Proceedings of the 2005 Winter Simulation Conference
M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines, eds.

EXPERIMENTAL INVESTIGATION OF THE IMPACTS OF VIRTUAL REALITY ON DISCRETE-EVENT SIMULATION

Justice I. Akpan
Roger J. Brooks

Department of Management Science
Lancaster University Management School
Lancaster, LA1 4YX
UNITED KINGDOM

ABSTRACT

This paper presents the results of experimental studies that were undertaken to test the impacts of Virtual Reality (VR) on Discrete-Event Simulation (DES). The experiments focused on spotting errors in the DES model. The models were developed in 2D and 3D/VR displays using WITNESS. The 2D display used icons and other visualization techniques that confine its scope to essentially flat 2D surface. On the other hand, the 3D display was represented by means of a three-axis spatial position (XYZ) plots, but appeared on a two-dimensionally mappings, otherwise known as 2.5D. The experiments involved paid participants who were recruited from amongst the staff and students of Lancaster University, UK. The results showed that it is easier, and faster to spot errors in 3D/VR model than in 2D. The findings also indicated that users can easily understand the modeled operation of 3D/VR display compared to 2D, irrespective of background or technical ability.

1 INTRODUCTION

The application of VR in DES is becoming popular. Unfortunately, there is little or no empirical evidence establishing clear benefits of 3D/VR modeling over the conventional 2D (Zutphen et al. 1996, Asthmeimer 1999). The current literature seems to be dominated by fairly superficial assessment of the novelty of 3D/VR simulation software, and 'propaganda of success stories' in an attempt to sell simulation solutions.

A number of speculative claims about 3D/VR based DES (VRSIM) have been identified. But, most of the claimed benefits also lack any scientific evidence. The possible impacts of VRSIM at various stages of DES modelling activities or stages (Smith 1999) remains blurred with existing studies.

This paper presents the experiments conducted at Lancaster University, UK, in 2004 to investigate the impacts of visual display (3D/VR and 2D) on spotting errors in DES model at the testing and validation stage of simulation modeling process.

The models used for the experiments were developed in 2D display and 3D/VR perspective (2.5D). Hence, the term 3D/VR model used for the experiments reported in this paper actually refers to the 2.5D display. The 2D display uses icons and visualization techniques that confine its scope to essentially flat 2D surface (see Figure 4). On the other hand, the 3D display was represented by means of a three-axis spatial position (XYZ) plots, but appeared on a two-dimensionally mappings. Although this display (see Figure 5) provides much of the outstanding visualization features of the 3D display, it is technically not a full 3D display but a quasi-3D or 2.5D display (Cleveland and McGill 1988, McAllister 1993). This is because it contains no real binocular stereographic depth effects.

The reasons for using the 3D/VR perspective rather than full VR version were as follows:

1. It was considered a necessary first step to compare two simulation model displays that appeared to be pretty similar, bearing in mind the great disparity between a 2D model display and a full 3D/VR version. This helped to measure what impacts the slight change in the visual display can have on users' performance in terms of spotting errors in simulation model.
2. The initial intention however, was to follow on with another experiment involving comparative evaluation of the 2D display and full 3D/VR version. But, this was not possible as the 3D/VR simulation modeling software was not readily available at the time of the experiments.

Notwithstanding the perceived limitations, the experimental outcome were convincing and clear, in terms of the effects of visual display on model accuracy; how easy it is, and the time taken to spot errors in simulation models.

The rest of the paper discusses the evaluation criteria based on some of the claimed benefits about VRSIM, the aims and objectives of the study, the experimental process, procedures and method. Finally, the paper analyses the experimental outcome and suggests areas for further research.

2 EVALUATION CRITERIA

This section outlines some of the claims about VRSIM that were tested during the experiments.

2.1 VRSIM Enhances Detection of Errors in DES Model

Current literature speculates that it is easier to spot errors in 3D model than in 2D. This implies that VR can enhance more accurate model, which is made possible by its excellent visualization capability (Kamat and Martinez 2000, Munro et al. 1999, McKay et al. 2002, Mesquita et al. 2000).

