Gregory Lea
5.16.2009

Final Report:

Identifying Success Measurements for Default Call Center Agents

Improving Training Materials via Evaluation

Introduction

I am the manager of the instructional design team within the Training Department at a medium-sized mortgage servicing company. My responsibilities include developing the design strategies and setting the goals that are the framework for the learning content that the team provides to the rest of the company. As the manager of the instructional design team it is also my responsibility to ensure that the team is producing effective training materials and that the training itself is producing employees who are prepared to function in their jobs. I believe learners should be evaluated to provide insight into the effectiveness of the training programs.

As an organization we generally perform basic (Level 1 and Level 2; see the *Literature Review* section for more information) evaluations of learners, but it concerns me that the Training Department does not employ a more robust evaluative system in which a baseline level of performance is established and evaluations, including levels 1-4, are compared to the baseline as a method of measuring the effectiveness of a particular curriculum or learning strategy.

Most likely, many factors contribute to an employee's poor performance in a training program. This project's scope does not extend beyond my area of influence which is the design of a training program. My goal is to define success measurements to be used to evaluate employees. I will use employee evaluations to help identify weaknesses in our training programs.

Note: The Instructional Design (ID) Team and the Training Department are both part of Learning & Development (L&D). To simplify this report "Training Department" is used throughout it to refer to all of these. However, in the survey questions, each team or department is referred to in the context that the person completing the survey will best understand.

Problem Statement

To effectively evaluate learners, the Training Department must first determine how learner success is measured and then record those measurements to establish a baseline against which future learners can be compared. This will allow the Training Department to more easily answer the following questions:

- Is a program effective?
- Which of the program's learning strategies are effective?
- Which of the program's learning strategies are ineffective?

- How does the outcome of one instance of a training program compare to another?
- Are alternative forms of delivery (e. g. instructor-led vs. computer-based) for the same program equally effective?

Background

This action research project will focus on gathering success measurements for one specific group of learners in my organization: Default Call Center agents. Default Call Center agents receive inbound calls from customers who have defaulted on their mortgage payments and are looking for payment options that will allow them to avoid foreclosure and keep their homes. These agents currently receive training first in the classroom and then in the "training bay." The training bay is a supervised, live environment in which new agents must demonstrate proficiency in the job before graduating to the Default Call Center. The classroom training is facilitated by a trainer from the Training Department. The training bay is facilitated by a Default Call Center supervisor. Both environments provide an opportunity for data collection which is explained later in this report.

The Default Call Center agent training was chosen because it is a high-visibility, high priority program that affects the organization on a large scale. There is a large amount of speculation in the organization that many of these agents perform at levels below standard, but it is not clear why. This is a good opportunity to determine if perception is reality. If it is, the results of this project will be used to help identify training methods and materials that contribute to the problem.

Research Questions

Trainers	• How can I identify questions or topics in the classroom training assessment for which learners consistently show deficits? • Are there any assessment methods not currently in use that could be easily implemented? The Training Department currently uses Level 1 and Level 2 evaluations during the classroom training phase. The assessments are used only to score the participant. The Training Department may be able to identify deficiencies in the training materials by analyzing assessments to determine where participants consistently perform poorly.

Training Bay Supervisors	• What methods of assessment are currently used in the training bay? • Can any existing processes be used to establish baseline success measurements in the training bay? • Are there any assessment methods not currently in use that could be easily implemented? The Training Department is unfamiliar with the processes employed in the training bay. This project will identify how agents are assessed and if any of these measurements may be used by the Training Department to evaluate its training programs. I believe the Training Department may be able to identify deficiencies in the training materials by identifying deficiencies in the training bay.
Managers	• What methods of assessment are currently used in the call center? • Can any existing processes be used to establish baseline success measurements in the call center? • Are there any assessment methods not currently in use that could be easily implemented? The Training Department is unfamiliar with the processes employed in the call center. This project will identify how agents are assessed and if any of these measurements may be used by the Training Department to evaluate its training programs. The Training Department may be able to identify deficiencies in the training materials by identifying areas of poor performance in call center measurements. Any deficiencies identified here can also be cross-referenced to those identified in the previous two phases. If a deficiency self-corrects, analyzing that portion of the classroom training may be a lower priority.

Table 1: Research Questions by Measurement Phase

Methods

This project was part case study and part action inquiry. It explored the evaluation possibilities of one training program which will be used as a model going forward. The Default Call Center agent training is the most frequently delivered, highest profile training program in the organization. This action research project provided the insight into the Default Call Center agent training sought by the Training Department and can be applied in other departments in the future.

Data collection consisted of a qualitative mixture of surveys and interviews. Three pairs of subjects were selected across two locations and presented with identical questions via surveys. The surveys were followed with interviews of each participant. The interviews were used to solicit additional data, clarify answers provided in the surveys, compare the responses provided by each subject, and clarify any discrepancies between responses from different subjects. Interviewees were able to reflect on their answers and make amendments or additions if necessary. The research questions identified in Table 1 were delivered via surveys to the participants (data sources) identified below in Table 2.

Data Sources

Research Questions	**Data Sources 1 & 2**	**Data Sources 3 & 4**	**Data Sources 5 & 6**
How can I identify questions or topics in the classroom training assessment for which learners consistently show deficits?	2 Trainers • Denver • Indianapolis	2 Training Bay Supervisors • Denver • Indianapolis	2 Managers • Denver • Indianapolis
What methods of assessment are currently used in the training bay?		2 Training Bay Supervisors • Denver • Indianapolis	2 Managers • Denver • Indianapolis
Can any existing processes be used to establish baseline success measurements in the training bay?		2 Training Bay Supervisors • Denver • Indianapolis	2 Managers • Denver • Indianapolis
What methods of assessment are currently used in the call center?	2 Trainers • Denver • Indianapolis	2 Training Bay Supervisors • Denver • Indianapolis	2 Managers • Denver • Indianapolis
Can any existing processes be used to establish baseline success measurements in the call center?	2 Trainers • Denver • Indianapolis	2 Training Bay Supervisors • Denver • Indianapolis	2 Managers • Denver • Indianapolis
Are there any assessment methods not currently in use that could be easily implemented?	2 Trainers • Denver • Indianapolis	2 Training Bay Supervisors • Denver • Indianapolis	2 Managers • Denver • Indianapolis

Table 2: Data Sources

Surveys were administered via email and included the following questions:

Trainer Survey

1. List any methods that you currently use to evaluate training participant performance (job tasks only). Answer the questions below for each included method.
    - Do you record and track this data?
    - If the data is recorded and tracked, how is this done?
    - How often do you do this?
    - Can this data be made available to the ID team?
    - Can you provide an example with this survey?
    - What is the minimal acceptable level of performance?

- What methods do you currently use to identify trends of poor performance in the classroom? Can you share these results?
- How are employees evaluated in the Call Center to determine if they are performing at an acceptable level?
- How are employees evaluated in the Training Bay to determine if they have performed well enough to move on to the Call Center?
- Do you have any suggestions that would help us identify additional opportunities to measure a training participant's performance in the classroom?

Training Bay Supervisor Survey

- List any methods that you currently use to evaluate employee performance (job tasks only). Answer the questions below for each included method.
    - Do you record and track this data?
    - If the data is recorded and tracked, how is this done?
    - How often do you do this?

    - Can this data be made available to L&D on a periodic basis?
    - Can you provide an example with this survey?
    - What is the minimal acceptable level of performance?
- What methods do you currently use to identify trends of poor performance in the Training Bay? Can you share these results with L&D?
- Does the Training Bay currently have a defined baseline level at which an employee should perform (job tasks only) prior to graduating from the Training Bay? Can you provide this with this survey?
- How are employees evaluated in the classroom to determine if they have performed well enough to move on to the Training Bay?
- How are employees evaluated in the Call Center to determine if they are performing at an acceptable level?
- Do you have any suggestions that would help L&D identify additional opportunities to measure employee performance in your department?

Manager Survey

- List any methods that you currently use to evaluate employee performance on the job (job tasks only). Answer the questions below for each included method.
    - Do you record and track this data?
    - If the data is recorded and tracked, how is this done?
    - How often does your department do this?
    - Can this data be made available to L&D on a periodic basis?
    - Can you provide an example with this survey?
    - What is the minimal acceptable level of performance?
- Do you currently analyze evaluations to identify trends of poor performance, such as analyzing QA scorecards, for employees in the Default Call Center? Can you share these results with L&D?
- Does your department currently have a defined baseline level of employee performance (job tasks only)? Can you provide this with this survey?

- How are employees evaluated in the classroom to determine if they have performed well enough to move on to the Training Bay?
- How are employees evaluated in the Training Bay to determine if they have performed well enough to move on to the Call Center?
- Do you have any suggestions that would help L&D identify additional opportunities in your department to measure employee performance for the purpose of evaluating training program effectiveness?

Data Collection

Interviews were conducted in person for the Denver participants and via phone for the Indianapolis participants. Interview questions and responses were recorded manually in the form of text notes by the researcher. At the conclusion of the second interview for each participant pair, the researcher transcribed the notes for the pair into a single text narrative that is representative of the clarified, collective response of the pair. The interview was used to compare responses and clarify discrepancies, so a collective narrative was appropriate.

