The Provoking Thing: A VR Relationship

Josephine Anstey, Dave Pape

Department of Media Studies,
University at Buffalo, Buffalo NY, 14260
jranstey@buffalo.edu
Electronic Visualization Laboratory
University of Illinois, Chicago, IL 60607
pape@evl.uic.edu

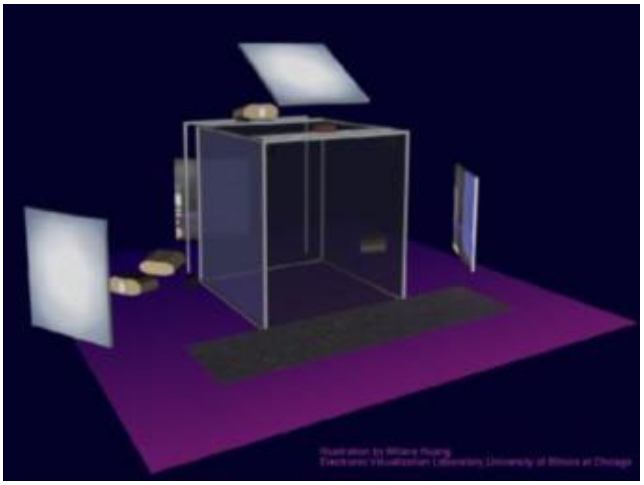
Abstract

The Thing Growing is a CAVE® virtual reality artwork designed to promote an emotional relationship between a user and a computer–controlled character, the *Thing*. The application uses literary and dramatic techniques to form a context for the growth of the relationship and as the basis for the construction of the computer character's intelligence.

Introduction

The Thing Growing is an art project; an immersive VR application centered on the creation of an emotional relationship between a user and a computer–controlled character, the *Thing*. The motivation for building the project was to bring a dramatic fictional experience into an interactive environment. One possibility of fictional narratives is to take the reader or viewer on an emotional roller–coaster ride. The reader or viewer identifies with the protagonist's feelings as they change in response to difficult or dangerous circumstances. Our goal was to create a similar experience in VR, but in this case the user *is* the protagonist and her own emotional responses become a major part of the fiction. Although several research teams have focused on using intelligent agents for interactive, narrative experiences, including the *Virtual Theater for Children* (Hayes Roth, Brownston, and Sincoff 1995), the *Alive Project* (Blumberg, and Galyean 1995), the *Oz project* (Matues 1997), we felt that they addressed the user too cerebrally to be useful role models for our application. In these projects the user is invited to create fiction with the agents; to direct them; or to marvel at simulations of autonomy and personality. We wanted to provoke a more basic emotional response in the user. Anstey's background is in fiction, and experimental video narrative (Kotz 1993). Therefore our approach to building the intelligence for the application was to use some basic narrative and dramatic strategies, and to focus on the agent's character and responsiveness to the user rather than on its own set of intelligent and autonomous behaviors.

The Thing Growing was built between 1997 and 2000 at the Electronic Visualization Laboratory, University of Illinois at Chicago for a CAVE (Cruz–Neira, Sandin, and DeFanti 1993), ImmersaDesk (Czernuszenko et al. 1997), or CAVE–like, VR display system. The CAVE is a room–sized, virtual reality theater. Computer generated images are rear–projected onto the walls and front–projected onto the floor. In a CAVE system, the position of the user relative to the VR world is tracked by a number of sensors. A sensor on the head allows the system to correctly calculate the 3D perspective of the real–time graphics as the user moves through the virtual environment. To create the stereo effect the user wears active–shutter glasses and the graphics system sequentially projects an image for the right and left eyes. The user carries a wand with a joystick for navigation and buttons which can be programmed for interaction. This wand is also tracked so that the system knows where the user's hand is. For *The Thing Growing*, we added a tracker on the other hand. Therefore the information that our application knows about the user is the position of her head and two hands, and whether she is using the joystick or buttons.



CAVE

Section one introduces the story *The Thing Growing* application was based on and the aims of the project. Section two briefly discusses some design choices; of the *Thing*'s visual appearance, voice and moods; and of the dancing activity which was the most significant trope used to build the relationship between user and *Thing*. Section three and four describe the intelligence implemented in the application both in the narrative strategies used and in the creation of the virtual character. We finish with a short assessment of the project.

