Text Analysis Using Large High-Resolution Displays

Sven Mayer[1,4*] **Lars Lischke**[1,5,6] **Valentin Schwind**[1,7] **Markus Gärtner**[2] **Eric Hämmerle**[3]

Emine Turcan[3] **Florin Rheinwald**[3] **Gustav Murawski**[3] **Jonas Kuhn**[2] **Niels Henze**[1,6]

[1] Institute for Visualization and Interactive Systems, University of Stuttgart, Stuttgart, Germany
[2] Institute for Natural Language Processing, University of Stuttgart, Stuttgart, Germany
[3] University of Stuttgart, Stuttgart, Germany
[4] Human-Computer Interaction Institute, Carnegie Mellon University, Pittsburgh, PA, USA
[5] Vrije Universiteit Amsterdam, Amsterdam, Netherlands
[6] Eindhoven University of Technology, Eindhoven, Netherlands
[7] University of Regensburg, Regensburg, Germany

ABSTRACT

Large high-resolution displays (LHRDs) are entering into our daily life. Today, we already see them in installations where they display tailored applications, e.g. in exhibitions. However, while heavily studied under lab conditions, real-world applications for personal use, which utilize the extended screen space are rarely available. Thus, today's studies of LHRD are particularly designed to embrace the large screen space. In contrast, in this paper, we investigate a real-world application designed for researchers working on large text corpora to support them in deep text understanding. We conducted a study with 14 experts from the humanities and computational linguistics which solved a text analysis task using a standard desktop version on a 24 inch screen and an LHRD version on three 50 inch screens. Surprisingly, the smaller display condition outperformed the LHRD in terms of task completion time and error rate. While participants appreciated the overview provided by the large screen, qualitative feedback also revealed that the need for head movement and the scrolling mechanism decreased the usability of the LHRD condition.

*Author contact: info@sven-mayer.com

MuC '19, September 8–11, 2019, Hamburg, Germany

ACM ISBN 978-1-4503-7198-8/19/09…$15.00
https://doi.org/10.1145/3340764.3340768

CCS CONCEPTS

• **Computing methodologies → Natural language processing**; • **Human-centered computing → Displays and imagers**; Empirical studies in HCI; • **Applied computing** → *Document management and text processing*; *Annotation*;

KEYWORDS

Large high-resolution displays, wall-sized display, application study.

ACM Reference Format:
Sven Mayer, Valentin Schwind, Markus Gärtner, Eric Hämmerle, Emine Turcan, Florin Rheinwald, Gustav Murawski, Jonas Kuhn, and Niels Henze. 2019. Text Analysis Using Large High-Resolution Displays. In *Mensch und Computer 2019 (MuC '19), September 8–11, 2019, Hamburg, Germany.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3340764.3340768

1 INTRODUCTION

In natural language processing (NLP) and digital humanities (DH), but also for the general public, text understanding is a central and widely performed task. In professional settings, entity recognition, and entity linking are important subtasks in the domain of NLP. Analyzed texts vary a lot in both size as well as genre. Examples of common (literary) text resources can be found in various online repositories, such as the Project Gutenberg[1], the CLARIN Virtual Language Observatory[2] or the Digital Library of TextGrid[3], showcasing texts ranging from letter size over newspapers or articles all up to complete books. Especially text corpora with dozens of documents pose a serious challenge to the human reader when the analysis is done unassisted.

A number of software tools have been developed to support users while analyzing textual documents. Over the last years, these tools all support different aspects of professional

[1]https://www.gutenberg.org
[2]https://vlo.clarin.eu
[3]https://textgrid.de/en/digitale-bibliothek

text processing, e.g., DBPedia Spotlight [7], and YODI [11]. Recently, Gärtner et al. [10] presented an integrated open source web-based application which uses entity recognition and entity linking to help users to gain deep text understanding by providing third-party knowledge in-situ. In detail, they present additional information for the named entities which were found on Wikipedia using the Google Knowledge Graph.

Previous work indicated that displaying large pieces of information at once supports a sensemaking process. Ball et al. [4] showed that extended display space enables faster exploration of geo-located data. Lischke et al. [20] analyzed overviewing text documents on six display sizes. Their results revealed that users extract relevant information efficiently even on large display spaces with a width of up to 4 m. Similarly, Gutwin et al. [12] highlighted the importance of seeing large pieces of information for rediscovering information in large page-based documents. While research has identified benefits of extended display space for artificial tasks [8, 20] and tasks involving data with an inherent spatial meaning [18], it is still an open question if large high-resolution display (LHRD)s can help users performing complex real-world tasks such as text exploration.

In this paper, we compare a desktop setup and an LHRD setup in a deep text understanding task, a common task in NLP, and DH. Presenting logical links within the text and information beyond the text helps the reader to understand the text in depth. LHRD have the potential to present more insights into a given text at the same time, enabling a better overview and links across more words and paragraphs. Thus, this paper investigates the research question if LHRDs can support deep text understanding tasks and outperform traditional screens. To compare these two setups, we use Gärtner et al.'s [10] web-based tool for entity linking and displaying large text documents (see Figure 1). Using this tool provides the advantage of having comparable presentations for the conditions. Surprisingly, the results show that users were not able to utilize the possible advantages of the LHRD and preferred the common desktop setup. We argue that while previous work focused on facilitating gaining an overview over a given dataset, in our setup a detailed view is more critical and thus the smaller display yields better results.

