Expertise in Cognitive Task Analysis Interviews

Danny S. M. Koh (dannykoh@alumni.cmu.edu)
Human-Computer Interaction Institute, Carnegie Mellon University
5000 Forbes Avenue, Pittsburgh, PA 15213-3890, USA

Kenneth R. Koedinger (koedinger@cmu.edu)
Human-Computer Interaction Institute, Carnegie Mellon University
5000 Forbes Avenue, Pittsburgh, PA 15213-3890, USA

Carolyn P. Rosé (cprose@cs.cmu.edu)
Human-Computer Interaction Institute, Carnegie Mellon University
5000 Forbes Avenue, Pittsburgh, PA 15213-3890, USA

David F. Feldon (david.feldon@usu.edu)
Science, Technology, Engineering, Education, Mathematics (STE2M) Center, Utah State University
6720 Old Main Hill, Logan, UT 84322-6720, USA

Abstract

Cognitive Task Analysis (CTA) interview technique is commonly used to elicit knowledge of subject-matter experts and to design instruction better focused on what experts don't know they know. However, the knowledge of how to conduct an effective interview is, itself, largely implicit. In this study we performed protocol analysis on a set of interview transcripts from an expert CTA practitioner to elicit the cognitive processes of conducting CTA interviews. We also consulted expert CTA practitioners to identify the strategies that they used during the interviews. We present key strategies that were employed by the expert CTA practitioners to ensure comprehensiveness and accuracy in the information collected, such as looking for perceptual cues (e.g. considering verbs such as "determine") to ascertain adequacy of SME's responses and selection of follow-up questions. We present a production rule model as a detailed description of the cognitive processes underlying expert CTA interviewing.

Keywords: cognitive task analysis; knowledge elicitation; protocol analysis; production rule model

Introduction

Subject Matter Experts (SMEs) in various professional fields are often consulted to inform the curriculum development for training purposes. However as the SMEs acquire expertise in a specific skill, less conscious effort is required in using the skill during problem solving. Therefore the SMEs might not be able to articulate the procedures completely and accurately through self-report (Feldon, 2010). The errors and omission of important information on the strategies in performing the tasks effectively would potentially result in poor curriculum development for the trainees (Feldon, 2007).

To circumvent the abovementioned problem, Cognitive Task Analysis (CTA) is increasingly used to ensure comprehensiveness and accuracy during the knowledge elicitation process (Clark, Feldon, van Merrienboer, Yates, & Early, 2008). CTA is a term that describes a list of techniques for eliciting knowledge, cognitive processes and goal structures from SMEs who are consistently successful when performing a specific task. Among all elicitation techniques, methods on expert interviews are most commonly used especially in complex domains (Yates & Feldon, 2011). Examples of real-world contexts using CTA interviews to develop training materials or decision aiding tools are firefighting (Klein, Calderwood & Macgregor, 1989), medicine and health (Clark, 2014), and academic fields (Feldon, Timmerman, Stowe & Showman, 2010).

While the findings from CTA interviews are beneficial in informing both instructional and system design, results may vary across CTA practitioners due to difference in expertise level and perspectives. Achieving mastery in conducting interviews requires considerable practices. While there are guides for conducting specific interview methods (Clark, 2014; Crandall, Klein & Hoffman, 2006), these guides are usually described at a broad level or in a step-by-step sequential order, which might be insufficient given the complex nature of the interviews. For instance, how does the CTA practitioner identify a decision point during the interview? When should the CTA practitioner continue or stop probing deeper into the decision point? What are the follow-up questions that the CTA practitioner should perform to ensure an effective interview? Failure to employ the effective CTA strategies would impede a comprehensive collection of information during the interview. In view of the difficulties in conducting CTA interviews, there is a need to better understand the cognitive processes supporting the performances of these interview methods.

