Doing It Right: Psychological Tests to Ensure the Quality of Scientific Visualization

Helmut Haase[†], Christoph Dohrmann

Fraunhofer Institute for Computer Graphics (Fraunhofer IGD),
Wilhelminenstr. 7, D-64283 Darmstadt, Germany
haase@igd.fhg.de – http://www.igd.fhg.de/∼haase

Abstract

This paper discusses a general scheme for determining the quality of scientific visualization systems. It presents psychological tests which have been performed in order to find quantitative relationships for this scheme, and discusses why simple numerical relationships may be hard to find.
Our work is motivated by research into interactive and immersive scientific visualization which pose two opposing demands: maximum image quality at sufficient frame rates. We believe that these two factors are also crucial for many other applications, e.g., virtual reality. For this scheme, special focus is on the user's perception of the system. Three components of Visualization System Quality are identified: Data quality, image quality, and interaction quality. In order to migrate from a merely qualitative to a more quantitative model, psychological tests were performed to measure the influence of frame rate and rendering mode on the perception of visualized threedimensional vector data. Results of the tests are presented and general suggestions for good perceptual tests are made.
We believe that the experience we gained will be of benefit to many who are interested in questions of visualization system quality.

0.0.0.1. Keywords: *Scientific Visualization, Realtime Interaction, Psychological Tests, Visualization System Quality*

1. Visualization System Quality for interactive and immersive scientific visualization

Many different techniques for scientific visualization are known, and each technique can be parameterized in various ways. But what is the "best" visualization for a given problem, i.e., for a particular dataset and a particular user with a particular degree of experience and particular goals using hardware with particular features?

This is a tough question. Doing good visualizations is very hard. Lately, some people have pointed out the importance of this question ([10], [18]). In [21], a case study is reported which addresses the question of visualization quality. In [16], a tool is presented which recommends visualization techniques for a given set of data.

We will address this topic by means of psychological tests. In the following, a general scheme for Visualization System Quality (VSQ) will be presented. By the term *visualization system* we mean a combination of hard- and software. Visualization systems can be classified according to various criteria, e.g., price, CPU speed, screen size, etc. We are mainly interested in user performance as a criterion for system quality. This approach is motivated by our previous work on Scientific Visualization, e.g., [7].

Increasingly, scientific data now is being visualized using Virtual Environment or Virtual Reality (VR)[1] techniques. This is due to easy availablitiy of hardware and software suitable for VR techniques as well as to the observable benefits of merging these two techniques to become Immersive Scientific Visualization.

Some applications which already utilize this synergy are reported in [2] (Virtual Windtunnel), [5] (several visualizations using the CAVE system), [20] (VR in dataflow systems), [15] (Responsive Workbench), and work by us: [12] (medical training, material testing, visualization of room accoustics, and 3D weather visualization), [13] (investigation of molecule datain VR), and [14] (coupling of general purpose VR and scientific visualization system). In all these examples, the opposing demands of image quality and rendering speed have been met in a certain way (see also [3]). Even though CPU performance and memory size increase drastically, the tension between image quality and frame rate will remain to be a crucial issue for all kinds of graphics systems. This is due to demands for an ever increasing number of details in the data, no matter if it is finite element meshes, realistic landscapes, or medical data.

The solution to this problem lies in dynamic adaption of data and image complexity as well as of the degree of interaction to the current situation. This situation consists, among other factors, of system load, of the kind and amount of data which is used by the system, as well as of the knowledge, strategy, and goals of the current user. It seems that there is need for a quality measure of graphics systems which allows to adjust system parameters depending on the situation in order to maximize the benefit of the system to each individual user. The perfect solution would be a quantitative measure for system quality allowing us to optimize system parameters according to the individual constraints of each possible user/data/system combination.

Starting from this observation, we developed a general scheme for determining the quality of a (3D) graphical system. It is presented in more detail in [13]. Visualization System Quality as described in the following is a special kind of Graphics System Quality since visualization systems are a special kind of graphics systems. Thus, VSQ describes the underlying system in terms of hardware and general performance but not the visualization technique used or the didactics of a visualization. This scheme still is only qualitative. We tried to determine some numerical input for this scheme by performing perceptual studies with test persons. Unfortunately, we discovered that valid, representative numerical results are not easy to obtain. Here we will report on the preliminary, qualitative scheme, on the tests that we performed, and on the implications of these tests.