2 RELATED WORK

The work presented in this paper is inspired by previous work in the areas of NLP, working with digital text documents as well as interacting with with LHRDs. In this section, we reflect on work from these three areas.

Natural language processing and Text Understanding

A number of tools in computational linguistics support users in text analysis. Here, a numerous tools are dedicated to support entity disambiguation and linking such as DBPedia Spotlight [7] and the YODI [11] module for GATE. Some tools also support users by highlighting the found entities and provide additional assistive functions or visualizations. For example, the Illinois Cross-Lingual Wikifier [29] directly links Wikipedia article to the named entities using hyperlinks. Moreover, TASTY by Arnold et al. [3] implements an as-you-type approach to interactive entity linking. It provides live outline of complementary information such as a picture or a linked article. Lastly, Gärtner et al. [10] presented their NALTool which supports users in deep text understanding. The tool support users with in-situ information about all named entities, with texts, additional images, and maps for places. Compared to similar tools, NALTool has the additional advantage that it supports a variety of screen sizes including LHRDs.

Working with Digital Text Documents

O'Hara and Sellen [23] compared reading on a paper to on-screen reading. The authors highlighted the importance of annotations and the spatial layout for reading. Alexander and Cockburn [1] argue that appropriate navigation techniques for efficient document interaction are required. Here, extended display space can provide more space for annotations while viewing the actual document and can lower the need for virtual navigation.

For faster document overview Cockburn et al. [6] proposed space-filling thumbnails of multipage documents to reduce scrolling. Based on these space-filling thumbnails, Gutwin et al. [12] showed that spatially stable overviews allow users to rediscover specific parts of large documents faster. Even if Gärtner et al.'s [10] web-based tool provides no spatially stable view, extended display space allows to show larger text parts at once and reduces the need for scrolling.

Interacting with Extended Display Space

Research identified various benefits of LHRDs. Through a lab study, Ball et al. [4] showed that participants were able to solve map-based tasks faster on larger displays than on smaller ones. The authors argued that the physical navigation is preferred over virtual navigation. Yost et al. [31] showed in line with the results presented by Ball et al. [4] that information visualizations can benefit from extended display space. Liu et al. [21] identified comparable benefits of LHRDs for abstract classification tasks. In line with these results, Andrews et al.[2] argued that extended display space enhances sensemaking even without any special tool.

In contrast to work showing benefits of LHRDs, Jakobsen and Hornbæk [17] controlled participants' locomotion and could not reveal a positive effect on physical navigation when interacting with an LHRD. Furthermore, Rädle et al. [24] showed that with an increase of displays size also

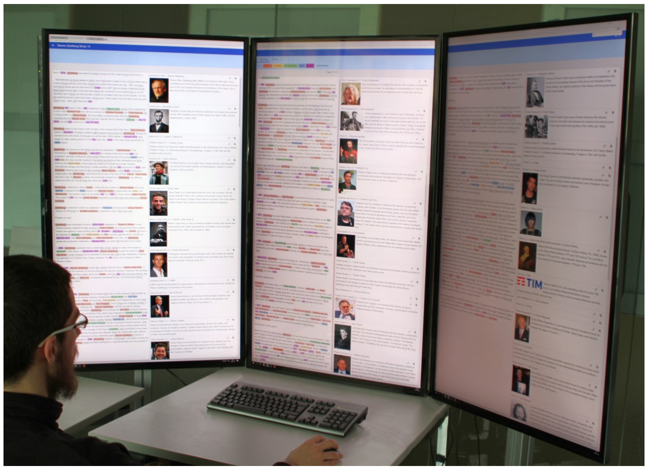


Figure 1: The study's apparatus in the *Large Display* condition solving questions on the "Steven Spielberg" text.

navigation speed increases and task load decreases. However, when the display is larger than a typical tablet, their data could not reveal a positive effect on navigation performance. Forlines et al. [9] explored the effects of group size and display configuration on visual search tasks. The authors varied the screen configuration between single screens, multiple screens, mounted horizontally or vertically. Furthermore, the authors asked single participants, pairs, and groups of four to perform the search tasks. While the authors showed that the error rate is decreasing for groups, they did not find an influence of the screen configuration. Tan et al. [28], compared a regular office display to a projected LHRD for a spatial orientation task and a text reading task. Thereby the authors kept the visual angle of the display constant. The authors could show a positive effect of the extended display space for the spatial orientation task. Moreover, instead of analyzing the benefits of LHRDs in controlled lab studies, Rajabiyazdi et al. [25] invited researchers to explore their own data on an LHRD. The results show that the LHRD enabled the researchers to gain more meaningful insights, than a regular office screen.

