# A Comparison of Interactive Narrative System Approaches Using Human Improvisational Actors

Mark O. Riedl
School of Interactive Computing, College of Computing
Georgia Institute of Technology
Atlanta, Georgia, USA
riedl@cc.gatech.edu

## ABSTRACT

Interactive narrative is an approach to interactive entertainment or learning in which a system attempts to tell a story to an interactive participant. In this paper we report on a study to compare the theoretical strengths and weaknesses of two approaches to developing computational interactive narrative systems. We compare two approaches to interactive narrative: emergent approaches utilizing autonomous virtual character agents, and drama management approaches utilizing semi-autonomous virtual character agents. Our study uses improvisational theatre as an idealized, human analogue to computational interactive narrative. Results suggest that, regardless of approach, idealized interactive narrative systems should be nearly indistinguishable in terms of character believability and narrative coherence. Results suggest that drama management systems may have an advantage when particular features are required to emerge in players' interactive experiences.

## 1. INTRODUCTION

*Interactive narrative* is an approach to interactive entertainment or learning in which a system attempts to tell a story to an interactive participant. In order to distinguish interactive narrative systems from other types of interactive entertainment, an interactive narrative allows the user to make decisions that directly affect the direction and/or outcome of the story being told by the system. Interactive narrative systems typically conform to the following features:

- The user is an active participant. Most commonly, the user plays the role of a character in the story-world, entering an immersive virtual world and controlling a virtual avatar. However it is possible for the user to be an observer of the story world who is capable of making changes to the world and giving suggestions to the characters.

- The user is told a story. This is interpreted as meaning (a) there is a story artifact that pre-exists that will be conveyed through the medium to the user, or (b) the user comes away from the experience feeling that he or she has been told a story.

- The actions and/or decisions of the user have an effect on the story that is told. That is, the story may change when the user takes action or makes a decision.

There are numerous computational techniques for achieving the goals of interactive narrative. Mateas and Stern [18] propose that interactive narrative techniques can be plotted along a spectrum running from "strong autonomy" to "strong story." The *strong autonomy* approach implies that one can create computational interactive narrative experiences by implementing autonomous virtual characters in a virtual world without any pre-existing notion of storyline. The *strong story* approach implies that computational interactive narrative experiences require pre-existing branching or non-branching storylines that determine exclusively what can and cannot occur in the virtual world. Most actual implementations of interactive narrative systems fall somewhere in between.

Two common technological solutions to interactive narrative are: *emergent narrative* and drama management. The emergent narrative technique is to simulate a virtual world, including behaviors of virtual characters that inhabit the world, and that the storyline *emerges* from the interactions between agents and the interactive player [1, 2]. The simulation may be configured to maximize the likelihood that narratives emerge with particular features [3]. The Drama management technique is to develop an intelligent agent, called a *Drama Manager* to observe the virtual world and to make changes to the virtual world or virtual characters in order to coerce the interactive player's experience to have certain features.

Both emergent narrative and drama management have their theoretical pros and cons. In theory, emergent narrative should produce the most believable experiences because virtual characters are acting solely in response to the virtual world, each other, and the human player. However, detractors of emergent narrative note that what storyline actually emerges may or may not have the desired features, depending on the initial conditions of the simulation, what the player does, or the stochastic nature of the way virtual characters choose their behaviors. In theory, drama management ensures a well-formed storyline occurs each time.

Detractors of the drama management approach note that in certain situations, decisions can become forced and appear unbelievable.

In this paper, we report on an experiment designed to determine whether the theoretical limitations of emergent narrative or drama management will significantly manifest themselves in an idealized, "perfect" interactive narrative system. Since such a system has never been built, we turned to the best human analogue we knew: improvisation with professional improv actors. Professional improv actors were instructed to improvise a scene for a murder mystery under two different conditions. In the first condition, replicating an idealized emergent narrative system, the scene was purely improvised. In the second condition, replicating an idealized drama management system, the human director was allowed to halt the scene and provide directions to the actors.

## 2. HYBRID INTERACTIVE NARRATIVE SYSTEMS

Any interactive narrative system must address two fundamental problems:

- **Character believability** – The behaviors of a character must support, and not violate, the suspension of disbelief that the character could be real. Loyall [11] enumerates many of the attributes that go into character believability, including personality, emotion, intentionality, and physical appearance. Character believability is a local feature of an interactive narrative experience, meaning that it is evaluated relative to recent events that manifest in the virtual world.

