



**University of  
Zurich**<sup>UZH</sup>

**Zurich Open Repository and  
Archive**

University of Zurich  
Main Library  
Strickhofstrasse 39  
CH-8057 Zurich  
[www.zora.uzh.ch](http://www.zora.uzh.ch)

---

Year: 2020

---

## Asymmetric effect of text-chart proximity on reading behavior

Yang, Chia-Kai ; Wacharamanotham, Chatchavan

**Abstract:** Statistical charts complement textual reports by visualizing overall patterns or relations in the data. However, layout algorithms may place charts far from their associated text. Such distant placement can cause reading difficulties, or worse, a misinterpretation. We conducted an eye-tracking experiment comparing reading behaviors in two proximity levels: Placing text and chart on the same page, versus placing them on two different pages. The results indicate that the proximity influences text-reading stronger than chart-reading behavior. We discuss design implications for text-chart layout algorithms and practices. This paper and all data and materials are freely available at <https://osf.io/xunt9>.

DOI: <https://doi.org/10.1145/3419249.3420184>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-190195>

Conference or Workshop Item

Accepted Version

Originally published at:

Yang, Chia-Kai; Wacharamanotham, Chatchavan (2020). Asymmetric effect of text-chart proximity on reading behavior. In: The 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society (NordiCHI '20), Tallinn, Estonia, 25 October 2020 - 29 October 2020.

DOI: <https://doi.org/10.1145/3419249.3420184>

# Asymmetric effect of text-chart proximity on reading behavior

Chia-Kai Yang  
Chat Wacharamanotham  
yang@ifi.uzh.ch  
chat@ifi.uzh.ch  
University of Zurich  
Zurich, Switzerland

## ABSTRACT

Statistical charts complement textual reports by visualizing overall patterns or relations in the data. However, layout algorithms may place charts far from their associated text. Such distant placement can cause reading difficulties, or worse, a misinterpretation. We conducted an eye-tracking experiment comparing reading behaviors in two proximity levels: Placing text and chart on the same page, versus placing them on two different pages. The results indicate that the proximity influences text-reading stronger than chart-reading behavior. We discuss design implications for text-chart layout algorithms and practices. This paper and all data and materials are freely available at <https://osf.io/xunt9>.

## CCS CONCEPTS

• **Human-centered computing** → *Empirical studies in visualization*.

## KEYWORDS

Statistical graphics, Multimodal reading, Eye tracking

### ACM Reference Format:

Chia-Kai Yang and Chat Wacharamanotham. 2020. Asymmetric effect of text-chart proximity on reading behavior. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society (NordiCHI '20)*, October 25–29, 2020, Tallinn, Estonia. ACM, New York, NY, USA, 9 pages. <https://doi.org/10.1145/3419249.3420184>

## 1 INTRODUCTION

Statistical graphics—in short, *charts*—can be used to effectively communicate quantitative information [4]. Charts are widely used in both business [6] and scientific articles. In fact, Smith et al. [50] found that the more rigor a scientific field is, the more proportion of page space is devoted to charts. Charts are complemented by expository *textual* paragraphs that allow the authors to develop their arguments by presenting their propositions in a logical sequence [33]. Well-designed text-chart constellations can enhance reading comprehension by engaging both verbal and visuospatial cognitive systems [51].

---

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

NordiCHI '20, October 25–29, 2020, Tallinn, Estonia

© 2020 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-7579-5/20/10...\$15.00

<https://doi.org/10.1145/3419249.3420184>

Previous research found that the wording of the textual captions and the use of textual annotations on charts influence the reader's interpretation [1, 31]. Another aspect that could be influential is the *proximity* between the chart and the associated paragraph of text. The proximity has been shown to influence reading strategy and comprehension in reading materials with pictures or schematic diagrams [5, 23, 28]. However, placing charts in the vicinity of the relevant text might not always be possible due to layout constraints. Additionally, some scientific publishers may use automatic layout algorithms to reformat the articles for e-reading, leaving text-chart placements out of the hands of the authors or professional graphic designers. A better understanding of the role of proximity is fundamental to develop better layout algorithms or other software tools that readers can use to facilitate reading statistical reports.

We present an eye-tracking experiment investigating the effect of text-chart proximity on reading behavior. We hypothesized that when text and chart are placed in proximity, readers would be able to exercise relevant cognitive processes more efficiently. We first motivate our work with a short survey of the status quo of the text-chart placements in CHI papers. Then, we review the related work. Afterward, we report an experiment that investigates the effect of proximity on reading. We conclude by discussing implications for manual and algorithmic layout design.

## 2 TEXT-CHART SEPARATION

Below, we use the term *pointer* to refer to a textual mention of a figure (e.g., “Figure X shows”). Guidelines, such as the Chicago Manual of Style, recommend placing charts in the vicinity of their pointers, preferably at the top or the bottom of the page [44, section 3.8]. However, the  $\LaTeX$  layout algorithm does not take into account the location of the pointers (`\ref`).  $\LaTeX$ 's greedy algorithm only uses the location of `figure` in the code and place the figure immediately or soon after it finishes producing a column of text. This process could result in a suboptimal layout that place charts on a page separated from their pointers. To fix these separations, authors or editors must manually move the `figure` environment in the document source code [42]. However, other layout constraints may still prevent charts from being located near their pointers [43]—e.g., space limitation or when multiple pointers refer to the same chart.