2.2 VR Enhances Users' Understanding of the Model

VR provides true to scale 3D graphics and animation, making simulation models easy to understand and invaluable for communicating new ideas and alternatives (Bennaton and Sivayoganathan 1995).

3 THE EXPERIMENTAL DESIGN

3.1 Aims and Objectives

This experiment was designed to answer two research questions relating to the claim that VRSIM enhances more accurate model than 2D display. The research questions and the hypotheses developed and tested are as follows:

3.1.1 Is It Easier to Spot Errors in 3D/VR Model than in 2D?

This Section defines the hypothesis that was formulated to answer the above research question.

The null hypothesis H_0 and alternative hypothesis H_1 were defined as follows:

- H_0: VRSIM does not make it easier to spot errors in simulation model
- H_1: VRSIM makes it easier to spot errors in simulation model

3.1.2 Does It Take Shorter Time to Spot Errors in 3D/VR Model Than in 2D?

Similar to Section 3.1.1 above, the following hypothesis was formulated to answer the above research question.
The null hypothesis H_0 and alternative hypothesis H_1 were defined as follows:

- H_0: VRSIM is not more efficient in spotting errors in simulation model
- H_1: VRSIM is more efficient in spotting errors in simulation model

Efficiency in this context refers to the time taken to spot errors in DES model. The shorter the time taken to spot the error in a particular display, the more efficient the display is assumed.

3.2 Recruitment of Participants

The students and staff members of Lancaster University were recruited from various academic disciplines for the experiment. A total number of 62 subjects from different backgrounds took part in the experiment as shown on Table 1. The reason for recruiting subjects from different disciplines was to observe any effects of backgrounds of subjects on their performances, bearing in mind the spread of simulation users from various backgrounds. There was no particular attention on the gender of subjects as such demographic details and classification was not essential.

Table 1: Background of Participants

Academic Discipline/Background	No of Subjects
Simulation	7
Other OR/MS	24
Management/Finance	12
Engineering/Science	6
Computing/IT	6
Arts/Humanities	7
Total	*62*

3.2.1 Educational Level of Participants

Figure 1 shows that, the academic/research staff and PhDs students made up 52% of the sample, while postgraduate (MSc, MA, MRES, etc.) and undergraduate students had 24% of the participants each.

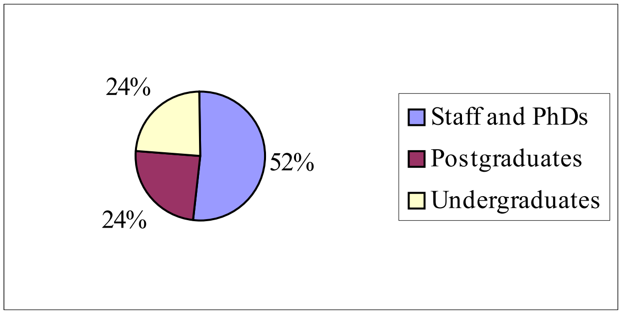

Figure 1: Subjects' Level of Education

3.2.2 Subjects' Experience in Simulation

Prior to the start of the experiment, the participants were asked to indicate their experience in building or using computer simulation. The reason was to observe any influence that users' past experience in simulation modeling can have on their performance during the experiment.

Figure 2 shows that just over 58% of subjects had experience in 2D simulation modeling or in the use of 2D models, while less than 2% of the subjects had any previous experience in the use or development of 3D models.

