EFFECT OF PROMPTING TECHNIQUES AND LEARNING STYLES ON REQUIREMENTS ELICITATION

A Thesis
Submitted to the Graduate Faculty
of the
North Dakota State University
of Agriculture and Applied Science

By

Bonnie Lynn Jan

In Partial Fulfillment of the Requirements
for the Degree of
MASTER OF SCIENCE

Major Program:
Software Engineering

July 2016

Fargo, North Dakota

North Dakota State University
Graduate School

Title

Effect of Prompting Techniques and Learning Styles on
Requirements Elicitation

By

Bonnie Lynn Jan

The Supervisory Committee certifies that this ***disquisition*** complies with North Dakota State University's regulations and meets the accepted standards for the degree of

MASTER OF SCIENCE

SUPERVISORY COMMITTEE:

Dr. Gursimran Walia

Chair

Dr. Kendall Nygard

Dr. Limin Zhang

Approved:

July 1, 2016 Dr. Brian Slator

Date Department Chair

ABSTRACT

Research efforts on improving requirements elicitation are focused on developing and validating better techniques for eliciting a comprehensive set of requirements. However, there isn't enough empirical evidence available to requirements analysts on selecting elicitation techniques that are the most appropriate for all technical and non-technical stakeholders of a project. This thesis focuses on the use of prompting techniques, which employ techniques found to have greater effectiveness and have been improved for overcoming cognitive limitations. Motivated by previous research, an empirical investigation was conducted on the effectiveness of *syntactic* vs. *task characteristics* prompting techniques and the impact *Learning Styles* (a cognitive psychology aspect) has on their usability. The results show greater effectiveness from the task characteristics technique and the detail- or holistic-oriented people using the techniques. These results can be used to identify prompting techniques improvements for comprehensive requirements elicitation from stakeholders of varying technical backgrounds.

ACKNOWLEDGEMENTS

I would like to thank everyone who participated in the research. I thank Dr. Gursimran Walia for his guidance and helping design the experiment, Dr. Mark Pavicic for contributing his time and assisting with participation, Anurag Goswami for his help in analyzing Learning Styles vs. requirement analysis, Dr. Brian Slator and Dr. Kendall Nygard for offering to answer questions in regards to research, and Annette Sprague and Brittney Knowlton for guidance with paperwork and graduation requirements. Thanks also to Dr. Nygard and Dr. Limin Zhang for being on my graduate committee.

DEDICATION

I dedicate this thesis to my family, particularly my parents Dr. Chao-Chien Jan and Dr. Grace Wei for helping me have an environment conducive for concentrating on my education and my siblings Joanna and Brian. Thank you to Steven La, who has supported and been a grounding force for me. I also dedicate this achievement to my friends – Dr. Obioma Ohia and Deepak Dawar – who have encouraged me throughout the process.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACRE..Acquisitions of Requirements

ACT..Active

FSLSM...Felder and Silverman's Learning Style Model

GLO ....Global

GLR..Goal Level Requirements

GUI ....Graphical User Interface

ILR ....Information Level Requirements

ILS...Index of Learning Styles

INT ....Intuitive

LS ....Learning Style(s)

NL ....Natural Language

PLR ....Process Level Requirements

PT ....Prompting Technique

RAD ....Rapid Application Development

REF ....Reflective

S ....Syntactic

SEM ....Semantic

SEN ....Sensing

SPT...Syntactic Prompting Technique

TC ....Task Characteristics

TCPT..Task Characteristics Prompting Technique

TLR ....Task Level Requirements

VER..Verbal

VIS ....Visual

LIST OF APPENDIX TABLES

1. INTRODUCTION

Requirements engineering is the first and an important step in the software development process for gathering and analyzing user needs. Software requirements are gathered from different stakeholders of technical and non-technical backgrounds and documented using natural language (NL). Due to the inherent ambiguous nature of NL, requirements gathering and analysis is especially prone to requirement problems [19]. Requirements can be missed, inconsistent, ambiguous, or just plain incorrect. Detecting problems early can help reduce the effort that is otherwise spent finding and fixing them during the later stages of software development, when they are much harder to find and fix. A major goal of the requirements engineering phase deals with the complexity of eliciting a complete set of requirements from a diverse set of stakeholders (e.g. end users, sponsors, customers) and to arrive at a good baseline set of requirements. An active research area in requirements engineering focuses on developing and empirically validating different types of requirements elicitation techniques that describe from whom and how requirements are elicited. While there has been a lot of research in the area, most of the traditional research is focused on using questionnaires or interviews to gather requirements. These techniques are effective at gathering requirements [11] but use close-ended questions that rely on an end user's *priori* knowledge and beliefs. As a result, these techniques lack a means of evoking mental prompts that would help elicitors and those filling the questionnaire avoid judgment biases or discover requirements that they may not have thought of otherwise. Therefore, there is ample scope for improvement to the requirements elicitation process.

Additionally, analysis of previous research revealed that traditional elicitation techniques and questions are highly context-dependent. Questions are set for a particular context and cannot be used interchangeably in reality for the information they acquire [10]. While elicitation techniques can be equally useful for simple, well-defined problems, differences include information acquired, information quality, and efficiency [1, 4, 10, 11, 23, 29]. Questionnaires are more efficient than most traditional elicitation techniques but aren't good at extracting tacit (non-recognized) or semi-tacit (implicit) knowledge [11]. Interviews are the most effective of traditional elicitation techniques at eliciting non-tacit (explicit) knowledge, but training in interviewing is an important factor for completeness, accuracy, and reliability [11]. Observation is time-consuming but can acquire tacit, semi-tacit, and non-tacit information. Extracting tacit information does, however, require using supplemental techniques.

Contrary to traditional elicitation methods, this thesis is focused on improving and validating a new class of requirement elicitation techniques called "*prompting techniques*". These techniques are context-independent and utilize prompts for targeting specific types of requirements information. The techniques use directed questions to overcome cognitive issues that can occur during requirements elicitation. More details on the prompting techniques and the directed questions used in these techniques appear in Section 2.

To empirically evaluate the validity of prompting techniques (PTs), the current study examines the effectiveness of syntactic (S) and task characteristics (TC) prompting techniques. Syntactic prompting technique (SPT) uses the interrogatories technique (questioning who, what, when, where, how, why), which is useful as building blocks for question construction but specific questions rely on the analyst. Task characteristics prompting technique (TCPT), on the other hand, makes use of a task model that offers guidance for interview content organized as organizational problems. TCPT also helps overcome cognitive issues with the use of strategies, such as building scenarios and generating counterarguments. Browne and Rogich [4] first introduced and studied the effectiveness of TCPT using SPT as a control.

Another aspect of this thesis investigates *Learning Styles* (LS) to further enhance the applicability of prompting based requirements elicitation techniques. LS's are the characteristic strengths of individuals on how they perceive and process given information. LS's have more recently been applied to research involving software requirements and has provided promising results [2]. There is evidence that LS does have an influence on efficiency and effectiveness [19]. Prior research involving LS has observed better efficiency and effectiveness in teams with different rather than similar LS for fault detection during software requirements inspections [19].

On the topic of requirements elicitation, researchers [2] have also used LS models to try to improve distributed requirement elicitation by forming an approach for choosing the best elicitation techniques for a distributed software project group. [2], however, lacks empirical evidence to validate improved requirements elicitation since it used a pilot study to determine correlations between LS and groupware tool selection. While [2] addresses LS for elicitation technique selection, the research in this paper studies LS and the effect on requirements elicitation responses. We hypothesize that an individual's LS can have an impact on their ability to effectively use a particular technique for eliciting and analyzing customer needs.

This research analyzes the effect of prompting techniques and stakeholders' LS on requirements elicited from them. A better understanding of the effect of LS on the communication of requirements would be beneficial to improving the requirement elicitation process with respect to completeness, quality, and quantity of requirements, with potential applications in helping tailor questions for different stakeholders' LS's. Since elicitation techniques can also have an effect on information acquired, it helps to be familiar with those and use a technique(s) that isn't as likely to receive responses that are too general. For this purpose, a modified version of the task characteristics prompting technique (TCPT) questions from [4] was used in questionnaire format for the controlled study of the effect of stakeholders' LS's on requirements elicitation. The syntactic prompting technique (SPT) was used as a control for TCPT's application, as it was in [4], to help further test the effectiveness of TCPT.

The goal of this research is to perform a comparative evaluation of TCPT and SPT while investigating the use of LS's for requirements elicitation. To accomplish this goal, a controlled empirical study was conducted at North Dakota State University (NDSU) with 97 students who used the TCPT questionnaire or the SPT questionnaire for a given scenario. Next, their responses were coded using generic requirement categories, tallied, and analyzed. Based on the results from the study, TCPT was more effective at eliciting requirements that also covered a wider breadth of requirements information without being more complex in usability. Also, the students' responses showed that those who like detail or are good at visualizing the big picture for a scenario were more effective with TCPT.

The remainder of this paper is structured as follows. Section 2 discusses related background work and motivations for this research. Section 3 is a discussion of how the background work was used to develop the research approach. Section 4 covers the design of the study for determining LS and evaluating responses elicited using the questionnaires. Section 5 is the data analysis and results. Section 6 discusses the results and their relevance for requirements elicitation. Section 7 includes suggestions for improvement and future work.