Some level of bias in qualitative research of this type is expected and, some researchers argue, unavoidable. Standards in qualitative research promote the researcher as an integral part of the data collection process who cannot, therefore, be completely separated from that process. In qualitative research knowledge is the result of researcher—subject interaction. Researcher bias, then, is a factor, no matter how diligently the researcher works to maintain neutrality. With this in mind, it is important that the researcher remain as neutral as possible, knowing that bias cannot be eliminated completely (Mehra, 2002). In this action research project, the researcher attempted to influence data collection as little as possible by recording and considering all responses and by asking open-ended, non-leading questions.

Data collection was limited to two identical groups in two locations. Each group had only one trainer, one training bay supervisor, and one manager. Because there were only two trainers, two training bay supervisors, and two managers, it was not possible to triangulate the data collection between participants with the same job responsibilities. Therefore, questions for each job were administered to all participants. For instance, managers were asked how trainees are evaluated in the training bay. This was not out of scope for a manager because the training bay falls under the manager’s area of responsibility. However, trainers and training bay supervisors were less knowledgeable about the areas outside of their daily responsibilities and yielded less helpful data.

Participants

The study's data sources consisted of two managers, two trainers and two training bay supervisors. Each pair has identical job responsibilities, but is responsible for one of two locations: Denver or Indianapolis.

Participant Overview

Participant Pairs	Responsibilities
1 Trainer: Denver 1 Trainer: Indianapolis	Trainers conduct one three week training program for the Default Call Center. It is for new agents and is required prior to being admitted to the training bay. At the conclusion of the session, trainers assess the trainees' performance in the classroom.
1 Training Bay Supervisor: Denver 1 Training Bay Supervisor: Indianapolis	Manages the training bay for the Default Call Center. The training bay is an interim environment for newly trained agents in which the agents receive additional training and support while taking live calls from customers. Supervisors assess the trainees' performance in the training bay.
1 Manager: Denver 1 Manager: Indianapolis	Manages the Default Call Center and is responsible for evaluating agents' performance after they graduate from the training bay.

Table 3: Participant Overview

Data Analysis

Qualitative data from the surveys and interviews were analyzed by reviewing the survey responses and the interview observations, coding the information extrapolated from the data, and summarizing the conclusions drawn from the extrapolated data in a text narrative. Additionally, examples of the evaluation instruments identified during the data collection phase were gathered and analyzed to clarify the type of data each contains and to confirm the description provided for each instrument in the surveys and interviews.

The findings identified in this phase of the action research project will be used to implement a more robust evaluation program that includes summative and confirmative evaluation. Summative evaluation is administered at the conclusion of the training program and is used to determine if a training program is immediately effective. Confirmative evaluation is used to assess the long-term effectiveness of a training program (Dessinger & Moseley, 2004). In both cases data is gathered about a trainee's performance and can be used to modify the program content if necessary.

Validity

Windsor et al. (2004) define face validity, content validity, and reliability as follows:

- Face Validity: "Extent to which the instrument appears to be measuring what it is supposed to measure."
- Content Validity: "Extent to which an instrument samples items from the universe of content desired."
- Reliability: "Reliability (consistency) is an empirical estimate of the extent to which an instrument produces the same result (measure or score), applied once or two or more times."

The data gathered with the research instruments is valid because it generated responses consistent with the subject of the project. The data is considered reliable because the data gathered was mostly consistent across multiple samplings and was verified via follow-up interviews.

Data analysis was performed by an internal subject matter expert (the researcher), providing validity for the data analysis phase. However, research was limited by the number of available subjects in each area. The company has only two trainers, two training bay supervisors and two call center managers split between two locations. Data was triangulated by sampling multiple job functions in each location, though it would have been preferable to survey three similar locations.

Schedule

The combined initial schedule and amended schedule is included in the table below. The table outlines the plan for the data collection, data analysis, and reporting phases of this action research project. This plan lists the six major milestones for this project—from building the data collection instruments to implementing an action plan. Approximately three weeks were allocated to collecting and analyzing the data prior to the April 20 due date for the data analysis report. The final step of the action research process—action planning—extends beyond the timeline for this class.

Phase	Tasks	Participants	Due Date
Build Instruments	1. Build Survey: Bus. Leaders 2. Build Survey: Trainers ~~3. Build Survey: Participants~~ *Training Bay Supervisors added* 4. Prepare Interview Questions	Researcher	3/22/2009
Data Collection	1. Conduct Surveys 2. Conduct Interviews	1. Business Leaders 2. Trainers ~~3. Course Participants~~ *Training Bay Supervisors added*	4/3/2009
Data Analysis	Analyze Survey and Interview Data	~~1. Instructional Designers~~ ~~2. Trainers~~ *Researcher*	4/13/2009
Data Analysis	1. Draft: Analysis Report Due 2. Final: Analysis Report Due	Researcher	4/20/2009 4/25/2009
Report Results	1. Draft: Final AR Report Due 2. Final: Final AR Report Due	Researcher	5/7/2009 5/12/2009 *5/16/09 report submitted*
Action Planning	TBD	TBD	TBD
Table Key: • ~~Strikethrough~~ = Items removed from the initial schedule • *Italics* = Items added to the amended schedule			

Table 4: Combined Initial and Amended Schedule

As the research progressed the scope of the project was modified to meet reasonable time frames for completion dependent on the time available to the researcher as well as the availability and appropriateness of the selected participants. Also the initial proposal did not include the training bay supervisors as survey and interview participants. Collectively, this resulted in three modifications to the project schedule and scope:

1. Course participants (trainees) were eliminated from the pool of survey and interview subjects. As new hires these subjects offered the least opportunity for discovery of existing methods of measurement and evaluation. These participants were eliminated to provide sufficient time to complete the other data collection and data analysis tasks listed in the table below.

2. Due to the availability of Training Department resources, the researcher performed all of the data analysis rather than include the trainers and instructional designers in this process. Otherwise, the research adhered to all time frames stipulated in the table.

3. The training bay supervisors were added after it was determined in the planning phase that their participation would most likely provide helpful insight into the processes utilized in the training bay.

Literature Review

The instructional design team's ultimate goal is to implement a process for evaluating the training programs produced by the team. The goal of this action research project is to identify success measurements for trainees to be used as a baseline during an evaluation program. For an evaluation program to be most effective, the evaluator must first identify what he or she will be evaluating against (i.e. success measurements). This review builds on that argument, providing further evidence that planning—including the identification of success measurements—and following a process are both critical elements of a successful evaluation program.

This literature review will help to determine best practices for program planning and implementation. An evaluation program will require additional time and effort from each participating party (i.e. training bay supervisors, default call center managers, trainers, and instructional designers). If an evaluation program is implemented without a clear process and purpose, the researcher runs the risk of losing the support of the participating parties. This could diminish or eliminate the possibility of the program being effective and useful.

Literature Review - Research Questions

The need for evaluation in training programs is widely accepted in the literature available for this topic. The general consensus is that training should be evaluated. However, recent trends place an emphasis on implementing a process to evaluate training. This includes planning and designing a program and identifying objectives and key performance indicators. This research project is important to ensure that an evaluation program is implemented efficiently and effectively. This review explores the following questions:

- Which learning theories and design models support evaluation?
- What is the role of evaluation in common learning theories and design models?
- What are common methods of evaluation?
- What should be evaluated?
- When should evaluation occur?
- How should an evaluation program be planned and implemented to ensure it is effective and useful?

Search Methods

Research was performed online using a combination of Google, Google Scholar, and the Auraria Library Website. Google Scholar was far more efficient for finding scholarly research. Subscription journals identified via Google Scholar were accessed via the Auraria Library. This consisted of three books including one owned by the researcher and two obtained via the Auraria Library. Search phrases included: evaluating corporate training, effective corporate training, key performance indicators, Kirkpatrick, ADDIE, Gagne, and adult learning theory.

Learning Theories

Evaluation is an element of many instructional design models and the theories behind those models. A quick survey of established educational theories reveals evaluation to be a critical element within each.

- In Robert Gagne's Nine Events of Instruction, feedback and assessment (steps 7 & 8) play a very significant role as reinforcers of learning (Kruse, 2004).
- Behaviorism places an emphasis on the evaluation of measurable, observable objectives to determine if a behavior has been learned (Ormrod, 2008).

- Cognitivism also places an emphasis on the evaluation of measurable, observable objectives, but it adds that reasonable inferences can be made about how a subject's mental processes contributed to learning (Ormrod, 2008).

- In constructivism the learner defines in what context they will learn, learning is based on personal experiences and preferences, and evaluation is usually an examination of what the learner has produced (Mergel, 1998).

- Adult learning theory contends that adults learn best when frequent practice and structured feedback is incorporated into their learning activities (Adult Learning Theory, 2004).

It is evident that the need for evaluation is supported by a variety of common learning theories. In *Instructional Design & Learning Theory,* Mergel (1998) links behaviorism, cognitivism, and constructivism to the modern practice of instructional design. Evaluation is easily integrated into behaviorist and cognitivist design. Because content is not pre-defined in constructivist training, it is more difficult to implement a formal method of evaluation such as a test. Instead, assessment is often based off of the materials produced by the learner. Mergel argues that instructional designers find it much easier to implement evaluation when the instruction is rooted in behaviorist or cognitivist principles.

