MEASURING THE EFFECTS OF INTERPERSONAL TRAINING FOR THE WORKFORCE

Mario Martinez
Department of Educational Leadership
University of Nevada, Las Vegas

Cecilia Maldonado
Department of Educational Leadership
University of Nevada, Las Vegas

Derek Lester
Department of Educational Leadership
University of Nevada, Las Vegas
MEASURING THE EFFECTS OF INTERPERSONAL TRAINING FOR THE WORKFORCE

Abstract

This study assesses the cumulative effect of a training curriculum focused on teamwork, which was delivered to technical employees in a multinational organization. Employees were given a 10-item survey prior to the training and four months after the survey. Several aspects of the literature were incorporated into the design of the curriculum and analysis of the results: content was tailored toward the employees’ environment, management was asked to promote the training to contribute to a positive climate, and assessment was deferred by four months to maximize the chance that employees might implement the lessons from the training into practice. The results for the 10-item survey show a significant pre- and post-training difference, and implications and findings are discussed and reconciled with the literature.
Introduction

The enduring references to and utilization of Kirkpatrick’s (1987) foundational work on the four levels of training evaluation (reaction, learning, transfer/behavior, and results) some fifty years after its derivation is impressive indeed. Kirkpatrick’s levels of evaluation continue to draw attention not only because of their simplicity (Alliger & Janak, 1989) but because there remain challenges to evaluation beyond reaction, particularly as it pertains to nontechnical training. The nontechnical training we refer to throughout this research pertains to any training that speaks to interpersonal exchanges and interactions between and among people or groups. This interaction and exchange may occur between employees working within the same organization; or it may refer to interactions between employees and customers. Martinez (2004) refers to training and education that enhance such skills as “people-related training,” and we will adopt that nomenclature here for ease of use and conciseness.

Professionals who conduct people-related training often distribute course evaluations at the conclusion of their courses. At best, these evaluations give an impressionistic sense of the learning that has taken place, with the promise and perhaps commitment of behavioral change. Whether that promised behavioral change leads to any true improvements in such areas as teamwork or interpersonal relationships is difficult to discern. In practice, pre- and post-testing of effectiveness in people-related training is rare and may be complicated by confounding variables that occur between the initial assessment, the treatment, and the post measurement. Thus, most evaluations of training that are disseminated by facilitators and professional trainers often only get at Kirkpatrick’s first level, which is simply a participant’s reaction to the training, as guided by that participant’s thoughts, feelings, and attitudes after having immediately received the training.

The spirit of the research we report here investigates whether people-related training led technical employees in one organization to more positively perceive the level of engagement and interaction between and among individuals and groups within their company. The results we report are only one component of a larger study, but they are important enough to merit individual analysis and interpretation.

In our research, we asked employees to rank characteristics to assess the internal organizational relationships between individuals and groups before and after the training, as a proxy for measuring the intended results of our training intervention. More importantly, our training efforts were intended to help employees contribute to the strengthening of those organizational characteristics exemplified by the questions in the assessment.

We developed and distributed a simple ten-item survey to assess employee perceptions prior to the training intervention. Employees then received a full day training session on topics in three areas: teamwork, basic communications skills, and relating one’s interactions and actions to the purpose of the organization. Each component of training was specifically designed to include questions to motivate employees to action (Kirkpatrick’s transfer/behavior level). For example, curricular design of the teamwork module covered standard elements of the topic, such as stages of teamwork and team interaction, but application questions were purposefully included as well. In the case of the teamwork curriculum, participants were asked to identify a specific action they could take (as a result of what they learned in the training) in the next week to begin improving
teamwork. They were also asked to identify a specific action they must stop (as a result of what they learned in the training) so as to improve teamwork.

Four months after we distributed the pre-training survey and conducted the training, we redistributed the survey. Our interest was in learning whether employee rankings on the pre-survey results differed from the post-survey results. We purposefully focused on nonexecutive employees (through survey distribution) so as to more accurately gauge perceptions of those who carry out the day-to-day operational work of the organization. Thus, our target respondents were technical employees and their immediate crew supervisors (all of whom work side-by-side with employees despite having some management responsibilities).

