Sample and Design

Two hundred two undergraduate students in business and psychology were recruited from two southeastern universities to participate in the study. The students were offered a chance to win a $50 cash prize in a random drawing after the study in exchange for their participation. In addition, all students received some extra course credit. Twenty-eight of these participants were involved in the pilot study; the remaining 174 subjects participated in the actual laboratory experiment. One subject was removed from the final analysis for non-compliant responding on the instrument. Therefore, 173 subjects were included in the final analysis. Demographic data for these participants are presented in Tables 1a through 1e (in Appendix A).

A 2 x 2 x 2 factorial design was used to test the hypotheses. The independent variables included the specificity of information (high vs. low) in the justification, the expertise of the trainer (expert vs. non-expert), and the involvement of the subjects in the outcomes (high vs. low).

Procedure

Overview. The experiment was conducted in the context of a training session. Participants were told that the purpose of the experiment was to test the relative effectiveness of two types of job training on a clerical task: the use of human trainers versus interactive, computer-based training. The rationale stated that,
A team of management researchers at Virginia Tech are interested in the relative effectiveness of different types of training for a particular job task. The two types of job training being compared are the use of “live” human trainers and interactive, computer training. They have designed an experiment to test the effectiveness of these two types of training and will need participants. If interested, you will be trained by one of these two methods to perform a clerical job task. The experiment will require at least one hour (but no more than 1.5 hours) of your time. If you participate, you will be entered into a lottery for a cash prize. The winner will receive $50. Your participation will help contribute to our understanding of job training.

Actually, there was no computer training group; all participants were assigned to a group led by a human trainer who implemented the experimental manipulations. This scenario was chosen for a rationale because of its plausibility, and the growing interest in computer technology. Most importantly, it was designed to provide a cover story that would misdirect student perceptions from the true purpose of the study and to provide a basis for the negative outcomes.

**Task.** A catalog searching task used in previous studies of organizational justice was used in this study (Cobb & Frey, 1996; Flinder, 1994; Greenberg, 1983; 1987). The task in this study required participants to act as mail order clerks where they processed customer orders that had been phoned in earlier to an office products firm. The actual task required the participants to transcribe all relevant product information to the appropriate order form, and to provide the missing product number which was found in the catalog.
This task was chosen not only because it simulates a realistic work situation, but because participants typically lack familiarity with it. Subjects, therefore, have no preconceived ideas about the various performance levels possible on this type of task. This was important because performance feedback was manipulated by the trainer as part of the experimental design.

Procedure. Participants, in groups of 15 to 25, were randomly assigned to a training session that served as an experimental cell across the three independent manipulations (expertise, outcome involvement, and justification specificity). Each group was scheduled to meet with a member of the research team at a designated classroom for assignment to a training condition, followed by the actual training session.

By way of overview, each training session was comprised of five distinct periods: the introduction, the training, the performance period, the reaction assessment period, and the debriefing session. In the introductory period, students initially met with the researcher/experimenter and received an introduction to the study, a rationale, and a training condition assignment. They also were provided a statement about the trainer’s qualifications. This served as the expertise manipulation. In the second “training” period, the trainer taught the students to perform the catalog task. In the performance period, the trainer conducted the outcome involvement manipulation, put the students to work at the task for a brief period of time, scored and assessed the group’s performance, and then conducted the justification specificity manipulation. In the reaction assessment period, the researcher relieved the trainer and administered a questionnaire to assess the participants’
reactions. In the final debriefing period, the trainer returned to the room to debrief the student participants about the true purpose of the experiment. These training periods are discussed in greater detail below.

Introduction period. In the introductory period, students met with the researcher/experimenter for assignment to a training condition. Upon arrival, the researcher provided a more detailed rationale and explained that this group had been assigned to a human trainer. He then provided a brief description of the trainer’s qualifications, which actually served as the expertise manipulation (high vs. low experience). This manipulation was conducted by the researcher to keep the trainer “blind” to the expertise condition. After briefing the students about the trainer’s expertise, the researcher called the trainer into the room, and excused himself.

