Comparison of Digital Human Model-Based Ergonomic Software Using Eye-Tracking Methodology – Presenting Pilot Usability Tests

Mária Babicsné Horváth[1(✉)], Károly Hercegfi[1], and Tamás Fergencs[2]

[1] Department of Ergonomics and Psychology, Budapest University
of Technology and Economics, Magyar Tudosok krt. 2, Budapest 1117, Hungary
{bhorvathmaria,hercegfi}@erg.bme.hu
[2] Information Studies Master's Program, Aalborg University,
A. C. Meyers Vænge 15, 2450 Copenhagen, Denmark
tfergel8@student.aau.dk

Abstract. Nowadays, analysis software for ergonomics are more and more wide spread among researchers and ergonomists. There are several different ergonomic software available on the market, and it would be important to know which one to choose in various applications. Although data on the capabilities of the software can be found easily, a comparison regarding their ease of use and the quality of their user interface cannot be found in the literature. In this article, a methodology for comparison is presented along with the results of a pilot tests. For the pilot tests, a cloud based software, ViveLab was chosen. We could draw conclusions for the software and the methodology, as well. Regarding the usability of the software, many suggestions can be made, however, not all function was tested. We could interpret the eye movements in the model space only with strong limitations. Based on the experiences of the pilot tests, we will insert a short subtask into the protocol of the real usability tests, when we will ask the participants not to move in the model space. In conclusion, except a few changes, the method can be considered as applicable for the further usability tests.

Keywords: Usability testing · Eye-tracking ·
Human Factors and Ergonomics (HFE) · Digital human model ·
Computer Aided Anthropometric Design (CAAD) ·
Software for ergonomic risk assessment · ViveLab

1 Introduction

Nowadays, analysis software for ergonomics are more and more wide spread among researchers and ergonomists. These can be used for anthropometric design (fit tests, reachability analysis, etc.), and ergonomic risk assessment [1]. These are most often used for evaluating manufacturing processes [2], or in the field of human-machine interaction, e.g., in vehicle design [3]. There exist several different ergonomic software available on the market, and it would be important to know which one to choose in various applications.

Although data on the capabilities of these software can be found easily, a comparison regarding their ease of use and the quality of their user interface cannot be found in the literature. Therefore, a comprehensive study about the usability of these software tools could mean scientific novelty. NB, in a research done in Sweden [4], digital human models were compared, and in another examination in the Czech Republic [5], production workplace models were tested, focused on comparing the carrying, lifting-lowering, and biomechanical conditions. There also exist a research for comparison of different lifting tools to estimate spine loads during static activities [6], however, in this case, they did not always use digital human model. Our research focuses on the different features of the different ergonomic software and the usability of the interfaces.

We developed a methodology for comparison. In this article, the methodology is presented along with the results of a pilot research. For the pilot, a cloud based software, ViveLab [7] was chosen, because it is a new, Hungarian software with notable features, and it was easily accessible for us. It is primarily used for risk assessment of industrial workplaces, ergonomic evaluation of products, reachability tests, and path review (with spaghetti diagram).

For the future tests we plan to use the current version of Jack [8], as probably the most known ergonomic software, and as many other software as we can reach e.g. RAMSIS, SAMMIE, DELMIA, Anybody, SANTOS, IMMA [9].

2 Methods and Tools

During the pilot phase of the research several methods were used. Before the usability test, three experts were interviewed. During the test, we apply eye-tracking method with think aloud technique, and, after that, each participant was interviewed.

2.1 Eye-Tracking

The eye-tracking methodology in human-computer interaction is a well-known tool for measuring usability or user experience [10]. Several researches was made in the field of web design [11] and other human-computer interaction fields [12, 13]. This method can give us additional information about the users' behavior [14]. Combining the conventional usability test with the eye-tracking method, more data can be acquired; furthermore, its visualization techniques, can support us to interpret these data in a relatively efficient way.

In our research, we used a monitor based Tobii T120 eye-tracking device. There are two cameras in the monitor, one is to record the participants' movements, gestures, and facial expressions, and another one to determine the gaze. The device also record the computer screen, and as a result we can get a video with the eye movement, heatmaps [15], AOI (Area of Interest) statistics [16, 17], and gaze plot diagrams [18].

