D 5.5

EVALUATION OF CINEMATOGRAPHIC SOLVER

Project Number	FP7-ICT-231824
Project Title	Integrating Research in Interactive Storytelling (NoE)
Deliverable Number	D 5.5
Title of Deliverable	Evaluation of Cinematographic Solver
Workpackage No. and Title	WP5 - Cinematography
Workpackage Leader	UNEW
Deliverable Nature	Report
Dissemination Level	Public
Status	Final
Contractual Delivery Date	30th June 2011
Actual Delivery Date	31st December 2011
Author(s) / Contributor(s)	Guy Schofield (UNEW), Stephen Lindsay (UNEW), Patrick Olivier (UNEW), Marc Christie (INRIA), Christophe Lino (INRIA)
Number of Pages	29

DOCUMENT HISTORY

Version	Date	Comment	Author(s)
1.0	30/12/2011		UNEW: Guy Schofield, Stephen Lindsay and Patrick Olivier; INRIA: Marc Christie and Christopher Lino

LIST OF ABBREVIATIONS

CGI	Computer Graphic Imagery
IS	Interactive Storytelling

TABLE OF CONTENTS

1.0 Introduction

As discussed in D5.1, filmmakers directing and editing narrative movies refer constantly to a standard vocabulary of cinematographic techniques. These conventions draw upon decades of experimentation and viewer familiarity and make up a widely understood cinematographic language which can be used to communicate subtle and complex information about the plot, the characters and the world they inhabit [Arijon76]. Some of these conventions are intended to expose the spatial and temporal configuration of the action. Techniques such as parallel editing and reverse camera configurations are designed to describe to the audience the space in which the action takes place and allow the viewer to easily distinguish between characters and follow dialogue. Others are used to impart information about the psychological and narrative make-up of the scene, suggesting character relationships, atmospheres and psychological states.

Throughout this project, we have continually referred to a particular scene from a well-known and critically acclaimed film, '1984' directed by Michael Radford [Radford84]. To recap, this film was chosen to serve as a test-bed and reference point throughout the development of the real-time camera control system for several reasons. The first of these was the subtle and understated nature of the scene. Very little occurs in terms of action on the part of the characters. The dialogue relates little to the plot, being a lunchtime conversation between friends: the characters simply sit, eat and discuss their work. The principle purpose of the scene seems to be to introduce the characters and how they relate to each other, establishing narrative threads which are developed throughout the remainder of the film. Much of this information is imparted through the way in which the scene is shot and lit. When we examined the original scene, subtle uses of the film frame emerged which seemed largely independent of the staging of the scene (these discussed at length in D5.1). Over the course of the Director Volumes system's development, we drew inspiration from these techniques and attempted to incorporate them into the system in a way that would allow enough flexibility to support a variety of directorial styles and narrative contexts and crucially could be transplanted from linear cinema and implemented in real-time computer graphic settings.

Figure 1: A still from Radford's 1984

This document represents an evaluation of the system both in terms of its ability to direct and edit scenes in a way which allows viewers to read complex narrative, psychological and spatial information about the scene and in terms of its success in altering narrative content through the adjustment of character relationships through cinematography. The document describes a user study in which volunteers were shown different versions of the scene, shot using different configurations of the system. These versions were compared to the original Radford version. Included is an analysis of the results of this study and a discussion of its implications for further development of the system.

1.1 The System

The real-time camera control system developed as part of the Cinematography Work Package of IRIS, draws upon the filmmaking conventions described above and applies them to the control of virtual cameras in 3D computer graphic environments. It aims to allow hands-off direction of Interactive Stories by automatically placing cameras in configurations which mimic cinematographic conventions and draw upon their descriptive and connotative power.

Much of the system is devoted to avoiding occlusion and placing cameras in the scene in such a way that conventional framing of characters (close-ups, apex shots etc) can be attained without interference from objects or other characters. This part of the system is described and evaluated in depth in D5.4. However, the system also attempts to communicate subtle information about the relationships between characters in a scene through further adjusting these cameras, once they have been placed. This filtering stage allows the author to specify rules relating to narrative dimensions. Currently the system is able to adjust for levels of affinity between characters, dominance of one character over others in a scene and isolation of a particular character[Lino10].

This adjustment is achieved through altering the characters' eye-lines and positions in the frame, replacing neutral values (such as symmetrical character placement in dialogues) with specific ones. Dominance, for example, is enforced by changing the height of the dominant subject's eye-lines in external shots, and using low angles in point-of-view shots of the dominant subject (see Figure 2, middle), causing them to appear larger and more important. Figure 2 shows examples of results for affinity, dominance and isolation in the 1984 canteen scene.

Figure 2: Shots generated using the system, highlighting Affinity (left), Dominance (middle) and Isolation (right)

Affinity.
Affinity between key subjects is accomplished by favouring balanced and symmetrical shots where similarity between key subjects can easily be established. Apex shots are preferred to external shots, symmetry is enforced in the framing and eye-levels are constrained at the same height both within shots and between shots. Reaction shots generally follow key subject utterances.

Dominance.
Dominance is accomplished by offering asymmetric views, in which dominant subjects appear larger, more often and for longer durations than dominated key subjects. As a consequence, external shots and subjective views of the dominant key subject are preferred over internal and apex shots. In terms of eye-levels, the dominant key subject gaze is generally higher than the dominated key subject. Another device that is considered is the use of high-angle shots of the dominated key subject with low-level shots of the dominant.

Isolation.
Isolation of a key subject is enforced by favouring shots displaying only that key subject (over apex and external shots), spending more time on him than other key subjects. In terms of composition, the key subject is generally framed within his environment (medium shot or long shot) with large empty space around or in front of him.

1.2 Challenges

As discussed in D5.1, certain challenges exist for IS authors and developers in simply transferring devices directly from cinema to real time computer graphic contexts. The languages of cinema have evolved in a specific context and filmmakers are usually able to anticipate the way their films will be experienced, whether in theatres or in viewers' homes. The most crucial of these differences is in the interactivity of the medium: film is traditionally a non-interactive, linear and stable media where devices of plotting and character development can be deployed by the director over time, secure in the knowledge that they will be experienced in sequence. This predictability affects not only the staging and action of the film but also allows fine control of cinematography and lighting. With characters moving between prearranged positions, cameras and lights can be placed to take best advantage of each new spatial and temporal configuration. In IS, this becomes more difficult as user decisions may affect the spatial arrangement of the entire scene.

Another important consideration in borrowing the languages of cinema for IS is the expressive capabilities of computer graphic characters (CGI) compared to human actors. The subtleties of trained actors' expressions and actions are difficult to transfer onto CGI characters, requiring a high degree of skill on the part of the animator. Low resolution characters may lack the necessary detail to be able to manifest shades of emotion as facial expressions. At the other end of the scale, Mori's[Mori70] 'Uncanny Valley' effect: one of strangeness or repulsion experienced by viewers encountering virtual characters, is a particular challenge to film makers and authors of IS, using high-resolution high quality models. Hollywood films such as The Polar Express and Tron Legacy offer interesting examples of the unsettling and distracting effects of using characters which appear almost, but not quite human.

