The ITB Journal

Volume 14 | Issue 1 Article 4

2013

How data-ink maximization can motivate learners – Persuasion in data visualization

Judith Gottschalk
Aalborg University, gottschalk.judith@gmail.com

Recommended Citation

Gottschalk, Judith (2013) "How data-ink maximization can motivate learners – Persuasion in data visualization," *The ITB Journal*: Vol. 14: Iss. 1, Article 4.
doi:10.21427/D7NX6P
Available at: https://arrow.tudublin.ie/itbj/vol14/iss1/4

How data-ink maximization can motivate learners – Persuasion in data visualization

Judith Gottschalk
Aalborg University, Denmark
gottschalk.judith@gmail.com

ABSTRACT

This paper discusses both the macro- and the micro-level of persuasion in data visualizations in persuasive tools for language learning. The hypothesis of this paper is that persuasive data visualizations decrease reading time and increase reading accuracy of graph charts. Based on Tufte's (1983) data-ink maximization principle the report introduces a framework for persuasive data visualizations on the persuasive micro level which employs Few's (2013) conception of de-emphasizing non-data and emphasizing data-ink. This way persuasive data visualizations for a performance optimizing tool called Learning Journey Online (cf. Gottschalk and Winther-Nielsen 2013) is presented.

1 Introduction

"Make more from your brain" – this slogan resembles the well-known heading of the Nike running App[4] and it promises optimization of brain resources. How can one make more of one's learning and consequently of one's brain? Can these be changed by a web interface keeping track of our learning progress? And what makes such an interface persuasive? This report deals with the latter question from a specific perspective: It asks how the visualization of learning data can be persuasive. What is needed for a data visualization to be persuasive and what helps users to easily understand such a visualization?

Over the past three years six partner institutions from four countries in Europe have been working on the development of Persuasive Learning Objects and Technologies (PLOT) in a Lifelong Learning Project which is called EuroPLOT (www.eplot.eu). EuroPLOT has been funded by the Education, Audiovisual and Culture Executive Agency (EACEA) of the European Commission from 2010 – 2013. Within the project PLOTLearner has been developed as new learning technology (cf. Winther-Nielsen ms, 1). One part of PLOTLearner is Learning Journey Online which attempts to measure learning progress in students by using IRT within the persuasive architecture of PLOTLearner. The guiding hypothesis which Learning Journey Online is based on is that gathering data from learning statistics through computational surveillance and data mining algorithms and their visualization can support persuasive teaching in systems for computer-assisted language learning.

Based on this hypothesis I present a framework for persuasive data visualizations in a computational surveillance tool that can predict the learning progress of a learner and enables teachers to keep track of their students' learning progress. The term surveillance is essentially defined by Fogg as "observe others' behavior" (2003: 46), but the concept has been

[4] http://nikeplus.nike.com/plus/products/gps_app/

developed further in Gottschalk and Winther-Nielsen (2013) as a persuasive principle which enables a feedback-loop to apply within CALL systems.

Currently PLOTLearner is being repurposed as an online application developed by Claus Tøndering under the name of Bible Learner Online (http://bibleol.3bmoodle.dk/; Winther-Nielsen 2013a: 22f). This online application will automatically upload data on learner performance to Learning Journey Online at regular intervals. In Learning Journey Online these data can then be used, stored in log files, as input for modeling learning statistics. Teachers, students, and peers can then use these continuously-uploaded and complete learning statistics to plot an individual's learning progress in a class and enhance the learning process on the basis of complete statistical evidence. With the new system I follow the ultimate goal of emulating the presence of an artificial tutor as a learning supervisor in an interactive, virtual environment. The plan is to expand Learning Journey Online to a full-fledged intelligent tutoring system in language learning; the architecture of such as system is described in Gottschalk (2012).

The present report and study are rooted in the experimentation of a course taught for the EuroPLOT project by Nicolai Winther-Nielsen at Fjellhaug International University College Denmark in Copenhagen and is based on a paper written by Gottschalk and Winther-Nielsen (2013) which has introduced Learning Journey Online at the International Workshop on EuroPLOT Persuasive Technology for Learning, Education and Teaching 2013 [IWEPLET 2013] which was held in Paphos, Cyprus, from 16.-17 September 2013. Learning Journey Online (http://statdb.3bmoodle.dk)[5] grew out of a study by Gottschalk (2013) which analyses the importance of feedback in language learning. The study gives evidence that students desire feedback and support in their learning progress. Consequently Learning Journey Online provides the student, on the one hand, with feedback on his or her learning and, on the other hand, still opens he field for self-directed learning. In this report I complement the findings in Gottschalk and Winther-Nielsen (2013) with details of how persuasive data visualization within Learning Journey Online can look like to develop a full-fledged performance optimization system which gives teachers and learners the possibility to keep track of the students' learning progress.

