Measuring the Impact of Technology Transfer Activity on Transportation Agencies

Transportation Research Board Paper 13-2601

Submitted July 31, 2012

Timothy Colling, Ph.D., P.E., Director (corresponding author)
Center for Technology & Training
Michigan Tech Transportation Institute
309 Dillman Hall
Michigan Tech University
Houghton, MI 49931
906-487-2102
Fax: 906-487-3409
Email: TKCOLLIN@MTU.EDU

John Kiefer., P.E., Research Engineer II
Center for Technology & Training
Michigan Tech Transportation Institute
309 Dillman Hall
Michigan Tech University
Houghton, MI 49931
906-487-2102
Fax: 906-487-3409
Email: JAKIEFER@MTU.EDU

Number of words: 7,166
Number of figures & tables: 1
ABSTRACT

The Kirkpatrick training assessment model is the prime method of assessing the impact of training events. This method was primarily developed for execution in a corporate environment and researchers may encounter difficulty employing it to assess the impact of non-compulsory technology transfer events. Complications of applying the Kirkpatrick method in these types of events can result from the following for factors: 1) the inability to predict likely behavior changes in participants, 2) the widely varying and sometimes unknown backgrounds of participants, 3) the lack of a control group for learning assessment tests 4) unavailability of participants to be observed for behavior changes.

This paper discusses a number of techniques that have been developed and have undergone field trials to supplement the Kirkpatrick assessment method and overcome these complications for non-compulsory technology transfer events. The techniques include; using personal information of participants to provide context in the analysis phase, the use of a crossover design in developing learning assessments; and the use of ethnographic research and data analysis techniques to provide a framework for Level Three and Level Four Kirkpatrick assessments.
Measuring the Impact of Technology Transfer Activity on Transportation Agencies

NEED FOR MEASURING IMPACT
Performance measurement is not a new concept, but with the passage of MAP-21 the concept is rapidly moving into the forefront in every activity that transportation agencies engage in. Transportation agencies are increasingly pressed to show impact and benefit for their programs and activities or face budget cuts and elimination of services in this time of economic stress. Technology transfer activities such as those performed by the Local and Tribal Technical Assistance Programs (LTAP/TTAP), the National Highway Institute (NHI) and others, have long been regarded as having a positive impact on the transportation community, however without a formal method to show impact, the benefit of the activity may be considered by some as subjective.

BACKGROUND
For over 30 years, the most commonly used body of work dedicated to the assessment of training and education programs was developed by Donald L. Kirkpatrick (1). Kirkpatrick developed a four level learning assessment model in 1954 when he was completing his dissertation at the University of Wisconsin. Kirkpatrick’s learning assessment model was not the original focus of his graduate studies, but rather a construct developed to assess the effectiveness of the training program that was the focus of his dissertation. In the decade following its original development, Kirkpatrick’s learning assessment model gained wider visibility in the education community after he was asked to author a series of journal articles for the American Society for Training Development (ASTD) describing the assessment model (2).

The framework for Kirkpatrick’s learning assessment model dates back to the 1950’s, but it was not published until Kirkpatrick authored the first edition of the book “Evaluating Training Programs – The Four Levels” in 1993. The original work was expanded upon when Donald Kirkpatrick and his son James Kirkpatrick co-authored a third edition of the elder Kirkpatrick’s first book “Evaluating Training Programs – The Four Levels”, which is the most complete work on the subject to date.

Assessments for each successive level of the model are required to show positive response before it can be concluded that the training event has a chance to influence the next level of the model. For example behavior changes can’t be attributed to a result of the training if participants did not learn (3). Generally speaking, the higher level assessments involve an increased difficulty in measurement.

Level One - Reaction
Reaction assessments seek to determine how participants initially respond to a training event and can discover factors that may have hampered or assisted participants in learning the material presented in a training event. These assessments are intended to provide an indication of the conditions that were present during the event that may contribute to or detract from learning. They can be helpful in determining areas that may need revision or more emphasis during the training.