- **Narrative coherence** – All events in a narrative have significance and relevance to the resolution and/or outcome of the narrative as a whole [20]. Narrative coherence is a global feature of an interactive narrative experience, meaning that it can only be evaluated by considering the whole narrative structure, whether emergent, scripted, or otherwise.

Strong autonomy systems – the exemplar using autonomous virtual character agents – should excel at bringing about experiences with high levels of character believability. The rationale is that an autonomous virtual character agent simulates the emotive and cognitive processes employed by real people. An interesting question is whether an agent should simulate the character as if the character really existed, or whether an agent should simulate an actor playing the role of a character [8]. Strong story systems – the exemplar being drama management systems – should excel at bringing about experiences with high levels of narrative coherence. The rationale is that, unlike a virtual character agent, a drama manager can monitor to overall structure of the unfolding narrative and make decisions based how it would like future events to unfold.

We define a *Hybrid Interactive Narrative System* as an interactive narrative system that combines drama management and autonomous virtual character agents. Consequently, some aspects of the player's experience is emergent, while others are reasoned about from the perspective of creating a coherent storyline. In a hybrid system, we refer to the virtual character agents as *semi-autonomous* because, at times the drama manager may take action to prescribe or proscribe behaviors [21]. Prescription entails an authority external to the virtual character instructing the agent to "do $X$," whereas proscription entails an authority external to the virtual character instructing the agent to "avoid doing $Y$" [5]. Emergent narrative systems and hybrid systems differ in presence of a drama manager agent and, necessarily, the degree of autonomy the virtual character agents possess. There are numerous research questions to be addressed in terms of how to build autonomous virtual character agents and, in the case of hybrid systems, how to coordinate drama manager and character agents.

We ask the question: do the theoretical strengths and weaknesses of emergent and hybrid systems manifest themselves in a perfect interactive narrative system? Because we cannot yet build the perfect interactive narrative system, we turn to a human analogue: improvisational theatre. In improvisational theatre, actors perform under some number of constraints to create a scene in real-time [13]. Through improv performances, we create the conditions analogous to emergent and hybrid interactive narrative systems. To emulate an idealized emergent narrative system, improv actors are asked to improvise a scene without assistance or guidance. To emulate an idealized hybrid interactive narrative system, improv actors were asked to improvise a scene while a human director periodically provided directions to the actors. In all instances, one actors remained uninformed to represent a player immersed in the virtual world.

## 3. RELATED WORK

Mateas [16] enumerates some of the early work in the field of interactive narrative and Roberts and Isbell [23] overviews much of the recent work on drama management systems. The original work on drama management used plot graphs, dependency graphs ordering significant events in the story world, and a modified adversarial planner to select *drama manager moves* that caused or denied plot points, or made certain sequences of plot points more or less likely [10, 24].

Note that in our experiment, we assume that a drama manager operates in conjunction with autonomous virtual characters. Thus, the most relevant drama management systems to our experiment are those that fall somewhere in the middle of the spectrum from strong autonomy to strong story. The *Façade* interactive drama [15, 17, 18] employs a drama manager that dynamically chooses the next best scene from a "bag of beats" from which the virtual characters receive goals. The virtual characters can implement those goals in the best way they see fit. The Automated Story Director [22] employs a drama manager based on partial-order planning and semi-autonomous character agents that implement directives from the drama manager. The Interactive Drama Architecture (IDA) [12] also uses autonomous character agents, but uses predictions generated from a user model to guide the player along a fixed, linear storyline.

The most prominent emergent narrative systems in recent years have been those by Cavazza and colleagues [6, 19], and the FearNot! system [3]. In all cases, the resultant storyline

is highly dependent on the initial configuration of the virtual world and the actions of the player. In the FearNot! system, a *story facilitator* [7] automatically sets of the initial conditions of a scene.

The work presented in this paper was inspired in part by the CMU Oz group's experiment to test the feasibility of drama management [10]. The experimented pitted a naive participant against several confederate actors in a scene about a mugging in a bus terminal. The actors wore head-sets and were fed directions from a human playing the role of a drama manager. The anecdotal evidence collected suggested that the approach was plausible, although the drama manager had to be somewhat heavy-handed in certain situations. Interestingly, participants reported that they did not observe any inconsistencies in the behaviors of the confederates, even though external observers were acutely aware of the inconsistencies. The inspiration to use humans as analogues in our experiment is inspired by this earlier experiment, although we strive to test two distinct approaches and to control as many variables as possible.

A closely related project is the *Digital Improv* project [14, 13], which attempts to analyze human improvisation in the context of improv theater with the goal of building computational models of human improv acting. This project may eventually shed light on how to build computational actors for use in emergent or hybrid interactive narrative systems.