 For the researcher or developer creating new systems for IS yet another problem exists. The experience of film is holistic, relying not on the separate function of each part of the medium but on action, staging, lighting, cinematography and sound working together. As most cinemagoers are aware, flaws in any of these factors can disrupt the entire viewing experience, either negating the effect of other factors or altering them at odds with the author's intention. This makes the development of test scenarios for new IS systems problematic. In our case, the success of a cinematography system affecting narrative dimensions also depends on a well realised plot, legible and aesthetically acceptable graphics and good sound. The omission of any of these factors may well radically affect viewers' readings of the scene.

2.0 Experiments

In order to evaluate the success of the system's camera solutions, we needed to examine how the results altered people's perception of the different relationships between characters. However we also needed to account for the impact of transferring the content of the film into CGI. To this end, we created a replica of the scene with cameras placed manually so as to closely mimic the original. High quality characters with lip-synch, commissioned earlier in the project were used within a 3D CGI environment which approximated the layout and atmosphere of the original film set. Edits, character actions and lighting were all rendered as closely as possible to the Radford scene and the original audio track was used. The scene was lit and rendered to video in 3DS Max. By comparing this version to the original, we hoped to identify and account for any changes in viewers' reading of the scene emerging from the transfer of a linear cinema piece with live actors to a CGI movie with virtual characters.

Next, for each narrative dimension we wished to test (dominance, affinity or isolation) we prepared two movies allowing comparison with the benchmark videos (the original Radford scene and our own CGI replica). In each case, the first of these comprised the scene with camerawork specified by

the Director Volumes system, incorporating adjustments to camera position, orientation and framing. This scene was lit and rendered in 3DS Max, using a standard 3 point lighting scheme, designed to light the characters evenly and similarly. Audio was provided once again by the original film's audio track.

The second video in each case was a more developed version where the camera positions were supported by manually adjusted lighting and a bespoke musical soundtrack intended to enhance the narrative dimension. Each movie was rendered using 3DS Max's Mental Ray renderer to allow integration of high quality indirect lighting and shadows, with techniques such as under-lighting and low-key backgrounds used to support the narrative dimension being addressed. To emphasise the intended effect even further, a musical score was composed and recorded using VST software instruments in Ableton Live.

The aim in each case was to both evaluate the success of the system's solutions and to establish how much effect different aspects of the video (i.e. soundtrack, cinematography, dialogue) had on the viewers' reading of the narrative dimensions under examination. However, we also identified a number of secondary goals. We were interested in whether the versions of the scene generated by the system would have a different narrative effect than the original Radford scene, distorting or changing viewer's perceptions of the plot. Lastly we were concerned what aspects of the video, (Soundtrack, camera, staging) did viewers identify as contributing to each narrative dimension.

2.1   Procedure

In preparing materials for the evaluation phase, we attempted to design a set of experiments which would allow both direct evaluation of the system's success in quantitative terms: that is to say, the success of its application of narrative dimensions - and that would also allow more qualitative results relating to the viewer's reading of the resulting movie. We designed an anonymous online survey relying on a combination of multiple choice answers and more open-ended text responses, all relating to the video versions of the scene we had prepared, also hosted online. This approach allowed easy anonymization and processing of the results. Results were collated through a series of spreadsheets allowing direct comparison of each viewing. Where possible, each multiple choice question was devised along a 7 point Likert scale, allowing simple conclusions of statistical significance to be extracted. A full list of the survey questions is included in section 7.

2.2   Design of the Survey

The survey itself consisted of 27 questions over 3 pages. A preliminary section was included to establish background information about each volunteer's level of familiarity with both the medium and the text of the scene, which we reasoned might have an effect on their reading of the videos. As the novel 'Nineteen Eighty-Four' is not only a widely read text but is studied in universities and schools, we realised that some viewers may have a far greater familiarity with either the film or the novel than others, which might affect the results. More broadly, we attempted to establish in the most simple terms whether viewers were film enthusiasts or not, which might affect their ability to read the subtle cinematographic effects we were attempting to implement.

To test the survey we ran a small pilot study involving 5 participants, viewing 3 videos each. The purpose of this study was to ascertain what size of sample was likely to generate statistically significant results and also identify any technical faults in the survey. After running several comparisons, we calculated that 10-15 views per video should generate the data we would require, which would likely necessitate multiple viewings per volunteer.

We secured 41 volunteers, from as wide a variety of backgrounds as possible. Their ages ranged between 21 and 65 and comprised 22 male and 19 female viewers. They came from a variety of backgrounds, including school teachers, administrators, art-gallery assistants, students and computing scientists. Most were native English speakers although also among the group were volunteers from Italy, Germany, Vietnam, Greece, Malaysia and Taiwan. We reasoned that diversity in the volunteer group's occupations and cultural backgrounds was important as it would allow a variety of readings of the videos which did not simply reflect the tastes of a single social group. These volunteers were recruited and contacted via email, each being asked to complete the online survey and watch 3 of the body of 8 videos, hopefully generating a total of 123 views, or around 15 of each video.

3.0 RESULTS

3.1 Background Information

Out of 41 volunteers, 38 completed the entire online survey, providing between 10 and 15 viewings of each video, with 3 leaving the online questionnaire incomplete. All 41 completed the initial background questions. The results were collated and each comparison was subjected to a Mann-Whitney U test to examine the statistical significance of the results. Of the 41, all professed to being either 'slightly interested or 'very interested' in film, with nearly ¾ of viewers 'very interested'. As we anticipated, many of the viewers were familiar with either the film or novel and a number had studied the novel either at school or in higher education.
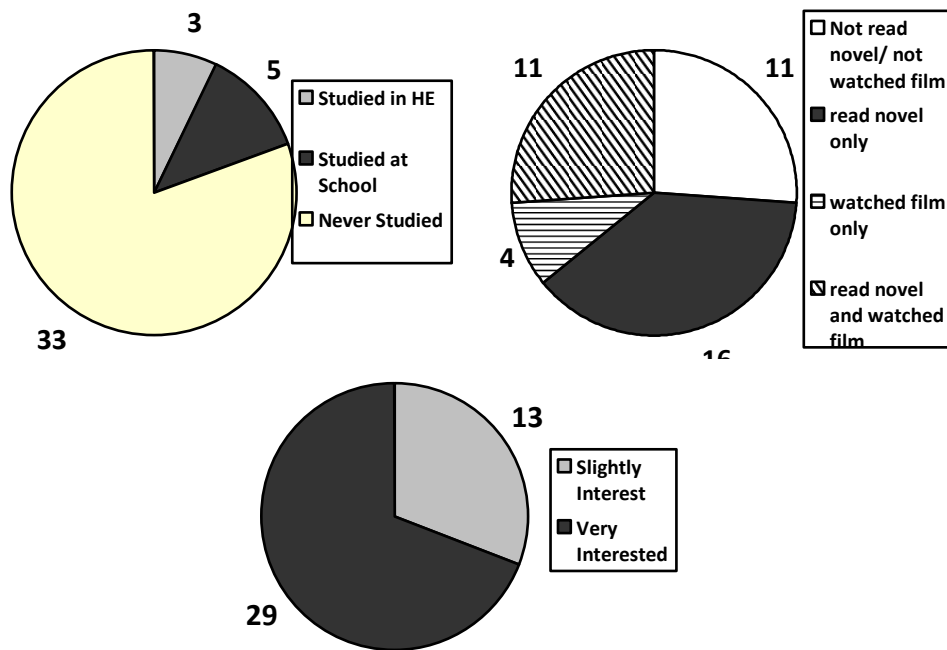
Figure 3: Viewers who had previously studied the novel (top left), viewers who had watched the film or read the novel (top right) and viewers' stated interest in film in general.