The remainder of this paper is organized as follows: In section 2 I give an introduction to Tufte's (1983) and Few's (2013) world of data visualizations. Section 3 reviews of a number of empirical studies on Tufte's data-ink ratio on which Few's work as well as my approach are based. I discuss in detail how a persuasive approach for data visualization needs to look like. Section 4 presents a concrete design approach to persuasive data visualization in Learning Journey Online which is in use in section 5.

[5] The system of Learning Journey Online can be accessed by using a guest account (login name guest, password start123) until February 2014. In case one desires to log onto the system later is asked to send a request to gottschalk.judith@gmail.com

2 The world of Tufte and Few: An introduction

Tufte (1983) has developed a framework for data visualization which is the result of extensive research. He has introduced a metric to measure the efficiency with which a data visualization presents its data (cf. Blasio and Bisantz 2002: p. 89): The key factors for them are the data ink ratio, chartjunk, data density and the lie factor (cf. Tufte 1983).

The concept of the data ink ratio applies when quantitative information is displayed in printed form. Some of the ink in such a visualization presents data and some presents visual content that is not data (cf. Few 2013: p. 96). Tufte defines the data-ink ratio as follows:

> "A large share of ink on a graphic should present data-information, the ink changing as the data change. Data-ink is the non-erasable core of a graphic, the non-redundant ink arranged in response to a variation in the numbers represented." (Tufte 1983: p. 93)

From this concept Tufte develops the design principle which he puts this way:

> "Maximize the data-ink ratio, within reason. Every bit of ink on a graphic requires a reason. And nearly always that reason should be that the ink presents new information." (Tufte 1983: p. 96)

Both Tufte (1983) and Few (2013) agree that the following list of non-data can be reduced from data visualizations: The third dimension of depth on pie charts or bar charts, Lightning effects that make pies or charts look three dimensional, Ornaments on titles, labels, backgrounds and graphs, Grid line in graphs (cf. Few 2013). Tufte uses the term chartjunk to refer to effects like grid lines, ornaments, decoration etc. For him these are non-informative visual elements which do not support the core purpose of display. Here it even coins the way in which chartjunk obscures the visual data. Tufte says: The interior decoration of graphics generates a lot of ink that does not tell the viewer anything new; it is often chartjunk (Tufte 1983: p. 107).

Within my attempt to develop a persuasive visualization in Learning Journey Online I follow the guiding principles developed by Tufte (1983) and described in Few (2013). I reduce the non-data content in the visualizations as much as possible. This is done to proceed with the enhancement of the data content with as much clarity and meaning as possible (cf. Few 2013: 87). In this my approach follows the procedure described in Few (2013):

1. Reduce the non-data pixels
a) Eliminate all unnecessary non-data pixels
b) De-emphasize and regularize the non-data pixels that remain.
2. Enhance the data pixels
a) Eliminate all unnecessary data pixels.
b) Highlight the most important data pixels that remain.

(cf. Few 2013: p. 98)

Few's conception of data enhancement which he complements with the de-emphasizing of non-data is in contrast to Tufte (1983) who completely deletes non-data ink.

3 Review of Tufte's data-ink ratio and Few's data enhancement principle

While Tufte's argumentation for the data-ink ratio is compelling he does not give empirical evidence confirming that the maximization of the data-ink ratio leads to better task performance (cf. Blasio and Bisantz 2002: p. 91). Nevertheless an extensive body of research exists on the question whether there is empirical evidence for better performance of visualizations using Tufte style. Carswell et al. (1991) have shown that three-dimensional effects on bar charts, pie charts or line charts make the accuracy with which these graphs are read decrease. Visualizations which use 3D-effects also lower the performance on tasks which involve the magnitude estimations and trend classifications of these types of graph (cf. Blasio and Bisantz 2002: p. 91).

A considerable amount of research was conducted on whether background images have an effect on preference, performance and recall of data visualization. Sorensen (1993) is one example but also Gillian and Richman (1994) and Gillian and Sorensen (2009) have investigated these effects. In Tufte (1983) and Few (2013) background images are regarded as chartjunk per se and both conclude that using background images is counterfactual to the principle of data-ink maximization. Gillan and Richman (1994) generally have shown that pictorial backgrounds increased response time and decreased accuracy. Additionally they have shown that specific kinds of pictorial backgrounds have more effect on specific types of charts (line chart vs. bar chart). The type of background also influences the performance in response time and accuracy on reading data visualizations. Gillan and Sorensen (2009) have based their work on the findings in Gillan and Richman (1994) and have investigated graphs which have no background, a pictorial background with circles or with rectangles and found that accuracy in reading the graphs was highest for the difference task when the features in the indicators and the background of a graph have differed (cf. Gillan and Sorensen 2009: p. 1096).

