Screen Play: Film and the Future of Interactive Entertainment

Andy Clarke

Kinonet Consultancy

78 West Kensington Court

Edith Villas

London W14 9AB

Tel: 44 (0)20 7602 3382

Email: andy@kinonet.com

Grethe Mitchell

Senior Lecturer in Interactive

Design (University of Lincolnshire

& Humberside)

Kinonet Consultancy

78 West Kensington Court

Edith Villas

London W14 9AB

Tel: 44 (0)20 7602 3382

Email: grethe@kinonet.com

Abstract

This paper looks at existing computer games and virtual environments from the perspective of film theory and practice. From this, we will draw conclusions about the ways in which the designers of computer games and virtual environments can use what has been discovered in the study of film to build more interesting, engaging and entertaining interactive narratives.

Introduction

Filmmaking has been called the art of telling stories through images; others have referred to it as "sculpting in time". In reality, it is both of these: a film is a series of images fixed in time.

This is true at the most fundamental level. We may refer to films as "moving pictures", but when one looks at a strip of film, it consists of a series of still images. Each of these is slightly different from those before and after it; when they are projected, our eyes take these still images and blurs them together - this is how we get the impression of movement.

On a filmstrip, we can see time recorded. We can measure time in frames - on a 35mm film, 24 frames equal one second - but we can also measure it in terms of distance. On the same size film, eighteen inches also equal one second.

Time, in film, therefore becomes something physical, something that can be cut and glued back together. I can take one section of film and place it after another shot - from a different place, or at a different time - and when it is projected, it forms one continuous sequence in viewing time even though the sequence may refer to different times in the story.

Even the simplest sequence is composed from a number of different shots, from various angles and locations. These shots add variety to the film, but

they are not used solely for this purpose - they also tell the story. A close-up of the hero allows us to see the expression on his face clearly, while a wider shot shows us where he is and what is around him. Indeed, a close-up does more than just allow us to see the expression on his face - by excluding all other detail, it *forces* us to look at his expression. The wide shot, likewise, forces us to look at the scene as a whole, rather than concentrate on one part of the scene. In a film, every detail is significant - film does not present reality, but rather an enhanced reality where every extraneous detail has been cut out and every object left in the frame has been considered on the basis of what it does for the narrative.

Yet meaning in film comes not only from what is in the frame, but also from how it is composed. A shot from high up, looking down on the hero makes him appear vulnerable. If this shot is held for a long time, then it implies that it is the viewpoint of someone watching him. These conventions form part of the "language" of film - the vocabulary of shots that the director can use, safe in the knowledge that the audience is familiar with them and know what they mean. There are also established conventions regarding styles of lighting, types of camera moves, the use of music and sound effects, etc.[1]

---

[1] A good introductory text on film language is Ralph Stephenson and Guy Phelps, *The Cinema as Art* (London: Penguin, 1989)

A further layer of meaning comes from how these individual shots are put together, one after another. David Mamet gives the example in his book *On Directing Film* of a documentary filmmaker using a shot of a bird snapping a twig and a shot of a fawn raising its head. He says:

> The two shots have nothing to do with each other. They were shot days or years, and miles, apart. And the filmmaker juxtaposes the images to give the viewer the idea of *great alertness* .... [The shots] are not a record of how the deer reacted to the bird. They're basically uninflected images. But they give the viewer the idea of *alertness to danger* when they are juxtaposed. That's good filmmaking.[i] (italics his)

By using the conventions of editing, the filmmaker can compress or expand time, jump between different places and times, switch between different viewpoints, or use close-ups to emphasise certain objects or actions, without confusing the audience. A cut implies connection and this, in turn, produces meaning - a meaning that may not have been present in either of the shots individually.

This ability to jump between different places and times means, for instance, that the filmmaker can produce suspense by cutting to another location and showing the bad guy waiting in ambush, or by cutting to some time in the past and showing the bad guy plant a bomb where our hero is now. In both

of these examples, the filmmaker reveals information to the viewer that is not available to the protagonist of that film. The opposite of this is when the filmmaker uses montage to show what is going on in the mind of the characters in the film. You can show a character sitting alone in a room, and then cut to a shot of someone else (in another location) to show that they are thinking of them. Alternatively, you can cut to the same room at some moment in the past; in this case, it is showing the character's memories.

In short, film has - over the years - developed a variety of techniques: camera angles, lighting, editing, use of sound etc. - with their readily understood conventions - to enhance narrative development. These techniques serve the narrative in compressing or expanding time, portraying emotion and feelings, emphasising objects and movement. The judicious combination of these techniques allows for powerful storytelling.