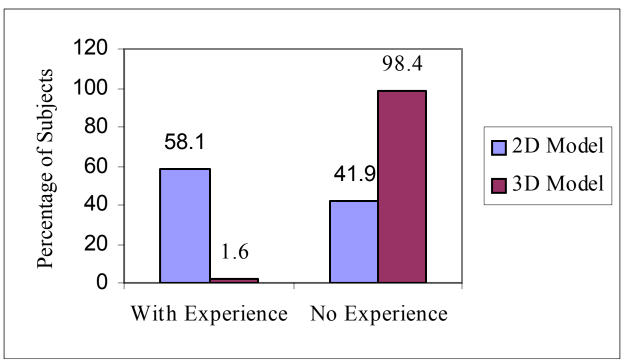

Figure 2: Subjects' Experience in Simulation Modeling

3.3 Method

This study adopted a structured and controlled strategy for the experiments. This implies the control and manipulation of variable of interest (Montgomery 1997) during experimentation. For example, participants were randomly selected to perform tasks on 2D or 3D displays. Participants from different academic areas or disciplines were fairly assigned to undertake the experiment in either of the model displays as shown in Figure 3. Furthermore, the maximum time in which participants were allowed to spot the error was set prior to the start of the experiment.

Finally, the experiment was conducted in a controlled environment so as to minimise any external factors (such as noise) capable of influencing the outcome of the experiment.

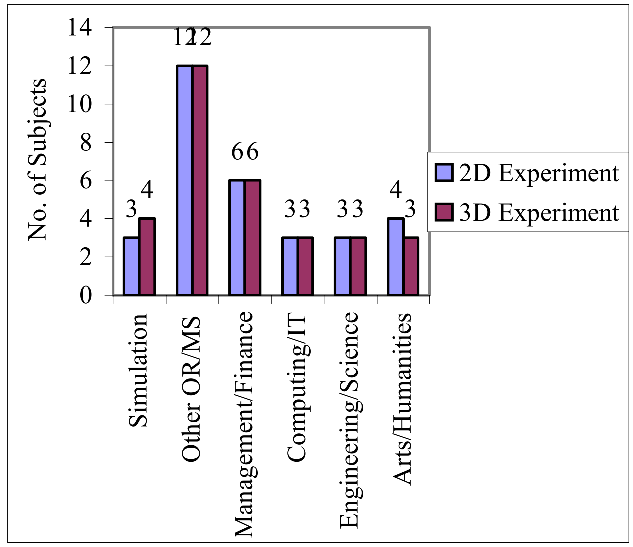

Figure 3: Assignment of Participants into Tasks

3.4 Equipments / Resources

The resources and other equipments used for the experiments are summarized in Table 3.

Table 3: Resources for the Experiment

Software	Hardware	Other Equipment
WITNESS Software	P4 Computer	Controlled Environment (Lab)
Simulation models (2D and 3D - 2.5D)		

3.5 The Experimental Model

The model used for the experiment was based on a car assembly factory simulation. This operation was chosen because of its interesting and good visual features for the experiment since visualization was a key feature in the study.

Moreover, it was necessary to select a product and operation that most people are familiar with, hence the choice of the car (as the product) and a queuing conveyor locomotive system in a car assembly factory.

3.5.1 Description of the Experimental Model

The model simulates a fictional car assembly process in which three components namely, 1 body, 2 doors and 4 tyres are assembled into a complete car. The model elements (e.g. machine and labour) and the parameters (e.g. cycle time, breakdown intervals, machine repair time, all in minutes) are shown in Table 4. The activities carried out by the machines in the factory are:

1. Each component is tested

2. The components are assembled into a car
3. The car is tested
4. Rejected cars are reworked.

Table 4: Model Elements and Parameters

Machine:	Test Tyre	Test Body	Test Door	Assemble Car	Test Car	Rework Car
Cycle Time:	12.5	40.0	25.0	35.0-45.0[a]	25.0	35.0
% Rejects:	15	5	5	-	40	2
Labour:	-	Insp[b]	Insp[b]	Assem[c]	Insp[b]	-
Breakdown Intervals[d]:	600	60	60	57	180	-
Repair Time[e]:	-	2, 10	2, 10	2, 9	2, 10	-

a. Uniform Distribution
b. Inspection Technician
c. Assembly Technician
d. Negative Exponential
e. Lognormal Distribution: μ, SD

3.5.2 The Model Displays

As mentioned in Section 1, the model was developed in two displays, 2D and 2.5D or 3D/VR perspective. Figures 4 and 5 show the screen shot of the 2D model before and at run time, while Figures 6 and 7 show that of the 3D/VR display.