Design Models

The most popular method used to build learning programs is ADDIE from the instructional systems design (ISD) model. ADDIE is a development process used to create effective learning materials (Januszewski, Molenda, & Harris, 2008). The letters in ADDIE represent the five steps of the ISD process: Analysis, Design, Development, Implementation, and Evaluation (Learning Theories Knowledgebase, 2008).

Figure 1: The ADDIE Model

In the ADDIE model each phase is the foundation for the next phase. Evaluation is the fifth and final phase of the process. It is used to test not only the effectiveness of the final instructional event but also each of the four previous phases in ADDIE (Januszewski, Molenda, & Harris, 2008).

This dichotomy is commonly referred to as formative and summative evaluation. In *The Training Evaluation Process*, Basarab and Root (1992) describe how formative and summative evaluation are applied during the construction of a training program and help to ensure that the final product adheres to the program's defined goals and objectives. Summative evaluation is used after a program's implementation to identify its effectiveness and value to learners and to the organization. Summative evaluation is also used to identify which portions of a training program work well and which do not, thereby providing valuable insight to the instructional designer for where improvements can be made.

While evaluation is built into the ADDIE model the process for implementing it is not defined within ADDIE. The most well-known evaluation model is Donald Kirkpatrick's Four Levels of Evaluation. Kirkpatrick uses four types of learner evaluation to assess and ensure the effectiveness of a training program. In *Kirkpatrick's Four Levels of Evaluation*, Winfrey (1999) argues that evaluation should always begin with level one and progress through each of the four levels over the course of the evaluative process. Each level builds on the data gathered in the previous level and provides a more complete understanding of the results of a training program.

Level 1 – Reactions

A level one evaluation is typically administered immediately after a training program. As the name implies it seeks to identify the general reaction of the learner to the training program. The level one evaluation might ask if the learner enjoyed the program, was comfortable in the training environment, and if they believe the instructor was well prepared.

Level 2 – Learning

Level two evaluations attempt to determine if the learner increased his or her knowledge or skills as a result of the training program. Typically, this evaluation is administered via test or self-assessment. The results are used to identify if the training program achieved its intended objectives.

Level 3 – Transfer

Level three evaluations seek to determine if learners are applying the new information, skills, and behavior presented in the training program to their jobs. Many training professionals believe this is the most compelling indicator of the effectiveness of a training program.

Level 4 – Results

The final level of evaluation in Kirkpatrick's model is used to measure a training program's effect on an organization's overall performance. This includes factors such as "increased production, improved quality, decreased costs, reduced frequency of accidents, increased sales, and even higher profits or return on investment" (Winfrey, 1999).

The goal of both ADDIE and Kirkpatrick is to improve training (Basarab & Root, 1992; Januszewski, Molenda, & Harris, 2008). Before either can be implemented it is important to establish a plan for the evaluation. A well designed evaluation plan increases the instructional designer's opportunities to identify effective and ineffective training strategies in the training program being evaluated (Basarab & Root, 1992).

Effective Evaluation in Practice

The objective for creating an evaluation plan is to provide a structure that will help the researcher determine if evaluation is necessary (Basarab & Root, 1992), and if it is, to identify weaknesses in the training program that require clarification or modification. The evaluation plan provides structure, which is necessary to determine if a training program is effective (Bushnell, 1990).

One evaluation method presented by Bushnell (1990) in *Training & Development Journal* that includes creating a plan and incorporates both ADDIE and Kirkpatrick is Input, Process, Output (IPO). The IPO model breaks the training process into three main phases: Input, Process, and Output. These phases provide a process for creating an evaluation plan, implementing the training and the evaluation, and then adjusting the training based on the results of the evaluation.

Evaluation occurs in and influences each phase of the IPO model. Bushnell (1990) argues that these steps help the instructional designer to define the purpose of the evaluation, create appropriate measurement instruments, define efficient, effective data sources, and then link the results of the evaluation back to the original purpose established in the first phase.

The value of the IPO model is also a theme in *Confirmative Evaluation: Practical Strategies for Valuing Continuous Improvement* by Dessinger and Moseley. Like Bushnell, Dessinger and Moseley (2004) argue that evaluation should occur throughout the planning (input), implementation (process), and adjustment (output) phases of a training program. But, they believe that evaluation in its common application is incomplete. Most design/evaluation models, like ADDIE, include only formative and summative evaluation.

Confirmative evaluation is a third form of evaluation that, unlike formative and summative evaluation, is used to assess the long-term effectiveness of a training program. The results of the confirmative evaluation, which show permanence, can be used to modify the program if necessary. The process with which confirmative evaluation is applied can be referred to as, “Plan, Do, Analyze, Improve” (Dessinger & Moseley, 2004). This process is essentially the same as the IPO process described previously. The difference is in the timing and the purpose. Confirmative evaluation is implemented up to 12 months after a training program is completed. Its purpose, among other uses, is to improve the quality of the training program (Dessinger & Moseley, 2004).

In *What should training evaluations evaluate?* Mann and Robertson (1996) argue that evaluation is very important for training programs, but that the benefits of different types of evaluation, specifically Kirkpatrick’s four levels, require elucidation. Corporations have decided that training is necessary, but they do not choose evaluative methods that truly predict future competencies in learners. Mann and Robertson believe corporations should implement more robust evaluation processes to identify effective and ineffective elements in their training programs and then allow revision to accommodate those results. Most corporations evaluate reactions (Kirkpatrick’s Level 1) and learning (Kirkpatrick’s

Level 2) at the conclusion of a training program. Mann and Robertson cite a study that contends this type of evaluation cannot provide insight into if learners will benefit in the long-term from the training. A more effective approach to evaluation, they argue, would be to measure the learner's self-efficacy in regards to the learning content at the conclusion of the training program. Self-efficacy, as supported by the study, is a better predictor of the long-term benefits of training.

Key Performance Indicators

Using Indicators to Measure Sustainability Performance at a Corporate and Project Level argues that properly identifying key performance indicators is critical to the prosperity of a corporation (Keeble et al., 2003). Likewise, identifying goals and objectives is an important part of any evaluation process. In ADDIE, IPO, and confirmative evaluation, emphasis is placed on planning the evaluation process before implementing it.

Part of planning is to define a training program's goals and objectives (Basarab & Root, 1992). Similarly, key performance indicators are defined as measurements of an organization's goals (Reh, 2009). The parallels here provide an opportunity to apply the principles of key performance indicators to evaluation. For instance, one approach to defining key performance indicators is to answer the following four questions:

1. *What is critical and relevant to the organization?*
2. *What commitment does the organization need to support?*
3. *How will they benchmark performance?*
4. *What do stakeholders expect of them?* (Keeble et al, 2003)

If an instructional designer substitutes "training program" for "organization" in these questions, he or she will have a good structure for building a training program's goals and objectives.

Literature Review - Conclusion

Learning theories, instructional design models, and research about key performance indicators all emphasize the need for evaluation of corporate training programs to ensure the quality of the materials and to determine the program's effectiveness. As this is generally accepted in corporate training, much of the research for this topic today explores the need for structure in evaluation programs similar to what ADDIE provides for instructional design.

Consistent with the researcher's thesis, this review supports the idea that identifying success measurements is an important part of the evaluative process.

However, this review reveals that planning and implementing a process are more important to the success of an evaluation program than was previously considered. Evaluation programs should be planned and the key phases and goals of the program should be identified and recorded prior to the implementation of the evaluation.

The type of learning theory used to design the training should also be considered during this planning phase. Different types of instruction, particularly instruction based on behaviorism and cognitivism, more easily support evaluation. Evaluators should also consider more than formative and summative evaluation during the planning of the program. Confirmative evaluation and the evaluation of self-efficacy are two examples of what can be evaluated outside of formative and summative evaluation.

While research revealed much about the planning and implementation stages of an evaluation program, it did not reveal much about what to do with the results of the evaluation. In some cases the implications of the results may be clear. A researcher may identify an assessment topic that was troublesome for learners and determine that the learning content does not adequately support that topic. In other cases the cause and effect may be less clear. A topic, its supporting learning content, the reinforcing activities, and the corresponding assessment may appear to be sound, yet learners consistently struggle with that topic. Future consideration for further research may be given to studying how to analyze and apply the results of the evaluation to the existing materials.

Findings

Research Questions:

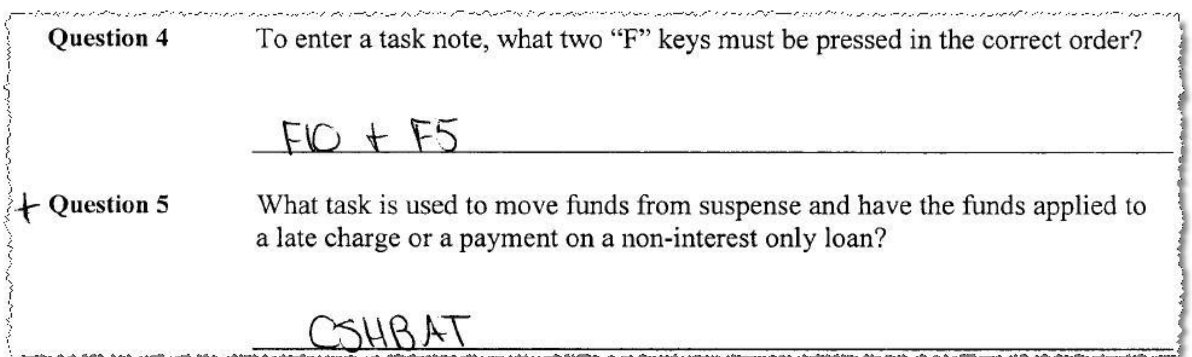
- How can I identify questions or topics in the classroom training assessment for which learners consistently show deficits?
- What methods of assessment are currently used in the training bay?
- Can any existing processes be used to establish baseline success measurements in the training bay?
- What methods of assessment are currently used in the call center?
- Can any existing processes be used to establish baseline success measurements in the call center?
- Are there any assessment methods not currently in use that could be easily implemented?