In this paper we describe the study that we conducted to elicit the cognitive processes applied when performing CTA interviews and the implications of the study. We will also present a production rule model for conducting CTA interviews, which may inform instructional design for teaching CTA interview technique, or to develop adaptive tools to facilitate CTA interviews.

Related Works

Among the numerous interview methods developed for CTA, the Critical Decision Method (CDM) and the Precursor, Action, Result, and Interpretation (PARI) method are most commonly adopted by CTA practitioners (Tofel-Grehl & Feldon, 2013). CDM (Klein et. al., 1989) is a type of incident-based interview where the SME tries to recall a specific event that is highly challenging and unusual. The PARI method (Hall, Gott, & Pokorny, 1995) on the other hand focuses on the challenges that occur in routine events by seeking an understanding of the four elements as described by the name of the method. Another method worth mentioning is the Concepts, Processes and Procedures (CPP) method (Clark, 2014) which is a variation of the PARI method. The information collected from the CPP method can easily be mapped to the requirements of the instructional design recommended by the Guided Experiential Learning (GEL) design (Clark, 2014). This framework facilitates the transition from knowledge elicitation to the implementation of instructional design.

Crandall et. al. (2006) provided a detailed guide for conducting CDM. There are four phases in the interview: the CTA practitioner starts by identifying a challenging incident where the SME's expertise has played a significant role in the outcome, followed by expanding the account of the story and creating a timeline of the selected incident. The next phase is to probe into critical decision points to identify the perceptual cues and alternative options when making a decision. Lastly the CTA practitioner will pose hypothetical "what-if" questions to determine how the decisions might change with varying conditions or situations. The guide also provides numerous strategies with examples that illustrate how to conduct an effective interview. For instance, the guide suggests that when the SMEs mention words such as ''We just knew …'' or ''My gut told me that …'', the content implies that that the SMEs might have omitted some parts of the procedure and therefore calls for the need to probe in deeper.

Instruction for performing CPP is also well articulated (Clark, 2007, 2014). At the start of an interview, the SME is required to outline the performance sequence of the key sub-tasks essential in performing a task. The SME will also be asked to come up with a few scenarios that a person would be able to solve if he has mastered the relevant skills and knowledge. The task sequence and scenarios will then be refined iteratively through interviews with other SMEs. Once the task sequence and scenarios have been finalized, the CTA practitioner will start delving into the subtasks to elicit the information that informs the decision process, as well as the essential declarative and procedural knowledge. Some of the types of information collected include the actions and decision steps for performing the subtasks, the cues and conditions that lead to the actions or decisions, tools and the required performance standards. The guide also illustrates the variety of questions that can be used to elicit the same type of information from the SMEs.

There are also guides for eliciting implicit knowledge that are method-independent. These guides provide strategies that can generally aid in creating an effective interview. For instance, Wilson, Holloway & Miller (2008) discussed strategies for building rapport with SMEs so that they feel more comfortable in sharing their own experiences, as well as tactics to ensure that information collected is not biased by the perception of the CTA practitioner. While these guides focus at the macro level, it is still useful for the novice CTA practitioners to appreciate how to conduct an effective CTA interview in general.

While the current guides for the interview methods are extremely beneficial to the CTA practitioners, the step-by-step presentation of the procedure can be somewhat misleading because interviews are often highly dynamic and might not follow any sequential flow at all. While the strategies and examples mentioned in these guides provide insights of when and what actions to take, how these strategies come about are not stated explicitly. If these strategies are largely based on the reflections of the expert CTA practitioners, there is a possibility that more strategies have not been tapped given the automaticity of expertise. For instance, what have not been adequately addressed in these guides are the conditions of when to continue probing deeper and when to move on with other aspects of the interview.

The need for an empirically driven approach to identify the knowledge and cognitive processes of performing an effective CTA interview serves as the key impetus for this research. Specifically, by conducting a CTA of the CTA interviews, we would be able to use the findings to better educate the novice CTA practitioners on how to detect and respond to various situations during the interviews.