We conducted our survey in an organization that provides highly skilled maintenance services within the aerospace and aviation industry. Although we focused only on one location for purposes of access and cost, the company has major operations around the globe. Any location at any given time may be providing service to customers who are from anywhere in the world. This study supports the body of knowledge that advocates for “training for impact” (Robinson & Robinson, 1989, p. 10) rather than “training for activity” (Ibid, p. 10). The former is results oriented and driven by business needs, helps the organization achieve its goals, provides people with the skills and knowledge they need to improve their performance, assesses readiness of the work environment to support learned skills, has management accepting the responsibility for a supportive work environment that encourages skill transfer, and has measurable results that can be tracked (Ibid, p.10).

Many studies in the field of training and development, whether they attempt to derive competency models or assess training at a particular level, disproportionately focus on management-level employees. We were interested in those frontline employees and their supervisors who provide the core services of the organization and, in many instances, come face-to-face with the customer. Given the above context, our research question for this applied research was simply: Does people-related training for frontline employees, which is geared toward increased engagement and teamwork, lead to measurable increases in those areas throughout the organization?

Although the intent of our research is applied in nature, we discuss relevant literature as it informed the context of the study and the research design. We follow with the results and corresponding analyses, and then conclude with final recommendations.

Study Context

The quest to prove training effectiveness has a long history in the research literature (Baldwin & Ford, 1988) and has been a salient issue with professionals in the field. Researchers have studied the effects of training in a variety of fields and in a variety of ways. Flavin (1997) used pre- and post-test measures to monitor the effect of cross-cultural training for nurses and found positive differences in terms of participant reaction to the training. Tansik and Driskill (1977) examined the changes in racial attitudes of supervisors at a military installation after a required training course and found small improvements. Spooner, Baker, Harris, Ahlgrim-Delzell, and Browder (2007) were able to create a true experimental design to test the effectiveness of a one-hour course intended to help professors enhance lesson plans for their students, with the results showing a significant difference between the control and experimental groups. In a
similar vein, Cauble and Thurston (2000) studied the effects of interactive multimedia training on social work students and found significant differences between pre- and post-test results measuring user knowledge and a feeling of competence.

Many of the studies that populate the research literature testing for the effects of an intervention deal with technical-type training and/or conduct their pre- and post-test surveys immediately before and after the training. The nature and length of training clearly presents opportunities for how effectiveness is measured, but longer term effects—especially as they pertain to behavioral change and implications for organizational results—are certainly more difficult to isolate. Furthermore, many facets of people-related training do not lend themselves to immediate evaluation. The information that people acquire in the training and then hopefully translate into behavioral change requires time, which means that results will not be immediate either. The literature does, however, provide insight into additional areas of training and the factors that affect it. In an intriguing meta-analysis, Colquitt, LePine, and Noe (2000) found that training motivation, at the individual participant level, explained incremental variance in training outcomes beyond cognitive ability. Reese and Miller (2006) found that a career development course increased self-efficacy for those who took the training. Indeed, the effects of training on self-efficacy are a common theme in the training literature (Libermann & Hoffman, 2008; Schwoerer, May, & Hollensbe, 2005).

Other studies relevant to our applied approach speak not to the training itself but outside factors that may affect it. Hattie, Biggs, and Purdie (1996) in a review of 51 studies found that training must be done in context. A higher degree of learning and transfer is probable when the design and conduct of training incorporates the contextual work situations and environment of those who undergo the training. The organizational climate and culture are also key factors to consider when determining whether skills acquired during training will be applied to tasks and one’s daily function (Tracey, Tannenbaum, & Kavanagh, 1995). Certainly, managers and those who commission the training contribute to the overall climate and culture of the organization and must somehow communicate the importance of training and how it factors into the success of the organization. In sum, the effectiveness of training is dependent on the content, its relevancy to the participant, and the environment in which the participant is to operationalize his or her newly acquired skills.