The use of editing enables time to be discontinuous - for the future, present and past to exist simultaneously e.g. in flashback or flashforward sequences, or parallel action sequences where the viewer is in two places at one moment in time. In film, "time" is never "real time", it is always being compressed and expanded at the service of the story.

The same could be said for the notion of "space" in film. Once again, space is flexible, compressible, expandable through the use of lighting, camera moves and angles, sound and editing. Space and perspective are also always

at the service of the narrative, there is no "real" or "true" space in film that exists outside of the drive of the narrative.

Aims

This paper looks at computer games and virtual environments from the perspective of film theory and practice. It acknowledges that there are clear differences between films and videogames, but asks whether we can learn how to create better virtual worlds and interactive narratives from the study of film techniques and of the language of film.

We are not seeking, however, to make simplistic parallels between the videogames being made now and the early days of film - such direct comparisons between old and new technologies are of limited use. Nor are we insisting that games adopt the techniques of film wholesale. There are fundamental differences between games and films - in what they aim to achieve, how they achieve it, the context of their consumption, and so on - and it is naive to think that games can use all of the same techniques of film, without modification.

What we intend to do is more subtle. We intend to look at existing computer games and virtual environments from the perspective of film language. Through this analysis, we hope to identify the ways in which computer games use film language and, more importantly, the points of

tension - those aspects of film language which computer games have modified or actively resisted. By doing this, we hope to suggest some ways in which more interesting, engaging, and satisfying games, virtual environments and interactive narratives can be produced.

For the purposes of this paper, our analysis of existing product will concentrate primarily on mainstream, mass-market, computer games. One reason for this is that we want to use examples that are easily available and viewable to those reading this paper. There is, however, another more significant reason.

Narrative techniques become conventions over time when a critical mass of the audience accept and understand what the conventions are doing. We therefore feel that it is important to look at those products which are aimed at the largest possible audience. By doing so, we are able to identify more clearly (and thereby discuss) those elements of the language of videogames which are already broadly understood (as well as those new elements which, through their exposure to a large audience, are most likely to become established).

We also believe that the conclusions we draw from our analysis of these mainstream videogames will be more widely applicable than those drawn by looking at other, more experimental, virtual environments seen only by a handful of people.

Narrative in Videogames

Having considered the techniques of narrative development and enhancement in film, we must now also look at what is understood by the term "narrative" in both film and videogames.

"Drama", as many writers have said, "is conflict". Normally this is used to say that drama can only be expressed through conflict, but the opposite is also true: as soon as there is a conflict, there is drama. All videogames, no matter how simple they are, have conflict - enemies trying to kill you, a hostile environment, or both - and they therefore have narrative. As soon as you have a goal or a mission, you have a narrative; as soon as you have a stable state that is disrupted, you have a narrative.

Here we are taking a broader - more archetypal - definition of narrative than that used by other writers on interactive fiction, such as Brenda Laurel[ii] and Janet H. Murray[iii]. They are looking for stories like those found in conventional theatre and literature - Shakespeare, Homer, Austen, Bronte, etc. - and this is reflected in the choice of both the products that they analyse and the conclusions that they draw from them.

When we say that videogames should adopt - and adapt - the storytelling techniques of film, it is not because we think that games should be like film - it is because there are clear formal similarities between the two media and

this affects both the stories that they tell and the way in which they tell them.

Both videogames and films are screen-based media - that is to say, they present their stories on a screen. A screen, whether it is a movie screen or a monitor screen, has certain properties: size, shape, luminescence, resolution etc. This means that it is better for some things rather than others: for instance, it is better used to present images than text.

It therefore seems inappropriate to look - as Janet H. Murray does - at text-based adventure games and text-based interactive narratives and use the conclusions she draws from them to make generalisations about *all* adventure games. The modern graphical adventure games, such as *Final Fantasy VII*, are not just text-based adventure games with graphics - the use of graphics allows a different type of interaction to text-based games and this, in turn, affects the type of story that they tell.

This has a parallel in film. A film is more than a literary or theatrical work presented or recorded on celluloid - the ability of film to tell a story in pictures (and to create meaning out of editing these images together) defines the type of stories that are appropriate to it and changes those stories adapted to film in a fundamental way.

We would argue that the designers of videogames have known right from the start that they could not have a conventional narrative (or tell it in a

conventional way) and so have developed alternatives to both of these. We will look at these alternative approaches to narrative - which include the sparseness and repetitiveness that we find in videogames - in our next section.

