Transforming Online Assessments: Creating Alignment with Objectives

Ken Thomas

Final Report

Version 2.0

INTE 5650 Policies & Planning for eLearning Programs

Background

The Client

The Client is a health care benefits management company with several contact centers in different states within the United States. As with most contact center environments, they suffer from heavy attrition, even during the current economic downturn. They find themselves hiring and training over 1,000 employees every year, including temporary and seasonal help, who may return year after year for peak months of call volume.

Problem description

The Client recently transitioned from printed assessments to online assessments in their multi-week New Hire Curriculum for Customer Service Representatives (CSRs). The goals driving this transition were to:

- Create standardized assessments that could be distributed across all sites immediately upon publication to the LMS.

- Replace manual scoring and reporting with automated scoring and recording through the LMS.

During their evaluation of the assessments, they found that although they met these important goals, there was no predictive validity to the assessments (i.e., a learner who performed well on the assessments did not necessarily perform well in the job, and learners who performed poorly on the assessments did not necessarily perform poorly on the job).

Without predictive validity, the assessments, no matter how easily administered, scored, tracked, and reported, provide no value to the organization. If this problem is left unresolved, the Training Department will continue to graduate unprepared CSRs to the floor, which will have a continued negative impact to contact center productivity and customer satisfaction.

Root cause

A one-week analysis of the job, New Hire Curriculum, and related assessments revealed the following root causes of the problem:

- The Terminal Learning Objectives (TLOs) of the New Hire Curriculum were defined to the wrong level of Bloom et al.'s Taxonomy of the Cognitive Domain (Bloom, 1956).
  - The primary job tasks of the CSR involve working with a customer on the phone while using a variety of systems to review the customer's record, resolve the customer's

problem, present the solution to the customer, and implement the agreed upon solution.
  - The TLOs derived from these tasks should be written to the Application and Synthesis level of Bloom et al.'s Taxonomy of the Cognitive Domain. Instead, they are written to the Knowledge and Comprehension level.
  - Assessments are intended to prove mastery of the objectives, but since the objectives are written to the wrong level and do not accurately reflect the job tasks, success on the assessment would not necessarily predict success on the job. For example, a newly hired CSR may be able to list the key events of the billing cycle successfully on the test, but may not be able to interpret an actual bill or explain that bill to a customer.

- Designers creating assessments are only using multiple choice questions in their assessments.
  - Although the Client's selected assessment development toolset (Articulate Quizmaker '09) supports 11 graded question types, they have decided that multiple choice questions offer the easiest reporting to interpret.
  - While some TLOs can be assessed using multiple choice questions, many would be better assessed through other interaction types. For example, the task of finding a customer's record would be better assessed through a scored simulation than a multiple choice question.

- Assessments are not representative of the current TLOs.
  - Designers find some TLOs are easier to generate questions to test, so create tests reflecting only these TLOs.
  - Designers often choose obscure "factoids" from the readings, rather than questions that relate to the objectives.

- Individual questions are poorly constructed.
  - Designers fall prey to mistakes made by juniors (loaded questions, use of "always" and "never," clues in the stem, non-parallel distracters, etc.).
  - Some designers are delighted by writing "trick questions" (questions intentionally designed to penalize the learner for not reading the question properly), and actually brag about their high failure rates.

My role

I was initially brought in by the client to help redesign their New Hire Curriculum. They have now asked me to expand my scope to include revising their assessment strategy and tools.

The Intervention - Overview

Goals for the intervention

The goals for the intervention are to:

- Define Terminal Learning Objectives (TLOs) and Enabling Learning Objectives (ELOs) based directly on a job/task analysis.

- Transform the current assessment approach of the Designers to:
  - Align assessment items to TLOs and ELOs (and ultimately the job/task).
  - Design representative assessments of the TLOs & ELOs.
  - Construct instructionally sound assessment items.