Summary

In summary, we show working with large text corpora is important for the computational linguistics domain. Furthermore, a number of tools support computational linguistics experts in their daily routines. Finally, previous work has identified various benefits of extended physical display space. However, these benefits seem not always transferable to complex real-world tasks. Hence, we explore the benefits of LHRDs for text reading and understanding in the computational linguistics domain.

3 STUDY

Deep text understanding tasks are common in NLP and DH. In such tasks researchers are interested in extracting insights from the text or to even understand the text in its' entirety. Here, researchers use software to visualize text annotation such as named entities but also enriching the text with additional information from external sources. Recently Gärtner et al. [10] published such a tool as open-source. To analyze the effect of the display size on a text understanding task, we conducted a lab study comparing a desktop setup with an LHRD setup (see Figures 1 and 4). To compare the two display sizes we designed a within-subject study. We used Display as an independent variable with two levels: *Small Display* and *Large Display*. We used two different texts in which participants were asked to analyze text. The order in which participants were introduced to the two differed displays was counterbalanced while we randomized the order of the presented texts.

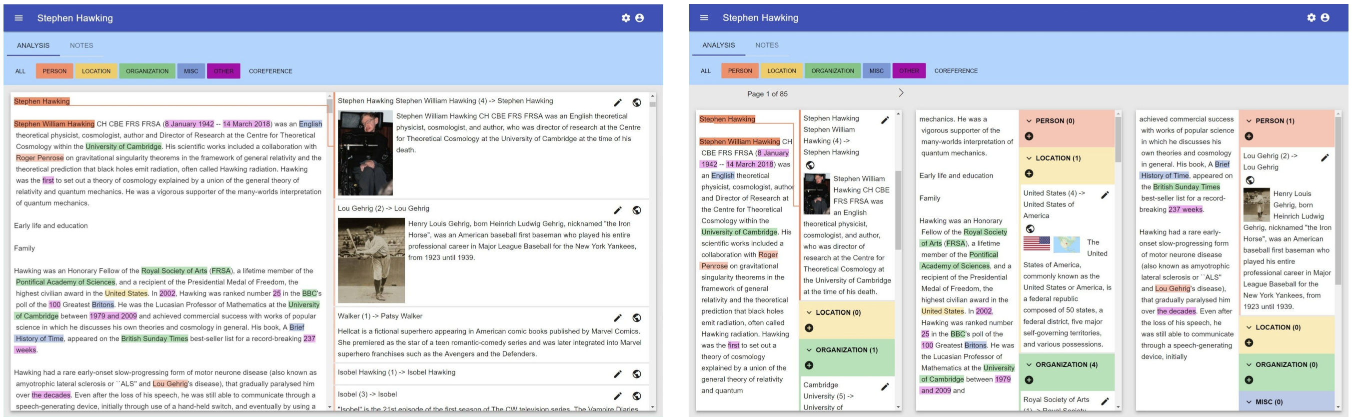
Task

Participants were asked to analyze two different texts with both around 7,000 words in total. We used the "Steven Spielberg" text from Wikipedia as the first text[4] and the Wikipedia article of "Stephen Hawking" as the second text[5]. Participants were asked to solve two tasks with each text. First, they performed a simple search task and second an aggregation task where they had to search for multiple clues in the text. The two questions for the "Steven Spielberg" text are as follows, question one "Which contending partner did Hawking recommend for one of the highest academic honors?" and question two "At which locations did Hawking live or stay for extended periods of time (not including short scientific events)?" For the "Stephen Hawking" text question one was "After founding of a separate studio, which of Spielberg's subsequent films was not released by it?" and question two "List examples for cooperation between Spielberg and George Lucas (including companies by Lucas)." In each condition, participants worked on one of the two texts.

Apparatus

In both conditions, we provided a regular office keyboard and mouse, placed on a desk for user input and asked all participants to sit on an office chair in front of this desk. In the desktop setup condition, a 28 in display with 4K (3840 × 2160 pixel) resolution, namely an ASUS PB287Q, was placed on the desk. The LHRD setup condition was designed according to findings by Lischke et al. [19]. We arranged three

[4]Wikipedia article of Steven Spielberg: https://en.wikipedia.org/w/index.php?title=Steven_Spielberg&oldid=826306514
[5]Wikipedia article of Stephen Hawking: https://en.wikipedia.org/w/index.php?title=Stephen_Hawking&oldid=830533697



(a) Small Display **(b) Large Display (scaled)**

Figure 2: (a) shows the layout for the *Small Display* condition. (b) shows a scaled version for the *Large Display* condition. In the study each of the three text and information view components were stretched out over one screen. Note: (b) the presented screenshot is not the real scale, to fit the screenshot into the paper and preserve readability we scaled the *Large Display* view down. In the study, each of the three columns displayed 1,000 words of the text.