The scheme does not explicitely comprise "hidden" features like memory size or processor speed, even though they may influence factors which we consider important for Graphics System Quality (GSQ), e.g., rendering speed. Rather, the scheme comprises features which are visible to the user directly.

The scheme consists of three main components: Data quality, image quality, and interaction quality. The details of these main components are summarized in Table 1.

Data Quality	Image Quality	Interaction Quality
• Data Semantics: 1. Only geometric or image data 2. Static semantics 3. Dynamic semantics • Object space resolution	• Image space resolution • Color space resolution • Rendering/shading (wireframe, flat, gouraud, phong)	• Rendering speed (static,sequence,realtime) • Degree of interaction (none, interactive, VR) • Number of users (none, one, multiple)

Table 1: *Components of Visualization/Graphics System Quality.*

In the following, the components of the GSQ scheme are explained in more detail.

Data quality stands for the semantics of graphic objects and resolution in object (data) space.

Image quality is a term which may be used with different meanings. First of all, its definition may depend on the application area, e.g., text, technical drawings, or video. For the sake of this paper, we are mainly interested in (sequences of) raster images of scientific 3D data. Furthermore, image quality can be investigated on a very basic level (spatial resolution of pixels, dynamic range, etc.), on higher levels (geometric distortion, feature distortion, etc.), or even on a very high conceptual level (style, redundancy, etc.). For an overview of the state-of-the-art of definition and measurement of image quality, see [8].

Image quality is an assembly of several factors, all influencing the amount of information which is accessible from a single image. For our purposes, the following characteristics seem most important for the definition of image quality: Resolution in image space, resolution in color space, and rendering/shading method.

Finally, *interaction quality* is an important component of the GSQ scheme. It comprises: Rendering speed, degree of interaction, and number of users.

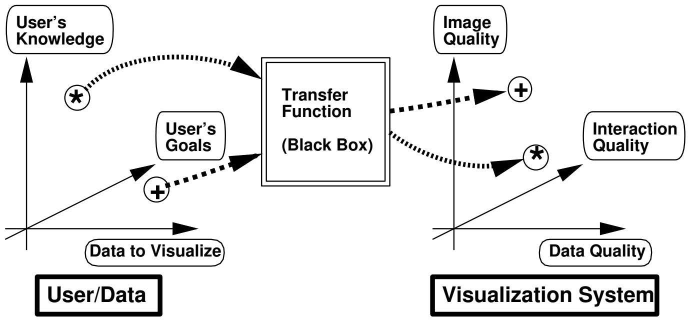

Figure 1: *Main factors influencing Graphics System Quality.*

In order to further elaborate this scheme and to develop a quantitative measure for Visualization/Graphics System Quality, we performed some psychological tests. These tests made clear that users have very different visualization and perception strategies depending on their knowledge and goals. The data to visualize as well as knowledge and goals of the user constrain the best choice of data quality, image quality, and interaction quality. This can be seen as a kind of mapping or transfer function from 3D to 3D (see figure 1). Unfortunately, this transfer function is unknown to us. Investigation to determine this transfer function is worth-while in order to achieve better visualizations. Therefore, further perceptual tests need to be performed.

2. Testing the influence of rendering speed and rendering mode on perception of 3D vector fields

The previous chapter outlined our view on Visualization/Graphics System Quality. It is obviously motivated by a great interest in the performance of a user of a graphics system. In general, a user of a graphics system may want a visualization of great complexity at the highest possible frame rate. Due to limitations in rendering speed and increasing amounts of data to be visualized, this demand cannot be fulfilled entirely. On the other hand, since human visual perception is limited in respect to spatial and temporal resolution [11], it could be assumed that parameters can be found which allow optimum human performance without exceeding the capabilities of a given graphics system. Exact numerical models of visual 3D perception are not available ([4] presents such a model for the limited task of text processing). Thus, rules or formulas for the optimal configuration of a visualization system are hard to specify. Psychological tests are one possibility to obtain a crude, simple model of the "visualization user".

