Enhancing User Cognition in E-learning Scenarios with 3D Web Applications: A Comparison of 2D and 3D Visualizations with HTML5

Gabriel Rausch
Bauhaus-Universität Weimar (Germany)
gabriel.rausch@uni-weimar.de

Abstract

Introduction
As the web has evolved into a mature platform for building Rich Internet Applications (RIA), technologies to render three-dimensional elements in the browser have become commonplace [1]. While it is tempting to believe that the visual potential of 3D on the web will enhance users' perception of information and facilitate new approaches to how we learn and gather information, numerous previously proposed approaches to 3D e-learning scenarios, such as Google's "Zygote Body", have not yet been widely adopted [2]. Building on the work of Tavanti and Lind, Cockburn's fundamental research established general behaviors regarding users' perceptions of 2D and 3D visualized information [3, 4]. Focusing on the web interfaces proposed by Jankowski suggests an approach that mixes 2D and 3D elements to boost users' cognition [5]. We developed the foundations of this project to explore the potential of browser-based user manual scenarios as a method to evaluate the effectiveness of user cognition in the presence of three-dimensional elements.

Test Setup and Methodology
Comparing different visualization technologies to evaluate the potential of 3D on the web, we created a series of browser-based user manuals for electrical and mechanical products in collaboration with our industry partners.
For each product, three test applications were prepared:

- *An interactive 3D version with live rendered models based on WebGL*
- *A 2D version with images and illustrations*
- *A video version to describe the user manual in animated sequences*

In order to both gain deeper insight intro the thoughts of users and to gather empirical qualitative data, a combination of methods from usability evaluation and field studies was applied, including the think-aloud-method and monitoring. The evaluation of the data follows the methodology set forth by Mayring [6].

Results

- *Spatial cognition: Participants coped with tasks more effectively while using the 3D interface than while using the video or 2D versions.*
- *Quality of visualization: Participants using the 3D interface were convinced by the visualization and felt entertained by the task procedure.*
- *Contextual References: representations information should be simplified.*
- *Usability: 3D interactions yielded difficulties for most of the test subjects, especially older participants.*
- *Accessibility: 3D on the web requires a high-level hardware and software setup. This technical condition limits the potential user base.*

1. Introduction

With the rise of ubiquitous access to digital information and the widespread adoption of interactive technologies at the consumer level, e-learning has assumed a central role in the field of education. As the technical underpinnings of e-learning continue to develop further, the occurence of new approaches for the presentation and consumption of information is all but assured. The latest HTML5 specification defines native browser support for three-dimensional elements rendered with WebGL, which allows publishers and administrators to visualize and arrange information in virtual 3D space without proprietary plugins such as Flash, Silverlight or Java Applets [1,3]. While it is tempting to believe that the visual potential of 3D on the web will enhance users' perception of information and facilitate new approaches to how we learn and gather information, numerous previously proposed

approaches to 3D e-learning scenarios, such as Google's "Zygote Body", have not yet been widely adopted [2]. The effectiveness of an interface is strongly predicted by users' spatial memory and aptitude. This conjecture has been confirmed in many studies conducted on a variety of interface types since the 1980s [4]. Building on the work of Tavanti and Lind [3], Cockburn's fundamental research analyzed the implications for spatial memory with regards to location learning. Cockburn suggested a 2D solution for navigation concepts and virtual finding scenarios. Jankowski's subsequent work proposed an approach that mixes 2D and 3D [5], especially for visualizing information in virtual interfaces that are already spatial in the real word. For the experiment described in this paper, the potential of browser-based user manual scenarios is explored as a method to evaluate the effectiveness of user cognition in the presence of three-dimensional elements rendered with the latest browser technologies. HTML5 and WebGL offer several advantages in visual quality and interaction performance, and as such, more effective user cognition when using this technological framework is assumed.

2. Methodology

In this study, different visualization technologies were presented to participants to gather qualitative data on emotional and cognitive parameters to evaluate the potential of 3D product manuals on the web. A series of browser-based manuals for electrical and mechanical products was created, where three test applications were prepared for each product manual. Each application contained the same product description, but was differentiated from others by core aspects of the visual and technical presentation, including:

- An interactive 3D version with live rendered models based on WebGL
- A 2D version with images and illustrations
- A video version to describe the user manual in animated sequences

2.1 Recruiting

Shoppers were approached at the point of sale and asked to complete a brief questionnaire to gather basis data. Participants in the study were then selected depending on the results of the introductory survey. The selection is intended to accurately represent the target group for selected products produced by industry partners, and to ensure participants have basic motivation and comptence to address the presented topic. In total, 21 participants were selected for this trial.

