The Application of Expert Review as a Formative Evaluation Strategy within an Educational Design Research Study

James L. Vesper
LearningPlus
USA
jvesper@learningplus.com

Thomas C. Reeves
The University of Georgia
USA
treeves@uga.edu

Jan Herrington
Murdoch University
Australia
j.herrington@murdoch.edu.au

Abstract: This paper describes how expert review strategies were applied to the testing and refinement phase of an educational design research project being conducted under the auspices of the World Health Organization (WHO). The overall goals of the educational design research project are to address the cold chain management challenge confronted around the world when vaccines and other perishable medicines are transported and stored, and to develop reusable design principles for the development of e-learning for the WHO. The paper focuses on the expert review process as it aligns with a design research approach, and in particular, on the instrument used to collect the expert opinion and data. Three highly qualified experts participated in this formative evaluation activity. The paper describes how their recommendations were analyzed and used by the WHO design team to refine their prototype design for a "Pharmaceutical Cold Chain Management on Wheels" e-learning course.

Introduction

When seeking to solve an education or training problem, there are many decisions that must be made that will influence the nature of any eventual learning solution such as the specific instructional strategies to be used and the type of delivery system to be deployed. Ideally, such important decisions will be based upon accurate and timely information that has been collected using a systematic approach to inquiry. Educational design research, defined as "a series of approaches, with the intent of producing new theories, artifacts, and practices that account for and potentially impact learning and teaching in naturalistic settings" (Barab & Squire, p. 2), is a unique type of inquiry increasingly being used to tackle complex education and training problems.

Educational design research (also referred to as *design-based research, design experiments*, and other names) (van den Akker, Gravemeijer, McKenney, & Nieveen, 2006) is recommended because it:

- Focuses on broad-based complex educational problems
- Requires collaboration between researchers and those directly involved with the problem of interest
- Integrates known and hypothetical design principles and technology in achieving a solution
- Utilizes rigorous and reflective inquiry to test and refine innovative learning designs and identify new design principles
- Involves improvement of the design through evaluation
- Contributes to both theoretical understanding while solving real world problems (Herrington, Reeves, & Oliver, 2010, p. 176).

Reeves (2006) presented a design-based research model consisting of four phases. Figure 1 shows the phases and how they inter-relate. This paper is focused on the part of the model shown in Figure 1 that involves the testing and refinement phase. Each cycle of the testing and refinement phase uses methods usually associated with formative evaluation within the context of instructional systems design, another widely used approach to addressing education and training problems (Reeves & Hedberg, 2003). The purpose of formative evaluation during educational design research is to provide the design team with the information required to make decisions and take the appropriate actions to refine or improve the prototype learning solution. Sample questions addressed during the testing and refinement cycles include:

- What can be done to improve the proposed visual design of the program?
- What can be done to improve the proposed interface design of the program?
- What can be done to improve the proposed instructional design of the program?
- What can be done to improve the proposed delivery system for the program?
- What can be done to improve the proposed content to be included in the program?

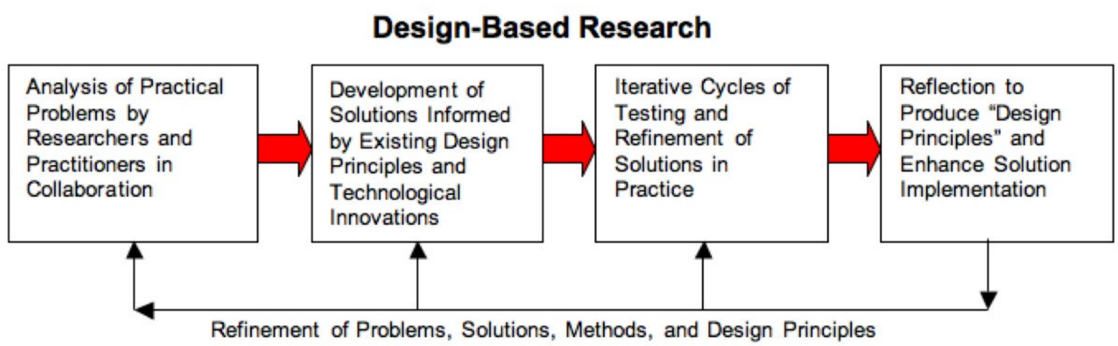


Figure 1: Four-phased design-based research model (Reeves, 2006, p. 59).

