback was given after this speech except general praise on how much they had improved. Questionnaires were filled (the same as at T1 and T2). This concluded the module on charismatic leadership, and other topics were then pursued in the course of this training program.

Measures

Tannenbaum and Yukl (1992) called for more research on behavioral outcome measures in management training. Our behavioral outcome measures were related to the principles mentioned above: eye contact, gestures, variations of speed and variation of loudness, orientation toward audience, repetition of vision, explaining significance of vision, value appeal, use of metaphors, increase of group self-efficacy, emotional appeal, positive statements, use of "we" form. Two raters coded these behaviors during presentation of the videos as well as after.

In addition to the principles that were taught to the participants, we also measured additional public speaking (rhetoric) variables that had not been taught during the training program, but are considered to be elements of effective public speaking (control variables; Duden, 1994; Fey, 1993; Jentsch, 1993): good structure of speech; good frame, that is, beginning and end well connected; use of rhetoric questions; simple and easy sentences; refraining from use of nonlexical utterances such as aah; clear pronunciation; relaxed posture; and artificial pauses for example after "but" and "therefore." Hence, we had 12 experimental variables related to the principles taught for visionary speech (although eye contact could not be coded from the videos and was, therefore, not measured); in addition, we had 8 control variables related to public speaking principles that were not trained.⁴

The raters received approximately 5 hours of training to use the coding form. Moreover, to increase agreement, the raters were given explicit definitions and examples of each principle of visionary speech or public speaking. The videos were randomly arranged so that the coders did not know whether these videos were from the first or third measurement wave. The raters were not blind to the theoretical difference between public speaking and charisma variables (they needed to understand what they were coding to be good raters). Because we were only in-

⁴Intercorrelation tables for both studies can be received from the first author.

terested in differences between before and after the training, only blindness with regard to the waves was needed. The coding form contained a 5-point response format (1 = *not true*, 2 = *not completely true*, 3 = *middle*, 4 = *true*, 5 = *completely true*) for assessing how often each principle was present in the speech. The independent ratings of the two coders were added together and divided by two to keep a 1–5 response format. Because we were only interested in pretest–posttest differences and because T2 was used for providing feedback to the participants, we only analyzed the measures of T1 and those of T3 (we only kept and coded the videos for these two waves). Interrater agreements of the video ratings were, on average, .83 for T1 trained items, .94 for T3 trained items, .84 for T1 control items, and .96 for T3 control items (ICC 3,k for averaging the scores). There were only two cases in which the two raters had an agreement below .70 for individual items (suggested by Klein et al., 2000 as a cut-off point), but given that in both cases T3 agreements were sufficiently above .70, we kept these variables in our analyses.

Feeling inspired: The peers/trainees were asked to rate subjectively how much they felt inspired by the speech of their peer. We used this as one criterion of performance because the inspirational speech was supposed to produce the appropriate feelings in other people belonging to this company.

In addition, to examine the transfer, we conducted telephone interviews 10 weeks after the training program to ascertain whether participants had implemented what they had learned at their work place. We asked if they had given any inspirational speeches and, if so, what the reaction of their audience had been. It should be noted that we do not assume to be able to test transfer effects with the telephone interviews—they are just descriptive add-ons to provide face validity of the training program.

Statistical Methods

We used repeated measurement multiple analyses of variance (RM-MANOVA) to test whether participants improved after training and whether there were differences between the trained and the control variables. Effect sizes were calculated with the formula $M(T3) - M(T1)/SD(T1)$ (Carlson & Schmidt, 1999). Results on the transfer interview are only reported descriptively.

Results

The RM-MANOVA for the trained items was significant with a Wilks-Lambda of 83.25 ($df = 1,24$, $p = .000$, partial $\eta^2 = 0.776$) for the within-

TABLE 1
*Means and Standard Deviations of the Trained and Untrained
 (Control) Variables, F-Values, and Effect Sizes (Study I)*

