A multi-level model for critical discussion of interaction design objects

Anthony Carton, South Dakota State University, United States

In this paper, I present a preliminary model for better understanding and explicit discussion of the goals, visual representation, and implementation of an interactive system to aid in the design and critique of design objects that lie increasingly towards the center of a continuum between digital and tangible. I then outline a recent pilot study involving university design students and present preliminary results before proposing future development of the model.

As the boundaries between design fields continue to blur, professionals and students alike need to strengthen their ability to critically discuss interaction design with the same confidence and skill at which they are able to critique the formal, rhetorical, and semiotic qualities of a designed object. To facilitate this discussion of the shaping of the behavior of an interactive system, I propose a model based on David Marr's Three-Level Model for analysis of information processing systems in the brain, specifically used to theorize about the visual processing system at the levels of computation, representation, and implementation.

In this pilot study, students have been asked to 1) critique an interactive system with comments noted for reference and comparison. 2) I then introduce the adapted multi-level model for analyzing an interactive system at each of the three levels of computation, representation, and implementation. Finally, students are asked to 3) critique a similar interactive system using this model to focus and guide their discussion. Further development of the adapted model is then recommended based on student responses.

Keywords: Interaction Design, Critique

Introduction

The purpose of this work is to present an adapted model for critical discussion of interaction design based on David Marr's *Three-Level Model* for understanding complex information processing systems in the brain. (Dawson, 1998) Results of a pilot study of student use of an adapted multi-level model to facilitate critical discussion of the behavior of interactive systems are presented. Twenty-two (22) sophomore graphic design students were first asked to discuss the interaction design of a dance-activated music video in an open response format. Students responded on the appropriateness of the content for the interaction, positive and negative responses to the experience, lack of discoverability, novelty of the interaction and attempted to describe their mental model of the workings of the system.

The same students were then introduced to an adapted version of Marr's *Three-Level Model* and asked to use it as a framework for discussion of the interaction design of the system, with interaction design defined as the shaping of the behavior of the system. (Löwgren & Stolterman, 2004) This time students responded by discussing the goals of the system, and the representation media of the input and the output. Some students communicated confusion with the model, especially in the representation and implementation levels. Going forward, further work is required to adapt the model for use by undergraduate design students. 1) The model will need to be easier to understand by students through short explanation, 2) the representation level may provide an outlet for student discussion of visual design aspects as part of a larger context of interaction, and 3) the role of the implementation level requires further examination in its role in the adapted multi-level model.

Marr's Three-Level Model

David Marr, a cognitive psychologist, developed the *Three-Level Model* to describe the processing of visual information in our brains by the visual processing system in response to what he saw as a deficiency in the way that systems were described. Marr felt that heuristic or experience-based descriptions of a task were being incorrectly accepted at theory, and that clear distinctions between what a system did and how it did it were necessary to theorize about a system. Through description at each of three levels, the computational level, the representational and algorithmic level and the hardware implementation level, the *Three-Level Model* describes

what the system does and how it does it. (Dawson, 1998) While intended for understanding complex information-processing systems in the brain, I immediately thought it might be helpful for students to better understand the interaction they were struggling to discuss during critiques, with a bit of adaptation.

Marr's computational level is the most abstract, addressing the system's goals, This level is about theorizing what the system does, what is being processed, and why. For explanation, Marr presents an analysis of a cash register. At the computational level, the goal of a cash register is arithmetic, and it does this because addition is the appropriate method for arriving at a total cost of items purchased.

Marr's representation and algorithmic level addresses *how* a system accomplishes its goals. For the cash register, we describe how arithmetic is accomplished. This requires the making of two choices. First a representation must be chosen for the input and output of the process. For the cash register, this means choosing a system of numbers, Arabic numerals in the US for example, to represent the input. It also means choosing an algorithm by which the transformation may actually be accomplished, so choosing a method for adding. The cash register accomplishes its arithmetic by adding least significant digits first, and carrying a value to the next significant digit if the sum exceeds nine.

At the hardware implementation level these goals, representations and algorithms are implemented physically. It is important to note that, according to Marr, the same algorithms may be implemented in different ways when different technologies are used. A digital cash register would implement the algorithm, the method of addition, with a configuration of wires and resistors, while a mechanical cash register would do so with a combination of gearing. (Dawson, 1998)

Adapted Multi-Level Model

This method of multi-level description was developed as a way to better understand visual information processing in the brain, but Marr's choice of the cash register example to explain the model is interesting for us as a good example of an interactive system. We supply input to the cash register, it performs some algorithm for calculation depending on the transaction, and then we are given output of the total due for payment. This is a strong interaction; we do not receive the same input back in a slightly different modality or format, and we aren't simply taken to a different place, but instead we are given new information as a result of the interaction with the system. As such, Marr's levels of explanation, especially the computational and the representational and algorithmic levels might provide a thorough means to discuss the interactivity of a system alongside the graphic qualities.