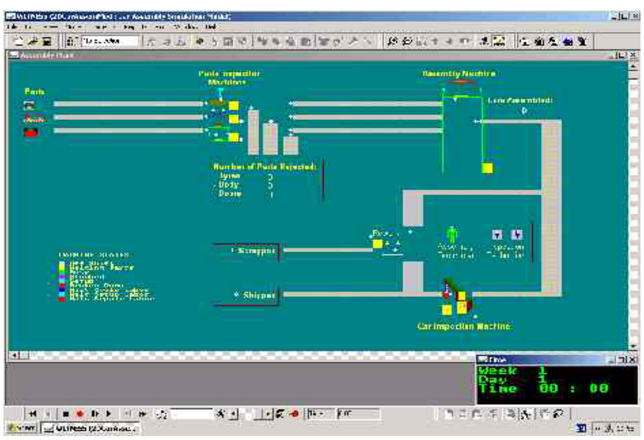

Figure 4: Screen shot of 2D display of the Car Assembly Model

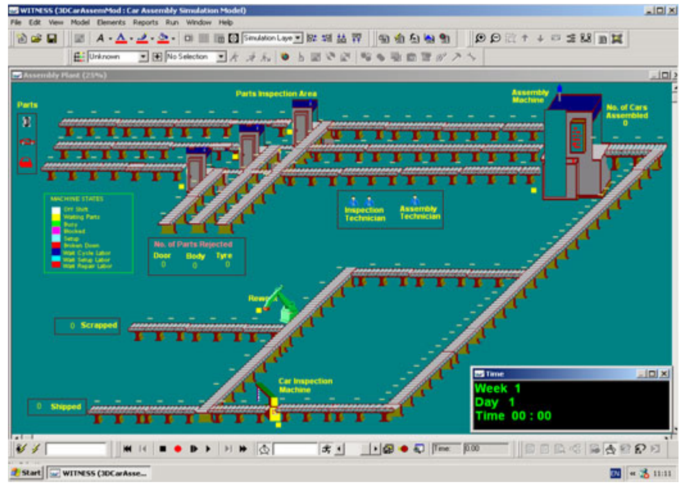

Figure 5: Screen Shot of 3D Display of the Car Assembly Model

The models were tested to ensure that they functioned correctly and according to specification after being developed. Also, it was essential to ensure that both models were identical in terms of technical functionalities and physical features. These requirements were mandatory to ensure unbiased outcomes from the experiments.

The following features were common to both displays:

1. Both model displays ran at the same speed and produced same results when run for the same duration (see Figures 6 and 7).
2. The models were designed to have similar physical characteristics.
3. The elements and parts in both models were differentiated using different scale, colour and shape.

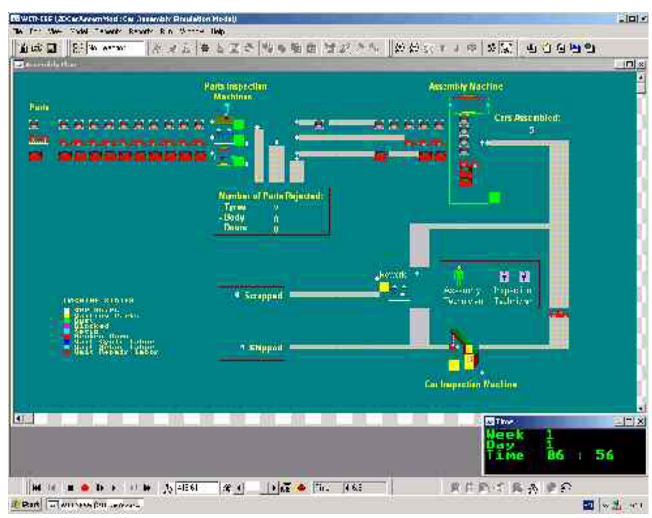

Figure 6: Screen shot of 2D Display of the Car Assembly Model (at a Runtime of 416.61 Minutes).