Therefore, we investigated the possibility of perceptual tests. We are interested in mathematical relationships of parameters for visualization and their influence on human performance in perceptual tasks. Thus, a large number of parameter combinations must be tested. This implies that, in order to gain results that can be trusted with a certain confidence, a large number of test subjects must be involved which makes the test very expensive. Normally, only industry can afford to carry out tests with hundreds of subjects.

But even if we could perform such tests, the quantitative results by themselves may not be sufficient. We want to generalize from the specific test conditions to more general phenomena and therefore observations of the bahaviour and personal comments of subjects must be recorded and analyzed. It may be that the numerical results were strongly influenced by factors that are unexpected and unknown to the test designer, for example different strategies in solving a perceptual task.

It also is more easy to look for new possible phenomena than to design a test which tries to verify a theory. To verify a theory with sufficient confidence, a large number of subjects is needed, maybe all with a similar background. To discover new phenomena, a smaller group with very different subjects can be sufficient. It should also be noted that there are different statistical methods that can only be used either to detect new phenomena or to verify expected phenomena. For example, analysis of variances should only be used to verify or falsify expected phenomena while cluster analysis should only be used to search for new phenomena.

2.1. Our tests

Since we have been concerned with Scientific Visualization for a couple of years, we chose perception of individual vectors as well as of vector fields as task for the tests. Our main goal was to classify the influence of frame rate and of rendering mode of the vectors on the perception of the angles of individual 3D vectors ('detail' test) as well as on the holistic perception of a complex 3D vector field ('overview' test).

We started by doing some research into literature on visual perception. Gibson's ecological approach to visual perception [9] is the most fruitful for the explanation of perception of 3D structures. Unfortunately, it does not contain numerical models. This approach identifies "gradients" and "invariants" as the most important factors for spatial relations and object recognition. Cue-theory[11], which is based on feature extraction for the "cues" stereo, shading, texturing, movement, and focus, should be considered as a subset of the bigger concept of the ecological approach. It clarifies the different aspects of an animated image which all should be seen as building blocks of an overall image quality. Most of the cues have already been tested seperately, but the effect of combinations of several parameters, which is essential in visualization systems, has not been examined in detail. The effect of combinations of different parameters on 2D data analysis currently is under investigation[19], especially in the case of dynamic presentations[17].

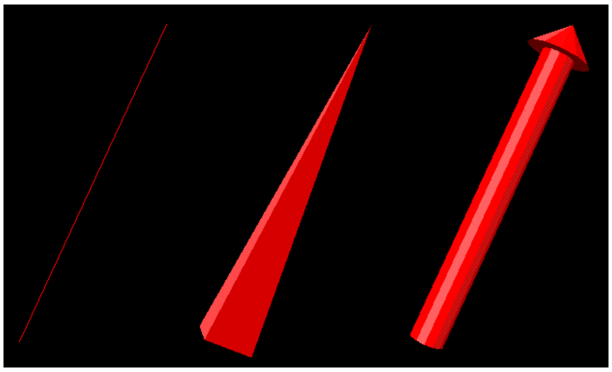

Figure 2: *The three investigated rendering modes: line, pyramid, detailed arrow.*

Our tests investigated four different frame rates and three different vector rendering modes (line, pyramid, and complex arrow). The pyramid consists of 5 polygons while the complex arrow has 34 polygons (see figure 2). Thus, complex arrows take considerabely longer to draw than pyramids while lines are the fastest mode. The intention was to obtain an overview of possible interactions between these two parameters, i.e., how to choose the best rendering mode for a given rendering speed and vice versa. Twelve subjects were used for these tests; they were computer science students with basic knowledge of computer graphics but no experience in visualization of vector data. All tests were performed using the ISVAS[†] 3.2 visualization system running on SGI hardware. The analysis of variances (ANOVA) was performed using the SAS/STAT software.

