Why Analysts Draw: The Effect of Diagrams Supporting UML Use Cases

Andrew Gemino, Simon Fraser University, 8888 University Drive, Burnaby, British Columbia, Canada V5A 1S6; E-mail: gemino@sfu.ca
Drew Parker, Simon Fraser University, 8888 University Drive, Burnaby, British Columbia, Canada V5A 1S6; E-mail: drew@sfu.ca

ABSTRACT

An experiment was undertaken to compare effectiveness of use cases with and without supporting use case diagrams. The Cognitive Theory of Multimedia Learning is used to hypothesize diagrams improve the effectiveness of use cases by providing visual cues aiding model viewers in selecting and integrating relevant information. The level of understanding developed by participants viewing either uses cases or use cases with a use case diagram was measured using comprehension, retention and problem solving tasks. Results support the hypothesis that participants developed a significantly higher level of understanding when viewing UML use cases with the support of a use case diagram. This suggests practitioners should consider combining a visual representation with use cases to achieve higher levels of understanding in persons viewing these descriptions.

Keywords: UML, Use Case Modeling, Conceptual Modeling, System Analysis

1. INTRODUCTION

Use cases in the unified modeling language (UML) are a popular modeling technique for system analysis and design (Burton Jones and Meso, 2006; Siau and Loo, 2006). The application of use cases by practitioners varies (Dobing and Parsons, 2006). However, use cases remain primarily text-based descriptions that provide a structured sequence of processes within a system (Jacobsen et. al., 1994). While text is a rich, familiar and expressive modeling language, text can also be ambiguous and difficult to conceptualize. It seems reasonable, given the popularity of use cases, to consider whether use case diagrams provide a significantly more effective method for communicating system analysis information than text-based use cases alone.

An experiment was undertaken to compare the effectiveness of use cases with and without supporting use case diagrams. The Cognitive Theory of Multimedia Learning (Mayer, 2001) is used to hypothesize that diagrams improve the effectiveness of use case delivery by providing visual cues aiding model viewers in selecting and integrating relevant domain information into effective cognitive representations. We take the view that techniques should be compared on how well they support the development of an understanding of the domain they represent (Gemino and Wand, 2003). To test understanding, we use a problem solving task (Bodart et. al., 2001; Gemino, 1999) that requires reasoning about the domain and focuses attention on higher levels of understanding.

2. BACKGROUND

Use Case Modeling

A Use Case is a description of a sequence of events in a system that produces a result as understood from a user's perspective. A use case presents the actions associated with a person's "use" of the system (Jacobson et. al., 1994). The use case is often an important part of object-oriented analysis methods (Dobing and Parsons, 2006; Siau and Cao, 2001).

Researchers that support the UML suggest the simplicity of the use case method is an asset. Kobryn (1999) suggests that use cases are simple and natural notations that are easy to understand for stakeholders, analysts and designers. While the use case is primarily text, the UML has developed an assortment of diagramming techniques that can potentially integrate with use case information. A key for the use case remains the lack of formalism (Jacobson et. al., 1999).

Previous UML Research

UML modeling has attracted significant research attention (Agarwal, 2003; Burton Jones and Meso, 2006; Evermann and Wand, 2005; Fedorowicz and Villeneuve, 1999; Siau and Cao, 2001; Siau and Loo, 2006). While much has been said about expected benefits of Use Case modeling (Jacobson et. al., 1999; Kobryn, 1999), little empirical research has been directed at these claims. Dobing and Parsons (2006), for example, found little empirical research on the effectiveness of Use Case modeling.

The UML approach is genrally accepted, however, it has critics. Douglass (1998) argued UML is large and overly complex. Halpin and Bloesch (1999) suggested UML models are designed for software engineering and are less suitable for validation of conceptual models. Dori (2003) has suggested that UML has difficulty in integrating structural and process elements of system designs. Dobing and Parsons (2006) suggested Use Case modeling faces two significant challenges. One challenge is that Use Cases tend to isolate stakeholders from object class models. This results in a lack of information on classifications and categories within the system. They argue that information in the Class Diagram is valuable in developing understanding and is not provided by Use Cases. A second challenge is the lack of formalism, which allows Use Cases to mix conceptual, design and implementation details in the same description. This mixture of design and conceptual elements may cause confusion for stakeholders and reduce the effectiveness of the stakeholder/analyst communication.