Training period. During the training period, the trainer introduced the students to the catalog task and then trained them to process the customer orders. Before beginning the actual instruction, the trainer indicated that the students should find an office products catalog and three large manila envelopes marked, “Training Phase”, “Work Period”, and “Post-work Period” at their desk. He instructed them to put aside all materials except the “Training Phase” envelope and catalog. He then had participants examine the contents of the envelope, which included transcripts for the training phase and an office products

---

1 The trainer was role-played by the principle investigator in each session. The researcher was role-played by one of several graduate students.
order form marked “SAMPLE”. The actual training involved a demonstration of the correct procedure for processing each order using an overhead projector. Correct processing required “clerks” to copy the order number, customer name, the product description, quantity, and price of the items ordered to the sample office products order form. The clerk then had to search for and provide the “missing” product number in the catalog to complete the processing of the order. Since the task was relatively simple, the trainer offered suggestions for performing the task most efficiently. After the trainer demonstrated the first order, the participants processed the second order on their own. After a brief check of the participants’ work, the trainer demonstrated a more difficult order. This was followed by participant processing of a similarly difficult order. The trainer again checked the group’s work a second time, answered any questions, and instructed them to return the training materials to the envelope.

Performance period. At that point, the trainer conducted the performance period which involved four distinct events: the trainer announced the outcome involvement manipulation; then put the students to work for a brief period; afterwards, he scored and assessed the group’s average performance; and finally, he implemented the justification specificity manipulation. These events are detailed below in sequence.

2 Pilot testing revealed that students perceived the task to be too simple to require “training”. By providing tips for faster processing, students realized the benefits of the trainer. Moreover, this addition helped preserve the cover story and rationale.
First, the trainer prepared the subjects for a 10 minute work period. He asked that each participant pretend to work as a member of a team comprised of the present participants, employed by a busy office products warehouse. He announced that their team’s average performance on the task would be compared to the average performance of all previous groups that had been trained by computer. The trainer then identified the criterion their performance would be judged by:

There are several ways we could assess your performance; we could look at the quantity of orders processed, the accuracy of the orders, and/or the legibility of your completed orders. All these criteria are important to this task. But due to obvious time constraints, we would be unable to score your forms according to all three elements, so I decided to consider quantity only. Therefore, the best training method will be determined according to which group averages the greatest number of processed orders.

Following the criterion announcement, the trainer introduced the incentives involved (outcome involvement manipulation), instructed them to remove the customer transcripts and office products forms from the envelope marked “Work Period”, and put them to work at the task in a 10 minute, timed performance period.

After the 10 minutes, the subjects were stopped, and their completed forms were gathered. The trainer indicated that the assessment of the forms would take 5-10 minutes,
and that the experimenter would help. In order to pass the time, the trainer asked the participants to complete an attendance form that also contained their lottery tickets. In addition, the trainer announced that as a favor to another professor, the principle investigator had agreed to ask that the participants complete a pretest of a questionnaire unrelated to this experiment found in the “Post-work Period” envelope.

At that point the trainer left the room and joined the experimenter in an adjacent room to score the forms. A simple average of the quantity of orders completed was obtained without regard to errors. After 6-7 minutes (the time for most students in a group of 15 to finish the survey and attendance form), the trainer returned to the room, to provide performance feedback.

Upon entering the room, the trainer asked if the attendance form had been completed by everyone, and then announced that with the help of the researcher, the forms have been scored and their team’s performance assessed. At that point, the researcher entered the room (exactly 20 seconds after the trainer entered) and asked the trainer if he could speed things along, so that there would be enough time to administer the

3 The pilot test revealed that students found it implausible that the trainer could adequately score the forms by himself in such a short period of time.