Workshop and instructor evaluations are a typical method for providing feedback in a reaction assessment. Evaluation instruments used for reaction assessments may require participants to respond to a series of short answer, multiple choice or Likert scale questions relating to the items of interest. Questions typically focus on the training topic, training environment, and instructor communication skills.
Level Two - Learning

Learning is a measure of the information that participants absorbed as a result of the training and result in a change in knowledge, attitude or skills (3). Learning assessments typically are in the form of tests given to participants before a training event, then directly after the conclusion of instruction. Comparing scores between pre-training and post-training tests gives a measure of the amount of learning that took place during the training. In the Kirkpatrick model, learning assessment scores can be tracked and recorded individually for each participant, or can be recorded as an aggregate for the entire class without the need for tracking individual pairs of scores.

Level Three - Behavior

Training can provide the first two of the four factors necessary for a change in behavior, which are:

1. A desire to change
2. Knowledge of what to do and how to do it
3. The correct work environment
4. A reward for change (3)

Positive behavior changes as the result of training are typically the desired target of training events. Behavior is typically assessed by interviews or direct observation of the subject before and after training.

Level Four - Results

The results assessment quantifies how the change in behavior observed in the Level Three assessment impacts the “business” where training participants are employed. Results are generally assessed in terms of dollars or some other overall measurement of benefit to the parent organization.

COMPLICATIONS WITH KIRKPATRICK’S METHODS IN TECHNOLOGY TRANSFER

Unknown Behavior Changes

Authors have criticized Kirkpatrick’s model, although there have been few suggested improvements or alternatives to it. Critics of Kirkpatrick’s learning assessment model contend that it is overly simplistic and fails to take into account factors related to the context of the training or factors unique to each participant in cases where the expected behavioral changes did not result. Bates suggests that this focus on a predetermined behavioral change fails to account for other factors that may produce a benefit when a training event did not create changes in behavior that it was designed to influence. These other unmeasured factors may produce value to the employer, but the Kirkpatrick assessment may discount or miss them because they were not the focus of the training event (1).

Predictability of behavior changes allows evaluation instruments to be developed prior to the training event and allows for a variety of pre-training assessments. In cases where it is difficult to know or predict the likely change in behavior that will result from the training event, static survey or interview questions are not likely to be a satisfactory evaluation tool. These evaluation instruments can only provide data on the topics that they were designed to investigate.
and are not likely to elicit data on unknown factors. In technology transfer activities, it is usually not known how training participants will use the information provided or what changes will occur as a result of attending.

There is a significant body of literature available that provides guidance for creating surveys when the topics of interest and the range of expected responses are known (4) (5) (6). However, there is very little literature that provides guidance on conducting interviews that can generate the type of data that is necessary for Kirkpatrick’s behavior assessments when the potential behavioral changes are unknown.

**Unknown and Variable Participant Background**

The focus of Kirkpatrick’s model is a corporate environment where training is compulsory, the training participant’s background is known and the training participants are observable for long periods of time. This is not the case for many technology transfer or research outreach programs designed to share innovative ideas and practices, and draw audiences from a diverse background of transportation agencies with diverse knowledge backgrounds.

Workshop and instructor evaluations in a standard Kirkpatrick evaluation typically used for reaction assessments assume that the audience’s education background is known and relatively homogenous, something that would be common in a corporate environment. Without a context for the audience’s knowledge and education background, reaction assessment responses can provide conflicting or misleading information. For example an engineer attending a training class for non-technical staff may indicate the class material was “too basic” on a reaction assessment, when in fact the class was taught at a very high technical level for the target audience. Failure to recognize the context with which the response was given can provide for an erroneous conclusion. Many times in these types of events it is not known who will be attending the training until the day of the event.

**Lack of A Control Group**

A major challenge with using pre- and post-training learning assessments is insuring that the testing instruments are balanced, so one assessment is not more difficult than the other. Administering identical learning assessments for the pre- and post-training assessment has the advantage of insuring that both testing instruments are exactly the same, however this can result in what is known as the practice effect, where participants tend to focus their attention during training on collecting the answers to assessment questions that they could not answer in the pre-training assessment, resulting in elevated post-training assessment results (7). Kirkpatrick recommends using a control group of individuals that do not attend the training to ensure that pre- and post-training assessments are balanced without repeating questions in both assessments (3). However, gaining access to a control group that accurately reflects the demographic makeup of participants in a training event is difficult outside of a corporate environment..