Separating Conceptual Modeling from Requirements Engineering

To understand how uses cases can be used, it is important to outline the system development process. The information system development process can be viewed as a series of increasingly formal representations ending in machine executable code (Wand and Weber, 1993).

Three generic roles in this development process include stakeholders, analysts and developers. The least formal representations of the system are the concepts held by stakeholders. In a standard development process, analysts interact with stakeholders to develop more formal representations of the system, often called conceptual models (Wand and Weber, 2002). These conceptual models are shown to stakeholders to validate analysts' perceptions of the system.

Conceptual models are a foundation for the development of formal functional requirements. Analysts develop functional requirements primarily to communicate system details with developers. In an ideal world, developers could use the formalized functional requirements to develop the eventual machine code for the systems (the system artefact).

Throughout the development process, the system representation grows increasingly formal and precise. The role of the analyst is to communicate system details in such a way as to develop a common understanding of the system between developers and stakeholders.

Analysts are involved in two distinct processes. The first involves interacting with stakeholders to develop an understanding of the system. This process is defined as conceptual modeling (CM) (Everman, 2005; Wand and Weber, 2002). It involves eliciting requirements, representing them, and having stakeholders interpret and validate these requirements. The second process formalizes this conceptual understanding into a set of functional requirements. This second process is defined as requirements engineering (RE). CM and RE are related processes that facilitate the common objective to reason and communicate about a domain. Because they are related, the same techniques are often touted for use in both

CM and RE, though the audience for the processes may be different. Use cases have often been suggested as useful tools for interacting with stakeholders, but use cases can also inform developers about process issues.

3. THEORETICAL FOUNDATIONS

Conceptual modeling involves the capture of domain information to develop understanding and support communication. Developing understanding of a system and its components is, therefore, a process of learning (Gemino and Wand, 2003). This is true for the person developing the model as well as the person viewing it. The design of conceptual modeling techniques may be informed, therefore, by theories of how humans develop understanding from the graphics and words they are presented with. Mayer (2001) suggests two contrasting views of learning –information acquisition and knowledge construction. These views are discussed briefly below.

Information Acquisition

Learning as information acquisition implies that learning is a process of adding to long-term memory. The model viewer receives information and stores it in memory. The responsibility for learning rests on the model creator to deliver appropriate information. The goal is to deliver required information efficiently. In this view, the conceptual model is a standard vehicle for efficient information delivery to the model viewer.

Knowledge Construction

An alternative view is knowledge construction. This view suggests knowledge is personally constructed. Two model viewers presented with the same conceptual model may come away with different learned outcomes. This occurs because the model viewers attempt to make sense of the information presented and integrate this information into a coherent mental representation. Knowledge construction suggests the model viewer is an active sense maker rather than a passive receiver of information.

A Model of Conceptual Modeling as Knowledge Construction

As a framework for reasoning about conceptual modeling, we use the model of knowledge construction (Gemino and Wand, 2003; 2005). In this model, the model viewer is constructing knowledge by actively organizing and integrating information with previous experiences. Three antecedents of the process are suggested: (1) content, (2) presentation method, and (3) model viewer characteristics. The content represents the domain information to be communicated. The presentation method is the way in which content is presented to the viewer. Viewer characteristics are attributes prior to viewing the content. These characteristics include knowledge and experience with the domain and with the modeling methods used to present information.

The construction process is where the sense making activity is hypothesized to occur. The results of knowledge construction are encoded into the long-term memory. The learning outcome modifies the model viewer's characteristics. Learning outcomes can then be observed, only indirectly, through learning performance tasks.

Cognitive Theory of Multimedia Learning

Messages that combine graphics and words are defined by Mayer (2001) as "multimedia messages". The Cognitive Theory of Multimedia Learning (CTML) provides a theoretical perspective on the level of understanding developed by a person viewing explanative material, such as an analysis diagram in requirements validation. The theory is based on work by Baddeley (1992) and Paivio (1986) and has been developed through a decade of empirical work (Mayer, 1989; Mayer, 2001).