Another aspect discussed in Gillan and Richman (1994) is the effect of x- and y-axis in data visualizations and how they affect reading accuracy and response time of graphs. Both ticks and axis serve as redundant data-ink in Tufte (1983) because they carry redundant information as the data they provide is also contained in the positions of the numerical labels. Based on a literature review Gillan and Richman (1994) doubt this because it has been pointed out in Sanders and McCormick (1987) that tick marks might be beneficial to graph readers. In another study which Gillan and Richman cite, Whitehurst, (1982), who gives empirical evidence based on performance times for using ticks and axis, too.

While Few (2013) in his framework for dashboard design agrees with Tufte that all kind of background should be regarded as non-data pixels as he coins it, he does not agree with Tufte on the point that grid lines and axis should be regarded as chartjunk. Instead he argues to de-emphasize and regularize non-data ink that remain in a chart (cf. Few 2013: p. 102). De-emphasizing axis and grid lines is, according to Few, a process which follows the reduction of non-data and is thought to be used when lines, borders and grid support the structure, organization and legibility a visualization. Here Few is in line with Gillan and Richman's study on the effectiveness of axis in data visualization.

4 Data-ink maximization: An approach towards persuasive data visualizations

In persuasive technology computers act as persuaders to help people change their attitudes or actions (cf. Gottschalk and Winther-Nielsen 2013; Fogg 2003). When applied to data visualizations this principle must be interpreted on two levels: A persuasive macro level and a persuasive micro level. On the persuasive macro level the data visualization is put into context with the other persuasive elements of the technology it occurs with while on the micro level the persuasion itself is enriched with aspects that make the visualization persuasive. In the development of persuasive data visualizations I follow two principles which utilize Tufte's and Few's principle of data-ink ratio maximization and data-ink enhancement respectively:

(1)

a. A visualization is persuasive on the **micro** level if it follows the principle of data ink maximization to decrease reading time of a visualization and increase reading accuracy.
b. A visualization is persuasive on the **macro** level if it utilizes persuasive principles like tailoring and surveillance to the degree that it does not interfere with the principle of data ink maximization.

The persuasive principle in (1a) boils down to the concept of usability: If a data visualization is easy to read and understand it can be regarded as persuasive. This simple principle is derived from Oinas-Kukkonen and Harjumaa (2008: pp. 166) who, present seven postulates for the evaluation of persuasiveness within ICT systems. One of these is the following: "Persuasive systems should aim at being both useful and easy to use" Oinas-Kukkonen and Harjumaa (2008: p. 168). Ease-of-use includes a number of aspects like convenience, ease of access, error-freeness, high information quality, a positive user experience, attractiveness and user loyalty. To make it short: if a system is useless or difficult to use it is unlikely that it could be persuasive (cf. Oinas-Kukkonen and Harjumaa 2008: p. 168). I will exemplify on a practical basis how the principle in (1a) is used to make Learning Journey Online persuasive on the micro level.

The predominant principle of persuasion that inspired Learning Journey Online on the macro level is surveillance and it is set out by Fogg (2003). The surveillance approach chosen in Learning Journey Online is uncovered monitoring. Gottschalk (2012) exemplifies that feedback is what most users of a software for CALL like PLOTLearner desire. Surveillance is a way to provide learners with feedback via detailed learning statistics. This is possible because it supports teachers to give their students such feedback and also gives learners the possibility to keep track of their learning progress (cf. Gottschalk and Winther-Nielsen 2013). Learning Journey Online also uses the persuasive principle of tailoring. The detailed feedback provided by learning statistics makes it possible for the learners to adjust their learning process according to their statistics. By choice of the student the teacher receives a detailed learning profile of the student so that he can adjust his teaching according to the statistics and needs of the student (cf. Gottschalk and Winther-Nielsen 2013).

Winther-Nielsen (ms a: 7) and Gottschalk and Winther-Nielsen (2013) describe persuasion on the macro level in PLOTLearner as being based on three levels in the development of ability and motivation through computer-assisted language learning. These levels are reduction, tunneling and tailoring. The most persuasive mechanism to increase the ability of a learner is tailoring. In such a persuasive technology the training is adjusted to the learners' knowledge level, age, learning style, progression, goals and other highly individual parameters which are related to vocational needs (cf. Gottschalk and Winther-Nielsen 2013). The persuasive architecture which Winther-Nielsen follows for the development of PLOTLearner and which is also employed within Learning Journey Online is exemplified in figure 1 below