**Access to Participants to Observe Behavior Changes**

Changes in behavior take place over a period of time ranging from immediately following the training event to several years later (3). Timing of behavior assessments can be critical when attempting to document a change. Assessments completed too soon may show little or no change in behavior because sufficient time has not passed for the participant to act on the new knowledge that they received from the training. Conversely, behavior assessments that are completed too long after a training event can be difficult to relate back to the training event.
INTRODUCTION TO ETHNOGRAPHIC RESEARCH TECHNIQUES

Ethnographic techniques provide a data collection and analysis methodology that can be used to complement Kirkpatrick’s learning assessment. These techniques can be used for data gathering on all levels of Kirkpatrick’s model, but are especially valuable for collecting data where the range of outcomes cannot be determined in advance of the assessment. Researchers in the field have described the prime role for ethnographic methods to “describe unexpected or unanticipated outcomes” (8).

The technique of ethnographic research is deeply rooted in the fields of social and human behavioral science, and traces its beginning to anthropological research, where investigators sought to understand and document cultural beliefs and attitudes in primitive societies. This technique has been widely applied in research areas where investigators have no preconceived ideas about how the target of the investigation behaves or functions. The usefulness of ethnographic techniques has allowed them to be applied to many other fields of study. Examples of ethnographic research techniques crossing over into fields other than the social sciences include medical research (9), education (10) and engineering (11).

The most significant work dedicated to ethnographic research is the Ethnographer’s Toolkit (12). This seven-book series consists of over 1600 pages of information on developing experimental designs, conducting field work and analyzing results from ethnographic research. The series was authored by Jean Schensul and Margaret LeCompte with contributions from Stephen Schensul.

 Ethnographic Research Designs

Ethnographic research designs are divided into three general groups: cross-sectional research, experiments, and controlled field studies. Controlled field studies are utilized when true control groups are not available. This design is considered a modification of the true experimental design and is sometimes referred to as a “quasi-experimental” design. Controlled field studies are typically used in practical applications where researchers want to know if a treatment or intervention creates a benefit, but where they cannot exercise control over the subject that may be found in a laboratory environment, such as a classroom or clinical environment (8). Controlled field studies are ideal for assessing the impact of technology transfer events.

 Ethnographic Interviews

There are three general classifications of ethnographic interview techniques: open-ended or "unstructured" interviews, structured ethnographic interviews (often referred to as ethnographic surveys), and semi-structured interviews (8).

Open-ended interview techniques are considered the most technically challenging form of interview data collection. In an open-ended interview the participant is not required to select from a series of predetermined responses, but rather can respond in any way they are comfortable. A researcher conducting an open-ended interview explores the research topic in depth, but also must be open to other topics as the interview proceeds. The primary purpose of this interview method is to provide a broad depth of knowledge where little is known. This technique also can be beneficial in building trust between the researcher and the subject. (8)

Open-ended interviews have several drawbacks. The interviewer has to be able to react and make decisions quickly when deciding to pursue new topics or lead the conversation back to the original topic. This "think on the fly" requirement of open-ended interviews can complicate the researcher’s job, because there is only limited time to consider the next interview topic and
the researcher needs to be able to quickly form an appropriately worded question. Data collected from open-ended interviews is subject to bias from the researcher. Bias includes unintentionally leading the interviewee or redirecting the conversation based on the researcher's view of the topic. Most interview subjects have a limited amount of time that they are willing to devote to interviews. Time limits complicate open ended interviews because subjects may speak at length on topics that are significant to them, but are of little importance to the researcher (8).

Structured ethnographic interviews restrict participants to a predetermined selection of questions and responses. Participants must choose from one of the offered responses and cannot formulate their own answers. The researcher must be fully familiar with the research topic and the participants to be able to create an appropriate range of responses for each question. Structured ethnographic interviews are different from standard surveys because their target is always the meaning of some element of culture. They are typically the last data collection technique used rather than one of the first data gathering techniques, which is the case in a standard survey questionnaire (8).