### 2.1 Text-chart separation in the proceedings of CHI 2018

The prevalence of text-chart separation is unclear. One way to estimate the *lower bound* prevalence of text-chart separation is to count the occurrences of charts that are placed on a different page

from their pointers. We chose the proceedings of the CHI 2018 conference [37] as a representative population for its large number of articles (655) and the diversity of research methodology—and therefore reporting styles. First, we used pdfsearch [34] to extract the pages that have a figure or a pointer. Then, an annotator visually inspected these pages from the PDF files and labeled whether a figure is, or contains a chart. Source code and data are provided in supplementary S1<sup>1</sup>.

Out of 665 articles, 338 contain at least one chart. There are 977 charts in total. 240 charts (24.56%) are placed on a separated page from their sole pointer. 156 charts (15.96%) are referred to by multiple pointers that are on at least two different pages.

In summary, at least 40.53% of charts from the CHI 2018 proceedings are separated from the pointer. Although many of these cases may be unavoidable due to multiple pointers, a sizable portion of charts could have been placed closer to their pointer for a better reading experience. This prevalence of text-chart separation motivated our work.

### 3 RELATED WORK

We draw theories from psychology and education and operationalization from eye-tracking studies.

#### 3.1 Theories from psychology and education

Reading articles containing both textual paragraphs and charts is cognitively demanding [40]. The two presentation modalities are processed differently in the working memory [48]. According to the Integrative model of Text and Picture Comprehension (ITPC), a successful comprehension requires three cognitive processes [40]:

- (1) The *selection* of the relevant pieces of information to pay attention to.
- (2) The *organization* of these pieces of information into a coherent cognitive structure in the working memory.
- (3) The *integration* of these cognitive structures with each other and with relevant prior knowledge activated from the long-term memory.

The selection process is difficult because readers can pay attention to either text or chart, but not simultaneously on both. Additionally, reading through the material in the *first pass* may not yield a thorough understanding. Therefore, people re-read the text or look back at the figure. All these subsequent episodes together are called the *second pass* [25]. The second-pass selection process is deemed to be more intentional than the first-pass [28, 39]. The organization and the integration processes are limited by the capacity of the working memory in which information decays rapidly.

#### 3.2 Operationalizations from eye-tracking studies

Eye-tracking has been used by researchers to identify the cognitive process happening while performing a task [24, 29]. Raw eye-tracking data consist of gaze points. Consecutive gaze points that occur on the same vicinity could be aggregated into a *fixation*. The jumps between fixations are called *saccades* [7, 22]. Further analysis could define Areas of Interest (AOIs) on the stimulus, and

consecutive fixations on the same AOI could be aggregated into a *dwell* [7, 22]. A saccade from one AOI to another is called a *gaze transition* [7, Figure 2].

The three cognitive processes of ITPC could be operationalized into eye-tracking metrics. The selection process is measured by the total dwell duration on AOIs [28]. Longer duration indicates that the reader pays more attention to that AOI [21, 28]. Mason et al. [39] used the second-pass dwell duration on AOIs to reflect intentional selection process. They compared reading a text with an abstract versus a concrete illustration. An effect of the illustration types on the second-pass duration was found on the text AOI but not on the illustration AOI. In other studies, a longer second-pass dwell duration also correlates with better text summarization [26] and learning outcome [3].

The organization and the integration process are constrained by the working memory. Such constraint has been studied by Bauhoff et al. [5] in a picture comparison task: They found that placing the pictures farther apart decreases gaze transitions but increases the average dwell duration on each picture. They explained that—as the distance increases—participants increasingly prefer remembering the features of each picture in their working memory to attention-switching. The negative correlation between the distance and the gaze transitions was also found in multimodal reading, such as text-illustration in newspapers [23], and text-diagram in learning material [28].

Several works suggest gaze transitions as a suitable proxy measure for the integration process. Hegarty and Just [20] showed that, when reading instructional material describing a mechanical system, study participants switched their gaze between sentences and parts of diagrams that are semantically related. These frequent gaze transitions positively correlated with successful comprehension. Similar positive correlations were reported in Mason et al.'s work [39] and in a study of text and scientific illustrations [19]. A study by Johnson and Mayer [28] compared learning a text and a diagram in two constellations. In one condition, the text was presented as a paragraph. In another, the text was annotated onto the diagram. They found that study participants made more gaze transitions in the annotation condition. However, the total dwell duration—which indicates the selection process—is the same in both conditions. These results suggest that the proximity between text and diagram may affect the *integration* process more than the *selection* process.

Previous works investigate text-chart reading in multiple aspects. Acartürk et al. [1] found that annotations on charts facilitate integration between the chart and the paragraph text. Ho et al. [21] also found a correlation between gaze transition and readers' prior domain knowledge. However, the effect of proximity presented in diagrams and illustrations has yet to be investigated for charts. Therefore, our work adds to the understanding of reading constellations of text and charts by investigating how different proximity of paragraph text and chart influences the working memory use and the cognitive processes according to the ITPC model.