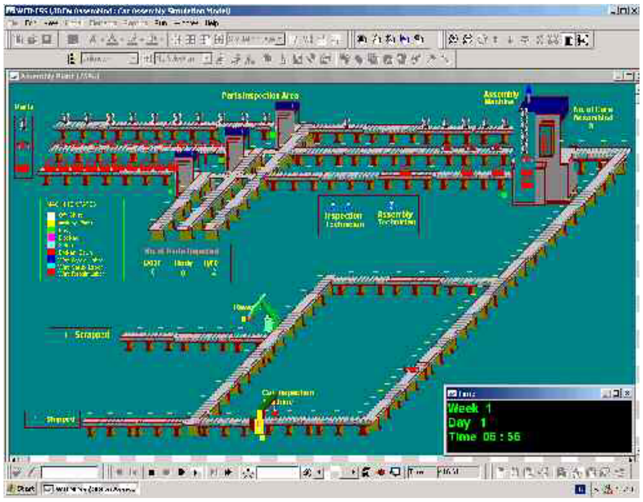

Figure 7: Screen shot of 3D display of the Car Assembly Model (at a Runtime of 416.61 Minutes).

3.6 The Errors Required to Spot

The errors to spot during the experiment were inserted into the models (both the 2D display and 3D/VR), which hindered the models from working correctly as described in Section 4.3. In order to ensure that the error did not completely alter the operational processes and functionality of the model, only one error was introduced into the model at a time. For example, the first experimental task was to spot the ‘routing error’ (see Section 3.6.1) in the 2D display, thereafter spot the ‘assembly error’ (see Section 3.6.2) in another copy of the same model. Similar tasks performance was applicable to the experiment on the 3D display (VRSIM).

3.6.1 Routing Error

This error involved wrong routing of the finished product (assembled cars) after quality inspection. In the correct version of the model, the cars that did not meet the required quality standard were reworked, with 98% to be sent back for re-inspection and ship, and just 2% scrapped. But, in the model with error, all the cars that needed rework were wrongly routed to scrap.
There were three possible ways to spot this error as follows:

1. Visually: The rework machine was idle throughout the operational cycle.
2. Idle Conveyor: The conveyor in which the reworked cars were to be re-routed for quality inspection was idle throughout the operation.
3. Statistically: The number of scrap cars was higher than the required 2% during the model runtime.

3.6.2 Assembly Error:

In this error, wrong number of components was used for the assembly operation. The assembly machine pulled five tyres instead of four (in addition to the correct number of other components) for assembly.

This error was to be spotted “visually” through the following ways:

1. Presumably this error could be spotted by observing the number of tyres at the machine during the cycle time (at the machine),
2. Carefully observing the frequent flagging of the assembly machine in “idle-state” due to shortage of the part (the tyre).

3.6.3 Classification of Errors

The errors were classified into two namely, easy error and hard error. The modality for the classification was based on feedback from different subjects during a trial experiment. Hence, the ‘Routing Error’ was classified as easy error, while the ‘Assembly Error’ was considered as ‘Hard Error’.

3.7 Experimental Processes and Procedures

The experiment was conducted with only one participant per session. Prior to arrival, each participant was randomly assigned to perform the tasks in either 2D or 3D simulation models using ‘precise sampling’ technique (Faulker 2000). This means that, all the subjects from a similar background cannot all undertake experimental activities on the same display. For example, of the 6 subjects from computing/IT background, 3 subjects were randomly assigned to perform the experimental tasks in 2D and the other three subjects on the 3D displays (see Figure 3). Overall, 31 subjects performed the experimental tasks in 2D and 31 in 3D. The tasks were set-up before the arrival of each subject.

Table 5: Experimental Tasks/Activities

Activities/Tasks During Experiment	Maximum Time Allowed
i. Read model description + Questions	5 minutes
ii. Spot First Error	10 Minutes
iii. Spot Second Error	10 Minutes

The following processes and procedures were adopted during the experiment:

1. Each participant was handed a set of instructions about the experiment at the start of the session. The instruction sheet explained the experimental

tasks, the time allocated for each task, and a description of how the model is supposed to work (see Section 4.3).
2. The experimenter guided the subject through to open and run the model for the first task, which was to spot the 'Routing Error'. The subjects were allowed to make several attempts to spot the error but subject to the maximum allowable time of 10 minutes.
3. The experimenter records the result of the first task (whether or not the error is spotted), and the time taken to spot the error.
4. Repeat steps (ii – iii) for the Assembly Error (see Section 6.4 for description of the error).
5. The experimenter records the result of the second task (whether or not the error is spotted) and the time taken to spot the error.