	M_{T1}	SD_{T1}	M_{T3}	SD_{T3}	F	d
<u>Trained items:</u>						
Gestures	2.840	1.134	3.620	.881	23.49**	0.69
Variation of speed	2.560	1.364	3.560	.821	12.63**	0.73
Variation of loudness	2.120	1.054	2.820	1.207	6.76*	0.66
Orientation toward audience	2.100	.935	3.220	.792	67.56**	1.20
Repetition	2.260	1.183	3.720	.891	37.66**	1.23
Value appeal	2.520	1.141	4.020	.907	50.94**	1.31
Increase of group self-efficacy	2.640	1.095	4.060	.961	30.37**	1.30
Use of "we" form	3.340	.943	3.880	.939	6.61*	0.57
Use of metaphors	1.880	.927	2.940	1.261	13.31**	1.14
Emotional appeal	1.740	.948	3.000	1.414	21.14**	1.33
Positive statements	3.220	.805	4.040	.776	23.82**	1.02
Explaining significance of vision	2.220	1.071	3.600	1.031	25.31**	1.29
<u>Untrained items:</u>						
Good structure of the speech	2.980	.984	3.300	.979	1.26	0.33
Good framework: Beginning and end well connected	2.360	1.327	3.100	1.242	2.91	0.56
Use of rhetoric questions	2.400	1.146	2.320	1.079	.08	-0.07
Simple and easy sentences	2.960	.957	3.560	1.054	5.84*	0.63
Refraining from use of nonlexical utterances	2.680	1.257	1.920	1.106	14.71**	-0.60
Clear pronunciation	3.180	1.154	3.720	.925	8.99**	0.47
Relaxed posture	3.180	1.050	3.880	1.054	8.65**	0.67
Artificial pauses	1.660	.800	2.140	1.005	4.04	0.60

* $p < .05$ ** $p < .01$

subject repeated measure factor. The univariate comparisons showed that all 12 items had significantly improved from before to after training (cf. Table 1). Thus, each criterion measure of good inspirational communication improved significantly. Unexpectedly, the repeated measurement factor for the nontrained (control) variables was also significant with a Wilks-Lambda of 5.422, ($df = 1,24$, $p = .029$, partial $\eta^2 = 0.184$). However, in the univariate tests of the eight control variables, only three showed a significant ($p < .05$) improvement and one a significant decrease after the training (cf. Table 1). The three control variables that showed a significant improvement in the training were simple and easy sentences, clear pronunciation, and relaxed posture. In addition, there was a decrease in the variable "refraining from use of nonlexical utterances"; in other words, there was a deterioration in this variable after training. The average effect size for the 12 trained items was $d = 1.04$ and for the 8 nontrained control items, $d = 0.32$. A RM-MANOVA testing the factors Time (repeated measures) and Training (trained vs. control variables) and Time \times Training leads to significant

TABLE 2
*Correlations of Trained Items with Feeling Inspired by Listeners
 (T3; Study I)*

Trained items	<i>r</i>
Gestures	.23
Variations of speed	.25
Variations of loudness	.44*
Orientation toward audience	.38
Repetition of vision	.08
Explaining significance of vision	.28
Value appeal	.21
Use of metaphors	.33
Increase of group self-efficacy	.39
Emotional appeal	.29
Positive statements	.14
Use of "we" form	.26

* $p < .05$

Time (Wilks-Lambda = 36.999; $df = 1,24$, $p = .000$, partial $\eta^2 = .607$), nonsignificant Training (Wilks-Lambda = 1.869; $df = 1,24$, *ns*, partial $\eta^2 = .072$) and significant Time \times Training (Wilks-Lambda = 40.345; $df = 1,24$, $p = .000$, partial $\eta^2 = .649$) effects. The latter effect (Time \times Training) is the most important for us, showing that with training, participants become better in the trained variables than in the control variables. Cohen's *d* comparing trained and control variables after training was .95. The subjective feeling of being inspired by the speech, measured by peer questionnaire responses, also increased significantly from before to after the training ($F[2,22] = 35.06$, $p < .01$, with an effect size of $d = 1.14$).

The correlations reported in Table 2 related trained items (measured with video ratings) to feeling inspired by the listeners (measured by the peer reactions). Although only one of these correlations was significant, the majority of the trained items explained the variance of feeling inspired with more than 5%—we assume that lack of power was the main reason the correlations were nonsignificant. The most important correlations to feeling inspired existed for variations in loudness, increase of group self-efficacy, orientation toward the audience, and use of metaphors (cf. Table 2—all of these correlations were above .30).

The transfer interviews revealed that 80% of the participants reported using inspirational speech at least once, with 16% once, 28% twice, 8% three times, and 28% more than three times over a period of 10 weeks. Of those who had given at least one inspirational speech, 83% reported that the speech was positively or very positively received by the audience.

Discussion

Study I showed that there was, indeed, an effect of the training program on the trained variables that was distinguishable from the effect found for the untrained control variables. However, there was also an unexpected effect that our action training had on the control variables. The construct of control variables is central to our design and, therefore, it is problematic in that there was an effect of training on these variables as well. A certain amount of change in these untrained control variables may have been due to a Hawthorne effect. An alternative interpretation is that our control variables were from the area of public speaking, and because participants were asked to give an inspirational speech, they may have thought about general public speaking principles as well and improved them during the course of the training. We were unable to decide between these two potential interpretations of this study. This convinced us that we needed to do a more sophisticated analysis of the control variables and needed to implement a better strategy for selecting them. An additional point of concern from our study was the small number of participants and the question of whether the results could be replicated and generalized. Therefore, we decided to conduct a second study that was supposed to overcome some of the problems identified in the first study.