Research Design

Our study design and the actual delivery and assessment of training were very much informed by the literature. In a sense, our approach was a combined research and action project, as we relied on published knowledge to guide our design and delivery of the content and the collection and analysis of results. Tracey, Tannenbaum, and Kavanagh (1995) found that work environment is critical to the transfer of training. To proactively address work environment, we contacted management and leadership of those to receive the training to secure commitment to the training. As a means to build awareness, we also stressed to management the need to explicitly support and encourage those who would be going through the training. Management took the additional step of communicating support by agreeing to address each group on the day of training, explaining how it fit into the organization’s overall goals.
Work environment was not the only consideration integrated into our training solution. Hattie, Biggs, and Purdie (1996) advocate conducting training in context, meaning that the curriculum and strategies and methods used within it should draw on tasks, terms, and situations familiar to the organization in which participants conduct their work. In an effort to contextualize the training, a needs assessment was conducted in the form of phone interviews with a quarter of the participants prior to conducting the training session. In these interviews, we asked participants about prior training experiences and personal examples in the workplace that might suggest positive and/or challenging team interactions. Through these interviews, we were able to modify the curriculum so that training tools (such as scenarios) used the language of the organization and reflected realistic situations that the participants might face during an actual workday.

The analysis strategy also benefited from the literature. Libermann and Hoffman’s (2008) investigation of training transfer of service quality training for bank clerks did not aim to conduct post-tests immediately after the training. The researchers chose to wait twelve weeks after the completion of the training to conduct a post-test, a logical choice since service training involves interpersonal interactions that can only be assessed after participants have a chance to implement them in the workplace. Following this lead, we distributed our post-test surveys well after the actual training (four months). Our content was comprised of teamwork, basic communications skills, and relating one’s interactions and actions to the purpose of the organization. All of these topics involve people-related knowledge and skills, so the lapse between pre- and post-testing was, like the Libermann and Hoffman (2008) study, appropriate. We also let the research question drive our analytical approach. Studies that compare pre- and post-results for the same group are what Cook and Campbell (1979) refer to as a one-group pretest-posttest design. Schwoerer, et. al. (2005) conducted a one-group pretest-posttest design to compare the self-efficacy of recruits who had undergone sales training, and as part of their design presented descriptive statistics and used a t-test to compare results between the two periods. Likewise, we were interested in pre- and post-training survey results between the time before training happened and four months after—comparisons that can also be made with descriptive statistics and a t-test.

After settling on the operational details of training curriculum, delivery and subsequent analysis, we finalized the logistics of survey design pertaining to pre- and post-testing. We assembled a list of twenty statements to gauge the effectiveness of communications that occur internally in the organization as well as communications with external customers. The instrument was an adaptation of one used in a prior study by Martinez (2008), in which he calculated Cronbach alphas for the resulting four factors he found after conducting a factor analysis on the data. Three of the four alphas were above the standard .7 threshold, indicating a strong degree of reliability, from an internal consistency perspective. Although the survey items were slightly revised to fit the context of the company we studied, the test results from the original instrument (Ibid.) combined with our joint deliberations on the items gave us confidence that our adapted instrument contained a sufficient level of reliability and face validity to proceed with our investigation.

Previous work addresses the internal and external dimensions that characterize an organization (Cameron & Quinn, 1999), and this was a fundamental concept embedded in
our instrument. We also wanted our questions to address the idea that individual and organizational functionality depends on both technical and interpersonal effectiveness, something found in the leadership (Blake & Mouton, 1985) and competency literatures (Green, 1999). Green’s competency model, for example, equates core competencies with the more technical aspects of the organizational environment. The model recognizes, however, that individual leadership characteristics (which are highly interpersonal in nature) are necessary for optimal organizational performance. The task versus people emphasis found in Blake and Mouton’s work is certainly one of the great contributions of the leadership literature and at the very least reinforces the notion that attention to technical and interpersonal aspects of the organization are both necessary. Although we did create an instrument with twenty statements, we focus here on the first ten items of our survey instrument. The training was explicit in that it concentrated on issues of teamwork and internal communication, something the first ten items of the survey were intended to gauge. Table 1 shows the ten items.