How can I identify questions or topics in the classroom training assessment for which learners consistently show deficits?

- ***Course Assessment***
- ***Job skill Role-play/QA Scorecard***

Course Assessment

Course assessments are administered at the conclusion of each classroom training session. Currently, the assessment is scored and returned to the trainees. Only the score is recorded in the Training Department's Curriculum Management System (CMS). Since the assessments are comprehensive and are administered at the conclusion of the three week training, the scores alone will not provide sufficient insight for how trainees are performing in the classroom for specific topics. It will be necessary to record the responses to each question individually to identify performance trends for specific topics. The trainers are willing to modify their process and provide copies of the completed and graded assessments to the Training Department (IDs) for further analysis.



Question 4 To enter a task note, what two "F" keys must be pressed in the correct order?

F10 + F5

Question 5 What task is used to move funds from suspense and have the funds applied to a late charge or a payment on a non-interest only loan?

CSHBAT

Image 1: Course Assessment

Job Skills Role-plays

Job skills role-plays are administered for specific topics during the classroom training. QA scorecards (formal scorecards used to evaluate call center agents) can be used by the trainer or a peer reviewer to evaluate trainee performance in the classroom. This is not implemented consistently in the classroom and the results are not recorded. Trainers typically have to rely on new hires to implement these depending on class size and time constraints. To use this as a summative and confirmative evaluation tool, trainers would need to use the QA scorecards with every role-play (role-plays are implemented at various intervals throughout the training), save the results, and provide the results to the Training Department.

Implementing QA scorecards into classroom role-plays is the least desirable option for gathering evaluative data. It would require a change to the existing process and could be difficult to implement effectively. In the call center, QA scorecards are administered by trained individuals in the QA Department and the results are compiled and reported by the same department. The QA Department is not available to provide similar functions in the classroom. Due to large class sizes, implementing this consistently in the classroom would require the scorecards to be administered mostly by peers. This presents a risk that the data would not be valid because it is not being administered by experts.

What methods of assessment are currently used in the training bay?

- ***Quizzes***
- ***Average Handle Time Report***
- ***Production Report***
- ***Error Report***
- ***QA Scorecard/QA Trending Report***
- ***Side-by-Side Observation/QA Scorecard***

Quizzes

Quizzes are used to evaluate an agent's performance before and after he or she enters the training bay. Currently, quiz results are not recorded or saved. These results may be useful to the Training Department as a verification of the results from the analysis of the course assessments.

Average Handle Time Report

In the Training Bay Supervisor Follow-up Interview it was reported that:

> *This report includes two measurements of an agent's efficiency receiving and concluding calls. Average Handling Time measures the time each agent spends on the calls he or she received the previous day. Wrap Time measures the time an agent spends after he or she hangs up with a caller before the agent receives a new call. The Wrap Time is typically used to record notes or perform other tasks resulting from the call. While these times will vary from call to call, the agent's average times should be within a standard range. A manager in the Default Call Center compiles and distributes this report daily to all agents including those in the training bay and those in the general population.*

In both manager interviews the managers reported that, “Average handle time is a combination of talk time + wrap time + hold time. Handle time measures the total time each agent spends on the calls he or she receives.”

Team Name	Agent Name	Inbound Calls	Outbound Calls	Total Login Time Seconds	Total Login Time hh:mm:ss	Total Handle Time Seconds	Average Handle Time hh:mm:ss	Total Handle Time hh:mm:ss	Total Idle Time Seconds
DEN_DF_WILLIAMS_FPS	[illegible]	38		32,353	08:59:13	24,158	00:10:36	06:42:38	1,361
DEN_DF_WILLIAMS_FPS	[illegible]	46		32,124	08:55:24	27,542	00:09:59	07:39:02	98
DEN_DF_WILLIAMS_FPS	[illegible]	42		32,459	09:00:59	27,880	00:11:04	07:44:40	964
DEN_DF_WILLIAMS_FPS	[illegible]	51		32,244	08:57:24	25,221	00:08:15	07:00:21	939
DEN_DF_WILLIAMS_FPS	[illegible]	29		32,318	08:58:38	22,998	00:13:13	06:23:18	969
DEN_DF_WILLIAMS_FPS	[illegible]	35		30,268	08:24:28	19,123	00:09:06	05:18:43	81

Image 2: Average Handle Time Report

Production & Error Reports

In the Training Bay Supervisor Follow-up Interview it was reported that, “These [production reports] are the same as the Average Handle Time Reports.” In the Manager Follow-up Interview it was reported that, “Production refers to the daily number of loan workout plans set up by an agent. This description is contrary to what was reported by the training bay supervisors, but was confirmed by both managers.”

Production and error reports provide a summary quantitative overview of what the agent produces (production) and how well he or she produces it (error). Low production may indicate that an agent is having difficulty navigating online systems and resources or recalling definitions and procedures presented during the training. Error reports are more specific and, therefore, may indicate precise topics where agent performance trends poorly. Both production and error reports indicate individual agent and summary group performance.

Daily HRG Deal Tracking

Mar-09

Loss Mitigation Resolutions by Processor ID

Loss Mit Removal Date: (All)

Count of Loan Number — Deal Type

Processor Name	Loss Mit Processor Id	FORBEARANCE	MODIFICATION	PAYOFF	REINSTATEMENT	SHORT SALE	DEED IN LIEU	FNMA CASH ADV	SETTLEMENT	Total Resolutions
[illegible]	AMZ		29							29
[illegible]	AEJ		34							34
[illegible]	AVD		1			35	2	3		4

Image 3: Production Report

F	G	H
LOSS MIT PROCESSOR ID	**COMMENT CODE DESCRIPTION**	**REASON**
J15	FEES AND COSTS NOT REQUESTED	
53A	INCORRECT FEES AND COST REQUES	
K50	FEES AND COSTS NOT REQUESTED	
FBN	FEES AND COSTS NOT REQUESTED	
V19	FEES AND COSTS NOT REQUESTED	
FBN	INCORRECT FEES AND COST REQUES	
E68	NO EXCEPTION APPROVAL	Rejection
FBN	NO EXCEPTION APPROVAL	Rejection
BGA	NO EXCEPTION APPROVAL	Rejection
FBN	NO EXCEPTION APPROVAL	Rejection
I1R	NO EXCEPTION APPROVAL	Rejection
H1H	NO EXCEPTION APPROVAL	Rejection
V34	NO EXCEPTION APPROVAL	Rejection
C60	NO EXCEPTION APPROVAL	Rejection
FBN	NO EXCEPTION APPROVAL	Rejection
V34	NO EXCEPTION APPROVAL	Rejection
DMA	SURPLUS REQUIREMENT NOT MET	Rejection
D5L	NO EXCEPTION APPROVAL	Rejection
J07	NO EXCEPTION APPROVAL	Rejection
I1R	NO EXCEPTION APPROVAL	Rejection
F75	SURVEILLANCE APPROVAL NOT REQU	Rejection

Image 4: Error Report

QA Scorecard/QA Trending Report

QA (Quality Assurance) scorecards measure recorded call quality for the following categories:

- Greeting
- Financials
- Program Knowledge
- Status Call Handling
- Closing
- Professionalism

An additional category for serious infractions or omissions by the agent, such as inappropriate behavior that “discredit” a call is also included. QA scorecards are implemented in the training bay by the QA Department, but only on a limited basis. Recorded calls are evaluated and scored five times per month for each trainee. Scores for each agent are also presented collectively to show how the entire group is trending via the QA Trending Report.

Quality Monitoring Form			
Agent Name:			
Caller Name:			
Evaluator:			
Date & Time of Call:			
Call ID:			
Section One: Greeting / Validation / Goal		Weighting	
1	Greeting	1,C,0	
2	Verification Process	3,C,0,NA,DC	
3	Purpose / Goal of Call	3,C,0,NA	
4	Confirm Assistance / Set Call Agenda	1,C,0,NA	
Section Two: Understanding Customer's Financial Picture			
5	Review the File / System Navigation *	3,C,0,NA	
6	Probing/Consultative Questions / Rephrase for Understanding *	3,C,2,0,NA,DC	
7	Accurately Complete Financials	3,C,2,0,NA,DC	
Section Three: Identify & Position Options			
8	Demonstrate Program Knowledge	3,C,2,0,NA	
9	Select Appropriate Program/Options	1,C,0,NA	
10	Effectively Present Program/Options *	4,C,2,0,NA,DC	
Section Four: Status Calls (steps 1-5 apply / 6 may apply / 7-10 NA)			
11	Provide Complete Status / Advance File	10,C,5,0,NA	

Image 5: QA Scorecard

Month End Quality Scores		Average Quality by	
Default			
	Februray 2009 Average	**Completed Evaluations**	
[illegible]	90.26%	34	
[illegible]	88.80%	45	
[illegible]	91.04%	54	
[illegible]	89.77%	56	
[illegible]	90.83%	47	
[illegible]	86.81%	26	
Overall Average - Denver	**89.59%**	**262**	**Total Evaluations**

Image 6: QA Trending

Side-by-Side Observation/QA Scorecard

In the Training Bay Follow-up Interview it was reported that, “The training bay supervisor also performs side-by-sides with each agent. This is an observation of the agent by the supervisor as the agent executes a live call. The agent’s performance is recorded using QA scorecards.” If the number of agents in the training bay is small (five or less) the training bay supervisor can administer them in the time allowed. If the number of agents is more than five, this will be difficult due to time constraints. Currently there are 18 agents in the training bay at each location.