The theory is focused on the interaction between a person and the information presented to him or her. The CTML suggests there are two pathways in cognition, verbal and visual. While independent, these channels communicate in working memory. When a person views presented material, relevant information from the verbal and visual channels is selected. This information is organized to create separate visual and verbal models. These two models then interact and are subsequently integrated with prior knowledge in long-term memory to develop new knowledge. An overview is provided in Figure 1.

In the CTML, an understanding of verbal and visual information is developed through three stages of memory. In the first stage, *sensory memory*, information is selected into one of the two dual coding pathways. The selected information is then developed into visual and verbal models in *working memory*. In the final stage, the verbal and visual models from working memory are integrated with *long-term memory* to create understanding. This describes the process of knowledge construction.

Learning Outcomes and Performance

The CTML has enabled Mayer to develop principles relating to the effective design of multimedia messages. He suggests the most effective communication occurs when verbal and visual pathways are utilized simultaneously. Mayer suggests three outcomes when presenting explanative material to people: 1) no learning, 2) rote learning and 3) meaningful learning. These outcomes are based on measures of two variables: retention and problem solving. Retention is the comprehension of material being presented. Problem solving is the ability to use knowledge gained to answer related problems not directly answerable from presented material. For example, if presented with an explanation of how a car's brake system works, a retention question might be "*List the components of a braking system*," but a problem solving question would be "*What could be done to make brakes more reliable?*" These problem solving task have been used by Bodart et. al (2001); Burton-Jones and Meso (2006) and Gemino (1999; 2004).

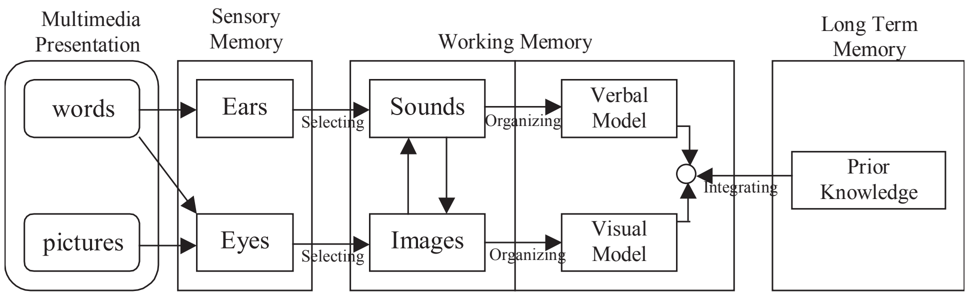
No learning occurs were retention and problem solving are low. Rote learning occurs where retention is high; however, problem solving measures are low. This indicates that although the material has been received, material has not been well integrated with prior knowledge. Meaningful learning occurs where retention and problem solving are high.

4. EXPERIMENTAL DESIGN AND HYPOTHESIS

Overview of Experimental Design

Three dependent variables (comprehension, retention and problem solving) will be measured. Other variables measured include prior knowledge of the domain,

Figure 1. The cognitive theory of multimedia learning (adapted from Mayer, 2001, p. 59)



knowledge of the modeling method, and participant demographics. Two treatment groups were compared using a single case. The first treatment was provided with a set of use cases describing a system. The second treatment was provided the same set of use cases along with a one-page use case diagram. The single page diagram shows the interaction between use cases and actors in the system as well as any interaction between use cases in the system. The following statement provides the underlying logic for conducting this experiment: If a participant is presented with a) a set of use cases and b) a set of use cases and a diagram relating these use cases, then the participant will gather a significantly higher level of understanding of the domain being presented with b) than with a).