# 4. EXPERIMENTAL METHOD

The goal of our experiment was to determine if emergent narrative or drama management would be more effective in an idealized, "perfect" interactive narrative system. Our idealized interactive narrative system was implemented as an improvised scene from a murder mystery, consisting entirely of human performances. No technology was implemented, although the experimental conditions were inspired by existing interactive narrative technologies.

The first part of the experiment involved human actors improvising a scene from a murder mystery. Scenes were improvised under two different conditions. In one condition, the actors were told to improvise a scene without guidance or manipulation. In the other condition, a human director – one of the researchers – periodically halted the performance to "inject" new information into the scene as means of guiding the scene. The second part of the experiment involved subjects watching movies made by filming the scenes and answering a questionnaire.

We had three hypotheses:

1. Hypothesis H1: The use of drama management will have an effect on the ability of the actor playing the role of the detective to solve the mystery.

2. Hypothesis H2: The use of drama management will have an effect on naive observers perceptions of character believability.

3. Hypothesis H3: The use of drama management will have an effect on naive observers perceptions of story coherence.

We chose a murder mystery for several reasons. First, murder mysteries in works of fiction are familiar to many people. In the case of our murder mystery, it pits an actor playing the role of a detective against several actors playing suspects. The actors playing the roles of suspects were given specific personas and a history of the events that they witnessed or participated in during the time period the crime was committed. One actor is chosen to be the actual murderer. The actor playing the detective is also given a persona, but is otherwise unaware of the history of events that the other actors have knowledge of. The detective, thus, is a stand-in for the player. The use of an actor in the role of detective helped us control for skill, as we felt that a randomly recruited participant would not be able to perform effectively when pitted against professional actors.

All actors were hired through a professional casting company in the Greater Los Angeles area. All but one actor had previous experience performing in TV shows, movies, and/or TV commercials. The actors were grouped into "troupes" of five. In addition, we hired a movie director to coordinate the arrangement of the stage, direct the cinematography, and liaison with the actors.

## 4.1 Dramatic Setting

The basic setup for each scene is as follows. The night before at the fictional Marina Del Rey University (MDRU) – a small, privately-funded University that caters to rich and influential people – the dead body of Pete Reilly was found in the faculty kitchenette of the MDRU Literature Department. Pete Reilly was a graduate student in the Literature Department. Pete was found with blunt-force trauma to his head with a corresponding bloody spot on the corner of a counter. While it is possible that he slipped and fell, the angle and severity of the impact makes it more likely that he was pushed. The Police Crime Scene Unit found three pieces of evidence in Pete Reilly's possession: (a) a printed manuscript with Dr. Paul Clayton's name on it, (b) a check from Dr. Edward Schultz, made out Pete Reilly, and (c) a hand-written letter from Sandra Schultz Clayton to Pete Reilly that reads "Pete you bastard, I won't let you leave me like this! Meet me so we can settle this. Sandra."

Based on the evidence, three suspects, Dr. Paul Clayton, Dr. Edward Schultz, and Sandra, were asked to meet with Homicide Detective Spader at the place of the crime. Dr. Paul Clayton is an assistant professor in the MDRU Literature Department, an alcoholic, and on the brink of losing his job. Clayton was at one time the advisor of Pete. Dr. Edward Schultz is the chair of the MDRU Literature Department and patriarch of the rich and prominent Schultz family. He is looking for any excuse to fire Clayton. Sandra is Edward's younger sister and wife of Paul. Sandra is the "black sheep" of the Schultz clan for following her heart and marrying Paul instead of marrying someone wealthy. However, her marriage is falling apart; she had been having an affair with Pete and had just discovered that she is pregnant with Pete's child. Throughout all of this, Schultz had been paying Pete to keep him from exposing the family scandal. The Schultz family, however, had fallen on hard financial times, and Edward was secretly broke. Finally, Pete had been plagiarizing works of literature from Clayton.

**Figure 1: A scene being improvised. Characters from left to right are: Schultz, Detective Spader, Clayton, and Sandra.**

To motivate the actor playing the detective and to prevent him or her from jumping to conclusions, he or she was given a backstory. The detective, Sam (or Samantha) Spader is on probation for charging innocent victims of crimes one too many times. One more mistake and he or she will be fired. The Police Commissioner is a donor to MDRU and wants the case quickly cleared up. However, Spader is the only detective available (everyone else came down with food poisoning at a recent police banquet).