3.2 Experiment 1: Radford Scene Vs Replica

First to be examined was the comparison of the Radford scene to our replica. These two videos were compared with particular reference to differences in viewers' readings which might affect the second experiment. We anticipated that replacing the actors with CGI characters might have a detrimental effect on viewers' ability to read detailed emotional relations between the characters and that the differences in detail and lighting might influence viewers' perception of the film's atmosphere.

Dominance

Initially we compared viewers' readings of each narrative dimension in both videos.
First, viewers were asked to score on a scale of 1-7 their reading of each character's level of dominance in the conversation. The exact question asked was 'Rank how dominant each of the characters appears to be in the conversation'.
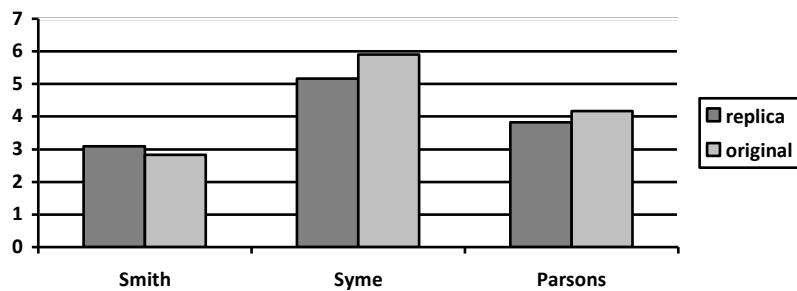
Figure 4: Dominance of characters in Replica and Original videos

Affinity

In this case viewers were asked 'On a scale of 1-7, how much affinity does each of the characters have for each other (with 1 being no affinity and 7 being a strong connection or liking for each other)'.
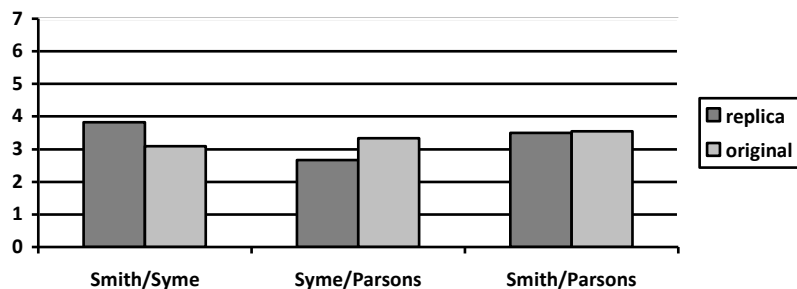
Figure 5: Affinity between characters in Replica and Original videos

Isolation

Here viewers were asked 'On a scale of 1-7 how isolated does each of the characters seem (1 being least isolated and 7 being very isolated)'.
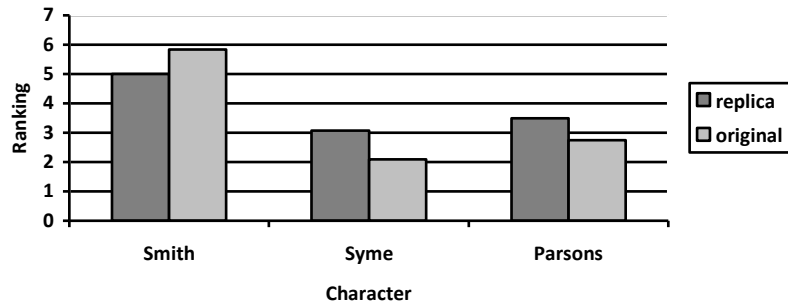
Figure 6: Isolation of characters in Replica and Original Videos

Initially these readings would seem to suggest a degree of similarity in viewers' readings of all the dimensions (see figures 4, 5, 6). The readings of dominance and isolation were broadly similar and more importantly, the order of dominance of the characters in the conversation was seemingly preserved. The same was true of isolation with Smith appearing the most isolated in both cases. In the case of affinity the ranking of connections in the conversation was slightly different, with Smith and Syme seeming the most connected in the replica but with Smith and Parsons appearing to have more affinity in the original. Perhaps surprising however, was the amount of disagreement between viewers of each video (see figure 7). In the readings of dominance in the two videos, viewers' responses covered the whole range of 7 possible answers. The responses were examined closely to ensure that no faults such as viewers mistakenly reversing the ranges had occurred, however nothing (such as rankings exactly contrary to expectation) was detected to suggest this.
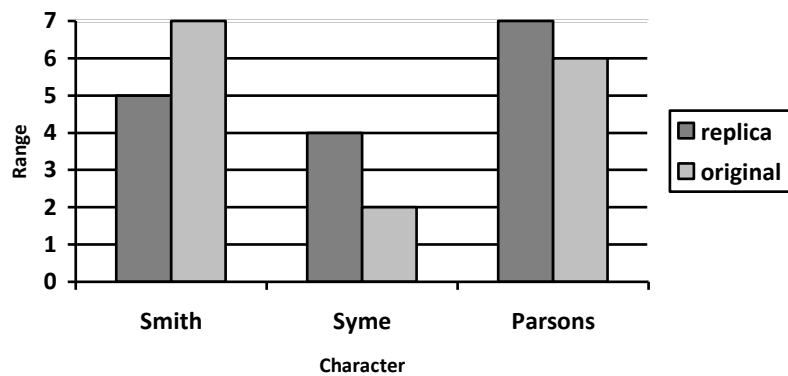

Figure 7: Range of responses from viewers in judging dominance

In order to try to understand the reasons behind viewers' decisions, we also asked them to rank how they thought different elements of the videos' production influenced their reading of the scene. We compared these rankings for each narrative dimension (see figures 8, 9, 10). As the replica video included the soundtrack from the original and the actions, camerawork and lighting also closely mimicked the original, we expected a high degree of similarity. We were also interested in whether these perceptions would differ significantly between dimensions.
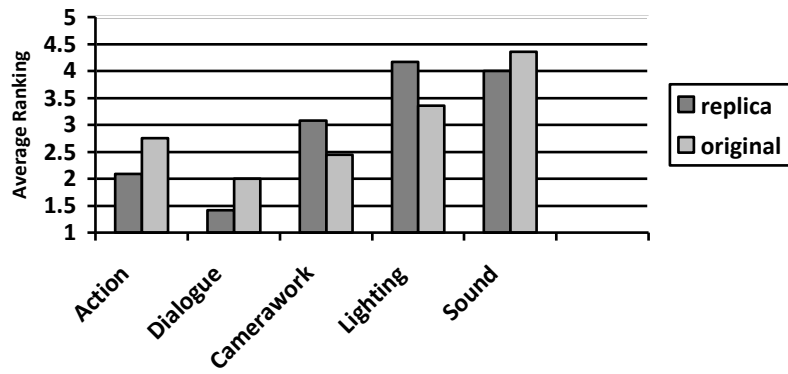
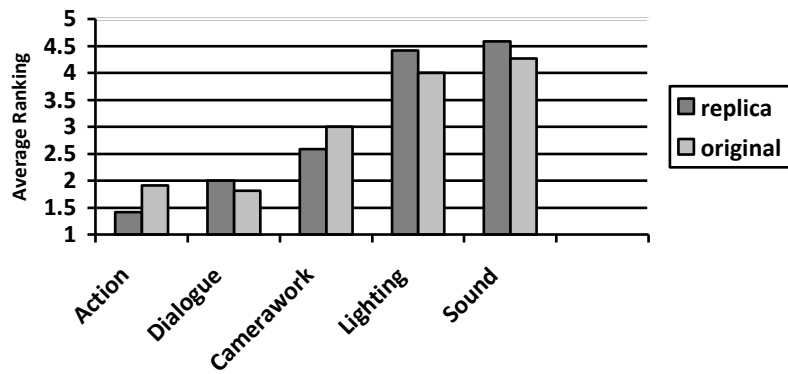**Figure 8: Ranking of the perceived importance of each factor in the reading of dominance**