Structured interviews have several drawbacks. They are focused survey instruments, and in order to be effective they require significant experience with the population or group of interest. They can only discover information about known phenomena and cannot be used to explore new topics or factors. Structured interviews have the same complexities that semi-structured interviews have in terms of generating an appropriate range of responses. Follow up probes cannot be used with structured interviews, further adding to their inflexibility (8).

Semi-structured ethnographic interviews combine the flexibility of open-ended interviews with the discipline and focus of a survey. Semi-structured interview questions are pre-written in advance of the interview, however the questions do not have structured responses, so the interview subject can reply in any way. Researchers conducting semi-structured interviews can follow up with questions or comments designed to get the subject to expound on the topic of discussion. Follow-up questions are sometimes referred to as "probes". Semi-structured interviews are used when the researcher has a strong understanding of the subject and the factors of interest, but is not sure of the possible range of outcomes that may result (8).

Semi-structured interviews are created by identifying topics of interest for the study, then formulating questions that address the topics of interest. Interview questions must be non-leading, unbiased, and presented in language that is meaningful to the participants. Semi-structured interview are typically in long-answer form so they cannot be answered "yes" or "no". Completed interview questions are recorded on an interview schedule that gives guidance to the interviewer. Interview schedules provide background information for the researcher and draw attention to items of interest for each interview question. Probe questions are also listed in an interview schedule.

Ethnographic research techniques are ideal for learning assessments. A controlled field study design with qualitative ethnographic components provides a methodology and structure to complete Level Three Kirkpatrick training assessments and can provide data for Level Four assessments. A semi-structured ethnographic interview allows information on a specific topic to be collected while providing enough flexibility to allow the subjects to respond in any manner they see as appropriate. The open-ended nature of this technique fits with Kirkpatrick Level Three and Level Four evaluations in cases where the behavioral changes or results are not easily predicted prior to the evaluation phase. The “self reporting” nature of ethnographic interviews also overcomes the issue of being able to observe participants for behavior changes, which is a major concern with Kirkpatrick methods in non-compulsory technology transfer events.
ANALYZING ETHNOGRAPHIC DATA

Analyzing quantitative data produced by ethnographic studies is straightforward and similar to working with other types of quantitative data from other sources. Analyzing qualitative data from an ethnographic study that consists of hours of recorded interviews and field notes can be a challenge for a researcher who is more familiar with quantitative data. Analyzing qualitative data has several distinct steps: coding, pattern analysis, and interpretation (8).

Coding

Analysis of qualitative ethnographic data starts with review of the data and identification of the themes that are present. Sub-themes are added as necessary to fully describe possible ranges of responses in the data. The process of developing themes is called "pre-coding". Themes emerge in the data by searching for common or repeating phrases, events, activities, behaviors or ideas among participants or data sets and from the researcher's background knowledge of the research topic. As themes are identified they are assigned an alpha numeric code that represents the theme. Each code is recorded in a code book that provides a definition or set of conditions that an item must possess to be subject to a specific code, describes the range of responses for a specific code, and provides examples of items that meet the definition for the code, including exclusions to the definition (8).

When themes have been identified, defined and assigned a code, the researcher begins the process of coding by reviewing the field notes and data sets that the study has generated and coding key data that fits the predefined criteria. Data elements such as an excerpt from an interview are "tagged" with the specific codes that the data elements fit. Tagging can be done either manually by physically sorting data elements based on the codes they relate to, or more commonly, through the use of software (8).

Codes in qualitative ethnographic research are analogous to numbers in quantitative research. Each code is a marker that represents an instance of a defined event or concept that is present in ethnographic data. Codes can be used to represent anything that the research wishes to identify for later analysis that may be pertinent to the research, including events, concepts and people. Ethnographic data is coded to reduce the volume of data and to allow for methodic analysis of its elements. Coding schemes can either be generated specifically for the research project or a researcher may adopt one of a series of generic coding schemes that have been developed for many areas of cultural, social, economic and political research (8).

