Evaluation Practices within Instructional System Process

Nature and Process of Design

George Mason University

Ahmad Almufarreh

Dr. Brenda Bannan

Evaluation is the systematic determination of worth and significance of a learning process or a designed product by comparing set of criteria against a set of standards (Williams, 2011). Although the evaluation stage is often listed last in many models of instructional system design, it is in fact ongoing process throughout the entire design process. This is what makes instructional system models an active process rather than just liner process. Thus, evaluation practices are performed during the entire life cycle of the design in order to fix faults and enhance a product quality in both formal and informal ways. In order to understand and explore both evaluation methods during the design process, this qualitative study has been proposed to facilitate meeting this goal. The study expands the research of the current literature. It accomplishes that by addressing and attempting to understand the real practices of evaluation activities during and after the design process. In order to do so, this research examined the perspective of practicing designers as well as observed their projects. Second, the study provided future directions and implications for both training and research.

Literature Review

It has been established that evaluation practices play an essential role in instructional systems design and development in both theory and practice. However, to gather more information about evaluation practices within instructional system design, it was necessary to carry out a review of the literature using the databases of the George Mason University Library and Google Scholar, as well as my own personal library. It is commonly recognized that evaluation indicates the degree to which a designed product achieves the intended outcomes. Hence, the general model of instructional design ADDIE (analysis, design, develop, implement, and evaluate) is complete with evaluation serving as the final phase of a design process

(Molenda, 2003). However, it is possible to see the larger view of evaluation as essential throughout all phases of design and to consider the idea that development is not a separate phase during the design process. According to Hannum and Hansen (1989), evaluation is an ongoing activity that takes place throughout all phases of the ADDIE model and simply culminates at the last stage. Thus, one carries out evaluation during every instance of feedback that might inform the design process or indicate weaknesses that emerge along the way.

There are various reasons that have been noted for the failure of instructional design organizations to conduct systematic evaluations. Many instructional designers do not believe in the necessity of conducting evaluation (Swanson, 2005). Another reason for not performing evaluations is the lack of confidence with regards to the positive impact of a designed product or a training program on organization (Spitzer, 1999). The absence of expertise and the lack of an organization culture that believes in the importance of evaluation also contribute to the lack of evaluation in the instructional design process (Moller, Benscotor, & Rohrer-Murphy, 2000). However, there are some efforts that support conducting evaluation, though most of them are retrospective in nature (Brown & Gerhardt, 2002). A study on a group of instructional designers was conducted to determine the evaluation efforts of instructional designers as part of a systematic instructional design. This study shows that more than 89 % of participants only performed end course evaluation, while around 70 % evaluated users' learning. However, less than 45 % used acceptable techniques to measure users' achievement, and only 20 % of practitioners surveyed correctly recognized the techniques for results evaluation (Moller & Maillin, 1996). As Brown and Gerhardt (2006) indicated, companies typically expend less effort in evaluating the instructional design process.

Various studies over the past years provide some evidences that evaluation activities are often omitted in many cases or do not represent a key aspect of design work (Williams, South, Yanchar, Wilson, & Allen, 2011). In this empirical literature, successful design is associated with elements such as openness to new ideas and appropriate analysis. However, rarely is it associated with formalized evaluation activities such as formative or summative evaluations (Cox and Osguthorpe, 2003; Pieter & Bergman 1995).

Formative Evaluation and Summative Evaluation

In a broad view, evaluation activities within the instructional design process can be divided into two groups, formative evaluation and summative evaluation (Wang & Wilcox, 2006). Formative evaluation is a type of evaluation intended to offer information on improving design and development (Scriven, 1991). The purpose of this type of evaluation is to determine weakness in design materials, methods, or objectives with the intention of developing solutions throughout all design phases. The practices and research of evaluation in instructional design has been dominated by the convention of paying more attention to the outcomes of completed programs and design or to methods of assessing outcomes (Brown & Gerhardt, 2002). This tendency to focus on the outcomes and methods has mostly led to the neglect of activities relating to formative evaluation. As a result, Brown and Gerhardt (2002) suggested a guideline for formative evaluation practice and an integrative model was subsequently developed. In addition, they presented a case study that demonstrates the model in use. The purpose of this guideline is to serve as stimulus for instructional designers and evaluators to dedicate equal amounts of attention to both summative and formative evaluation activities. Thus, it is expected that engaging instructional designers in formative evaluation within the design process will improve the quality of design.