2.2   ViveLab

ViveLab is a digital human model-based software for ergonomic analysis. The software was released in 2015. This is a cloud-based software, which is one of its main advantages. It means the shared model spaces (so called "virtual labs" or simply "labs") can be available from all countries, only a utility software, Citrix Receiver must be installed. Everyone can register on the webpage [7] and get a one-hour trial license.

After login, we can create labs, and work with coworkers in the same lab at the same time. The built-in human model has a database of accurate body dimensions. We can adjust the percentile, age, somatotype, and acceleration of our human mannequin. The parameter of acceleration means we can adjust the birth year of the human, since, over the decades, the later they are born the higher they will be. There is an opportunity for import xsens motion capture file, and we can create our own animation manually as well. The software includes three implemented risk assessment methods (RULA, OWAS, NASA-OBI), two implemented standards (ISO 11226, EN 1005-4), and two other analysis techniques (reachability zone, spaghetti diagram). After analyzing the human motion and/or postures, we can generate risk assessment documents and ask for statistics.

3   Protocol of the Pilot Tests

The pilot research was made exactly like the planned real tests. We defined a divided user profile, and the participants were recruited in accordance of it: University students and university teachers both were among the participants. Six person was participated, three students and three teachers.

In the case of the actual six participants, the students had much more active experiences with CAD programs than the teachers had, which could give us better task completion time or error-free rate. On the other hand, the teacher participants usually had much more knowledge about risk assessment and anthropometric fit.

The protocol of the pilot tests was the following. After the calibration of the eye-tracker, the participants have to complete the given tasks. The tasks for the participants were the following:

1. Ergonomic verification of a product in the design stage. The aim is to analyze the reachability and the differences in the measurements of the components.
2. Ergonomic risk assessment for a body posture in the use of a product or for a particular posture during work.

The particular tasks targeted ergonomic analysis of an adjustable height table. The required steps were the following:

- Opening the given ViveLab project.
- Importing the given object to the lab as a model of a machine.
- Placing a human model.
- Setting the body posture of the human model to "stand – typing on computer".
- Placing the human model near the table. The forearm should be over the table top.

- Adjusting the properties of the human model to the specific parameters: work clothes; female; Swedish; database: Bodyspace; age: 27; percentiles: 5%. Participants did not have to change other attributions.
- Setting the height of the table top to touch it with the forearm of the human. Determining whether the end stops of the adjustable table allow this proper height or not.
- Modifying the properties of the human to another specific parameters: male; percentiles: 95%; without any changes in other attributions.
- Repeating the task in accordance to the new human: Setting the height of the table top to touch with the forearm of the human; determining whether the bumpers allow this to be carried out.
- Ascertaining the anthropometric compliance of the table.
- Opening the RULA risk assessment tool.
- Determining the load caused by the angle of the neck based on the RULA tool.

The simplicity of the task (assessing just a simple adjustable table, not a complex machine), and the level of detailing the steps in the protocol were necessary to carry the sessions in a comfortable period and with a reasonable effort by novice users as well. However, we consider these steps as representative for practical, industrial usage of the main feature of the tested software and the other pieces of software to be tested later.

4 Results of the Pilot Tests

As the results of the tests, we have identified many usability problems regarding the software. We gained two types of data: qualitative and quantitative. The qualitative data came from the eye-tracking visualization techniques (heatmaps, gaze plots) and interpretation of the statistics of AOI. The quantitative data came from the task completion time and the success rate.

### 4.1 Qualitative Results from the AOIs, Heatmaps, and Gaze Plots

The first subtask was opening the software, which was almost the easiest subtask, though there was one problem regarding the completion time. According to the fast and wide saccades, we could identify a searching behavior of the participants. This means they assumed that the program will indicate how much time left or the status of the loading.

Importing the CAD model was the next subtask. This subtask was one of the hardest. Based on their previous experiences with other similar programs (CAD programs), all participants searched for a file menu. Therefore, they tried to find this tool mainly the top left corner and under the "Main" menu as we can see on the heatmaps. (See Fig. 1.)