Intervention components & rationale

The main components of this intervention are the job/task analysis and an assessment development workshop for Designers. The following table illustrates the relationship between these components, the root causes of the problem, and goals of the intervention:

Root Cause	Goal	Intervention Component
The Terminal Learning Objectives (TLOs) of the New Hire Curriculum were defined to the wrong level of Bloom et al.'s Taxonomy of the Cognitive Domain.	Define Terminal Learning Objectives (TLOs) and Enabling Learning Objectives (ELOs) based directly on a job/task analysis.	Job/Task Analysis (output: TLOs & ELOs)
Designers creating assessments are only using multiple choice questions in their assessments. Assessments are not representative of the current TLOs. Individual questions are poorly constructed.	Transform the current assessment approach of the Designers to: • Align assessment items to TLOs and ELOs (and ultimately the job/task). • Design representative assessments of the TLOs & ELOs. • Construct instructionally sound assessment items.	Designing Instructionally Sound Web-Based Assessments (Workshop)

Paradigm shift

While both components represent a paradigm shift to how the Training Department designs and develops training, they have embraced the analysis approach (and results), and welcomed a new approach to assessment.

Note: *It should be noted that while this represents a paradigm shift for the Training Department, the rest of the organization does not require any level of change management communication or effort. The next level of management did need to approve funding.*

Stakeholders and risk

The biggest risk to successful implementation was the reaction and support of one of the key stakeholders of the project: the Client's Director of Instructional Design, who designed the current curriculum without proper analysis (stating he did not have time to do an analysis or properly define objectives) and who made the decision to only use multiple-choice questions because "other question types are too hard to write" and "multiple choice questions yield better reports."

In the Innovation-Decision Process Model (Rogers, 1995), this stakeholder failed to properly execute the "Knowledge" stage of the model, and has based the remaining stages on an improper understanding of assessment development. The fear was that his ego would prevent him from admitting such a foundational and fundamental mistake, and any attempt to point out the failed reasoning risked disaster for the project.

Luckily, being able to arm him with findings and recommendations allowed the Director to present the strategy to his managers, who agreed to support funding for the initiative.

The Intervention – Job Task Analysis

Methodology   The job task analysis consisted of the following activities:

- Review of the existing job descriptions

- Focus group discussion with selected supervisors

- Observations of Customer Service Reps (CSRs) ranging from new graduates to tenured top performers

- Review of quality ratings across the CSR population

- Focus group discussion with Quality Coaches

- Breakdown of all call types and call sub-types

- One-week session with top performing CSRs to validate findings and document the subtasks for all calls

Output Sample A: The following table is a sample output of the analysis – the call
Call types   types/sub-types breakdown:

Call Type	Call Sub-Type
Document Request	**Account Maintenance** • ID Card request • Other replacement fulfillment materials **Claims** • Medical Explanation of Benefits (EOB) request • Pharmacy Explanation of Benefits (EOB) request **Letters/Correspondence** • HUD Letters • Other custom generated letters • Premium account history request • System generated replacement
Prescription Benefits	**Prescription Benefit Level** • Prescription plan benefits / cost explanation • OOP (out of pocket) / TROOP (true out of pocket) • Pharmacy lookup • Drug lookup / Tier – Formulary **Prior Authorization Status**
Medical Benefits	**Facility/Physician** • Check Prior Authorization Status for Medical Services • Facility/Physician lookup & Par (Participating) Status

Call Type	Call Sub-Type
	(Primary Care Physician (PCP) or Specialist)
	Medical Benefit Level • Limited serviced (e.g., number of visits, number of trips, days of hospital stays) • Medical plan benefits / cost explanations (copay, coinsurance, cost sharing) • OOP (out of pocket) max **Passport** • Activate/Deactivate Passport • Passport status **Riders** • Dental • Deluxe (dental, hearing, & vision) • Fitness
Account Maintenance	**Auxiliary Representative** • Request to add, change, or delete an authorized representative **Deceased Member** • Erroneous date of death (DOD) • Notification of date of death **Demographics** • Change date of birth (DOB) • Change incorrect/changed name • Change permanent address • Change email address • Change mailing address • Change temporary address • Change phone number • Change language preference (spoken/written materials) **Effective dates (send to Enrollment group)** • Change enrollment effective date • Change disenrollment effective date **General** • Set do not call, do not mail, and/or do not email **Low Income Subsidy (LIS)** • Update/change LIS **Primary Care Physician (PCP)** • Change Primary Care Physician

Call Type	Call Sub-Type
Enrollment/Disenrollment	**Application status** • Status of application • Denial reason • Missing information **Disenrollment** • Request to disenroll or withdraw application • Disenrollment status • Reinstatement • Rider (Covered in Benefits - Riders) **Enrollment** • Rider (Covered in Benefits - Riders) **Late Enrollment Penalty (LEP)** • Complete attestation • General questions **Membership status** • Effective date • Plan change • Reinstatement **Prospects** • General questions
Premium Billing	**Payment method** • Change method of payment (MOP) & MOP discrepancy • Electronic Funds Transfer (EFT) mis-key changes • One-time EFT • SSA issues **Payments** • Misapplied (split payment) • Missing payments • Premium refund status • Request premium refunds • Status • Education
Prescription Claims	**Coordination of Benefits** • Claims processing issues **Coverage determination** • Administrative override • Clinical override