50 in 4K screens in a bow shape screen band configuration without any gaps between the screens behind the desk. The bow setup brings the screens on the sides closer to the participant, which helps reading the text on the outer edges. The screens were arranged in a portrait mode resulting in a close to squared area, as this is more common and preferred according to results by Lischke et al. [19]. Thereby the display had a size of approximately 2×1.1 m and a resolution of 3840×6480 pixel. In both conditions, we used the NLATool by Gärtner et al. [10] to support the user. The NLATool provides entity recognition and co-reference resolution to support linguists during analysis tasks, which require deep text understanding. The graphical user interface (GUI) of the NLATool is designed with LHRD in mind and thus has built-in support for multiple screens. In the *Small Display* condition, the user is presented with one area for text view and one for the additional information view. In the *Large Display*, the user has a text and information view per screen, see Figure 2. The "text view" showed the text and its' annotations, while in the "additional information view" the user can acquire information beyond the text such as photos and maps but also a written summary with more information. This information can be used to possibly better and faster solve the text analysis task. The NLATool implements a scrolling for the single screen view and a page-turning visualization approach [22] for multiple screens. For the page-turning, the user has to click one of dedicated next and previous page which is available on each screen. By pressing the next button, the most right screen presents the next page while the content of the most right screen is then shown on the screen on the next screen to the left and that content is getting moved to the next and so on, until the content of the last screen disappears. The page-turning is happening on all screens at the same time; thus, no delay occurs.

Procedure

After welcoming the participants, we explained the procedure of the study and asked them to fill in an informed consent form. Then, we introduced them to the NLATool on the screen which they used for the first task. We let the participant try out the NLATool with the content of a Wikipedia page about "Barack Obama" until the participant was confident enough in using the tool. After that, we started the main part of the study by introducing the participant to the first text and providing the associated questions. The questions have been shown to the participant on an extra piece of paper one after the other. After the participants felt confident in answering both questions, we introduced them to the second condition by providing the second text and the respective two questions. Finally, we conducted semi-structured interviews.

Measurements and Data Collection

To compare the two conditions, we gathered and analyzed the following data:

Task completion time (TCT) [min]. The time between the text was loaded, and the participant answered all questions.

Error rate [#]. The number of wrong answers in respect to the question about the presented text. The correct answers were provided by computational linguistics experts.

Perceived Quality of the Result. After performing the task in one condition, we asked all participants to rate their performance on a Likert-Scale from 1 ("failure") to 7 ("perfect").

Perceived Work Load (raw TLX). After performing each task, we asked the participants to rate the workload with the raw NASA-TLX questionnaire [13, 14].

Interface Usability (SUS). To understand the influence of the display size on the perceived GUI quality, we asked each participant to rate the usability using the system usability scale (SUS) [5].

AttrakDiff. To understand the attractiveness when using the two display sizes, we asked participants to fill in an AttrakDiff [15].

Semi-structured interview. At the end of the study, we discussed both study conditions and the used software tool in a semi-structured interview. We discussed preferred screen sizes and screen layouts to understand users' requirements for such software tools.

Participants

To ensure that participants are trained in analyzing documents, we recruited participants from our university's linguistics department. In total, 14 experts working in humanities and computational linguistics took part in the study (8 female, 6 male). The age range was between 22 and 39 years (M = 28.1, SD = 5.6). Six of our participants used a single screen setup at their work desk, while the remaining 8 participants used two or more screens for work. Six of the participants were employed at the linguistics department as student assistants, and the other eight were Ph.D. level or higher also employed in the linguistics department. When asked to rate how much they are used to work with text on a 7-point Likert scale (7 = "strongly agree"), the average rating was 6.1 (SD = .9). Moreover, 5 participants stated that they analyzed more than 100 texts, another five more than 50, three stated more than 10, and only one participant had only analyzed more than 5 texts. Most of them did manual annotation tasks which can take days and even weeks with a single text. We reimbursed the participants with €20.

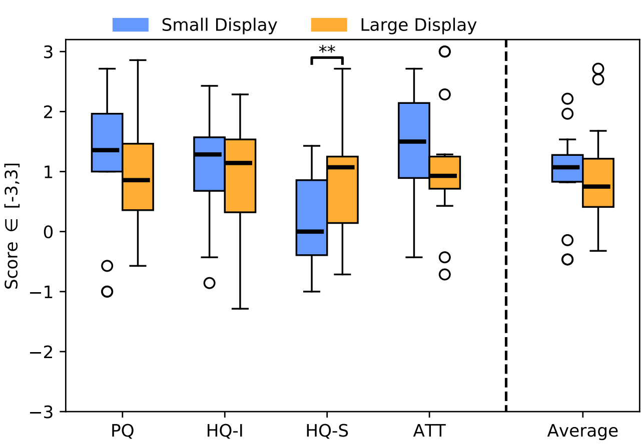


Figure 3: The AttrakDiff results of the four categories Pragmatic Quality (PQ), Hedonic Quality-Identity (HQ-I), Hedonic Quality-Simulation (HQ-S), and Attractiveness (ATT) for the three CommunicationPatterns, with ** $p < 0.01$. All scales range between -3 and 3.