According to two participants' individual gaze plots showed in Fig. 2, we can see the eye movements on the horizontal menu bar. The particular participants returned and searched again on the same part of the user interface. That means they were sure that
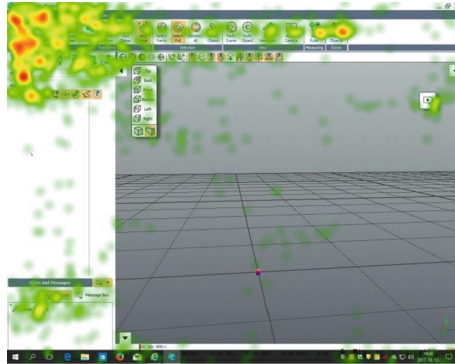
Fig. 1. Importing the CAD model – aggregated heatmap for all participants.
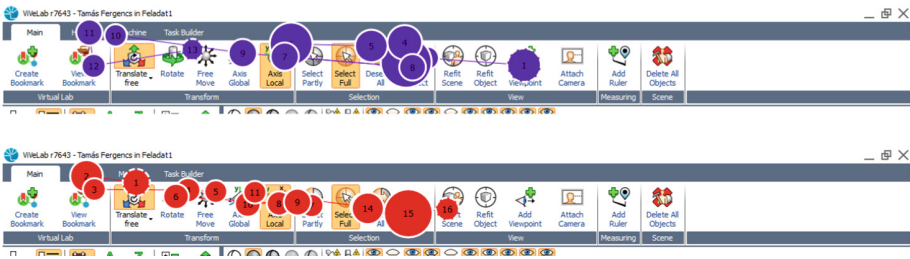
Fig. 2. Importing the CAD model – gaze plots: saccade regressions for two participants.

the menu should be there, but they could not find it, or the pictograms were not evident enough. Some of them tried the "drag and drop" method unsuccessfully.

Two of the participants found the "Add machine" menu, however, they did not click on it. After the moderator informed the participants that the software considers the CAD model as a machine, all of them found the correct user interface element in 8 s.

Inserting a human model was the next problematic subtask. The label of the correct button was "Diving suit" and not "Add human" or "Create human". This was confusing especially for the students. The Heatmap showed by the left image of Fig. 3 also confirms, with the red parts, the long fixation time and the inactivity, which indicates the difficulty of interpretation of the object.

The AOI analysis of the fixation counts also confirms the difficulty of this task. Right image of Fig. 3 displays the selected AOIs, where the red one is the menu for creating a human model. Let's see a particular participant's data as an example: the time to first fixation was 2.84 s, the number of fixations was 36, and 17 visits were counted. So, the participant looked at the right part relatively early, however, they clicked on it later, after many visits and fixations.

The next problem was to change the posture of the human model. In average, the students were faster than the teachers. As we can see from the two aggregated gaze
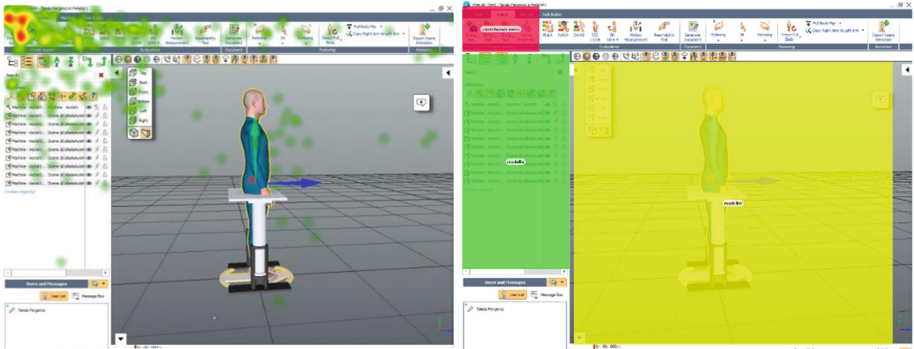
Fig. 3. Creating a human model – aggregated heatmap for all participants (left); selected AOIs (right). Red AOI: create human menu; green AOI: model tree; yellow AOI: model space. (Color figure online)

plots for teachers and for students in Fig. 4, it can be stated, that all participants searched in the human model context menu, however, the teachers first scanned the "Human Properties" tool. The real user interface elements to use are located in the main menu.