Methods

The first phase of the research was to find out how the CTA interviews should be conducted based on a literature review. Although the elicitation approaches for different interview methods are largely similar (Yates & Feldon, 2011), we focused on just CPP method for consistency of this research. CPP method was chosen because it has a clearer mapping of the elements collected from the interview to the information required for a GEL training design. The alignment of the elicited knowledge and the parameters for instructional design would allow potential follow-up of the findings from this research to instructional development for teaching purposes. The findings were developed into a set of coding scheme that identified the list of actions that a CTA practitioner should perform during the interview. 31 action categories were identified and classified into the following five phases of the interview: Introduction, Task Overview, Detailed Subtask Sequence, Deepening Decision Points and Generic Strategies for Effective Interview. The coding scheme was then used for protocol analysis during the second phase of the study.

The second phase of the study was to determine how the interviews were actually performed by the expert CTA

practitioner through protocol analysis. Protocol analysis is another knowledge elicitation technique by reviewing the verbal reports or transcripts to elicit the thought sequences (Cooke, 1999). Transcripts of past interviews were used for the protocol analysis, which provided an accurate account of what actually happened during the interviews between the expert CTA practitioner and the SME of the task of interest. The transcripts also provided insights of what might possibly be the triggers that an expert CTA practitioner seeks in deciding which action(s) to perform during the interviews. Audio recordings of eight interview sessions conducted by an expert CTA practitioner were used for the protocol analysis. These interviews were conducted with six SMEs to understand how they conduct scientific inquiry in the field of biological sciences and the results were developed into instruction for a laboratory-based undergraduate biology course (Feldon et. al., 2010). The audio recordings were transcribed on a turn-by-turn basis, where all the comments made by one person (expert CTA practitioner or SME) at an instance were constituted as one response. The expert CTA practitioner made 473 responses throughout the interviews.

All responses made by the expert CTA practitioner were coded using the coding scheme developed in the first phase of the study. As it was common for the CTA practitioner to make multiple comments or ask several questions all at a time, each response could be coded with multiple action categories. Two researchers coded the same set of transcript independently and computed the inter-rater agreement for both sets of coding. Given the huge list of categories (31 action categories) and the classification of each expert CTA practitioner's response was non-exclusive, we decided to compute the inter-coder agreement only for categories that occurred frequently in the transcript ($N > 10$) instead. The mean kappa coefficient was established at 0.55. While the inter-rater agreement was generally moderate (Landis & Koch, 1977), the researchers continued to resolve any discrepancies in the coding through discussion and the finalized coding scheme was used to code the rest of the transcripts.

Concurrent with protocol analysis, a series of informal interviews were conducted with three expert CTA practitioners as well. The expert CTA practitioners were familiar with the CPP method with years of experience. The objectives of the informal interviews were to validate the findings from the protocol analysis and also to explore additional insights that might not be explicit from the transcripts or literature.

Results and Discussion

The following section describes some of the key strategies used by the CTA practitioners that facilitate the knowledge elicitation during the interview process.

Focus on task procedures and decision criteria, with eliciting options being implicit

Figure 1 shows the total number of instances that the action categories in each phase of the interview occurred in the transcripts. The actions used by the CTA practitioner during the interviews were mainly to elicit detailed step-by-step description of the actions and decision steps the SMEs performed to complete the subtasks (Detailed subtask sequence: 70% of the actions, $N = 473$, $n = 329$), and also to probe in deeper into the decision points that the SME made (Decision Points: 31% of the actions, $N = 473$, $n = 148$). The result is congruent with the literature findings, where the phases "Detailed Subtask Sequence" and "Deepening Decision Points" are continuously repeated throughout the interviews until the procedure of all subtasks have been captured. The continuous process of seeking the subtask sequences and deepening into the decision points results in the higher occurrences of the related action categories as compared to other phases of the interview.