On the other hand, summative evaluation refers to an evaluation conducted to determine whether or not design outcomes are achieved (Scriven, 1991). Following this line of thought further, Wang and Wilcox (2006) presented a number of reasons leading instructional designers to conduct summative evaluation. First, so doing connects all ADDIE phases with organization objectives. More importantly, summative evaluation can help individuals identify areas of the completed interventions that do or do not meet the expectations of stakeholders. Last but not least, undertaking summative evaluation provides opportunities to make future improvements. Researchers have been attempting to understand summative evaluation for quite some time. Many studies have indicated that Kirkpatrick's four-level evaluation was the first taxonomy created specifically for summative evaluation (Alliger & Janak, 1989; Holton, 1996; Wang & Wang, 2005). The function of the four-level evaluation is to provide functional evaluation techniques or steps for evaluation practices. According to Wang and Wilcox (2006), the purpose of all classification systems in evaluation is to help understand and conceptualize the functions, nature or purposes of evaluated design from different perspectives. Therefore, the practical goal of utilizing evaluation practices is to create the best design under a given set of conditions.

Evaluation Challenges

In practice, design process activities seem to be conducted differently. Instructional designers believe that what is explained in textbooks and formal design processes often differs from what they actually do during design process (Cox & Osguthorpe, 2003; Christensen & Osguthorpe & Osguthorpe, 2004). Thus, there is still much to be identified in the areas of the instructional designers' activities regarding the different forms of evaluation (Allen et al., 2011). Prior studies of design practices, particularly those using qualitative methods, have begun to contribute valuable information on many aspects of instructional design to the literature. For

instance, Ertmer et al (2008) discussed the knowledge of expert and novice instructional designers in conjunction with decision-making. In addition, the studies of practical use of instructional design theories and the challenges instructional design practitioners face have shown light on the work of instructional design (Viaacher -Voreman & Gustafson, 2004; Liu et al, 2002; Yancher et al., 2010). Such studies have offered a clearer understanding and created learning experiences surrounding the actual work of instructional designers. This outcome is quite valuable, as the more is known about the actual work of design practices, the more applicable scholarships can be produced (Smith & Bloing, 2009).

While there are various articles that have reported many techniques on data assessment, Williams et al. (2011) stated that these studies lack the perspective of instructional designers as well as insight on practices regarding evaluation within the design process. Thus, there is a shortage of understanding as to the ways in which evaluation practices might differ from what is presented in textbooks, and one can see there is a need for more data to improve evaluation activities during the design process.

Research Questions

The goal of this research is to explore and understand the formal and informal evaluation activities during the design system process from the perspective of instructional designers. The following research questions have been proposed in order to facilitate meeting this goal:

1. How do instructional designers incorporate evaluation activities throughout their daily design practices?
2. What is the perspective of instructional designers regarding the use of evaluation activities within the design process?

These questions will expand the research of the current literature. First, the study will accomplish this goal by addressing and attempting to understand the real practices of evaluation activities during and after the design process. In order to do so, this research will examine the perspective of practicing designers as well as observe their projects. Second, the study should provide future directions and implications for both training and research.

The selection of these research questions is indicative of my own biases derived from my instructional design experience. Furthermore, I firmly believe that the more attention is given to evaluation activities, the more professional designs will be produced.

Methods

Research Participants

The participants in this study included current and prior students in an instructional system design Master program at an American university; they were all working as instructional designers. Participants were selected because of my bias, which is consistent with and congruent to research, that they practice evaluation activities different than what is presented in textbooks and theories. The decision to interview those participants was intentional and convenient as well. I believe that since the participants are either studying for or already graduated from the program, they have had valuable experiences with real evaluation activities.