Table 1
List of Items to Rate Pre- and Post-Training

1. Team members communicate in a positive, verbal manner.
2. Business processes within the company/organization are efficient.
3. If two people in a team disagree about something, they are usually able to work it out without involving management.
4. The people I work with are very capable of doing their jobs; their skills and abilities match their job tasks.
5. Employees in our organization relate well to their managers.
6. I understand how my job contributes to keeping customers happy and satisfied.
7. People in my organization work to solve problems instead of taking things too personally when a disagreement occurs.
8. If I see an area that needs improvement, I feel like I can make a suggestion and people/managers will listen to my suggestion.
9. People communicate well across different departments/groups.
10. People who work in my group/company have high levels of training, expertise, or experience in the area they work.

Importantly, our intent at this stage of the research project was not to run any analysis to establish groupings or construct validity for the survey items, but we did design the questions within the context of the literature (as noted above). We also reviewed and drafted the instrument with an eye toward simplicity and thus face validity (does this seem like a reasonable measure), particularly because many of the technical employees in the training had to feel comfortable with the language and tone of the questions. The odd numbered items, as we conceived them, speak to interpersonal communication within the organization while the even numbered items relate to technical aspects of internal functionality that may influence communication and teamwork. Our chief concern, however, was that the questions in total represented a sense of how people might feel about teamwork and communication, and whether they conveyed a sense of understanding of how individuals’ jobs contributed to the organization.
We deemed our ten survey item list as a reasonable starting point for the research, given our audience and intent. We did not wish to overwhelm our audience with too many questions. Those receiving the training were not used to taking surveys, answering emails, or working on a computer as a routine part of their jobs. Thus, we developed a user-friendly questionnaire, with the intent of handing out the survey in person to immediately collect the results. Employees were given instructions pre- and post-training and asked to answer the questions honestly. We explained that we were interested in their opinions regarding their organization. Employees were asked to rate the ten statements (Table 1) about their organization, according to the following 5-point scale, which ranged from: (5) Very strongly describes my company to (1) Does not describe my company. Respondents were given the choice to opt out, or not rate a statement.

### Analysis and Results

We collected 103 surveys from training participants prior to the actual training intervention. The surveys were distributed the day of the training, but prior to the presentation of any content. The 103 responses include technical employees and immediate supervisors but not the eight members of the executive management team of the company under study. Four months later, upon our return, we gathered 96 post-training surveys. Again, the surveys were distributed and collected in person, prior to a subsequent training session. Given that we had collected seven more surveys in the pre-training phase, we randomly eliminated seven surveys from the pre-training data to simplify our analysis. Thus, out of 142 non-executive employees, we analyzed pre- and post-training surveys from 96 employees, accounting for 67.6% of our target population.

We began the analysis with a descriptive look at the items on the survey, to get an initial sense of whether participants felt that the training resulted in improvement on the associated survey measures. Table 2 shows the mean and standard deviations for each survey item, pairing pre-training survey results with post-training survey results. Each result is for the total group (n=96), as our analysis focuses on the group level. The pre-training descriptive statistics for each item appear first, followed by the post-training statistics.
Table 2
Paired Sample Descriptive Statistics