Computational Level

Adapted for critical discussion of interaction design, the Computational Level deals with system goals. Students are asked to talk about the overall goal of the site. Questions like, "What does the site do? What does it allow them to do? How does it respond to their interactions?" help students to formulate an answer at this level. In the second introduction video, students were given a description of the goals of the previous video they were asked to watch, relating it to their previous discussion.

Representation Level

At the Representation Level, students were given a short explanation of the concept of representation and then asked to discuss the input and output being represented in the Bear71 documentary. They are also give examples both from Marr's cash register example and from the previously viewed sprawl2 video.

Implementation Level

At the implementation level students were asked to discuss the medium rather than the goals or the representations that are moving in our out of the system. For instance, in the sprawl2 video, the medium of the input was the webcam footage of their bodies, or in many cases their hands. The video reassured students that advanced knowledge of computer vision was not required but that a computer is able to recognize their movement and adjust video playback and rhythmic skipping accordingly.

Table 1.1: Adapted Multi-level Model and Student Discussion Questions

Computational Level	*Representation Level*	*Implementation Level*
What is the overall goal of the system?	What concept is being represented for the input?	What is the medium for the input?
How does the input help accomplish the goal?	What concept is being represented for the output?	What is the medium for the output?
What strategy might the system use to achieve the goal?	How is the website converting these concepts from input to output?	What transformation has to happen that connects these media together?

Method

In this preliminary study, twenty-two (22) sophomore graphic design students in an introductory web design course at a university in the Midwestern United States were divided into pairs and asked to discuss the interaction design of two systems, an interactive dance video and an interactive documentary. While multiple definitions and descriptions of interaction design have been formulated as the field continues to develop, for this study interaction design was defined for students as the shaping of the behavioral aspects of the system. (Löwgren & Stolterman, 2004) Students were told that these elements complement each other but that they both require adequate means for critique. For many of these students, this was the first time interaction design as an area of design has been introduced. For their discussion of the second system, students were presented with the adapted multi-level model, an adapted version of Marr's *Three-Level Model* and asked to use it as a means to discuss the interaction design of an interactive system.

Of the twelve student pairs, one group failed to submit the second critical discussion, while two other groups failed to view all of the introductory materials before completing the exercise. As a result these three student pairs were excluded from consideration. The two student groups that did not watch the introductory video for the second exercise were confused, frustrated and visibly upset in their second videos. Further work developing materials that can support a single mode of instruction if students are to be presented with just a written project description may be required.

Stimuli / Interactive Systems

Part 1: Pretest

Sprawl 2, a dance-activated music video
By AATOAA studio
www.sprawl2.com/

Sprawl 2 functions on several levels, it's a quality music video without the interactivity, a non-interactive version is also available online. The video takes in both webcam footage and mouse interaction, and then transforms that interaction into directional information that is then used to control the speed and direction of the playback of the music video to encourage fans to dance along with the video. The band Arcade Fire is known for their experimental videos and online experiences.

Part 2: Post Test

Bear71, an interactive documentary
By Jeremy Mendes and Leanne Allison with interactive design,
development and programming by Jam3
http://bear71.nfb.ca/

Bear71 is the story of a bear that is tagged and tracked throughout her life in Banff national forest in Canada. The story discusses the difficulty for survival of the grizzly bears native to the wilderness there, and how the encroachment of humans has had multiple negative effects. In addition to the interaction with the map of the national park, the system takes the input of your webcam feed, and broadcasts it to any other users viewing the interactive video at that time. As a result the user is added to the story in real time as a visitor to the national park. When viewing the interactive film, students reported encountering other students interacting with the system at the same time. This capture of user video data is added to a matrix of video feeds, some prerecorded of animals in the park, some live feeds of other users, allowing you to encounter other tourists like yourself in the bear's habitat.

Results

Part 1: Pretest Results

In their response, students almost universally commented on the novelty of the interactive dance video format in one way or another with most students commenting on the ability to move in front of the webcam and to control the speed of playback.

Students in four groups (G1, G3, G9, G12) responded on the appropriateness of the content with comments like this one from group 12, "the music worked well with the glitchy site because it was choppy and robotic itself." Students familiar with the band also said that this type of video is typical for their style and works well for that band.