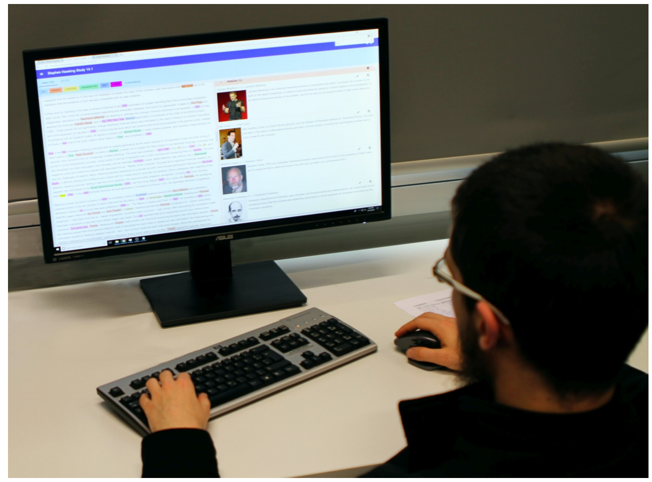


Figure 4: The study apparatus in the *Small Display* condition solving the questions on the "Stephen Hawking" text.

4 RESULTS

We used paired t-tests for parametric data and Friedman tests for non-parametric data to reveal the effects of display size.

Task completion time (TCT)

As the normality assumption was not violated ($p > .087$), we conducted a paired-sample t-test to determine whether Display significantly influenced the TCT. Our analysis revealed a statistically significant effect of Display on TCT; $t(13) = 2.187, p = .046, d = .84$. On average participants took only 19.6 *min* (SD = 5.6) to complete the tasks on the *Small Display*, while it took them 25.8 *min* (SD = 8.8) to complete the tasks on the *Large Display*.

Error rate

A Friedman test revealed a statistically significant effect of Display on error rate; $\chi^2(1) = 5.333$, $p = .021$, $W = .1$. The *Large Display* was with an average of M = .14 (SD = .36) more prone to wrongly answered questions than the *Small Display* with M = .07 (SD = .27).

Perceived Quality of the Result

A Friedman test showed no statistically significant effect of Display on perceived quality; $\chi^2(1) = .143$, $p = .706$, $W = .21$. The score of the *Large Display* ($M = 5.3$, $SD = 1.4$) was less than the *Small Display* ($M = 5.4$, $SD = .6$).

Perceived Work Load (raw TLX)

As the normality assumption was not violated ($p > .087$), a paired-sample t-test was conducted to determine whether Display significantly influenced the raw TLX. There was a statistically significant effect of Display on raw TLX; $t(13) = 2.256$, $p = .042$, $d = .62$. The *Large Display* shows on average a statistically higher task load index with $M = 8.04$ ($SD = 2.5$) than the *Small Display* with $M = 6.4$ ($SD = 2.8$), see Figure 5.

Interface Usability (SUS)

As the normality assumption was not violated ($p > .112$), we conducted a paired-sample t-test to determine whether Display significantly affected the SUS. There was no statistically significant effect of Display on SUS; $t(13) = -.578$, $p = .573$, $d = .17$. The *Large Display* had an average SUS score of $M = 73.9$ ($SD = 16.$) and, thus, less than the *Small Display* with $M = 76.6$ ($SD = 14.9$). Thus, both systems get an "acceptable" on acceptability scores and the letter grade for both is a "B-" [27].

AttrakDiff

As all assumptions of normality were not violated (all $p > .05$), we conducted five t-tests to determine whether Display significantly influenced the AttrakDiff scores, see Figures 3 and 6, one for each scale: Pragmatic Quality (PQ), Hedonic Quality-Identity (HQ-I), Hedonic Quality-Simulation (HQ-S), and Attractiveness (ATT). We used an additional Friedman test for the overall AttrakDiff score.

There was a statistically significant effect of Display on HQ-S; $t(13) = 2.256$, $p = .041$, $d = .75$. The *Large Display* scored on average with $M = 2.6$ ($SD = 1.1$) more errors than the *Small Display* with $M = 1.5$ ($SD = 1.$).

There was no statistically significant effect of Display on PQ, HQ-I, ATT, and the overall AttrakDiff score ($t(13) = -.544$, $p = .595$, $d = .17$; $t(13) = -.793$, $p = .442$, $d = .18$; $t(13) = -.817$, $p = .428$, $d = .22$; $t(13) = .098$, $p = .923$, $d = .02$; respectively).

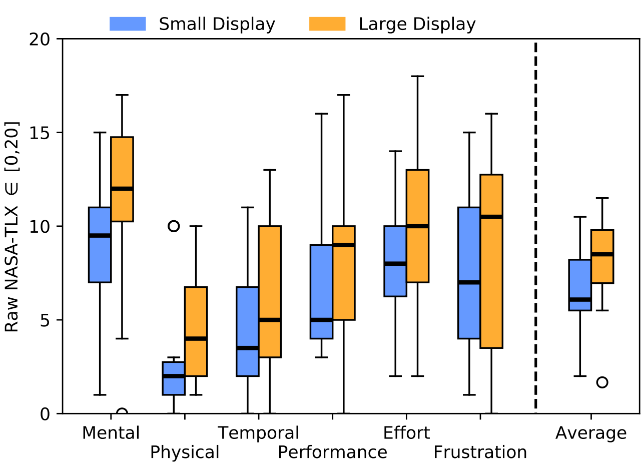


Figure 5: The raw TLX results of the six categories Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, and Frustration, as well as, the avarage.