4 This questionnaire contained some individual difference measures thought to be important to the ELM. These measures were included for exploratory purposes and were not formally analyzed for this study.
questionnaire. This announcement was performed loud enough to be heard by all participants. The trainer looked at his watch and nodded approval. The researcher then exited the room. Subsequently, the trainer announced that he would like the participants to complete a job training evaluation form while he simultaneously provided their performance feedback. He indicated that the evaluation form is usually done last, but if they start now if would enable everyone to finish the experiment on time. While passing out the form, he announced that the research team also wanted him to explain the reason for using the catalog task. He then recited a brief monologue about the appropriateness of the catalog task with regard to the increasing number of mail order firms, and the fact that the task could be easily learned within the time constraints of the experiment. If students looked up while the trainer was speaking, he suggested that they keep working while he talked. The monologue was thus delivered to give everyone a chance to get started on the evaluation form and to correct off-task behavior (i.e., participants who stop working to listen to the trainer) prior to the performance feedback and justification.

The purpose of the evaluation form was to provide a distraction for the subjects during the justification. It was felt that since a low involvement condition would be difficult to obtain in an experiment of this nature (where performance is assessed), it was necessary to provide a distraction to all participants that would tax, but not overwhelm their ability to process the information. The rationale was that those most involved in terms of outcomes would be more motivated to resist the distraction and listen carefully to the justification. According to the ELM, central processing is a function of motivation
(i.e., involvement) and ability. Distraction is a commonly employed technique to reduce one's ability to process new information (e.g., Foti & Hauenstein, 1993; Petty & Brock, 1981). This enhancement would provide a greater opportunity for peripheral processing.

After the talk about the appropriateness of the catalog task, the trainer provided the feedback and the justification. He announced:

Now as for your team’s performance — going strictly by the criterion “average quantity of completed orders”, your team performed better than the computer trained groups. Your team averaged ________ [actual average performance], while the computer trained groups averaged ________ [subtract 0.5 from the above average]. But, I decided to use an additional criterion to assess your team’s performance. After doing so, your team average dropped to ________ [subtract 0.5 from the computer group average], so it turns out that the computer groups did better. Let me give you an explanation for why I used another criterion.

Following that announcement a justification for the change in the scoring procedure was given (specificity manipulation). The trainer then thanked the participants for their time, announced that the experimenter needed the remaining time to administer a questionnaire, and excused himself. The experimenter immediately entered the room and administered the questionnaire containing the dependent measures.
Reaction assessment period. The experimenter distributed envelopes marked “Questionnaire Materials” to the participants and indicated that in order to assess their reactions to the job training simulation, he needed them to complete a final questionnaire.

The experimenter then instructed the participants to open the envelope and stated,

> You will find a questionnaire and a card marked “Human Subjects Research Committee”. The card is a routine form required by the department of management to ensure that all participants are treated fairly. I forgot to include these in the last session, so let’s complete that form first while it’s on my mind. Just complete the form and place it in this departmental envelope that I’m passing around now. No names on the card please; this must be done anonymously.

After all participants completed the HSRC card, they were instructed to place it in a large, manila, Management Department envelope which the experimenter passed around the room. The last student was instructed to seal the envelope and pass it back to the

---

5 This statement was included to enhance the perception that this form was a routine departmental requirement, which helped preserve the study rationale and avoid contamination.

6 These cards had a form number at the bottom corresponding to a number embedded in the questionnaire. This allowed the HSRC card to be later matched with the questionnaire.
experimenter, who then put it aside. Afterwards, the experimenter asked the subjects to turn their attention to the questionnaire. He instructed them to carefully read all instructions and respond to all items on the questionnaire. He then asked them to return the questionnaire to the envelope upon completion and to quietly remain seated until all others had finished. After most of the questionnaires had been completed, the trainer (actual experimenter) returned to the room to deliver a debriefing session.

In the debriefing session -- using open-ended questions -- the trainer assessed and recorded the student reactions to the experiment, especially with regard to the effectiveness of important manipulations. In addition, he explained that the experiment was actually intended to assess their reaction to distributive rule changes and the explanations given for such changes. It was explained that the “computer trained group” and the human subjects committee did not really exist. He explained that the card was used as just one measure of their reaction to the situation. It was emphasized, however, that their input was very valuable in understanding how distributive changes affect workers. Furthermore, he announced that everyone would be entered into the second lottery for $100 (that was originally promised only to the winners in the high outcome involvement condition), in addition to the $50 lottery entry promised to everyone as an

7 This was done primarily in the pilot test phase to assess the impact of the manipulations and to determine if changes needed to be made. In later sessions, the trainer asked for their general reaction.
incentive for participation. Finally, the trainer asked that all participants promise not to
discuss the experiment with anyone.