Two pilot tests were designed and performed on a few subjects; the results of these tests led to two more elaborated tests which were performed on all twelve subjects. For example, in the pilot test the users interacted with the system in order to select a point of view. We soon realized that this leads to test results which are mainly influenced by the strategy of each user instead of the parameters which were of original interest to us (e.g., rendering speed and rendering mode).

For the sake of brevity we will only report on one of the two final tests, the so-called 'detail' test.

† ISVAS is registered trade mark of Fraunhofer IGD, Darmstadt, Germany

2.2. 'Detail' Test

In the 'detail' test, four vectors (all of them rendered in the same of the three possible rendering modes) were shown on a flat surface (see figure 3). The scene was rotating about the vertical axis at one of the four possible speeds (2, 6, 12, or 18 frames per second) at one degree per frame.

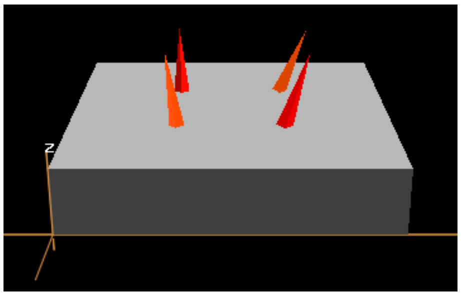

Figure 3: *Detail test: which of the four vectors is most orthogonal to the surface?*

The subject's task was to find the one vector which was most orthogonal to the surface as well as to estimate the exact angle between this vector and the surface. This experiment was motivated by the need to determine the correctness of boundary conditions in finite element simulations.

After a number of training cases to make the subject acquainted with the task, the test itself started. It consisted of 12 test cases. The whole session took approx. 60 minutes. For each test case we recorded

1. the 'most orthogonal' arrow as perceived by the subject,
2. the subject's confidence in this decission,
3. the estimated angle of the most orthogonal arrow to the plane,
4. the number of frames that were viewed until the correct arrow had been selected, and
5. misc. comments of the subjects.

We say that items 1 and 3 measured 'objective' entities since we have no reason to assume that these answers were biased by expectations. On the other hand, we say that items 2 and 4 measured 'subjective' entities since we believe that these results are influenced by the subjects' expectations. For example, if the subject expects that at a given frame rate it is more difficult to estimate an angle, the subject will take longer to decide on this angle and the number of viewed frames (4) will be increased.

Now, let us discuss the test results. First, the resulting number of frames (4) are shown in figure 4.

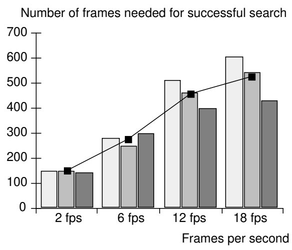

Figure 4: *Number of viewed frames for the frame rates (2, 6, 12, 18 fps) and the three rendering modes (left to right in each block: line, pyramid, arrow).*

It may be seen that for higher frame rates, more frames are viewed. This is due to the fact that the whole scene needs to be viewed for at least 20 seconds which is equivalent to 40 frames at 2 fps or to 400 frames at 20 fps (increasing number of frames with increasing frame rate).

By the way: If we use 'viewing time in seconds' instead of 'frames per second' in figure 4, the mean values of each block are decreasing from 2 to 18 fps. In this case, we get approx. 70 seconds for 2 fps, followed by 45, 37.5, and 30 seconds viewing time for 6, 12, and 18 fps. So the overall viewing time is decreasing and thus performance of a user is increasing with higher frame rates.

But we can observe something much more interesting: at 2 and 6 fps, all three rendering modes seem to be more or less identical, but at 12 and 18 fps, there is an advantage of arrows over pyramids over lines.

Source	DF	F-Value	p
FPS	3	29.14	0.0001
REND	2	0.81	0.4458
FPS*REND	6	0.49	0.8121

Table 2: *ANOVA for number of frames.*

The overall ANOVA for the number of frames leads to the numbers given in table 2. We can see that frame rate (FPS) has a strong (F=29.14) and significant (p=0.0001) influence on the number of frames. So the first observation from the graph is confirmed: with increasing frame rate, the number of frames increases significantly. On the other hand, rendering mode (REND) is not very strong and not significant. To be significant, p should be < 0.05 or at least < 0.1. So our second observation about arrows being better than pyramids and lines is only a trend but not a significant result. Finally, the interaction of these two variables is even less significant.