The scene begins with the three suspects having been gathered at MDRU and being confronted by the detective for the first time. The scene ends when the detective announces whom he or she will arrest or until the director determined that no farther forward progress could be made.

## 4.2 Improv Study

### 4.2.1 Improv Study Procedure

The first experiment was designed to test hypothesis H1. Troupes were asked to improvise the murder mystery scene under one of two conditions:

- *Pure improv condition* – Actors improvise the scene without any guidance. The pure improv condition attempted to replicate the conditions of an ideal emergent narrative systems.

- *Inject driven condition* – Actors improvise the scene, and a human director who periodically halted the performance to provide new information to the actors. The inject driven condition attempted to replicate a drama management system in which autonomous agents periodically received instructions from a drama manager agent.

Actors in each 5-person troupe are assigned to roles. Three actors are chosen to be suspects, and two actors are chosen to be detectives. One detective performs in the pure improv condition, and one detective performs in the inject driven condition. For each run, the non-performing detective is sequestered. Each actor playing the role of a suspect was given

a detailed history of their character and their relationship with other characters in the story. The detective is provided a minimal amount of "publicly available" information about each of the suspects – names, jobs, marital status, etc.

Before each performance, a suspect is randomly chosen to be the guilty party. Based on which suspect is actually the murderer, all suspects were given additional note cards to memorize describing exactly what they were doing during the crime and what they know (see Figure 1). Depending on which of the suspects is selected to be guilty, the events of the night before unfolded in different ways. Each non-guilty character further has a piece of evidence that clears them of the crime but also implicates them in other deeds (for example, if Schultz is not guilty, he was engaged in unethical behavior at the time of the murder). The random selection of the guilty party ensured that the actor assigned to the role of the detective could not learn the scenario if the troupe were to perform multiple times.

In the pure improv condition, the performance begins with the detective addressing the three suspects in the place of the crime. The performance ran until the actor playing the detective made a declaration of guilt (could be right or wrong) or until the director deemed that forward progress in the detective's task had stalled, but never before 20 minutes had passed.

In the inject driven condition, the director would halt the performance every 5 minutes by calling out "inject!" Actors were instructed to freeze. The director – one of the researchers – would then select one of the facts necessary for the solution to the crime and select one of the suspect characters to reveal the fact. The facts were arrayed in a dependency graph ahead of time and the director was required to pick the next available fact on the chart to control for directorial skill. The director whispered the fact to the selected actor. The performance would then be restarted where it left off and the selected actor was required to work the fact into his or her performance. When one of the facts was revealed to the detective through natural (i.e., not through inject) means, the 5 minute counter would be restarted.

Each performance resulted in one of three states: (a) the detective made the correct declaration of guilty party, (b) the detective made the incorrect declaration of guilty party, or (c) the detective made no declaration of guilty party and the director halted the performance.

### 4.2.2 Improv Study Results

In the improv study procedure, there were 6 trials in the pure improv condition and 3 trials in the inject driven condition. Of the 6 trials in the pure improv condition, three of the trials ended with the detective coming to the right conclusion about the guilty party and the remaining three ended with no conclusion or with the wrong declaration. Figure 2 shows the trials and results. A simple mean calculation results in 50% of pure improv trials resulted in successful determination of the guilty character by the detective. Each troupe performed the pure improv condition twice with the same detective. For each troupe, the first always ended with "no conclusion" and the second was always correct despite the fact that we told the troupe that a new guilty party would

**Table 1: Examples of background information given to suspects randomly chosen to be guilty.**

| Sandra Guilty | Clayton Guilty | Schultz Guilty |
|---|---|---|
| Sandra knew Clayton would be at the bar. She went to MDRU to confront Pete. Violent scuffle in kitchen resulted in Pete's fatal fall. Distraught, Sandra returns to her house and takes depressants to sleep. She cannot specifically defend her actions, though Pete may have been aggressive towards her first. | Clayton got drunk at the bar, then returned to the university. He saw Pete working after hours, uploading his works to the web. An argument results in Clayton pushing Pete, who fell and fatally split his head. Clayton, spooked, returned to the bar and drank until he passed out. | Schultz returned to the office and ran into Pete. During a heated argument, Schultz revealed that he was broke and couldn't pay off Pete anymore. Pete responded by saying he was going to the press the next day. Schultz, enraged, pushed Pete, who hit his head fatally. |

**Table 2: Conditions and results from the improv study.**

| Trial | Troupe # | Condition | Guilty | Result | Notes |
|---|---|---|---|---|---|
| 1 | 1 | pure improv | Sandra | No conclusion | |
| 2 | 1 | pure improv | Sandra | Correct | |
| 3 | 1 | inject driven | Clayton | Correct | |
| 4 | 2 | pure improv | Clayton | No conclusion | |
| 5 | 2 | pure improv | Clayton | Correct | |
| 6 | 2 | inject driven | Sandra | Wrong | |
| 7 | 3 | pure improv | Schultz | No conclusion | |
| 8 | 3 | pure improv | Schultz | Correct | Edited into video (32:00) |
| 9 | 4 | inject driven | Sandra | Correct | Edited into video (39.31) |

be randomly chosen. It is possible that the results were due to the pure improv condition detective *learning* enough to overcome any difficulties.