A variety of formative evaluation techniques can be used at different points of the testing and refinement cycles to provide information to the e-learning program designers and developers. For example:

- Experts can review sketches, instructional strategies, and activities;
- Potential learners can review detailed sketches, prototypes, and activities; and
- Actual learners or proxies can use and evaluate a "pilot" program (Dick, Carey, & Carey, 2008).

Each review cycle is intended to generate information that can be considered and incorporated by the design team into the learning solution. This paper describes how an expert evaluation was performed as part of an educational design research project being done in association with the World Health Organization (WHO).

Background

With the increasing development of biotech medicines and the growing use of vaccines, there is a greater concern in how these time and temperature sensitive pharmaceutical products (TTSPP) are transported, stored, and distributed to the end users (Milstien, Kartoğlu, & Zaffran, 2006). If a typical TTSPP is exposed to high temperatures, it can deteriorate, resulting in a lack of effectiveness and an increase in impurities. For many products, like human insulin and certain vaccines, freezing can cause immediate damage to the molecule rendering it inactive. There are individual and public health implications to this: a person's diabetes may not be controlled if they try to use insulin that was frozen; a national immunization campaign may be waste of time and money and result in illness and death if frozen vaccines were used (Ewbank & Gribble, 1993).

To keep these TTSPP at the proper temperatures (typically 2 to 8 degrees centigrade), a cold chain is utilized. A cold chain is the integrated system of equipment (e.g., shipping containers, refrigerators, trucks), procedures, records, and activities used to handle, store, transport, distribute, and monitor time-temperature sensitive products (Taylor, 2001). The allusion to a chain is very apt. As with a physical chain, a cold chain is only as strong as its weakest link. People are a critical element in the cold chain. Not only must people develop and execute procedures, people need to understand and control the risks that products may be exposed to. People also must quickly – and appropriately – address deviations or incidents that may occur so as not to expose the TTSPP to additional significant risks.

To help develop this type of expertise in those directly (e.g., vaccine manufacturers, public health professionals) and indirectly (e.g., packaging developers, engineers who design electronic temperature monitoring instruments) involved with pharmaceutical products, the World Health Organization's Global Learning Opportunities for Vaccine Quality (WHO GLO) developed a unique training course, Pharmaceutical Cold Chain Management on Wheels (PCCMoW), that takes 15 carefully selected participants on a bus trip in Turkey where they could make direct observations at the storage, warehousing, distribution and health care facilities that they visited as they physically travel with mentors down the length of the cold chain. (Turkey has been the location of the learning event because of the availability of sites to visit, the tradition of hospitality to guests, the course leader's network of contacts, and a local travel coordinator.) Throughout the course, guided observation exercises take place at the visited facilities. Participants are provided with notes and tools to support their critical observations. Participants interact with operational staff and management at these facilities. Presentations and group discussions take place on the bus, in restaurants, and in the open air before and after the visits to the facilities (Vesper, Kartoğlu, Bishara, & Reeves, 2010). Approximately 75 people (as of June 2011) have participated in the PCCMoW course, a very small number compared to the thousands of people world-wide who could benefit from gaining expertise in this field.

To increase the number of people who have expertise in handling TTSPP, in 2010, WHO began working on an e-learning solution to develop the expertise of those handling time/temperature sensitive pharmaceutical products. This project is anticipated to provide educational benefits to the users of the learning solution, and, if successful, result in health benefits for those who would use TTSPP. The first and second authors of this paper are members of the design team.

Methods

To structure and guide the testing and refinement of this learning solution, a plan was developed encompassing three rounds of formative evaluation:

- Experts examining early sketches of the course and its learning activities,
- Potential learners using a working prototype (an "alpha" review), and
- Actual learners participating in initial or pilot version of the course (a "beta" review).

This paper describes the first of these three rounds, expert review.

A fundamental question in the expert review process is answering the question, "who is an expert?" Tessmer (1994) gave very general guidance when he wrote, "The expert may be a content expert, teacher, technician, or subject sophisticate" (1994, p. 4). Dreyfus and Dreyfus (2005) described an expert as someone who doesn't have to ponder the options and how to accomplish them; he or she has a deep understanding of what to do and how to do it. They also wrote that an expert is someone who "knows what" (having high levels of declarative knowledge), "knows how" (procedural knowledge), and "knows when and where" (contextual flexibility). Ericsson, Prietula, and Cokely (2007) maintained that expert performance requires 10 years or 10,000 hours of deliberate practice.