Source	DF	F-Value	p
REND (FPS= 2)	2	0.11	0.8995
REND (FPS= 6)	2	0.55	0.5837
REND (FPS=12)	2	1.22	0.3068
REND (FPS=18)	2	1.13	0.3342

Table 3: *ANOVA for number of frames for each individual rendering speed.*

We were looking for some influence of the rendering mode on user behaviour, so we were not satisfied with these results. Next, we investigated if the influence of rendering mode would be approximately the same for all rendering speeds. Table 3 shows the results of this ANOVA.

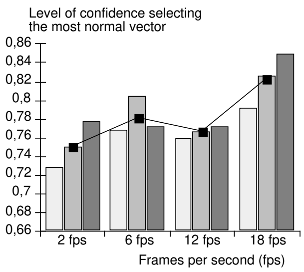

Figure 5: *Level of confidence for the frame rates (2, 6, 12, 18 fps) and the three rendering modes (left to right in each block: line, pyramid, arrow).*

We can see that for increasing frame rates, the rendering mode has increasing influence on the time subjects are examining the scene. At 2 fps, the three rendering modes are more or less identical in terms of user performance, but at 18 fps, arrows seem to be clearly superior to pyramids or to lines, even though this influence is still not signficant in our test (p=0.3342 is still too large).

The second subjective result (level of confidence, 2) is shown in figure 5. It can be seen that generally, confidence is higher with arrows than with pyramids or lines and that confidence is higher at high frame rates. ANOVA is not given here but again, the results are not significant.

Next, let's have a look at the 'objective' results of selection error: figure 6 shows the error in selecting the most orthogonal vector (1).

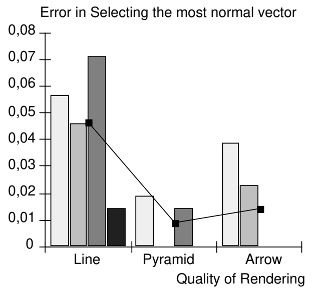

Figure 6: *Error in selecting the most orthogonal vector: the rendering modes and four frame rates each (left to right in each block: 2, 6, 12, 18 fps).*

We can see that for these 'objective' results, lines are not well suited. Pyramids are slightly better than arrows. Also, high frame rates (dark bars) are better for test results than low frame rates (white bars) for the given task. ANOVA confirms that these results are not by chance (see table 4), and that the influence of rendering mode is much more significant than the one of frame rate.

Source	DF	F-Value	p
FPS	3	1.54	0.2067
REND	2	4.66	0.0111
FPS*REND	6	0.48	0.8193

Table 4: *ANOVA for selection error.*

Finally, figure 7 shows the 'objective' results of errors on estimation of angle. We can see that pyramids perform better than lines or arrows but it's hard to tell what frame rates are best: for lines, low frame rates (white) are better than high frame rates (dark), but for arrows, the highest frame rate (18 fps, dark) is best.

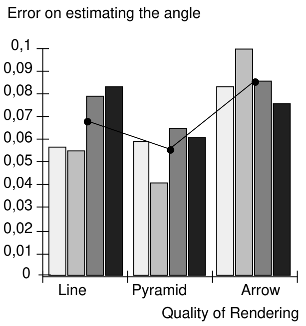

Figure 7: *Error in estimating the angle: the rendering modes and four frame rates each (left to right in each block: 2, 6, 12, 18 fps).*

Concluding the discussion of this test, it can be seen that for feature extraction (error in selection, figure

6) high frame rates are better than low frame rates while for exact measurements, this is not so clear. In fact, if lines are used as drawing primitives, low frame rates are superior to high frame rates.