Narration in Videogames

Before we begin this section, it is necessary to first make the distinction within narrative between plot and narration[2]. The plot is, if we put it bluntly, "what happened"; the narration is, on the other hand, the way these events are presented[3]. As a result, we can say that every version of Cinderella has the same plot, but a different narration.

As we have shown in our introduction, the telling of a story relies on the manipulation of time and space. At the simplest level, this is just cutting out the boring bits - we don't describe every event of Cinderella's night at the ball (or every inch of the ballroom), nor do we describe every day that the prince spends searching for Cinderella. In each of these cases, we simply move from one significant detail to another.

This is, in many ways, the secret of storytelling. The storyteller knows that if they include the important details, then the story will make sense to the

---

[2] Alternative words used by other writers for these concepts include story and discourse.

[3] The use of the term "narration" does not necessarily imply the spoken word - it is used for all of the techniques used by the storyteller.

listener; similarly, the listener knows that if a detail is included, it is significant.

In videogames, as opposed to film, there is continuity of time and space. Continuity of time and space in videogames is evident in that the action always takes place in a single environment within a single time frame, unlike film which compresses or expands time, or where differing moments in time may be presented within the same sequence, and where the action can move from one (far distant) place to another with one cut.

Videogames present the user with a continuum of space - one normally moves from one place to another in quite a pedestrian fashion (not much teleporting here) and most often the user will have to retrace all their steps to go down a different corridor, alley, path etc. This is unlike film where a cut will suffice to transport the action back to the starting point.

The action in most videogames takes place in the present - there is no flashback (or flash forward) or parallel story. The user lives the game in a form of "real time" (the time of the game) from which there is no escape other than by reaching the end of the game or through a decision to exit the game prematurely.

We believe that this use of the continuity of time and space in videogames makes it difficult for the techniques of storytelling outlined previously (i.e. the manipulation of time and space combined with the emphasis of

significant detail) to take place - techniques which we believe are essential for the existence of an engaging narrative experience in screen-based media

Once a game adopts the continuity of time and space, the only way left to tell the story is through the design of the world and of the objects within that world. The world of the game ceases to be an environment for the narrative and, in a very real sense, *becomes* the narrative. When we enter the world of *Doom* or *Quake*, there is nothing to tell us where we are or what to do. Unsure, we simply explore and find that there are monsters. Our mission becomes clear: to survive and escape. We keep moving forward and find a locked door - a secondary mission becomes obvious: we must go back, retracing our steps, and find a key....

What is apparent here is that no matter how large and open the terrain for these games, it is always just a maze and the design of this maze is the design of the story. There are no climaxes in the narrative of *Doom*, in the sense that the climax of *Star Wars* represents both an emotional and a narrative climax to that film - there are only rooms with a lot of monsters, and the climax to the game is just the room with the most monsters.

In a film, the director creates suspense by showing the viewer what is waiting ahead to ambush our hero (who remains unaware of the danger). Once a game insists that all the action must be seen from one angle, in one space and in one time frame, this type of storytelling is impossible. There may be suspense in *Quake*, but this comes from a familiarity with the

design of the game: once a player has been ambushed in a certain type of location - say an exposed bridge across a pool, with no escape route - they will approach similar locations with greater tension and suspense and greater caution.

The design of objects within the world - and their placement - is another storytelling technique. The reason that we can spot a key in *Quake* is because it is out of all proportion with its environment and floating in mid-air - it is not like in a film, where we would just show a realistic key, lying on the floor, in close-up. Similarly, when an object is important in a videogame, it is prioritised simply by being the only interactive object in a room.

In essence, these games show a symbol of a key rather than a key; the symbol of a switch, rather than a switch. This creates a very real tension within these games between realism and artificiality, and places the designers of these games on the horns of a dilemma as they attempt to create more meaningful and sophisticated narratives. If the designer chooses to increase the narrative content of a world by including a greater number of interactive objects, rendered in greater detail, then the ability to prioritise them - simply on the basis of design or position - decreases.

For instance we are able to find the clues in *Myst* because there are so few objects in the world: everything that is interactive is a clue (or part of a puzzle). We can find the clues hidden in the books because there are only a

dozen or so books in the bookcase (and those that aren't burnt have only a handful of pages). If we try, however, to create a realistic world - with a realistic amount of books in the bookcase - the clues would simply get lost.

