Impact of Incidental Visualizations in Primary Tasks

João Moreira
Visualization and Multimodal Interfaces Group
INESC-ID
Instituto Superior Técnico, Universidade de Lisboa
Rua Alves Redol, 9, 1000-029, Lisboa, Portugal
joaopedro.moreira@outlook.pt

Daniel Gonçalves
Visualization and Multimodal Interfaces Group
INESC-ID
Instituto Superior Técnico, Universidade de Lisboa
Rua Alves Redol, 9, 1000-029, Lisboa, Portugal
daniel.goncalves@inesc-id.pt

***Abstract*—Nowadays, the amount of information available to us is increasing every day, generated from different devices and sensors all around us. However, making its use efficient and effective can be hard. Information Visualization (InfoVis) is the area that can help use that information. In most cases, solutions based on InfoVis presuppose the user's total focus. However, in our daily routine, we execute a myriad of tasks of different complexities, from turning on a switch to cooking a meal. During those tasks, a visualization could give relevant data, previously gathered to us. On those cases, it would be important that the visualization did not interfere with the primary task. For visualizations, where their understanding in not the main aim, but instead an aim that appears during another task, we give the name "Incidental Visualizations". In this paper we present a study where we tested the impact of Incidental Visualizations (IncidentalVis) on primary tasks. We studied the quality of perceived information by users, how this visualizations affect the time to realize a primary task and the cognitive load induced by their presence. We searched for effects for different combinations of task complexity of primary tasks and different idioms to understand how different combinations can be, or not, workable on this context. Our results tell us that introducing a visualization in a primary task has a significant impact on it, on the three parameters mentioned. This means that the execution of the primary task will be less efficient and less effective than before. The results also show that changing the idiom of a visualization does not change, significantly, the time a primary task takes and the cognitive load induced. Quality of perception is the only parameter where the results change, significantly, when we change the idiom.**

***Index Terms*—Information Visualization, Incidental Visualization, Casual Visualization, User Studies**

I. Introduction

Every day, we have information available to use generated by smart devices, smart apps, Internet of Things, Big Data, etc. Besides that, the amount of it is increasing, even if we do not notice it. Having such a variety of information can make it hard for us to choose and filter data to generate useful knowledge, either by the lack of systems that can process that data or by the reduced available time we can use to find out how. InfoVis can help us by having visualizations that do that job for us. It is a very attractive idea because of their capacity of allowing us a better understanding of raw data using different idioms to show us information. But, there are cases where the understanding of a visualization is not our main task. In our daily routine, we execute different primary tasks, more easily than others. During those tasks, supplying relevant information, previously gathered, in the form of a visualization could have advantages. In those cases, it would be important that the visualization did not interfere too much with the primary task. Let's imagine we are searching a book on a shelf. Next to it, we have a visualization showing us the amount of books read in the present year, tagged by category. The focus is to search for the book, but could it be possible to perceive the information from the visualization without loosing track of what we are doing? Will the visualization distract us and upset the efficiency of the task? Could these visualizations coexist with our primary tasks? The subject of our investigation shows up.

We will call these visualizations "Incidental Visualizations" (IncidentalVis). Incidental, because we perceive it by accident. We want their perception not to be the focus of a person, but instead an aim that appears incidentally during a task. We already have things in our lives that transmit information, incidentally. A clock, for example, is constantly telling what time it when we want to know and when we do not. There are cases when we look to a watch without wanting (incidentally) and in that moment it gave us an useful piece of information.

Our goal was to study the impact of Incidental Visualizations in the performance of users while they execute primary tasks. We also defined important minor goals: a user should understand with sufficient quality the information transmitted by a visualization while he executes a primary task which means, the process should be effective; the attention needed to understand a visualization should not be too much high as to not disturb the primary task; the information transmitted in a visualization should be casual so that people with less or no experience at InfoVis can understand it better.