Use of side-by-sides was only reported by one of the two supervisors. The supervisor that reported using the side-by-sides as an evaluation tool also reported that use of this is inconsistent and the results are not recorded. Like the job skills role-plays in the classroom training, these may be difficult to implement and would require peer evaluators.

Can any existing processes be used to establish baseline success measurements in the training bay?

Since the training course assessment has already been identified as a useful success measurement, and the training bay quizzes assess similar skills and knowledge, the quizzes likely will provide little additional insight into trainee performance not already contained within the course assessments.

The Average Handle Time Report may be a useful method of evaluating training content. If a high-percentage of agents is not achieving the desired handling times, it may indicate a lack of call control skills or basic call center knowledge. This could indicate that an agent is having difficulty navigating online systems and resources or recalling definitions and procedures while he or she is on the phone with a customer. While the Average Handle Time Report will not help identify specific training topics that should be analyzed, it may be useful as an indicator that a more rigorous examination of the potential causes for slow times is required.

The Production Report and Error Report may also be useful methods of evaluating training content, but are not preferred. The Production Report lacks specificity. While it may indicate a potential deficiency, it would offer few clues as to what the deficiency is. The Error Report contains much useful information, but it largely duplicates the information contained within the QA scorecards and the QA Trending Report.

Implementing QA scorecards into training bay side-by-sides is not preferred. Like using QA scorecards with the classroom role-plays, the QA Department is not available to perform this function. The scorecards would have to be administered by peers (novices) which may provide unreliable results.

What methods of assessment are currently used in the call center?

- ***Average Handle Time Report***
- ***Production Report***
- ***Error Report***
- ***QA Scorecard/QA Trending Report***

Methods for measuring success for agents in the call center are identical to the methods for measuring success for agents in the training bay with two exceptions:

- Quizzes are administered in the training bay and not to the call center.
- Side-by-side observations are administered in the training bay and not to the call center.

Can any existing processes be used to establish baseline success measurements in the call center?

The Average Handle Time Report is less useful in the call center for evaluating training content. A call center agent must show proficiency with system navigation and department procedures to graduate from the training bay. Unlike for trainees poor handling times are less indicative of a training issue for an experienced agent.

As in the training bay, the Production Report and Error Report may be useful methods of evaluating training content, but are not preferred.

In one Manager Survey the manager responded that, "We use QA scorecards for calls and adherence to P&Ps [Policies & Procedures], but don't use them for trending opportunities." Though these reports are not used for comprehensive performance trending opportunities by the call center, they present the most thorough summary of an individual agent's on-the-job performance. The sections of the report are broken down to specific tasks (i.e. did the agent "effectively present Program/Options?"). Each task is scored based on written guidelines that are provided to the agents during training and then compiled into the QA Trending Report. The Training Department can use the QA Trending Report to identify topics where a higher percentage of agents are performing poorly. If further detail is required, the individual QA scorecards are available for analysis.

Are there any assessment methods not currently in use that could be easily implemented?

- ***Periodic Survey***

Periodic Survey

Currently, there is no instrument in place for trainers, training bay supervisors, or managers to record their general perceptions of the successes or failures of the training. Trainers do provide a subjective verbal review of each participant to call center supervisors at the conclusion of the training, but this is neither formal nor consistent. It is also not recorded. For the purpose of evaluating the training content, this offers no benefit to the Training Department in this format.

During the interview process all participants concluded that a simple survey of the participants' perception of the group's overall performance for each training topic may be useful. This would best be administered quarterly, allowing the participant to base his or her responses on multiple training sessions. Trainers, training bay supervisors, and call center managers are agreeable to implement such an instrument on a quarterly basis.

Conclusion

Environmental Impact

Establishing a more robust evaluation process for the organization's training programs will help the Training Department to better evaluate the effectiveness of the courses it provides, implement changes to improve ineffective courses, and ensure that courses are delivered with the most cost-effective method. This project will help the researcher to be a more effective manager and instructional designer by providing insight into which learning strategies and materials work and which do not. Ineffective content can be modified and effective content can be promoted.

This action research proposal is the first step toward establishing an evaluation process that will become standard practice for all training programs in the organization. Overall, the impact of this project has been positive. Business leaders and training bay supervisors believe the effort is an attempt to improve training which will ultimately help them as managers who are accountable for the performance of the call center. Trainers see this as an opportunity to improve the classroom experience for trainees, which, consequently, will improve the classroom experience for the trainers.

The next phase of the project is to implement an action plan, putting the findings of the project in motion, establishing a baseline for performance on the identified measurement instruments, and creating an evaluation program. This runs the risk of exposing multiple weaknesses in the existing training programs, which could reflect poorly on the Training Department. This is only a minor concern. Deficiencies should be minimal and of an amount that would be expected of any program. The issue is not that there are deficiencies in the training materials; it is that they should be indentified and corrected.

Implications for Practice

It would have been preferable to collect data via a focus group consisting of all of the project participants. Much of the time allocated to conducting the interviews and data analysis was spent determining if two seemingly different measurement instruments were actually the same instrument referred to by different names. A focus group could have clarified issues like this much more quickly.

Additionally, it would have simplified the research to gather and analyze the existing measurement instruments, such as the Average Handle Time Report, used by the participants to evaluate their trainees/employees sooner. Several of the measurement instruments identified by the survey participants, such as those that track attendance and punctuality, were not relevant to training content and could have been eliminated earlier in the project allowing the researcher to focus on the instruments that showed greater potential.

Revisiting the issue of bias in qualitative research, while it may not be possible to eliminate researcher bias completely, it could be advantageous to present the project's findings to the research participants and a third-party, such as an instructional designer, for further analysis. This could be used to verify the results and to ensure that researcher bias did not influence the results negatively.

The research revealed suitable results consistent with the original thesis. Success measurements for call center agents are available, but they have not been consistently recorded or made available outside of the call center. The most surprising finding was the idea that a periodic survey could be administered to trainers, training bay supervisors, and managers to gain further insight into which training content may require further evaluation. This was surprising because it is simple and easy to implement, but had not been previously considered. This also supports research that contends that evaluation is most effective and efficient when the program adheres to a structured process in which goals and objectives are defined during the planning phase. In this case action research helped to reveal which success measurements are most accessible and which would be more difficult to use.

The most useful measurements identified via the data analysis are those that are already in place and show sufficient insight for potential training deficiencies. This consists of the QA scorecards and QA Trending Reports administered to the call center and training bay by the QA Department, the Average Handle Time Reports compiled for the training bay, and the course assessments administered during the classroom training.

There is already an existing process in place to administer the QA scorecards and QA Trending Report in the call center and training bay. The QA Trending Report provides both a summary and detailed analysis of agent performance and can be made available to the Training Department. Similarly, average handle time is an existing process that is compiled for the training bay by call center management. A process to administer the course assessments is already in place as well.

Using course assessments for evaluative purposes would require minor changes to that process. Trainers would be required to save graded copies of each assessment and make those available to the instructional design team within the Training Department. The trainers are agreeable to the possibility of this change and it is not viewed as an impediment to the evaluation program.

It is recommended that a process be established to gather and analyze the QA scorecards, QA Trending Reports, Average Handle Time Reports, and Course Assessments on a quarterly basis. Additionally, the Training Department will implement a quarterly survey of trainers, training bay supervisors, and call center managers to gauge their perception of the effectiveness of the training materials. Part of the implementation process will include the acceptance of the new process by all involved parties. This is important to ensure that the instruments identified in this report continue to be implemented and the results are recorded, saved and made available to the Training Department.

Additional research may be necessary in two areas. First, it may be helpful to survey the reactions of trainees to the learning content and learning strategies included in the classroom and training bay curriculum. Trainee reactions were excluded from this project because they were deemed less helpful for identifying success measurements in the call center. But, when a baseline performance level is established as a result of this project, it may be useful to research what role learner motivation plays in learner success. Motivation plays an important role in the effectiveness of a training course and is partially determined by a learner's perception of the relevancy and usefulness of a training program (Ormrod, 2008). Training program effectiveness could be influenced as much by learner motivation as it is by the quality and relevance of the training content.

A second opportunity for further research is to identify industry standards for call center performance and then compare the performance baseline established as a result of this project to the newly identified industry standards. This would allow the organization to compare the organization's performance overall to the performance of similar organizations.