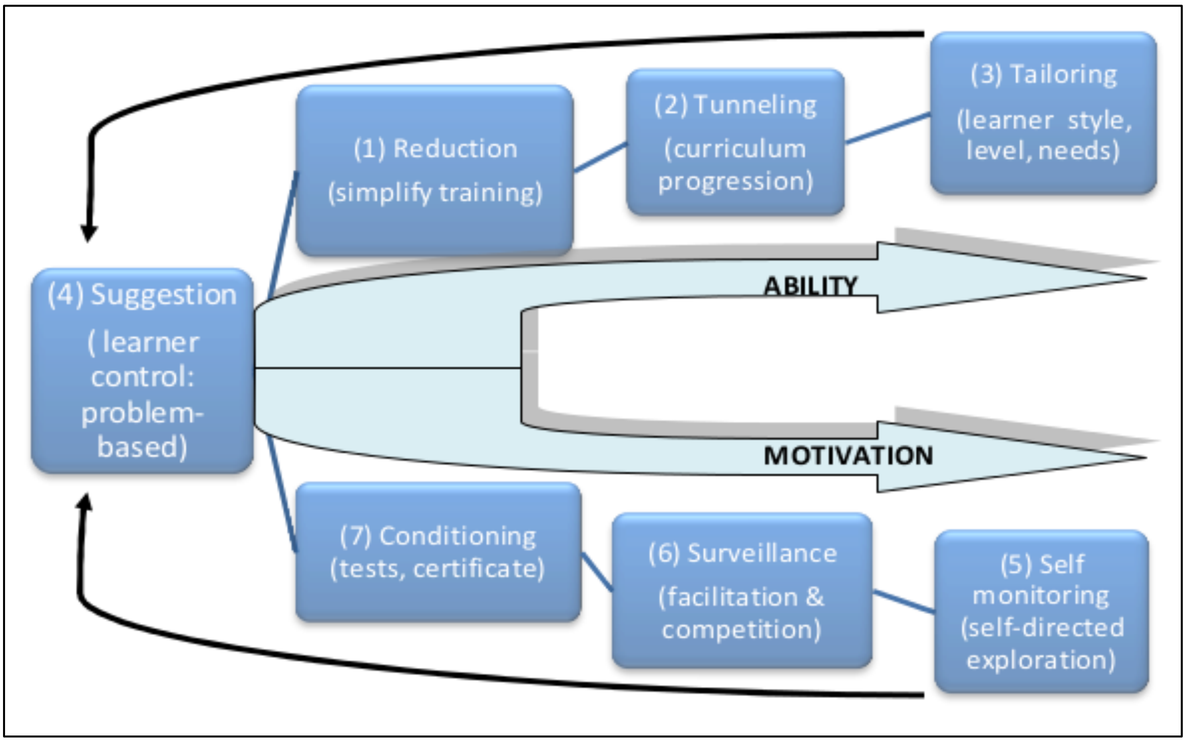

Figure 1: Persuasive architecture (Winther-Nielsen ms a:)

The persuasive data visualizations in Learning Journey Online support tailoring by giving detailed feedback based on learning statistics will make it possible to adjust the whole learning process according to the statistics results and the specific needs of the student. The other branch of Winther-Nielsen's architecture is *motivation* with similar functions and set up in a parallel track to focus on increasingly persuasive feedback. The data visualizations in Learning Journey Online are persuasive because the enable the learner to self-monitor their learning progress and to be intrinsically motivated to changer their learning habits. The ultimate system for language learning described in Gottschalk (2012) and envisioned by Winther-Nielsen uses artificial intelligence and natural language processing to record the individual's processes and outcomes and measure performance on language learning tasks. Mainly it is the addition of a further feedback element to the learning circle which makes the system persuasive on the macrolevel. In Learning Journey Online the robust theory of design for learning introduced in Laurillard (2012) has been used as pedagogical approach (cf.

Gottschalk and Winther-Nielsen 2013). It is this elements within the architecture of Learning Journey Online which makes it persuasive on the macro level. Gottschalk (2013) points out in some detail that what most users of persuasive tool for language learning like PLOTLearner and Learning Journey Online desire feedback. As has been shown in the empirical study by Gottschalk (2013) what most learners desire is to have their learning progress visualized for themselves. In the environment of PLOTLearner learners act self-directetly and therefore it is very important for them to keep track of their learning progress (cf. Gottschalk 2013). I use Laurillard's model of conversational theory to visualize the processes involved in learning from en external environment (cf. Winther-Nielsen *ms* a) and to add a feedback-loop to Learning Journey Online. A sketch of this approach is given in figure 2 below:

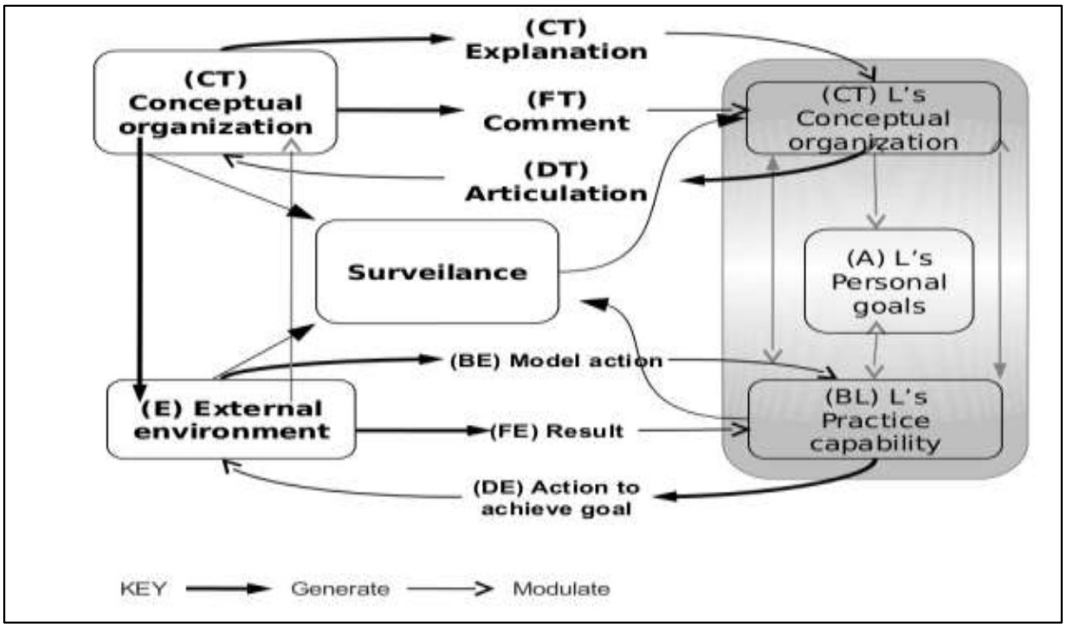

Figure 2: Surveillance added to the model of Laurillard (2012: 60)

In the approach developed by Laurillard (2012: 60) learners use their personal goals and the current organization of learning spaces for the selection of a desired practice which generates learning actions on the external environment (cf. Gottschalk and Winther-Nielsen 2013). Learners use actions modelled by the teacher or even results of their own actions to modulate and build practice and capability. It is also possible to learn through direct communication. In this case exlanations and comments from the teacher or the learning environment enable the learner in her development of a conceptual organization of the learning environment. In figure 2 this is indicated by the elements (CT) and (FT) (cf. Gottschalk and Winther-Nielsen 2013). Specifically the feedback-loop, which is created by adding the learning circle to Learning Journey Online, creates persuasive data visualization on the macro level. What gets obvious at this stage is that the persuasiveness of visualization on the macro level is a means to transport information and, in this way, to contribute to the general persuasive architecture of an ICT system.

5 Persuasive data visualizations in action

Learning Journey Online uses a number of different parameters to measure the learning performance of a student: The parameters ability, logarithmic likelihood and difficulty have their origin in the formal framework of Item Response Theory [IRT] (cf. Gottschalk and Winther-Nielsen 2013). Especially the parameter proficiency, a persuasive label I am using for the concept of automatization, is rooted within psycholinguistic research on second and foreign language acquisition (cf. DeKeyser and Criado 2012). Gottschalk and Winther-Nielsen (2013) describe in length how ability, logarithmic likelihood and difficulty are calculated within IRT statistics. The bullet chart in figure 2 shows how many right answers a student gave for every completed section compared to the overall sum of all answers. The line graph visualizes how proficiency, has developed over time. (cf. figure 3).

Since learners can solve a great number of exercises to the different grammatical aspects of Biblical Hebrew a visualization is needed which allows several of graphs in conjunction. A bullet chart is specifically designed to serve this need (cf. Few 2013: p.151). While especially Tufte (1983) mentions that color of graphs should be considered as chartjunk Few (2013) explains that hues other than gray can be used in visualizations if they do not disturb the reader. In my visualization of the bullet chart I follow this advice.

The main reason for this is persuasion on the micro level: The use of the color blue should give the user a positive impression without disturbing her too much from the data. Contrary to Tufte's approach to the maximization of the data ink ration the bullet chart and all other charts have at least an x-axis with ticks. This is motivated by the principle in (3a) which refers to persuasion on the micro level: All means to decrease the reading time of a graph and to decrease the reading accuracy are considered as persuasive. Following Few's conception of de-emphasizing non-data ink the axis and ticks are however in light gray to emphasize the data in the visualizations. Although the original concept of the graphs for Learning Journey Online included using an image of Mount Sinai to give learners the impression of being on a journey, I decided against using pictorial images in my data visualizations.