Hypothesis

Mayer's (2001) multimedia principle suggests a potential for higher levels of understanding from use cases associated with diagram than use cases alone. The multimedia principle therefore enables us to suggest the following hypothesis:

H1: *According to the multimedia principle from the CTML, participants viewing a set of use cases with an associated use case diagram will develop a higher level of understanding of the domain than participants viewing use cases alone.*

5. METHOD

An empirical procedure was developed to test the hypothesis above. The procedure was based on Mayer (1989, 2001), and used for system analysis by Bodart et. al (2001), Burton-Jones and Meso (2006), Gemino (1999, 2004), Gemino and Wand (2005).

Participants

Forty-nine upper level business students took part in the study. All students had taken a system analysis course and had basic familiarity with use case models. Females accounted for 20 of the 49 participants (41%) of participants. Participation was voluntary. An incentive of $15 was provided for the top four performers. The average time to complete the study was 45 minutes. All participants were at an introductory level in business process design, and had no particular experience with object oriented analysis. A pre-test was given to measure experience with system analysis and the business domain used in the analysis as well as other demographic variables.

Materials

One case including five use cases and one use case diagram was used in the experiment[1]. The use cases and use case diagram were created using an approach described in Dennis and Wixom (2000). The text description was provided by the Voyager Bus company case in Bodart et. al. (2001).

Procedure

Participants were randomly assigned into two treatment groups. An envelope was given to each participant containing a pre-test, five use cases (plus diagrams if necessary), experimental tasks (comprehension, retention and problem solving) and a posttest. Participants worked independently and first completed the pre-test followed by the three experimental tasks and finally the post-test.

The first task was a twelve question multiple choice comprehension task (True, False, Uncertain). After the comprehension task, participants were instructed to put away the use cases and diagram (if provided). Participants were than given 6 minutes to complete a retention task, which asked participants to write down everything they knew about the processes in the use cases. This task was followed by four problem solving questions used by Bodart et. al. (2001). Participants were given 2 minutes to write as many answers as possible to each problem solving question.

Measures

Learning performance was measured using three variables: comprehension, retention and problem solving. Comprehension was the number of correct answers out of a possible of 12 questions. Retention and problem solving scores were coded by two individuals. The retention score was created by giving one mark for each complete and correct idea statement expressed by the participant. There was a maximum of 20 idea statements identified in the use cases. The problem solving score was created by giving participants one point for each acceptable response to the problem solving questions. The Pearson correlation between coders for retention was 0.88 and for problem solving questions 0.90. Differences between independent ratings were then discussed, and a final score for retention and problem solving was established.

6. RESULTS

Preliminary Tests

Since the sample size was relatively small, it is important to establish the homogeneity of variances before ANOVA analysis. Levene statistics for each of the dependent measures indicate that the hypothesis of equal variances is not rejected across any of the variables at the 0.05 level.

Domain and modeling experience were collected in the pre-test and used as covariates in an ANCOVA analyses. Both domain and modeling method experience were found to have insignificant influences on the dependent variables. This result may be due to the uniformly low levels of experience held by participants. While it seems likely prior domain experience and modeling method experience are related to the dependent measures, the factors, as measured in this study, had no significant effect in this study and were excluded in further analysis.

Results

The means and standard deviations of the dependent measures (comprehension, retention and problem solving) across the two treatment groups are provided in

Table 1. Means and Std. Dev. across treatments for dependent measures

Dependent Measure	**Case: Voyager Bus**			
	Treatment Groups			
	Without Diagram n=25 Means (SD)	**With Diagram n=24 Means (SD)**	**Difference between means (With-Without)**	**Sig.**
Comprehension	7.627 (.321)	8.139 (.328)	.0512	0.271
Retention	7.877 (.541)	9.670 (.552)	1.793	0.025*
Problem solving	12.174 (.824)	14.568 (.841)	2.394	0.045*

* *significant at the 0.05 level*

Table 1 below. The results show little difference across treatment groups for comprehension measures. Note that participants had full access to use cases during the comprehension test. Since the information was available in either treatment, the diagram had little effect in basic comprehension.

Retention and problem solving measures showed differences in the anticipated direction. The size of the effects was approximately 20%. This is measured by dividing the difference between the "with" and "without" diagram scores and then dividing the result by the score for the without diagram group.