Three experts were recruited for this study. Selection criteria for the "ideal" expert reviewers included:

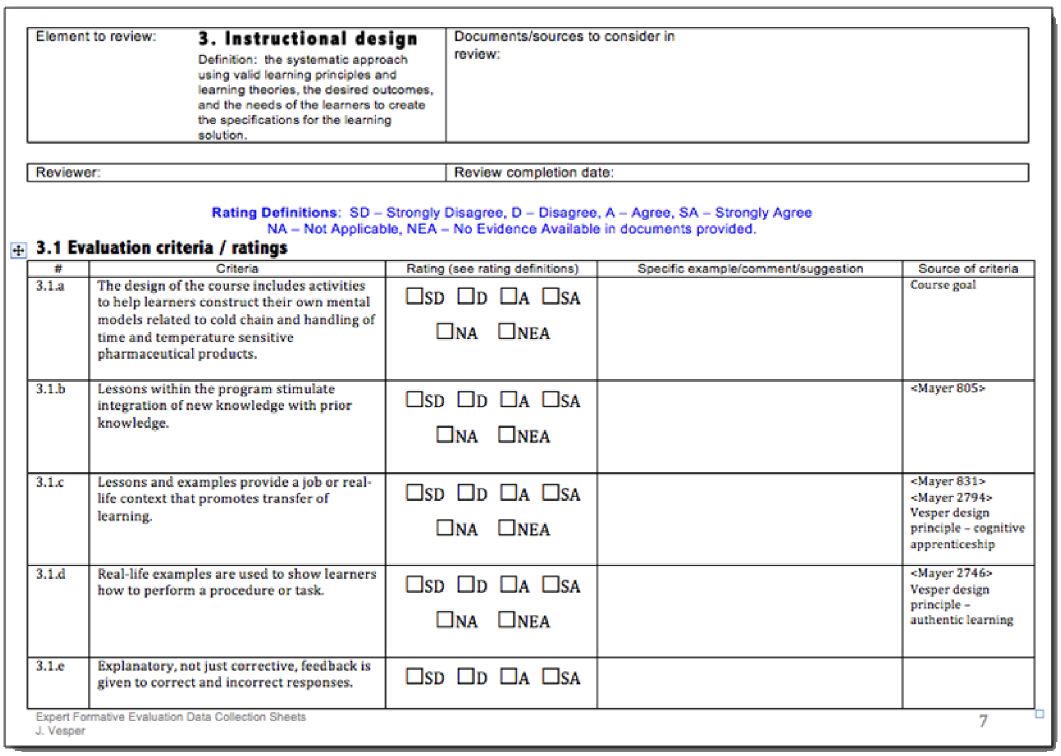
- Training and experience as instructional designers or as graphics/interface designers;
- Experience in a range of e-learning projects;
- Experience in designing and developing e-learning projects used in the life-sciences, particularly in the pharmaceutical/medical device industries and in healthcare delivery, and
- Experience living and working in other than North American cultures.

One expert ("RG") who was selected is a graphic designer with extensive experience (>20 years) in designing interfaces and the "look and feel" of e-learning programs that have been used by large, international pharmaceutical manufacturers such as The Lonza Group and Pfizer. This expert reviewed the overall visual design and the user interface design. Two experienced instructional designers were also used as experts. Each has more than 30 years working as an instructional designer producing a variety of learning solutions including those used by American Red Cross Blood Services, The Lonza Group, and Pfizer. One of the instructional designers ("BA") has developed e-learning courses and simulations used to train physicians and military medics. The other instructional designer ("SM") has also worked in healthcare training settings. The instructional design experts reviewed only the user interface design and the instructional design. All three experts are familiar with Good Manufacturing Practice (GMP) regulations and requirements that apply to the manufacturing and handling of medicines. Also, all experts primarily work with adult learners. One of the three experts has worked in a culture other than North America.

A protocol was created for the experts to use in reviewing the design documents. The first version developed gave relatively little specific guidance to the experts, intending them to look at the materials through the lens of their expertise. Testing the protocol revealed that having too little structure would not generate information that was meaningful and comparable. A second, more structured worksheet was created but it was judged to be too long from a practicality point of view. The third and final version had three sections covering:

- Overall visual design – the collection of visual elements such as drawings, photos, formats, arrangements, fonts, type sizes, colors, and symbols used in the learning program.
- Interface design – the methods, mechanisms, and "tools" used by the user to interact with the learning program and control movement through the program.
- Instructional design – the systematic approach using valid learning principles and learning theories, the desired outcomes, and the needs of the learners to create the specifications for the learning solution.

Each section of the protocol included a number of specific evaluation criteria; rating options; a space for the reviewer to list specific examples, comments, or suggestions; and the reference source of the theoretical basis for each of the criteria. Figure 2 shows a small portion of the protocol/data collection sheet used.