2. BACKGROUND WORK AND MOTIVATION

This section discusses the motivation for studying the effects of prompting techniques (PTs) and LS's in relation to requirements elicitation. It also provides background information relevant to the research methods employed in subsequent sections of this document.

2.1. Research Motivation

The potential usefulness of LS's in software development is a small but growing area of research. Previous research has shown that teams composed of inspectors with different LS's performed better at fault detection during inspections [19]. Software engineers with non-technical backgrounds have also performed better than those with technical backgrounds, which could be due to LS [7]. These results provided the motivation to further explore potential applicability of LS's on improved effectiveness in other areas of software development.

Inspections are meant to catch faults to reduce more work in later stages of software development, and in [19] the inspections occurred after there was a set of requirements specifications. Requirements elicitation takes place before inspections and involves acquiring important information about a project to be undertaken. This makes requirements elicitation a suitable stage for exploring improved effectiveness due to LS's, since the earlier reduction of faults reduces more time-consuming changes in later stages.

In respect to requirements elicitation, the application of LS's has been explored for improving distributed requirements elicitation. The usefulness of LS's was examined for selecting elicitation techniques most appropriate for all stakeholders and analysts of a project [2]. A better understanding of the effect of LS's on the communication of requirements could be beneficial to developing methods to improve the requirements elicitation process. There are reviews of existing requirements elicitation techniques [10, 24] and numerous research on improving requirements elicitation through the use of different techniques [1, 2, 4, 11, 29]. However, there isn't much research on the applicability of LS's in improvements in software development.

An ad hoc review of requirements elicitation techniques was conducted to have a better understanding of existing techniques and their observed shortcomings and advantages. It was decided that a questionnaire format would be utilized along with the use of prompting techniques. A questionnaire format is more efficient for the large number of participants in the study and also prevents influencing responses. The structured format of PT questions would make a questionnaire more effective than if it were unstructured [10], and the generic questions save time by being usable for different scenarios. Additionally, the strategies used by TCPT help overcome the traditional focus of non-tacit (explicit) and semi-tacit (implicit) knowledge to consider tacit knowledge (non-recognized knowledge) as well. PTs are also useful if requirements have to be elicited from distributed stakeholders. The cognitive issues PTs help overcome can occur in both local and distributed requirements elicitation.

The main motivation of this research is to study the effect of PT and LS on requirements information provided by stakeholders of a project. It can help with understanding if people with certain LS's provide more information in the different requirements categories. The usefulness of TC can also be validated and improved on with further studies for more effective and efficient requirements elicitation. The results of this study can be beneficial to future research on developing requirements elicitation techniques or questions for eliciting more requirements. With the questionnaires incorporated, the research can also help determine if a suggested improvement in requirements elicitation techniques, as in [4], can improve the quantity of requirements elicited by several if not all LS's.

2.2. Related Work

This section covers research on elicitation techniques and LS's that are relevant to and motivated the study.

2.2.1. Review of Requirements Elicitation Techniques

New ways to elicit requirements is constantly evolving. Dieste and Juristo's systematic review of studies on elicitation techniques served as the starting point for the ad hoc review of requirements elicitation techniques [10]. Their systematic review provided an overall review for our purposes and consistent evaluation of 30 empirical studies on elicitation techniques, which together cover 43 elicitation techniques. The elicitation techniques were tested in efficiency, effectiveness, and completeness. Though interviews were found to be the most effective, they aren't always the most efficient. Some observations were that structured are better than unstructured interviews for needs elicitation [1, 4] and hierarchical structuring techniques provide more information than unstructured interviews [5, 6, 9]. For our study, there was more interest in the reviews of questionnaire elicitation techniques and structured interviews research.

Some elicitation techniques, such as unstructured interviews and laddering, are equivalent in effectiveness when used by experienced and inexperienced respondents [1, 31]. However, the knowledge elicitation techniques acquire can differ. Unstructured interviews are good for early stages of requirement gathering to cover general questions; structured interviews are goal-oriented and a systematic exchange [1]. Elicitation techniques also differ in information quality and efficiency. Interviews are more effective at requirements elicitation, but questionnaires can be more efficient. A drawback of questionnaires, or other paper-based methods, is that they give the impression of communicating with someone with shared knowledge [29]. Implicit information about tasks may not be communicated without clarification through requirements discussion and evaluation, using interviews or face-to-face interactions. However, those methods can influence responses and hinder eliciting unexpected requirements. Moore and Shipman [29] suggested a graphical user interface (GUI) with built-in tools for graphics creation for diagrams to overcome assumptions of shared knowledge.

The review in [10] also tested for factors that can have relevance during elicitation, such as elicitor type factor [22], respondent experience factor [16], and task type factor [18]. This indicates already existing interest in factors that can influence the development of questionnaires. LS, instead, is a factor considered and evaluated in this paper. An ad hoc review of elicitation techniques was conducted, because [10] identified a large portion of empirical elicitation techniques research but only up to 2005.There are different types of knowledge elicitation techniques want to acquire. Implicit information is semi-tacit knowledge. Non-tacit knowledge is explicit information that is easier to extract. Tacit knowledge is non-recognized knowledge that isn't easily verbalized. Eva examined the requirements results elicited using different techniques in a study of the acquisitions of requirements (ACRE) framework, a disciplined Rapid Application Development (RAD) [11]. These requirements results are shown in Table 1. Useful methods for acquiring the different types of knowledge are in Table 2. The listed techniques don't include the ones that aren't useful alone.

These tables show that different techniques help elicit different types of requirements information, in precision and type of knowledge. While the listed techniques are more effective for different levels of precision or knowledge, that doesn't mean they are efficient. The methods for eliciting non-tacit requirements have the drawback of being more time consuming. In fact, observation was the method that is able to retrieve information from all knowledge areas the best. Nevertheless, it's one of the most inefficient methods.

Table 1. Requirements from Elicitation Techniques (Based on results from [11])

Elicitation Technique	Requirements Obtained
Brainstorming	Only high level requirements
Task analysis	Precise requirements
Scenario analysis	Precise requirements
Critical success factor analysis	Precise requirements
Prototyping	Very fine-grained requirements; finding missed and taken-for-granted requirements

Table 2. Better Methods for Acquiring Requirements Knowledge (Based on results from [11])

Knowledge Area	Method
Tacit	Observation Scenario analysis Task analysis Prototype provided to user Personal Construct Theory techniques
Semi-tacit (implicit)	Prototyping
Non-tacit (explicit)	Structured interviews Record searching Observation

Browne and Rogich created a context-independent structured interview technique – the task characteristics (TC) technique [4]. There are several useful models for structuring interviews [1, 4]. TC uses the requirements determination task model to address organizational goals, business processes, tasks to achieve goals, and information determining task behaviors. Their technique addresses cognitive issues of the common difficulties with capturing a stakeholder's conceptualization with strategies. Table 3 includes strategies from [3] and [4] for eliciting stakeholders for requirements. The other two difficulties that TCPT doesn't address are problem structuring issues and communication issues.

Table 3. Strategy Types for Eliciting Requirements (Borrowed from [3, 4])

Strategy	Description
Scenario building	Ask user to imagine or construct a scenario in domain and respond as user would in that situation
Conditionalizing	Using “if-then” clause to modify an assertion to limit or clarify its applicability
Elaborating with Instances	Asking user to illustrate a point by providing examples
Hedging	Asking a user to design contingency plans or fallback positions
Generating counterarguments	Asking user to argue against the conclusion they first reached
Generating arguments	Asking user to make more arguments favoring position, or different kinds of arguments
Feedback	Asking user for feedback or providing them with feedback on what was said
Summarization	Asking user to summarize what they said, or providing summarization for them
Directed questions	Stimulate user’s memory, causing associations to be made and causing user to think of things they otherwise may not
Flow chart	Represents events as a linear process, reducing working memory demands and allowing all parties to comment on correctness of flow
Evocative knowledge map	Represents knowledge and beliefs in a non-linear diagram, reducing working memory demands
Influence diagram	Represents influences on process steps or stages, reducing working memory demands and allowing agreement among all parties as to influences
Decision map	Captures decision maker’s mental model of a decision, task, or environment. Reduces demands on working memory and allows discussion among all parties involved
Affinity diagram	Organizes and categorizes information, beliefs, and/or arguments
Note board	Organizes and categorizes information, beliefs and/or arguments, with the added advantage that the individual pieces of information are portable

The use of a context-independent technique has the advantage of applicability to various tasks without requiring analysts to have extensive knowledge of the context [1]. Context-dependent techniques may be able to elicit more relevant requirements, but techniques generic enough to be applied to multiple scenarios are more efficient. Context-independent techniques require less time to develop or modify questions or to train the stakeholders requirements are elicited from. Context-dependent techniques may require training, since it's an important factor for the effectiveness of stakeholders eliciting requirements (e.g. analysts or knowledge engineers) when developing or adapting questions or interviewing for completeness, accuracy, and reliability [1]. Generality is the drawback of the TC technique. However, the drawback can also make stakeholders think outside the confines of their responses to specific, context-dependent questions. PTs, such as TC, can be context-dependent or context-independent. The PTs in our study are context-independent.

With the use of substantive and procedural prompts, the TC technique can elicit specific types of requirements or acquire them through the use of strategies. Generic requirements categories are adapted in [4] to measure the usefulness of TC. These are covered further in Section 4.5.2. The generic requirement categories are also a useful guide for developing questions that can cover important knowledge to be acquired, to know when to stop eliciting, and to confirm what knowledge has been acquired.