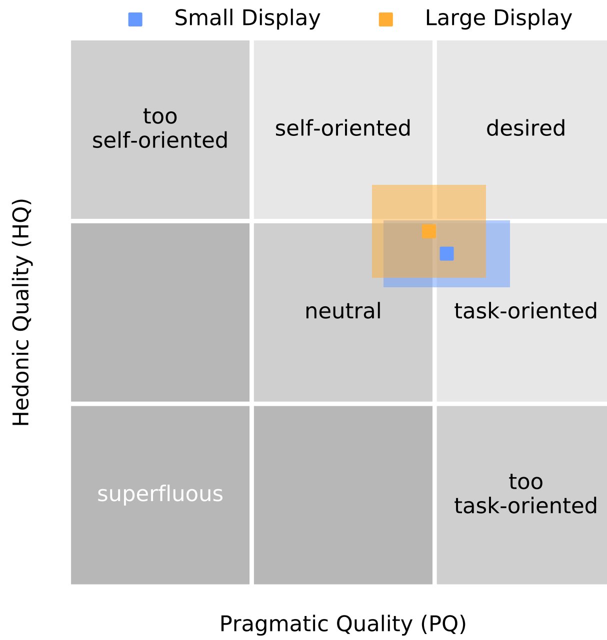


Figure 6: Portfolio presentation graph comparison of the AttrakDiff, with Hedonic Quality (HQ) = Hedonic Quality-Identity (HQ-I) + Hedonic Quality-Simulation (HQ-S).

Qualitative Feedback

Overall, we recorded on average 8 min of post study interview material per participant. In the interviews, six participants stated that they preferred working with the LHRD and seven participants favored the desktop-sized display. Participants preferring the LHRD argued that this display format provides a better overview over the text and thereby allows faster text skimming. Participants favoring the small display format mentioned that the GUI structure was clearer and that document scrolling, which caused high effort on the LHRD was more natural. Additionally, participants raised navigation issues due to the large display space. Finally, three participants mentioned ergonomic constraints, such as high physical demand when using the LHRD due to required head movements or inadequate head positions to read the text.

5 DISCUSSION

The results show that the participants achieved a higher performance on the small display condition. They were 24% faster and more accurate using the small display. This difference in performance is in contrast to the feedback from the participants. They did not perceive any difference in their personal performance. Also, the perception of having a better overview of the text on the large display is not reflected in the quantitative results. However, participants feedback regarding the GUI and ergonomic issues when using the large display explains the lower performance partly when using the large display. Surprisingly, the ergonomic constraints using the large display are not reflected in the perceived workload. Furthermore, possible distortions due to the viewing angle might have influenced the performance of the participants [30].

The higher rated hedonic quality indicates novelty factor for the large display. Furthermore, it shows that the participants perceived the large display as interesting and stimulating. Thus, participants perceived the large display as potential supportive for text exploration. The NLATool [10] used in our study was designed based on semi-structured interviews with experts from computational linguistics. Thus, Gärtner et al. [10] followed the human-centered design cycle [16] to achieve a well-designed system. However, the results clearly show that more research is needed to design GUI for LHRD and ergonomic guidelines are required to design more efficient LHRD workplaces. Moreover, it is important to investigate how scrolling might have affected the results. While we choose a page-turning visualization approach [22], we also could envision a scrolling were all pages scroll at the same time. This would reduce additional searching after a page-flip occurred but can raise visual complexity. Thus, a thorough investigation of possible scrolling mechanisms is needed to understand these factors fully.

We acknowledge that all participants were not familiar with the NLATool or using LHRD. This possibly influenced the outcome, and the performance with LHRD could improve while using the system in long-term. However, domain-specific applications in which users gain detailed insights are not yet studied on LHRDs. Thus, this paper presents a first step in gaining detailed insights which need to be studied further.

In contrast, to findings by Lischke et al. [20] who found that a $3.11m$ wide display wall is beneficial for a visual search task we cannot confirm these findings with our similar setup and even found smaller displays to be beneficial. While our LHRD was only $2m$ wide, Lischke et al. [20] also studied smaller screens and found them not as beneficial. However, the search task of Lischke et al. [20] and our first question was similar as a sniped of information needed to be found. Additionally, both implementation did utilize scrolling that present text is larger than with the display. However, while participants in the study reported by Lischke et al. [20] had only to compare features visually, in our task participants had to read the whole text and extract the required information. We here argue that LHRDs outperform smaller screens when getting an overview or initial skimming of new textual information. However, when focused work is needed reducing the amount of presented information to concentrate on the specific task outperforms the utilization of screen space. A high cognitive workload might cause this due to attention switches between different areas [26].