<table>
<thead>
<tr>
<th>Item Pairing</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pair 1</td>
<td>3.64</td>
<td>.95</td>
</tr>
<tr>
<td>SMEAN(pre1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMEAN(post1)</td>
<td>3.80</td>
<td>.65</td>
</tr>
<tr>
<td>Pair 2</td>
<td>3.65</td>
<td>.84</td>
</tr>
<tr>
<td>SMEAN(pre2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMEAN(post2)</td>
<td>3.86</td>
<td>.72</td>
</tr>
<tr>
<td>Pair 3</td>
<td>3.75</td>
<td>1.07</td>
</tr>
<tr>
<td>SMEAN(pre3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMEAN(post3)</td>
<td>3.75</td>
<td>.97</td>
</tr>
<tr>
<td>Pair 4</td>
<td>4.07</td>
<td>.89</td>
</tr>
<tr>
<td>SMEAN(pre4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMEAN(post4)</td>
<td>4.28</td>
<td>.67</td>
</tr>
<tr>
<td>Pair 5</td>
<td>3.90</td>
<td>.93</td>
</tr>
<tr>
<td>SMEAN(pre5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMEAN(post5)</td>
<td>4.09</td>
<td>.75</td>
</tr>
<tr>
<td>Pair 6</td>
<td>4.76</td>
<td>.64</td>
</tr>
<tr>
<td>SMEAN(pre6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMEAN(post6)</td>
<td>4.77</td>
<td>.42</td>
</tr>
<tr>
<td>Pair 7</td>
<td>3.46</td>
<td>1.01</td>
</tr>
<tr>
<td>SMEAN(pre7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMEAN(post7)</td>
<td>3.62</td>
<td>.72</td>
</tr>
<tr>
<td>Pair 8</td>
<td>3.75</td>
<td>1.12</td>
</tr>
<tr>
<td>SMEAN(pre8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMEAN(post8)</td>
<td>3.88</td>
<td>1.01</td>
</tr>
<tr>
<td>Pair 9</td>
<td>3.07</td>
<td>.97</td>
</tr>
<tr>
<td>SMEAN(pre9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMEAN(post9)</td>
<td>3.39</td>
<td>.84</td>
</tr>
<tr>
<td>Pair 10</td>
<td>4.01</td>
<td>.90</td>
</tr>
<tr>
<td>SMEAN(pre10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMEAN(post10)</td>
<td>4.13</td>
<td>.75</td>
</tr>
</tbody>
</table>

Two important observations emerge from the descriptive statistics. First, the standard deviation for every question was greater in the pre-training survey results than the post-training results. This result indicates a definite effect from the training in the sense that it signifies a convergence of how participants view teamwork, communication, and how they might attach meaning to their jobs. The obvious follow-up question is whether that convergence is in a positive or negative direction. An analysis of the paired means provides the answer. On every survey item the mean value for the group increased between the pre- and post-value (except for item 3, which remained the same), indicating increasingly positive views of organizational teamwork and communication, and how participants attach meaning to their jobs.
While the descriptive statistics are encouraging and indicate the training intervention was effective, we also ran a statistical paired t-test. Again, we analyzed the pre- and post-survey results for the entire group, and for this test, all items were pooled together. The collective analysis of the total group for all questions is simply intended to discern, at a macro-level, whether the training intervention resulted in any differences across the target people-related areas, from a statistical standpoint. The null and alternative hypotheses are thus:

\[ H_0 = \mu T_1 - \mu T_2 = 0 \]
\[ H_A = \mu T_1 - \mu T_2 \neq 0 \]
\[ \alpha = .05 \]

where,

\[ \mu T_1 = \text{the group mean for all survey items prior to the training} \]
\[ \mu T_2 = \text{the group mean for all survey items after the training} \]

Table 3 shows the results of the paired sample t-test for all combined survey items, for the group as a whole. As Table 3 shows, the paired test between pre-training and post-training yields a significant difference between the two timeframes, at an alpha of .05.

<table>
<thead>
<tr>
<th>Paired Samples Test</th>
<th>Paired Differences</th>
<th>95% Confidence Interval of the Difference</th>
<th>t</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Deviation</td>
<td>Std. Error Mean</td>
<td>Lower</td>
<td>Upper</td>
</tr>
<tr>
<td>Pair 1 PreAll - PostAll</td>
<td>-.151</td>
<td>.722</td>
<td>.073</td>
<td>-.297</td>
<td>-.004</td>
</tr>
</tbody>
</table>

Table 3 provides additional evidence that the training intervention had an effect on the target group. The statistical significance of the comparison test reinforces the suggested pattern of the training effect by the descriptive statistics, clearly pointing to similar analytical conclusions.