Another interesting result is that for both cases, pyramids seem to be the best choice for drawing primitives. This is an important finding: the most complex rendering objects need not necessarily be best suited for a given task. Since pyramids are faster to draw than detailed arrows, there seems to be no need for detailed arrows in a visualization system if only tasks similar to our tests are to be solved. But please note that we do not claim that pyramids are superior to detailed arrows in all or even in most cases; to make such a provocative statement, much more test cases would be needed.

From our observations during the tests as well as from comments of test subjects we learned something that we did not expect in the beginning: even in a quite homogenious group of subjects as in our test (all 20 to 30 years old, computer scientists with computer graphics backgroud, all from Darmstadt, etc.) there were different strategies in solving the tasks of the test. For example, in order to find the most orthogonal vector, some subjects would try to perceive the whole scene at once and see which vector would 'jitter' fewest while others would try to determine each angle sequentially and then calculate which one is most normal to the surface.

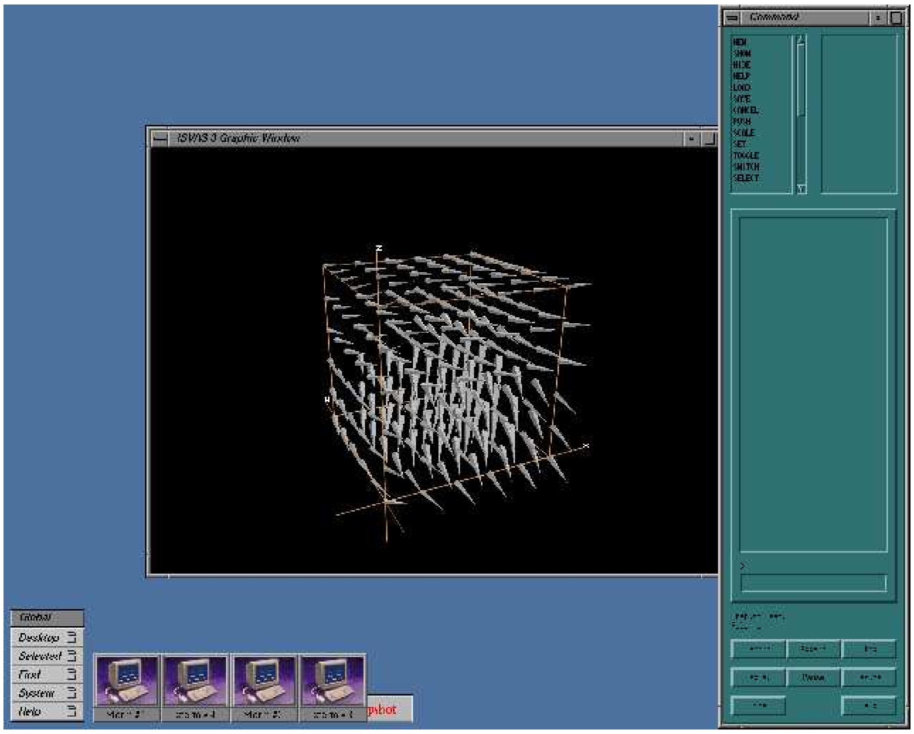

Figure 8: *Test environment and one view of the second test ('overview').*

In the second test ('overview'), a regular grid of vectors was given (6 x 6 x 6). Figure 8 shows a screen shot with test environment and the vector field in the middle. For the tests, a 21 inch monitor was used. As shown in the figur, vectors were not identical but changed slowly in lengthand direction from one corner of the cube to the other. Test subjects had to find a discontinuity of this function. Here, we made a similar observation: Some subjects would try to perceive the whole scene at once in a holistic manner while others would try to use certain orientations of the dataset to 'take a bearing of' rows of vectors. These different strategies also lead to different preferences: The holistic group needed a certain frame rate to get a good 3D impression of the scene while the second group prefered lower frame rates where there is more time to investigate the data set in a 'stable' orientation.

In conclusion, some interesting things that we learnd from the test are:

- Personal preferences and objective performance need not be correlated.
- Personal strategies for interaction and investigation of a scene influence the effects of visualization parameters (e.g., of frame rate and of rendering mode).
- Humans are pretty good in detecting a feature and much worse in estimating a numerical value.
- The formulation of a task and the goal of the viewer determine his perception by making him select alternative strategies.
- Kind and amount of data influence the suitability of different graphical primitives for scientific visualization.