When we provide a greater number of interactive objects (rendered in greater detail), the significant details of our narrative get lost in the "noise" of realism - the storytelling gets diluted and the player is no longer able to sense the connection between the significant objects (and thus follow the story). They may even be unable to tell the significant objects from the non-significant ones.

Videogame designers have chosen the convention of continuity of time and space, and this means that they are unable to emphasise objects significant to their narrative through the storytelling techniques outlined above. They have therefore adopted other methods for prioritising these significant objects. These include the design of the worlds and the objects within them, the sparseness of these worlds, etc. We can think of these as alternative storytelling techniques.

As we have shown in our introduction, the filmmaker has a wide range of techniques at their disposal, and this gives them a great deal of control over the meaning that they present. By locking themselves into the convention of the continuity of time and space, the techniques available to the game designer are, by comparison, extremely limited and lacking in subtlety. Thus it is difficult for the designer of the game to tell a sophisticated narrative.

It is therefore important to see that the lack of a complex, sophisticated and meaningful narrative in videogames is not a consequence of sparseness within the worlds, or of the way in which certain objects have to be made to make them visible to the player. In stead, it is a result of the use of the convention that the game adopts: the continuity of time and space. This is an important detail as it means that we cannot produce better narratives merely by adding more objects - if anything, this will reduce the ability of these games to tell stories.

As the use of the continuity of time and space is such a limitation on the designer of a game, we must ask why this convention is still so strong. To do this, we must look more closely at the relationship between viewpoint and identification, and the opposition that exists between immersion and narrative.

Viewpoint and Identification, Immersion and Narrative

At the Visual Narrative Matrix conference, held at the Southampton Institute in November 1999, we presented a paper entitled *Playing with Film Language*, in which we analysed the viewpoint used in three contrasting videogames (*Alien Trilogy*, *Tomb Raider* and *Resident Evil*) and used this to draw conclusions about the player's sense of identification and sense of being part of a narrative. The conclusion that we came to was that as the sense of immersion increases, the sense of narrative decreases,

and that this is a direct and inevitable result of the type of viewpoint used in the game.

Games such as *Alien Trilogy*, *Doom* and *Quake* use a first person viewpoint - that is to say, the user sees the action as if through the eyes of their character. In these, the player has a strong sense of identification - "I am the character" - but a weak sense of being part of a narrative. Their impression is that they are just exploring the world (and in most cases, just killing everything that they see).

In a game such as *Tomb Raider*, where the camera follows behind the player's character as it moves, there is a weaker sense of identification - it is no longer "I am the character", but rather "I control the character". The sense of narrative is slightly stronger, however, and we classed its as being "I control a character in a world".

In a game such as *Resident Evil*, which uses a static views of the action (cut together somewhat like a film), there is a stronger sense of a narrative but the player feels more like they are just watching the character in this narrative, and less like they are controlling them.

We would now like to clarify and classify these ideas. We now feel that what we referred to then as identification is, in fact, immersion - in a sense, what we did was use a film term for a virtual reality concept. The viewer of a film knows that they are sitting in the auditorium and that the character is

"in" the world of the film - yet in spite of this separation, they feel moved by the events on the screen and by what happens to the protagonist. This sense of involvement is not immersion, because the viewer knows that they are not - and cannot be - in the world of the film. In stead, it is identification - a vicarious sense of immersion. This is not to say that a form of immersion does not exist in film - the first person point of view shot provides this. Likewise, identification occurs in games (though this tends to be less strong than the sense of immersion).

The above conclusions lead us to propose the following device for analysing the relationship between narrative, identification and immersion: a triangle with empathy, narrative and immersion at the three corners. By immersion, we mean the sense of "being there"; by narrative, we mean the sense of a story told. In the use of the word 'empathy', we wish to encompass both the sense of identification (as in film) and a sense of "emotional immersion" - in other words, caring about the characters in the narrative.