Browne and Rogich [4] determined that TC performed better than the other PTs – S as control and semantic (SEM) devised by [27] that uses questions based on knowledge structures – in their controlled experiment. In the categories for its task model, TC had significantly more requirements and a greater number of process level requirements (PLR) and information level requirements (ILR). There wasn't a noticeable difference in breadth of requirements elicited or significant qualitative difference in requirements elicited.

2.2.2. Studies of Applicability of Learning Styles

Kolb [25] introduced the concept of LS's and is recognized with development of the first LS measurement instrument. Over the years, psychologists have developed a variety of LS models [8, 17, 30] and validated the use of LS's in engineering education [14]. Previous research revealed that the *Felder and Silverman's Learning Style Model* (FSLSM) is the most advanced and widely used to measure the LS's through an instrument called the *Index of Learning Styles* (ILS) [12-15].

This research utilizes the FSLSM, to capture most important LS preferences among individuals [14] and then classifies characteristic strength and preference of an individual across four LS dimensions. These dimensions relate to the way individuals perceive and process information. The two dimensions which relate to perceiving information are: a) Sensing/Intuitive; and b) Visual/Verbal. The remaining two dimensions – Active/Reflective and Sequential/Global – relate to information processing. Brief descriptions of the four dimensions of the LS model borrowed from LS studies in academia [14] are listed below:

- *Sensing* (SEN) Learners: Observant and careful with details, so they might be slow in completing tasks. Individuals with sensing preference like working with facts, concrete content, data, and experimentation. LS theory suggests that they reach a solution by using existing methods and are good at remembering facts. Sensing people face issues with understanding words, which requires them to read information presented to them again and again.
- *Intuitive* (INT) Learners: Individuals who are oriented towards concepts and theories. Intuitive people are innovative and don't like repetition, so they like complications in their tasks. They don't like details, so they are fast in completing tasks but might be careless. Intuitive people like to perceive information via symbols, which makes them comfortable translating words.

- *Visual* (VIS) Learners: People who prefer information in the form of pictures, diagrams, flowcharts, or video demonstration, so they may forget if someone has said something to them.
- *Verbal* (VER) Learners: People who tend to prefer and remember written or spoken explanations. They can benefit from discussion and are effective in explain things to others.
- *Active* (ACT) Learners: People who like to work in groups and tend to try things out. Active learners are practical and, hence, they cannot learn much from lectures or other forms of passive information.
- *Reflective* (REF) Learners: Individuals who tend to work alone and learn by thinking about information first. Reflective learners are better at theories and cannot learn when they are not given an opportunity to think.
- *Sequential* (SEQ) Learners: Tend to learn in a logical fashion. They follow linear reasoning in small steps during tasks and like gradual progression of difficulty. Sequential people can work with partial understanding of information, which makes them a good at analysis.
- *Global* (GLO) Learners: Tends to jump in large steps to complete tasks, which makes them divergent in their thought process. They might be able to solve a complex problem but be unable to explain how they solved it. Global people may have trouble understanding partial information.

The ILS instrument is an online questionnaire with 44 questions to measure LS's on the four dimensions and has been used in past studies [20, 21, 26, 32, 33]. Each of the four dimensions in ILS has 11 questions with two options (*a or b*) corresponding to one of the categories in that dimension. For example, in the Visual/Verbal dimension, if a person selects ten (10) answers that favor the VIS category and one (1) towards the VER category then the score will be nine (9) (i.e. 10-1) towards the VIS category. The symbol ‘X’ on the top of the score represents the preference towards a category in a LS dimension. The number of answers favored for a category is termed Actual Score (actual score of VIS category is 10 and for VER category is 1 in the example) in our research. A score between 1-3 on ILS represents that a person is balanced and is balanced towards both the categories in an LS dimension. A score between 5-7 and 9-11 states that the person has a moderate and strong preference, respectively, toward a category in a LS dimension. The ILS instrument has been empirically validated for its reliability and construct validity [15]. Figure 1 shows a sample ILS output.

```
ACT                                         X               REF
     11   9   7   5   3   1   1   3   5   7   9   11
                         <-- -->

SEN                 X                                       INT
     11   9   7   5   3   1   1   3   5   7   9   11
                         <-- -->

VIS       X                                                 VRB
     11   9   7   5   3   1   1   3   5   7   9   11
                         <-- -->

SEQ                           X                             GLO
     11   9   7   5   3   1   1   3   5   7   9   11
                         <-- -->
```

Figure 1. Example Result of the Questionnaire on the ILS

The applicability of LS to software development is a still new and growing area of research. Goswami and Walia did a study that addressed how inspectors are different in how effective they are during inspections [19]. LS is applied to study the effectiveness of fault detection during inspections performed by inspectors with different LS in a team and individually. The results show that teams composed of inspectors with different LS perform better in efficiency and effectiveness, and an evaluation of how LS categories do at requirements inspection showed greatest effectiveness and efficiency for ACT-SEN-SEQ LS individually. However, the controlled studies were small. There were also very few individuals who were VER in the Visual/Verbal dimension. Areas where the research could be further explored include having a larger dataset to study that include technical and non-technical inspectors that could provide more revealing results. Goswami and Walia's 2013 study [19] itself was motivated by results showing that software engineers with non-technical backgrounds perform better, so the effectiveness could be due to LS [7].

There has been prior research on applying LS's to requirements elicitation but in the form of determining suitable elicitation techniques for distributed teams [2]. Aranda et al. addressed how unfamiliarity with a requirements elicitation technique can negatively impact a team member's participation. Factors in addition to individual team members' LS's were groupware tools, common elicitation methods for distributed teams, and groupware tool preference. Some shortcomings were that the pilot group was small and there were difficulties in selecting an appropriate requirement elicitation technique that involved determining how to weigh stakeholder preferences. There wasn't empirical evidence to validate improved requirements elicitation. Our research tries to investigate the impact of LS's on the effectiveness of the requirements elicitation process when using the prompting based techniques.

3. RESEARCH APPROACH

This section describes the system scenario that was used to evaluate the efficacy of requirement elicitation techniques and the modification made to the prompting techniques for evaluation.

3.1. Development of Scenario

Development of a scenario for the PT questionnaires was inspired by [4], which used a brief explanation that a supermarket chain wanted to create an online shop for grocery shopping. For the scenario in [4], the user was told an analyst will ask them questions to determine what the new system needs, therefore determining requirements. Initially, the scenario developed for our study was a similarly simple introductory scenario for a virtual reality head-mounted display for gaming.

NewVRTech Requirements Elicitation Task

NewVRTech is an international company that is planning to provide an immersive gaming experience for its customers with a virtual reality head-mounted display. The device will comfortably supplement a customer's gaming experience and utilize natural body movement when a customer is exploring their virtual environment. However, the virtual environment shouldn't interfere with a customer's view of real life obstacles around them while using the device.

NewVRTech is first planning to produce the product for distribution nationally. NewVRTech first needs a set of requirements to guide development of the device. As the manager at NewVRTech to head the project, you're responsible for defining the virtual head-mounted display's requirements. A requirements engineer from the development department will provide you with questions to help you communicate what should be included in the system.

Figure 2. NewVRTech Scenario with Image from [28]

The above scenario was eventually replaced with a more generic game development scenario that was a better fit to the course objectives, outside the control of the researchers, and is discussed in Section 4.3. Both of the scenarios were kept concise so participants could think creativity in their responses for the needs of the software system.

3.2. Modification to Questions

The initial thought was that the elicitation technique should be one that is more effective at eliciting requirements from the different knowledge types, so it's easier to notice differences in elicited requirements that can be due to participants having different LS's. Since the study would have student participants, it was decided that a questionnaire would be used to fit time constraints. A questionnaire wouldn't be as effective as interviews and differs from the manner of elicitation in [4], but it gives more assurance that responses won't be influenced by the interviewer in the case of the latter [29].

It was decided PTs would be used for the questionnaire questions, because their structured format and directed questions add effectiveness to requirements elicitation. Also, the TCPT strategies add additional effectiveness with how they help overcome cognitive issues, one of the three common difficulties relevant to research along with problem structuring and communication issues. TCPT and its strategies would help elicit tacit in addition to non-tacit and semi-tacit knowledge. It also helps that the questions are generic enough to be useful for different scenarios, which was helpful when there was a change in scenarios for the study.

Since [4] already had questions listed for SPT and TCPT, those could be used for our study for validation of their findings. There were modifications to the questions for TCPT, however, to incorporate more strategies. In [4], TCPT's strategies included: directed questions, counterargument, summarization, feedback, and scenario building. The adapted TCPT utilizes the following strategies from Table 3 in its substantive and procedural prompts: directed questions, arguments, counterargument, summarization, feedback, scenario building, hedging, conditionalizing, elaboration, flow chart, decision maps, and influence diagrams. Table 4 is a mapping of the strategies utilized in our adapted TCPT, which is the TCPT referred to in the rest of this thesis. The strategies that ask for inclusion of visual responses may or may not have been used by the students, because it would depend on what they decide to include.