As pointed out in Gillan and Richman (1994) the use of pictorial images does increase the reading time in specific circumstances. While a pictorial background in some situations might support fast reading and reading accuracy (cf. Gillan and Sorensen 2009) it was difficult to decide whether a picture of Mount Sinai might support the two persuasive measures. My overall conclusion was that the picture would not support persuasion on the micro level.

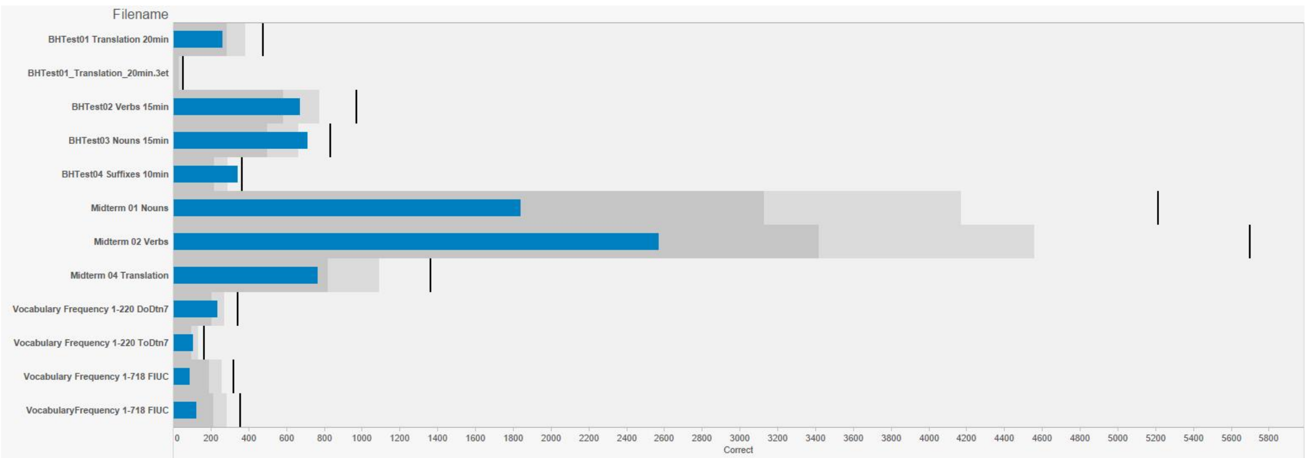

Figure 2 Persuasive Bullet Chart

To nevertheless support the impression of a 'journey' the parameters, accuracy, difficulty and right answers per minute are visualized as line diagrams. The line graph in figure 3 visualize the continuity with which learners have learned; with the peaks and valleys the line graphs display they give the impression of a journey even without the image in question. In a final version of Learning Journey Online also the statistical IRT model will be visualized as line chart. At the time of writing this was not possible, due to a lack of data. A specifically persuasive data visualization is given in figure 4. It displays the development of how many right answers a student has given within a minute.