I was introduced to participants through a colleague who has taken a number of instructional design courses with the participants. All participants who agreed to participate where invited by email. All meetings occurred in a Starbucks except in the case of one participant who preferred to be interviewed in her office. The meetings were cordial and all participants seemed eager to be a part of my research. I deemed having a colleague who has a

relationship with participants as an advantage because I was quickly able to establish a rapport with them.

The participants had varying years of experiences, which give me a broader perspective on the evaluation activities used by both novel and expert instructional designers. The first participant is Ethan who is a senior human resource development. He is in charge of high-level program designs and contracting as well as instructional design. He supervises a team of two other designers who work with him, basically design mandatory and other programs for government agency. The second participant is Sarah who is a training and development specialist. She has been in the training and development field for about 7 years now. The third participant is Joseph who has five years experience as instructional designer in a private organization. The fourth participant is Megan a chief learning officer at the economic development administration. She has 16 years of experience as instructional designer.

Data Collection

The methods of data collection for this study consist of interviews and observation. The form of the interview was an open qualitative style, and while there was a specified interview protocol, it was only used to provide a loose guide for the process. The protocol stipulated a few questions related to the research questions as well as follow-up questions to delve deeper into the participants' comments. At the end of the first interview, participants were asked to show a few samples of their current or previous works in order to observe both formal and informal evaluation activities that took place both during and after the design process. In addition, participants agreed to be interviewed for a second time in order to comment on the first interview

as well as to respond to any unresolved issues. Detailed information about the interviews is presented in the following tables:

First interviews		
Participant	Interview duration	Date
Ethan	54 minutes	11/08/2015
Sara	44 minutes	11/08/2015
Joseph	74 minutes	11/20/2015
Megan	63 minutes	11/28/2015

Second interviews		
Participant	Interview duration	Date
Ethan	18 minutes	12/06/2015
Sarah	14 minutes	12/06/2015
Megan	22 minutes	12/08/2015
Joseph	--	--

The interviews were recorded in order to facilitate data analysis. The interview conductor listened to the recordings and produced transcripts after the interviews occurred. In the process of transcribing, interviewer changed the names and other identification factors to protect the anonymity and confidentiality of the participants. All participants were assured of their confidentiality and informed about the recording of the conversation both orally at the start of the interview and in writing through the consent form that they signed. The participants were given a

copy of the consent form and assured that they had the right to withdraw from the study at any time.

Data analysis

The interview transcripts and observation notes were used to perform analysis on the data. The first step in data analysis was to examine it using open coding. I read all transcripts and observation notes for a sense of the whole experience and identified initial topics of relevance. According to Maxwell (2013), open inductive coding is used to ensure any new ideas were deducted. These first inductive codes were then inspected to reveal themes, and then reassembled in a matrix to analyze patterns across participants. Then illustrative quotes were selected form the transcripts to clarify the developed themes. Themes also were compared and contrasted to each other to look for connections among them. Moreover, themes were considered in the light of the whole thematic structure. This overview offers an accurate depiction of the procedure of data analysis for this study.

Trustworthiness and limitations

The major limitation of this study is the nature of its short time, needing to be completed in one academic semester. However, the validity issues were addressed using Creswell's (2013) checklist of eight validation strategies. The first strategy is engaging in long fieldwork. Although this was not s possibly due to the study limitation, this threat was somewhat mitigated by my instructional design background and having being introduced to participants by their classmates. According to Creswell (2013), one of the reasons for long engagement is to understand the culture and build relationship with participants. In addition, the second strategy is triangulation. To address triangulation I interviewed several participants with a diversity of working experience as the main sources, compared the interviews with observing participates professional works, and

following up interviews to check the results and for further clarifications. The critical friend review was instrumental in addressing the third strategy. Looing for discrepant information and negative cases is the forth strategy. When employing this strategy I was able to confirm or change my preconception. The fifth strategy is to clarify my bias as a researcher. During the study process, I examined my biases by writing reflective and identity memos. The sixth strategy is checking my data by participants. The following up interviews allowed me to check my results and paraphrased quotes with my participants. Recording the interviews, word for word transcriptions, and work observation notes assisted with the seventh strategy, which is obtaining rich and thick information. The final strategy is an external audit, which is not being met due the time limitation.