4 RESULTS AND DISCUSSION

This section presents the main results from the experiment and also tests the relevant research hypothesis (as defined in Section 6.2).

4.1 Does VRSIM Enhance Spotting Errors in Model than 2D Display?

The effect of the type of model display on spotting errors in simulation model was based on the percentage of participants who spotted the errors (Routing Error and Assembly Error) in either the 2D and the 3D displays of the simulation model.

Figure 8 shows the number of subjects (%) who spotted the errors within the time limit of 10 minutes. The 3D users achieved a 100% success in spotting the 'Routing Error', while only 81% of the 2D users spotted the same error. Similarly, 97% of the subjects who used the 3D display spotted the 'Assembly Error' while only 45% of the 2D users spotted the same error.

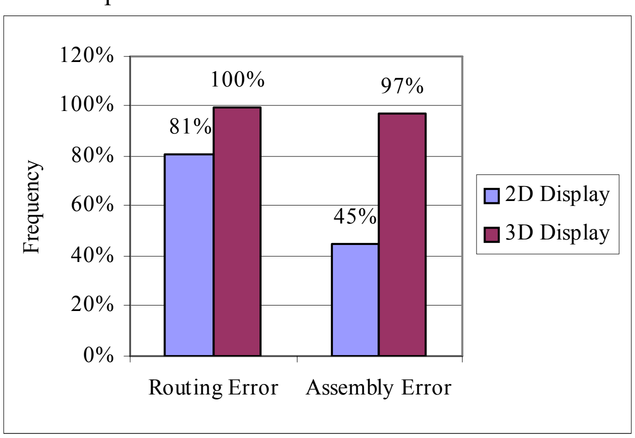

Figure 8: Subjects Performance on Spotting of Errors

• **Test of Significance/Hypothesis**

Further analysis was carried out to test the statistical significance (or not) of the differences in performance between users of 3D and 2D display. The 'Independent-Sample estimation' (see George and Mallery 2003) was used to compare the mean performance of subjects on the two displays. The results (see Table 6) indicate that, the performance of the 3D users was statistically better than users' performance on the 2D display at less than 5% level of significance ($p < 0.01$).

Similar results were also recorded on subjects' performance in spotting the 'Assembly Error'. The results also indicate that, the difference in performance between the two displays was statistically significant at less than 5% level ($p < 0.01$).

Table 6: Subjects' Performance in Spotting Errors

Types of Error	Categories of Error	2D	3D	P-value
Routing Error	Easy Error	81%	100%	0.01
Assembly Error	Hard Error	45%	97%	0.01

On the basis of the above statistical evidence, we reject the null hypothesis (see Section 3.1.1) and accept the alternative hypothesis that VRSIM makes it easier to spot errors in DES model, thereby enhancing more accurate model.

4.2 Time Taken to Spot the Errors

Figures 9 and 10, and Table 7 show the average time taken to spot the errors in the simulation model. The maximum allowable time of 10 minutes (in real-time) was recorded against every unsuccessful attempt.

The results indicate that subjects who spotted the 'Routing Error' on the 3D model spent less time (96.5 seconds on the average) compared to those who performed the same task using the 2D display (246.0 seconds).