<sup>1</sup>Supplementaries are available at <https://osf.io/xunt9>

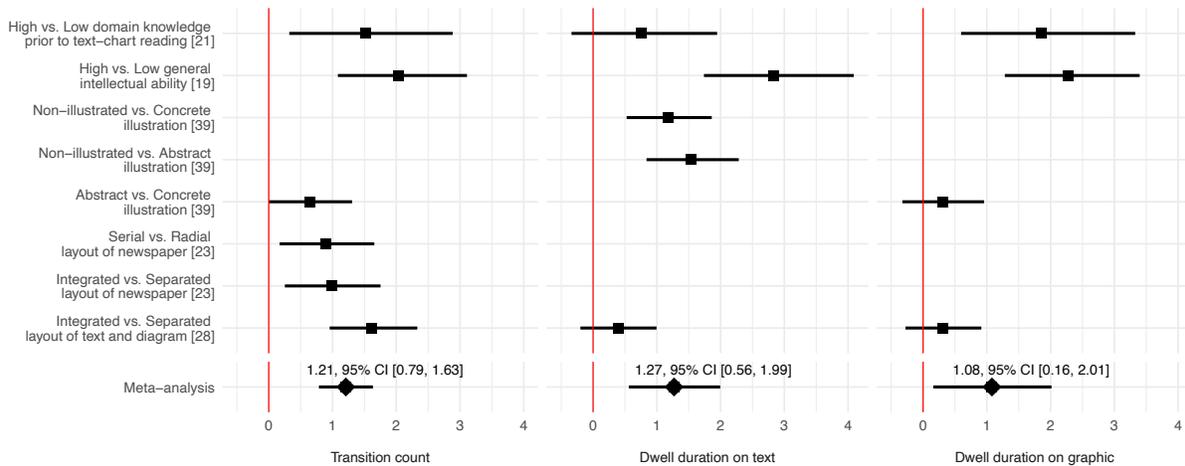


Figure 1. Meta-analysis of the effect size of various eye movement measures in text-graphic reading. The effect size is shown in Cohen’s  $d$  with 95% confidence interval.

## 4 META-ANALYSIS

To inform our *a priori* power analysis and to contextualize the effect sizes for the metrics in our experiment, we surveyed the effect sizes from relevant eye-tracking studies. We started the search process from Ho et al.’s work [21] that investigated the effects of prior knowledge in reading a paragraph of text with two charts. We then retrieved other citing/cited articles to create a corpus of literature. From this corpus, we selected the eye-tracking experimental studies that focused on reading text and graphics—including charts, diagrams, etc.—with a task that requires participants to understand the reading material.

Based on the retrieved papers, we focused on three eye movement measures: (1) gaze transitions between text and graphics, (2) dwell time on text, and (3) dwell time on graphics. We used only the studies that provided adequate information to derive the Cohen’s  $d$  effect size and its confidence interval in the text or the table. We used the ESCI software [13] to calculate the meta-analysis overall effect sizes. Details on the original measures, their statistics, and our calculation are in supplementary S2.

As shown in Fig. 1, the meta-analytic effect sizes is largest in the dwell duration on text, following by the gaze transition count, and the smallest one is the dwell duration on charts. However, the effect size on gaze transition count is the most precise among the three. We interpret the meta-analytic effect sizes using Cohen’s reference points for  $d$  [10, Chapter 2]: 0.2 = small, 0.5 = medium, and 0.8 = large. The gaze transition count and the dwell duration *on text* seem to have around medium to large effect, while the dwell duration *on graphics* could only have a small effect. In summary, the meta-analysis suggests that we should conduct *a priori* power analysis based on the dwell duration on graphics, and that we should expect a relatively larger effect on the dwell duration on text and the gaze transition count.

## 5 EXPERIMENT

To investigate the effect of text-chart proximity, we conducted an IRB-approved eye-tracking experiment<sup>2</sup>. In each experimental condition, participants read a short report that resembles the Results section in scientific articles to answer an inference question. To answer correctly, they needed information from both the *key text* paragraph and the *chart*. There were two LAYOUT conditions: In the *Near* condition, the key text and the chart were placed on the same page. In the *Far* condition, they were located on different pages.

We hypothesized that **the text-chart proximity would allow readers to intentionally direct more attention to the relevant content during the selection process, to reduce working memory use, and to increase integration**. In line with Mason et al.’s findings [39], we also hypothesized the effect to be **more prominent in texts** than in charts.

### 5.1 Eye-tracking measures

To gain insights into the *selection*, *organization*, and *integration* process, we calculated four eye-tracking measures—as discussed in subsection 3.2—on the key text and the chart AOI.