Two groups (G1, G3) viewed the interaction negatively, saying that the interactive dance video was annoying, irritating or boring. Sitting in front of a computer and having to wave at the webcam in order to keep the video feed going is a chore, and that the novelty wore off quickly. One student notes, "You would have to be really into the music to want to keep doing it."

Students experienced confused discoverability, at times not knowing what to do. Three groups (G2, G3, G12) were confused by the dual input modes or were not sure their actions had any effect on playback. Group 2 comments that the tutorial screen works differently than the video. Groups 3 and 12 both mentioned a steep learning curve.

Though no students used the term "mental model," many of them attempted to explain their own thoughts on how the interactive system worked. Within these responses, several themes emerged.

Only one group (G8) seemed to possess an appropriately matched mental model of the interaction, noting that movement from dance activates the music video, and recognized the directional nature of hand movement and clicking, as well as speed being variable as well.

Five groups shared this matched mental model of the interaction, but incorrectly believed that interaction was limited to hand interaction with the webcam. Groups 8 and 9 suggest that the

webcam is searching for and tracking hands as a way to move the video forward, one member of group 9 suggested a limitation of the webcam only being able to track one or two hands but not four.

Two groups (G2, G7) possessed an incorrect mental model of the interaction. Group 2 believed that faster hand movements resulted in louder music, while group 7 believed that the site was matching their gestures with images that contained similar movements and gestures.

One interesting note, group 1 comments on the social aspects of participating in an interactive dance video, stating that unknowing people will like not understand what is going on, and respond negatively to the interaction. One student tells her partner that if he weren't there she would probably be dancing, suggesting that she knows it isn't just about waving her hands across the screen, but instead that the goal is to dance with the video.

Two groups (G5, G6) also discussed the graphic design of the site even though all groups were asked to focus on the interaction design. Group 5 perceived the site as being very well laid out and neat, while a student from group 6 commented sarcastically about keeping things hip and cool, but that the function of the system was not understandable as a result, communicating confusion and frustration.

Part 2: Post Test Results

In part two, students were asked to answer these questions explicitly. Five of nine student groups (G2, G3, G6, G8, G12) went through the adapted multi-level model level by level using the associated questions introduced in the introductory video and project description. Four student groups (G1, G5, G7, G9) chose to use the levels only as a loose model, while no student groups ignored the levels all together.

At the Computational Level all groups identified the overall goal of the system being to create empathy with the animals in Banff National Forest. Some groups mentioned emotional appeals but all groups identified this as the main goal.

The groups also all mentioned their movement as being the input that accomplished the goal of creating empathy between then and the animals. Group 12 adds that the narrator's female voice was the input that accomplished the goal. Though incorrect, this representation of the content does likely aid in the accomplishment of the goal. Group 2 presents an especially good response, "our movement of the mouse is what is driving our movement through the landscape, allowing us to be able to act as if we are the bear, or with the bear along their life journey."

At the Representation Level, the input to the system was the student's mouse movements, their arrow key input, and the web cam footage that we being recorded. As a result the concept represented in the input would be their own exploration of the map, and their participation in the storyline. Group 2 was likely the closest, but this topic seemed to be almost universally difficulty for the groups. For instance, while group 2 suggested that the concept represented in the input was the navigation through the map, groups 6 and 12 suggested it was the concept of the bear's story, represented through the narrator's voice.

Similarly, the output of the system is the visual combination of the student's movement with a pre-recorded set of movements and video playback, thus the concept represented in the output is the story of bear71 *including users as in a dual role as active participants and as the bear itself resulting in a mixture of empathy and co-experience.* Students in group 2 supposed that a sense of the landscape was the concept. Groups 6 and 9 felt as though the output was blended with the input, while group 12 suggested that the concept represented in the output was the bear's life in general.

At the Implementation Level, the input and output for the system Bear71 are mouse interaction, arrow button interaction and webcam footage, and each of these have a medium. The medium of mouse movement is the coordinate system upon which the spatial data is represented,

similarly the arrow button interaction and webcam footage. The media of the output is the audio and video playback of the system.

Students in four groups (G2, G3, G8, G9) all mentioned the video feed and/or mouse movement, while two groups (G9, G12) mentioned the interactive map itself. Students in three groups (G3, G8, G12) all identify the medium of output as the video playback, while the other student groups did have difficulty. One group (G9) confuses goals with output medium, saying that the output is our caring more about the animals.

When asked about the transformation that needs to happen to connect the input medium with the output medium all groups are able to formulate an answer. Group 2 discusses mouse input, while four groups (G3, G6, G8, G12) all discuss webcam input and video playback.