Independent Manipulations

The three manipulations of the independent variables in this study were
justification specificity, source expertise, and outcome involvement.

Justification specificity. In this study there were two levels of justification
specificity. The high specificity justification was an account that provided specific
premises to explain the basis for the change in the distributive criterion. The low
specificity, justification was an account that provided less specific or vague premises to
explain the basis for the change in the distributive criterion.

Unlike the typical ELM study, messages in the accounts literature are given as a
way of explaining apparent norm violations that are usually associated with angry
reactions and perceptions of unfairness. In developing the accounts for this study, it was
important that the context or logic behind them be perceived as reasonable; that the norm
violation was justifiable. In other words, the actual reasoning behind both justifications
needed to be “strong”. By maximizing and minimizing the specificity in the message, the
participant's ability to perceive the “strength” was then either enhanced or diminished. If
the reasoning behind the justification was weak, then any counterarguments developed by
the subject might overwhelm the justification leading to message rejection. Thus, a weak
rationale behind the justifications would actually produce a “boomerang effect” where a
highly specific justification leads to reduced account acceptance.
Originally, it was planned to cue subjects that they would be assessed on the quantity of orders processed in the 10 minute work period, and then change the distributive criterion to consider quality of work (i.e., errors) as well. However, three waves of pretesting revealed that students held fast to a strict, literal interpretation of the distributive criterion. In fact, they perceived the trainer’s establishment of the distributive rule to be a “contract”. Any deviation from that contract (e.g., the elimination of orders with errors) was rejected. In other words, no justification for this norm violation was acceptable. In order to provide a sufficient rationale for establishing an elaboration to the “contract”, it was decided that the justification imply that some students in the group had cheated by duplicating processed orders (which would inflate productivity estimates if not discovered), and that these duplicates would not count. By citing a norm violation by some students, this rationale allowed for a change to the contract. The pilot test indicated that this manipulation was successful.

This manipulation was conducted by the trainer at the end of the performance feedback session. The following accounts were delivered in the high specificity and low specificity justification conditions respectively.

High specificity. After going through the catalog orders, I noticed that most of you made some minor errors, such as an incorrect product description, or a misspelled customer name. The computer trained groups also made similar mistakes. But strangely, some of you made a more serious kind of error which we did not find with the computer trained groups. Some of you
simply copied the same order several times, which at first gave me a false
count of your productivity. I decided that it would not be fair to allow these
duplicated orders to count. So, after subtracting all orders that were
duplicated from the total number of orders completed, it turned out that the
average computer trained group performed better.

Low specificity. After going through the catalog orders, I noticed that most
of you made some minor errors. These were small mistakes that we
expected on a task like this. The computer trained groups also made similar
mistakes. But strangely, some of you made a more serious kind of error
which we did not find with the computer trained groups. In other words
more than one of you made a duplication error, which at first gave me a
false impression of your groups’ overall productivity. I decided that it would
not be fair to allow all the processed orders to count. So, after discounting
any completed orders with this type of problem, it turned out that the
average computer trained group performed better.

Notice that both of the above accounts addressed the same issue. However, the
first example was quite specific. It explained the types of errors found, and the explicit
method used to re-determine the winners. Most importantly, it specifically identified and
explained the problem of duplicated orders. In contrast, the second example was much
more vague in explaining these issues. In fact, it identified but did not specifically explain the “duplication error”.

Previous research has reported that the length of the message can impact persuasion by serving as a peripheral cue (Wood, Kallgren, & Preisler, 1985). To control for this potential confound, both accounts were made equivalent in length.

**Source expertise.** Expertise has typically been manipulated by Petty and his colleagues by varying the relevant credentials of the message source. Expert sources are generally introduced as having a great deal of experience and knowledge with professional credentials. In contrast, non-expert sources are described as having little relevant knowledge or experience regarding the task at hand. For example, Petty et al. (1981) manipulated expertise by announcing that a tape-recorded persuasive report on university curriculum policy was either produced by the Carnegie Commission on Higher Education chaired by a professor from Princeton University (high expertise), or by a local high school class (low expertise).