Call Type	Call Sub-Type
	Denied • Prescription denial inquiries **Direct Member Reimbursement (DMR)** • DMR requests/general education **Explanation of Benefits (EOB)** • Education **Copay** • Incorrect copay issues
Medical Claims	**Appeals** • Appeal status • Denied claim appeals • PreAuthorization/ notification appeals **Coordination of Benefits** • Primary vs. secondary **Cost sharing** • Balance billing issues • Payment responsibility **Denied** • Medical denial questions **Payments** • Direct Member Reimbursements (DMR) **Explanation of Benefits (EOB)** • EOB/PRA claim discrepancy
Other	**Appeals and Grievances (A&G)** • Quality of Care (QOC) grievance • Quality of Service grievance • Standard appeal • Expedited appeal **Website assistance** • Locked out • Navigation

Output Sample B: Objective clusters   The following list shows how the learning objectives are clustered:

Call-derived objectives *(i.e., TLO will relate directly on the call type)*
• Document Request

- Prescription Benefits
- Medical Benefits
- Account Maintenance
- Enrollment/Disenrollment
- Premium Billing
- Prescription Claims
- Medical Claims
- Other

Call-related objectives *(i.e., TLO supports more than one specific call type)*
- Warm/cold transfer
- Access and logon to systems
- Navigate systems
- Access, navigate, and use call-support tools (wizards, job aids, & macros)
- Find and use latest process & procedure documentation
- Use phone

Non call-related objectives *(i.e., TLO supports a job task that is not related to handling phone calls)*
- Schedule PTO
- Use Outlook
- Send, receive, and organize emails.
- View updates

Output Sample C: Sample TLO

The following block is a sample breakdown of a call-related TLO (clustered into phases of the call flow):

Call Type:	Drug Lookup
Indicator/driver:	Member has a new prescription and wants to know if it's covered, and if so what the copay will be.
TLO:	Handle a drug lookup call.
ELOs:	

ELOs:

Open the call
- Greet the caller.
- Classify a call as a drug lookup.
- Provide acknowledgement.

Verify HIPAA

- Identify applicable HIPAA verification procedure.
  – Find member account.
  – Identify originator type.
  – Determine appropriate HIPAA verification requirements.
- Implement verification procedure.
- Determine if caller is qualified to receive drug lookup information.
  – If not, recommend Secure Horizons website.

Investigate the call
- Identify Contract Number.
- Identify appropriate resources.
- Identify resolution options.
- Access the appropriate formulary in RxWeb.phs.com
  – Access RxWeb Formulary.
  – Access MAPD Formularies.
  – Select the member's contract.
  – Determine which formulary to use.
  – Select appropriate formulary.
- Find the requested drug.
- Interpret drug page.
  – Determine if drug is in formulary.
  – Determine drug's tier.
  – Determine if PA is required (coverage determination).
- Determine LIS Subsidy Level.
- Price drug.

Communicate resolution
- Present the resolution.
  – Present tier.
  – Present covered/not covered.
  – Present alternatives.
  – Provide any special instructions (prior authorization required and how to obtain, how to appeal a rejected PA, amount limitations, etc.).
- Offer options.
- Gain agreement.
- Review and set expectations.

Close the call
- Provide offer of further assistance.
- Use standard closing remarks.
- Document the call (ongoing throughout call).

The Intervention – Workshop

Overview   Although a full workshop was not approved, I was given a two hour block to discuss the findings of the job task analysis to the Instructional Design Team, and was able to lead a discussion on assessment practices, encouraging a shift from their current approach to a more "aligned" approach.