2.3. Some things we learned about tests in general

Testing the effect that different parameters of a graphics system might have on a user's performance involves detailed preparation of the psychological test. This includes precise ideas about parameters to be tested, the kinds of subjects (e.g., experts or laypersons), test data, perceptual tasks, and performance indicators. These items are not independent of each other.

The number of parameters (e.g., rendering mode and rendering speed) and the number of values of each parameter (e.g., 8, 12, 16 frames per second for rendering speed) that will be tested have a great influence on the number of tests that must be conducted to gain significant results. Experienced psychologists claim that significant test results require the number of subjects to be in the order of 10 times the number of all combinations of parameters. This means that for our "detail" test where we tested 3 x 4 = 12 parameter combinations, approx. 120 test subjects would be needed.

Statistical significance also is influenced by the homogeneousness of the subjects. If all subjects have very similar background, there will be a higher confidence level for the test results, but the results will only be valid for a smaller subset of possible users. If the subjects are very different from each other, results will be less significant but apply to a wider range of persons.

Also, to avoid learning effects, there should be as many datasets as the number of trials for each subject. There need to be some additional training data sets for each subject to make it familiar with the task before the actual test starts. All datasets should be of similar difficulty. Since it may be impossible to get enough sets of real world data which fulfill all these criteria, it can be better to generate synthetic data. This also allows a precise variation of the parameters under investigation.

The task must be apropriate for the test subjects. E.g., the task "determine if there is a tumor visible in a medical data set" can be suitable for an experienced physician but not for a layperson.

The solution of a task should be easy to verify for the test personnel. Errors introduced by verbalization should be minimized. An example for this is the perception of angles. If test subjects are asked to sketch a perceived angle onto a piece of paper, this result generally is much better than a verbal estimation of the angle.

Some general recommendations concerning psychological tests with graphics systems:

1. Make a hypothesis and select a task which will be suitable to evaluate this hypothesis.
2. Select a simple task which is not influenced by too many variables; try to eliminate any side effects which may influence the perfromance of your subjects.
3. Use a large number of subjects in order to minimize personal differences.
4. Use a homogenious group of subjects with similar background.
5. Design and carry out preliminary tests in order to improve your test setup, take plenty of time for this.
6. Try to get the help of someone who has experince in psychological tests in case you don't have this experince yourself.
7. Know your statistical methods and your software packages for analyzing test results well.

3. Conclusion and Future Work

In this paper we have presented a qualitative scheme for Visualization Sytem Quality, we have reported on perceptual tests that were performed in order to gain more quantitative input for such a scheme, we have discussed the results of these tests, and we have given hints to test design in general. Especially the strong influence of user goals and user knowledge on the best settings for Visualization System Quality in each case has been demonstrated.

There is need for further investigation and classification of rendering methods and other factors influencing Visualization System Quality. Only a quantitative metrics of Visualization System Quality will allow us to build systems which can automatically select the best combination of image quality and interaction quality on a given hardware configuration for each specific set of data and for each individual user.

If we want to do it right, if we want to ensure the quality of scientific visualizations, the role of human perception in scientific visualization must be further investigated.

4. Acknowledgements

We wish to thank all our subjects for participation in the tests, all those who contributed to the ISVAS visualization system (Kennet Karlsson, Peter Fritzen, etc.), and Prof. Dr.-Ing. J. Encarnação for providing the environment which made this work possible.