Figure 9: Ranking of the perceived importance of each factor in the reading of affinity
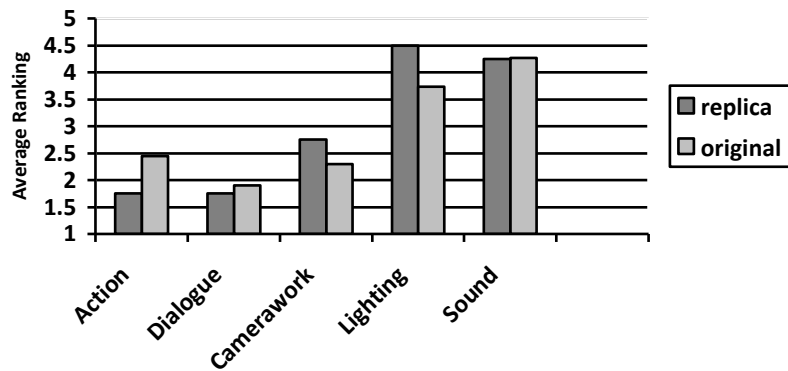
Figure 10: Ranking of the perceived importance of each factor in the reading of isolation

Encouragingly, the rankings of each factor were broadly similar over the two videos, suggesting that in all three cases, lighting and sound were ranked as least important, with action and dialogue most crucial in establishing character relations. These results (although based on a relatively small sample size) suggested that viewers would be able to read our CGI versions of the scene with similar results to the experience of watching the original. Interestingly, the range of responses varied far more greatly in viewers' perception of the original than in their viewing of the CGI replica and in both cases viewers' perception of the importance of camerawork varied enormously.
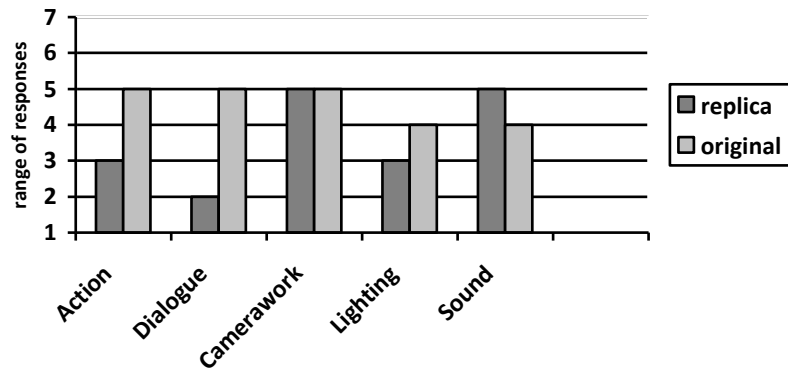
Figure 11: Range of responses in judging the perceived importance of each factor in the reading of dominance

3.3   Experiment 2: Multiple Videos

We next extended the experiment to include comparison of the videos generated using the Director Volumes system. Once again, each result was subjected to a Mann-Whitney U test to check for statistical significance. Throughout the survey, viewers were asked consistent questions on each video in order to allow examination not only of readings of narrative dimensions in the videos designed to emphasise them, but to allow comparison with the rest of the videos.

Dominance

As in the previous experiment, viewers were asked to score their reading of each character's level of dominance in the scene. Once again, the exact question asked was 'Rank how dominant each of the characters appears to be in the conversation'. The results from the videos designed to emphasise the dominance of the character Syme, were compared with a) the original Radford scene, b) our replica and c) the video designed to suggest affinity through camerawork. Our hypothesis was that the 'dominance camera' video would show an increase in Syme's dominance and the 'dominance camera/lighting/score' video would show a further increase, with the 'affinity camera' video showing less dominance than all the others.
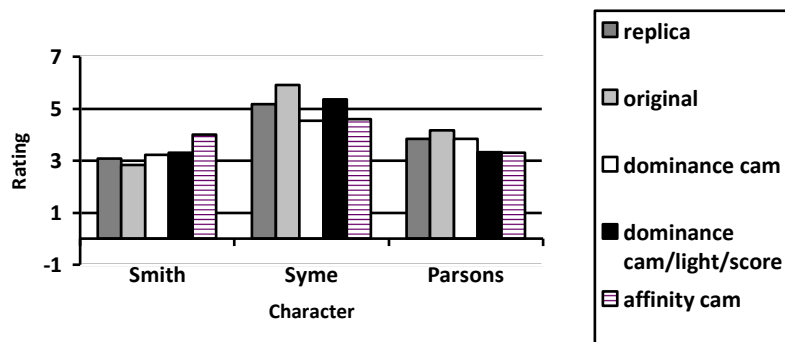

Figure 12: Dominance ratings in multiple videos

The result of the comparison once again showed a lack of conclusive statistical significance when subjected to a Mann-Whitney test, with no real increase in Syme's dominance over the original and replica versions of the scene (see figure 12). Indeed, the video with camera work for dominance actually showed a slight decrease. The ranking of dominance in the character relations was roughly preserved with Syme being consistently judged to be slightly dominant in the conversation. Surprisingly, the version of the video designed to show affinity between Syme and Smith also seemed to have little effect on viewers' perception of his dominance of the conversation.