During posturing the human model, a software error was occurred. In isometric view, after opening the side toolbar, the ratio of the model space had change, and the coordinate system of the selected models didn't function until the next zoom in or zoom out operation.

Adjusting the table surface to the correct height subtask was the most time consuming. The longest time was more than 5 min. The default setting is the "Select full" interaction mode, which means that the users can shift and rotate the whole model, but they have to click on the "Select partly" button to move only one part, or they can select the part from the model tree. In this subtask, the students examined the model tree and the horizontal menu bar as well, and the teachers focused only on the menu bar.

Table 1 shows the AOI data of this subtask. According to the data, it can be stated, most of the participants did not think that the solution is in the main menu. We selected three areas to analyze: menu bar, model tree, model space. The fixation count of the model space AOI was always at least two or three times higher than the fixation count of the menu bar AOI or the model tree AOI. This result refers to the fact that the participants intended to complete the task using the mouse, for example double click or right click on the parts. Those who chosen the "Select partly" command needed help from the moderator.

The RULA risk assessment tool was easy to find, mostly because during the previous subtasks they had seen it. The determination of the angle of the neck was relatively easy subtask, but one problem has occurred. The participants firstly searched for the angle directly on the human model, but later they found the solution in the RULA tool. For two participants, the View More Options tab was not opened, making it difficult to find information about the body parts (see Fig. 5).

Fig. 4. Gaze plots of two types of solution strategy for adjusting the posture of the human body: searching for a variety of areas, primarily the Human Properties window (left, showing gaze plots of two teacher participants), and focusing on the menu bar (right, showing gaze plots of two student participants).

Table 1. Adjusting the table surface to the correct height– AOI data.

		Time to first fixation (s)			Fixation count (piece)			Visit count (piece)		
		Menu bar	Model tree	Model space	Menu bar	Model tree	Model space	Menu bar	Model tree	Model space
Student	1	33.38	35.25	0.03	41	92	221	17	21	22
	2	26.66	19.91	0	67	64	190	13	24	17
	3	49.67	30.53	0	1	24	322	1	8	15
Teacher	1	4.43	75.49	0	40	64	261	10	9	21
	2	3.69	47.94	0	60	5	206	21	3	17
	3	23.13	35.14	2.23	77	94	582	22	21	28

Deleting the models and closing the program was the final subtask. All of the participants completed this task successfully, but two of them missed the function of the "DEL" button and none of them used the more time-effective "Delete all objects" command.
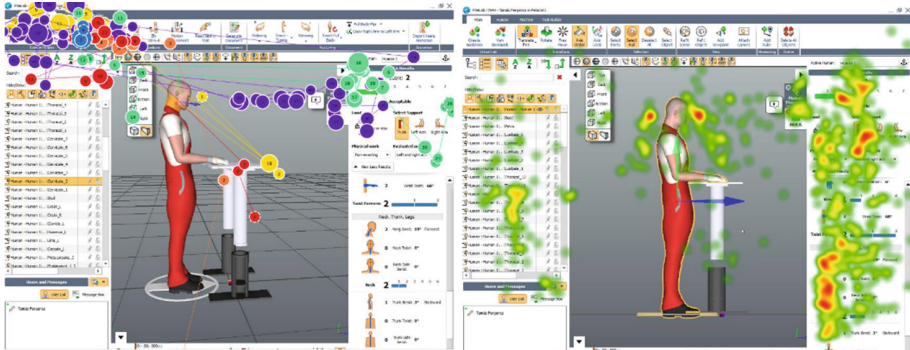
Fig. 5. Finding RULA risk assessment tool – aggregated gaze plots (left); determination of the angle of the neck – aggregated heatmap (right).