6 CONCLUSION

In this paper, we compared using a common 24 *in* desktop screen to using an LHRD with a size of $2 \times 1.1\ m$ for a complex text analysis task. In contrast to previous works, we use an application which is not tailor-made to utilize the extra screen space, however, was designed to adapt its GUI to support working on LHRDs. The tool supports computational linguistics experts in deep text understanding. Therefore in our study, we asked computational linguistics experts in the field to solve deep text understanding tasks. Our analysis revealed that the small display outperforms the LHRD in both TCT and error rate. Related work mainly used LHRDs to support the user in their tasks by presenting an overview; however, in the used text understanding tool a detailed view is more important than an overview. Thus, we argue that small displays are beneficial to gain detailed insights and on the other hand LHRDs surpass small displays in explorations tasks and tasks where an overview is important.

7 FUTURE WORK

To gather comparable quantitative results, we conducted a controlled lab study with an artificial instance of a task commonly performed CL. To fully understand the influence of the screen size, it would be helpful to conduct an in-situ study with CL experts. Rajabiyazdi et al. [25], showed that experts benefit from extended display spaces when performing real-world analysis tasks involving visual content. However, it is still unclear how this can be translated to text-based analysis. To explore this, experts should use the various screen sizes and the NLATool for their daily work over multiple days. Therefore, the tool would also need to support analyzing cross-document research questions.

Considering previous work by Ball et al. [4] and Jakobsen and Hornbæk [17], future work should also insvestigate if and how body movement effects analyzing large text documents. Body movement could help users to stay engaged and active. On the other hands, ergonomic requirements have to be carefully designed to prevent users from working over a longer time in unnatural body postures.

ACKNOWLEDGMENTS

This work was financially supported by the German Research Foundation (DFG) within Cluster of Excellence in Simulation Technology (EXC 310/2) at the University of Stuttgart, through project C04 of SFB/Transregio 161, and project INF of SFB 732.

REFERENCES

[1] Jason Alexander and Andy Cockburn. 2008. An Empirical Characterisation of Electronic Document Navigation. In *Proceedings of Graphics Interface 2008 (GI '08)*. Canadian Information Processing Society, Toronto, Ont., Canada, Canada, 123–130. http://dl.acm.org/citation.cfm?id=1375714.1375736

[2] Christopher Andrews, Alex Endert, and Chris North. 2010. Space to Think: Large High-resolution Displays for Sensemaking. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. ACM, New York, NY, USA, 55–64. https://doi.org/10.1145/1753326.1753336

[3] Sebastian Arnold, Robert Dziuba, and Alexander Löser. 2016. TASTY: Interactive Entity Linking As-You-Type. In *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: System Demonstrations*. http://aclweb.org/anthology/C16-2024

[4] Robert Ball, Chris North, and Doug A. Bowman. 2007. Move to Improve: Promoting Physical Navigation to Increase User Performance with Large Displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07)*. ACM, New York, NY, USA, 191–200. https://doi.org/10.1145/1240624.1240656

[5] John Brooke. 1996. SUS-A quick and dirty usability scale. *Usability evaluation in industry* 189, 194 (1996), 4–7.

[6] Andy Cockburn, Carl Gutwin, and Jason Alexander. 2006. Faster Document Navigation with Space-filling Thumbnails. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '06)*. ACM, New York, NY, USA, 1–10. https://doi.org/10.1145/1124772.1124774

[7] Joachim Daiber, Max Jakob, Chris Hokamp, and Pablo N. Mendes. 2013. Improving Efficiency and Accuracy in Multilingual Entity Extraction. In *Proceedings of the 9th International Conference on Semantic Systems*.

[8] Vishal Doshi, Sneha Tuteja, Krishna Bharadwaj, Davide Tantillo, Thomas Marrinan, James Patton, and G. Elisabeta Marai. 2017. StickySchedule: An Interactive Multi-user Application for Conference Scheduling on Large-scale Shared Displays. In *Proceedings of the 6th ACM International Symposium on Pervasive Displays (PerDis '17)*. ACM, New York, NY, USA, Article 2, 7 pages. https://doi.org/10.1145/3078810.3078817

[9] Clifton Forlines, Chia Shen, Daniel Wigdor, and Ravin Balakrishnan. 2006. Exploring the Effects of Group Size and Display Configuration on Visual Search. In *Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work (CSCW '06)*. ACM, New York, NY, USA, 11–20. https://doi.org/10.1145/1180875.1180878

[10] Markus Gärtner, Sven Mayer, Valentin Schwind, Eric Hämmerle, Emine Turcan, Florin Rheinwald, Gustav Murawski, Lars Lischke, and Jonas Kuhn. 2018. NLATool: An Application for Enhanced Deep Text Understanding. In *Proceedings of the 27th International Conference on Computational Linguistics: System Demonstrations (COLING '18)*.

[11] Genevieve Gorrell, Johann Petrak, and Kalina Bontcheva. 2015. Using @Twitter Conventions to Improve #LOD-Based Named Entity Disambiguation. In *Proceedings of the 12th European Semantic Web Conference on The Semantic Web. Latest Advances and New Domains*. Springer-Verlag, New York, NY, USA, 16. https://doi.org/10.1007/978-3-319-18818-8_11

[12] Carl Gutwin, Andy Cockburn, and Nickolas Gough. 2017. A Field Experiment of Spatially-Stable Overviews for Document Navigation. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 5905–5916. https://doi.org/10.1145/3025453.3025905

[13] Sandra G. Hart. 2006. Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 50, 9, 904–908. https://doi.org/10.1177/154193120605000909

[14] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. *Advances in Psychology* 52 (1988), 139 – 183. https://doi.org/10.1016/S0166-4115(08)62386-9

[15] Marc Hassenzahl, Michael Burmester, and Franz Koller. 2003. AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In *Mensch & Computer 2003*. Springer, 187–196.