Table 4. Strategies Used in Task Characteristics Prompting Technique (Adapted from [4])

Question	Strategy(ies)
Overview of System	
What would customers want the system to do?	Substantive prompt
How will it perform better than other similar systems?	Procedural prompt - Generating arguments
What are examples of how it does these better?	Procedural prompt - Elaborating with instances
If it does these things better, how will that impact a customer's experience?	Procedural prompt - Conditionalizing
Why might the system not work better than other similar systems?	Procedural prompt - Generating counterarguments
Why would your customers not want to use the system?	Procedural prompt - Causal counterargument
What can be done to overcome these negatives? (consider previous two answers)	Procedural prompt - Causal counterargument
What if these strategies for overcoming negatives don't work?	Procedural prompt - Hedging
What would your company want the system to do?	Substantive prompt
Summarize everything the system should do. You may include any use-case diagrams (or other notation).	Procedural prompt - Summarization, Feedback
How to Use the System	
What does a customer need to do to use the system? • Consider including flow charts or decision maps. • What influences each step?	Substantive prompt Procedural prompt • Flow chart • Elaborating with instances, Scenario building, Influence diagram

Table 4. Strategies Used in Task Characteristics Prompting Technique (Adapted from [4]) (continued)

Question	Strategy(ies)
How to Use the System	
Can you think of situations in which the customer would have problems using the system? • Consider previous answer and diagrams • Expand on diagrams where necessary	Procedural prompt - Scenario building • Feedback, Scenario building • Feedback, Flow chart, Decision map
What can be done to overcome these problems? • Consider scenarios and previous flow charts	Procedural prompt - Casual counterargument • Feedback, Flow chart, Scenario building
Summarize steps for using the system. Include any use-case diagrams.	Procedural prompt - Summarization, Feedback
Support	
What people or departments are needed to support customer's use of the system?	Substantive prompt
Describe and detail the tasks these people or departments must do • Consider including a flow chart, use-case diagrams, and/or outlining influences for different individuals involved	Substantive prompt Procedural prompt - Flow chart, Influence diagram, Decision map (for use-cases)
What feedback should the system provide for performing these tasks? Use-case diagrams or feedback scenario examples can be included.	Substantive prompt Procedural prompt - Decision map, Scenario building
During Use of System	
Can you think of a situation in which the customer would have to make a decision or choice when using the system? • A decision map can be included	Procedural prompt - Scenario building • Decision map
What does using the system allow people to do that they couldn't without the system?	Procedural prompt - Casual counterargument
What functions are people able to perform currently that they can't with the system?	Procedural prompt - Scenario building, Causal argument
Information Necessary	
What information must a customer supply to the system to use it?	Substantive prompt
Are there any additional supplies that would aid the use of the system (e.g. movement tracking software, certain type of computer, etc.)?	Substantive prompt
What information must the system supply to the customer?	Substantive prompt
What information must the system display to the customer?	Substantive prompt

4. EXPERIMENT DESIGN

The main goal of this research was to perform the comparative evaluation of two different PTs based requirements elicitation techniques. We also evaluated the impact of LS's of the stakeholders involved during the requirements elicitation process on the usability of prompting based requirements elicitation techniques.

4.1. Research Questions

The research questions investigated in this study are discussed as follows:

Research Question 1: Which prompting technique (Syntactic - S vs. Task Characteristic - TC) is more effective at eliciting requirements?

Research Question 2: Do LS's of individual subjects affect their ability to provide requirements during requirements elicitation?

Research Question 3: Which prompting technique (Syntactic - S vs. Task Characteristic - TC) is perceived as more useful by subjects at eliciting requirements?

4.2. Participating Subjects

The subjects participating in this study were ninety-seven (97) computer science undergraduate students enrolled in the Software Development for Games course at North Dakota State University during Spring 2016 semester. The course required students to work on several game-related projects, wherein they would elicit requirements and then implement those requirements later in the semester working individually or in groups. The subjects were randomly divided into two groups – an experiment group that used TCPT and a control group that used SPT. The division of subjects – 56 students who took TCPT and 40 took SPT – doesn't include the person who took both.

4.3. Artifacts: Prompting Techniques and System Scenario

The subjects were provided the following system scenario that required them to use a prompting technique – S or TC depending upon the group they were assigned – to elicit requirements related to stakeholders, how to use the system, support needed to use the system and other pertinent details related to the development of an educational game project. The scenario was kept concise to enable subjects to channel their creativity in eliciting the needs for such a software system.

> A project is to be defined for use by the game programming class. The purpose of the project is to help the students learn about developing a new software product in a real-world manner. The project will include the actual development of a simple game. What can we say about this game?

The PTs (TC and S) used to elicit requirements are shown in Appendix A. A high level overview of these techniques appears in the background section. Each technique contained a set of questions (e.g., *Who uses the system*? *How do they (users) affect the system*?) that required subjects to answer in context of the educational game project scenario described above. The TCPT was developed as a context-independent requirements elicitation technique and contains two types of directed questions: ones that are aimed at eliciting specific types of requirements (e.g., *What must the customer do to support the system?* – this set of questions is similar to the SPT) and to elicit requirements that are otherwise not brainstormed or gathered because of the cognitive biases or obstacles. This set of questions is based on theoretically and empirically validated reasoning strategies (e.g., *What kind of things can people do now that they might not be able to do when using the system?*) that tries to evoke their thoughts and help them reduce judgment bias with the goal of eliciting a larger set of requirements.

4.4. Experiment Steps

The experiment procedure contained the following steps during the course of the study:

Using prompting techniques to elicit requirements: Participants were instructed by the instructor to answer questions specific to their assigned technique in context of eliciting and clarifying requirements for an educational game project scenario, which would in turn impact their course deliverables. The subjects were randomly assigned to a control group (SPT) and an experiment group (TCPT) to document their answers to a specific set of questions. During this step, subjects individually answered the questions and reported their answers. This was done as a take-home assignment wherein the subjects were instructed not to consult with each other and answer the questions using the knowledge of the requirements elicitation techniques.

Post-study survey: At the end of the study, we asked the students to reflect on their requirements elicitation experience. We also discussed the issues they may have faced when using the PTs to elicit requirements for the educational game project scenario. The survey is provided in Appendix B.

Learning Styles questionnaire: At the conclusion of the experiments, all participants were given the Felder Silverman's LS questionnaire. The questionnaire can be accessed at: https://www.engr.ncsu.edu/learningstyles/ilsweb.html. Participants answered all 44 multiple choice questions and, the LS results were generated for each participant on the ILS scale. For each dimension on ILS (Active/Reflective, Sensing/Intuitive, Visual/Verbal, and Sequential/Global), the participant has a score towards one category. Hence, only four LS categories – one from each dimension – form the LS of an individual with a score of 1-3, 5-7 or 9 -11.

4.5. Data Collection and Evaluation Criterion

This section describes the data collected during the experiment steps and the criterion for evaluating the data prior to analysis. This section also discusses the process of transcribing textual data into a number of relevant requirements using a coding scheme and the process of inclusion and exclusion of individual responses.

4.5.1. Coding Scheme Development

For a coding scheme, the generic requirement categories from [4] were used, since these were already developed and adapted for use in coding responses for the questionnaires, which are mostly the same in our adaptation. The primary difference is that their coding scheme is applied to verbal answers, whereas ours is applied to textual answers.

4.5.2. Requirement Categories

Generic requirement categories are divided into goal level requirements (GLR), process level requirements (PLR), task level requirements (TLR), and information level requirements (ILF). The main categories are based on the task model used for TCPT. These categories are then further broken down into subcategories. A list of generic categories and their descriptions is in the Table 5.

Table 5. Generic Requirement Categories (Borrowed from [4])

Generic Requirement	Description
Goal Level Requirements	
Goal State Specification:	Identifying the particular goal state to be achieved.
Gap Specification:	Comparing existing and desired states.
Difficulties and Constraints:	Identifying factors inhibiting goal achievement.
Ultimate Values and Preferences:	Stating the final ends served by a solution.
Means and Strategies:	Specifying how a solution might be achieved.
Causal Diagnosis:	Identifying the causes of the problematic state.
Knowledge Specification:	Stating facts and beliefs pertinent to the problem.

Table 5. Generic Requirement Categories (Borrowed from [4]) (continued)

Generic Requirement	Description
Goal Level Requirements	
Perspective:	Adopting an appropriate point of view on the situation.
Existing Support Environment:	Description of the existing technological environment that can be applied to support the system to be developed.
Stakeholders:	Organizational units, customers, suppliers, competitors.
Process Level Requirements	
Process Description:	A series of steps or tasks designed to produce a product or service.
Process Knowledge Specification:	Facts, rules, beliefs, algorithms, and decisions required to perform a process.
Difficulties, Constraints:	Factors that may prohibit process completion.
Task Level Requirements	
Task Description:	Identification of the sequence of actions required to complete a task.
Task Knowledge Specification:	Facts, rules, beliefs, assumptions, algorithms, and decisions required to perform a task.
Performance Criteria:	Statement that associates an outcome with specific conditions, actions, and constraints.
Roles and Responsibilities	Individuals or department who are charged with performing tasks or steps within tasks.
Justification:	Explanation of why specific actions are or are not to be taken.
Information Level Requirements	
Displayed Information:	Data to be presented to end-users in paper or electronic format.
Interface Design:	Language and formats used in presenting “Displayed Information.”
Inputs:	Data that must be entered into the system.
Stored Information:	Data saved by the system.
Objects and Events:	Physical entities and occurrences in the world that are relevant to the system.
Relationships Between Objects and Events:	Description of how one object or event is associated with another object or event.
Data Attributes:	Characteristics of objects and events.
Validation Criteria:	Rules that govern the validity of data.
Computations:	Information created by the system.