In addition to the three levels, students also discussed the topics of interaction as a means for engaging the viewer, additional features, confusion due to poor discoverability, and confusion in relation to the use of the adapted multi-level model.

Four groups (G1, G5, G6, G12) all discussed interaction as a means for engaging the viewer. Group 1 stated that the interaction facilitated multi-tasking tendencies for their generation, affording them the ability to watch different feeds and jump back to the narrative at will. Four groups (G3, G5, G6, G12) expressed sadness or other emotional experiences, and group 6 found this to be in line with the goals of a documentary, but more impactful as a result of the interaction. A student from group 5 claims to have gasped when the bear dies.

Groups 2 and 8 experienced confusion, and seemed to have trouble discovering how to use the system, and were unsure if their actions had an affect.

Discussion

In Part 2, students seem to have grasped the goals and input of the system, however the question of strategy to achieve the goal with the available input was much more difficult for the student pairs to answer. An appropriate answer would have been something along the lines of, “by affording the user the opportunity to explore the forest from the point of view of the bear, while also occasionally being shown their own likeness as one of the objects of surveillance of the system, users can develop a sense of empathy of the animals in the habitat.” Most students responded with first a long pause, and then struggled to communicate an answer. Group 8 guessed that the system could track where their mouse was moving across the map, or that the map always shows you where the bear is.

At the representation level, the groups also had difficulty differentiating between the representation of the input and that of the output. Interestingly, the groups that chose to discuss graphic design, or the visual aspects of the system in the computational level did a good job describing the way concepts were represented, but they did so during their discussion of the computational level. In the future it might make sense to more explicitly encourage students to discuss the visual elements or graphic design of the system at this representational level, as it likely would give them increased confidence and a better understanding of the purpose of the representation level.

Interestingly, during their broad discussion of the representation level, group 5 identifies the concept of surveillance as being represented through the video instead of any of the other concepts discussed by other groups. Though they do not use the provided questions to formulate their responses, theirs is among the best formed of the student pairs. Perhaps as students learn to integrate these questions into an understanding of each level, the answers may become more holistically formed.

For the majority of students, the implementation level was a difficult one. In Marr’s *Three-Level Model*, the hardware implementation level theorizes about the hardware upon which the representations and algorithms run, however, this may not be relevant to an undergraduate student in design. As a result, in the adapted multi-level model, hardware implementation is adjusted to deal with the medium or mode of the input and output, as a way to better allow students to understand the requirements of an interactive system that employs more than mouse

interaction. This level looks to give students an opportunity to think about the different modes of input and output that are becoming possible. More development is required on this level, as evidenced by several group's total inability to address this level.

Students seem to develop a better grasp of the input and output when asked to describe the transformation that must occur to the input to create the output. Group 8 responded, "they take your feed and they show it off to others so you can see your feed with other people's feed." This group hinted at video as the medium of the input and the output, and the statement shows a better understanding of what is happening to the video as input, and what is going in the system to create the output.

Three groups exhibited difficulty with the adapted multi-level model in addition to the two who did not watch the interactive videos. Group 3 seemed to have a good deal of difficulty with the model. When they came to the question about the strategy of the system, they respond, "I don't even know what that means." The course management system was checked to ensure that at least one member of the team viewed the introductory videos, as they were instructed to work as a team. Group 3 did view the video, but has a difficult time with the model. Group 9 states that the notes included in the assignment description are very confusing and that accept for the video they would have been lost.

Group 7 spent a good deal of time describing the formal elements of the map, talking about the representations in place. They then move on to the representational level where they begin talking about their interaction with the environment and the reasons they do so. Seemingly, a total failure of the adapted multi-level model to work for these two students, they almost completely switch the two levels. This further supports the notion that a representation level needs to provide students an outlet to talk about the graphic design of the interactive system.

Conclusion

Overall, the adapted multi-level model is a positive start as the majority of students were able to use it, formulating a response to many of the associated questions. The responses for the goals of the system seem especially useful for the students in their ability to critique the interactive system.

Going forward, further development of written materials for the explanation of the 3 levels is required as students apply the discussion of an interactive system for critical discussion. In addition, the value of the third level, the implementation level likely needs further investigation. Does the difficulty and confusion that some of the students experienced warrant the opportunity to discuss extra modalities of interaction?

REFERENCES

Dawson, M. (1998). *Understanding Cognitive Science*. Blackwell Publishing.

Löwgren, J., & Stolterman, E. (2004). *Thoughtful Interaction Design: A design perspective on information technology.* MIT Press.