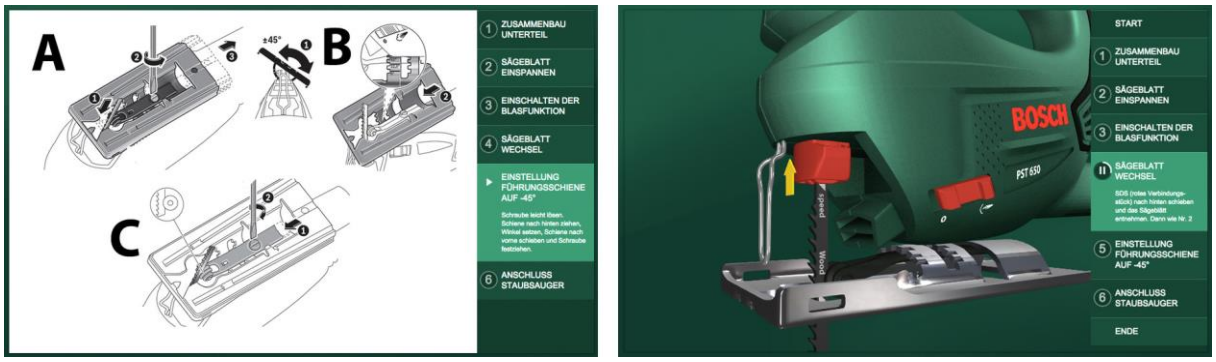

Figure 1: Screenshot of 2D (left) and 3D (right) manual for a jigsaw

2.2 Procedure

A pre-experiment and post-experiment questionnaire established the scope of demographic data, as well as participants' attitudes and expectations toward computers, e-learning and the Internet. In the core evaluation, candidates were requested to perform a task, such as to assemble a jigsaw, and provided with a test manual. In order to gain deeper insight into users' thoughts and to gather further qualitative material, a combination of methods from usability evaluation and field studies was applied, including the think-aloud-method and monitoring.

3. Data Analysis

According to Mayring, the first step to analyze interview material is to transcribe it [6]. Similar methodology is also found in Strauss' more complex method [7], as well as the simple 8-step method from Chi [8]. After segmenting the sampled protocol, a coding scheme is chosen. Generally this

coding is intended to reduce the amount of transcribed data and reduce the number of statements derived from them [9]. Specific phrases from the transcript are assigned to keywords from the chosen coding scheme. In order to find tendencies that 3D with WebGL support user cognition, two main coding categories formalize both pro and contra statements.

4. Results

4.1 Spatial Cognition

Concurring with previous studies, this experiment showed that user cognition of spatial elements is more efficient when objects are presented in virtual 3D space, as opposed to visualization as a video or are series of 2D images. Participants provided with a 3D manual successfully completed tasks unambiguously faster and more accurately. In particular, the group provided with 2D manuals struggled with static images and limited views, which often led participants to attempt to perform the task by trial and error, as the limited set of images could not provide all possible points of interest. This result substantiates Jankowski's work that suggest visualizing objects in 3D space when they are spatial in the real word already [5]. The group provided with video manuals, intended as a hybrid visual approach between 2D and 3D, were generally more successful at task completion in comparison to the 2D group. The cognitive burden to translate an animated object from a video to a real object is less complex than from still images. Yet difficulties for the video group were found in controlling time-based parameters, such as to pause, resume or repeat the video between task steps, as well as the limited, pre-defined point of views.

4.2 Quality of Visualization

The visual attraction of the manual is an emotional factor that influences users' satisfaction, which in turn affects motivational aspects inherent to intensively concentrating on a task. To visualize the objects using a modern approach built on WebGL, we developed a render algorithm to optimize the quality of the output result. As the technical details of the implementation are outside of the scope of this paper, the render algorithm will be described in future work. In comparison to 2D illustrations and the previous approaches undertaken by Moritz [10], which investigated the impact of 3D in the browser with Flash and Java Applets, the quality of visualization has been highly refined using the technical framework based on WebGL. Participants were strongly positively effected by the visual quality of the 3D manual, and many found the virtual representation of the objects to be photorealistic, while some participants had even speculated that the content was being produced with a camera in real-time.

4.3 Usability

Aside from standard issues with web usability, interaction with virtual 3D interfaces also presents substantial usability risks, especially if the target group is not experienced with computer games. The affordance of interaction was generally unclear in the 3D manual. Some participants initally addressed the 3D model as a static image; only after an information tooltip visualized could they determine the behavior of 3D interaction. As such, a navigation element was specifically developed to minimize some pitfalls of poor usability. Building on the "common sense approach" to usability described by Krug [11] this element was assembled to allow for simple control of the point of view and to switch between different task descriptions based on 2D elements. Despite this element, some participants, especially those that were older and less experienced with digital interfaces, still encountered difficulties in interaction, which correlates to the results described by Cockburn on the impact 3D vs. 2D in location learning in virtual interfaces [4].