[16] ISO 9241-210:2010(E) 2010. *Ergonomics of human-system interaction – Part 210: Human-centred design for interactive systems*. Standard. International Organization for Standardization, Geneva, CH.

[17] Mikkel R. Jakobsen and Kasper Hornbæk. 2015. Is Moving Improving?: Some Effects of Locomotion in Wall-Display Interaction. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 4169–4178. https://doi.org/10.1145/2702123.2702312

[18] Lars Lischke, Jan Hoffmann, Robert Krüger, Patrick Bader, Paweł W. Wozniak, and Albrecht Schmidt. 2017. Towards Interaction Techniques for Social Media Data Exploration on Large High-Resolution Displays. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '17)*. ACM, New York, NY, USA, 2752–2759. https://doi.org/10.1145/3027063.3053229

[19] Lars Lischke, Sven Mayer, Katrin Wolf, Niels Henze, Harald Reiterer, and Albrecht Schmidt. 2016. Screen Arrangements and Interaction Areas for Large Display Work Places. In *Proceedings of the 5th ACM International Symposium on Pervasive Displays* (2016-01-01) *(PerDis '16)*. ACM, Oulu, Finland, 228–234. https://doi.org/10.1145/2914920.2915027

[20] Lars Lischke, Sven Mayer, Katrin Wolf, Niels Henze, Albrecht Schmidt, Svenja Leifert, and Harald Reiterer. 2015. Using Space: Effect of Display Size on Users' Search Performance. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '15)*. ACM, New York, NY, USA, 1845–1850. https://doi.org/10.1145/2702613.2732845

[21] Can Liu, Olivier Chapuis, Michel Beaudouin-Lafon, Eric Lecolinet, and Wendy E. Mackay. 2014. Effects of display size and navigation type on a classification task. *Proceedings of the 32nd annual ACM Conference on Human Factors in Computing Systems - CHI '14* (2014), 4147–4156. https://doi.org/10.1145/2556288.2557020

[22] Catherine C Marshall. 2009. Reading and writing the electronic book. *Synthesis lectures on information concepts, retrieval, and services* 1, 1 (2009), 1–185. https://doi.org/10.2200/S00215ED1V01Y200907ICR009

[23] Kenton O'Hara and Abigail Sellen. 1997. A Comparison of Reading Paper and On-line Documents. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI '97)*. ACM, New York, NY, USA, 335–342. https://doi.org/10.1145/258549.258787

[24] Roman Rädle, Hans-Christian Jetter, Jens Müller, and Harald Reiterer. 2014. Bigger is Not Always Better: Display Size, Performance, and Task Load During Peephole Map Navigation. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 4127–4136. https://doi.org/10.1145/2556288.2557071

[25] Fateme Rajabiyazdi, Jagoda Walny, Carrie Mah, John Brosz, and Sheelagh Carpendale. 2015. Understanding Researchers' Use of a Large,

High-Resolution Display Across Disciplines. In *Proceedings of the 2015 International Conference on Interactive Tabletops & Surfaces (ITS '15)*. ACM, New York, NY, USA, 107–116. https://doi.org/10.1145/2817721.2817735

[26] Umar Rashid, Miguel A. Nacenta, and Aaron Quigley. 2012. Factors Influencing Visual Attention Switch in Multi-display User Interfaces: A Survey. In *Proceedings of the 2012 International Symposium on Pervasive Displays (PerDis '12)*. ACM, New York, NY, USA, Article 1, 6 pages. https://doi.org/10.1145/2307798.2307799

[27] Jeff Sauro and James R Lewis. 2016. *Quantifying the user experience: Practical statistics for user research.* Morgan Kaufmann.

[28] Desney S. Tan, Darren Gergle, Peter Scupelli, and Randy Pausch. 2003. With Similar Visual Angles, Larger Displays Improve Spatial Performance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03)*. ACM, New York, NY, USA, 217–224. https://doi.org/10.1145/642611.642650

[29] Chen-Tse Tsai and Dan Roth. 2016. Illinois cross-lingual wikifier: Grounding entities in many languages to the English Wikipedia. In *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: System Demonstrations*.

[30] Daniel Wigdor, Chia Shen, Clifton Forlines, and Ravin Balakrishnan. 2007. Perception of Elementary Graphical Elements in Tabletop and Multi-surface Environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07)*. ACM, New York, NY, USA, 473–482. https://doi.org/10.1145/1240624.1240701

[31] Beth Yost, Yonca Haciahmetoglu, Chris North, and Chris North. 2007. Beyond Visual Acuity: The Perceptual Scalability of Information Visualizations for Large Displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07)*. ACM, New York, NY, USA, 101–110. https://doi.org/10.1145/1240624.1240639