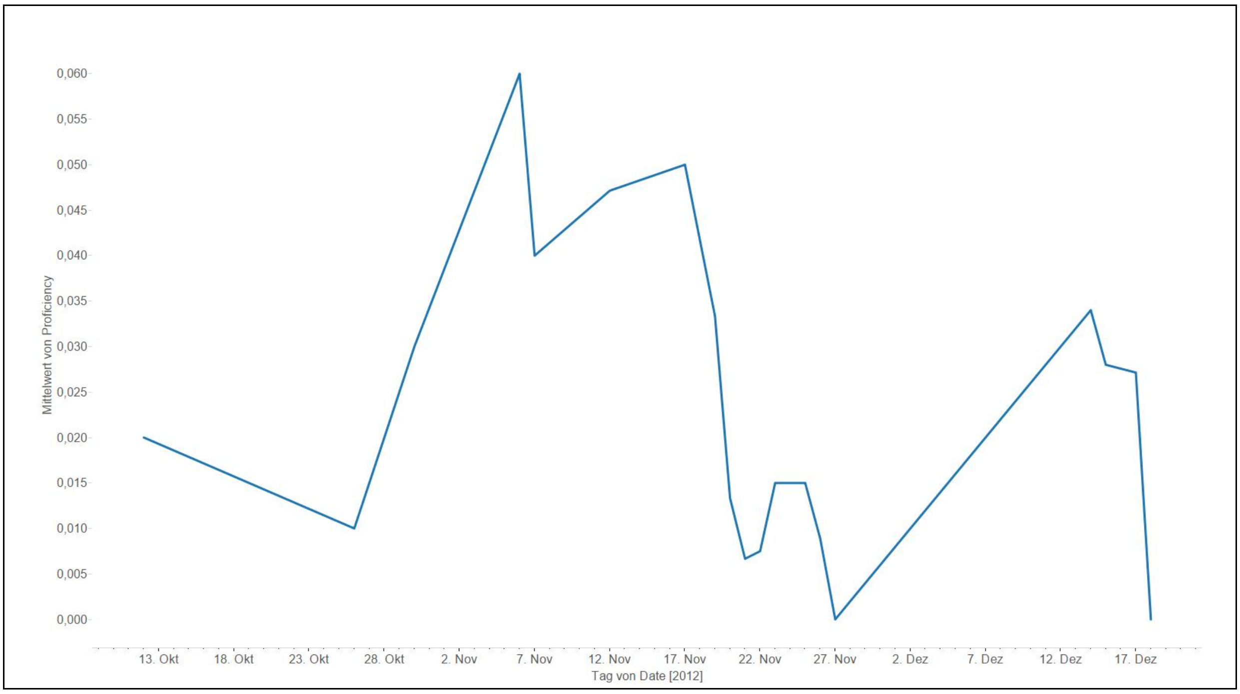

Figure 3 Persuasive Line Chart Profiency

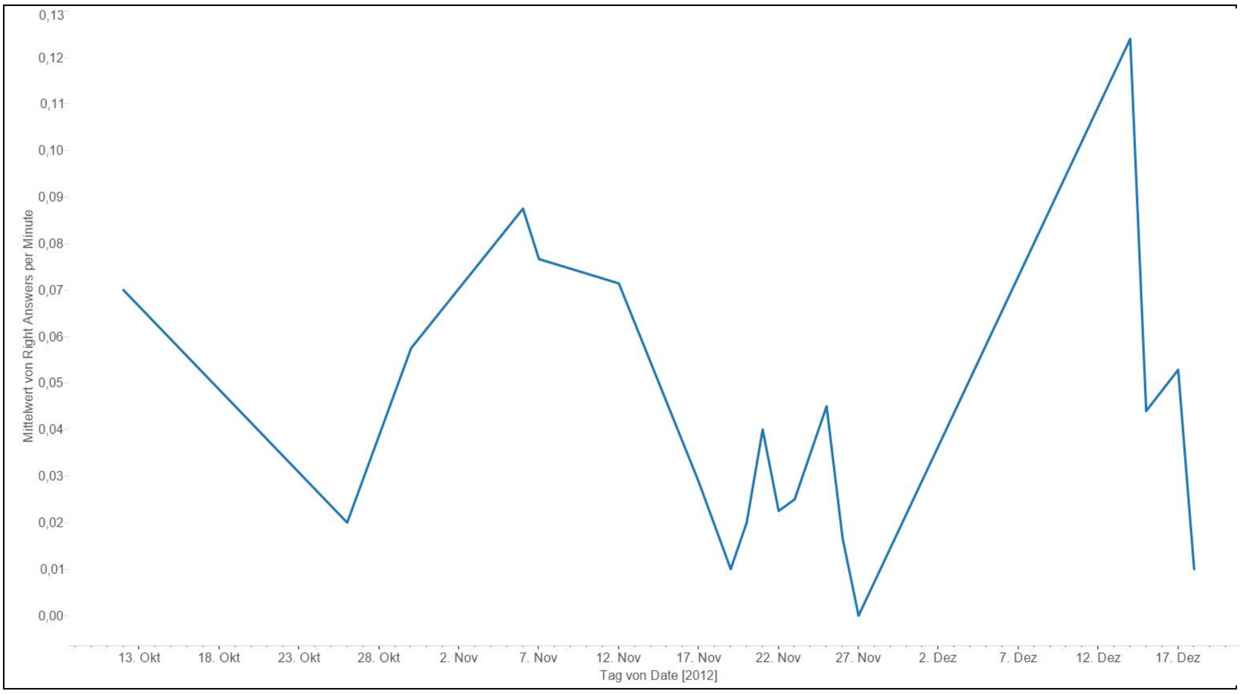

Figure 4 Line chart for right answers per minute

While in earlier approaches to data visualizations in Learning Journey Online and PLOTLearner the measure of how fast learner was able to solve a grammatical exercise, I meanwhile use the measure of right answers per minute in Learning Journey Online. The reason for this have to do with the motivational character of this measure: Although the speed with which a student has solved an exercises gives inside of how well he already mastered the acquisition of a new language (cf. DeKeyser and Criado 2012) this measurement is not persuasive when displayed on a graph. The reason is that a graph with many valleys would in this case display good learning progress while usually low values, especially when displayed in a graph are regarded as bad. Another aspect is that the visualization of processing speed would be the only one having a decreasing direction on the graph; this would make the understanding of the data visualizations specifically difficult. Therefore I have chosen the measure right answers per minute within Learning Journey online.