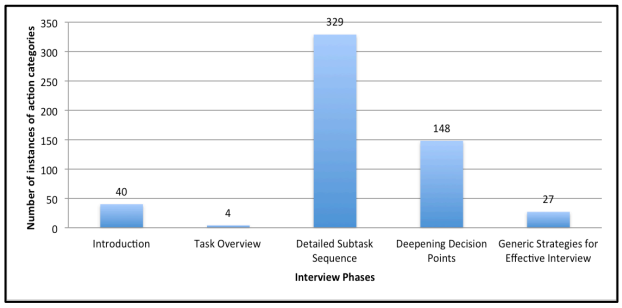

Figure 1: Number of instances of all the action categories in each interview phase. The actions executed by CTA practitioner were mainly to capture a detailed sequence of the subtasks and delve into the decision points.

We delved in deeper to look at the count of each action category for the two most frequent interview phases. Figure 2 shows the count of the ten action categories that the expert CTA practitioner would take in order to capture the step-by-step procedure of performing the subtasks. The result shows that the expert CTA practitioner asked the SMEs to list the steps required in performing each subtask ($n = 78$) and verify the information provided ($n = 171$) much more often than the other action categories. Figure 3 shows the count of the six action categories that the expert CTA practitioner would take in order to identify the options available to SMEs when making a decision and also the criteria for choosing between the options. The result shows that seeking criteria for choosing between alternatives ($n = 68$) is more often performed as compared to other categories.

Given that SMEs often have difficulty articulating the procedural knowledge of performing a certain task (Koedinger & Anderson, 1990), it is logical that the CTA practitioner would spend a substantial amount of time seeking a detailed step-by-step procedure of performing specific subtasks and eliciting the criteria when making a

decision. What's interesting is that despite the frequent action of trying to elicit the criteria in choosing between options, the action of identifying the various options is surprisingly infrequent. We expect that in order to identify the criteria for choosing between the options, the CTA practitioner has to firstly seek the list of options available. The action of seeking the list of alternatives is however infrequently performed (n = 19), suggesting that the options might have already been provided even without having to ask the SMEs explicitly. Another takeaway is that CTA interview is usually dynamic and opportunistic, and the actions performed by the CTA practitioner do not always follow a linear progression. The result also suggests that there should be other perceptual cues that the CTA practitioner might be looking for to determine whether the SMEs have provided alternatives and other options in performing the task.

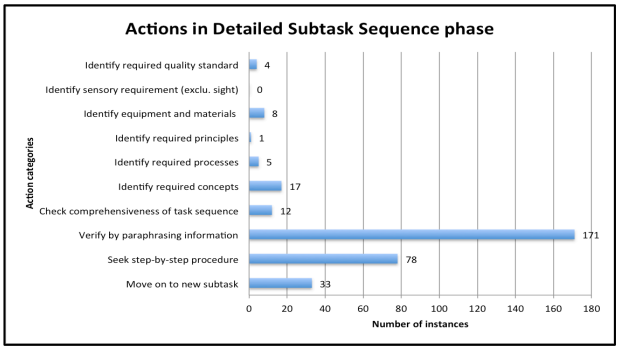

Figure 2: Count of each action category related to gathering a detailed subtask sequence. Seeking a step-by-step procedure of the subtask and verifying the information are the most common actions performed by the CTA practitioner.

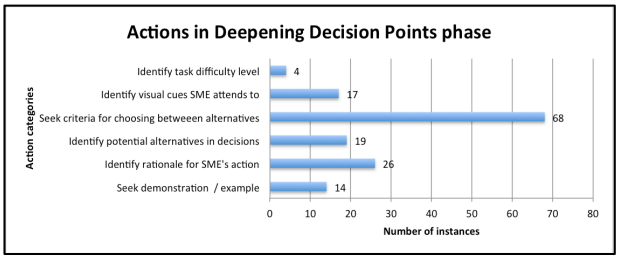

Figure 3: Count of each action category related to deepening decision points. Seeking the criteria for choosing between alternatives is the most frequent action performed by the CTA practitioner.