Study II

Study II was done for three reasons: First, we wanted to replicate the findings of Study I. Second, we attempted to ask the generalization question: Can the results be replicated with managers from another industry? Third, we thought it necessary to perform a more sophisticated selection of the control variables.

Methods

Sample and Procedure

Similar to the sample of Study I, 22 midlevel managers participated in two separate management training courses given by the first author in a commercial setting. Participants had academic degrees. In contrast to Study I, the participants were from a different industry, which was struggling industry wide at the time of the study; they were managers in a large, economically stable but shrinking German construction company with worldwide activities. The age range was from 27 to 56 years; four participants were women.

We used the same training procedure as in Study I except that, following the wishes of the participating company, participants only gave two (instead of three) inspiring visionary speeches, one time before the training and one time afterwards. Participants did not complete questionnaires, but we used the same analysis of video recordings as in Study I.

Measures

We used primarily the same measures for the trained variables as in Study I, except that we added two coding variables and slightly modified one. We measured "using the vision to organize the talk" and "feeling inspired by speech" in this study (in contrast to Study I, we instructed the trained raters to code feeling inspired), and we changed "explaining significance of vision" in Study I to "presenting the vision clearly" in this study. The biggest difference between the two studies can be found, however, in how we developed and selected the untrained control variables. We developed 14 items from descriptions of two training seminars on presentation techniques (public speaking training programs; Kissling-Institut, 1998; Olbort, 1998). These 14 items were then given to 10 psychology graduate students who independently coded how similar these control variables were to the experimental variables measuring principles of inspirational speech. Interrater agreement (ICC 2) of these ratings was .72. Based on these ratings, we chose those items that were coded to be least similar to principles of inspirational speech. The average similarity of the items, thus chosen, was 2.57, and the nonchosen items were coded with an average of 3.37 (on a scale of 1–5, with 5 being the most similar to the principles of inspirational speech).

A short form of the seven control variables, thus chosen, is presented in Table 3. Coding of the videos was performed in the same way as in Study I. Interrater agreements of the video ratings were, on average, .90 for T1 trained items, .86 for T2 trained items, .86 for T1 control items, and .71 for T2 control items (ICC 3, *k* for averaging the scores). There were two cases in which there were agreements below .70 (.65 and .68) in the trained items but given that, in both cases, agreements at the other time were sufficiently above .70, we kept these variables in our analyses. In the control (untrained) variables there were, however, three cases in which T2 agreements were lower than .70; (these were .66, .62, and .41); however, we continued to use these items because there was no agreement lower than .70 at T1.⁵

⁵When we excluded the variable with the lowest agreement of .41—use of building up importance—from the analyses, the same results were found as when it was included.

TABLE 3
*Means and Standard Deviations of the Trained and Untrained
 (Control) Variables, F-Values, and Effect Sizes (Study II)*

	M_{T1}	SD_{T1}	M_{T2}	SD_{T2}	F	d
<u>Experimental (trained) items:</u>						
Gestures	3.02	1.04	4.30	.87	48.71**	1.23
Variation of speed	3.50	1.23	4.36	.68	14.01**	.70
Variation of loudness	3.34	1.29	4.23	.70	12.98**	.69
Orientation towards audience	2.73	1.04	4.02	.75	25.99**	1.24
Repetition of vision	2.50	1.62	4.05	1.43	13.80**	.96
Vision presented clearly	2.61	1.46	3.98	1.28	13.40**	.94
Vision used to organized speech	2.45	1.40	3.93	1.37	17.46**	1.06
Value appeal	2.80	1.11	4.09	.85	33.23**	1.16
Increase of group self-efficacy	2.93	1.32	4.32	1.03	31.75**	1.05
Use of "we" form	4.27	1.04	4.59	.85	5.33*	.31
Use of metaphors	2.77	1.37	3.93	1.16	10.95**	.85
Emotional appeal	2.41	1.17	3.89	1.00	60.23**	1.26
Positive statements	2.30	1.27	3.68	.95	28.67**	1.09
Inspired listener	3.02	1.32	4.52	.84	25.67**	1.14
<u>Control (untrained) items:</u>						
Good closure of speech	2.07	.90	2.11	1.05	.04	.04
Good organization, such as a, b, c	1.18	.48	1.82	1.09	5.74*	1.33
Use of building up importance	1.11	.53	1.39	.65	2.39	.53
Use of polarized contrasts	1.00	.00	1.13	.54	1.41	.24 ^a
Examples from practice	1.45	.58	1.48	.55	.14	.05
Visualization	1.73	1.21	1.37	.56	.76	-.30
Combines serious/factual with witty and comical information	1.11	.53	1.10	.25	.03	-.02

^a SD_{T1} could not be used because it was zero; therefore, in this case we used SD_{T2} to calculate d .