In this study the source expertise condition was implemented by the researcher/experimenter near the end of the introductory period. In the high expertise condition, the researcher provided the following description of the trainer.

Mr. [trainer’s last name] has a quite a bit of experience in job training. He worked for the human resources training division Anderson Consulting for four years. He has worked with clients such as IBM, AT&T and Hewlett-Packard. He is currently, pursuing his Ph.D. in
management while consulting on the side. He is serving as the trainer of this project because of his expertise in training.

In the low expertise condition, the researcher announced,

Unfortunately, the professional trainer who has been working with us could not be here today, so I asked [first name of the trainer] to fill in. He is a graduate student who is helping conduct the research.

Although, he is not really a trainer, and has no experience in job training, I am sure he’ll do an adequate job.

Since the role of the trainer required the greatest preparation time, it was played by the principle investigator in all conditions. In order to maintain “blind” conditions to avoid experimenter effects, the researcher announced the expertise condition in the introductory period while the trainer was absent. The trainer (i.e., the principle investigator) was not blind, however, to the outcome involvement and specificity conditions. In order to minimize any potential experimenter effects, the trainer concentrated on delivering both of these conditions exactly the same way each time. Pilot testing confirmed that the expertise manipulation was successful.

**Outcome Involvement.** Most of the studies of the ELM conducted by Petty and his colleagues manipulated “outcome involvement” (Johnson and Eagly, 1990). This construct has generally been operationalized by indicating that an undesirable event will either affect (high involvement) or not affect (low involvement) the message recipients. For example, a typical manipulation involves the argument that comprehensive senior exams should be given as a requirement for graduation from college. The high
involvement condition recommended an implementation date that would affect the participants, whereas, the low involvement condition stipulated a date that would not affect them. In this case, the manipulation can be classified as either outcome or issue involvement. This study also manipulated outcome involvement.

In this study, the outcome involvement manipulation occurred just prior to the performance task following the training phase. Participants in the high involvement condition were told,

Before you begin, I want to make things interesting for you. If this group averages more completed orders than the computer trained groups, then each of your names will be entered into a second lottery; but this lottery will be worth $100. That means you will have two chances to win some cash if you outperform the computer trained groups, and you could win up to $150.

Participants in the low involvement condition were told nothing about this extra incentive. Thus non-winners in the high involvement condition would lose an opportunity to win $100 cash by the change in performance criteria, whereas non-winners in the low involvement condition would lose nothing.

**Manipulation Checks**

Manipulation checks were provided for each of the treatments (expertise, outcome involvement, and justification specificity) using a 7 point Likert-type scale (anchored by “Strongly disagree” and “Strongly agree”). Each measure was checked for internal
consistency using Chronbach’s alpha. These coefficients are presented in Table 2 (in Appendix A) along with their intercorrelations with other measures in the study. Each measure is presented below with its respective items.

**Justification Specificity.** The measure for the manipulation check on the specificity level of the justification included the following two items.

“The trainer’s explanation for how he determined the winning group was specific.”

“The trainer’s explanation for how he determined the winning group was vague.” (reverse scored)

Internal reliability for this measure was adequate (α = .86).

This manipulation check (justification specificity) had the greatest potential to create demand effects, because the items suggested that the explanation for how the winners were chosen had some importance to the study. Furthermore, it could have suggested to subjects that the trainer’s change in the distributive rule was actually a planned event. For this reason, this measure was placed near the end of the Likert-type items in the questionnaire.

**Expertise.** A three item measure was used as a check for the expertise manipulation. These items assessed expertise in terms of knowledge and experience. These items included:

“Generally speaking, the trainer was knowledgeable.”

“I believe the trainer lacked expertise.” (reverse scored)
Internal reliability for this manipulation check was adequate (‡ = .79).