Pattern Analysis

Pattern analysis is the process of determining the relationship between coded items in qualitative research. Pattern analysis starts by arranging like-coded items into a hierarchy of groups. Schensul and LeCompte describe pattern analysis as:

"Something like the middle stages of assembling a jigsaw puzzle; once the player has found all of the orange pieces and all of the blue pieces, for example, or all of the pieces with a particular pattern on them, he or she then can begin to assemble those pieces into a coherent chunk of the design portrayed in the completed puzzle." (12)

Patterns can be discovered by several different methods including: declaration, frequency, omission, similarity, co-occurrence, sequence and hypothesis. Discovery of a pattern by declaration is common when open-ended or semi-structured interviews are used. In declaration a participant identifies the pattern for the researcher. The researcher is then able to verify the pattern through other data. Frequency analysis is another common method of identifying patterns. Events, items or responses that appear with regularity in data can point a
researcher to the existence of a pattern. A less frequently used pattern discovery method is omission. Researchers using this discovery technique look for items that are notably missing in the data set. A topic that is considered taboo in a modest or conservative society is likely to not be mentioned during interviews. Co-occurrence of coded data items can be an identifier of a pattern. When two coded items appear in concert with one another and each is a discrete item, a pattern may exist with the two items. Data items that appear in a temporal sequence may be indicative of a pattern. One event may precede or trigger a second event or the two may be unrelated. Researchers with significant background in a research area may be able to discover patterns using their previous working knowledge of the subject. For example, a researcher who is a member of a Native American Indian tribe may have insight regarding how a different tribe functions (12).

Computer Assisted Analysis

Recording field notes, reduction of data, coding and pattern analysis are all processes that historically required a significant effort when researchers relied on physical paper data management systems. Researchers employing ethnographic techniques have developed software to aid in the entire process of ethnographic research. The general term for the use of these software tools is Computer Assisted Qualitative Data Analysis (CAQDA). There are over 20 CAQDA software programs and there is a “Center for Computer Assisted for Qualitative Data Analysis” in England that is funded by the UK Social and Research Council (13).

CAQDA software can assist researchers in recognizing patterns in the raw data that is coded and stored in a database. Field notes, interview transcripts and other qualitative data from ethnographic research can be reduced to transcribed text documents that can be imported directly into a CAQDA software package. Data that has been entered into a CAQDA software package can be interrogated using advanced Boolean search functions that are capable of identifying instances of items within the respective data sources. These features are useful for identifying relationships between coded data. A group of coded data that exists as a pattern is stored as a “node” (14). Nodes can be further coded to identify key relationships between the elements in a node or relationships to other nodes. Nodes can also be organized and mapped to develop interpretations that support or disprove a hypothesis. Developing relationships (cause, effect, contingent upon) between nodes creates a model that describes the object of the research.

CAQDA software started in 1981 with the release of “NUD*IST” – the first qualitative data analysis package. In the 1980s and 1990s many other CAQDA software packages appeared, including “Ethno” and “Altis”. Several papers were published in the 1990s regarding use of CAQDA software, including articles questioning its use (15) and many more supporting its use (16) (17) (18). In 1998, the book “Computer Analysis and Qualitative Research” formalized the practice of utilizing CAQDA software into a series of methods (19).

The CAQDA software package used most for research is Nvivo 8. This software has been used for numerous research projects involving the analysis of qualitative data and is licensed by over 200,000 researchers in over ninety countries (20). The use of CAQDA software does not create quantitative data out of qualitative data, but it can aid in the development of cognitive constructs from the data and can aid in pattern recognition rather than relying on the researcher to develop these.
MODIFIED KIRKPATRICK ASSESSMENT FOR TECHNOLOGY TRANSFER
APPLICATIONS

Field Trials of the Modified Kirkpatrick Method
The modified Kirkpatrick assessment presented in this paper was originally developed for a FHWA-funded research project to assess the impact of two traffic safety training workshops, one targeted at local elected and appointed officials, and one targeted at local agency technical staff. Two separate field trials were assessed using the method, one for each training program, with each training program consisting of four repeat workshops each. The assessment methodology was developed out of necessity for that project because suitable methods were not readily available. In depth information, including examples of assessments instruments are included in the original project report (21).