Examining the range of responses, once again we found a large disparity in viewers' readings of each scene (see figure 13). The original scene showed slightly more agreement on each character's place in the conversation with the exception of the protagonist Smith, who attracted rankings of every level. The CGI movies however attracted a wide range of ratings on each character.
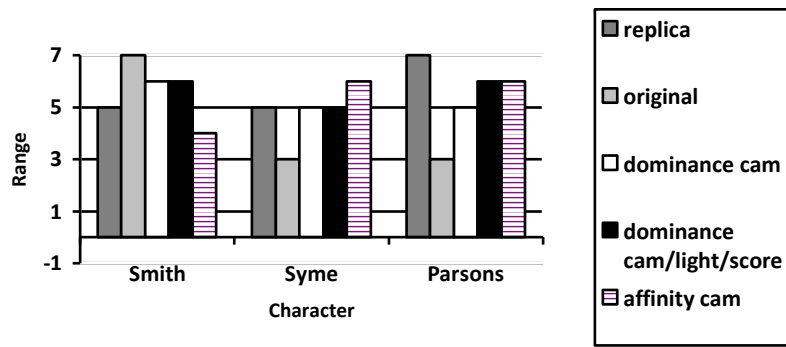
Figure 13: Range of responses in multiple videos

Upon examining the results of viewers' readings of the importance of each factor of the video production it was apparent that to some extent they agreed with the video comparison of the original and replica versions, with action and dialogue taking precedence over lighting and sound (see figure 14). Once again however, the range of responses was so great as to negate conclusions of statistical significance (see figure 15).
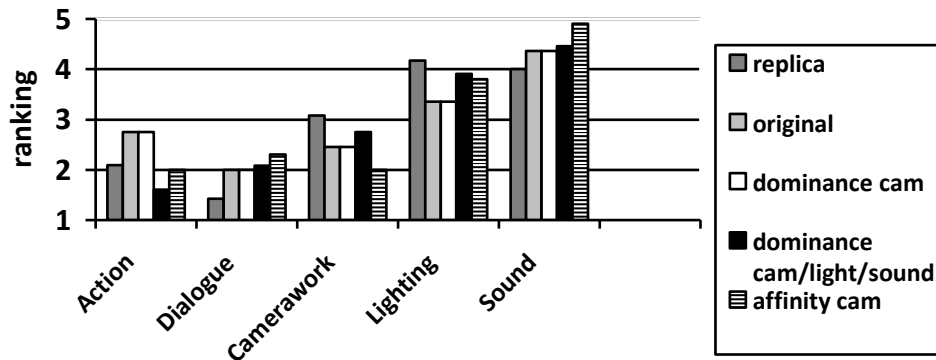
Figure 14: Ranking of the perceived importance of each factor in the reading of dominance
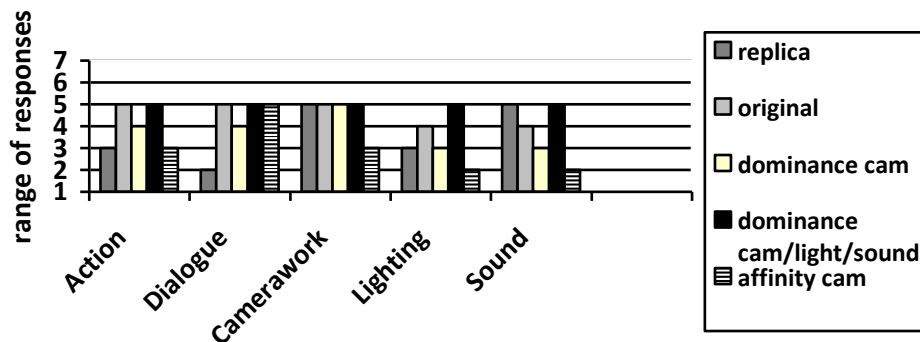
Figure 15: Range of responses in judging the perceived importance of each factor in the reading of dominance

Affinity

The analysis of the effect of attempting to heighten a sense of affinity between Syme and Smith was achieved in the same fashion as the previous experiment. In this case viewers were asked 'On a scale of 1-7, how much affinity does each of the characters have for each other (with 1 being no affinity and 7 being a strong connection or liking for each other)'. Videos compared here are the replica, original, a version with camerawork to suggest affinity and one that also featured supportive lighting and a musical score. Also included for comparison is a video to suggest the opposite: isolation of Syme.
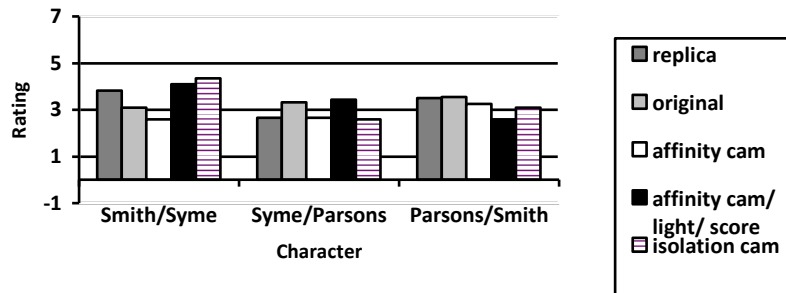

Figure 16: Affinity ratings in multiple videos
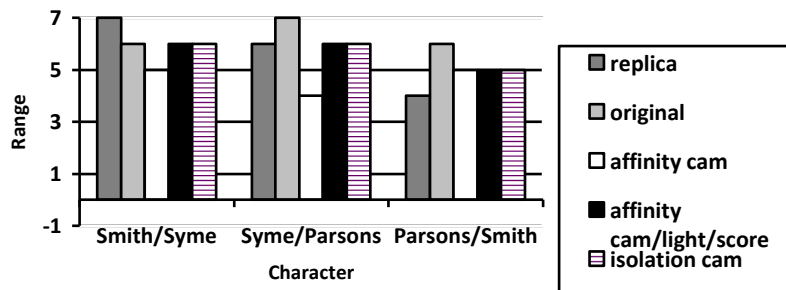

Figure 17: Range of responses in multiple videos

As can be seen from the results (see figure 16), there was a slight increase in the perceived affinity between Syme and Smith in the video designed to emphasise this with camerawork, lighting and sound, however, once again, the version with adjusted camerawork alone showed a slight drop in this dimension compared to the original and replica versions. When subjected to a Mann-Whitney test, these variations were too slight to accord a high level of statistical significance. The range of responses varied enormously from viewer to viewer, with visitors coming to very different conclusions, even on the original film (see figure 17).
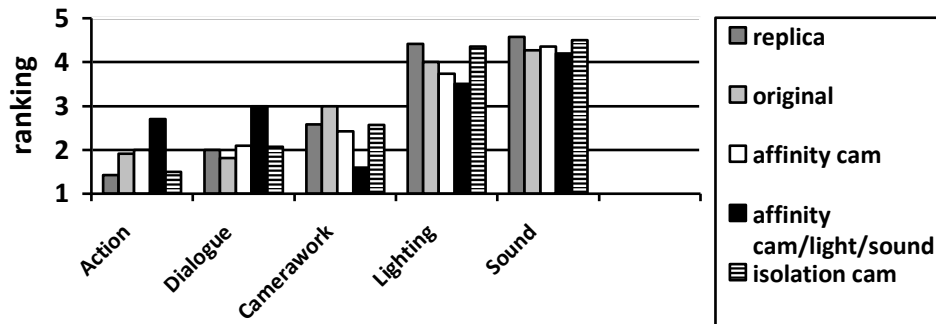
Figure 18: Ranking of the perceived importance of each factor in the reading of affinity

Once again, there was some agreement on the factors contributing to the characters' sense of affinity (figure 18). Interestingly, the dialogue seemed to be more influential in the video designed to emphasise affinity with lighting and sound, although the sound and lighting themselves were still rated as largely unimportant.
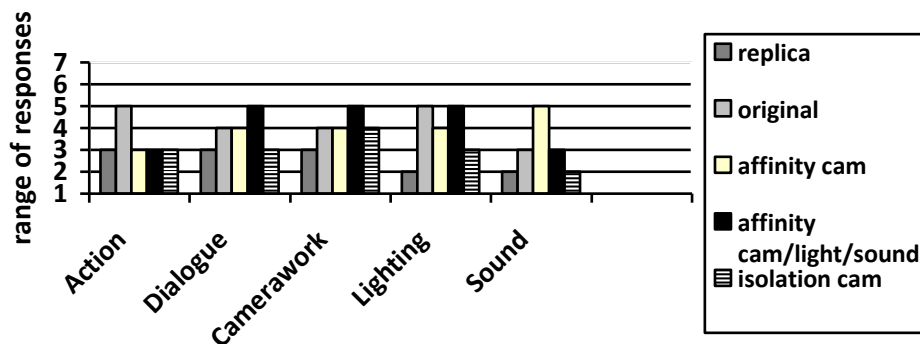

Figure 19: Range of responses in judging the perceived importance of each factor in the reading of affinity

The range of responses in judging the affinity of the characters was broader in viewers of the 'affinity camera/light/sound' video, especially with regard to the dialogue, camera and lighting (figure 19).