Importance of verification and accuracy

Another interesting finding is that the most frequently used action performed by the CTA practitioner is to paraphrase the information provided by the SME for verification during the interview. While the existing CTA guides (Clark, 2007, 2014) have also suggested the need to verify the information provided by SME, this action is however not explicitly emphasized and is only required to perform a few times throughout the interview. The finding from the protocol analysis shows that verification was done almost immediately after each SME's response. Another method of verification is by challenging the SMEs' assertions. After the SMEs described the task sequence (i.e. "First I do this and then I do that …"), the CTA practitioner would pose an alternative situation that seemed practical but not supported by the assertion (i.e. "So meaning you would just do X and not Y, is that correct?"). The challenge allows the SMEs to reconstruct the task sequence and at the same time identify previously omitted decision point. Although the action of frequent verification could be attributed to the CTA practitioner's personal style of conducting the interview, it seems to be a useful strategy in ensuring the accuracy and completeness of the information collected. By verifying the information right after every SME's response, the SMEs can also correct any misconception that the CTA practitioner might have. This is especially important as the CTA practitioner is usually not an expert in the task of interest (to avoid bias) and requires SMEs to guide him in understanding the process of performing the task of interest. We also observed that by paraphrasing the information provided by the SME, the CTA practitioner was able to monitor his evolving understanding by integrating presented knowledge with his prior knowledge and seek explanation from the SME when a knowledge gap was spotted.

Perceptual cues in determining adequacy of SME's responses and follow-up actions

During the discussion with the expert CTA practitioners, they shared one useful strategy in identifying a decision point, which was to look for action verbs in SMEs' responses. In order to look for potential cues that might indicate the presence of decision points, we grouped the CTA practitioner's responses according to the coded action categories and reviewed the SMEs' responses prior to each CTA practitioner's action. The findings suggest that there were certain keywords and phrases that indicated decision points in SMEs' responses. Particularly, the occurrence of considering verbs such as "identify", "determine" or "decide" usually led the CTA practitioner to ask about alternatives, or delving into the criteria if the alternatives had already been provided by the SMEs. There were also some phrases that are useful in determining whether certain questions have been answered adequately by SMEs. For instance, phrases such as "First … Second … Third …", "There are X ways to do so …" tended to appear in SME's responses when the CTA practitioner asked for options or alternatives in performing a task. Similarly, keywords such as "If … then …" occurred very frequently when CTA practitioner asked for the criteria in choosing between the options. Absence of these words usually implied inadequacy in SMEs' responses to CTA practitioner's questions. These perceptual cues might be a useful strategy especially for novice CTA practitioners to determine when to continue probing or move on to the next task.

Comprehensiveness of information collected

Although the action of identifying the various options when making a decision was less often used by the CTA practitioner, we noticed that there were several instances where the CTA practitioner continued to probe for more options (e.g. "Are there anymore methods that you can think of apart from those that you have mentioned?") even though SMEs had already provided some options. The rationale for the CTA practitioner to repeatedly ask for options is that the SMEs only examine the best options and unconsciously omit others that are less effective for the situation (Klein, Wolf, Militello & Zsambok, 1995). While omitting less effective options is important in performing the task, sometimes the CTA practitioner would want the SMEs to list all potential options during the interview. The presence of the options would then facilitate in eliciting the criteria that distinguish the various options. While the effectiveness of repeatedly seeking for options requires further research, this strategy does ensure comprehensiveness in the options elicited from the SMEs.

Precisely Specifying Expert CTA Interview Processes in a Production Rule Model

Based on the findings from the literature, protocol analysis and expert interviews, a production rule model was developed to precisely specify cognitive processes inherent in expert CTA interviews. The production rule model is made up of a series of actions that characterize alternative interview moves and a set of "IF-THEN" statements that characterize when an expert decides to evoke an interview move. Table 1 shows a subset of the production rules model that focuses on eliciting the subtask step-by-step procedure and seeking decision points. Interview guides identified from literature survey generate about 50% of the production rules, whereas the findings from the protocol analysis and expert interviews are represented by 88% of the production rules. The greater contribution from the protocol analysis and expert interviews suggests that this study may have been effective in identifying more insights and strategies on how to conduct an effective CTA interview.