References

Adult Learning Theory. (2004). Retrieved February 21, 2009, from National Central Regional Educational Library Web site: http://www.ncrel.org/sdrs/areas/issues/methods/technlgy/te10lk12.htm

Basarab, D. J., & Root, D. K. (1992). *The training evaluation process*. Norwell, MA: Kluwer Academic Publishers.

Bushnell, D. S. (1990).Input, process, output: A model for evaluating training. *Training & Development Journal*. *44(3)*, 41-43.

Dessinger, J. C., & Moseley, J. L. (2004). *Confirmative evaluation: Practical strategies for valuing continuous improvement* .San Francisco, CA: John Wiley & Sons, Inc.

Januszewski, A., Molenda, M., & Harris, P. (2008). *Educational technology: A definition with commentary*. Mawah, NJ: Erlbaum.

Keeble, J. J., Topiol, S., & Berkeley, S. (2003). Using indicators to measure sustainability performance at a corporate and project level. *Journal of Business Ethics*. *44*, 149-158.

Kruse, K. (2004). *Gagne's nine events of instruction: An introduction.* Retrieved February 21, 2009, from e-LearningGuru.com Web site: http://www.e-learningguru.com/articles/art3_3.htm

Learning Theories Knowledgebase. (2008). *ADDIE model.* Retrieved February 21, 2009 from: http://www.learning-theories.com/addie-model.html

Mann, S., & Robertson, I. T. (1996). What should training evaluations evaluate?. *Journal of European Industrial Training*. *20(9)*, 14-20.

Mehra, B. (2002). Bias in qualitative research: Voices from an online classroom. *The Qualitative Report*, *7(1)*, Retrieved May 15, 2009, from http://www.nova.edu/ssss/QR/QR7-1/mehra.html

Mergel, B. (1998, May). Instructional design & learning theory. Retrieved March 29, 2009, from Penn State York Web site: http://www2.yk.psu.edu/~jlg18/506/Instructional%20Design%20by%20Mengal.pdf

Ormrod, J. (2008). *Human Learning*. Upper Saddle River, NJ: Pearson Education, Inc.

Reh, F. J. (2009). Key performance indicators (KPI): How an organization defines and measures progress toward its goals. Retrieved March 26, 2009, from About.com Web site: http://management.about.com/cs/generalmanagement/a/keyperfindic.htm

Windsor, R., Clark, N., Boyd, N. R., & Goodman, R. M. (2004). *Evaluation of health promotion, health education, and disease prevention programs*. New York: McGraw-Hill.

Winfrey, E.C. (1999). Kirkpatrick's Four Levels of Evaluation. In B. Hoffman (Ed.), *Encyclopedia of Educational Technology.* Retrieved February 21, 2009, from http://coe.sdsu.edu/eet/Articles/k4levels/start.htm

Appendix: Data Collection Instruments

Manager Survey: Indianapolis

1. Please indicate your name, department and position.

David Hughes Default Call Center Strategy Manager

2. List any methods that you currently use to evaluate employee performance on the job (job tasks only). Answer the questions below for each included method.
    a. Do you record and track this data?
    b. If the data is recorded and tracked, how is this done?
    c. How often does your department do this?
    d. Can this data be made available to L&D on a periodic basis?
    e. Can you provide an example with this survey?
    f. What is the minimal acceptable level of performance?

Method 1: Attendance & Schedule Adherence a. Company does b. Reports (Attendance & Status) c. Review daily d. Reports are available to management e. Link below: http://my.lehman.com/sites/MYD_27119/Management_Contact_Center_Reports/default.aspx f. Company standards – <3 occurrences (tardies / absences)
Method 2: Plans/P2P Per Day a. Company does b. Reports from Dana Wilson & Daily Scorecard (MTD results) c. Review daily d. Dana's reports & Daily Scorecard can be provided e. Yes f. no minimum standard in TB / compare to average - FPS I & II: approx 4-5 P2P's a day – FPS III & IV: approx 3 plans a day
Method 3: AHT & Wrap Time a. Yes b. Productivity Reports c. Review daily d. Reports are available to management e. Yes (link above)

f. 12-minutes in TB / target <10-minutes in production / must balance w/need to deliver one-call resolution

These are three. There are more.

3. Do you currently analyze evaluations to identify trends of poor performance, such as analyzing QA scorecards, for employees in the Default Call Center? Can you share these results with L&D?

Absolutely. We review all production reports—QA Trending Reports, Status & Production Reports, Error Reports, Daily Scorecard, Dana Wilson's reports—as well as internal performance assessments—Repayment Plan checklists, QA Form (side-by-sides, recorded calls, live monitoring). These are mostly company results available to management.

4. Does your department currently have a defined baseline level of employee performance (job tasks only)?
   a. Can you provide this with this survey?

Yes. I believe this was already provided to you.

5. How are employees evaluated in the classroom to determine if they have performed well enough to move on to the Training Bay?

Schedule Adherence & Attendance
Average Handle Time & Wrap Time
Production (plans, P2P's, calls) – reports (quantity) & checklists (quality)
QA Results – overall quality of calls (right resolution achieved)
Quizzes – working knowledge of module topics
Overall questions tapering down to one-off situations (demonstrating working knowledge)

6. How are employees evaluated in the Training Bay to determine if they have performed well enough to move on to the Call Center?

Demonstrate a working knowledge of all covered modules.
Achieving minimum production requirements.

7. Do you have any suggestions that would help L&D identify additional opportunities in your department to measure employee performance for the purpose of evaluating training program effectiveness?

Appendix: Data Collection Instruments

Per our conversation last week, Greg, I would like to further explore building into the training material & classes techniques & discussions around maximizing efficiencies to serve both the business needs (quality) as well as the call center needs (quantity). In other words, focusing not only the effectiveness (quality) of the agent's performance, but also their efficiency (quantity)—without at all sacrificing effectiveness. I am thinking of things like "Call Control" & "Multitasking" (talking & typing). Thanks, Dave

Appendix: Data Collection Instruments

Manager Survey: Denver

1. Please indicate your name, department and position.

Jason Cramer, Default Call Center, Denver Site Manager

2. List any methods that you currently use to evaluate employee performance on the job (job tasks only). Answer the questions below for each included method.
    a. Do you record and track this data?
    b. If the data is recorded and tracked, how is this done?
    c. How often does your department do this?
    d. Can this data be made available to L&D on a periodic basis?
    e. Can you provide an example with this survey?
    f. What is the minimal acceptable level of performance?

Method 1: Deal Production a. Yes b. Deal Tracking Reports c. Daily/Monthly d. Yes e. Attached in email as "Copy of Daily HRG Deal Tracking 2009_03" Look at Forbearance Starts tab f. There is no "minimum", but new ICP will have thresholds for payouts that start at 30 per month.
Method 2: Phone Statistics (these are log in / log outs, schedule adherence, tardies, absences, etc) a. Yes b. Workforce Management tracks this and creates reports c. Daily d. Yes, let me know if you really want to see these. I don't really see L&D influencing this aspect of their job e. See above f. 3 occurances in rolling 12 months, 94% schedule adherence
Method 3: Average Handle Time a. Yes b. Workforce Management Reports c. Daily d. Yes e. Yes

f. Handle time is a combination of talk time + wrap time + hold time. There is no “magic number” but I train agents that they should not stand out from their peers. Avg for department right now is about 10-11 mins

3. Do you currently analyze evaluations to identify trends of poor performance, such as analyzing QA scorecards, for employees in the Default Call Center? Can you share these results with L&D?

A little. We use QA scorecards for calls and adherence to P&Ps, but don’t use them for trending opportunities.

4. Does your department currently have a defined baseline level of employee performance (job tasks only)?
   a. Can you provide this with this survey?

We are developing a scorecard and are using a rolling 6 month performance of each agent that encompasses metrics as well as “intangible” items such as strengths and opportunities.

5. How are employees evaluated in the classroom to determine if they have performed well enough to move on to the Training Bay?

There is a test

6. How are employees evaluated in the Training Bay to determine if they have performed well enough to move on to the Call Center?

There is a test

7. Do you have any suggestions that would help L&D identify additional opportunities in your department to measure employee performance for the purpose of evaluating training program effectiveness?

We need to focus on deal production. I would like to see a pseudo “sales” training to sell customers on workout options. This should always be in the back of agents minds on every call. Additionally, I would like to see uptraining developed for specific areas such as modification talk-offs, short sale talk-offs, etc.

Appendix: Data Collection Instruments

Manager Follow-up Interview

Question 2: List any methods that you currently use to evaluate employee performance on the job (job tasks only). Answer the questions below for each included method.

a. ***Do you record and track this data?***
b. ***If the data is recorded and tracked, how is this done?***
c. ***How often does your department do this?***
d. ***Can this data be made available to L&D on a periodic basis?***
e. ***Can you provide an example with this survey?***
f. ***What is the minimal acceptable level of performance?***

Agents are evaluated based on their average handle time and production.

Follow-up Question: What is Average Handling Time?
Average Handle Time is a combination of talk time + wrap time + hold time. Handle Time measures the total time each agent spends on the calls he or she receives.