Isolation
The final dimension to be examined was isolation. Here viewers were asked 'On a scale of 1-7 how isolated do each of the characters seem (1 being least isolated and 7 being very isolated)'. Viewers were presented with videos generated using the Director Volumes system, one with the camerawork adjusted to emphasise Syme's isolation and another with camerawork, lighting and a soundtrack to further heighten this dimension. These were compared with the original and replica videos and with the 'affinity camera' video, the hypothesis being that these would provide very different readings.
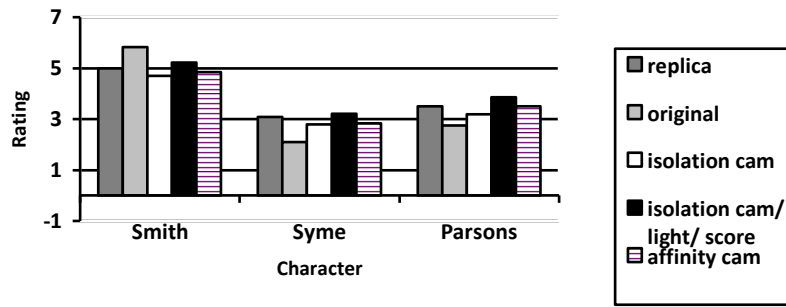
Figure 20: Isolation ratings in multiple videos

Once again, the responses were too close to provide a significant statistical difference with a wide range of responses for each video (see figure 20), however there was a slight increase in Syme's isolation over the original and the video with adjusted camerawork, lighting and sound also seemed to show Syme more isolated than the replica.
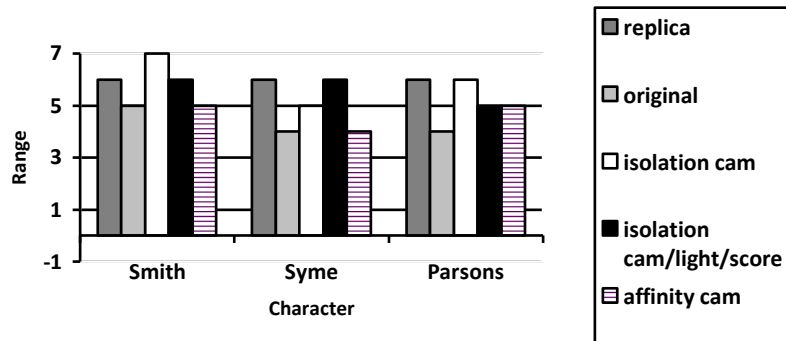
Figure 21: Range of responses in multiple videos

The factors which viewers identified as influencing their readings of the characters' isolation followed the pattern established in the other experiments with sound and lighting viewed as less important than action and dialogue (see figure 22).
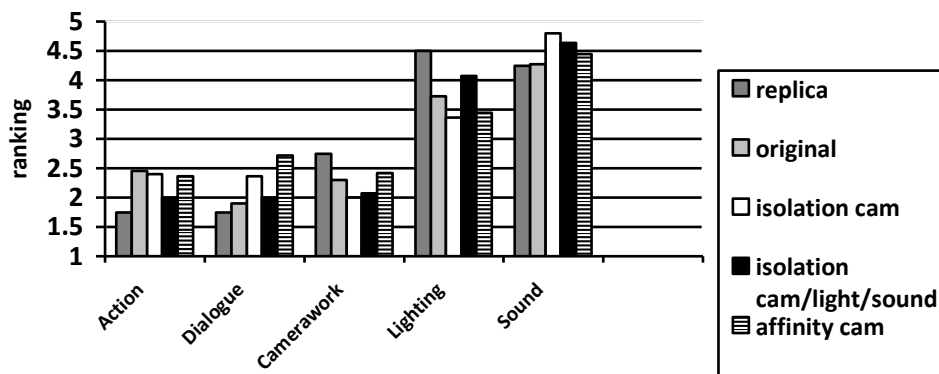
Figure 22: Ranking of the perceived importance of each factor in the reading of isolation

Interestingly the range of responses on the ranking of these factors was somewhat smaller from viewers of the isolation camera video than some others (see figure 23).

Figure 23: Range of responses in judging the perceived importance of each factor in the reading of isolation

3.4   Further Readings of the Scene

In addition to the multiple choice elements of the survey, we also asked viewers two open-ended questions concerning the content of the scene. These were placed in the survey immediately after viewers had watched the first video, before exposure to other ve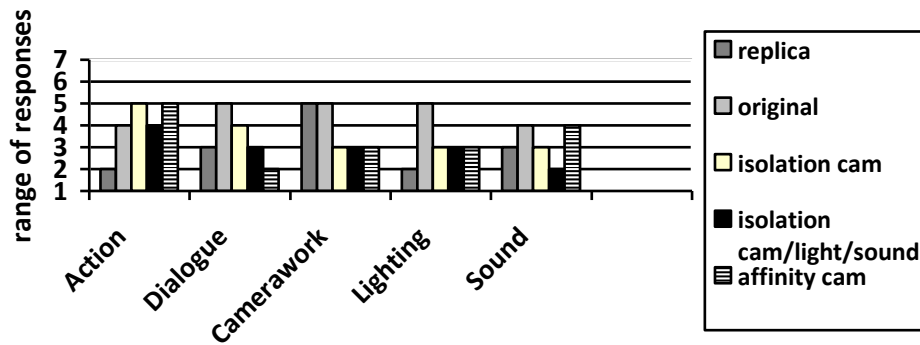rsions. The purpose of the questions was to ascertain in each case whether there were any significant differences in their readings of each video as a whole and to allow viewers to make observations on the content of the scene. The questions were as follows.

- 'In the box below, describe briefly what happened during the scene.'
- 'In the box below, describe the atmosphere of the scene.'

These questions were examined for more qualitative information such as identification of themes, speculation as to the emotional states of characters and instances of descriptive language relating to the characters or environment.

We anticipated that the conversion to CGI might have an effect on viewers' ability to read emotional connections between characters. However, we found that this was not the case. When asked to comment on the atmosphere of the scene, nearly all viewers were able to make strong statements on the emotional state of the characters. 36/40 responses contained statements relating to the characters' emotions or the emotional atmosphere of the scene. These included the following:

> *'A Very scary, intimidating atmosphere. People seem to be afraid of talking'*
> *'A Smith was creating some tension and awkwardness, like wariness but Parsons balanced this out a little.'*
> (Both from responses to the 'dominance camera' video).