4.5.3. Unit of Requirement

To determine requirements from textual data collected from participant responses to the PT questionnaires, it was necessary to define what a requirement is. In the study, a requirement was defined as: *A statement for a singular necessary attribute of a system.*

Participant responses were coded into requirements based off this definition. Therefore, a one-sentence response to a question could be coded as providing more than one requirement. Each determined requirement was then coded to one category it best fits according to the generic requirement categories. A partial sample of how requirements were coded is shown in Figure 3. The first answer provided a stakeholder and also two instances of ultimate values and preferences in the highlighted portions. In the second answer, the requirements gained could be considered ultimate values and preferences, goal state specification, or gap specification. The chosen category is gap specification, because they answer how the system compares to other similar ones and answer the question in that manner.

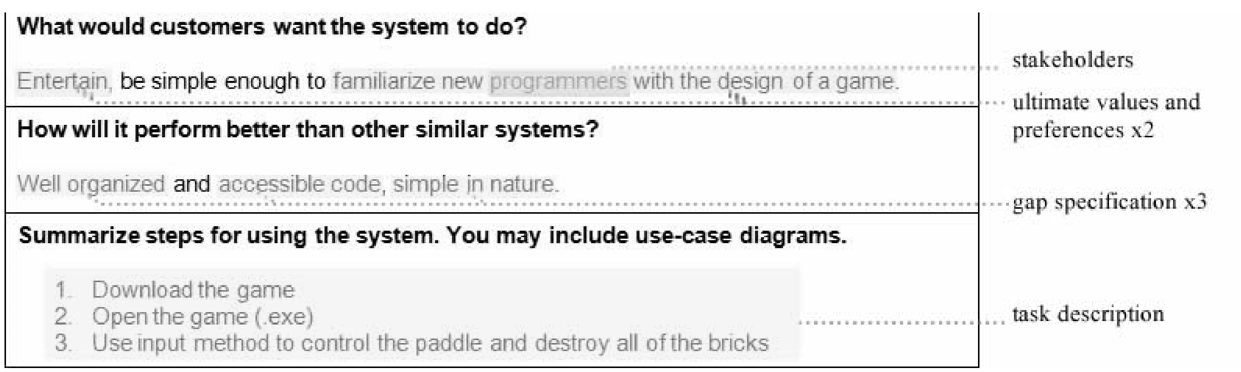

Figure 3. Questionnaire Response Coding Sample

4.5.4. Exclusion/Inclusion Criterion

Some data had to be excluded so that only useful data was evaluated to address the research questions. The three forms of data acquisition were the PT questionnaires, post-study survey, and LS questionnaire. In some cases, all three forms of data from some participants were excluded.

Due to some students' confusion about the scenario and PT questionnaires, there were responses that weren't based on the provided scenario. For the sake of experimental control, their data was excluded, so evaluated data only included responses based on the scenario. A second reason to exclude data was if a participant filled out both S and TC questionnaires, since answering one prompting technique questionnaire may affect their responses on the other version. Students who submitted both S and TC aren't included in the separate counts of S and TC participants. Thirdly, some participants left many unanswered questions. To limit the impact on quantitative data on requirements elicited, participant data was excluded if they left five (5) or more questions unanswered. While participants did include vague responses, data from these students weren't excluded since that could be indicative of the effectiveness of a questionnaire or how people of certain LS process information during requirements elicitation.

5. RESULTS AND ANALYSIS

This section includes analysis of the data collected from the two prompting techniques and LS questionnaires and survey on the prompting technique questionnaires. Section 5.1 evaluates the data collected from the prompting techniques questionnaires. Section 5.2 discusses observations about LS's and the collected data. Section 5.3 looks at participants' opinions of the prompting techniques questionnaires they respectively took.

5.1. Evaluation of Prompting Technique Data (RQ 1)

Part of the research goal was to evaluate differences in requirements information acquired through different questionnaire prompting techniques, as in [4]. In addition to further studying the effects on information acquired using different prompting techniques, this is useful for analyzing differences that could correlate to differences in LS's.

An overview of the coverage of requirement categories is shown in Figure 4, which depicts the mean number of requirements information elicited for each category from S and TC questionnaires.

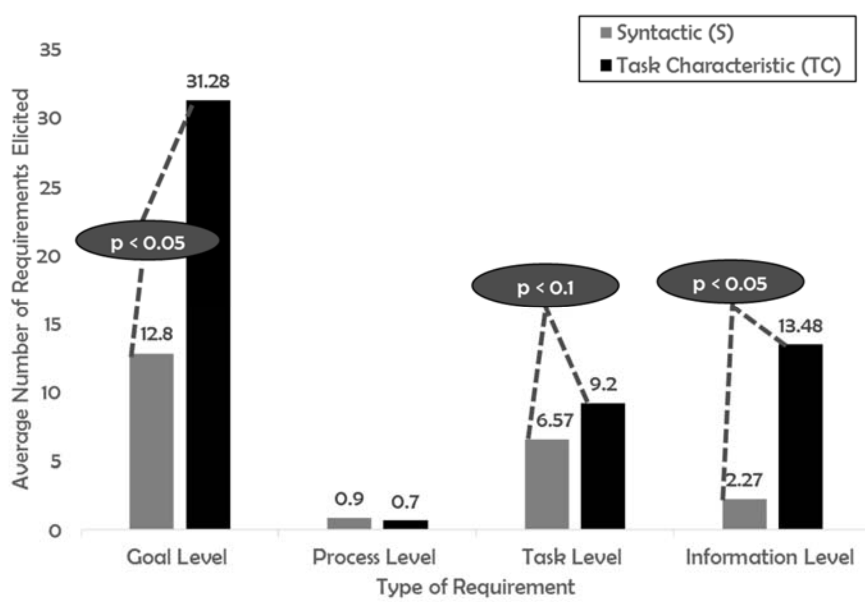

Figure 4. Comparison of Syntactic and Task Characteristic Effectiveness

An independent samples t-test evaluated the effectiveness of S and TC techniques by statistically comparing whether the average number of information elicited for each requirement category – GLR, PLR, TLR, or ILR – for each technique is significantly different. An alpha value of 0.05 was used to test the statistical significance. From the data acquired, the following observations can be made:

- TC resulted in more requirements for GLR, TLR, and ILR but in fewer PLR. The results from an independent samples t-test confirmed that TC helped subjects elicit a significantly larger number of GLR, TLR and ILR information as compared to S prompting technique.
- Interestingly, both S and TC resulted in the least amount of information related to PLR. The results from an independent samples t-test showed no considerable difference between SPT vs. TCPT in their effectiveness at eliciting information related to PLR.
- The largest difference in number of requirements between S and TC occurs in GLR – an average of 13.48 requirements for TCPT vs. an average of 2.27 requirements for SPT – which was found to be statistically significant ($p < 0.05$). A more detailed analysis on the sub-categories for GLR, ILR, TLR and PLR are provided in the following paragraphs.

To understand the relative effectiveness of prompting techniques (S vs. TC), Figure 5 compares the average number of requirements found by control group (SPT) subjects and the average number of requirements found by experiment group (TCPT) subjects across all 27 sub-categories (10 sub-categories within GLRs; 3 sub-categories within PLRs; 5 sub-categories

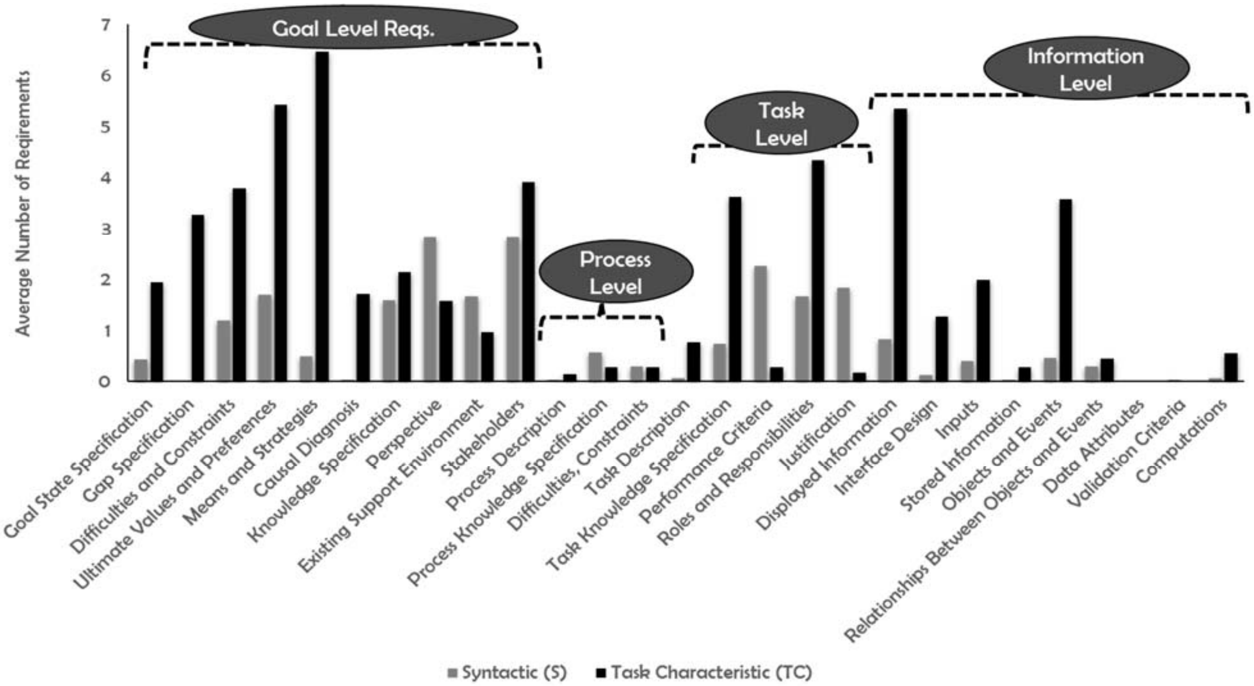

Figure 5. Comparison of Syntactic and Task Characteristic Effectiveness Across Sub-Categories

within TLRs; 9 sub-categories within ILRs). The major observations from Figure 5 are discussed as follows:

- The significantly higher effectiveness of TC technique (vs. S technique) is not restricted to a single sub-category but rather distributed across a majority of the sub-categories. This reinforces that, TC was significantly more effective at evoking requirements information from stakeholders that are otherwise overlooked when using the S technique.
- Looking at the number of sub-categories (e.g., 10 and 9 for GLR and ILR respectively vs. 3 and 5 for PLR and TLR), it was expected that the prompting techniques would help uncover larger number of information related to GLR and ILR as compared to PLR and TLR. The result from this study also showed that the largest difference between the techniques was found in evoking GLR and ILR.

Consequently, there was not an observable difference in discovering PLR, which had the fewest number of sub-categories.

- In terms of the effectiveness of TC in terms of the sub-categories; around 20% of goal level requirements were related to means and strategies (i.e., *specifying means and strategies for goal achievement*) whereas around 40% of information level requirements were mostly related to displayed information (i.e., *information related to data that needs to be presented to end-users of the system*).

Finally, we also wanted to investigate if the requirements information elicited from the control and experiment group were distributed across the subjects to understand the usefulness of technique across subjects. Figure 6 shows the variability in the effectiveness of each subject within control (S) vs. experiment group (TC) across all four requirement categories (GLR, PLR, TLR, and ILR).

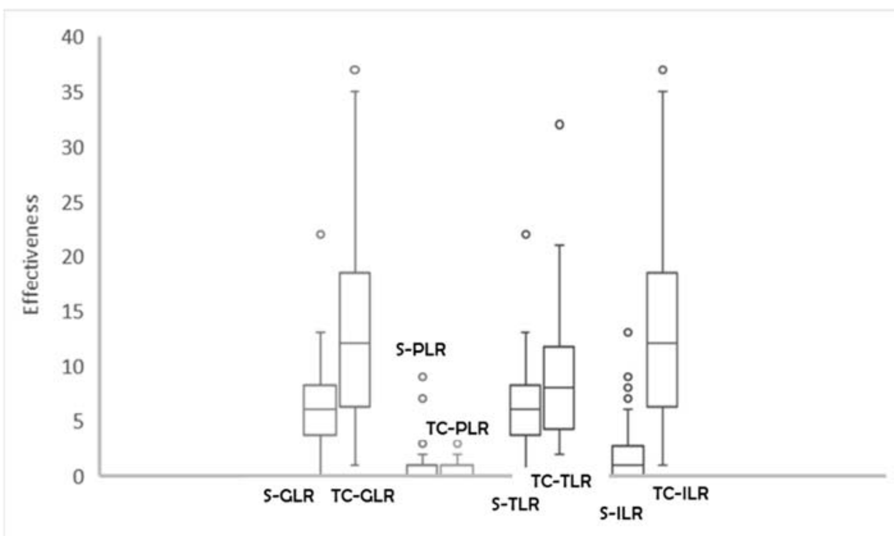

Figure 6. Variability in the Subjects' Responses using S vs. TC technique

Based on the results in Figure 6, there are very few outliers across both techniques and four requirement types. As expected, on average, the subjects using TC found more information related to requirements types as compared to the subjects using S technique.

These observations also supplement the findings in [4]. In their research, the TCPT introduced in [4] was also found to result in more requirements than SPT. However, there were significantly more GLR and ILR from this study instead of PLR and ILR. These differences could be partially due to using different scenarios, experiment setting, and different interpretations of the generic requirement categories when coding. PLR also has the fewest sub-categories. There was also a noticeable difference in breadth of requirements elicited from the study, with better coverage of the generic requirements by TCPT, which wasn't observed in [4].

5.2. Observations about LS and TC prompting technique (RQ 2)

In previous research, Goswami and Walia found significant correlations between the LS of software engineers and their performance during the software engineering task of software inspections [19]. Additionally, previous research in requirements elicitation and gathering has shown that, end-users, requirements engineers, and other stakeholders involved during the requirements discovery and analysis have varying LS's that should be taken into consideration when selection the requirements elicitation techniques [2]. The results from Section 5.1 show that TCPT is significantly better at evoking a more comprehensive set of requirements information as compared to SPT. Therefore, this section extends the previous research to analyze the correlations between LS of subjects and their effectiveness at identifying relevant requirements information using TCPT.

We analyzed the impact of certain LS's, or combination of LS categories for each participant, on requirements elicitation effectiveness. To perform this analysis, the participants were divided into different clusters, where each cluster represents a certain combination of LS categories. Next, the average effectiveness of subjects in each cluster, based on the total number of requirements found by participants in each cluster, were calculated and compared across different clusters.

Using six LS categories across three LS dimensions, we created eight clusters of possible LS combinations: 1) ACT-SEN-SEQ, 2) ACT-INT-SEQ, 3) REF-SEN-SEQ, 4) REF-INT-SEQ, 5) ACT-SEN-GLO, 6) ACT-INT-GLO, 7) REF-SEN-GLO and 8) REF-INT-GLO. Since only two participants were VER learners, VIS and VER didn't record any differences and weren't included in the LS combinations. Using the LS results, participants were grouped into the eight clusters. For example, a participant who has preference towards REF, SEN, and SEQ category across three dimensions can be placed in REF-SEN-SEQ cluster. Similarly, all participants were grouped into the clusters based on their ILS score sheet.

Figure 7 shows the average effectiveness of each cluster – a combination of LS categories – calculated by averaging the number of relevant requirements information reported by subjects belonging to each cluster and for each of the four requirement categories of GLR, PLR, TLR, and ILR.

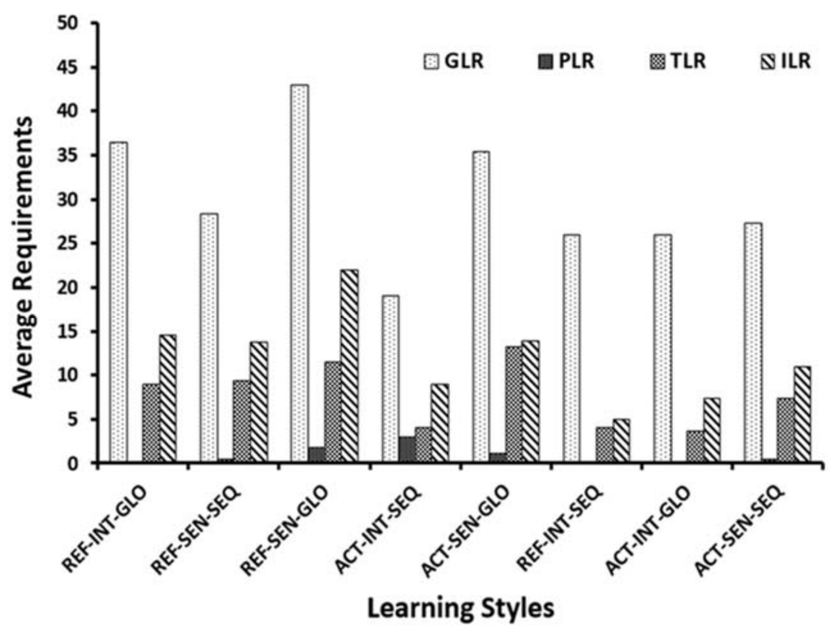

Figure 7. LS vs. TC Effectiveness

The major result from Figure 7 is that, the subjects with REF-SEN-GLO LS's had the maximum average effectiveness at discovering GLR and ILR's. Also, ACT-SEN-GLO subjects reported a large number of TLR, ILR and GLR's. Therefore, SEN and GLO learners seem to be especially effective at discovering a large number of requirements across all requirement categories.

This result is consistent with the theoretical underpinnings of SEN and GLO learners. Based on the cognitive psychology research, SEN learners are observant and careful with details due to which they seem to be better suited at discovering more requirements when prompted with specific set of questions. Similarly, GLO learners tend to jump in large steps to complete their task which makes them divergent in their thought process and are able to solve a complex problem which aligns with their abilities to make them good at requirements analysis.

5.3. Perceived Usefulness of Prompting Techniques (RQ 3)

While quantitative data analysis showed that, TCPT outperformed SPT, we also wanted to analyze the subjects' perceived usefulness of each technique. To perform this analysis, subjects rated the technique that they used on seven different attributes using a 5-point likert scale (1- very low; 3- medium; 5- very high). The attributes and an average rating of subjects across each attribute is shown in Table 6.