Triangle of narrative forces

Narrative

Immersion   Empathy
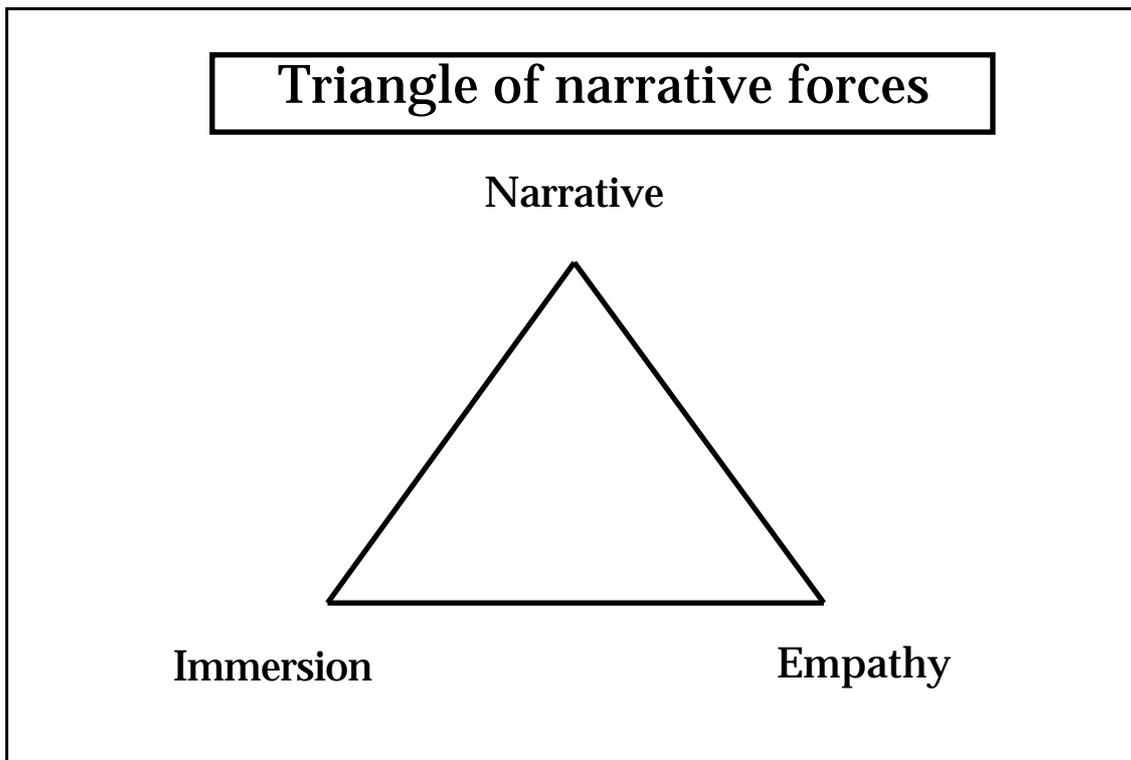
Viewpoint is the tool that the designer of a videogame, virtual environment or film uses to position the viewer/user within this triangle, but the closer they position them to one corner or side of this triangle, the further they are from the others.

A film positions the viewer predominantly towards the empathy/ narrative edge. The viewer of a film experiences a strong sense of narrative and cares about the protagonists in that narrative - characters who undergo evolution and transformation in the course of the story - but they have little real sense of immersion: they always know that they are in an audience watching the action on a screen.

Even so, the viewer of a film sometimes gets a sense of visceral thrill - of "being there". A big budget action movie will use a variety of techniques - including a first person viewpoint, hand-held camera, rapid editing, hyper-real sound - to give the viewer a sense of being part of the action: "a rollercoaster ride".

Most games and virtual environments position their user firmly in the immersion corner of our triangle. The user has a strong sense of "being" in the world of the game, but there is no engaging narrative, nor are they made to care about the characters in that narrative.

The sense of wonder and excitement that we experience when we enter a virtual environment and have a sense of "being there" - of immersion - cannot be underestimated. It has great seductive power. Yet we would argue that it is the power of immersion that has lead videogames to adopt conventions which prevent it from maturing as a medium.

Continuity of time and space has become a convention in videogames because it provides a quick and easy way to achieve this powerful sense of immersion. Yet it seems as though the designers of these games have gone no further than this.

Film does not rely upon the fact that it is a spectacle to entertain the audience - it also engages them in a story. In a way, it is misleading to refer to an action movie as a rollercoaster ride - even these films involve

characters (who we care about) who undergo transformation through the course of the action (and as a result of the action).

A videogame, however, relies solely on the "fairground trick" of immersion. The events in an action movie provide a backdrop to the evolution of the protagonist - and ideally form a parallel to this evolution; the events in a videogame are, by contrast, merely tests of skill that allow the player to progress further into the game. They do not lead to the evolution of the player or of their character. Indeed, it is very difficult to perceive the events in a videogame as a narrative, rather than just as isolated challenges.