27/32 of the viewers who watched the CGI videos mentioned unease, tension or fear in describing the characters' relationships while 7/8 who watched the original film made similar observations. Interestingly, 1 viewer of the original film made the opposite observation, commenting that the scene was,

> '*not very emotional or dramatic, just a normal chat as I would envision it to happen in a hospital'.*

A number of viewers also made reference to complex emotional states. Examples included:

> *'Although everyone is clapping, they do not seem happy'* (dominance camera video)
> *'It seems quite tense as everyone is aware that they are being watched and it seems that Syme is just saying what he thinks he should be'* (dominance camera/lighting/sound video)

Encouragingly, many of these statements originated from viewers who had watched the versions directed by the Director Volumes system. There was little deviation from the types or number of statements made by viewers of the original, replica or system-generated versions and there seemed to be no contradictions in viewers' understanding of the spatial or narrative configurations of the scene.

4.0 Discussion

4.1 General Performance

The survey we conducted was designed to explore and evaluate the success of the Director Volumes system in its ability to direct scenes in a way which would allow viewers to understand plot, spatial arrangement and character relations. Our inclusion of the original film scene in each part of the test was an attempt to continually gauge the success of the system's solutions in comparison to an example of 'best practice': an excerpt from a critically acclaimed film featuring well-known and well-respected actors.

The most surprising feature of the results was the wide range of responses to each question in the survey and the general lack of agreement on subtle character relations, even with regard to the original film; however the results of our comparisons in terms of legibility and atmosphere were encouraging. Viewers were able to gain the same highly detailed information about plot and characters from the Director Volumes system's solutions as from the replica and original versions of the scene. When asked to comment on the content and atmosphere of the scene, there was little discernible difference in responses whether viewers were watching the original or Director Volumes systems' versions. At the very least this suggests that the system was able to provide credible and legible camera positions, framings and edits which did not differ from viewers' expectations in a way which disrupted their ability to read the scene.

4.2 Performance of Narrative Dimensions Filters

The results of the system's ability to affect narrative dimensions through adjusted camerawork were less conclusive. The wide range of responses to each question suggested that viewers were coming to very different conclusions about the relationships between the characters. In each case, there was some agreement that the most important factors in determining character relations were the action and dialogue with camerawork being ranked as slightly less important. Interestingly, this phenomenon was not confined to viewers' readings of the CGI version but was also pronounced in responses to the original. This perhaps suggests either limitations in the viewer group's abilities to discern these factors or flaws in the design of the study itself rather than problems with the system's solutions. A number of factors may have contributed to this.

Content of the Scene

In considering the results of the survey it is important to bear in mind that none of the videos could be described as a 'control', 'neutral' or 'null' rendering of the narrative dimensions we were examining. The staging, characters and dialogue were the same in each video and although the camerawork and lighitng was altered in different videos, the basic content of the scene was the same. We chose the 1984 scene because we believed it offered a scenario in which the narrative content was sufficiently ambiguous to provide leeway for the cinematography system to make changes, however, as borne out by the opinions of viewers in each experiment, dialogue and action contributed enormously to their perception of the characters' relationships, effectively rendering any version we could produce, already somewhat loaded.

Exposure

A factor which also may have had an effect was the amount of exposure to the scene experienced by the participants. To keep the sample to a manageable size, we relied on each participant viewing 3 videos in fairly rapid succession. It is possible that a much larger study where subjects watched only

one video might have slightly different results, as in our survey, subjects were surely becoming familiar with the text of the scene by the third viewing.

Familiarity with the Text.
All of the viewers in our study professed at least a slight interest in film in general and nearly ¾ had either read the novel 'Nineteen Eighty-Four' or watched the film. Several viewers referred to details of the novel which were not mentioned in the scene (for example viewers referred directly to The Ministry and big Brother, themes not mentioned directly in the videos), displaying a familiarity with the plot beyond the short videos we showed. It seems likely that this may have introduced a slight bias to the viewers' perceptions of the narrative and character relationships in the film.

Sample Size
Our choice of sample size was based on a small pilot study which yielded results which suggested promisingly conclusive statistical outcomes. Each video was viewed a maximum of 15 times, consequently a small number of outlying responses might have had a disproportionate effect on the results. Increasing the sample size might have increased the statistical significance of the results. However, the wide range of responses to many of the questions in the survey suggests that even a vastly increased sample size may have had little effect in generating useful results.

Mode of Viewing
Due to the technology we used to host the survey, it was not possible to time each participant's progress through the videos and questionnaire. As we allowed subjects to pause and rewind each video at will, the viewing experience of each subject was slightly different. It would have been interesting to explore whether viewers who took this opportunity came to different conclusions about the narrative than others who simply watched each video once from start to end. We also allowed a degree of freedom in how volunteers watched the videos. Anecdotally, we are aware that some watched at home, others at their desks at work etc. Screen size, sound quality and environment were all factors beyond our control in this instance and may well have had some effect.

5.0 Conclusion

Ultimately the results yielded from the survey are encouraging, especially in terms of the system's ability to direct scenes in a way which allows complex narrative information to be imparted to the viewer. The survey has also generated useful information to continue development of the system, especially with regard to the narrative dimensions filters.

Our attempts to evaluate the Director Volumes system met with challenges typical of media research of this type. The indivisibility of different factors such as content, production, viewers' backgrounds and context made isolating and quantifying subtle differences in viewer experience extremely challenging. Decoding complex character relations and plot information from a short video clip is not a trivial task and the factors discussed above may all have contributed to the wide range of responses submitted by viewers to many of the questions.

An important factor, again relating to the inter-connectedness of factors in considering cinema was the conditions under which viewers watched the videos. The conceptual framework of the Director Volumes system is intended to be flexible in terms of integration with different types of interactive narratives, the specification of which fall largely outside the scope of this part of the project. However, to compare viewers' readings of the narrative effect of the system's solution, we had to create a replicable test which could take into account multiple viewings of the same solution, hence our reliance on pre-rendered video. Based on similar research [Vorderer11] we are aware that

viewers experiencing the effect of the Director Volumes system in the context of an Interactive Story where they are actively engaged, might have a very different experience of the narrative than our subjects who were exposed to a linear rendering of a scene.

5.1 Approaches for further studies

Having analysed and discussed the results, we were able to identify a number of factors to consider for further evaluation and development of the system, which might assist in overcoming the challenges of this type of survey. A simple measure which might eliminate hard-to-detect biases on the part of the viewers would be basing the test-bed scenario on a less familiar text. Unfamiliar characters and plot might well offer a more controlled environment for exploring viewers' understanding of the scene. The 1984 scene was selected at the beginning of the project as it offered a compelling and unusual plot and characters and the lack of dynamic action provided an opportunity for cinematography to come to the fore. In selecting an alternative, these considerations would again be important: a difficult task would be identifying a text which was unfamiliar to viewers but which provided well-defined and interesting characters, scenarios and action.

Controlling the circumstances in which viewers experienced the video might also have some effect on the outcome of the results. It is possible that distraction or lack of attention on the part of viewers may have contributed to the wide range of responses. In designing the survey we attempted to provide an online solution which could be easily completed by volunteers at their leisure, however providing a quiet space free from distraction with high quality audio/visual equipment might eliminate any uncertainty surrounding the quality of viewer's responses.