Alignment   The theme of the discussion was alignment. The following graphic served as the master organizing image for the discussion:
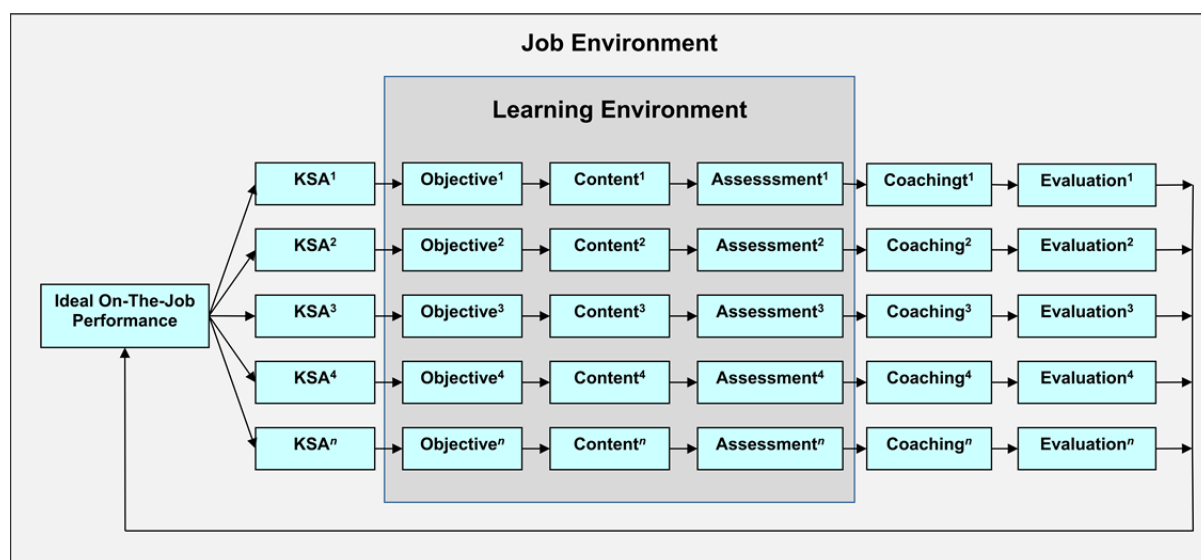
The graphic is interpreted from left to right as follows:

- First, define the ideal on-the-job performance as it is or should be performed within the job environment.

- Break down this ideal performance into the Knowledge, Skills, and Attitudes (KSAs) required to successfully perform the job.

- Translate these KSAs into Terminal and Enabling Learning Objectives.

- Design and build courseware to achieve these objectives.

- Build assessment items that not only reflect the content, but also directly link back to the objective, the KSA, and ultimately the ideal on-the-job performance.

- When the learner is released from training, they should be coached and evaluated/rewarded based on the defined ideal behavior.

Inferences from the alignment model

Once the designers understood the alignment model, we were able to discuss the strengths and weaknesses of various assessment strategies.

We then pulled up some of their current multiple choice questions and tried to apply the alignment to determine what on the job performance they supported and what objective they demonstrated mastery of. They realized that by selecting "factoids" from the content (which was not built based on analysis in the first place), they had created assessments that failed to demonstrate mastery of objectives and did not adequately qualify a learner's ability to perform the job.

Getting designers out of their box

With the foundations laid for a philosophy of aligning assessments to the job and derived objectives, I shifted to discuss how their current assessment development tool could be leveraged without significant additional learning to achieve this approach. The key strategies we brainstormed included:

- Instead of thinking of interactions as individual units, consider presenting a series of interactions together to simulate the job with greater realism.
  - For example, a formulary call ("Is drug ABC covered under my plan?") requires the CSR to:
    - Access a customer's record.
    - Navigate to a particular screen.
    - Interpret data on that screen.
    - Determine whether or not the drug is covered.
    - Tell the customer whether or not the drug is covered.
  - Currently, this is assessed using the following multiple choice question:
    What is the tool you will use to determine whether or not a member's drug is covered by their plan?
    A. Drug list
    B. Drug lookup
    C. Rx.com
    D. Formulary
  - A new approach could be a series of screens as follows:
    - Leverage screen captures and a combination of hot spot and fill-in-the-blank interactions to create basic simulations of accessing the member's record and navigating to their formulary.
    - Have the learner be able to toggle back and forth between the member's formulary screen and the simple

  2 choice multiple choice question, "Is this member's drug covered by their plan? Yes/No.
- Provide an open text field for the learner to type the sentence they would use to provide the answer to the member.

- Include links to references available on the job.

- Leverage live role-plays for hard-core communication skills that are challenging or impossible to assess using the online assessment tools (e.g., deescalating an angry caller).

- Make sure the assessment measures mastery of all the course's TLOs in a representative and distributed way.