- Talk time measures the amount of time the agent actually spends communicating with customers
    - No baseline*
- Wrap Time measures the time an agent spends after he or she hangs up with a caller before the agent receives a new call.
    - Baseline = 30 sec.
- Hold time is the amount of time that the agent placed a caller on hold for the day.
    - Baseline = 1 hour/day

*Previously the baseline for average handle time was 10 minutes for an experienced agent and 12 minutes for an agent in the training bay. Management is transitioning from a purely quantity based measurement standard to a quality + quantity based standard. Essentially, management does not want to establish a strict baseline for average handle time as this may impact call quality. Average call times for agents are now viewed collectively. Any outliers are further analyzed and addressed as needed, but the agents' average handle time is not compared to a static measurement.

Follow-up Question: What is Production?
Production refers to the daily number of loan workout plans set up by an agent. This description is contrary to what was reported by the training bay supervisors,

but was confirmed by both managers. The types and amounts of plans depend on the level of the call center agent.

- FPSI & FPSII – 5 promise to pays per day; 5 repayment plans per day
- FPSIII & FPSIV – 3 loan workout plans per day

The difference in the output above is based on the severity of default for the calls routed to the level three and level four agents and the resulting complexity of the workout plans the higher level agents use to resolve the defaults. Production reports are generated daily and show a current month to previous month comparison for agents individually and collectively.

Question 3: Do you currently analyze evaluations to identify trends of poor performance, such as analyzing QA scorecards, for employees in the Default Call Center? Can you share these results with L&D?

Management reviews many reports including QA Trending Reports, Status & Production Reports, and Error Reports, as well as internal performance assessments—Repayment Plan checklists, QA Form (side-by-sides, recorded calls, live monitoring).

Follow-up Question: Describe the Error Reports?
The description provided for this report is consistent with that provided by the training bay supervisors:

There are two versions of Error Reports: HMP and Repayment Plans. These grade a new agent on his or her ability to correctly process the most common loan workout options. The report lists the steps required to process each and grades the actual loan workout options processed by the agent for completeness and correctness relative to each step of the process. Essentially, this tells the agent and the agent's supervisor what the agent does incorrectly. These reports can be provided to the training department.

Follow-up Question: Describe the QA Reports?
QA scorecards are utilized by the QA Department to measure recorded calls. QA scorecards measure call quality for the following categories:

- Greeting
- Financials
- Program Knowledge
- Status Call Handling
- Closing

- Professionalism

An additional category for serious infractions or omissions, such as inappropriate behavior that "discredit" a call is also included.

Management and training bay supervisors may score live calls. The QA Department scores 5 agent calls, selected at random per month. Management and training bay supervisors do not score agents consistently with the QA scorecards. Thus, no baseline numbers for sample calls scored by management or training bay supervisors in a given time frame is available.

Question 4: Does your department currently have a defined baseline level of employee performance (job tasks only)? Can you provide this with this survey?

Follow-up Question: Can you describe the outline that you mentioned in the survey?
This is a checklist that each agent must complete to graduate from the training bay. Essentially it is a list of topics for which the agent must have demonstrated mastery prior to graduating to the general call center. The training bay supervisors use the quizzes and QA scorecards to evaluate mastery of each topic.

Question 5: How are employees evaluated in the classroom to determine if they have performed well enough to move on to the Training Bay?

Follow-up Question: How are employees evaluated in the classroom to determine if they have performed well enough to move on to the Training Bay?
One manager responded that this process is the same as the evaluation methods described in question 2. [I believe he misread the question. It has been verified with both trainers and discussed the response with the manager and concluded that performance is only evaluated via a test and trainer perception in the classroom.]

Question 6: How are employees evaluated in the Training Bay to determine if they have performed well enough to move on to the Call Center?

Follow-up Question: How are employees evaluated in the Training Bay to determine if they have performed well enough to move on to the Call Center?
Trainees are evaluated via quizzes. The criteria for graduation are a combination of the agents' performance on the quizzes and an informal assessment of their performance by the training bay supervisor.

Appendix: Data Collection Instruments

Question 7: Do you have any suggestions that would help L&D identify additional opportunities in your department to measure employee performance for the purpose of evaluating training program effectiveness?

Follow-up Question: Would it be possible for managers to complete a survey periodically that elicits the manager's perception of the agents' performance?
Yes.

Appendix: Data Collection Instruments

Training Bay Supervisor Survey: Indianapolis

1. Please indicate your name, department and position.

Ta-Tanisha Thames , Training Bay Supervisor In Indianapolis, IN

2. List any methods that you currently use to evaluate employee performance (job tasks only). Answer the questions below for each included method.
    a. Do you record and track this data?
    b. If the data is recorded and tracked, how is this done?
    c. How often do you do this?
    d. Can this data be made available to L&D on a periodic basis?
    e. Can you provide an example with this survey?
    f. What is the minimal acceptable level of performance?

Method 1: Schedule Adherence a. yes b. workforce management c. every day d. yes e. yes f. you are allowed less than 3 Occurances
Method 2: Average Handling Time & Wrap Time a. yes b. by reports c. daily d. yes e. yes f. 12 mins
Method 3: Quiz a. yes b. this is not tracked given back to agent c. d. e. f. There is no “magic number” but I train agents that they should not stand out from their peers. Avg for department right now is about 10-11 mins

3. What methods do you currently use to identify trends of poor performance in the Training Bay? Can you share these results with L&D?

When taking quizzes, Error reports, Production reports, Schedule Adherence reports. Yes we can share results.

4. Does the Training Bay currently have a defined baseline level at which an employee should perform (job tasks only) prior to graduating from the Training Bay?
    a. Can you provide this with this survey?

Yes we have a defined outline on what needs to be mastered before leaving the training bay. Yes we can provide the survey.

5. How are employees evaluated in the classroom to determine if they have performed well enough to move on to the Training Bay?

QA results, monitoring calls, training bay checklist, production reports, interaction during question and answer period, how they handle calls.

6. How are employees evaluated in the Call Center to determine if they are performing at an acceptable level?

Same answer for question 5

7. Do you have any suggestions that would help L&D identify additional opportunities to measure employee performance in your department?

Call control

Appendix: Data Collection Instruments

Training Bay Supervisor Survey: Denver

1. Please indicate your name, department and position.

Sonja Williams, Training Bay Supervisors Denver Branch

2. List any methods that you currently use to evaluate employee performance (job tasks only). Answer the questions below for each included method.
    a. Do you record and track this data?
    b. If the data is recorded and tracked, how is this done?
    c. How often do you do this?
    d. Can this data be made available to L&D on a periodic basis?
    e. Can you provide an example with this survey?
    f. What is the minimal acceptable level of performance?

Method 1: Shedule Adherence a. Yes b. It is tracked by workforce management c. Daily d. Yes e. Yes f. Less than 3 Occurances
Method 2: Quiz a. No – Quizes are given back to the agents to use as tools on the job b. c. d. e. f.
Method 3: Average Handling Time & Average Wrap Time a. Yes b. Reports c. Daily d. Yes e. Yes f. 12 minutes

3. What methods do you currently use to identify trends of poor performance in the Training Bay? Can you share these results with L&D?

> Error reports, Repayment Plan checklists, Productions Reports, Schedule Adherence Reports. Yes we can share those results.

4. Does the Training Bay currently have a defined baseline level at which an employee should perform (job tasks only) prior to graduating from the Training Bay?
    b. Can you provide this with this survey?

> We have a guide in which we utilize. Yes I can provide this with the survey.

5. How are employees evaluated in the classroom to determine if they have performed well enough to move on to the Training Bay?

> Its based on a combination of schedule adherence, production results, training bay checklists, QA results, monitoring calls, Quizes, The questions they ask and how they handle their calls.

6. How are employees evaluated in the Call Center to determine if they are performing at an acceptable level?

> They are evaluated by the same principals as listed above with the exception of not requiring quizes.

7. Do you have any suggestions that would help L&D identify additional opportunities to measure employee performance in your department?

> Assisting in structuring modules and testing materials.

Appendix: Data Collection Instruments

Training Bay Supervisor Follow-up Interview

Question 2: List any methods that you currently use to evaluate employee performance (job tasks only). Answer the questions below for each included method.

- ***a. Do you record and track this data?***
- ***b. If the data is recorded and tracked, how is this done?***
- ***c. How often do you do this?***
- ***d. Can this data be made available to L&D on a periodic basis?***
- ***e. Can you provide an example with this survey?***
- ***f. What is the minimal acceptable level of performance?***

Follow-up Question: Explain what the Average Handling & Wrap Time Report is and what type of information is provides.
Average Handling & Wrap Time Report – This report includes two measurements of an agent's efficiency receiving and concluding calls. Average Handling Time measures the time each agent spends on the calls he or she received the previous day. Wrap Time measures the time an agent spends after he or she hangs up with a caller before the agent receives a new call. The Wrap Time is typically used to record notes or perform other tasks resulting from the call. While these times will vary from call to call, the agent's average times should be within a standard range. A manager in the Default Call Center compiles and distributes this report daily to all agents including those in the training bay and those in the general population.

Follow-up Question: When do you administer the quizzes?
Most of an agent's time in the training bay is spent taking actual calls. Twice per day agents participate in 30 minute classroom sessions that focus on a single topic. Quizzes are administered at the beginning of the six week training bay session and then again at the end of the session. The final quizzes are used to evaluate the agent's readiness to graduate to the general call center. Each location uses the same quizzes. Currently the quizzes are graded and returned to the agents. The quiz scores and individual responses to each question are not currently recorded.