Although not part of our initial tests, we subjected the survey item results to additional comparisons. For example, we looked at each question individually and ran paired sample t-tests. We also grouped the odd-numbered questions together and ran paired sample t-tests and did the same thing for even-numbered questions. In the individual item analysis, we only found a significant difference between pre- and post-training results for one question (question 9) at the .05 level, though at a .1 level the majority of items showed significance. The pooling of odd- and even-numbered questions also showed significance at the .1 level but not at the .05 level.
In sum, the complementary results from the descriptive statistics and the paired sample $t$-test for all items pooled together indicate that the training intervention produced the intended results. The less conclusive but still suggestive follow up analysis on individual items and pooled odd- and even-number items also offers clues. The individual items when taken as a whole—for the entire group—cumulatively measure the effects of the training intervention that is not as readily detectable when isolating the items and subjecting them to analysis. It is also fully possible, in reflecting on the results of the pooled odd- and even-numbered analysis, that the survey instrument is in fact measuring one or two constructs (i.e., there is a relationship between items that possibly group together and measure the same construct) that when analyzed together produce significance. The results certainly show, though, that there is a cumulative effect, as measured by all ten items, of the training intervention on improving teamwork within the organization.

Conclusions

The results of our study empirically reinforce much of the conceptual wisdom in the literature on many counts. First, the significance of our analysis underscores the need to conduct people-related training within a context that is meaningful to the participants receiving that training (Hattie, Biggs, & Purdie, 1996). Robinson and Robinson (1989) support this concept. They state that “business results occur when skills taught in a training program are applied on the job, yielding improved performance” (p.11). While general content in areas of teamwork and interpersonal interactions are integral to the functionality of any organization in today’s world, customizing that content to the participants’ environment is perhaps even more important for those who do technical work.

The nature of work for our participants is, by definition, quite tangible in practice, as they perform maintenance services on equipment and machinery. The abstract nature of people-related training heightens the need to make the curriculum concrete and relevant to any technical employee so that it more readily meshes with their natural inclinations for thinking, learning, and acting. As a result of the needs assessment we conducted, we were able to take great care to customize the content by integrating role plays, scenarios, examples, and discussion questions which reflected the participants’ contextual work environment. We utilized industry related language and examples throughout the course workbook The purpose, of course, was to teach the concepts inherent in teamwork-related curriculum as it applied to their environment, helping the transfer of learning. One of the vehicles facilitators and trainers can use to create the appropriate curriculum context is to interview participants prior to the workshop, which is how we were able to customize our curriculum. These interviews provided a great deal of “intelligence” that helped us more readily connect with the participants during the delivery process by integrating their language and everyday situations into the formal and informal aspects of the training day.

Second, we believe our results were significant because we were sensitive to the cultural and climate issues that are so important to the transfer of training results (Tracey, Tannenbaum, & Kavanagh, 1995). We gained leadership and management support for the training, and top-management introduced every session and stressed the reasons why the training was taking place. In every session, we explained after this introduction that the
leader would not be participating in this particular session because we wanted uninhibited discussion, sharing, and learning among employees. Participants were receptive to this explanation since we addressed it openly and directly, so as to avert a feeling that management is mandating training but not participating in it. We felt strongly that management should set the tone at the beginning of each session, and by doing so it would more strongly contribute to a positive climate whereby employees would be more receptive to the training and understand its worth.

Finally, by following Libermann and Hoffman’s (2008) lead, we did not rush to evaluate the training. Any evaluations immediately after training would have only gotten to Kirkpatrick’s reaction level. The results of people-related training become manifest over time, if that training is effective. By waiting four months after the training to disseminate our evaluation, we consciously complied with Libermann and Hoffman’s approach to, in effect, let the training “take root” if it was going to. Of note also is Kirkpatrick’s (1998) list of guidelines for evaluating behavior. In this list he states in the second guideline to “allow time for behavior change to take place” (p. 49).

In studies of training or any other intervention, there will always exist threats to validity (internal or external) in terms of proving whether improvements were in fact linked to the initial intervention. Many of these threats are inherent to the study of training—especially people-related training—but the clues that our analysis provide suggests that training does make a difference. We also consulted the literature in terms of methodological design and analysis to strengthen the claim that training was a factor in improvement if the results bore it out (which it did, as given by our analysis). Clearly, our survey showed significant differences pre- and post-training, on a cumulative level. Further analysis remains to determine whether individual constructs exist, and to distill how and if such constructs result in different effects for how employees put the concepts into practice to better themselves and the organization they work for.
References