We realized user tests with 20 users, all executed in a controlled environment. We created tasks of three different complexities, easy, medium and difficult and we chose eight idioms for tabular data. First, we analyzed the performance of users when they execute tasks with no visualization. Then, they executed the tasks with all idioms, one at a time. For each idiom: we created two questions; we noted down the total time of each task and we used the paper and pen form NASA-TLX to measure cognitive load [1]. This three parameters allowed us to study the quality of perceive information, how the total time of the task changes and the cognitive load induced by visualizations.

Our paper has the following structure: presentation of the related work that complemented our own, either by giving ideas for us to use or by giving possible future work after this one; description of our study, starting at the planning, execution and results analysis; statistic processing; discussion of results got, their impact and possible future work.

II. Related Work

D. Huang executed a study where the result was a table with different design dimensions when are creating visualizations for personal information [2]. *Actionability* is the degree at which a visualization will guide future actions (we may or may not use this directly since we did not test if an incidental perception could change behavior). *Attentional* Demand is the amount of attention a visualization demands from a person (here we have a more direct relation since we said previously that we want a visualization not to disturb too much a primary task).

One example of a system with a low level of attention required is the *ShutEye*, developed by Jared S. Bauer [3]. We use it to promote us awareness for activities that contribute for good habits of sleep. It used a perispherical view on the wallpaper of a smartphone to supply recommendations about sleep health.

W. Willett, Y. Jansen and P. Dragicevic executed a relevant study about visualization placement [4]. They present two categories of visualizations: embedded and not embedded. The visualizations not embedded are the ones where data is not next to the thing the data is about. These can be: situated and not situated. A visualization is "situated" if we see it on an external device next to the thing the data of the visualization is about, for example, a tablet. A visualization is "not situated" if we see it on an external device not next to the thing the data of the visualization is about, for example, a Desktop in a different room. If a visualization is "embedded", then we see the data directly on the thing the data is about, either by virtual reality or physical visualizations. On our study, we used situated visualizations since we wanted them to coexist with the primary tasks.

A system implemented by J. Rodgers and L. Bartram explored the supply of ambient and artistic feedback integrated on the environment [5]. The authors believe the subtle use of these techniques can reduce attention and interaction demands from visualizations. Specifically on their study, they try to make people know their impact in resource consume, like water and electricity. As we did, they wanted to reduce the cognitive load of visualizations.

III. Methods

We wanted to test the impact incidental visualizations have in different tasks. We made the validation of this aim by testing different situations where we changed characteristics of tasks and visualizations.

A. *What did we test?*

We needed to see what happened when a visualization is present during the execution of primary tasks. First, we need to know what happens when users executed them without visualizations. This gave us base values for us to use as references when comparing with the cases the visualization is present. With these values got, we then tested different tasks with growing complexity and in each one of them we tested eight different idioms.