4.2 Quantitative Results from the Task Completion Time and the Success Scale

In order to quantify the efficiency completing the tasks, we applied two metrics: the task completion time and the level of the subtask's success. The level of success scale was a Likert scale from 0 to 3.

In our research, we defined the levels of the success scale as the followings:

3 – The participant could complete the task without help.
2 – The participant could complete the task with little help from the moderator or the use of the help system.
1 – The participant could complete the task with significant help from the moderator.
0 – The participant could not complete the task, intervention of the moderator was necessary.

Regarding the task completion time, the data of the first student participant during posturing the human model was inaccurate because of the previously mentioned software error. All the other data are correct. According to the task completion time and the success scale, it can be stated, the hardest subtask was to import a CAD model. The second hardest subtask was to move the table top.

The participants acquired routine already in these short sessions: it can be concluded from the task completion time and the successes of the two repeating subtasks. They completed the second similar subtask much earlier and they were more successful. The experience in other similar program was also important. Those who wasn't experienced or have not used these kind of software recently finished the tasks in average later than those who had fresh experience.

5   Discussion and Conclusion

In conclusion, the pilot research was useful for both purpose. We could draw conclusions for the software and the methodology as well. Regarding the software, many suggestions can be made, however, not all function was tested. Based on the task completion time, which was around 15 min, the complexity of the task was correct, and the continuation of this list of steps is recommended, however, it can be expanded with more tasks while the complexity of the task will not significantly change. Additional subtasks would probably take more task completion time, but would also give more results.

During the evaluation and the post processing, a significant problem had observed regarding the methodology. We could not interpret the visualizations of the eye movements in the model space in spite of the simple objects and subtasks. (See Fig. 6.) The reason for this problem is the nature of the model space. The participants can rotate the whole model and the virtual space too. In a usability test every user could rotate, shift and zoom as they like, therefore will not be one same view, which means we have to evaluate for each participant separately. Furthermore, evaluation has to be based on dynamic videos of screen- and eye-tracking recordings instead of static screenshots.
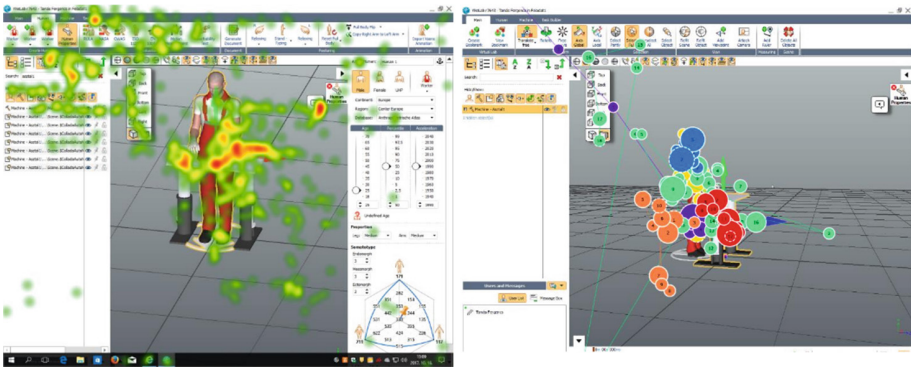
Fig. 6. A problem during the evaluation. Heatmaps and gaze plots are not interpretable in the model space because of the potential movements. Aggregated heatmap (left); aggregate gaze plots (right).

After this experience of the pilot, we will insert a new, short phase into the protocol of the planned usability tests, when the participants will be asked neither to move in model space nor to move the model.

For the further usability tests, it can be stated, the number of the involved participants (six) is correct, however, at least one unexperienced student participant is recommended, thus the correlation between the age, the experience and the task completion time can be observed. For instance, this is confirmed by the fact that the students had found the function of change the body posture of the human model faster than the teachers.