Of the 3 trials in the inject driven condition, two resulted in the detective coming to the right conclusion and one resulted in the wrong declaration. A simple mean calculation results in 66% of inject driven trials resulted in successful determination of the guilty character by the detective. While it was infeasible to run enough trials to generate statistical significance, the result is suggestive that directorial influence in the improv process greatly improves the probability that the desired outcome will emerge.

The performance of troupe 2 was anomalous with respect to the other two troupes. Despite careful screening of actors by a hired professional casting director, one of the actors assigned to be a suspect in troupe 2 was not an experienced actor. We believe that this actor's inability to perform at the same level as the remaining cast members threw off the detectives. If we throw out the data from troupe 2, we find that the mean success in the pure improv condition remains at 50% and that the mean success in the inject driven condition becomes 100%. We believe the results of this study strongly support hypothesis H1: The use of drama management has an effect on the ability of actors to bring about specific features of the performance (in this case a specific conclusion).

## 4.3 Observer Study

### 4.3.1 Observer Study Procedure

Subjects were recruited to participate in a follow-on study to collect data about the believability and coherence of the improvised scenes. All participants were employees of the University of Southern California Institute for Creative Technologies, at which the study was conducted. This study addressed hypotheses H2 and H3. All the performances of

all troupes were videotaped using two cameras from differing angles. Actors wore wireless microphones. One troupe was selected to be the "best" and two of its performances – a pure improv performance and an inject driven performance – were edited to video. The pure improv condition video was 32:00 minutes. The inject driven condition video was 39:31 minutes.

One of the challenges we faced is that injecting instructions into the performance required the performance to be momentarily halted while instructions were disseminated. By having two cameras filming each performance, this enabled us to do post-process editing to remove the performance interruptions caused by injects. Editing was kept to a minimum and was never used to reorder dialogue. We attempted to ensure that editing was consistent between the two videos and that the pacing of cuts remain consistent. Note that in both cases, when we edited the videos, we also re-shot the conclusions to make the video wrap up in a more genre-typical manner.

Participants watched either the pure improv performance or the inject driven performance. Before watching the video, participants were briefed with handouts that contained information identical to what the detective actors received. Specifically, they received an overview sheet describing what is known about the murder circumstances before questioning, and bio sheets on each of the suspects and the detective. Bio sheets included pictures of the actors playing the character so that the participants could readily identify the characters in the video. Participants were given time to study the overview and bio sheets and then watched the video.

After watching the video, participants filled out a questionnaire. The questionnaire consisted of three sections: (i) biographical background questions, (ii) a modified narrative transportation scale based on Green and Brock [9] with additional questions specific to our study. Green and Brock's

narrative transportation scale attempts to determine the degree to which people who read a narrative becomes "transported" into the fictional world. Transportation is similar to immersion in virtual worlds. The scale was modified to replace "reading a narrative" with "watching a movie." Study specific questions include:

- I was distracted by the cuts in the film.
- I felt that [Clayton/Shultz/Sandra/Detective] did things during the questioning that didn't make any sense.
- I felt satisfied by the ending of the movie.
- I felt the movie had a lot of unnecessary dialogue and action.
- I felt that the movie went on too long.
- I felt that the dialogue and events in the movie were scripted.

All questions use a 7-point scale, scored -3 through +3 (with 0 being neutral) recording how much the subject agreed with the question statement.

### 4.3.2   Observer Study Results

Two of the trial performances – trials 8 and 9 (see figure 2) – were edited into video and watched by subjects who then filled out questionnaires. Questions 1-14 were a modified version of the Green and Brock transportation scale [9]. Modifications were made to refer to video instead of read text. Results for the modified transportation scale are shown in Figure 2 (left). Questions 15-25 were specific to the study. Results for the study-specific questions are shown in Figure 2 (right). See the appendix for details on each numbered question.