B. *Complexity of a Task*

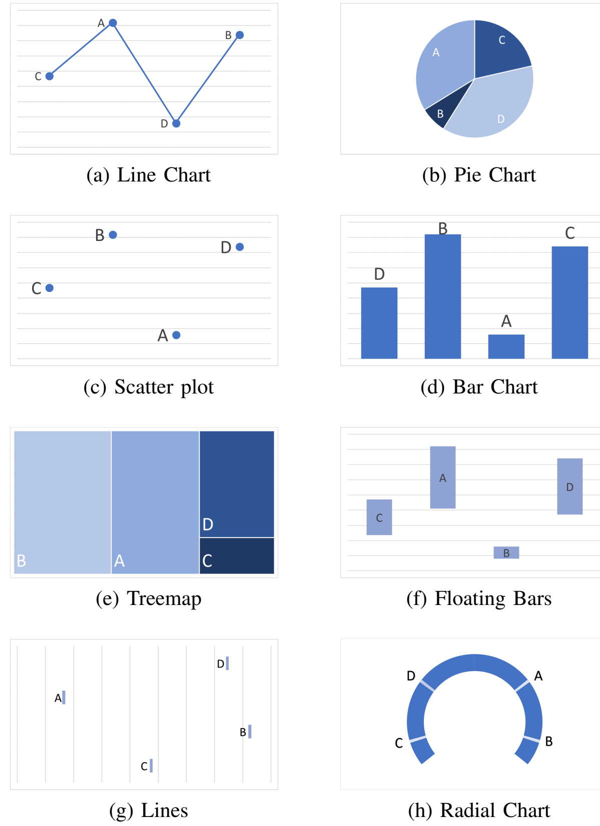
As said above, we only focused on three complexities for a task: easy, medium and hard. However, this gave us a challenge. How to put a task inside one of these categories? We searched articles that could help us in this matter, but we only found one. Donald Campbell tell us we only have four categories of complexity for all tasks [6]. The author starts by describing which attributes of complexity each task can have. It is four, which give us 16 different categories for all different attributes (4 categories x 4 attributes). Never the less, the author mentions that no combination can be strictly more or less complex than another one. Therefore, what we chose the category that fit the most our idea and then we changed the attributes so that difficulty increased. The category chosen was the "Problem Tasks". These have multiple path to get to a specific goal. The other types have over one goal, so they were not appropriate for our study. We were not interested in study tasks with over one goal. We then defined out three complexities where the complexity grew by adding attributes. Easy Complexity: tasks that have only one way for we to execute to achieve its goal. Medium Complexity: tasks that have over one way for we to execute and each leads to a goal. Hard Complexity: tasks that have over one way for we to execute, but only one leads to pretended goal.

IV. Idioms vs. Task Complexity

As previously said, we will have three different complexities of tasks. These tasks should be primary tasks, which means, they should somehow simulate real-life scenarios. However, we had trouble defining such cases for two major reasons. First, to simulate a real scenario of a primary task, like turning on a switch, we would need to implement a prototype in people's life. Unfortunately, the time given for this investigation was not enough to do this with 20 users. Second, incidental visualizations should coexist normally during a primary task. If a person is being tested, we cannot ensure she with perform as in a real case scenario. So, what we did was making our primary tasks to help us study the important aspects of the investigation (mainly the impact cause by these visualizations) instead of trying to simulate a real-life primary task. Our easy complexity task was: "flip the switch". The only thing the user needed to do was: enter the room; flip the switch and leave the room. This is an easy task because it only has one way to achieve one goal. Our medium complexity task was: "from six books, choose two and change their position and after that choose another two and change also their position". As before, the user entered the room, executed the task and then left. This is a medium task because it has different ways of doing it and each one achieves its goal. Our hard complexity task was: “from six books, choose one, then open at a random page such as the number has three digits, then from a set of pool balls (1 to 15, the white one being the 0) pick the ones with the same digits as the page and put them in the box next to them”. This is a hard task because it has different ways of doing it and only one achieves its goal.

The idioms chosen for our study were: Line Chart 1a; Pie Chart 1b; Scatter plot 1c; Bar Chart 1d; Treemap 1e; Floating Bars 1f; Lines 1g and Radial Chart 1h. We chose them for two reasons: they are common on InfoVis and they allow us to create simple data for them. To avoid uncontrolled variables in our study and to help us focus on what was important to test, we created each visualization with each idiom with the same principle. Each visualization had only three things: the idiom, four data points and one letter next to each point. The data behind each visualization was the same. We tried to followed the idea of Zachary Pousman, John Stasko and Michael Mateas by making the information transmitted more casual [7]. The letters changed between visualizations. This avoided the users to memorize an order that would be the same every time. The main goal on each visualization was to memorize the four letters (the user decided the ordering of the letters). The questions made for each visualization were always about the positioning of the letters in terms of value. For example, if the letter A corresponded to the highest value, then B, then C and then D, when I could ask “What is the value in position 1?”. Position 1 is always the highest value and so on. As we said, the tests did not simulate real-life scenarios because that would require us to implement something in the routines of 20 users, which was not workable for us.