* $p < .05$, ** $p < .01$

Results and Discussion

The within-subject repeated measure factor of the RM-MANOVA for the trained items was significant (Wilks-Lambda $F = 44.577$, $df = -1, 21$, $p = -.000$, partial $\eta^2 = .680$). Univariate comparisons showed that all 14 items had significantly improved from before to after the training (cf. Table 3). Thus, each criterion measure of good inspirational communication improved significantly. This includes the variable coded as being inspired by the speech, which also increased significantly from before to after the training (cf. Table 3; with an effect size of $d = 1.26$). As expected, the RM-MANOVA for the nontrained (control) variables was not significant ($F = 1.583$, [$df = 1, 21$, $p = .222$, partial $\eta^2 = .070$]). Of the seven univariate tests of these control variables, only one showed a significant ($p < .05$) improvement after the training (cf. Table 3). The average effect size for the 13 trained items

was $d = 0.98$ and for the seven nontrained items $d = 0.27$. A RM-MANOVA testing the factors Time (repeated measures) and Training (trained vs. control variables) and Time \times Training leads to significant Time (Wilks-Lambda = 39.167, $df = 1,21$, $p = .000$, partial $\eta^2 = .651$), significant Training (Wilks-Lambda = 235.424, $df = 1,21$, $p = .000$, partial $\eta^2 = .918$) and significant Time \times Training (Wilks-Lambda = 35.150, $df = 1, 21$, $p = .000$, partial $\eta^2 = .626$) effects. The latter effect (Time \times Training) showed that, with training, participants became better in the trained variables than in the control variables. Comparing trained and control variables after training (the sum of all trained and untrained variables divided by number of items respectively and comparing the means, divided by their common standard deviation) produced an effect size of $d = 6.28$.

Overall Discussion

Our two studies suggest that action training to improve one aspect of charismatic leadership—being inspirational—is successful and produced positive changes in objectively recorded behaviors. These results could be replicated in managers from a different industry. In addition, Barling et al. (1996) showed that another aspect of charismatic leadership—intellectual stimulation—could also be trained and Dvir, Eden, Avolio, and Shamir (2002) recently demonstrated that a general training to become more charismatic worked as well. Taken together, these studies speak for the feasibility of using training to promote charismatic leadership (cf. Conger, 1992). Furthermore, there is descriptive, self-reported evidence in Study I that a large majority of participants used the inspirational speech after completing the training program.

There is clear evidence that the trained variables improved much more than the untrained variables. In both cases, there were obvious differences in the effect sizes comparing T1 and T2. Trained variables were much more effective (averaged 1.04 and .98) in contrast to the control variables (.32 and .27 for Study I and II, respectively). The average effect sizes were also remarkably similar in the two studies. Thus, there is evidence that the training worked to improve the targeted behaviors but did not systematically influence the untargeted behaviors.

Study II helps us to understand an issue left open by Study I. It was unclear from the results of Study I whether the significant increase of the control variables was due to a Hawthorne (or placebo effect) or whether these results were due to the co-change of those control variables as a result of the training because they were not carefully selected a priori. For example, learning to use variation in loudness (one of the experimental variables) may have led to clearer pronunciation (one of the control

variables in Study I that improved because of the training). The control variables in Study II were selected empirically and this procedure ensured that we only used those variables that were clearly distinguishable from visionary and inspirational speeches. Because the control variables did not change significantly as a result of the training, the conclusion is supported that once variables are carefully selected, Hawthorne effects are negligible.