Results

In identifying emerging themes, I focused on material from the observation and transcript notes that appeared to have the potential to provide new insights on the evaluation practices for instructional designers. Some themes confirmed my preconception while others opposed them. I presented a few descriptive quotes with pseudonyms in order to clearly represent themes.

Technology use involves evaluation theme

The use of technology during the different stages of design required several evaluation decisions. Participants compared relative strengths of alternative technologies and attempted to identify which users' needs could be served by different technologies. As Sarah said *"the more tools you bring in, the more you have to evaluate"*. Thus, instructional designers believe that the more technology they can apply to meet learners' needs, the greater the number of evaluation decisions they need to make. Joseph clearly stated, "*I use Captivate for creating this tool and let's say I use Articulate. Let's say I use Photoshop and Illustrate and all these other things*.

Then you're looking at so many different components that you are having to evaluate whether or not Photoshop is better or is Captivate better". Therefore, even if instructional designers do not use a formal evaluation process to measure alternative technologies, their decision to use certain alternative technologies is intended to meet learners' goals and needs.

Evaluating students learning theme

Participants in the study often looked for methods to assess how learners are learning. Sarah, for instance, descried her evaluating role when she observed training sessions by saying, *"I'm always thinking to myself—are we evaluating them correctly, whether or not they're learning, and are they learning anything? Is this beneficial to them?"*. Moreover, Joseph succinctly pointed out, "*When the training is actually released, you certainly want to know how much they're learning and you've built in some evaluation ability, some capability to measure that such as tests, quizzes or other evaluation activities"*. In addition to the final evaluation after the training was released, Megan clearly said, "*In the design and development process, you're always going back to your learning objectives, trying to justify everything, trying to test whether or not you're really acting on those learning objectives"*. Therefore, instructional designers perform evaluation practices during the design phases by going back to the learning objectives. They also design products that create opportunities for leaners to be evaluated. These evaluation opportunities help them modify the instructions and assist learners in adjusting the activities of learning.

Designers evaluate informally theme

Instructional designers used various evaluation concepts such as feedback, options, analysis, testing, quality, review, and evaluating. Joseph, for instance, mentioned, *"The upfront analysis, audience analysis, objective statements, and the enabling objectives. All that stuff was*

absolutely critical … Every time we got to a decision point, we'd go back to that and say, what are we trying to do? How would this impact them?". Moreover, Megan stated that when she is signing a contract, all touch points are negotiated up front, including the review and feedback process. Following this train of thought, Ethan clearly illustrated that design documents must be presented to SMEs to get their review and approval. Additionally, when Sarah described her evaluation practices in her current project, she said, *"We did a lot of email back and forth. Hey, we have this idea, what do you think about it? I would bring like a printout when I would go to meet with them, just to meet with them. Get their quick feedback. It was basically like a touch point. Do you think we're on the right track?"*. Thus, participants considered the importance of evaluation practices as they sought to design high-quality products. However, they refer to their own use of evaluation even if they didn't usually evaluate their work officially.

Team member evaluation theme

Some participants clarified that evaluation of team members is required to develop a good instructional design. Megan for instance, stated, *"Although I work by myself on a project, I'm still within a team. And what I'm finding is that we all have strengths and we all have weaknesses. My weaknesses could be somebody else's strength"*. Thus, instructional designers often discuss their role or ideas with others in order to enhance the quality of the project, not just to identify their own weakness. On the other hand, while Ethan practiced team evaluation in many projects he conducted, he noted that instructional designers often have no time to carry out team evaluation. Furthermore, Joseph noted that team evaluation depends on the project. For instance, in most routine projects and mandatory training, I do not need to evaluate my team members or myself. Therefore, team evaluations allow instructional designers to share insights

and receive feedback. However, due to the project type or time constraints some instructional designers do not conduct team evaluation during the design process.