The result in Table 6 shows that, SPT was rated higher on simplicity and easy to use. These are expected, because SPT has fewer prompts, is simpler in design, and lacks built-in checks to avoid judgment bias and obstacles. Overall, the subjects rated both techniques similarly. Therefore, combining the results from quantitative and qualitative data analysis, TCPT is significantly more effective than SPT while not being harder to use and apply.

Table 6. Subjects' Rating of S vs. TC Prompting Technique

Attribute	S	TC
Simple	4	3
Easy to Understand	3	3
Easy to use	4	3
Intuitive	4	3
Comprehensive	4	3
Useful	3	3
Overall opinion	3	3

There were common comments and difficulties the subjects had regarding the questionnaires. The most common complaint was that the scenario or system definition was too vague (44% who took SPT; 39% who took TCPT). The second most common complaint was that the questions were confusing (50% who took SPT; 27% who took TCPT). Suggestions for both SPT and TCPT included having clearer questions, more detail and instructions for the questionnaires, and providing a more detailed scenario. There were also the suggestions to provide examples of categorizing requirements for SPT and having the questionnaire not as long and less repetitive for TCPT. Only a few subjects mentioned repetitiveness (19% for SPT; 9% for TCPT) and clearer questions (25% for SPT; 21% for TCPT).

Subjects' qualitative feedback indicates that those who used SPT provided more input on difficulties. The higher rating on *simplicity* and *easy to use* for SPT could be due to the format of the questions, rather than the subjects' ease in responding to them. Also, fewer subjects took the SPT than the TCPT questionnaire.

6. CONCLUSION AND DISCUSSION OF RESULTS

This section answers the research questions postulated in Section 4 using the evidence reported in Section 5. Additional discussion of results and its relevance to research and practice is provided in this section.

Research Question 1: Which prompting technique (*Syntactic* - S vs. *Task Characteristic* - TC) is more effective at eliciting requirements?

Based on the results shown in Figure 4 and 5, TCPT helped subjects report a significantly larger number of requirements belonging to GLR, TLR, and ILR categories. Regarding PLR, there was no observed difference which could stem from the fact that PLR had the fewest sub-categories and corresponding prompts (questions) to answer during the experiment. We also know from the survey responses that a large number of the subjects thought the scenario or system definition were too vague (44% who took SPT; 39% who took TCPT) and the questions were confusing (50% who took SPT; 27% who took TCPT). Therefore, TCPT elicits a greater number and breadth of requirements but can still be improved to elicit more PLR. However, the lower number of PLR by TCPT in the experiment compared to the significantly greater number of PLR in [4] may be due to PLR having fewer sub-categories, a different scenario and experiment setting, and perceived scenario vagueness and confusion. The last two can affect the subjects' ability to address PLR without a thorough base understanding of the scenario, since there wasn't the opportunity for clarification to address those concerns as with interviews.

Research Question 2: Whether LS's of individual subjects affect their ability to provide requirements during requirements elicitation?

Based on the results in Section 5.2 and Figure 7, clusters of subjects with SEN and GLO learning preferences were more effective at discovering a large number of requirements across all requirement categories. The SEN preference is made up of people who are detail-oriented, so they may be better at thinking of and describing requirements. On the other hand, INT learners don't like repetition, which was a common complaint about the PT questionnaires. The concise scenario left room for creativity with responses, but the dislike of details may have resulted in the better effectiveness by those with SEN. People with GLO learning preferences likely are effective, because their divergent thought process, in addition to the strategies used, help them come up with tacit knowledge. They are able to see the big picture with their way of thinking. Having room for creativity may have hindered SEQ learners, since they like to see the logic in what they're doing and proceed that way. However, SEQ learners are also good at analyzing from partial information.

Research Question 3: Which prompting technique (*Syntactic* - S vs. *Task Characteristic* - TC) is perceived more useful by subjects at eliciting requirements?

Based on the results in Section 5.2 and Table 6, while SPT was found to be simpler and easier to understand, as expected due to fewer prompts and lack of inbuilt checks to challenge conventional wisdom, the difference was not significant. Overall, participants rated both the techniques equally useful. The subjects did provide some qualitative feedback on improving the usefulness of PTs. Specifically, S and TC participants overall similarly thought the scenario and clarification of what the "system" in the questionnaires referred to was vague. These were the most common difficulties. Confusion over the questions and what they were asking for was another common complaint. More students who took SPT had problems with understanding questions than those who took TCPT, which is understandable since SPT questions are more generic. Repetition of questions was another common observation, but that was a minority of the students who didn't have their data excluded – 19% for SPT and 9% for TCPT.

Based on the results from this experiment, it can be concluded that TCPT can help requirement engineers, analysts, and other stakeholders gather information related to customer needs that are otherwise not considered when using standard PT, as shown in results, or using a standard questionnaire, which is established in literature review. Additionally, the results provided useful insights into the learning strengths of stakeholders that can be manipulated to help uncover more comprehensive set of requirements. While the results from this experiment were of interest and motivate further experimentation, we were able to identify some active research directions to improve the usage of TC prompting techniques in requirements engineering research and practice.

7. FUTURE WORK

In terms of future research investigation, improvements can be made to the questionnaires. The current format doesn't make it easy for stakeholders to include diagrams. Providing more space for them to create diagrams in a print form of the questionnaire could help, since people may not be experienced with creating those in Microsoft Word. Another option would be to provide the questions to stakeholders through a GUI like the one in [29] that includes tools in the interface that aid in their creation. Providing TCPT through a GUI would likely increase effectiveness and possibly efficiency when creating the diagrams. It would also be a useful improvement for distributed requirements elicitation. Participants also may not have experience giving requirements. Since there is a coding scheme, it isn't necessary for stakeholders to have experience giving requirements to answer the questionnaires. LS's may also have a greater impact on their effectiveness than experience [7]. Experience giving requirements can, however, make the coding of responses go faster. A possibility for further testing of the PT questionnaires with experienced stakeholders would be having a study with participants from industry or in a class that covers and practices requirements elicitation. A third improvement can be inclusion of more or better strategies to elicit PLR. Also, based on student opinions, more explanation on answering the questionnaires and clearly specifying the system in mind, or including it in place of "system," would help stakeholders using them.

With regard to stakeholders' LS's, there can be changes so TCPT can accommodate the different LS's better and be more balanced. Using a GUI to distribute the questions to stakeholders provides more of a visual component for VIS learners. Currently, the TCPT questionnaire and original interview version are both primarily VER-oriented, unless people provide diagrams. A GUI like that in [29] would provide a better VIS experience by having the graphical interface and making it easier to create diagrams. Finding where there can be a reduction in repetitive-seeming questions can be helpful to INT learners. Accommodating the different LS's can make TCPT better at effective requirements elicitation.

The coding scheme can also be improved. Browne and Rogich [4] suggested after their results that there could be more categories for the coding scheme. The new categorization could be more specific with less broad categories to notice more differences in the effectiveness of TCPT and in what areas.

Currently, the research of TCPT has mainly involved the quantitative analysis of how TCPT performs against other PTs and differences in the quantity of requirements. A method to determine the accuracy of requirements could be developed as another way of analyzing TCPT's performance. In order to do this, either more training for applying the coding scheme consistently by the analyst may be required, unless there's a way to do so with little or no training. Potential steps would be to: (1) elicit requirements from stakeholders for a given scenario, (2) code responses from the questionnaire into requirements, (3) return to the stakeholder a summary of requirements based off their questionnaire responses (participant determines accuracy) or the analyst doing the coding determines requirements completeness from participant-provided information (coder determines accuracy).

REFERENCES

[1] Agarwal, R. and Tanniru, M.R. “Knowledge Acquisition Using Structured Interviewing: An Empirical Investigation,” *J. Management Information Systems*, vol. 7, 1990, pp. 123-140.

[2] Aranda, G.N., Vizcaino, A., Cechich, A., and Piattini, M. “A Cognitive-Based Approach to Improve Distributed Requirements Elicitation Processes,” *4th IEEE Conf. Cognitive Informatics, 2005 (ICCI)*, 2005, pp. 322-330.

[3] Browne, G.J. and Ramesh, V. “Improving Information Requirements Determination: A Cognitive Perspective,” *Information & Management*, vol. 39(8), 2002, pp. 625-645.

[4] Browne, G.J. and Rogich, M.B. “An Empirical Investigation of User Requirements Elicitation: Comparing the Effectiveness of Prompting Techniques,” *J. Management Information Systems*, vol. 17(4), no. 4, 2001, pp. 223-249.

[5] Burton, A.M., Shadbolt, N.R., Hedgecock, A.P. and Rugg, G. “A Formal Evaluation of Knowledge Elicitation Techniques for Expert Systems: Domain 1.” *Proc. Expert Systems ’87 Research and Development in Expert Systems IV, The Ann. Technical Conference British CS Specialist Group on Expert Systems*, D.S. Moralee, ed., 1987, pp. 136-145.

[6] Burton, A.M., Shadbolt, N.R., Rugg, G., and Hedgecock, A.P. “The Efficacy of Knowledge Acquisition Techniques: A Comparison Across Domains and Levels of Expertise,” *Knowledge Acquisition*, vol. 2, 1990, pp. 167-178.

[7] Carver, J. “The Impact of Background and Experience on Software Inspections,” *Empirical Software Engineering*, vol. 9(3), 2004, pp. 259-262.