An ANOVA was applied to test the significance of these differences. Results, provided in the final column of Table 1, suggests significant differences for both problem solving and retention measures at $\alpha = 0.05$ level. This result provides evidence to support hypothesis H_1. These results suggest that although the content across treatments was the same, the organization provided by the use case diagram enabled participants with access to the diagram to build a more sophisticated mental model. Note that although the sample size is relatively small, the effect size is relatively large.

These results suggest that diagrams, even simple diagrams such as the use case diagram provided in this experiment, have measurable effects on viewer understanding.

7. CONCLUSIONS

Mayer's Cognitive Theory of Multimedia Learning has been used to demonstrate that humans process both text and pictorial images together to develop a deeper understanding than either media alone. Text descriptions often accompany drawings, but little formal research has been done to clarify the potential effect of exhibits. This paper describes an experiment which evaluates the level of domain understanding for subjects when drawings are used to support use cases compared to the use cases alone. The results support the hypothesis that a use case diagram has a significant positive effect of the level of understanding developed by a person viewing use cases.

Future research can be directed more closely on what diagram elements are most effective in supporting use cases. In addition, more empirical evidence is required to understand the effectiveness of use case modeling. While the text based approach has some excellent features and has appealed to practitioners, it is clear that diagrams are an important component for communication. More needs to be understood about this relationship if we are to make use case modeling an even more effective communication tool for stakeholders and developers.

ACKNOWLEDGMENT

The authors wish to thank Liliana Petrescu for her work in collecting and coding experimental data. This research was supported by the National Science and Engineering Research Council of Canada.

REFERENCES

Agarwal, R., P. De, and A.P. Sinha, "Object Oriented Modeling with UML: A Study of Developer Perceptions", Communications of the ACM, Vol. 46, No. 9ve, September 203, p. 248-256.

Baddeley, A. D., "Working Memory", Science, 1992, vol. 255, pp. 556-559.

Bodart, F., A. Patel, M. Sim, and R. Weber, "Should Optional Properties Be Used in Conceptual Modelling? A Theory and Three Empirical Tests," Information Systems Research, Vol. 12, No. 4, (December 2001) 384-405.

Burton-Jones, Andrew and Meso, Peter, "Conceptualizing Systems for Understanding: An Empirical Test of Decomposition Principles in Object-Oriented Analysis," Information Systems Research, Mar 2006, Vol. 17 Issue 1, pp. 38-60.

Dennis, A., and Wixom, B., System Analysis and Design: An Applied Approach, John Wiley and Sons, New York, 2000.

Dobing, B. and Parsons, J. " How UML is Used", Communications of the ACM, Vol. 49, No. 2, May 2006, p. 109-113.

Dori, D. "Why Significant UML Change is Unlikely", Communications of the ACM, Vol. 45. No. 11, November, 2002, p. 82-85.

Douglass, B. P. "UML for Systems Engineering" Computer Design's: Electronic Systems Technology & Design, (37:11), 1998. p. 44-49.

Evermann, J. Thinking Ontologically - Conceptual versus Design Models in UML. In: Rosemann, M. and Green, P. (eds.) Ontologies and Business Analysis. Idea Group Publishing, 2005, 82-104.

Evermann, J. and Wand, Y, "Ontology-Based Object-Oriented Business Modelling: Fundamental Concepts, Requirements Engineering Journal, Vol 10, No 2, 2005, 146-160.

Fedorowicz, J. and Villeneuve, A. 0, "Surveying Object Technology Usage and Benefits: A Test of Conventional Wisdom", Information & Management, (35:6), 1999. p. 331-345.

Gemino, A., Empirical Comparison of System Analysis Techniques, Ph.D. Thesis, University of British Columbia, Vancouver, British Columbia, June 1999.

Gemino, A. Empirical Comparisons of Animation and Narration in Requirements Validation, Requirements Engineering Journal 9 (3) (2004) 153-168.

Gemino, A. and Y. Wand, Evaluating Modeling Techniques Based on Models of Learning, Communications of the ACM 46 (10) (2003) 79-84.

Gemino, A. and Y. Wand, "Dimensions in Experimental Evaluation of Conceptual Modeling Techniques", Requirements Engineering Journal, vol. 9, 4, 2004, 248-260.