The use of the design, called nonequivalent dependent variable design by Cook and Campbell (1979, p. 118) or internal referencing strategy by Haccoun and Hamtieux (1994), had several advantages. We propose that our design is pragmatically applicable for evaluating company trainings because pseudotraining control groups are unlikely to be used. Our design deals with the problems of testing and Hawthorne effects adequately. The nonsignificant pre-post differences of the control variables in Study II suggest that there were no Hawthorne effects producing changes in the control variables. The design has the advantage to enable the testing of Hawthorne effects because simply participating in a training program, self-reflecting, and concentrating on one's own leadership behavior is applicable to all variables, the trained and the untrained ones. Study II also suggests that knowledgeable raters could reliably code the theoretical similarity of items to a construct—in our case, they could rate how similar public speaking constructs were to a visionary inspirational speech. We suggest the use of such a procedure in the future if a nonequivalent dependent variable design (Cook & Campbell, 1979) or an internal referencing strategy (Haccoun & Hamtieux, 1994) is employed. When using this design, it may also be useful for knowledgeable raters to code how well potential control variables allow the evaluation of Hawthorne or placebo effects (e.g., because people are just thinking or talking amongst themselves about their performance in a given task area)—this should be maximized—and to what extent they would be directly affected by the experimental manipulation. The latter could be optimized by taking items that are not affected by the experimental manipulation but that are nevertheless from the same general domain that is being trained.

We think, therefore, that this methodological design has the potential to be utilized more extensively. Many practitioners do not use control group designs because the rationale of a randomized and even quasi-experimental control group is difficult to communicate and meets resistance by managers. Firms and managers are hesitant to accept a random selection of participants to the groups and certainly are against pseudotraining control groups. The design used in our study is less invasive and a comparison between those that were trained and the control variables examines the arguably most important threats to internal validity: testing

and Hawthorne effects. Our average effect size comparing T1 and T2 of the trained aspects was 1.02 in Study I and .98 in Study II, which is rather on the high side compared with other training programs (Burke & Day, 1986; Carlson & Schmidt, 1999). The design can be employed whenever the evaluation is based on some kind of performance test. For example, when evaluating a lecture, it would be possible to compare questions on declarative knowledge that has been taught with questions on knowledge that has not been taught. Similarly, evaluations of skills training programs may take advantage of the nonequivalent dependent variable design (e.g., in the area of social skills training).

The approach of action training is interesting in its own right. However, our studies were not designed to test different training concepts against each other. Although we theoretically contrasted the action training approach from behavioral modeling (Latham & Saari, 1979)—an approach that is highly influential and of proven usefulness (Burke & Day, 1986; Tannenbaum & Yukl, 1992)—we do not view these two approaches as mutually exclusive. There are far more commonalities than differences. One of the differences may be that the action theory approach does not explicitly employ behavioral models. However, we do not want to argue that behavioral models are unimportant but rather that most have already been exposed to them.

Managers usually know of good inspirational speakers and charismatic leaders, but this remains mainly abstract knowledge, not connected to their actual behavior. Thus, the training of adults needs to provide good reasons for change, to motivate people to change their actions, and to connect the abstract knowledge of inspirational communication with the appropriate (speech) acts. It should be added, however, that participants watched each other and, informally, we observed that some form of emulation appeared (e.g., when participants learned to use good metaphors from each other or to visualize their visions). A second difference may be that the action training approach is more likely to use negative feedback than the behavioral modeling approach. Feedback is deliberately given in a matter-of-fact way along the list of principles presented to the participants. Action training maintains that feedback should provide full information, which requires knowing what went well and what did not. Future studies should compare action training with other training forms, such as behavioral modeling training as was done by Debowski et al. (2001) and Wood, Kakebeeke, Debowski, & Frese (2000) for training computer skills.

There are, of course, a number of limitations of these studies. The most important one is the lack of measuring long-term objective effects of the training on the units' performance and organizational commitment by the subordinates. We were not able to measure such data within

the setting of our studies. Other limitations were that our design did not allow controlling for history, maturation, and regression effects (Campbell & Stanley, 1963). A true control group design would have allowed testing for these effects. However, we think that these effects are unlikely to have played a large role in our studies. History and maturation need a certain time to unfold their effects—a study on the direct effects of training on performance usually only covers a brief amount of time—in our case, 1½ days for the training module. (Of course, history and maturation effects could have played a role for the transfer questions; however, these were just add-ons and were not central for our studies). Regression effects appear only when there has been some kind of selection, for example, if training would have only been given to “needy” managers. This was not the case in our studies; in contrast, the managers in both companies had demonstrated high potential and needed to participate in this training program to be able to advance their careers. Therefore, any regression effects should have made people become worse over time, and this would have made it more difficult to get significant results. Therefore, we think that the design used in our studies allows researchers to deal with the most important internal threats without using a control group.

One strength of these studies is that behavioral data was coded before and after the intervention in a real life training course and, therefore, goes beyond self-report data or data provided by the supervisors. In terms of Kirkpatrick's (1976) four evaluation criteria, we have systematically looked at Criterion 3: being able to demonstrate behaviors taught in the training. Our measures were better than self-reports by participants or reports by supervisors who may bias their reports because of their knowledge that a person participated in a training program. Thus, we think that in spite of their shortcomings, our studies showed that one factor of charismatic leadership—communicating a vision—can be improved with action training and that some training evaluation designs may be more useful to be practically employed within the constraints of commercial settings.