A more wholly qualitative approach might allow a deeper understanding of how viewers form their impressions of the dimensions we were attempting to affect and might generate insights into the reasons why certain questions generated such a wide range of responses.

Lastly, a good way of more accurately verifying results might be to engage in a far more extensive survey, greatly increasing the sample size. However, as discussed, the ranges of results suggest that simply increasing viewer numbers in isolation might serve only to increase the logistical overhead of the study without yielding more conclusive results. It seems likely that only by extending the survey itself, with more detailed and in-depth questioning and by eliminating factors such as viewer distraction and bias from the text could we more definitively prove success or failure on the part of the narrative dimensions filter.

6.0 References

[Arijon76] Arijon, D. *Grammar of the Film Language*. New York: Hastings House Publishers. (1976).

[Radford84] Radford, M. (Director). 1984 [Motion Picture]. (1984).

[Mori70] Mori, M. The Uncanny Valley, *Energy,* 7(4), pp. 33-35 (1970)

[Lino10] Lino, C., Christie, M., Lamarche, F., Schofield, G., and Olivier, P. A Real-time Cinematography System for Interactive 3D Environments. *In Proceedings of Eurographics/ACM SIGGRAPH Symposium on Computer Animation.* (2010)

[Vorderer11] Vorderer, P., Klimmt, C., Roth, C. and Vermeulen, I.E., 2011. The Audience Embedded in the Narrative: An Experiment in Interactive Storytelling. *In Proceedings of the Conference of the International Association for Media and Communication Research (IAMCR) - "Audience" Section, Istanbul, Turkey,* (2011).

6.0 Appendix: Survey Questions

Included here is the full text of the survey questions. These questions were built into an online survey hosted at http://www.surveymonkey.com

1. Have you previously seen the film '1984', directed by Michael Radford?
- No
- Yes, once
- Yes, more than once

2. Have you previously read the novel 'Nineteen Eighty-Four' by George Orwell?
- No
- Yes, once
- Yes, more than once

3. Have you studied either the film '1984'or the George Orwell novel 'Nineteen Eighty-Four' at school or during higher education?
- No
- Yes, at school
- Yes, in higher education

4. How interested are you in watching films in general?
- Not at all interested
- Not very interested
- Slightly interested
- Very interested

5. We'd now like you to watch the first video in the email we sent you and answer a few brief questions. Feel free to pause, rewind or watch the video more than once if you like. The video depicts a conversation between three characters named below. Which video did you watch?
- A
- C
- D
- F
- G
- I
- J
- K

6. In the box below, describe briefly what happened during the scene.

7. In the box below, describe the atmosphere of the scene.

8. Rank how dominant each of the characters appears to be in the conversation
- Very Subservient
- Quite Subservient
- Slightly subservient
- Neither subservient nor dominant
- Slightly dominant
- Quite dominant
- Very dominant

9. What led you to this decision? Please rank the aspects of the video which most affected your reading of which character was dominant, with 1 being the most important aspect and 5 the least.
- Action
- Dialogue

- Camerawork
- Lighting
- Sound

10. On a scale of 1-7 how isolated do each of the characters seem (1 being least isolated and 7 being very isolated).
- 1(Not at all isolated)
- 2
- 3
- 4
- 5
- 6
- 7 (Very Isolated)

11. Again, what led you to this decision? Please rank the aspects of the video which most affected your reading of which character was isolated, with 1 being the most important aspect and 5 the least.
    - Action
    - Dialogue
    - Camerawork
    - Lighting
    - Sound

12. On a scale of 1-7, how much affinity does each of the characters have for each other (with 1 being no affinity and 7 being a strong connection or liking for each other).
- 1(No affinity)
- 2
- 3
- 4
- 5
- 6
- 7 (Strong Affinity)

13. Again, what led you to this decision? Please rank the aspects of the video which most affected your reading of which characters had an affinity for each other, with 1 being the most important aspect and 5 the least.
- Action
- Dialogue
- Camerawork
- Lighting
- Sound

14. We'd now like you to watch the second video in the email we sent you and answer a few brief questions. Feel free to pause, rewind or watch the video more than once if you like. The video depicts a conversation between three characters named below. Which video did you watch?
- A
- C
- D
- F
- G
- I
- J
- K

15. Rank how dominant each of the characters appears to be in the conversation
- Very Subservient
- Quite Subservient
- Slightly subservient
- Neither subservient nor dominant
- Slightly dominant
- Quite dominant

- Very dominant

16. What led you to this decision? Please rank the aspects of the video which most affected your reading of which character was dominant, with 1 being the most important aspect and 5 the least.
- Action
- Dialogue
- Camerawork
- Lighting
- Sound

17. On a scale of 1-7 how isolated do each of the characters seem (1 being least isolated and 7 being very isolated).
- 1(Not at all isolated)
- 2
- 3
- 4
- 5
- 6
- 7 (Very Isolated)

18. Again, what led you to this decision? Please rank the aspects of the video which most affected your reading of which character was isolated, with 1 being the most important aspect and 5 the least.
  - Action
  - Dialogue
  - Camerawork
  - Lighting
  - Sound

19. On a scale of 1-7, how much affinity does each of the characters have for each other (with 1 being no affinity and 7 being a strong connection or liking for each other).
- 1(No affinity)
- 2
- 3
- 4
- 5
- 6
- 7 (Strong Affinity)

20. Again, what led you to this decision? Please rank the aspects of the video which most affected your reading of which characters had an affinity for each other, with 1 being the most important aspect and 5 the least.
- Action
- Dialogue
- Camerawork
- Lighting
- Sound

21. We'd now like you to watch the third video in the email we sent you and answer a few brief questions. Feel free to pause, rewind or watch the video more than once if you like. The video depicts a conversation between three characters named below. Which video did you watch?
- A
- C
- D
- F
- G
- I
- J
- K

22. Rank how dominant each of the characters appears to be in the conversation

- Very Subservient
- Quite Subservient
- Slightly subservient
- Neither subservient nor dominant
- Slightly dominant
- Quite dominant
- Very dominant

23. What led you to this decision? Please rank the aspects of the video which most affected your reading of which character was dominant, with 1 being the most important aspect and 5 the least.
- Action
- Dialogue
- Camerawork
- Lighting
- Sound

24. On a scale of 1-7 how isolated do each of the characters seem (1 being least isolated and 7 being very isolated).
- 1(Not at all isolated)
- 2
- 3
- 4
- 5
- 6
- 7 (Very Isolated)

25. Again, what led you to this decision? Please rank the aspects of the video which most affected your reading of which character was isolated, with 1 being the most important aspect and 5 the least.
    - Action
    - Dialogue
    - Camerawork
    - Lighting
    - Sound

26. On a scale of 1-7, how much affinity does each of the characters have for each other (with 1 being no affinity and 7 being a strong connection or liking for each other).
- 1(No affinity)
- 2
- 3
- 4
- 5
- 6
- 7 (Strong Affinity)

27. Again, what led you to this decision? Please rank the aspects of the video which most affected your reading of which characters had an affinity for each other, with 1 being the most important aspect and 5 the least.
- Action
- Dialogue
- Camerawork
- Lighting
- Sound