1. Basic Story and Project Goals

The Thing Growing was originally a short story Anstey was developing. Its central idea was to represent a relationship that was cloying and claustrophobic but emotionally hard to escape. The goal of the VR application was to create the conditions for such a relationship to evolve between the user and the computer controlled character, the *Thing*. The title refers to the *Thing*'s insidious and growing dominance over the user's emotional life (within the scope of a 15 minute virtual experience!). Three acts and the varying behavior of the *Thing* are designed to stimulate a sequence of emotions in the user; from interest and affection to annoyance and frustration, then to a sense of loyalty, pity or kinship with the creature. Finally the user is presented with a choice, to kill the *Thing* or let it live. If the user shoots it, it doesn't die – rather it tells the user she has failed the test and must return to the beginning. If the user does not shoot, the *Thing* assumes that she loves it and wants to live with it forever. The killing/not killing can be read metaphorically – either the user is violently ending the relationship or remaining in a relationship which is far from perfect. However, the consequences of the decision reveals there is no escape and refers to the problem of repeating patterns in our relationships.

The challenge of the project was to create scenarios and possible interactions between the user and virtual character that would promote the development of a virtual relationship along these lines. Successive drafts of the project aimed to accommodate different user's responses. We were not interested in designing a system that could always elicit the same emotional response from all users, but in creating a story that would react with its own internal logic to the any responses that were evoked.

Michael Mateas contrasts the criteria the scientific community use to build and judge AI with the criteria used to build and judge a cultural production that may use AI in some way (Mateas 2000). He suggests that important issues for the AI community are: task competence; objective measurement; generality and realism. For cultural production the issues are: poetics/aesthetics; audience perception; specificity; artistic abstraction. *The Thing Growing* is a cultural production. The computer controlled character the *Thing*, is not designed to accomplish a well–defined and specific task. We did not systematically record, measure or analyze the users' emotional responses to it, or run the application with a defined range of different users. We were not trying to build a system that could create intelligent characters for interactive narratives in general. The application was not designed to look real. Instead our goal was to create an involving, challenging, subjective experience for an audience that would encounter the application in an exhibition or museum context. The creation of the character was driven by audience perception of its intelligence, responsiveness, and believability. We assessed the work during development and at completion by observing people as they interacted with the application and talking to them afterwards. This qualitative assessment was a vitally important part of refining the dramatic impact of our narrative. – although the user is a protagonist in our piece we want to shape her participation in a way that furthers the basic story we are presenting. Most importantly our goal was to use immersive graphics, sound, and an intelligent system to allow the user to explore the terrain of a specific relationship and "engage the audience in specific processes of interpretation" (Mateas 2000) with respect to that relationship.

2. Design Choices

The graphic design of the virtual character's body is very simple and non–photorealistic. It is a collection of transparent pyramids – one for the head, one for each arm, one for the body and several for a tail. The head has eyes but no mouth, however the head flashes gently when the *Thing* is speaking in order to give a visual as well as audio clue. The pyramids are not attached to one another, which avoids the problem of body parts joining up badly when the character moves. We used a simple motion–tracking system to animate it. The life–like movement that results from the motion tracking creates a strong illusion of an autonomous being formed from the collection of primitive shapes. The *Thing*'s body movement are also used to convey its emotional state – for example it is abrupt and jerky when it is mad, flowing and at ease when it's happy. All the motion–captured movements for the *Thing* are pre–recorded. Scott McCloud suggests that viewers can more easily identify with simply drawn, iconic, cartoon characters (McCloud 1993). In the same way we believe that this simply designed character will be able to stand in for any significant other in the user's life (spouse, sibling, parent, child) so that elements of the user's own relationships will seep into the emotional narrative.

User dancing with the Thing

We made the assumption that the user would react emotionally to a computer character. Human beings personify and react emotionally to their cars and computers, we extrapolated that they would be equally

An early design decision was to use a dancing activity to build the relationship between the user and the *Thing*.

3. Narrative Strategies

A simple narrative provides a framework for the application and the intelligence system. The narrative structure is quite controlled and keeps elements like pacing and surprise in the hands of the author not the user – this strategy maximizes the dramatic tension of the piece (Anstey, Pape and Sandin 2000). *The Thing Growing* has a bridge structure with three acts and the experience lasts about 15 minutes. The first act introduces the situation and the characters, the second act is concerned with their evolving relationship and problems that arise, the third act contains the denouement. The order of the acts is fixed but within each act there are interactive episodes. The user is therefore not in control of the entire progress of the story but experiences a measure of free will within the acts. Following a Hollywood film formula the application has plot points at the transitions between the acts designed to surprise the viewer and to provide dramatic revelations or reversals (McKee 1997). The plot point between act one and two is the *Thing's* announcement that it loves the user. The plot point between act two and three is that the *Thing* and user are suddenly transported into a new environment and confronted with an enemy. Each act is also has the goal of eliciting a specific emotional response or set of emotional responses from the user.