- (1) **Second-pass dwell duration on each AOI:** Re-reading text or looking-back at a figure indicates *intentional* part of the selection process. In the same manner as Mason et al. [39], the second-pass dwell duration was calculated per participant as a percentage of the total reading duration in each condition. This normalization allows comparing across participants without being confounded by individuals’ baseline reading speed.
- (2) **Total dwell duration on each AOI:** The total dwell duration captures the overall selection process, both the first- and the second-pass [28, 39]. Following Mason et al. [39], it was normalized per participant in the same manner as the

<sup>2</sup>IRB: The Human Subjects Committee of the Faculty of Economics, Business Administration and Information Technology at the University of Zurich (OEC IRB # 2018-032)

second-pass dwell duration above. The total dwell duration is a percentage of time—to the entire reading duration—that the readers pay attention to the AOI.

- (3) **Average dwell duration per AOI visit:** The average dwell duration per visit on either text or chart AOI indicates the use of the working memory during the *organization* and *integration* process. Following Bauhoff et al. [5], the average dwell duration was used without normalization. Therefore, this measure could be sensitive to outliers when summarizing the data for each participant in each condition. To address this problem, instead of the arithmetic mean, we used the geometric mean ( $\sqrt[n]{x_1 x_2 \dots x_n}$ , where  $x_i$  denotes the dwell duration for a single AOI visit) as the central tendency statistics for summarization. This practice follows the recommendations in [2, 15, 46].
- (4) **Gaze transition count:** The frequency of the integration process between text and chart is indicated by the number of gaze transitions as defined in [19, 20, 28, 39].

Our hypothesis on the selection process predicted that the *Near* condition will promote intentional attention on key text and chart. Therefore, **the *Near* condition should yield a longer second-pass dwell duration on both AOIs.** We also measured the total dwell duration on these AOIs to provide a context of the overall selection process. On the working memory, our hypothesis predicted that in the *Near* condition, the working memory usage should be reduced. Hence, **the average duration per AOI visit should be lower in the *Near* condition.** The hypothesis also predicted that both effects should be more prominent in the text than the chart. This should reflect in **interaction effects between AOI and LAYOUT.** Lastly, our hypothesis predicted a higher integration in the *Near* condition, which should yield **a higher gaze transition count in the *Near* condition.**

## 5.2 Design

We used a within-subjects experiment design to avoid the systematic influences from working memory, intelligence, and reading skills. There was one independent variable, LAYOUT: *Near* and *Far*. Half of the participants read in the *Near* condition first. To prevent learning effects, we created two versions of the report text as stimuli for each condition (section 5.4) The order of experimental conditions was counterbalanced across participants with a  $2 \times 2$  Latin Square—this requires participants in multiple of 4.

## 5.3 Participants

To determine the sample size, we conducted *a priori* power analysis with the G\*Power software [17] in the matched pairs *t*-test mode. The effect size was estimated using the mean of the meta-analytic effect size of the dwell duration on graphic—which is the smallest one among the three measures (section 4). The input parameters for the power analysis were Cohen's  $d = 1.08$ ,  $\alpha = .05$ , and power = .95. This power analysis suggested a sample size of 14 or more.

Since the eye tracker might be incompatible with some participants (typically 10 to 20% cannot be tracked reliably [27]), we over-recruited in total of 24 participants. We recruited participants from our campus using posters and online forums. To ensure an appropriate background in reading scientific reports, we pre-screened

them in an interview. According to the fixation mapping quality indicated by Tobii Pro Lab software (section 5.7), only the data from 21 participants were usable. Although the dataset was not fully counterbalanced, we tested for order effects and found them to be not statistically significant (supplementary S4). In summary, we used the data from 21 participants: 13 female and 8 male, on average 25.2 years old  $\pm$  SD = 4.13. They reported on average  $3.62 \pm 1.56$  years experience in reading scientific papers and on average  $3.50 \pm 1.84$  years with reading statistics. All had a normal or corrected vision.

## 5.4 Stimuli: The reading material

Ideally, the reading material should be a scientific paper with Introduction, Method, Results, and Discussion sections. However, in a pilot study, we used 4-page research papers as stimuli. None of the participants was willing to continue after the first condition. Therefore, we attempted to balance ecological validity and internal validity by designing the stimuli as described below. All versions of the stimuli are provided in supplementary S3.

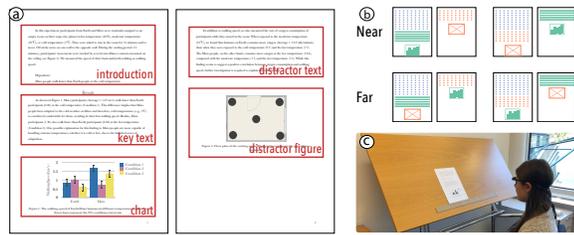
**5.4.1 Medium:** Each two-page stimulus was printed in color, double-sided, on a sheet of paper. We used print-out, instead of a screen, because previous works showed that scientists prefer paper to screen for in-depth reading [16, 36]. The physicality of paper allows participants to easily flip between pages [49]. We used double-sided printing to prevent the participants from manipulating the layout by simply laying two sheets of paper next to each other.

**5.4.2 Domain:** To avoid biases from participants' prior knowledge or personal preferences, we created reports from a fictitious replication experiment by two different labs. The experiments compared the walking speed of future Earthlings with Martians, in three weather conditions. This framing allowed for potentially different results from the two versions of the report that each participant read.