REFERENCES

- Anderson JR. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.
- Awamleh R, Gardner W. (1999). Perceptions of leader charisma and effectiveness: The effects of vision content, delivery, and organizational performance. *Leadership Quarterly*, 10, 345–373.
- Bandura A. (1986). *Social foundations of thought and action*. Englewood Cliffs, NJ: Prentice-Hall

- Barling J, Weber T, Kelloway EK. (1996). Effects of transformational leadership training on attitudinal and financial outcomes: A field experiment. *Journal of Applied Psychology, 81*, 827-832.
- Bass BM. (1990). *Bass and Stogdill's handbook of leadership*. (3rd ed.). New York: Free Press.
- Baum JR, Locke EA, Kirkpatrick SA. (1998). A longitudinal study of vision and vision communication to venture growth in entrepreneurial firms. *Journal of Applied Psychology, 83*, 43-54.
- Bruner JS. (1966). *Toward a theory of instruction*. Cambridge: Harvard University Press.
- 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.
- Carlson KD, Schmidt FL. (1999). Impact of experimental design on effect size: Findings from the research literature on training. *Journal of Applied Psychology, 84*, 851-862.
- Carroll JM, Mack RL. (1984). Learning to use a word processor: By doing, by thinking, and by knowing. In Thomas JC, Schneider ML (Eds.), *Human factors in computing systems* (pp. 13-52). Norwood, NJ: Ablex.
- Carroll JM, Mack RL, Lewis CH, Grischkowsky NL, Robertson SR. (1985). Exploring a word processor. *Human-Computer Interaction, 1*, 283-307.
- Conger J. (1992). *Learning to lead*. San Francisco: Jossey-Bass.
- Conger JA, Benjamin B. (1999). *Building leaders*. San Francisco: Jossey-Bass.
- Conger JA, Kanungo RN. (1998). *Charismatic leadership in organizations*. Thousand Oaks, CA: Sage.
- Cook TD, Campbell DT. (1979). *Quasi-experimentation: Design and analysis for field settings*. Chicago: Rand McNally.
- Cook TD, Campbell DT, Peracchio L. (1990). Quasi experimentation. In Dunnette MD, Hough LM (Eds.), *Handbook of industrial and organizational psychology, 2nd ed.*, (Vol. 1, pp. 491-576). Palo Alto, CA: Consulting Psychologists Press.
- Debowksi S, Wood RE, Bandura A. (2001). Impact of guided exploration and enactive exploration on self-regulatory mechanisms and information acquisition through electronic search. *Journal of Applied Psychology, 86*, 1129-1141.
- Den Hertog DN, Verburg RM. (1997). Charisma and rhetoric: Communicative techniques of international business leaders. *Leadership Quarterly, 8*, 355-391.
- Dormann T, Frese M. (1994). Error training: Replication and the function of exploratory behavior. *International Journal of Human-Computer Interaction, 6*(4), 365-372.
- Driskell JE, Willis RP, Cooper C. (1992). Effects of overlearning on retention. *Journal of Applied Psychology, 77*, 615-622.
- Duden. (1994). *Reden gut und richtig halten. (Talks—giving them well and correctly.)* Mannheim, Germany: Dudenverlag.
- Dvir T, Eden D, Avolio BJ, Shamir B. (2002). Impact of transformational leadership on follower development and performance: A field experiment. *Academy of Management Journal, 45*, 735-744.
- Fey H. (1993). *Sicher und ueberzeugend praesentieren. (Presenting confidently and convincing.)* Berlin, Germany: Walhalla.
- Frese M. (1995). Error management in training: Conceptual and empirical results. In Zuccheromaglio C, Bagnara S, Stucky SU (Eds.), *Organizational learning and technological change* (pp. 112-124). Berlin: Springer.
- Frese M, Albrecht K, Altmann A, Lang J, Papstein PV, Peyerl R, et al. (1988). The effects of an active development of the mental model in the training process: Experimental results on a word processing system. *Behaviour and Information Technology, 7*, 295-304.