Implications

One of the limitations of this study is that the protocol analysis was only based on one expert CTA practitioner. The production rule model would certainly be more representative of the general CTA interviews if the protocol analysis was conducted on at least three expert CTA practitioners. This limitation was mitigated by the additional expert interviews and some of the strategies were in fact observed in both protocol analysis and expert interviews. Another limitation of this study is the insufficiency of data for other action categories, which impedes the identification of more strategies and potential triggers for all action categories.

Nevertheless, the production rule model is still useful in translating into a more effective instructional design for teaching CTA. For example, given that the CTA practitioner's actions are dependent on the SME's responses, novice CTA practitioners should practice using the transcripts of expert CTA practitioner's past interviews. A snippet of the SME's responses could be given to the novice CTA practitioners who would then be asked to select the follow-up questions that they find most appropriate. Selecting follow-up questions based on SME's responses help to train the ability to identify the perceptual cues required for evoking specific questions. Novice CTA practitioners should also be exposed to challenging situations such as SME having difficulty in articulating their

Table1: Subset of the production rule model[1] in eliciting subtask step-by-step procedure and deepening into decision points. The production rule model is made up of a series of actions and a set of "IF-THEN" statements.

ELICIT SUB TASK STEP-BY-STEP PROCEDURE
• Seek for a step-by-step procedure of the subtask. Explore the subtasks one at a time. The current focus is to capture the procedure of the subtasks as comprehensively as possible. It is not necessary to cover all subtasks within a single interview. • IF SME's response suggests having to make a decision (use of considering verbs such as "think", "decide", "choose"), THEN ask for the list of options or alternatives. o IF SME has provided an elaborated account of at least one action/decision that he has to make, THEN clarify with SME again to make sure all alternatives have been discussed. o IF SME responds with the following keywords such as "First … Second ... Third ..." or "There are X ways …", THEN it suggests that SME has provided a list of options. Continue to probe for alternatives to ensure that the list is exhaustive. o IF SME has provided a comprehensive list of options and alternatives, THEN begin "Seek decision criteria".
SEEK DECISION CRITERIA
• Seek for criteria for choosing between alternatives. o IF the criteria given by the SME are too abstract (usually the absence of "if.." or "if … then …"), THEN rephrase the question on the selection criteria and ask SME to elaborate on the criteria. o IF the SME is having difficulty coming up with the criteria (long pause or "that's a hard question"), THEN rephrase the question on the selection criteria. o IF the SME responds with the following keywords such as "If … then …" or "whether ..." and the criteria are clear and observable, THEN it suggests that SME has provided the criteria for choosing between the options.

[1] Contact the authors for the production rule model.

knowledge. Exposure to challenging situations allows novice CTA practitioners to know how to react to similar situations in future. Apart from informing the instructional design, the production rule model could also be used to develop aiding tool for novice CTA practitioner during actual interview sessions. Such aiding tool recommends appropriate follow-up questions to the CTA practitioners throughout the interview so that the information collected would be comprehensive and accurate.

Conclusion

This research aims to identify the cognitive processes and effective strategies used in performing the CTA interview. The findings suggest that CTA interviews are highly dynamic and the types of questions asked need not necessarily follow a sequential order. Based on the current guides for conducting CTA interviews, results from protocol analysis and insights from expert CTA practitioners, a production rule model was developed. These production rules, with there detailed if-parts, provide a precise indication for critical dependencies between interview actions and interviewee responses. While a strict sequential order is not necessary, it is important, for example, that a CTA practitioner would need to elicit a comprehensive set of decision options before asking the SME to come up with the criteria that distinguish which option/s to choose. More generally, the model summarizes some of the key insights from our CTA of CTA interviewing. One particular insight was the use of perceptual cues to determine the adequacy of SME's responses and the selection of follow-up questions during the interview. For instance, keywords such as "if ... then ..." are used to determine whether the SME has adequately provided the criteria for choosing between alternatives during decision-making. Absence of these keywords usually implies that the CTA practitioner has to repeat the question again to the SME.