**5.4.3 Manipulation:** Each report resembled an excerpt from a research paper. It contained three paragraphs, one chart, and one diagram. The first paragraph was an introduction, which was placed to induce a reasonable text-chart separation in the report. Both the key text paragraph and the chart, that are important for the task, were placed on the same page in the *Near* condition. They were on different pages in the *Far* condition, printed on different sides of a sheet of paper, to impose a switching effort when participants read them together. The other paragraph and the diagram were distractors that described other aspects or results of the experiment that are irrelevant. These distractors were added to make all versions of the stimuli have the same AOI geometry as shown in Fig. 2a.

**5.4.4 Comprehension questions:** We designed the following questions to ensure that the participants read both stimuli in a goal-directed manner—similar to when scientists read papers [9].

- (1) identify the conclusion in the text about a particular experimental condition in the stimuli,
- (2) identify the evidence for that conclusion,
- (3) rate their confidence for that conclusion, and
- (4) make their own conclusion—which may agree or disagree with the text.



**Figure 2.** The reading material (a), overlaid with AOI names in red. Four versions (b) were created for counterbalancing. The key text is shown with solid lines. To avoid droopy eyelids, participants read on a tilted table (c).

**5.4.5 Text-chart dependency:** To successfully answer the questions above, participants needed the following pieces of information from each stimulus report:

- (1) the location (Earth or Mars),
- (2) the weather conditions (Cold, Hot, or Moderate),
- (3) the average value of the measure for each population and condition, and
- (4) their confidence intervals.

The confidence intervals were presented only in the chart. However, in the chart, the weather conditions were indicated by aliases (e.g., "Condition 1"), which required participants to look for the corresponding pointers in the text.

**5.4.6 Extraneous variables:** As shown in Fig. 2b, we counterbalanced the paragraph order (key text vs. distractor text) between participants by creating two versions of the content for each experimental condition. The order of labels and the colors of the bars were randomized to avoid memorizing.

## 5.5 Apparatus

Eye movements were collected using Tobii Pro Glasses 2: sampling rate 100 Hz, precision 5.4 mm, corresponding to 1.27 lines of printed stimuli text. Five participants who normally wear spectacles used an appropriate pair of correction lenses with the eye tracker. To reduce the likelihood of droopy eyelids, we used a tilted desk (45 degrees, Fig. 2c).

## 5.6 Procedure

**5.6.1 Pre-experiment:** We introduced an overview of the study and asked participants to sign an informed consent sheet. Then, to ensure that participants knew how to interpret the 95% confidence interval error bars, we asked them to read a 1-page refresher. Afterward, the participants confirmed their knowledge by sorting five bar charts according to their certainty of the difference in chart. These materials are provided in supplementary S3. If necessary, the experimenter explained the concepts until participants understood before continuing.

**5.6.2 Eye-tracker setup:** The participants wore the eye tracker which was then calibrated with Tobii's one-point procedure. We asked them to familiarize themselves by reading a half-page text out loud. We monitored this step and made necessary adjustments

to ensure data quality and participants' comfort. They could freely move their heads and arms in our setup.

**5.6.3 Main task:** We designed the task to invoke scientists' reading behavior. Such reading is goal-directed and strategic instead of front-to-back [9]. Before reading, participants were instructed that, after reading, they will be asked the comprehension questions (section 5.4.4). We further motivated them by offering an extra incentive if they correctly answered these questions. We suggested that they should spend 6–8 minutes but could take as much time as they needed. After reading, they answered the comprehension questions without referring back to the stimulus. Then, we re-calibrated the eye tracker before starting the second condition with the same task sequence.

**5.6.4 Post-task:** After both conditions, we conducted a retrospective think-aloud protocol [35] on the recorded eye-tracking video. Participants were asked to watch their own eye-tracking video (together with the experimenter) and say whatever came into their mind, including what they were looking at or looking for while performing the task. They were particularly encouraged to explain the reason they move fixations from one region to another. The experimenter took notes and recorded audio. Then we debriefed and gave a compensation equivalent of 35 USD (25 base + 10 extra). Everyone received the same amount regardless of their answers. The study took around one hour for each participant.

## 5.7 Data analysis

**5.7.1 Eye tracking data.** We used Tobii Pro Lab (v1.55.5126) for fixation filtering and snapshot-mapping with the Tobii I-VT (Fixation) filter (minimum duration: 60 ms). The software automatically identified and mapped fixations to a 2D snapshot of each page (1653 × 2339 pixels). We compared the mapped data with the front-camera video and manually fixed incorrect mappings and then exported the data for analysis with a script written in R [45]: supplementary S4.

Each fixation was mapped onto one of the five AOIs, as shown in Fig. 2a. To avoid spurious gaze transitions between the adjacent AOIs, there is a gap of at least 45 px (5.7 mm), corresponding to the white space on the stimuli. Fixations outside of the five AOIs were removed. We grouped consecutive fixations on the same AOI as dwells [22]. Since people are unable to derive new information from the text in less than 250 ms, and from the chart in less than 100 ms [20], dwells shorter than these thresholds were removed, and adjacent dwells on the same AOI were merged. Then we calculated the four measures on the two AOIs of our interest, the *key text* and the *chart* as described in section 5.1. When the data is skewed, we applied a log-transformation to meet the normality assumption.