The goal of act one is to stimulate feelings of well–being, pleasurable anticipation and affection for the *Thing*. The elements used to stimulate these feelings are cheerful, cartoon–like graphics; surprise, introduced when the user opens a box and produces an explosion releasing large rocks and the *Thing* into the environment; and the *Thing*'s delight at being free, and its expression of gratitude and immediate affection for the user.

Act two is designed to stimulate conflicting feelings about the *Thing*. The main device for stimulating these emotions is the dance activity described above. The range of emotions we are trying to stimulate include; affection for the *Thing*, exasperation and annoyance; a sense of being trapped or bullied; resistance to bullying; a sense of failure or not being good enough. The narrative comprises a series of role reversals to encourage these feelings. For example at one point the *Thing* asks the user to copy its movements, at another it will copy the user's movements.

Act three is designed to stimulate feelings of loyalty, pity, fellow–feeling with the *Thing*, aggression and empowerment. It culminates with the presentation of the choice to kill or not kill the Thing. These feelings are evoked by introducing four new characters, the *Thing's* cousins, who threaten the user and the *Thing* and become their common enemy. The user is given a weapon to fight these creatures and then has the option of turning it on the *Thing*.

The Thing Growing was built using XP (Pape et al. 1998), a VR authoring toolkit designed to facilitate the construction of art applications in the CAVE. The toolkit handles a number of activities common to VR environments, such as assembling objects into a world, collision detection, navigation, detecting events and passing messages in response to them. It provides a framework for extension; application–specific classes may be added to define behaviors for objects or characters. It provides a text file system to rapidly assemble virtual scenes: all the models, objects, their locations and behaviors are described in the text file along with messages to be passed between objects. *The Thing Growing*'s narrative structure was created with the text files; scripted sequences were intercut with interactive episodes; the narrative flow as a whole was produced using triggers based on time, user proximity, or the completion of specific events. The text file served as production manager for the story and could easily be edited and changed. Multiple triggers were sometimes used to avoid situations where the user could get stuck at a point in the narrative unless she performed a specific action, a strategy similar to that of Kidsroom (Bobick 1999). The narrative is comprised of sequences that explicate the story, periods of interactive possibility, and transitions.

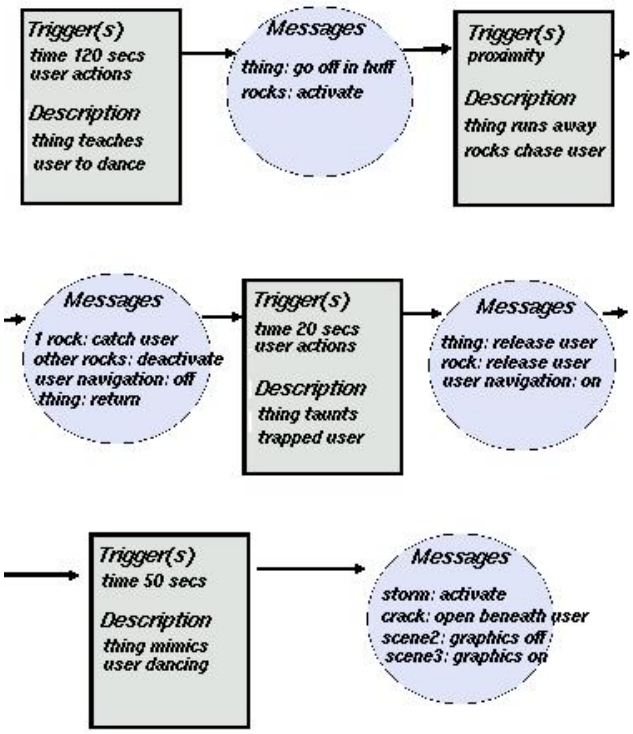


Figure 1

Figure 1 is a simplified diagram of the narrative sequencing in act two of *The Thing Growing*. Each gray box represents a section of the narrative, and contains the kind of triggers that are used to move the narrative on and a description of the activity taking place. Each blue circle contains the messages activated by the triggers. These messages contain instructions which set up the next narrative section. The act begins with an interactive sequence as the *Thing* teaches the user to dance (the details of this procedure are given in the section 4), the dance activity may continue for a maximum of 120 seconds before messages are sent to start the second