- Frese M, Zapf D. (1994). Action as the core of work psychology: A German approach. In Triandis HC, Dunnette MD, Hough LM (Eds.), *Handbook of industrial and organizational psychology*, (Vol. 4, pp. 271–340). Palo Alto, CA: Consulting Psychologists Press.
- Frijda NH. (1986). *The emotions*. Cambridge: Cambridge University Press.
- Geyer ALJ, Steyrer JM. (1997). Transformational leadership and objective performance in banks. *Applied Psychology: An International Review*, 46, 397–420.
- Gibson JJ. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.
- Goldstein IL. (1974). *Training program development and evaluation*. Belmont, CA: Wadsworth.
- Greif S, Keller H. (1990). Innovation and the design of work and learning environments: The concept of exploration in human–computer interaction. In West MA, Farr JL (Eds.), *Innovation and creativity at work* (pp. 231–249). New York: Wiley.
- Haccoun RR, Hamtieux T. (1994). Optimizing knowledge tests for inferring learning acquisition levels in single group training evaluation designs: The internal referencing strategy. *PERSONNEL PSYCHOLOGY*, 47, 593–604.
- Hacker W. (1992). *Expertenkönnen. Erkennen und Vermitteln* [Expert skills. Identification and teaching]. Göttingen: Hogrefe.
- Hacker W, Skell W. (1993). *Lernen in der Arbeit* [Learning at work]. Berlin: Bundesinstitut für Berufsausbildung.
- Heimbeck D, Frese M, Sonntag S, Keith N. (2003). Integrating errors into the training process: The function of error management instructions and the role of goal orientation. *PERSONNEL PSYCHOLOGY*, 56, 333–361.
- Hofsten CV. (1985). Perception and action. In Frese M, Sabini J (Eds.), *Goal directed behavior: The concept of action in psychology* (pp. 80–96). Hillsdale, NJ: Erlbaum.
- Holladay SJ, Coombs WT. (1994). Speaking of visions and visions being spoken: An exploration of the effects of content and delivery on perceptions of leader charisma. *Management Communication Quarterly*, 8, 165–189.
- Holding HD. (1965). *Principles of training*. Oxford, UK: Pergamon Press.
- Hollon SD, DeRubeis RJ. (1981). Placebo–psychotherapy combinations: Inappropriate representations of psychotherapy in drug–psychotherapy comparative trials. *Psychological Bulletin*, 90, 467–477.
- House R, Spangler WD, Woycke J. (1991). Personality and charisma in the U.S. presidency: A psychological theory of leader effectiveness. *Administrative Science Quarterly*, 36, 364–396.
- Howell JM, Frost PJ. (1989). A laboratory study of charismatic leadership. *Organizational Behavior and Human Decision Processes*, 43, 243–269.
- Jentzsch K. (1993). *Rhetorik*. (Rhetoric.) (2nd ed.) Frankfurt, Germany: Haag & Herchen.
- Judd CH. (1908). The relation of special training and general intelligence. *Educational Review*, 36, 42–48.
- Kirkpatrick DL. (1976). Evaluation of training. In Craig RL (Ed.), *Training and development handbook* (pp. 18.1–18.27). New York: McGraw Hill.
- Kirkpatrick SA, Locke EA. (1996). Direct and indirect effects of three core charismatic leadership components on performance and attitudes. *Journal of Applied Psychology*, 81, 36–51.
- Kissling-Institut. (1998). Unpublished seminar papers. Zurich, Switzerland
- Klein KJ, Bliese PD, Kozlowski SWJ, Dansereau F, Gavin MB, Griffin MA, et al. (2000). Multilevel analytical techniques: Communalities, differences, and continuing questions. In Klein KJ, Kozlowski SWJ (Eds.), *Multilevel theory, research, and methods in organizations* (pp. 512–553). San Francisco: Jossey-Bass.