Acknowledgements

The authors thank Dr Richard E. Clark and Dr Kenneth A. Yates for sharing their invaluable insights on CTA. We also thank DSO National Laboratories (Singapore) for the support to pursue this research.

References

Clark, R. E. (2007). *The Use of Cognitive Task Analysis and Simulators for after Action Review of Medical Events in Iraq* (Tech. Rep. W81XWH-04-C-0093). Fort Detrick, MD: U.S. Army Medical Research and Materiel Command.

Clark, R. E. (2014). Cognitive task analysis for expert-based instruction in healthcare. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of research on educational communications and technology*. New York, NY: Springer.

Clark, R. E., Feldon, D. F., van Merrienboer, J. J. G., Yates, K. A., & Early, S. (2008). Cognitive task analysis. In J. M. Spector, M. D. Merrill, J. J. G. van Merriënboer, & M. P. Driscoll (Eds.). *Handbook of research on educational communications and technology* (3rd ed.). New York, NY: Macmillan/Gale.

Cooke, N. J. (1999). Knowledge elicitation. *Handbook of applied cognition*, New York, NY: John Wiley.

Crandall, B., Klein, G. A., & Hoffman, R. R. (2006). *Working minds: A practitioner's guide to cognitive task analysis*. Cambridge, MA: MIT Press.

Feldon, D. F. (2007). The implications of research on expertise for curriculum and pedagogy. *Educational Psychology Review*, *19*(2), 91-110.

Feldon, D. F. (2010). Do psychology researchers tell it like it is? A microgenetic analysis of research strategies and self-report accuracy. *Instructional Science*, 38(4), 395-415.

Feldon, D. F., Timmerman, B. C., Stowe, K. A., & Showman, R. (2010). Translating expertise into effective instruction: The impacts of cognitive task analysis (CTA) on lab report quality and student retention in the biological sciences. *Journal of research in science teaching*, *47*(10), 1165-1185.

Hall, E. P., Gott, S. P., & Pokorny, R. A. (1995). *A procedural guide to cognitive task analysis: The PARI methodology* (Rep. No. AL/HR-TR-1995-0108). TX: Brooks Air Force Base, U.S. Air Force Armstrong Laboratory.

Klein, G. A., Calderwood, R., & Macgregor, D. (1989). Critical decision method of eliciting knowledge. *IEEE Transactions on Systems, Man, and Cybernetics*, 19, 462-472.

Klein, G., Wolf, S., Militello, L., & Zsambok, C. (1995). Characteristics of skilled option generation in chess. *Organizational Behavior and Human Decision Processes*, 62, 63-69.

Koedinger, K.R., & Anderson, J.R. (1990). Abstract planning and perceptual chunks: Elements of expertise in geometry. *Cognitive Science*, 14, 511–550.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 159-174.

Tofel-Grehl, C., & Feldon, D. F. (2013). Cognitive Task Analysis–Based Training A Meta-Analysis of Studies. *Journal of Cognitive Engineering and Decision Making*, *7*(3), 293-304.

Wilson, L. T., Holloway, P., & Miller, B. (2008). Mini Guide to Eliciting Implicit Knowledge. Retrieved from http://www.knowledgeharvesting.com/documents/Mini%20Guide%20to%20Eliciting%20Implicit%20Knowledge.pdf

Yates, K. A., & Feldon, D. F. (2011). Advancing the practice of cognitive task analysis: a call for taxonomic research. *Theoretical Issues in Ergonomics Science*, *12*(6), 472-495.