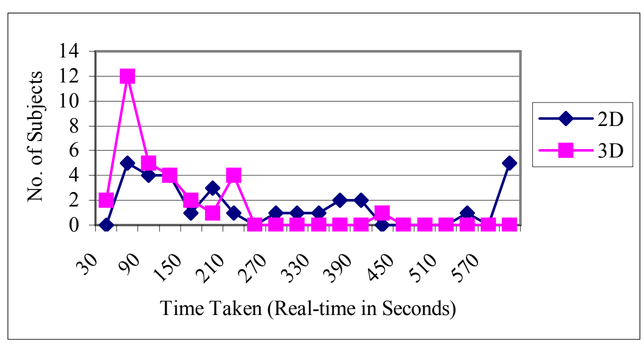

Figure 9: Time Taken to Spot the Routing Error

These results indicate that it took more than twice the time to spot the 'Routing Error' in the 2D display than in a similar model with 3D display (246 seconds > 96.5 seconds).

The results recorded for the 'Assembly Error' was followed a similar trend. On the average, it took more than thrice the time to spot the 'Assembly Error' in 2D model (462 seconds < 129 seconds in real-time) than in the 3D display (see Table 7).

Table 7: Average Real-Time Taken to Spot Errors and Significance Levels

Types of Error	Mean Time - 2D (in Seconds)	Mean Time - 3D (in Seconds)	P-value
Routing Error	246.0	96.5	0.001
Assembly Error	461.5	129.2	0.001

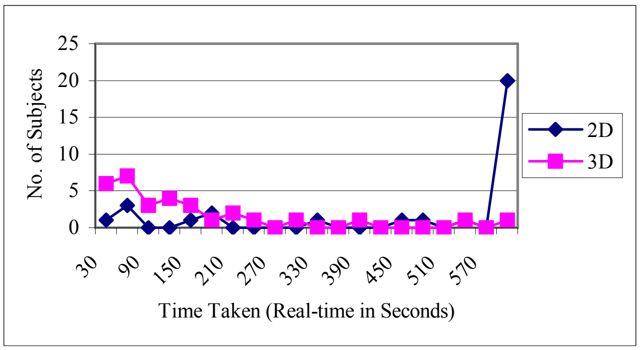

Figure 10: Time Taken to Spot the Assembly Error

- **Test of Significance/Hypothesis**

The results on Table 7 show that the difference in the average times spent to spot the 'Routing Error' on the 2D display was significantly higher than the time spent to spot the same error in the 3D model at less than 5% level ($p< 0.01$). Similarly, the difference in time taken to spot the 'Assembly Error' in the two displays was also statistically significant at less than 5% level ($p < 0.01$).

On the basis of this result, we reject the null hypothesis (see Section 3.1.2) and accept the alternative hypothesis that VRSIM is more efficient in spotting errors than 2D as it takes less time to spot errors in VRSIM model than in 2D.

5 CONCLUSIONS AND FURTHER RESEARCH

5.1 Conclusions

This study has provided interesting insights into the application of VR in DES modeling. The results show that it is easier to spot errors in 3D/VR model than in 2D. Moreover, it takes less time to spot error in 3D/VR model than in 2D. This implies that VRSIM enhances more accurate DES model than 2D display.

Interestingly, the effect of 3D display in spotting errors in the model is more pronounced if the error is generally difficult. For example, the percentage of 'hard' or 'difficult' errors spotted on 3D/VR model was significantly higher than in 2D display (97% > 44%) compared to the spotting the easy error (100% > 81%) respectively. This suggests that, the more difficult an error is, the more helpful it is to use VRSIM.

Another implication of the result is that, VRSIM enhances understanding of the modeled operation. This enables the user to easily detect errors in the model. Furthermore, the outstanding visualization capability of the 3D display makes it easier to spot a more subtle or difficult errors that may lurk hideously in the 2D display.

Regarding the time taken to spot the errors, it generally took takes less time to spot both errors in the 3D display than 2D. However, the time spent to spot the hard error in the 2D display was a lot longer than in the easy error. For example, it took about 2.6 times longer (on the average) to spot the easy error in the 2D model, and 3.6 times longer for the hard error. These results suggest that, the harder the error, the more essential it is to use 3D display as it saves significant amount of time.

5.2 Future Research

This study has produced impressive results to evaluate the claims of VRSIM. However, there is need for a careful interpretation and generalization. It is needful to carry out more scientific experiments using different applications, scenarios and involving VRSIM practitioners and users in order to compare the outcomes.