- Kluger AN, DeNisi A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119, 254-284.
- Kozlowski SWJ, Salas E. (1997). A multilevel organizational systems approach for the implementation and transfer of training. In Ford JK, Kozlowski SWJ, Kraiger K, Salas E, Teachout MS (Eds.), *Improving training effectiveness in work organizations* (pp. 247-287). Mahwah, NJ: Erlbaum.
- Latham GP, Saari LM. (1979). Application of social-learning theory to training supervisors through behavioural modelling. *Journal of Applied Psychology*, 64, 239-246.
- Locke EA, Latham GP. (1990). *A theory of goal setting and task performance*. Englewood Cliffs, NJ: Prentice-Hall.
- Lowe KB, Kroeck KG, Sivasubramaniam N. (1996). Effectiveness correlates of transformational and transactional leadership: A meta-analytic review of the MLQ literature. *Leadership Quarterly*, 7, 385-425.
- Martocchio JJ. (1994). Effects of concepts of ability on anxiety, self-efficacy, and learning in training. *Journal of Applied Psychology*, 79(6), 819-825.
- Miller GA, Galanter E, Pribram KH. (1960). *Plans and the structure of behavior*. London: Holt.
- Neisser U. (1985). The role of invariant in the control of movement. In Frese M, Sabini J (Eds.), *Goal directed behavior: The concept of action in psychology* (pp. 97-109). Hillsdale, NJ: Erlbaum.
- Niemiec RP, Sikorski MF, Clark G, Walberg HJ. (1992). Effects of management education. *Evaluation and Program Planning*, 15, 297-302.
- Olbort K. (1998). Unterhalten Sie Ihre Zuhörer [Entertain your listeners]. *Bank Magazine*, 3, 88-89.
- Parry SB, Reich LR. (1984). An uneasy look at behavior modeling. *Training and Development Journal*, 38(3), 57-62.
- Revens RW. (1982). *Active learning*. London: Blond and Briggs.
- Sackett PR, Mullen EJ. (1993). Beyond formal experimental design: Toward an expanded view of the training evaluation process. *PERSONNEL PSYCHOLOGY*, 46, 613-627.
- Salas E, Cannon-Bowers JA. (2001). The science of training: A decade of progress. *Annual Review of Psychology*, 52, 471-499.
- Schwertfeger B. (1997). Auf der bremse. (On the breaks.) *Wirtschaftswoche*, 11, 170-173.
- Semmer N, Frese M. (1985). Action theory in clinical psychology. In Frese M, Sabini J (Eds.), *Goal directed behavior: The concept of action in psychology* (pp. 296-310). Hillsdale, NJ: Erlbaum.
- Semmer N, Pfäfflin M. (1978). *Interaktionstraining*. [Interaction training] Weinheim, Germany: Beltz.
- Shamir B, Arthur M, House RJ. (1994). The rhetoric of charismatic leadership: A theoretical extension, a case study, and implications for research. *Leadership Quarterly*, 5, 25-42.
- Skell W. (1972). Analyse von denkleistungen bei der planung und praktischen durchführung von produktionsarbeiten in der berufsausbildung. [Analysis of thoughts in the planning and practical performance of production workers in their trade education]. In Skell W (Ed.), *Psychologische Analysen von Denkleistungen in der Produktion* [Psychological analysis of thoughts in production] (pp. 13-100). Berlin: VEB Deutscher Verlag der Wissenschaften.
- Analyse von Denkleistungen bei der Planung und praktischen Durchführung von Produktionsarbeiten in der Berufsausbildung
- Tannenbaum SI, Yukl G. (1992). Training and development in work organizations. *Annual Review of Psychology*, 43, 399-441.

- Volpert W. (1971). *Sensumotorisches lernen* [Sensorimotor learning]. Frankfurt: Limpert; Fachbuchhandlung für Psychologie.
- Volpert W, Frommann R, Munzert L. (1984). Die Wirkung allgemeiner heuristischer regeln im lernprozeß-eine experimentelle studie [The effects of general heuristic rules in the learning process: An experimental study]. *Zeitschrift für Arbeitswissenschaft*, 38, 235-239.
- Von Papstein P, Frese M. (1988). Transferring skills from training to the actual work situation: The role of task application knowledge, action styles, and job decision latitude. In Soloway E, Frye D, Sheppard SB (Eds.), *CHI'88 Conference proceedings. Human factors in computing systems. Special issue* (pp. 55-60). Washington: Association for Computing Machinery.
- Wood RE, Kakebeeke BM, Debowski S, Frese M. (2000). The impact of enactive exploration on intrinsic motivation, strategy, and performance in electronic search. *Applied Psychology: An International Review*, 49, 263-283.

APPENDIX

Curriculum of the Training Module: Inspirational Speech

Day 1, 9:00-11:30

- Short introduction into the concept of "vision"; each participant developed a vision
- Preparation of first inspirational speech (10 minutes)
- First inspirational speech, videotaped, participants filled out questionnaire (T1)

Day 1, 11:30-13:00

- Intensive lecture on "What is a vision," examples of visions for departments, discussion of the principles of a good vision
- Each participant's vision was discussed and enhanced
- Lecture on theory and empirical studies of charismatic leadership
- Presentation of the principles of a good inspirational speech that are based on theory and empirical studies
- Individual preparation of a second inspirational speech (10 min)

Day 1, 14:30-18:30

- Second inspirational speech, videotaped, after each speech, participants filled out questionnaire (T2)
- Videotape is shown and trainer and participants provide positive and negative feedback
- Principles for transfer are discussed and participants are asked to think of a time when they could communicate their vision to their employees

Day 1, 19:30-21:00: Other program, unrelated to charismatic and visionary leadership

Day 2, 8:00–9:30

- Preparation of third presentation of inspirational speech (10 min)
- Inspirational speech is videotaped, participants filled out the questionnaire (T3)
- After this, continuation with other modules of the course