[8] Charkins, R., O'Toole, D.M., & Wetzel, J.N. (1985). “Linking Teacher and Student Learning Styles with Student Achievement and Attitudes,” *J. Economic Education*, vol. 16, no. 2, 1985, 111-120.

[9] Corbridge, C., Rugg, G., Major, N.P., Shadbolt, N.R., and Burton, A.M. "Laddering: Technique and Tool Use in Knowledge Acquisition," *Knowledge Acquisition*, vol. 6(3), 1994, pp. 315-341.

[10] Dieste, O. and Juristo, N. "Systematic Review and Aggregation of Empirical Studies on Elicitation Techniques." *IEEE Trans. Softw. Eng.*, vol. 37, March/April 2011, pp. 283-304.

[11] Eva, M. "Requirements Acquisition for Rapid Applications Development," *Information & Management*, vol. 39, December 2001, pp. 101-107.

[12] Felder, R.M. "Are Learning Styles Invalid? (Hint: No!)," *On-Course Newsletter*, 27 Sept 2010.

[13] Felder, R.M., & Henriques, E.R. "Learning and Teaching Styles in Foreign and Second Language Education," *Foreign Language Annals*, vol. 28, 1995, pp. 21-31.

[14] Felder, R.M., & Silverman, L.K. "Learning and Teaching Styles in Engineering Education," *Engineering Education*, vol. 78(7), 1988, pp. 674-681.

[15] Felder, R.M., & Spurlin, J.E. "Applications, Reliability and Validity of the Index of Learning Styles," *Int'l Journal of Engineering Education*, vol. 21, 2005, pp. 103-112.

[16] Fowlkes, J.E., Salas, E. and Baker, D.P, "The Utility of Event-Based Knowledge Elicitation," *Human Factors*, vol. 42, 2000, pp. 24-35.

[17] Friedman, P., & Alley, R. "Learning/Teaching Styles: Applying the Principles," *Theory into Practice*, vol. 23(1), 1984, pp. 77-81.

[18] Goodrich, V. and Olfman, L. "An Experimental Evaluation of Task and Methodology Variables for Requirements Definition Phase Success," *Proc. 23rd Ann. Hawaii Int'l Conf. System Sciences*, Kailua-Kona, HI, 1990, pp. 201-209.

[19] Goswami, A. and Walia, G. "An Empirical Study of the Effect of Learning Styles on the Faults Found During the Software Requirements Inspection," *2013 IEEE 24th Int'l Symp. Software Reliability Engineering (ISSRE)*, Pasadena, CA, 2013, pp. 330-339.

[20] Goswami, A., Walia, G., & Singh, A. "Using Learning Styles of Software Professionals to Improve their Inspection Team Performance," *27th Int'l Conf. Software Engineering and Knowledge Engineering (SEKE)*, Pittsburgh, PA, 2015, pp. 680-685.

[21] Goswami, A., & Walia, G.S. "Improving the Cost Effectiveness of Software Inspection Teams: An Empirical Investigation," *26th Int'l Conf. Software Engineering and Knowledge Engineering (SEKE)*, Vancouver, 2014, pp. 735-739.

[22] Griffin, A. and Hauser, J.R. "The Voice of the Customer," *Marketing Science*, vol. 12(1), 1993, pp. 1-27.

[23] Hudlicka, E. "Requirements Elicitation with Indirect Knowledge Elicitation: Comparison of Three Methods," *Proc. 2nd IEEE Int'l Conf. Requirements Engineering (ICRE)*, Colorado Springs, CO, 1996, pp. 4-11.

[24] Ishtiaq, M., Choudhry, F., Awan, F., and Khanum, A. "Analysis & Selection of Requirements Elicitation Techniques for OSSD," *Int'l Journal of Computer Science and Information Security*, vol. 10(7), 2012, pp. 12-22.

[25] Kolb, D.A. *Experiential Learning: Experience as the Source of Learning and Development (Vol. 1)*. Englewood Cliffs, NJ: Prentice-Hall, 1984.

[26] Kuri, N.P., & Truzzi, O.M.S. "Learning Styles of Freshmen Engineering Students," *Proc. 2002 Int'l Conf. Engineering Education (ICEE)*, Arlington, 2002.

[27] Lauer. T.W. “Development of a Generic Inquiry Based Problem Analysis Method.” Working paper, David D. Lattanze Center for Executive Studies in Information Systems, Loyola College, Baltimore, MD. 1994.

[28] Microsoft Corporation. Microsoft HoloLens. 2016. Published [Online]. Available: https://www.microsoft.com/microsoft-hololens/en-us. [Accessed: Nov. 2015]

[29] Moore, J.M. and Shipman, F.M. “A Comparison of Questionnaire-Based and GUI-Based Requirements,” *Proc. 15th IEEE Int’l Conf. Automated Software Engineering (ASE)*, Grenoble, 2000, pp. 35-43.

[30] Myers, I.B., McCaulley, M.H., and Most, R. *Manual: A Guide to the Development and Use of the Myers-Briggs Type Indicator*. Palo Alto, CA: Consulting Psychologists Press, 1985.

[31] Pitts, M.G. and Browne, G.J. “Stopping Behavior of Systems Analysts during Information Requirements Elicitation.” *J. Management Information Systems*, vol. 21, 2004, pp. 203-226.

[32] Rosati, P. “Specific Differences and Similarities in the learning Preferences of Engineering Students,” *29th Annu. Frontiers in Education Conf. (FIE)*, San Juan, Puerto Rico, 1999, pp. 17-22.

[33] Smith, N.G., Bridge, J., & Clarke, E. “An Evaluation of Students’ Performance Based on Their Preferred Learning Styles,” *Proc. 3rd Annu. Conf. UNESCO Int’l Centre for Engineering Education (UICEE)*, Melbourne, Australia, 2002.

APPENDIX A. PROMPTING TECHNIQUES

Table A1. Syntactic Prompting Technique (Borrowed from [4])

Users of the System
Who uses the system?
What kinds of things do they do?
Where do they do them?
Why do they do them?
Overview of System
Who is affected by the system?
What kinds of things affect them?
Where are they affected?
When are they affected?
Why are they affected?
How are they affected?
Who affects the system?
What kinds of things affect the system?
Where do they affect the system?
When do they affect the system?
Why do they affect the system?
How do they affect the system?

Table A2. Task Characteristics Prompting Technique (Adapted from [4])

Overview of System
What would customers want the system to do?
How will it perform better than other similar systems?
What are examples of how it does these better?
If it does these things better, how will that impact a customer's experience?
Why might the system not work better than other similar systems?
Why would your customers not want to use the system?
What can be done to overcome these negatives? (consider previous two answers)
What if these strategies for overcoming negatives don't work?
What would your company want the system to do?
Summarize everything the system should do. You may include use-case diagrams (or other notation).
How to Use the System
What does a customer need to do to use the system? • Consider including flow charts or decision maps • What influences each step?
Can you think of situations in which the customer would have problems using the system? • Consider previous answer and diagrams • Expand on diagrams where necessary
What can be done to overcome these problems? • Consider scenarios and previous flow charts
Summarize steps for using the system. You may include use-case diagrams.

Table A2. Task Characteristics Prompting Technique (Adapted from [4]) (continued)

Support
What people or departments are needed to support customer's use of the system?
Describe and detail the tasks these people or departments must do • Consider including a flow chart, use-case diagrams, and/or outlining influences for different individuals involved
What feedback should the system provide for performing these tasks? Use-case diagrams or feedback scenario examples can be included
During Use of System
Can you think of a situation in which the customer would have to make a decision or choice when using the system? • A decision map can be included
What does using the system allow people to do that they couldn't without the system?
What functions are people able to perform currently that they can't with the system?
Information Necessary
What information must a customer supply to the system to use it?
Are there any additional supplies that would aid the use of the system (e.g. movement tracking software, certain type of computer, etc.)?
What information must the system supply to the customer?
What information must the system display to the customer?

APPENDIX B. POST-STUDY SURVEY

New System Questionnaire Survey

NAME ___________________________

This survey is for the evaluation of the requirements elicitation questions you answered for your new system. Your responses on this survey will **not** affect your course grade.

1. Have you ever filled out a requirements questionnaire or had to list requirements information prior to the questionnaire for the new system?

   Please circle your answer: Yes No

Please indicate your *level of agreement* with each of the following statements by picking one of the five options for each question.

2. Rate the questionnaire with respect to the following attributes:

Attributes	Very Poor	Poor	Neither Good nor Poor	Good	Very Good
Simple	1	2	3	4	5
Easy to understand	1	2	3	4	5
Easy to use	1	2	3	4	5
Intuitive	1	2	3	4	5
Comprehensive	1	2	3	4	5
Useful	1	2	3	4	5
Overall opinion	1	2	3	4	5

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
3. The questionnaire helped me consider issues I hadn't thought of.	○	○	○	○	○
4. I had a better understanding of the system after filling out the questionnaire.	○	○	○	○	○
5. The questionnaire helped me communicated my vision for the system.	○	○	○	○	○
6. The effort spent on the questionnaire is valuable and worthwhile in outlining goals for creating the system.	○	○	○	○	○

What were the difficulties faced when using the questionnaire to document requirements?

____

____

____

What improvements or changes would you suggest for the questionnaire?

____

____

____

Are there any questions you would've wanted to add or take away?

____

____

____