Question 3: What methods do you currently use to identify trends of poor performance in the Training Bay? Can you share these results with L&D?

Follow-up Question: What are the Production Reports?
These are the same as the Average Handling Time Reports.

Follow-up Question: Describe the Error Reports?

Appendix: Data Collection Instruments

There are two versions of Error Reports: HMP and Repayment Plans. These grade a new agent on his or her ability to correctly process the most common loan workout options. The report lists the steps required to process each and grades the actual loan workout options processed by the agent for completeness and correctness relative to each step of the process. Essentially this tells the agent and the agent's supervisor what the agent does incorrectly. These reports can be provided to the training department.

Follow-up Question: Are there any other methods that you think are relevant?
The training bay supervisor also performs side-by-sides with each agent. This is an observation of the agent by the supervisor as the agent executes a live call. The agent's performance is recorded using QA (Quality Assurance) scorecards. QA scorecards are utilized by the QA Department to measure recorded calls. In the training bay side-by-sides are performed only when time allows. This is based on call volume and the number of agent's who are currently under supervision in the training bay.

Question 4: Does the Training Bay currently have a defined baseline level at which an employee should perform (job tasks only) prior to graduating from the Training Bay? Can you provide this with this survey?

Follow-up Question: Can you describe the outline that you mentioned in the survey?
This is a checklist that each agent must complete to graduate from the training bay. Essentially it is a list of topics for which the agent must have demonstrated mastery prior to graduating to the general call center. The training bay supervisors use the quizzes and QA scorecards to evaluate mastery of each topic.

Question 5 & 6: No follow-up questions

Question 7: Do you have any suggestions that would help L&D identify additional opportunities to measure employee performance in your department?

Follow-up Question: Would it be possible for the training bay supervisors to complete a survey at the conclusion of each training bay session that elicits the supervisor's perception of the agents' performance?
Yes. The supervisor's may not always be able to provide quantifiable records of an agent's performance, but because they work with the agent's eight hours a day for six weeks they often understand where agents are struggling.

Appendix: Data Collection Instruments

Trainer Survey: Indianapolis

1. Please indicate your name.

Alan Taylor

2. List any methods that you currently use to evaluate training participant performance (job tasks only). Answer the questions below for each included method.
    a. Do you record and track this data?
    b. If the data is recorded and tracked, how is this done?
    c. How often do you do this?
    d. Can this data be made available to the ID team?
    e. Can you provide an example with this survey?
    f. What is the minimal acceptable level of performance?

Method 1: course assessment a. yes b. in CMS c. at the end of each class d. yes e. yes, all scores are available in CMS f. 80%
Method 2: class participation Q & A; daily debrief a. n/a b. n/a c. daily d. no e. no, I compose questions based upon the previous day's learning and tailor them for each group f. no standard; however, if a participant answers incorrectly, I coach them to use the available resources to determine the correct answer
Method 3: participate in subjective review of each participant with supervisors a. verbally b. n/a c. after each class d. n/a e. since review is verbal, no data is available f. n/a

Appendix: Data Collection Instruments

3. What methods do you currently use to identify trends of poor performance in the classroom? Can you share these results?

Daily debrief and Q&A; end of course assessment available on department G: drive; also make determinations of ability to navigate various systems trained in class based upon participant questions, primary observation; additionally I have participants use the facilitator’s workstation (and projector) to navigate systems using role play-type questions

4. How are employees evaluated in the Call Center to determine if they are performing at an acceptable level?

Department score cards including QA scores, schedule adherence, # of promises to pay, repayment plan starts, etc.

5. How are employees evaluated in the Training Bay to determine if they have performed well enough to move on to the Call Center?

Since I’m currently not involved in the training bay, unable to address this question

6. Do you have any suggestions that would help us identify additional opportunities to measure a training participant’s performance in the classroom?

Provide a computer-based assessment which would more accurately track success and coaching opportunities; provide additional behavior-based assessment involving role play, “what if” scenarios, system navigation involving a customer scenario which would include a step-by-step review regarding using available systems in the correct order in recording customer comments, financials, workout eligibility, etc.

Appendix: Data Collection Instruments

Trainer Survey: Denver

1. Please indicate your name.

Jered A. Wagner

4. List any methods that you currently use to evaluate training participant performance (job tasks only). Answer the questions below for each included method.
    a. Do you record and track this data?
    b. If the data is recorded and tracked, how is this done?
    c. How often do you do this?
    d. Can this data be made available to the ID team?
    e. Can you provide an example with this survey?
    a. What is the minimal acceptable level of performance?

Method 1: Job Skill Role plays a. Informally, yes b. I will provide the training bay supervisor with an end of class summary as to issues that remain to be focused upon going forward. c. Whenever such issues need to be identified which is dependant upon participant performance. d. I have no formal record of prior communication but I can gladly address questions pertaining to my process. e. When my current class concludes on 4/07/2009, I can provide a sample at that time. f. I use the Quality Assurance standards to gauge if the minimal acceptable level of performance is being met.
Method 2: End of class assessment a. Results are recorded within CMS b. Manually inputting into CMS whenever a class concludes. c. Whenever a training session concludes d. It can be as all assessments are held in storage for a one year period. e. I can provide either copies of the assessments from the last HRG new hire class or if desired, I can provide the full record of all assessments from the last year. f. 80% for passing grade.
Method 3: a. b.

c.
d.
e.
f.

3. What methods do you currently use to identify trends of poor performance in the classroom? Can you share these results?

The answers provided for question number two of this survey would seem to lend themselves to question number three. I'm uncertain as to what additional information would fit this question more specifically.

4. How are employees evaluated in the Call Center to determine if they are performing at an acceptable level?

Monthly Quality Assurance reviews (five calls reviewed per employee, per month); call center statics tracked (handle time, wrap time, etc.)

5. How are employees evaluated in the Training Bay to determine if they have performed well enough to move on to the Call Center?

Sonja Williams would need to address this more readily as new hires have been graduated on some occasions out of necessity to compensate for capacity restrictions (the bay can only hold so many new hires and larger classes may compel an early "graduation")

6. Do you have any suggestions that would help us identify additional opportunities to measure a training participant's performance in the classroom?

Pre assessment to determine skill set leading into job skill learning process. Revised assessment process to better capture job skill use rather than information retention.

Appendix: Data Collection Instruments

Trainer Follow-up Interview

Question 2: List any methods that you currently use to evaluate training participant performance (job tasks only). Answer the questions below for each included method.

a. ***Do you record and track this data?***
b. ***If the data is recorded and tracked, how is this done?***
c. ***How often do you do this?***
d. ***Can this data be made available to the ID team?***
e. ***Can you provide an example with this survey?***
f. ***What is the minimal acceptable level of performance?***

Follow-up Question: Are the individual scores for each question also available in the CMS?
Comprehensive assessments are administered at the conclusion of each three week training class. The score is entered into the CMS [commonly referred to as an LMS]. Only the final score is recorded in the CMS and not the individual responses to each question. The original, graded copies of the tests are stored for one year and are available to the instructional design team. 80% is considered a passing score for the class, but each question is individually scored and graded, which is indicated on the original cop y of the test by the trainer.

Follow-up Question: Is employee performance during the job skill role-plays recorded and saved?
Additionally, each trainee participates in periodic job skill role-plays throughout the duration of the class. Trainee performance is not formally recorded for these role-plays. In some cases, the QA scorecard is used as to grade the performance of a trainee during the role-plays. In a class with less than five participants the QA scorecard is administered by the trainer. In classes with more than five participants the scorecard is typically administered by peers. These scorecards are not saved. They are returned to the participants.

Follow-up Question: Can you explain more about the Subjective Review that you referred to in the survey? Is this a formal review or just something that is provided informally as needed?
Trainers typically provide a subjective review of each participant’s performance to the participant’s supervisor at the conclusion of the training class. This is neith er formalized nor recorded. The review is the trainer’s informal impression of the trainee’s performance during the three week training class.

Appendix: Data Collection Instruments

Question 3: What methods do you currently use to identify trends of poor performance in the classroom? Can you share these results?

Follow-up Question: Are QA Scorecards and subjective reviews performed only for individuals or are they compiled to reveal an overall group performance? The group performance would reveal a "trend" or an indicator that multiple learners are consistently performing poorly on the same topic.
This is also informal and is mostly conveyed via the subjective review described in the previous question. Trainers verbally assess the learners based on the leaner's performance on the QA scorecards and during the role-plays. Nothing is recorded and results are not compiled to reveal a group's performance.

Question 4: How are employees evaluated in the Call Center to determine if they are performing at an acceptable level?

Follow-up Question: Are these the same QA Scorecards that are used occasionally in the classroom?
Yes. Agents are assessed using QA scorecards up to five times per employee per month. The call center also uses schedule adherence reports (Average Handle Time) and promise to pay and repayment plan (production reports).

Question 5: How are employees evaluated in the Training Bay to determine if they have performed well enough to move on to the Call Center?

No Follow-up Question
Not sure (both trainers).

Question 6: Do you have any suggestions that would help us identify additional opportunities to measure a training participant's performance in the classroom?

Follow-up Question: Would it be possible for trainers to complete a survey periodically that elicits the trainer's perception of the agents' performance?
Yes