**Outcome involvement.** A four item measure was developed to assess the level of involvement in the training simulation. Internal reliability proved to be poor for this global measure of involvement (‡ = .61). Reliability analysis suggested that there may be two types of measures among the four items; a general measure of involvement (three items) and a more specific measure of outcome involvement (one item). When broken out this way, internal reliability improved for the general measure of involvement (‡ = .70). Items comprising the general involvement measure include items 1, 2 and 4. Item 3 assessed the importance of the outcomes.

“I felt personally involved in this experiment.”

“Participation in this experiment was important to me.”

“The lottery cash prize was important to me.”

“I found myself quite motivated to perform well on the task.”

**Central versus peripheral processing.** It is only assumed that when involvement is high, justification recipients will process the message centrally; and when it is low, they will process it using the peripheral channel. Even though planned involvement manipulations can lead to relative differences in cognitive motivation, the low involvement manipulation might still be high enough to stimulate central processing. Therefore, it is necessary to include some indicator of the relative effort expended on message processing in the involvement conditions. Two items were included to help identify which channels were used in processing the message.
“I tried to closely evaluate the trainer's feedback regarding my group's productivity.”

“I spent little effort thinking about the trainer's feedback about my group's performance.” (reverse scored)

Internal reliability for this measure was also adequate (ζ = .79).

**Dependent Measures**

The dependent measures for this study were developed using items from previous studies (where possible) which had demonstrated adequate reliabilities. Since most of these studies have relied on one or two-item measures, additional items were developed to insure multiple indicators for each construct. All dependent variables were measured using a seven point Likert-type scale (anchored by “Strongly disagree” and “Strongly agree”), with the exception of commitment to the task, commitment to the trainer, and complaints (these scales are discussed in the Behavioral Outcomes section). Most of the measures were pre-tested on a representative student population, and checked for internal consistency using Chronbach’s alpha (ζ). Graduate students and randomly selected participants in the pretesting were interviewed to assess their thoughts and perceptions of the measures to insure face and content validity. After pretesting, measures were revised and included in two laboratory training sessions that served as a pilot test for the study (all final measures were tested in the pilot study). The alpha coefficients and the intercorrelations for these measures are included in Table 2 (in Appendix A).

**Justification acceptability.** Three items were used to assess the acceptability of the justification for the change in the distributive criterion. The first two statements in the
following item pool were adapted from Bies, et al. (1988) measure of account adequacy. They reported a reliability index of .92 for the two items.

“I found the trainer’s explanation for how the winner was chosen to be adequate.”

“The explanation the trainer provided about determining which group won was not sufficient.” (reverse scored)

“The explanation given by the trainer for choosing the winning group was acceptable.”

Internal reliability for this measure was high (α = .93).

This measure, like the justification specificity manipulation check, had the potential to create demand effects, because it cues the participant that the change in the distributive criterion was a planned event. For that reason, it was also placed near the end of the Likert-type items.

Fairness outcomes. Three types of fairness were assessed in this study: Interactional fairness, procedural fairness, and distributive fairness. Measures for each of these constructs are discussed in turn.

Interactional fairness. A three item measure was used to assess interactional fairness. The first two items in the following item pool are adapted from Bies and Shapiro’s (1987) measure of interactional fairness. Reported reliabilities were adequate (Chronbach α = .74 in study 1; Chronbach α = .80 in study 2).

“The trainer treated me fairly during the experiment.”
“Overall, the trainer acted openly and honestly.”

“It seemed that the trainer failed to treat me with consideration.” (reverse scored)

Internal reliability for this measure was adequate (‡ = .75).

Procedural fairness. A four item measure of procedural fairness was used. The first two items below were adapted from a measure of procedural justice used by Bies et al. (1988). The reported reliability in that study was high (‡ = .97).

“Overall, the decision-making process used by the trainer regarding how best to evaluate my group’s performance was fair.”

“The trainer’s decision about the winning group was made in a way that was unfair.” (reverse scored)

“I believe the process the trainer used to decide my group’s success was fair.”

“The procedures used by the trainer to determine whether or not my groups won were fair.”

Internal reliability for this measure was also high (‡ = .94).