Training Participant Data Collection
Registration of training participants prior to the training event is an important consideration for technology transfer and other training events that are going to be assessed for impact. Pre-registration provides time for participant contact information and basic demographic information (type of agency the person works for, job title, etc.) to be properly collected and entered into a database for later use. During registration participants can also be assigned unique identification numbers that can be printed on evaluation instruments. The numbers link to all contact, attendance, and survey and interview data.

Tracking assessment instruments with discretely tagged identification numbers allows individual responses to be linked to the individual providing the feedback. This link provides context of the participant’s background without the stigma of a participant seeing their name related to their specific responses. Identification number tagging is specifically important when there are discrete groups of participants attending an event where it may be beneficial to analyze the group separately from the main body of the participants. In the field trials of this assessment method, one of the training events was aimed at non-technical elected and appointed officials, however, many engineers attended the event in an attempt to determine how technical topics were being delivered to non-technical audiences. Without the ability to selectively remove the engineers’ responses the collected assessment data would have presented conflicting information.

Reaction Assessments
Workshop and instructor evaluations are intended to be completed by participants during or immediately following the training event. It is recommended that the evaluations be distributed at the beginning of the training event so that participants can take notes during the event as material is presented. In the field trials of the assessment method, basic questions relating to the venue, the presentation skills and clarity of the instructor, and overall agenda were included, which is common for a Kirkpatrick level one evaluation. Additional open-ended questions were included that asked participants to identify key concepts or ideas that they would consider taking back and implementing at their agency. Responses to the workshop and instructor evaluations should be used as a general indicator of the participant’s interest in the topics covered and provide direct information on understanding issues or concerns that may have an influence on learning during the event. Open-ended questions serve as a secondary source of data for possible behavioral changes to follow up on in later stages of evaluation.
Learning Assessments

Pre-training assessments and post-training assessments should be designed to evaluate knowledge gained by the participants in key training objective areas during the training event. Comparing the pre-training assessment results to a participant's post-training assessment results gives an indication of knowledge transferred to the participant during the training event.

The use of a crossover design for learning assessments can eliminate the need for an outside control group, which may not be practical in most technology transfer events. A crossover design allows participants to act as a control group for each other (22). This method is efficient because no effort is expended gathering and testing an outside control group. The only additional effort required for this method is related to organizing and managing two separate sets of pre- and post-training assessments so that approximately half the participants receive each of the versions of the assessment and the additional tracking.

A crossover design allows questions on the pre-training assessment and the post-training assessment to be unique from the participant’s perspective. The use of unique questions prevents participants from searching for answers that they did not know in the pre-training assessment in order to answer them correctly the next time they are presented, which has been shown to inflate post-training evaluation scores and create a positive bias in the assessment results due to the practice effect (7).

A crossover design requires the pool of learning assessment questions to be randomly divided into two equal sized sets (question set 1 and set 2). The two question sets are combined in different orders to create two assessment tests (Form A and Form B). Figure 1 illustrates the combination of test questions to create the learning assessments.

During delivery of the learning assessments, half of the audience receives assessment Form A and the other half receives assessment Form B. This results in half the population answering question set 1 as a pre-training assessment and the other half answering question set 1 as a post training assessment. The two test forms can be compared by aggregating participant correct scores for a specific question when it is asked as a pre-training assessment vs. when it is asked as a post-training assessment question. This will determine if there is bias in one of the question sets.

Figure 1: Workshop learning assessment instrument development.
Questions should be randomly assigned an order that they will appear in the assessment forms, and the correct answer for each question should be randomly assigned an order (A, B, C, D) in the response list. Both of these steps are important to avoid any unintended patterns in the evaluation instruments. It is recommended to use multiple choice responses for ease of interpretation and grading where possible and to include a “don’t know” response to learning assessment questions to discourage participants from guessing or leaving questions unanswered.