References

1. Astheimer, P., Dai, F., Göbel, M., Kruse, R., Müller, S., Zachmann, G.: *Realism in Virtual Reality*, In: Thalmann, N., Thalmann, D. (Eds.): Artificial Life and Virtual Reality, John Wiley & Sons, 1994, pp. 189-210

2. Bryson, S., Levit, C.: *The Virtual Windtunnel*, IEEE Computer Graphics and Applications, 12, 4, pp. 25-34, 1992

3. Bryson, S., Levit, C.: *Lessons learned while implementing the virtual windtunnel project*, Visualization '92, Tutorial # 2, 4.1–4.7, 1992

4. Card, S.K., Moran, T.P., Newell, A.: *The psychology of computer human interaction*, Lawrence Erlbaum Associates, Inc. Publishers, Hilldale, 1983

5. Cruz-Neira, C., Leight, J., Papka, M., Barnes, C., Cohen, S.M., Das, S., Engelmann, R., Hudson, R., Roy, T., Siegel, L., Vasilakis, C., DeFanti, T.A., Sandin, D.J.: *Scientists in Wonderland: A Report on Visualization Applications in the CAVE Virtual Reality Environment*, Proc. IEEE Symposium on Research Frontiers in VR, San Jose, pp. 59–66, October 1993

6. Encarnação, J.L., Astheimer, P., Felger, W., Frühauf, Th., Göbel, M., Müller, S.: *Graphics & Visualization: The Essential Features for the Classification of Systems.*, In: Proc. International Conference on Computer Graphics ICCG'93, Bombay, India, February 1993

7. Frühauf, T., Göbel, M., Haase, H., Karlsson, K.: *Design of a Flexible Monolithic Visualization System*, in: Rosenblum, L., Earnshaw, R., et al. (eds.): Scientific Visualization - Advances and Challenges, Academic Press, London, 1994, pp. 265-286

8. Gerfelder, N., Müller, W.: *Quality Aspects of Computer Based Video Services*, in: Proc. of the 1994 European SMPTE Conference, Cologne, Germany, pp. 44–67, Sept. 1994

9. Gibson, J.J.: *The ecological approach to visual perception*, Houghton Mifflin, Boston, 1979

10. Globus, A., Uselton, S.: *Evaluation of Visualization Software*, Computer Graphics, pp. 41-44, May 1995

11. Groß, M.: *Visual Computing - The Integration of Computer Graphics, Visual Perception and Imaging*, Springer, 1994

12. Haase, H., Göbel, M., Astheimer, P., Karlsson, K., Schröder, F., Frühauf, Th., Ziegler, R.: *How Scientific Visualization can benefit from Virtual Environments*, CWI Quarterly, The Netherlands, 1994

13. Haase, H., Strassner, J., Dai, F.: *Virtual Molecules, Rendering Speed, and Image Quality*, in: Göbel, M. (ed): Virtual Environments '95, Springer Verlag, Wien, pp. 70-86, 1995

14. Haase, H.: *Symbiosis of Virtual Reality and Scientific Visualization System*, Computer Graphics Forum, Vol. 15, No. 3, Proceedings Eurographics '96, Poitiers, France, August 1996

15. Krüger, W., Bohn, C.-A., Fröhlich, B., Schüth, H., Strauss, W., Wesche, G.: *The Responsive Workbench: A Virtual Work Environment*, Computer, Vol. 28, No. 7, pp. 42-48, July 1995

16. Lange, S., Schumann, H., Müller, W., Krömker, D.: *Problem-oriented visualization of multi-dimensional data sets*, Proc. International Symposium on Scientific Visualization '95, Chia, Italy, World Scientific Pub., Singapore/London, pp. 1-15, 1995

17. Rheingans, P.: *Dynamic Exploration of Multiple Variables in a 2D Space*, Ph.D. Thesis at the Department of Computer Science, University of North Carolina at Chapel Hill, 1993

18. Robertson, P.K., Silver, D.: *Visualization Case Studies: Completing the Loop*, IEEE Gomputer Graphics and Applications, pp. 18-19, July 1995

19. Rogowitz, B.E., Treinish, L.A.: *Using perceptual rules in interactive visualization*, Proceedings of the IS&T/SPIE Symposium of Electronic Imaging Science and Technology, San Jose, Feb. 1994

20. Sherman, W.R.: *Integrating Virtual Environments into the Dataflow Paradigm*, Fourth Eurographics Workshop on Visualization in Scientific Computing, Abingdon, UK, April 1993

21. Treinish, L., Silver, D.: *After the Storm: Considerations for Information Visualization*, IEEE Gomputer Graphics and Applications, pp. 12-15, May 1995