Evaluation for stakeholders theme

Conducting any evaluation requires clarification with regards to what is being evaluated or what is of greatest interest to the stakeholders (Williams and Grajam, 2010). Stakeholders have values and needs that shape the criteria for judging instructional design. Indeed, the following quote illustrates how stakeholders shape the evolution practices. Ethan reported, "*Whatever the stakeholders want, that's what we're measuring. It's my job to figure out what they really want*". In addition, Sarah noted, "*Many design decisions are made in a particular way because our customers requested that*". Although participants did not indicate or mention evaluation theories or literature, their evaluation practices and design decisions are based on stakeholder' values criteria and standards recommended by many evaluation theorists (Williams, 2011).

Discussion

My research questions were "How do instructional designers incorporate evaluation activities throughout their daily design practices?" and "What is the perspective of instructional designers regarding the use of evaluation activities within the design process?". In this section I will briefly discuss the answers to these questions along with several implications of the results for instructional designers. As I illustrated by the quotes in this paper, I noticed that instructional designers included evaluation practices in their professional work. They make practical and professional judgments in all stages of the design process. However, evaluation comprises different forms and words that separate instructional designers from the formal field of evaluation. Participants in the study illustrated many ways of using evaluation practices during

their design process. For instance, they always evaluated learners' needs and determined methods to meet those needs by carefully selecting design options. They evaluated the technologies they used, team members, stakeholders' needs, and learning theories to adjust the design development process. Thus, it might seem obvious that instructional designers perform many types of informal evaluations by making tactical decisions throughout all design stages.

In addition, participants considered many informal evaluation practices as ordinary or typical aspects of the profession. They rarely used or discussed the word "evaluation" unless they were prompted to talk about it. Participants stated that they typically spend time thinking about ways of selecting appropriate technologies, addressing learners' needs, and improving their designs. However, they view these tasks as ordinary aspects of designing instructions instead of considering them as formal evaluation practices.

Additionally, even when formal evaluation was not feasible, participants were willing to identify and negotiate with stakeholders, conduct needs analysis, make modifications based on feedback, and complete formal evaluation activities in order to produce high quality instructions. Thus, it is obvious that the more the instructional designers knew about formal evaluation practices, the more effectively they were able to integrate them in their professional work, though often in indirect ways.

Following this idea further, the participants pictured their evaluation works as internal. They evaluated themselves and their team members. They also evaluated customers' needs or accepted and incorporated the need assessments of supervisors. They evaluated which technologies, theories, and quizzes would be best to employ in their designs. However, participants believed that using a formal external evaluator in the design work required additional funding, which most clients are not likely to provide. Therefore, focusing on informal

developmental evaluation would strengthen the designers' works even if external evaluators were not included.

If they were to look at the results of the study, instructional designers would notice that colleagues perform informal evaluations in their own sittings while progressing through the various design stages. Therefore, instructional designers should share their knowledge and become explicitly part of design and implementation teams in order to expand their skills and benefit their understanding. In the same way, instructional designer researchers and theorists might benefit from understanding the actual practice of instructional designers and how evaluation theories are used in the professional realm. In fact, further research is needed to understand how other instructional designer use evaluation practices to enhance their works. Additionally, instructional designers could study some evaluation principles in order to integrate these principles into their design practices. Then they can reflect on ways of implementing these principles to improve their evaluation and design practices.

Conclusion

The experience of participants in this study does not reflect the experience of all instructional designers. However, the participants indicated important issues to be explored in the future. They believe evolution practices should be central to the field of instructional system design; however, they think a good evaluation must be conducted formally through a few experiments during the final stage of the design process, though they rarely perform these evolution studies. Additionally, they perform many informal evaluation practices to guide critical design decisions throughout the design process without officially identifying such practices as evaluation.