Conclusions

In the essay *Cyberspace: First Steps*, **Michael Benedikt** says that:

> [I]n patently unreal and artificial realities such as
> cyberspace, the principles of ordinary space and time can,
> in principle(!), be broken with impunity. After all [....] the
> modern worlds of fantasy fiction, movies, and cartoons,
> are all replete with violations of the logic of everyday
> space and time: disappearances, underworlds, phantoms,
> warp speed travel [....] But let us notice two things: first,
> that there is a limit to how frequent and severe such

transgressions can become before credibility, orientation

and narrative power begin to be lost; and second, that

myth and fiction do not contain violations of ordinary

spatiotemporal logic but *descriptions* of such violations.[iv]

(exclamation mark and italics his)

This quote is informative for a number of reasons. Benedikt is right to point out that film presents descriptions of violations of space and time, but omits two very important details. The first is that he confuses *what* is presented with *how* it is presented. A film may cut together two shots taken in different places and times to describe (say) a teleportation, but in a very real sense, this cut is also a transgression of the rules of space and time.

The second detail that he omits is that film has an established set of conventions to allow these distortions of space and time to be understood by the audience. The audience knows what is happening in a film such as *Pulp Fiction* or *Reservoir Dogs*, with their complicated structure of flashforwards and flashbacks.

In her book, *The Pearly Gates of Cyberspace*, Margaret Wertheim says that:

So conditioned are we moderns to think of space as a

continuous *all-encompassing three-dimensional void* that

it is difficult for us to imagine any other view.[v]

(italics hers)

In a sense, this is the core of the problem: we think of the virtual environments that we create - or the world inside the videogame - as a space, rather than a storytelling medium.

The language of film (or any other visual medium, for that matter) is constructed on the basis of an agreed set of conventions between the makers and their audience: the director uses a shot because they know that it conveys a certain meaning; the viewer likewise knows what the shot means and so understands the message. Videogames have not yet developed as rich or as subtle a shared vocabulary of *storytelling* conventions as film.

It is also important to realise that what constrains present videogames and prevents the evolution of this storytelling language - the continuity of time and space - is just another convention. The "world" of these games (and everything within it) exists solely as a 3D model within the computer, to be rendered on the fly. This model can be viewed and rendered from any angle - in other words, our virtual camera can be anywhere with equal ease. This means that the same 3D engine used to produce flowing tracking camera moves could also render cuts and close-ups.

In essence, what we are arguing for in this paper is more authorial presence. As we have shown with our triangle of narrative forces, you cannot use a first person viewpoint (and continuity of time and space) to provide a strong sense of immersion and still expect to tell a strong narrative and move the player emotionally - this is impossible. We feel it is essential for

the designers of videogames (and of virtual environments) to take responsibility for the narratives that they are trying to tell, in other words, to become "authors" of their work.

A powerful and moving narrative will not emerge magically out of adding more detail to the worlds. Indeed, as we have shown, the conventions which make storytelling currently possible in videogames - the sparseness of the worlds and artificiality of the objects within them - means that this approach is counterproductive.

In stead, we suggest that videogames draw upon the conventions of film, modifying them where appropriate. This seems natural given the fact that they are both screen-based (and time-based) media and both tell their stories predominantly through images rather than through text or dialogue; the language of film is also well established and widely understood.

When we guide Lara Croft across a narrow bridge in *Tomb Raider*, we get no sense of danger or of nervousness. Now imagine the effect of showing a close-up of her foot stepping onto the bridge, a close-up of one of the ropes holding the bridge stretching and snapping, a shot of some bits of the bridge falling off and into the chasm. In the same way that the shots described by David Mamet describe "great alertness", these shots describe "a weak bridge".

We believe that it is only through an understanding of the triangular relationship between narrative, immersion and empathy - and of the way that the conventions of film language can be used to position and move the viewer within this triangle - that videogames can become a powerful storytelling medium in their own right, capable of engaging a wide audience (of all age groups and both sexes), and truly claiming to be a "grown up" medium.

References

[i] David Mamet, "Storytelling" in *On Directing Film* (London: Faber and Faber, 1992), p. 3

[ii] Brenda Laurel, *Computers as Theatre* (Reading, Massachusetts: Addison Wesley, 1993)

[iii] Janet H. Murray, *Hamlet on the Holodeck - The Future of Narrative in Cyberspace* (Cambridge, Massachusetts: MIT Press, 1997

[iv] Michael Benedikt, "Cyberspace: Some Proposals" in Michael Benedikt (ed.), *Cyberspace: First Steps* (Cambridge, Massachusetts: MIT Press, 1997), p. 128

[v] Margaret Wertheim, *The Pearly Gates of Cyberspace - A History of Space from Dante to the Internet* (London: Virago, 1999), p. 97