section of the narrative. However, during the dance activity the application also counts how many times the user disobeys the *Thing*'s injunctions to dance – disobedience consists of not dancing or repeatedly navigating away from the *Thing* as it attempts to teach – a maximum number of disobediances will also trigger the messages that start the second narrative section. In the second section, the *Thing* runs away from the user and hides, meanwhile rocks in the environment come alive and stalk the user – a simple algorithm is used for the rocks' movement and they move with reference to the proximity of the user and each other until one of them is near enough to trap the user. This triggers messages that cause one rock to catch the user while the other rocks are deactivated; a message is sent to the mechanism that controls the user's ability to navigate and it is turned off; and a message is sent to the *Thing* to reenter the scene. The third narrative section consists of the *Thing* taunting the user who is now trapped under the rock. Triggers for this section are time or the user spontaneously dancing to please the *Thing*. In the fourth narrative section the *Thing* mimics the user's movements – data from the tracking system is relayed to the *Thing*'s body parts to create this effect. This section has a fixed time limit. of 50 seconds before triggering messages that end act two and start the transition to a completely new graphic environment for act three.

4. Creation of Virtual Character

The XP system was extended to build the *Thing's* intelligence, which is a simple, hierarchical, finite state machine related to the narrative. Examples of states at the highest level in the hierarchy are EMERGE _FROM_BOX, TEACH_DANCE, MIMIC_USER. The narrative flow moves the *Thing* from state to state. In the EMERGE_FROM_BOX state the *Thing* performs a scripted sequence of actions. Its behavior as it emerges from the box it is trapped in and first meets the user is always the same; it says the same lines and moves in the same way. However, the TEACH_DANCE state describes an interactive episode and is more complex. It has two sub–states: TEACH USER TO DANCE and REACT TO USER WHO IS RUNNING AWAY. A change from one state to the other is triggered by the user – if she attends to the *Thing* and tries to dance the *Thing* is in the TEACH USER TO DANCE state, if she navigates away from the *Thing* it changes to the REACT TO USER WHO IS RUNNING AWAY state. Both these states are sub–divided into further states with a basic rule system for the Thing to follow. Figure 2 is a simplified diagram of the sub states and rules for the TEACH_USER_TO_DANCE state.

First the *Thing* **teaches a new step** to the user. The dance step is one of the short, pre–recorded, motion–captured movements. The *Thing* performs it while humming a related pre–recorded sound file. Next the *Thing* **watches while the user dances**. At the same time the application is **checking the state of the tracking system** to assess the user's performance. This checking process determines what the *Thing* will do next. It will make a comment recognizing that the user either **danced correctly**, **danced badly** or **didn't dance**. If the user has **danced correctly** the *Thing* make its comment and proceeds to **teach a new step**, and so the process repeats. If not the Thing will **repeat the step together with the user,** while simultaneously **checking the state of the trackers**.