References

1. De Magistris, G., et al.: Dynamic control of DHM for ergonomic assessments. Int. J. Ind. Ergon. **43**, 170–180 (2013). https://doi.org/10.1016/j.ergon.2013.01.003
2. Santos, J., Sarriegi, J.M., Serrano, N., Torres, J.M.: Using ergonomic software in non-repetitive manufacturing processes: a case study. Int. J. Ind. Ergon. **37**, 267–275 (2007). https://doi.org/10.1016/j.ergon.2006.10.022
3. Bergman, C., Castro, P.R., Högberg, D., Hanson, L.: Implementation of suitable comfort model for posture and motion prediction in DHM supported vehicle design. Procedia Manuf. **3**, 3753–3758 (2015). https://doi.org/10.1016/j.promfg.2015.07.816
4. Sundin, A., Örtengren, R.: Digital human modeling for CAE applications. In: Handbook of Human Factors and Ergonomics, pp. 1053–1078. Wiley, Hoboken (2006). https://doi.org/10.1002/0470048204.ch39
5. Polášek, P., Bureš, M., Šimon, M.: Comparison of digital tools for ergonomics in practice. Procedia Eng. **100**, 1277–1285 (2015). https://doi.org/10.1016/j.proeng.2015.01.494
6. Rajaee, M.A., Arjmand, N., Shirazi-Adl, A., Plamondon, A., Schmidt, H.: Comparative evaluation of six quantitative lifting tools to estimate spine loads during static activities. Appl. Ergon. **48**, 22–32 (2015). https://doi.org/10.1016/j.apergo.2014.11.002
7. ViveLab Ergo. http://vivelab.cloud/
8. Blanchonette, P.: Jack Human Modelling Tool: A Review. Air Operations Division Defence Science and Technology Organisation DSTO-TR-2364, Air Operations Division DSTO Defence Science and Technology Organisation, Fishermans Bend (2010)
9. Berlin, C., Adams, C.: Production Ergonomics: Digital Human Modeling (2017). https://doi.org/10.1007/978-3-642-21799-9
10. Poole, A., Ball, L.: Eye tracking in human-computer interaction and usability research: current status and future prospects. In: Encyclopedia of Human-Computer Interaction (2006). https://doi.org/10.4018/978-1-59140-562-7
11. Romano Bergstrom, J.C., Olmsted-Hawala, E.L., Jans, M.E.: Age-related differences in eye tracking and usability performance: website usability for older adults. Int. J. Hum. Comput. Interact. **29**, 541–548 (2013). https://doi.org/10.1080/10447318.2012.728493
12. Józsa, E., Hámornik, B.P.: Find the difference! Eye tracking study on information seeking behavior using an online game. J. Eye Track. Vis. Cogn. Emot. **2**, 27–35 (2012)
13. Michalski, R.: Information presentation compatibility in a simple digital control panel design: eye-tracking study. Int. J. Occup. Saf. Ergon. **24**, 395–405 (2018). https://doi.org/10.1080/10803548.2017.1317469
14. Wang, J., Antonenko, P., Celepkolu, M., Jimenez, Y., Fieldman, E., Fieldman, A.: Exploring relationships between eye tracking and traditional usability testing data. Int. J. Hum. Comput. Interact. 1–12 (2018). https://doi.org/10.1080/10447318.2018.1464776
15. Tula, A.D., Kurauchi, A., Coutinho, F., Morimoto, C.: Heatmap explorer: an interactive gaze data visualization tool for the evaluation of computer interfaces. In: Proceedings of the 15th Brazilian Symposium on Human Factors in Computing Systems, pp. 24:1–24:9 (2016). https://doi.org/10.1145/3033701.3033725
16. Conradi, J., Busch, O., Alexander, T.: Optimal touch button size for the use of mobile devices while walking. Procedia Manuf. **3**, 387–394 (2015). https://doi.org/10.1016/j.promfg.2015.07.182

17. Orquin, J.L., Ashby, N.J.S., Clarke, A.D.F.: Areas of Interest as a signal detection problem in behavioral eye-tracking research. J. Behav. Decis. Making (2016). https://doi.org/10.1002/bdm.1867
18. Räihä, K.-J., Aula, A., Majaranta, P., Rantala, H., Koivunen, K.: Static visualization of temporal eye-tracking data. In: Costabile, M.F., Paternò, F. (eds.) INTERACT 2005. LNCS, vol. 3585, pp. 946–949. Springer, Heidelberg (2005). https://doi.org/10.1007/11555261_76