Element to review:	**3. Instructional design** Definition: the systematic approach using valid learning principles and learning theories, the desired outcomes, and the needs of the learners to create the specifications for the learning solution.	Documents/sources to consider in review:

Reviewer:	Review completion date:

Rating Definitions: SD – Strongly Disagree, D – Disagree, A – Agree, SA – Strongly Agree
NA – Not Applicable, NEA – No Evidence Available in documents provided.

3.1 Evaluation criteria / ratings

#	Criteria	Rating (see rating definitions)	Specific example/comment/suggestion	Source of criteria
3.1.a	The design of the course includes activities to help learners construct their own mental models related to cold chain and handling of time and temperature sensitive pharmaceutical products.	☐SD ☐D ☐A ☐SA ☐NA ☐NEA		Course goal
3.1.b	Lessons within the program stimulate integration of new knowledge with prior knowledge.	☐SD ☐D ☐A ☐SA ☐NA ☐NEA		<Mayer 805>
3.1.c	Lessons and examples provide a job or real-life context that promotes transfer of learning.	☐SD ☐D ☐A ☐SA ☐NA ☐NEA		<Mayer 831> <Mayer 2794> Vesper design principle – cognitive apprenticeship
3.1.d	Real-life examples are used to show learners how to perform a procedure or task.	☐SD ☐D ☐A ☐SA ☐NA ☐NEA		<Mayer 2746> Vesper design principle – authentic learning
3.1.e	Explanatory, not just corrective, feedback is given to correct and incorrect responses.	☐SD ☐D ☐A ☐SA		

Expert Formative Evaluation Data Collection Sheets
J. Vesper
7

Figure 2: Example of a section of the expert review protocol and data collection sheet.

The evaluation criteria – what each expert was to review and comment upon – came from a variety of sources such as the goals of the course, draft design principles (one of the key elements in a design research study (Herrington, et al., 2010)), and published/online resources (Clark & Mayer, 2008). Additionally, two open-ended questions concerning strengths and areas needing enhancement were asked at the end of the protocol. The graphics designer was asked to view the overall visual design and the interface design; the instructional designers reviewed the interface design and instructional design. Each expert received a package of materials that included:

- WHO confidentiality agreement;
- Course Information Material – An overview to the learning solution, description of the need, course goals, learning objectives, description of intended use, abbreviated audience analysis;
- Current drawings/sketches of the interface available at the time;
- Documentation concerning learner activities and assignments; and

- Evaluation protocol and data collection sheet.

Prior to beginning their review, the researcher phoned two of the experts to familiarize them with the project, the use of materials and data collection sheets; instructions to the third reviewer were provided by email. Each expert performed their review independently; the reviews took between 3-5 hours to perform and document.

Findings

The expert reviewers provided their responses using the data collection sheet. Where there were multiple reviewers for a section, the separate responses were compiled into an integrated document and analyzed. Themes from the reviewers were identified. A weighting was given based on the number of reviewers who made the comment; those themes commented on by multiple reviewers received a higher weighting. A report was developed for the WHO project leader and design team that summarized the evaluation findings and made recommendations. Appendices provided transcripts of all the comments from the expert reviewers.

After the findings in the written report were shared with the WHO project director, a teleconference was held to debrief the feedback and make decisions about moving forward with the prototype design. The findings were much too lengthy to include in this paper. Table 1 provides a synopsis of the findings along with a summary of the reactions of the design team to the findings and the decisions made based on the results of the expert review. Table 1 is not a verbatim transcript of the decisions, but is meant to convey the flavor of the interactions among the members of the design team in reaction to the expert review results.