6 Conclusions

The core of the architecture of the new persuasive performance optimizing system, Learning Journey Online, is surveillance combined with tailoring. These two persuasive mechanisms act on the macro level of persuasion and they are complemented with thorough data visualizations which act on the persuasive micro level. In these visualization s the principle of data-ink maximization developed by Tufte (1983) and the principle of data-ink emphasizing and the de-emphasizing of non-data ink presented in Few (2013) are employed. With its persuasive data visualizations the system employs a feedback-loop persuading users to change their learning behavior.

Acknowledgements
This research was funded by the Education, Audiovisual and Culture Executive Agency (EACEA) of the European Commission through the Livelong Learning Program with grant #511633. I would like to thank Claus Tøndering for his supervision on the implementation of Learning Journey Online. Nicolai Winther-Nielsen receives my greatest gratitude for his supervision on various projects. I am thanking Morten Aagaard for discussions on persuasive data visualizations and thank Verena Fischer for indispensable language corrections. Finally Brian Nolan receives my thanks for publishing this work.

References

Behringer, R. / Sinclair, G. (2013). *IWEPLET 2013 Proceedings.* EuroPLOT Retrieved December 08, 2013 from http://www.lulu.com/shop/reinhold-behringer-and-georgina-sinclair/iweplet-2013-proceedings/paperback/product-21189131.html.

Blasio, A. / Bisantz, A. (2002). A comparison of the effects of data-ink ratio on performance with dynamic displays in a monitoring task. International Journal of Insdustrial Ergonomics 30, 89 – 101.

Carswell, C.M. / Frankenberger S. / Bernhard, D. (1991). Graphing in depth: perspectives on the use of three dimensional graphs to represent lower-dimensional data. Behaviour and Information Technology 10 (6), 459–474.

DeKeyser, R. M., / Criado-Sánchez, R. (2012). Practice in second language instruction. In C. A. Chapelle (Ed.), The Encyclopedia of Applied Linguistics. Oxford: Wiley-Blackwell.

Fogg, B.J. (2003). *Persuasive Technology: Using Computers to Change What We Think and Do*. San Francisco, CA: Morgan Kaufmann.

Gottschalk, J. (2012). The Persuasive Tutor: a BDI Teaching Agent with Role and Reference Grammar Language Interface. *ITB Journal,* 23, 31-51.

Gottschalk, J /Winther-Nielsen, N. (2013). Persuasive skill development: On computational surveillance strategies for modeling learning statistics with PLOTLearner. In R. Behringer / Sinclair, G. (Eds.), (pp. 109-116.).

Few, S. (2013). Information Dashboard Design. Displaying data for at-a-glance monitoring.Analytics Press, Burlingane, CA.

Gillan, D.J. /Richman, E.H. (1994). Minimalism and the syntax of graphs. Human Factors 36 (4), 619–644.

Gillian, D.J. / Sorensen, D. (2009). Minimalism and the Syntax of Graphs: II. Effects of Graph Backgrounds on Visual Search. Proceedings of the HUMAN FACTORS and ERGONOMICS SOCIETY 53rd ANNUAL MEETING – 2009.

Sanders, M.S. / McCormick, E.J. (1987). Human factors in engineering and design (6th ed.). New York: McGraw-Hill.

Laurillard. D. (2012). Teaching as a Design Science: Building Pedagogical Patterns for Learning and Technology. London: Routledge.

Oinas-Kukkonen, H. / Hasle, P. / Harjumaa, M. / Segerståhl, K. / Øhrstrøm, P. (Eds.). (2008). Proceedings of Persuasive Technology: Third International Conference. Oulu, Finland, June 4–6, 2008. Lecture Notes in Computer Science. Springer.

Oinas-Kukkonen, H. / Harjumaa, M. (2008). A Systematic Framework for Designing and Evaluating Persuasive Systems. In H. Oinas-Kukkonen et al. (Eds.): Proceedings of Persuasive Technology: Third International Conference : LNCS, (pp. 164 – 176). Springer.

Sorensen, D.L. (1993). The psychological processes of embellished graph reading. Unpublished master's thesis, University of Idaho, Moscow.

Tufte, E.R. (1983). The Visual Display of Quantitative Information. Cheshire, CT: Graphics Press.

Whitehurst, H.O. (1982). *The visual display of quanitative information.* Cheshire, CT: Graphics Press.

Winther-Nielsen, N. (ms a). The Corpus as Tutor: Data-driven Persuasive Language Learning. Submitted to Digital Humanities Quarterly.Winther-Nielsen, N. (ms b). PLOTLearner as Persuasive Technology: Tool, Simulation and Virtual World for Language Learning. Submitted to IWEPLET 2013 (http://www.iweplet2013.eu/)