Distributive fairness. According to Greenberg (1986), there are two types of distributive fairness: fairness associated with the actual distributive decision itself, and fairness associated with the outcome of the decision such as pay or some other reward. Both types of distributive fairness were assessed in this study (hereafter referred to as “distributive fairness: assessment”, and “distributive fairness: outcomes”. Each measure included three items.
The three items below were concerned with distributive fairness: assessment. These items focused on the participant’s reaction to the trainer’s assessment of their group’s performance. The first item was adapted from a measure developed by Cobb and Frey (1996), with reported high internal reliability (指望 = .94)

“My trainer’s overall evaluation of my group’s productivity was fair.”

“The trainer’s assessment of our performance was fair.”

“The trainer’s final decision regarding my group’s performance was fair.”

Internal reliability for this measure in this study was high (指望 = .97).

The next three items pertained to distributive fairness: outcomes. These items focused on participant reaction to the incentives or outcomes allocated to them.

“The amount of the lottery cash prize given for this experiment was fair.”

“The rewards for this experiment were lower than what I should’ve received.” (reverse scored)

“Generally speaking, the outcomes I received from participating in this study were unfair.” (reversed scored)

Internal reliability for this measure was adequate (指望 = .79).

Satisfaction outcomes. Two measures were included to assess respondent satisfaction with the trainer and the task respectively.

Satisfaction with the trainer. The following four items were designed to assess the overall satisfaction participants experienced working with their trainer. The first item was
based on a measure used by Cobb and Frey (1996); it had high internal reliability ($\alpha = .89$).

“I was satisfied working for my trainer.”

“In general, I was happy with the way the trainer conducted himself.”

“Overall, I thought my trainer was less than satisfactory as a leader.”

(reverse scored)

“My experience with this trainer was acceptable.”

Internal reliability for this measure was also high ($\alpha = .88$).

Satisfaction with the task. The following three items were designed to assess each participant’s satisfaction with the catalog ordering task.

“In general, I found that working on the catalog task was a satisfactory experience.”

“Working on the assigned task left much to be desired.” (reverse scored)

“I felt that the catalog task was enjoyable.”

Internal reliability for this measure was adequate ($\alpha = .81$).

**Behavioral outcomes.** Three behavioral outcome measures were also included in the questionnaire. These measures differ from the others because they provide an opportunity for the participant to make behavioral choices with regard to potential participation in future studies, or to register a complaint regarding the present experiment.

Commitment to the task. This behavioral variable was measured by asking the respondents near the end of the questionnaire if they would be willing to participate in
further training studies in the near future using the catalog task. A five item scale (anchored by “Not at all” and “Extremely”) was used to assess the extent of their interest. It was stated that if enough interest was generated, they would be contacted again.

Commitment to the leader. This measure was assessed following the above item on the questionnaire. The participants were asked, “Given the opportunity would you be willing to work with the same trainer again?” Again, the five item scale (anchored by “Not at all” and “Extremely”) was used to assess each participant’s willingness.

Complaints. Complaints were assessed by providing the participants an opportunity to register a complaint with a departmental committee that purportedly monitors ethical treatment in experiments. A postcard-sized form entitled Department of Management, Human Subjects Research Committee, was provided by the researcher with the questionnaires. It stated,

This card is being provided to you by a committee established within the Department of Management with the purpose of insuring all participants in experiments conducted by its faculty are treated ethically and fairly. Should you feel that you were not treated in such a manner, we ask that you report it on this card, so that appropriate, corrective steps can be taken. Your response will be confidential.

Respondents were asked to check one of two boxed anchored by items stating, “I experienced no problems in this experiment.,” or “I do not believe I was treated
appropriately during this experiment.” Furthermore, if respondents were reporting a problem, they were asked to rate the degree to which it was felt they were treated unethically, using a five point scale (anchored by “Slightly” and “Extremely”). This measure was treated as a six item scale by considering the check box labeled, “I experienced no problems in this experiment.” as an end point.

In order to minimize demand effects that might ordinarily be created by this type of measure, the researcher attempted to create the perception among the participants that this form was routine by indicating that he had almost forgot to include it the last session. Two steps were taken to reduce respondent inhibition. First, the “no problems” option was included on the card; and second, participants were instructed to place their card in a sealed departmental envelope distributed by the researcher immediately after completion.