We executed the study with 20 users. With the three complexities of a task and the eight idioms we had 24 tests per user plus the three tests without visualizations, which gives us 27 tests in total. Each user executed all 27 tests in a row. However, for each user, the order of appearance of the visualizations, for each task, was different. This prevented the users to vitiate the results. A user will always vitiate the result of its first test, because we will probably be nervous, even if he knows what will happen and what evaluation he will have. Evey test began and ended at the same physical place. There was no time limit in the realization of each test. We said the users to do the primary tasks as efficient as they could to avoid loosing too much time looking at the visualizations. We mention again, that the fact of this being a simulation, we dealt with the problem of the users knowing the visualization was part of the study. In real-life scenarios, incidental visualizations need to blend in in a way they almost “disappear” to our perception. So telling the users to focus on the task was a way to make them simulate that. The method of each test was: each user was on the physical point with their back turned to the place where the task happened; then, he received our signal to start the task, and the time started; the users executed the task and he would try to understand the visualization; he then returned to the starting point so that the time stopped; he answered the two questions and finally, he answered the questions of the NASA-TLX form.



(a) Line Chart (b) Pie Chart (c) Scatter plot (d) Bar Chart (e) Treemap (f) Floating Bars (g) Lines (h) Radial Chart

Figure 1: Idioms used in the study

A. *Results*

On this section, we will present the results of the study, every statistic processing made and our conclusions.

*1) **Idioms vs. Comprehension**:* On the easy task, the Scatter Plot had the highest percentage of right answers (97,5%) and the Pie Chart had the lowest (75%). On the medium task, the Radial Chart had the highest percentage (95%) and the Pie Chart had the lowest (77,5%). On the hard task, the Radial Chart had the highest percentage (95%) and the Pie Chart had the lowest (72,5%). We can see that we got mostly good results, showing it was possible, in every case, executing the primary task and understand most of what the visualizations transmitted. It is curious to see that the Pie Chart shows up always as the worst idiom.

Our null hypothesis was: the quality of perception is independent of the idiom of the visualization. To see if there were are significant difference between the results got, for all tasks, and since it was a categorical variable, we used the Chi-Square test. We verified that we could reject our null hypothesis with an α of 0,05, on every task, with p-values of 0,031 (easy complexity task), 0,004 (medium complexity task) and 0,014 (hard complexity task). Because every value was lower than 0,05, this means that changing the idiom of the visualizations impacts the quality of perception. However, we wanted to understand which idioms were effectively more different on their results. We applied the Bonferroni adjustment to our α, and we got the new α of 0,002. We executed the Chi-Square test to ever combinations of a pair of idioms, in every task, to see if there was a p-value lower than 0,002, but there was not. This means that, **in every task, there was an impact when we change the idiom, but we could not know exactly where**.

*2) **Idioms vs. Task Time**:* On the easy, medium and hard tasks, the test without a visualization was always the fastest (3,82s, 7,06s and 9,15s respectively). On the easy task, the test with the Treemap was the slowest (4,59s). On the medium task, the test with the Floating Bars was the slowest (9,41s). Finally, on the hard task, the test with the Line Chart was the slowest (14,93s).

Our null hypothesis was: the time to execute the task is independent of the idiom of the visualization (including not having a visualization). To see if there were any significant differences between the results got, for all the tasks, we started by using the Shapiro-Wilk test to see if these were following a normal distribution. On the tasks easy and hard, the results did, so we used the parametric test One-Way ANOVA. On the medium task, the results did not, so we used the non-parametric test Kruskal-Walis. On all tasks, we verified we could reject the null hypothesis on every task, with p-values of 0,0004 (easy complexity), 0,02 (medium complexity) and 0,00000002 (hard complexity) respectively. Our post-hoc tests revealed, for every task, significant differences when comparing the case without visualization with each of one. Instead, when we compare the idioms, in every task, between themselves, there are no significant different. This means, **introducing visualizations on primary tasks had a significant impact on the primary tasks, while changing them does not**.