Table 1: Synopsis of Expert Review Finding and Resultant Design Decisions

Aspect of design	Expert Feedback	Design Team Reactions and Decisions
Visual design	All three reviewers liked the look and "playfulness" of the visual design, but they questioned how this look aligned with the content of the program (cold chain management) and the learners (e.g., public health personnel in developing countries). RG commented that making changes to the eventual e-learning program would be time/resource consuming because of the nature of the illustrations. Cascading style sheets with simpler graphics were recommended; this would also speed the loading of the web pages for those with slower internet connections.	The "fun" and "playful" aspects of the design are intentional. This e-learning is meant to mimic the real world bus course that includes many "fun" and "playful" aspects. Cartoons and bright colors, hand writing, etc. were purposely chosen to support this goal. The other e-learning courses we will develop such as legislation for clinical trials and GCP inspections will not have such a "look". But on the other hand, we will always push the limits of creativity in any design in which we get involved as long as it does not detract from the learning goals.
Interface design	While visual design is more concerned with how elements like color, font, drawings, photos, and their arrangements contribute to the look and feel of the program, interface design is focused on the "tools" that the learner uses in moving through and around the e-learning program. All three expert reviewers felt that users would find the interface, as depicted on the sketches, difficult to use because the screens contain a great deal of information resulting in a high "cognitive load." They also said there was insufficient consistency in the sketches reviewed, and users faced different navigation systems in different sections of the program, also increasing cognitive load. As the cognitive load of the interface increases, the cognitive load that can be devoted to learning decreases. The reviewers noted that the prototype design lacks any clear indication of how far the learner has progressed through the material and/or how much farther they have to go.	The sketches shared with the experts were intended to depict a range of design options rather than fixed aspects of the eventual user interface. More user interface consistency will be built into the screens associated with different papers of the program. Once a working prototype of the program is ready, it will be subjected to rigorous usability testing including cognitive walk-through methods. The major goal of the usability testing will be to reduce the cognitive load of the graphical user interface on the learners. Several strategies will be added to the e-learning program to ensure that the learners know where they are, what they have accomplished, and how much remains to be completed.

Discussion

The project leader and design team carefully considered the expert review report with recommendations, and responded with many more detailed explanations regarding the concept/operational mechanics of the screen designs and interaction and changes that would be made than are represented in Table 1. A challenge in conducting the testing and refinement phase of educational design research is that any prototype that is evaluated essentially is a “snapshot” taken at one specific point in the overall development cycle. For this expert review, early design sketches and descriptions were used (they were the only such documents available), and as such, they did not completely represent the full functionality that was in mind of the developers. The sketches in some cases showed all the potential functionality of that screen, not specifically what the learner would see or experience during a transaction.

In any case, the purpose of formative evaluation activities during the testing and refinement phase is to enable people to make informed decisions, specifically in this case, concerning enhancements in the visual design, interface design, and instructional design of the e-learning solution. Based on the results of the expert evaluations, a number of recommendations were presented to the project leader and design team. These recommendations led to decisions that did not involve an extensive redesign, but rather, identified specific changes and additions that will improve the eventual program. Navigation and usability issues will be carefully studied in the next cycle of the testing and revision phase using a working prototype with real learners.

References

Barab, S. A., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *Journal of Learning Sciences*, *13*(1), 1-14.

Clark, R. C., & Mayer, R. E. (2008). *E-learning and the science of instruction: Proven guidelines for consumers and designers of multimedia learning* (2nd ed.). San Francisco: Pfieffer.

Dick, W., Carey, L., & Carey, J. O. (2008). *Systematic design of instruction* (7th ed.). Boston: Allyn & Bacon.

Dreyfus, H. L., & Dreyfus, S. E. (2005). Expertise in real world contexts. *Organizational Studies*, *26*(5), 779-792.

Ericsson, K. A., Prietula, M. J., & Cokely, E. T. (2007). The making of an expert. *Harvard Business Review*, *85*, 114-121.

Ewbank, D. C., & Gribble, J. N. (Eds.). (1993). *Effects of health programs on child mortality in Sub-Saharan Africa*. Washington, DC: National Academies Press.

Herrington, J., Reeves, T. C., & Oliver, R. (2010). *A guide to authentic e-learning*. New York: Routledge.

Milstien, J., Kartoğlu, Ü., & Zaffran, M. (2006). Temperature sensitivity of vaccines. *WHO Vol. WHO/IVB/06.10*. Geneva: World Health Organization.

Reeves, T. C. (2006). Design research from a technology perspective. In J. van den Akker, K. Gravemeijer, S. McKenney & N. Nienke (Eds.), *Educational design research*. London: Routledge.

Reeves, T. C., & Hedberg, J. C. (2003). *Interactive learning systems evaluation*. Englewood Cliffs, NJ: Educational Technology Publications.

Taylor, J. (2001). Recommendations on the control and monitoring of storage and transportation temperatures of medicinal products. *The Pharmaceutical Journal*, *267*(28), 128-131.

Tessmer, M. (1994). Formative evaluation alternatives. *Performance Improvement Quarterly*, *7*(1), 3-18.

Vesper, J. L., Kartoğlu, Ü., Bishara, R., & Reeves, T. C. (2010). A case study in experiential learning: Pharmaceutical cold chain management on wheels. *Journal of Continuing Education in the Health Professions*, *30*(4), 1-8.