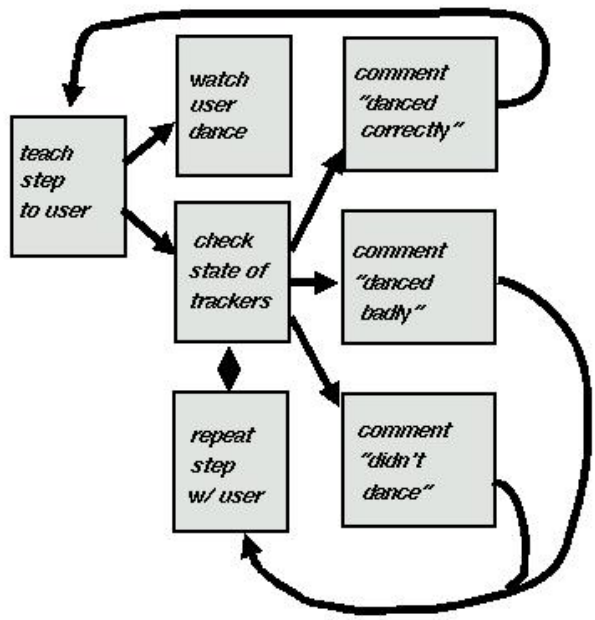


Figure2: TEACH_USER_TO_DANCE

For each of the sub–states described in figure 2 there are a number of similar actions that the Thing can perform – by action we mean the *Thing's* body parts executing its motion–captured motion while a sound file plays. This adds variety to the creature's responses and prevents repetition. There is not just variety in the movement or words of the phrase, but also in the mood of the action. Before deciding on an action the *Thing* therefore has to check to see what its mood is. The *Thing*'s mood changes depending on the point in the narrative and feedback from the user with some random noise thrown in. In the dancing section it starts out happy. A user who dances enthusiastically keeps the *Thing* happy or makes it manic. A user who moves more sluggishly may change the mood to sad or even angry. So although the *Thing* may acknowledge that the user is doing the dance step correctly it does so in a critical rather than praising manner. As previously mentioned we count the number of times that the user does not comply with the *Thing*, by not dancing or by running away. As this count increments it triggers alterations to the *Thing's* mood making it sad then mad. The *Thing*'s mood also deteriorates over the timed period of the dancing. However enthusiastic the dancer is, it tends to become sad or angry and complains that she is not trying hard enough. The system also progressively

judges the correctness of the user's dancing more harshly. The inevitability of the *Thing's* worsening mood is demanded by the story structure. The dance activity is broken by the *Thing* running off in a huff, upset by the user's inability to please it. The intelligence is designed to reach this point by one path or another – either the user disobeys too many times or the timed dance sequence is over. The reason that the *Thing* leaves, dramatically speaking, is to simulate a rift in a relationship.

To summarize, the *Thing's* behavior consists of a library of about 500 actions. When the piece is running, the *Thing*'s intelligence selects an appropriate action according to the point in the narrative, the user's actions, and the *Thing*'s own emotional state. It sends messages to its voice and body parts to execute that action. One of the major headaches of writing for the *Thing* was to ensure that anything it said fit in with the last thing it said and the next thing it was going to say. This meant keeping track of the hundred of different phrases and the multiple ways they could be combined. The constraints of the narrative and basic routines like the one described above made this possible. The XP textfile made it easy to change the way the actions were organized and combined if phrases did not fit together well.

In many ways it is fortuitous that our virtual character is dominating because a very pro–active character can both "tell" and control the story – the user blames her lack of control on this character and on her own inability to wrest control from it not on the limitations of the program. And of course the program is limited. We cannot write code for every possible action a user may make. We need to find ways to constrain those actions into a manageable subset that the program can react to. Pinning down the human subject and her responses was not a one shot deal. We consistently tested the application with a variety of users and watched what they did. Then we adjusted, refined, and added to the *Thing*'s functionality so that it had responses that could fold the users' different reactions back into our narrative thread. This iterative process also honed the story since watching users made it very clear when people were not really "getting" the narrative we wanted to send.

Assessment of Application

Robert McKee suggests that a good movie creates a rhythm of rising and falling tension, which allows the audience to reach tremendous emotional moments with a clarity that is absent from real life (McKee 1997). *The Thing Growing* is similarly designed to allow the user to take part in a model of a dysfunctional relationship, that is simplified and for that reason more easy to grasp. Much of the *Thing*'s behavior models childish power plays. It is designed to engage at a level beneath that of polite adult intercourse. It pouts, whines and threatens when it doesn't get its own way. It flatters outrageously and insults viciously. It gloats when the user is in its power. To express its love it copies the user or demands to be copied. All of these behaviors are designed to conflate self and other. In the *Bonds of Love*, Jessica Benjamin suggests that during the process of differentiation from the mother, the child's task is not merely to establish that it is separate but that a step of mutual recognition must occur as the child realizes that the other is also a subject. "And mutual recognition is perhaps the most vulnerable point in the process ..." (Benjamin 1988) She discusses the sadistic or masochistic positions that the evolving self may become stuck if unable to negotiate this point : "If I completely control the other, then the other ceases to exist, and if the other completely controls me, then I cease to exist." (Benjamin 1988) It is this emotional territory that the *The Thing Growing* is designed to explore.