*3) **Idioms vs. Cognitve Load**:* On the easy, medium and hard tasks, the test with the Pie Chart had always the highest percentage of cognitive load (26,42%, 28,50%, 33,83% respectively), and the test without the visualizations had always the lowest percentage of cognitive load (10% on all tasks). It is curious to see that the Pie Chart once again shows up always as the worst idiom.

Our null hypothesis was: the cognitive load felt during the task is independent of the idiom of the visualization (including not having a visualization). To see if there were any significant differences between the results got, for all the tasks, we started by using the Shapiro-Wilk test to see if these were following a normal distribution. On all tasks, the results did not, so we used the non-parametric test Kruskal-Walis. On all tasks, we verified we could reject the null hypothesis, with p-values of 0,00001 (easy complexity), 0,000001 (medium complexity), 00000002 (hard complexity) because they were lower than 0,05. Our post-hoc tests revealed, for every task, significant differences when comparing the case without visualization with the cases with idioms. Instead, when we compare the idioms, in every task, between themselves, there are no significant different. **This means, introducing visualizations on primary tasks had a significant impact on the cognitive load of the primary tasks, while changing them does not**.

V. Discussion and Conclusion

Even if our study did not simulate a real-life scenario, we still got good conclusions about the generic impact our incidental visualization have on users when they execute primary tasks. The quality of perception was always high. Was always above 70%, which shows these visualizations can coexist with primary tasks, by being effectively informative. Although we could not understand which idioms made the bigger differences, we still think the idiom used matters when we want to increase perception quality.

The other two parameters will have a special note here. About the time to execute each task. We verified that there is only impact when we introduce visualizations and not when we change the idiom on them. Why did this happen? We concluded that these results make sense. Since we said to the users the most important thing was to execute effectively the tasks, they probably sometimes sacrificed the quality of perception of the visualization. Finally, the cognitive load. The results were the same as the time parameter. Impact on introduction, but now on change. Why did this happen? Probably because we made our visualization simplistic. Every one was similar in color and in construction, which may have made the users feel the same cognitive load over the different tests.

Incidental Visualizations still have much to explore. In our study, we focused only in trying to realize the impact they have on primary tasks of different complexities. We reached to the conclusion that there is definitely an impact when they are coexisting with primary tasks. However, with our idioms and experimentation design we could not know which idioms could work better, at least for tabular data ones.

VI. Acknowledgements

This work was supported by national funds through Fundação para a Ciência e a Tecnologia (FCT) with reference UID/CEC/50021/2013.

References

[1] “Nasa task load index,” https://humansystems.arc.nasa.gov/groups/tlx/.

[2] D. Huang, M. Tory, B. A. Aseniero, L. Bartram, S. Bateman, S. Carpendale, A. Tang, and R. Woodbury, “Personal visualization and personal visual analytics,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 21, no. 3, pp. 420–433, March 2015.

[3] J. S. Bauer, S. Consolvo, B. Greenstein, J. Schooler, E. Wu, N. F. Watson, and J. Kientz, “Shuteye: Encouraging awareness of healthy sleep recommendations with a mobile, peripheral display,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI ’12. New York, NY, USA: ACM, 2012, pp. 1401–1410. [Online]. Available: http://doi.acm.org/10.1145/2207676.2208600

[4] W. Willett, Y. Jansen, and P. Dragicevic, “Embedded data representations,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 23, no. 1, pp. 461–470, Jan 2017.

[5] J. Rodgers and L. Bartram, “Exploring ambient and artistic visualization for residential energy use feedback,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 17, no. 12, pp. 2489–2497, Dec 2011.

[6] D. J. Campbell, “Task complexity: A review and analysis,” *The Academy of Management Review*, vol. 13, no. 1, pp. 40–52, 1988. [Online]. Available: http://www.jstor.org/stable/258353

[7] Z. Pousman, J. Stasko, and M. Mateas, “Casual information visualization: Depictions of data in everyday life,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 13, no. 6, pp. 1145–1152, Nov. 2007. [Online]. Available: http://dx.doi.org/10.1109/TVCG.2007.70541