The proposed learning assessment questions should be user tested prior to finalizing the assessment instruments. A group of individuals who have no knowledge of the subject matter in the training event, and ideally have a similar education background as the target audience, should complete both the pre- and post-test instruments and have their responses scored. User testing will determine if the questions in the testing instruments are not clear or if they include leading language which may allow correct answers to be intuited by comparing user test results to the possibility of a random correct response. Questions that have an unusually high positive answer rate by the test group should be reexamined and revised as necessary.

Each set of pre- and post-tests tests should be coded with the participant’s identification number, the testing form code (Form A or Form B) and date of the technology transfer event. This information should also be recorded in a master database for later analysis. Coding of forms allows pre- and post-training learning assessment scores to be related back to an individual.

Pre-training learning assessments should be distributed as soon as participants enter the training session and collected before the session begins to preclude participants from gleaning information from the instructor when completion of the assessment form. Post-training assessments should not be distributed to participants until after the end of the training session or should be distributed in a sealed envelope that is not to be opened until after the session is complete.

During the field trials of the learning assessment, moderators indicated to the participants the intent of the evaluation, and assured them that regardless of responses the assessment would not reflect positively or negatively on the participant or instructor. This step was necessary to eliminate the possibility of participants tailoring their answers to get a desired result based on an assumption of gain.

**Kirkpatrick Level Three - Behavior and Level Four - Results Assessments**

Ethnographic techniques, such as ethnographic surveys, and semi-structured ethnographic interviews provide a data collection and analysis methodology that fits well into the Kirkpatrick assessment model, specifically for Level Three and Level Four Analysis. The techniques are well suited for learning assessments because they can be delivered in a post-training only format which is usually necessary for non-compulsory technology transfer events. The techniques also overcome the issue of being able to observe behavior changes by using the participant to self-inform on their own activity, which is a practice that has been used in anthropologic studies for decades.

**Follow Up Survey**

Post-training follow-up surveys should be delivered several months after a technology transfer event to gather general information about the participants and to identify attitudinal and behavior changes that researchers anticipate would likely result from the event. Ideally, follow-up surveys should be delivered after participants would likely act on information discovered during the technology transfer event, but before too much time has elapsed where the details of the event are not easily associated with the change in attitude or behavior. Unfortunately there is
no set guideline for this timing and optimal timing may vary greatly depending on the technology transfer topic. During the field trial of this methodology it was assumed that this period would be between six to twelve months after the event.

Follow-up surveys can be conducted as a live interview or as a static instrument for the participant to complete on their own. Follow-up surveys delivered with a live interviewer typically have higher response rates since it is necessary for the interviewer to call the participant; however, they require significant resources to complete. If surveys are delivered as a static instrument for self-completion, they should be provided in electronic and printed formats to accommodate user preference and account for the possibility that some participants may not have access to the Internet.

During the field trials of the learning assessment, follow-up surveys were delivered between six and eleven months after the training event was complete as a static survey instrument. These surveys provided direct data and also provided insight into areas where follow-up was warranted during live, semi-structured ethnographic interviews that followed the survey.

**Ethnographic Interviews**

Semi-structured ethnographic interviews should be used to search for attitudinal and behavioral changes that participants experienced as a result of the training that could not be easily predicted. Interviews are also critical in determining the results of behavior change for Level Four analysis. Ideally, every participant in a technology transfer event should undergo ethnographic interview. In reality however, this is usually not possible due to time constraints and participant willingness to undergo the process. Researchers should strive to collect as much feedback from participants as time and budget allows.

During the field trials of this methodology approximately 20% of the participants underwent full ethnographic interviews. Because only a sample of participants were interviewed, candidates were randomly selected. Interview questions were limited to keep interview times to less than 60 minutes since researchers assumed that participants would be less likely to participate and less likely to provide open-ended, honest feedback if the interview length was longer.

Semi-structured ethnographic interviews start with the development of primary questions and associated follow-up probes that are carefully worded to elicit discussion and detect changes in attitude, behavior or thinking as a result of the training. During the field trials the interview questions were reviewed and revised after a series of mock interviews with the research staff. Support information for all the questions was documented in an interview schedule and an interview protocol was developed to guide each interview so that they would be delivered consistently.