Brainstorm

Designers had been asked in advance to bring assessments they were either currently working on or the assessment from their previous project. We then conducted a brainstorming session to discuss how to bring their current assessment into alignment.

The biggest challenge with this activity was that the designers did not have defined objectives for their courses. We were able to overcome this by defining draft objectives based on the job/task or policy they were designing training for, then working from them.

Each designer was able to present a strategy for leveraging multiple interaction types, clustering interactions together, and getting closer to the objective in context of the job than their previous factoid recall only multiple-choice questions.

Closing

As a closing, each designer created a brief learning plan that outlined their strategy to learn how to create the required new interactions using Articulate Quizmaker to create an aligned assessment.

Evaluation Plan

Perfect world – predictive validity	In a perfect world, the success of this intervention would be a pure Level 4, where we could measure the change of the assessment's ability to predict learner success or failure on the job, where this intervention would be a dependent variable.

However, because of the other initiatives being implemented concurrently (e.g., redesign of the New Hire Curriculum), it will be challenging to separate the impact of this portion of the intervention.

Note: *That's not to say I won't include the predictive nature of the assessment in the overall evaluation approach.*

Measuring face validity

The primary evaluation strategy for this intervention (at least by the end of the current semester of ePlanning) will be to measure the face validity of the final assessments created by the Designers after receiving the TLOs and ELOs, and completing the Designing Instructionally Sound Web-Based Assessments Workshop.

An evaluation form will be created to measure basic criteria of an instructionally sound assessment (Thomas, 1988), such as:

- Alignment of question type to objective (e.g., using a simulation to measure a system action), and ultimately to the job itself
- Assessment's distributed representation of the objectives
- Clarity of the question/interaction

Evaluation Results

Current status   The biggest challenge to formally evaluating the current assessments being rolled out after the workshop is that the courseware being assessed still does not have formally defined objectives (which is required in the current evaluation strategy).

Upon informal comparison of the previous assessments and the revised assessments clearly demonstrate a vast improvement of the designers' approaches to assessment. The revised assessments focus on the procedures being performed (e.g., clustered interactions, simple simulations, toggling between screens and interactions requiring interpretation of the screen) rather than the previous versions of simple factoid recall multiple choice questions.

Next steps   The Director of Instructional Design has asked to review the next round of courseware developed by each of the designers, ensuring that they have defined objectives based on the ideal on-the-job performance of the job. Then he will review their assessment strategy by each TLO. (He has promised to share these results with me, and to provide access to examples – he is reserving direct coaching as part of his role rather than mine.)

When the New Hire Curriculum is redesigned in 1st quarter 2011, I will work together with the Instructional Design Team again, and we can revisit alignment as we build the courseware and assessments based on the recently conducted job task analysis. We will also be tracking performance of new hires as they transition to the job for 6 months, allowing us an opportunity to formally measure the predictive validity of the new assessments based on alignment.

Lessons Learned

Reflections　　　　Working with this client is getting better every day as my primary point of contact sees me more as a partner than a competitor. The more he is able to demonstrate success to his managers and peer-clients, the happier they have been.

I usually enjoy much more of a partnership with my clients, so feel somewhat empty not being directly involved with some of the "fun" parts of this intervention (e.g., not being the one who will review the designers' next round of objectives and assessment development). I guess I need to better appreciate that this project has really been a growth opportunity for the director even more than for his designers.

I also have to learn to compromise more quickly when the client is not supporting the full solution I recommend.

If my goal was to improve the assessment strategy of the Instructional Design Team, I'd have to say this has been achieved with overwhelming success. I just need to learn how to let go of the control before the intervention is completely implemented in the case where the client ultimately needs to own the solution and the success.

Bibliography

Bloom, B. S. (Ed.) (1956). *Taxonomy of educational objectives: Handbook I: Cognitive domain*. New York: David McKay Co.

Rogers, E.M. (1995). *Diffusion of innovations, 4th ed.*. New York: Free Press.

Surry, D. W., & Ely, D. P. (2002). Adoption, diffusion, implementation, and institutionalization of educational technology. In Reiser, R. A., & Dempsey, J. V. (Eds.), *Trends and issues in instructional design and technology*. Upper Saddle River, NJ: Merrill/Prentice Hall.

Thomas, K.D. (1988). Developing certification tests. Unpublished course developed for Federal Bureau of Prisons.