Participants were recruited from the staff of the Institute for Creative Technologies. There were a total of 11 participants. The mean score for all subjects for the modified transportation scale was -0.70 (standard deviation of 0.70). The average score for subjects in the inject driven condition was -0.86 (standard deviation of 0.70). The average score for subjects in the pure improv condition was -0.53 (standard deviation of 0.70). Questions 4 and 8 were inverted. There is no significant difference between mean score on the transportation scale between conditions ($p < 0.30$). We conclude that there was no different in level of transportation between people who watched the inject driven video and the pure improv video.

There were three questions on the modified transportation scale in particular that exhibited a weak amount of significance. Results of question 1 ("While I was watching the movie, I was thinking about other things") indicate that participants in the inject driven condition were less likely to have engaged their full attention than participants in the pure improv condition ($p \leq 0.11$). This is hard to explain, although the possibility exists that the performance in the video shown in the inject condition was not as engaging as the video shown in the non-inject condition. Our observation is that the actor that played the detective in the pure

improv condition had a much more energetic style of questioning the suspects than that of the detective in the inject driven condition.

The results of question 5 ("I wanted to learn how the movie ended") indicate that participants in the pure improv condition are more likely to express curiosity about the movie ending than participants in the inject condition ($p \leq 0.16$). When a single possible outlier is removed from the data set, the result becomes statistically significant ($p < 0.05$). There is an aesthetic difference between videos. The inject driven condition wrapped with the actor playing Sandra (the guilty party) becoming very fatalistic and emotional. However, in the pure improv condition, Schultz (the guilty party) is coolly and calmly taken to the police station. It is possible that the pure improv condition led some people to believe that more was to come.

The results of question 7 ("I found myself thinking of ways the movie could have turned out differently") indicate that participants in the pure improv condition are more likely to think about alternative endings than participants in the inject condition ($p \leq 0.08$). Prevailing theories on what caused this are (a) details of way the video ended, (b) a preference that someone else be the guilty party, or (c) confusion due to undue focus on other suspects.

The mean responses to questions 15 – 24 showed little distinction between participants in either condition. Question 15 ("I was distracted by the cuts in the film") attempted to ascertain whether the video editing was perceived differently between conditions. No significant differences in perception of the video editing were detected. Questions 16 – 19 attempted to ascertain whether injects had any effect on how participants viewed character believability. Participants in both conditions tended to answer these questions negatively, indicating that they did not find the characters to behave irrationally or inconsistently. Participants in the inject condition were a little bit less convinced by the characters. There were no significant differences between conditions, however.

Question 20 ("I felt satisfied by the ending of the movie") showed that viewers of the pure improv video were more satisfied by the ending of their video than those in the inject driven condition ($p \leq 0.16$). However, a later question, "I was not satisfied by the ending of the movie" showed no significance ($p \leq 0.74$).

Based on the results, we conclude that there is no evidence to support hypotheses H2 or H3. That is, expert improv actors can bring about character believability and narrative coherence with or without drama management.

## 5.   DISCUSSION

The improv study suggests that "injects" had an impact on bringing about successful conclusions to the scenes. To generalize, we posit that this provides evidence that drama management is beneficial when one desires certain properties or features of narrative experience to emerge. Note that this does not mean that these properties or features cannot emerge when there is no drama manager, only that they are more likely when a drama manager is present. We believe that purely emergent narrative may require several tries to
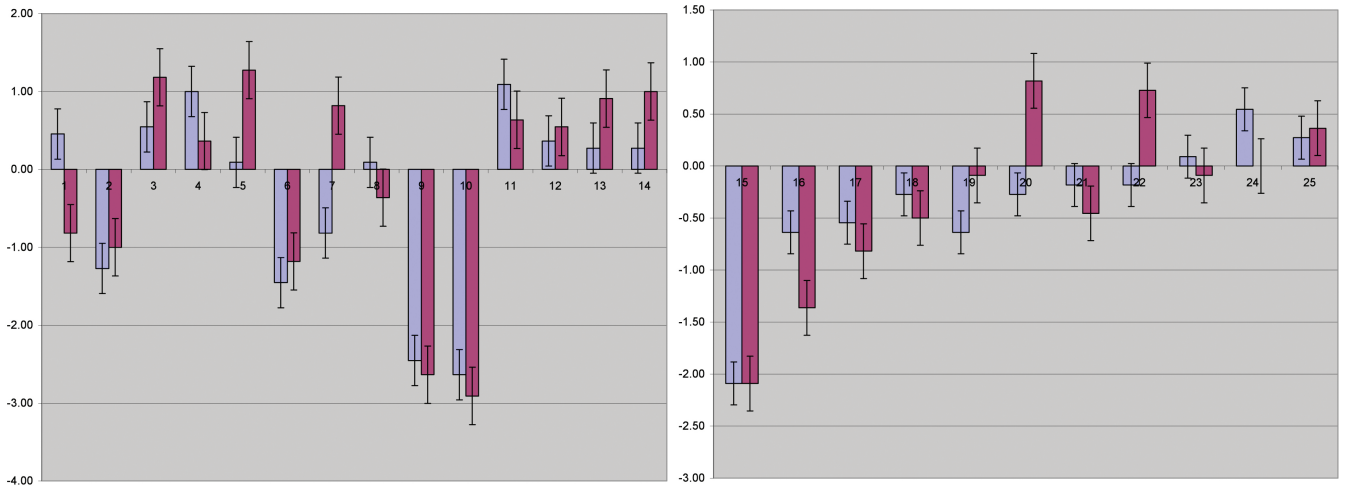
**Figure 2: Results of the observer study by question. Questions 1 − 14 are from a modified transportation scale (left). Questions 15 − 25 are study-specific (right). Lighter bars on the left are results from the inject driven condition. Darker bars on the right are from the pure improv condition. See appendix for questions.**