While this study provides a general observation regarding evaluation practices within the design process and how instructional designers view them, further studies are needed to offer insights into specific evaluation forms and tasks, including the provision of practical methods to measure the learning outcomes during the design process. Additional studies will also allow the field to better understand and evaluate values or the standards of different stakeholders. I hope this study will enhance instructional designers' knowledge regarding the evaluation practices during the design process.

References

Alliger, G. M., & Janak, E. A. (1989). Kirkpatrick's levels of training criteria: Thirty years later. *Personnel psychology*, *42*(2), 331-342.

Brown, K. G., & Gerhardt, M. W. (2002). Formative evaluation: An integrative practice model and case study. *Personnel Psychology*, *55*(4), 951-983.

Christensen, T. K., & Osguthorpe, R. T. (2004). How Do Instructional-Design Practitioners Make Instructional-Strategy Decisions?. *Performance Improvement Quarterly*, *17*(3), 45-65.

Cox, S., & Osguthorpe, R. T. (2003). How do instructional design professionals spend their time?. *TechTrends*, *47*(3), 45-47.

Ertmer, P. A., Stepich, D. A., York, C. S., Stickman, A., Wu, X. L., Zurek, S., & Goktas, Y. (2008). How instructional design experts use knowledge and experience to solve ill-structured problems. *Performance Improvement Quarterly*, *21*(1), 17-42.

Gibby, S., Quiros, O., Demps, E., & Liu, M. (2002). Challenges of being an instructional designer for new media development: A view from the practitioners. *Journal of Educational Multimedia and Hypermedia*, *11*(3), 195-219.

Hannum, W. H., & Hansen, C. (1989). *Instructional systems development in large organizations*. Educational Technology.

Holton, E. F. (1996). The flawed four-level evaluation model. *Human resource development quarterly*, *7*(1), 5-21.

Molenda, M. (2003). In search of the elusive ADDIE model. *Performance improvement*, *42*(5), 34-37.

Moller, L., Benscoter, B., & Rohrer-Murphy, L. (2000). Utilizing performance technology to improve evaluative practices of instructional designers. *Performance Improvement Quarterly*, *13*(1), 84-95.

Moller, L., & Mallin, P. (1996). Evaluation practices of instructional designers and organizational supports and barriers. *Performance Improvement Quarterly*, *9*(4), 82-92.

Pieters, J. M., & Bergman, R. (1995). The empirical basis of designing instruction. *Performance Improvement Quarterly*, *8*(3), 118-129.

Scriven, M. (1996). Types of evaluation and types of evaluator. *American Journal of Evaluation*, *17*(2), 151-161.

Smith, K. M., & Boling, E. (2009). What Do We Make of Design? Design as a Concept in Educational Technology. *Educational Technology*, *49*(4), 3-17.

Spitzer, D. R. (1999). Embracing Evaluation. *Training*, *36*(6), 42-47.

Swanson, R. A. (2005). Evaluation, a state of mind. *Advances in Developing Human Resources*, *7*(1), 16-21.

Visscher-Voerman, I., & Gustafson, K. L. (2004). Paradigms in the theory and practice of education and training design. *Educational Technology Research and Development*, *52*(2), 69-89.

Wang, G. G., & Wang, J. (2005). Human resource development evaluation: Emerging market, barriers, and theory building. *Advances in Developing Human Resources*, *7*(1), 22-36.

Wang, G. G., & Wilcox, D. (2006). Training evaluation: knowing more than is practiced. *Advances in Developing Human Resources*, *8*(4), 528-539.

Williams, D. D., South, J. B., Yanchar, S. C., Wilson, B. G., & Allen, S. (2011). How do instructional designers evaluate? A qualitative study of evaluation in practice. *Educational Technology Research and Development*, *59*(6), 885-907.

Yanchar, S. C., South, J. B., Williams, D. D., Allen, S., & Wilson, B. G. (2010). Struggling with theory? A qualitative investigation of conceptual tool use in instructional design. *Educational Technology Research and Development*, *58*(1), 39-60.