The transition counts were compared using a paired *t*-test to compare two dependent samples (*Near* vs. *Far*). Unlike other comparisons, there is no text vs. chart factor because transition counts are the transitions between text and chart; thus no need for a complex model.

For the rest, we used mixed-effects ANOVA with REML. The fixed effects were AOI, LAYOUT, and AOI × LAYOUT interaction. Participants were modeled as a random slope. The model was fit with the REML method, and assessed with Type III ANOVA to

ensure that the interaction effect (if present) was accounted for. The chi-square statistics in ANOVA are calculated from the Wald chi-square test. It is the standard option for a linear mixed model with REML [38]. Computing F-statistics on such a model is only an approximation and could be prohibitively time-consuming [18, section 7.2.3]. Each ANOVA was followed-up with a planned contrast analysis which compared the *Near* and *Far* condition separately for the text and chart AOI.

Furthermore, to compare the results with the meta-analysis (section 4), we reported the contrast results in Cohen's  $d$ . The charts showed results with confidence intervals. We selectively included numerical results in the text for readability.

**5.7.2 Other data.** Audio recordings from the retrospective interview were partially transcribed and analyzed by one of the authors using the thematic analysis method with bottom-up analysis and semantic themes [8]. Answers to the comprehension questions were not evaluated because previous works suggest that the performance on such questions depends heavily on the prior knowledge [21].

## 6 RESULTS

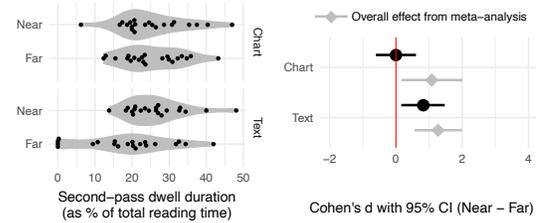
### 6.1 Effect of LAYOUT on eye movement behavior

For each test, we plot a pair of charts, as shown in Fig. 3. In each subfigure, the left chart shows the actual distribution of the data. For transparency, we plot the raw data as beeswarm and show the shape of distribution with a violin plot. The right chart shows the effect in a standardized format (Cohen's  $d$ ). This format allows comparison with the results from the meta-analysis, when available. It also facilitates future usage of our results, e.g., for planning a replication experiment. In the right chart, the confidence intervals—visualized as error bars—should be interpreted continuously as a range of plausible values of the estimation [12, p. 79]. Since they are calculated from the mixed effect model from the within-subjects data, whether they overlap or not is irrelevant for inferences [14, Rules of Eye 5, p. 177].

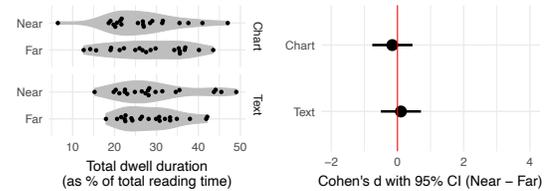
**6.1.1 Second-pass dwell duration on each AOI (Fig. 3a):** A mixed-effects ANOVA showed a significant interaction between AOI and LAYOUT,  $\chi^2(1) = 4.92, p = .02$ . According to the planned contrasts, the difference was apparent *only on the key text*. In the *Near* condition, participants made longer second-pass dwell durations than that in the *Far* condition, Cohen's  $d = 0.83$ , 95% CI [0.16, 1.47]. This effect is slightly smaller than the effect size from the meta-analysis. The effect on the chart, however, was not supported by our data—the 95% CI is centered around zero.

**6.1.2 Total dwell duration on each AOI (Fig. 3b):** A mixed-effects ANOVA showed no significant interaction between AOI and LAYOUT,  $\chi^2(1) = 0.35, p = .55$ . Similar to the above, the effect of LAYOUT was absent on the chart. However, the effect on the key text was smaller and the 95% CI was around zero.

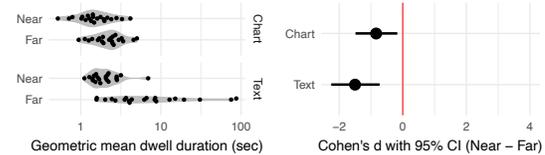
**6.1.3 Average dwell duration per AOI visit (Fig. 3c):** The right-skewed data were log-transformed to meet the normality assumption. A mixed-effects ANOVA showed a significant interaction between AOI and LAYOUT,  $\chi^2(1) = 8.90, p < .01$ . According to the planned contrasts, the average dwell duration on the key text was longer in



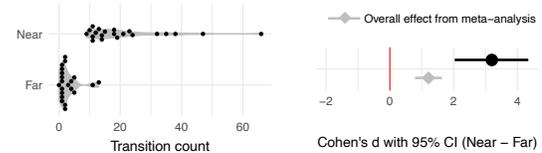
(a) The second-pass dwell duration on each AOI was longer in the *Near* condition, but only on the text-reading.