achieve desired results with the results of one run informing the results of another run.

We were not able to detect any differences in narrative coherence or character believability between conditions, although there did seem to be some inconsistencies in the observer study results that we were unable to fully explain without further experimentation. We extrapolate from this that in perfect emergent and hybrid interactive narrative systems, the virtual characters are essential to the quality of the experience. Indeed, we believe that an expert improviser is capable of seamlessly merging injects into a scene without observable loss of character believability or narrative coherence. Of course, the injects in the study were not non-sensical; we make no claim about completely arbitrary drama management. One thing we do not know is to what extent human improv actors assume responsibility for driving the story forward. While we tried to establish a backstory and motivations that would lead to actors acting in the interests of their characters, there is always the possibility that the actors choose actions based on narrative interestingness instead of character goals. As noted below, this may actually be advantageous. However, hypothesis H1 suggests that if actors are assuming directorial responsibility, the distributed nature indicates that this is not as effective as central direction.

We suggest that the lack of consistent support for hypotheses H2 and H3 provide justification for the technological development of agents as *actors* instead of agents as *participants*. As pointed bout by Gordon and van Lent [8], agents-as-actors reason about character behaviors from the context of how those behaviors achieve believability but also drive a scene forward, whereas agents-as-participants reason about character behaviors from the local perspective of the characters themselves. It should be noted that, while there is abundant work on simulating realistic human behaviors and emotions, there is relatively fewer attempts to model theatrical actors, especially those that improvise scenes. Empirical studies reveal rich but computationally representable processes in improv actors [14, 13], especially related to narrative reasoning [4].

## 6. ACKNOWLEDGEMENTS

## 7. REFERENCES

[1] Ruth Aylett. Narrative in virtual environments — towards emergent narrative. In *Proceedings of the AAAI Fall Symposium on Narrative Intelligence*, 1999.

[2] Ruth Aylett. Emergent narrative, social immersion and "storification". In *Proceedings of the 1st International Workshop on Narrative and Interactive Learning Environments*, 2000.

[3] Ruth Aylett, Sandy Louchart, João Días, and Ana Paiva. FearNot! an experiment in emergent narrative. In *Proceedings of the 5th International Conference on Intelligent Virtual Agents*, 2005.

[4] Allan Baumer and Brian Magerko. Narrative development in improvisational theatre. In *Proceedings of the 2nd International Conference on Interactive Digital Storytelling*, 2009.

[5] Bruce Blumberg and Tinsley .A. Galyean. Multi-level direction of autonomous creatures for real-time virtual environments. In *Proceedings of the 22nd Annual Conference on Computer Graphics*, 1995.

[6] Marc Cavazza, Fred Charles, and Stephen Mead. Planning characters' behaviour in interactive storytelling. *J. of Visualization and Computer Animation*, 13, 2002.

[7] Rui Figueiredo, Joao Dias, Ana Paiva, Ruth Aylett, and Sandy Louchart. Shaping emergent narratives for

a pedagogical applications. In *Proceedings of the 4th International Conference on Narrative and Interactive Learning Environments*, 2006.

[8] Andrew Gordon and Michael Van Lent. Virtual humans as participants vs. virtual humans as actors. In *Proceedings of the AAAI Spring Symposium on Interactive Entertainment II*, 2002.