Gemino, A. and Wand, Y., "Simplicity versus Clarity: An Empirical Comparison of Mandatory and Optional Properties in Conceptual Modeling", Data and Knowledge Engineering, 55, 2005, p. 301-326.

Halpin, T. and A. Bloesch, "Data modeling in UML and ORM: A Comparison", Journal of Database Management, (10:4), 1999, pp. 4-13.

Jacobson, I., Ericsson, M., & Jacobson, A, The Object Advantage: Business Process Reengineering with Object Technology. Reading, MA: Addison-Wesley, 1994.

Jacobson, I., Booch, G. and Rumbaugh, J., The Unified Software Development Process, Reading, MA, Addison Wesley, 1999.

Johnson, R. A., "Object-Oriented System Development: A Review of Empirical Research", Communications of the Association for Information Systems, Volume 8, 2002, p. 65-81.

Kobryn, C, "A Standardization Odyssey", Communications of the ACM, (42:10), 1999 p. 29-38.

Mayer, Richard E., "Models for Understanding", Review of Educational Research, Spring 1989, Vol. 59, No. 1, pp 43-64.

Mayer, Richard E. and J.K. Gallini, "When is an Illustration Worth Ten Thousand Words?" Journal of Educational Psychology, 1990, Vol. 82, no.4, p. 715-726.

Mayer, Richard E., Multimedia Learning, Cambridge University Press, New York, 2001.

Mellor, Stephen J, "Make models be assets" Association for Computing Machinery. Comm. of the ACM; Vol. 45, No. 11, Nov 2002, p. 76-78.

Paivio, A. Mental representations: A Dual Coding Approach, Oxford, England: Oxford University Press, 1986.

Siau, K. and Q. Cao, "Unified Modeling Language (UML) - a complexity analysis", Journal of Database Management; Vol. 12, No. 1, Jan-Mar 2001, p. 26-34.

Siau, K. and Loo, PP, "Identifying Difficulties in Learning UML", Information and Management, (23,3), Summer 2006, p. 43-51.

Wand Y., and R. Weber, "On the Ontological Expressiveness of Information Systems Analysis and Design Grammars", Journal of Information Systems, vol. 3, 1993, pp. 217-237.

Wand, Y. and R. Weber, "Information Systems and Conceptual Modeling a Research Agenda," Information Systems Research, December, 2002, pp. 203-223.

ENDNOTE

1 The exhibits used can be accessed online at http://parker.bus.sfu.ca/irma.pdf

Related Content

Information Dissemination Mechanism Based on Cloud Computing Cross-Media Public Opinion Network Environment
Ping Liu (2021). *International Journal of Information Technologies and Systems Approach (pp. 70-83).*
www.irma-international.org/article/information-dissemination-mechanism-based-on-cloud-computing-cross-media-public-opinion-network-environment/278711

Climate Change as a Driving Force on Urban Energy Consumption Patterns
Mostafa Jafariand Pete Smith (2018). *Encyclopedia of Information Science and Technology, Fourth Edition (pp. 7815-7830).*
www.irma-international.org/chapter/climate-change-as-a-driving-force-on-urban-energy-consumption-patterns/184478

A Machine Translation System from Indian Sign Language to English Text
Kinjal Mistree, Devendra Thakorand Brijesh Bhatt (2022). *International Journal of Information Technologies and Systems Approach (pp. 1-23).*
www.irma-international.org/article/a-machine-translation-system-from-indian-sign-language-to-english-text/313419

Viewpoints on Business Process Models
Giorgio Bruno (2018). *Encyclopedia of Information Science and Technology, Fourth Edition (pp. 788-798).*
www.irma-international.org/chapter/viewpoints-on-business-process-models/183790

A Systematic Review on Prediction Techniques for Cardiac Disease
Savita Wadhawanand Raman Maini (2022). *International Journal of Information Technologies and Systems Approach (pp. 1-33).*
www.irma-international.org/article/a-systematic-review-on-prediction-techniques-for-cardiac-disease/290001