**Interview Mechanics**

When possible, all ethnographic interviews for a specific study should be conducted by the same research staff to ensure consistency with the interview protocol. The researcher conducting the interviews should also not be involved or associated in any way with development of the training materials or presenting the workshop. Isolating the instructor from the interview staff removes the “need to please” by interview subjects who may be sympathetic to the instructor or may be not willing to reveal negative information about the technology transfer event to someone involved in the presentation of the event.

In order to elicit truthful and honest responses, interview candidates should be assured that anything they say during the interview will have no effect—negative or positive—on the
interviewer or the instructor. Interview candidates should not have their responses directly attributed to them or their agency in any report as another measure to protect their identity and ensure honest responses.

During the field trials of this methodology, interviews were conducted over the phone. Phone interviews were more efficient from a time standpoint and minimized the impact to participants’ lives. Initial telephone contact with interview candidates was used to explain some basic details of the interview, to build trust with the interviewer, and to schedule the actual interview. Dialogue during the initial phone conversations was semi-scripted to deliver the same message repeatedly without revealing any leading information that may influence the interview. Interviews were conducted over the telephone at a time convenient for the subject. Interviews started identically with the interviewer reading scripted background information from the interview protocol and asking permission to record the interview for future transcription and analysis. Interviews were recorded using a phone line digital voice recorder. The recorded audio files were cataloged based on the interview subject’s database identification number.

Recordings of interviews can be transcribed into text using voice recognition software such as Dragon 10: Naturally Speaking®. Current voice recognition technology requires the software to be “trained” to a specific voice before it can accurately transcribe spoken conversation. Training a piece of software to recognize a voice involves speaking predetermined phrases to identify inflection and dialect of the speaker. Direct voice recognition cannot accurately be used for transcription of recorded interviews because completing the voice training process is not practical for each interview subject. In the field trials, the research team overcame this problem by training the voice recognition software to a transcriptionist’s voice. The transcriptionists listened to recorded interviews with a set of headphones while repeating the dialogue into a microphone for the software to interpret. As the software created a transcription of the spoken interview, the transcriptionist was able to visually verify the transcribed text and correct errors as they occurred. Review of the audio recordings and transcripts provided the opportunity to hear voice inflection, hesitancy, agitation, excitement, and other features that would not be detected from transcripts alone.

**Data Analysis**

Consideration should be given to the use of CAQDA software for analysis of interview and survey data. CAQDA programs, while slow to learn, greatly speed the process of pattern analysis and coding. All data collected for the field trials of this methodology were input into the program Nvivo 8®. This allowed interviews, surveys, learning assessments, instructor reviews and demographic information to be tied together in an all-encompassing database. The same researcher should ideally complete pre-coding and coding of data to reduce variability in coding methodology.

A custom coding scheme, developed for field trials of this methodology, was designed to identify and classify instances of attitudinal or behavioral change that could be attributed to information learned from the workshops, however, there are many generic coding schemes that have been produced that are suitable for use (12).

**RESULTS**

Recording and tracking professional background and demographic information of technology transfer event participants was shown to provide vital context in the analysis phase of an impact assessment.
The use of a crossover design in learning assessments was successful in measuring participant learning without the use of a control group or being subject to the practice effect due to reuse of assessment questions. The technique allowed question set bias to be assessed and corrected if necessary.

The use of ethnographic survey and interview techniques combined with qualitative data analysis techniques has been shown to complement the Kirkpatrick training assessment method well for assessing the impact of technology transfer events. During the field trials of this methodology, ethnographic techniques successfully identified significant evidence of behavioral shifts and positive outcomes as a result of a series of technology transfer events. The interview techniques were critical in collecting data for Level Three and Level Four Kirkpatrick assessments to quantify the impact of the events (21).

Ethnographic research and data analysis methods provide well-documented, structured, and rigorous techniques which can be used to modify the Kirkpatrick training assessment method from a tool that was primarily developed for a corporate environment, to a tool that has application in measuring non-compulsory technology transfer events.
Bibliography


Qualitative Social Research.


12
13