4.4 Accessibility

The technical requirements necessary to render and access the 3D manuals specify an advanced hardware and software setup. Even though HTML5 and WebGL are specified to be platform agnostic and perform across multiple operating systems and browser environments, the minimum technological requirements are not yet met by the majority of our target group. This is especially true for older participants, who reported home computer setups that would not support access the 3D test application. This issue was addressed in this experiment by pre-assembling a suitable hardware and software enviornment purpose-built for this trial. The second issue is the general problem of making multimedia content in the browser accessible to those with disabilities. In contrast to accessbility technologies for other media, such as subtitles in videos or alternative descriptions for images for screen readers, there are currently no existing approaches for providing support for 3D to those with visual impairment other than providing a simple fallback text message.

4.5 Contextual References

Ariely suggests that a high-performance and comprehensive interaction can improve a user's overall satisfaction and postively effect the motivation to continue to attempt a task. [13]. In this approach, secondary elements were excluded to preserve a balance between performance and interaction. As an example, the task scenario of assembling a jigsaw is depicted without a human actor or a hand showing the installation. Such additional objects greatly increase the amount of polygons and scene complexity, which in turn negatively effects render performance, however omitting these contextual references can cause difficulties for the correct completion of certain tasks.

5. Conclusions and Suggestions for Further Research

The approach to visualize a digital 3D user manual directly in the browser using HTML5 and WebGL has demonstrated positive results, while also revealing insights to the effectiveness of spatial user cognition and how the quality level of a visualization enhances user satisfaction. The innovative and experimental character of 3D in the browser has shown its potential to engage users' attention, and in doing so, to invoke users' emotion in a manner that positively influences cognitive processes. The majority of participants demonstrated positive emotional feedback, while only a small number of older, more conservative users found the 3D manual to be superfluous. Less positive aspects of this approach stem from broader usability issues that are not limited to web technologies. In order to avoid numerous commonplace usability pitfalls, interaction affordances for the user must be improved, and developers must ensure that simple, clean 3D applications for education become more widely available. The relevance of these applications in a commercial space is especially relevant when considered with the context of Davis' Technology Acceptance Model, where the factors influencing a user's acceptance of perceived usefulness and ease-of-use are strongly predicted by social parameters including norms, brands and images [14]. However, the topic of accessibility for users with disabilities is a more challenging task requiring the main actors in browser specification and browser production to develop consistent fallbacks and alternative output formats.

References

[1] Khronos Group, "Khronos Releases Final WebGL 1.0 Specification at," 2011.
[2] R. Zeiger, "Google Body becomes Zygote Body; built on open source 3D viewer," Google Open Source blog, 09-Jan-2012. [Online]. Available: http://google-opensource.blogspot.de/2012/01/google-body-becomes-zygote-body-built.html. [Accessed: 27-Jan-2012].
[3] M. Tavanti and M. Lind, "2D vs 3D, Implications on Spatial Memory," in Proceedings of the IEEE Symposium on Information Visualization 2001 (INFOVIS'01), Washington, DC, USA, 2001, p. 139.
[4] A. Cockburn, "Revisiting 2D vs 3D implications on spatial memory," in Proceedings of the fifth conference on Australasian user interface-Volume 28, 2004, pp. 25–31.
[5] J. Jankowski and S. Decker, "A Dual-mode User Interface for Accessing 3D Content on the World Wide Web," in Proceedings of the 21st International Conference on World Wide Web, New York, NY, USA, 2012, pp. 1047–1056.
[6] P. Mayring, Einfuehrung in die qualitative Sozialforschung. Weinheim: Beltz, 2002.
[7] J. Corbin, Basics of qualitative research. Los Angeles: SAGE, 2013.
[8] Michelene T. H. Chi, Quantifying qualitative analyses of verbal data, University of Pittsburgh: Lawrence Erlbaum Associates, Inc., 1997.
[9] L. M. Given, The SAGE Encyclopedia of Qualitative Research Methods. SAGE Publications Inc., 2008.
[10] F. Moritz, Potentials of 3D-Web-Applications in E-Commercein Computer and Information Science (ICIS), 2010 IEEE/ACIS 9th International Conference, 2010, pp. 307–314.
[11] S. Krug, Don't make me think, revisited. Berkeley, Calif: New Riders, 2002.
[12] Tim Berners-Lee, World Wide Web Consortium Launches International Program Office for Web Accessibility Initiative, 1997.
[13] D. Ariely, Controlling the information flow, Journal of Consumer Research, vol. 27, no. 2, pp. 233–248, 2000.
[14] F. Davis, A technology acceptance model for empirically testing new end-user information, Massachusetts Inst. of Technology, 1985.