Furthermore, as mentioned in Section 1, the experiments reported in this paper compared the performance of users on 2D display and 2.5D or 3D/VR perspective display. Despite the clear and convincing results from this experiment, it is still necessary to conduct another experiment using 2D display and a full 3D/VR application.

ACKNOWLEDGMENTS

The authors wish to thank all the participants in the experiment.

REFERENCES

Astheimer P. 1999. A business view of virtual reality. *IEEE Computer Graphics and Applications.*

Barnes, M. 1997. An introduction to Quest. In *Proceedings of the Winter Simulation Conference*. Eds. S Andradottir et al.

Bennaton, J and Sivayoganathan, K. 1995. The usefulness of 3D discrete event Simulation. In *Proceedings of the 31st International MATADOR Conference*, 20th – 21st April, Manchester.

Faulker, 2000. *Usability engineering*. Macmillan.

George, D and Mallery, P. 2003. *SPSS for Windows step by step: A simple guide and reference,* 11.0 Update. th Edition. Pearson Education Inc.

Hlupic, V. 2000. Simulation software: An Operational Research Society survey of academic and industrial users. In *Proceedings of the 2000 Winter Simulation Conference*. Eds. J A Joines, et al.

Hurrion**,** Robert D. 1981. An visual interactive visual simulation system for industrial management. In *European Journal of Operational Research*, Vol 5. No.4 pp138 – 144.

Hurrion**,** Robert D. 1986. Visual interactive modelling. In *European Journal of Operational Research*.

Hurrion**,** Robert D. 1985. Implementation of a visual interactive consensus decision support system. In *European Journal of Operational Research.*

Jacob, R. K, Deligiannidis, L and Morrison, S. 1999. A software model and specification language for non-WIMP user interfaces. *ACM Transactions on Computer-Human Interaction*.

Jones, N. 1992. Manufacturing simulation using virtual reality. Master's Thesis.

Kamat, V R and Martinez, J. C. 2000. 3D visualization of simulated construction operations, *Proceedings of the 2000 Winter Simulation Conference*. Eds. J A Joines et al.

Kirkpatrick, P. and Bell, P. C. 1989. Visual interactive modelling in industry: Results from a survey of visual interactive model builders. *In Interfaces.*

McKay, K. et al. 2002. An aggregate resource model for the Provis digital mock-up within a distributed manufacturing planning system.

Mesquita, R. et al. 2000. *Usability engineering: VR interfaces for the next generation of production planning and training systems*, London.

Montgomery, D. C. 1997. *Design and analysis of experiments*. New York: Wiley.

Munro, A. J. et al. 1999. *Social navigation of information space*. Springer, London.

Shannon, R. 1975. *Systems simulation: The art and science.* Prentice-Hall, Englewood Cliffs, NJ.

Smith, R. 2000. Simulation: The engine behind the virtual world. In the *Simulation 2000 series*.

Smith, S and Duke, D. J. 1999. Virtual environments as hybrid Systems. Paper presented at *Annual Conference, Eurographics*, UK Chapter, Cambridge.

Zutphen, J et al. 1996 Computational design: Simulation in virtual environments [online]. Available via `<http://www.ds.arch.tue.nl/Researc/publications/rob/DDSS96_finale.htm>`, [Accessed: 06.01.03]

AUTHOR BIOGRAPHIES

JUSTICE I. AKPAN studied Software Development at Leeds Metropolitan University, UK, where he received his MSc. Currently, he is finishing his PhD Thesis on the impacts of Virtual Reality on Discrete-Event Simulation at Lancaster University. He is a member of British Computer Society, UK and has worked as Software Developer, Web Programmer, and Developer of Virtual Learning Environment (VLE) in the USA and UK. Email to: <J.Akpan@lancaster.ac.uk> Website at: <www.lancs.ac.uk/postgrad/akpan/>.

ROGER J. BROOKS is a lecturer in the Department of Management Science at Lancaster University. He has research interests in the methodology and the applications of simulation. Email to: <Roger.Brooks@lancaster.ac.uk>, website at <http://www.lancs.ac.uk/staff/smarb/>.