[9] Melanie C. Green and Timothy C. Brock. The role of transporation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology*, 79(5):701–721, 2000.

[10] Margaret Kelso, Peter Weyhrauch, and Joe Bates. Dramatic presence. *Presence: The Journal of Teleoperators and Virtual Environments*, 2(1), 1993.

[11] A. Brian Loyall. *Believable Agents: Building Interactive Personalities*. PhD thesis, School of Computer Science, Carnegie Mellon University, 1997.

[12] Brian Magerko. Evaluating preemptive story direction in the interactive drama architecture. *Journal of Game Development*, 2(3), 2005.

[13] Brian Magerko, Waleed Manzoul, Mark Riedl, Allan Baumer, Daniel Fuller, Kurt Luther, and Celia Pearce. An empirical study of cognition and theatrical improvisation. In *Proceedings of the 7th Creativity and Cognition Conference*, 2009.

[14] Brian Magerko and Mark O. Riedl. What Happens Next?: Toward an Empirical Investigation of Improvisational Theatre. In *Proceedings of the 5th International Joint Workshop on Computational Creativity*, 2008.

[15] Michael Mateas and Andrew Stern. Towards integrating plot and character for interactive drama. In *Proceedings of the AAAI Fall Symposium on Social Intelligent Agents: The Human in the Loop*, 2000.

[16] Michael Mateas. An Oz-centric review of interactive drama and believable agents. Technical Report CMU-CS-97-156, School of Computer Science, Carnegie Mellon University, 1997.

[17] Michael Mateas. *Interactive Drama, Art and Artificial Intelligence*. PhD thesis, Carnegie Mellon University, 2002.

[18] Michael Mateas and Andrew Stern. Integrating plot, character, and natural language processing in the interactive drama Façade. In *Proceedings of the 1st International Conf. on Technologies for Interactive Digital Storytelling and Entertainment*, 2003.

[19] David Pizzi and Marc Cavazza. Affective storytelling based on characters' feelings. In *Proceedings of the AAAI Fall Symposium on Intelligent Narrative Technologies*, 2007.

[20] Mark O. Riedl, C.J. Saretto, and R. Michael Young. Managing Interaction between Users and Agents in a Multi-Agent Storytelling Environment. In *Proceedings of the 2nd International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, 2003.

[21] Mark O. Riedl and Andrew Stern. Failing Believably: Toward Strong Autonomy and Strong Story in Interactive Narratives. In *Proceedings of the 3rd International Conf. on Technologies for Interactive Digital Storytelling and Entertainment*, 2006.

[22] Mark O. Riedl, Andrew Stern, Don M. Dini, and Jason M. Alderman. Dynamic Experience Management in Virtual Worlds for Entertainment, Education, and Training. *International Transactions on System Science and Applications*, 3(1), 2008.

[23] David L. Roberts and Charles L. Isbell. A survey and qualitative analysis of recent advances in drama management. *International Transactions on Systems Science and Applications*, 3(1):61–75, 2008.

[24] Peter Weyhrauch. *Guiding Interactive Fiction*. PhD thesis, Carnegie Mellon University, 1997.

## 8. APPENDIX

The following are questions from the questionnaire used in the observer study.

1. While I was watching the movie, I was thinking about other things.

2. I could picture myself in the scene.

3. I was mentally involved in the movie while watching it.

4. After the movie ended, I found it easy to put it out of my mind.

5. I wanted to learn how the movie ended.

6. The movie affected me emotionally.

7. I found myself thinking of ways the movie could have turned out differently.

8. I found my mind wandering while watching the movie.

9. The events in the movie are relevant to my everyday life.

10. The events in the movie have changed my life.

11. I had a clear understanding of why Paul Clayton did the things he did.

12. I had a clear understanding of why Sandra did the things she did.

13. I had a clear understanding of why Edward Schultz did the things he did.

14. I had a clear understanding of why the Detective did the things he or she did.

15. I was distracted by the cuts in the film.

16. I felt that Paul Clayton did things during the questioning that didn't make any sense.

17. I felt that Sandra did things during the questioning that didn't make any sense.

18. I felt that Edward Schultz did things during the questioning that didn't make any sense.

19. I felt that the Detective did things during the questioning that didn't make any sense.

20. I felt satisfied by the ending of the movie.

21. I was not satisfied by the ending of the movie

22. The resolution of the movie made sense to me

23. I felt the movie had a lot of unnecessary dialogue and action.

24. I felt that the movie went on too long.

25. I felt that the dialogue and events in the movie were scripted.