There are several reasons for involving the user in this provoking relationship. Some people may fall into similar patterns in their own relationships, these people may be able to see the patterns of the relationship more clearly in this fictional setting and assess their part in its creation more objectively. The virtual relationship is a safer space to play out issues of domination and control than real life. Other people may be very immune to the kind of relationship that the *Thing* inflicts. The virtual environment gives them the opportunity to step into an unfamiliar psychological pattern and may give them insight into other people's more troubled relationship patterns. McKee suggests that the real self of a movie character is revealed as they are put under pressure that forces them to make moral or ethical choices. In this virtual environment we put the user under pressure so that, possibly, she reveals herself to herself. She may also role play in the environment making choices that she would not make in her real life and allowing emotions that she would normally censor.

How can we assess whether The Thing Growing meets the goals outlined above, or, indeed, that any art experience conveys its essence to the audience? The judgment is necessarily subjective. It depends heavily on the particular audience member. We have shown *The Thing Growing* in various states of development to approximately 500 people. Our observations of people interacting with the application, and feedback from users tells us that at the very least it is engaging and entertaining for most people. Some reactions and comments indicate that for some people it does hit deeper into the psychological realms we are interested in exploring. All our evidence is anecdotal, but users have specifically said that the *Thing* feels like a being who is present and alert to them. They say it reminds them of their spouse or child. The way people talk about the *Thing* also indicates that they are engaged with it, reacting emotionally and also judging its behavior. The following comments were made by users who we interviewed after a show in March 2001 at the Electronic Visualization Laboratory in Chicago.

> It's so real dancing there in my face, insulting me making me mad. Actually I took a couple of swings at it. But I got my revenge with the laser gun, which was really cool. And I killed off its whole family.

The character was very manipulative.

I was talking back to it. It said, "You're doing a bad job." And I said, "I'm trying." I became very interactive, which I think you should do, because if you don't you won't enjoy the [experience].

References

J. Anstey, D. Pape, D. Sandin, "The Thing Growing: Autonomous Characters in Virtual Reality Interactive Fiction," in the Proceedings of IEEE Virtual Reality 2000, New Brunswick, NJ, March 18 – 22, 2000

J. Benjamin, The Bonds of Love: Psychoanalysis, Feminism and the problem of Domination, Random House, NY, 1988

B. Blumberg, and T. Galyean, "Multi–Level Direction of Autonomous Creatures for Real–Time Virtual Environments," In Proceedings of SIGGRAPH 95, ACM SIGGRAPH, Los Angeles, CA, August 1995. pp. 47–54.

A. Bobick, S. Intille, J. Davis, F. Baird, C Pinhanez, L. Campbell, Y. Ivanov, A. Schütte, A. Wilson. "The KidsRoom: A Perceptually–Based Interactive and Immersive Story Environment," Presence Vol 8, Number 4, August 1999. pp. 367–391

C. Cruz–Neira, D. J. Sandin, and T. A. DeFanti, "Virtual Reality: The Design and Implementation of the CAVE", in Proceedings of SIGGRAPH '93 Computer Graphics Conference, ACM SIGGRAPH, August 1993, pp. 135–142.

M. Czernuszenko, D. Pape, D. J. Sandin, T. A. DeFanti, G. L. Dawe, M. Brown. "The ImmersaDesk and Infinity Wall Projection–Based Virtual Reality Displays." In Computer Graphics, Vol. 31 No. 2, May 1997, pp. 46–49.

B. Hayes Roth, L. Brownston, E. Sincoff, "Directed Improvisation by Computer Characters," Stanford Knowledge Systems Laboratory Report KSL–95–04, 1995.

L. Kotz, "Anything but Idyllic: Lesbian Filmmaking in the 1980s and 1990s," Sisters, Sexperts, Queers, Arlene Stein ed. Penguin Books, NY, 1993, pp 67– 80

M. Mateas, "An Oz–Centric Review of Interactive Drama and Believable Agents," Technical Report CMU–CS–97–156, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA. June 1997.

M. Mateas, "Expressive AI". In Electronic Art and Animation Catalog, Art and Culture Papers, SIGGRAPH 2000. New Orleans, LA, August 2000.

S. McCloud, Understanding Comics, Tundra Publishing, Northampton, MA, 1993

R. McKee, Story: Substance, Structure, Style, and the Principles of Screenwriting, Harpercollin, 1997

D. Pape, T. Imai, J. Anstey, M. Roussou, T. DeFanti, "XP: An Authoring System for Immersive Art Exhibitions," In Proceedings of Fourth International Conference on Virtual Systems and Multimedia, Gifu, Japan, Nov 18–20, 1998