(b) The effect on total dwell duration on each AOI was very small or absent on both the key text and the chart.



(c) The average dwell duration per visit on each AOI was shorter in the *Near* condition, and the effect was more prominent on text-reading.



(d) Gaze transition count was higher in the *Near* condition.

**Figure 3. The effect of LAYOUT on eye movement measures.** For each measure, two charts are demonstrated side by side. Left: The distributions of the measure in *Near* and *Far* conditions are shown in a beeswarm plot (black dots) and a violin plot (grey shades). Right: The results of planned contrast analysis comparing the *Near* and the *Far* condition in Cohen's  $d$  with confidence intervals (black), accompanied by effect size from the meta-analysis (grey).

the *Far* condition: Cohen's  $d = 1.50$ , 95% CI [0.73, 2.25]. The effect on the chart was smaller: Cohen's  $d = 0.84$ , 95% CI [0.17, 1.49]

**6.1.4 Transition count (Fig. 3d):** A paired  $t$ -test on the log-transformed data showed that the difference was statistically significant,  $t(20) = 10.41, p < .001$ , Cohen's  $d = 3.20$ . Participants in the *Near* conditions made 6.34 times more transitions of those in the *Far* condition (as the ratio of geometric means with 95% CI [4.38, 9.19]). Our effect was higher but less precise (i.e. longer 95% CI) than the effect size from the meta-analysis.

## 6.2 Retrospective think-aloud

Based on the qualitative results from the thematic analysis, the following four themes represent participants' perceptions of and experiences of the reading in both conditions.

**6.2.1 Making incorrect connections:** In our stimuli, there were two paragraphs of text discussing the results of the experiment. One of them is a distractor text, which is neither relevant to the chart nor essential to the inference question. In the *Far* condition, the distractor text is on the same page as the chart. In this condition, seven participants mentioned that they were confused about how the distractor text related to the chart. Four of them attempted to connect the distractor text to the chart. For example, after seeing an attempt to make such a connection in the video, P01 said that *“the relationship between [a variable mentioned in the distractor text] and walking speed [in the chart] is not clear.”* In contrast, no participants mentioned the attempt to connect the distractor text to the chart in the *Near* condition.

A similar mistake also occurred in the Introduction paragraph. The chart in our stimuli required participants to look for the meaning of each bar color in the key text. However, six participants in the *Far* condition said that they used the order of conditions that was mentioned in the Introduction paragraph of the stimuli (by either recalling from memory or by rereading it). In the *Near* condition, only one participant mistakenly used the Introduction for the mapping.

Both incorrect connections above were mentioned by more participants in the *Far* condition. These results indicate that the text-chart distance interfered with the selection process.

**6.2.2 Mapping the bar color to key text:** In the *Near* condition, 16 participants mentioned that they checked the key text to map each bar color to its meaning; only 9 mentioned this in the *Far* condition. Without looking up the key text for the mapping, seven participants in the *Far* condition and one in the *Near* condition mentioned confusion about color mapping. Surprisingly, for the mapping in the *Near* condition, three other participants indeed invested time to directly compare numerical results (in the text) to the height of bars (in the chart). These results indicate that more integration occurred in the *Near* condition.

**6.2.3 Differences between modalities in the working memory load:** When reading the stimuli in the *Far* condition, P11 recalled the chart from the working memory: *“when I saw the numbers [in key text], I was thinking about the [chart], trying to see what they could have referred to.”* In contrast, P14 mentioned in the *Near* condition that she re-read the key text: *“I reread the [key text] again and again... I was trying to read and memorize it so I can picture the [key text] with the [chart] by not looking at [them]...”*. This quote indicates that several passes of reading is required to retain the text in working memory.

**6.2.4 Ease of reading:** Out of 21 participants, four mentioned that the stimuli in the *Near* condition were easier to read, whereas none was in favor of the *Far* condition. P15 and P26 explained that they did not like to flip pages too often when reading the stimuli in the *Far* condition.

## 7 DISCUSSION

To recap, we hypothesized that **the text-chart proximity would allow readers to intentionally direct more attention to the relevant content during the selection process, to reduce working memory use, and to increase integration**. We also hypothesized the effect to be **more prominent in texts** than in charts.

### 7.1 The selection process

The results showed that the *Near* condition yielded a higher second-pass dwell duration, which is attributed to the intentional part of the selection process. A long second-pass dwell duration could indicate either a difficulty in understanding [47] or a deeper cognitive processing [21, 23, 26], which concerns with pattern recognition and meaning extraction [11]. For our study, the difficulty is ruled out by the fact that we designed the stimuli to be at the same level of difficulties and counterbalanced them in both conditions. Participants also reported in the retrospective think-aloud that the stimuli in the *Near* condition were easier to read. Therefore, it is likely that the *Near* condition led to a deeper cognitive processing.

The increase in the second-pass dwell duration in the *Near* condition could be put into context by comparing with Johnson and Mayer's study [28]. In their studies where the stimuli were a diagram with textual annotation vs. a bare diagram separated from a text paragraph, the effect of proximity on the selection process was not statistically significant—unlike in our study. A possible explanation could be that the distance between text and chart is larger in our stimuli—being on different pages in the *Far* condition—than the distance in Johnson and Mayer's study, in which all stimuli were on the same screen.

Surprisingly, we found this effect asymmetric. It occurred on the texts, but not on the charts. We surmise that it is easier for readers to select relevant pieces of information in charts because visuospatial processing is more efficient than verbal processing. In the text modality, the lack of prominent visual cues might have hindered the *selection* process.

### 7.2 Working memory usage

From our results, the participants in the *Near* condition had a shorter average dwell duration per AOI visit. This result suggested that the *Near* layout was less demanding for the working memory, as predicted by our hypothesis. Although this difference was present in both modalities, the effect size was smaller in the chart modality. The different influence between modalities was also supported by the results from the retrospective interview. A reason could be that encoding charts in visuospatial working memory are less burdensome than encoding texts in verbal memory.

### 7.3 The integration process

Our results supported the effect of the proximity on *integration*. Placing text and its relevant chart in the vicinity of each other leads to more gaze transitions than placing them far apart. The retrospective interview also showed that in the *Near* condition, more participants connected the information in both sources, and fewer participants were confused. This proximity even encouraged participants to manually map the numerical results in the text to each bar in the chart. Although this additional mapping was not

necessary for the task in our study, the affordance for a better integration could be helpful in the more complex materials. In contrast, when the chart was on a different page, participants were more likely to recall from their memory or use the information from the wrong paragraph of text.

## 7.4 Overall

Taking all quantitative results (from eye-tracking data) and qualitative results (from retrospective interview data) into consideration, presenting a constellation of relevant text and chart on the same page facilitates readers in selecting relevant pieces of information to read, lightens their working memory load, and leads to a higher integration between the two modalities. Placing the chart on a different page from the text makes the article more difficult to read. Additionally, the reduction in the integration process could result in making decisions based on a single modality: Readers may lose either the holistic view afforded by the chart, or the logical arguments afforded by the text.

## 8 IMPLICATIONS

Placing charts far from their textual references reduces the *integration* process. This also makes text-reading more difficult by hindering the *selection* process and increasing the working memory use. However, the effect is asymmetric: the proximity benefits text-reading more than chart-reading.

### 8.1 For manuscript authors and graphic designers

Our research suggests that text and chart should be placed in proximity to each other—at least on the same page or screen. When the proximity has to be compromised (e.g., one chart with multiple pointers on different pages), authors and designers should focus on helping readers in their selection process—i.e. making it easier to find relevant pieces of information *in the text* to pay attention after the readers have looked at the chart. This can be achieved by strategically using typographical cues, such as *italics* or **bold face**, to highlight concepts that require connection to the chart. Nevertheless, the heavy use of these cues inside a paragraph could be distracting. Another alternative could be to annotate interesting features in the chart with labels—e.g., a, b, c—and use those labels in the text—e.g., “as Figure 1a shows...”. We believe that these features will be increasingly more important, especially for publications that are intended to be dynamically formatted for different media or to be laid out by an algorithmic system, e.g.,  $\LaTeX$  and the ACM Publishing System (TAPS).

### 8.2 For research

In our work, the key text and the chart in the *Far* condition are on different sides of a sheet of paper. Since some participants mentioned page-flipping to be a nuisance that changes their reading behavior, placements on the pages that are farther away (e.g., five pages away) could aggravate the effect. However, readers may adopt an alternative strategy, e.g., flipping a stack of pages together in one action, to maintain the same amount of effort. Further research is needed for this aspect.

The asymmetric effect suggests that automatic placement algorithms (e.g., in  $\LaTeX$ ) should take this asymmetric difficulty into account. Future research could include text complexity (e.g., [41]) as a factor to optimize the placement of charts relative to their textual references. The asymmetric effect also suggests that interactive systems that aid multimodal readings (e.g., [30, 32]) should focus on associating the chart back to the textual references. The challenge of such systems for real-world use would be to resolve the many-to-one references between text and chart. The increasing prevalence of low-cost, non-intrusive eye-trackers and augmented reality systems could allow such systems to take readers’ gaze history, and to provide real-time, augmented reading assistance.

## 9 LIMITATION

Although we planned our study based on a power analysis using the effect sizes from prior work, a larger study with more participants could improve the precision of the effects. Longer reading materials and more complex charts may also yield more pronounced effects. To facilitate replications, we publish study protocols, materials, raw and parsed gaze data at <https://osf.io/xunt9>.

## 10 CONCLUSION AND FUTURE WORK

Statistical charts complement the expository text in quantitative reports. However, typographical constraints may prevent them from being located in the proximity of each other. We surveyed the CHI 2018 proceedings and found that about 40% of charts are on a different page from the pointer in the text. Therefore, we conducted an eye-tracking experiment to investigate the effect of text-chart proximity on the cognitive processes. The results indicated the benefits of proximity on attention selection, working memory use, and integration. The results also indicated that the effect is asymmetric: While the *selection* cognitive process and the *working memory use* in text-reading were affected, such effect was less apparent or even absent in chart-reading.

Our study suggests several directions for future work. One could further investigate the asymmetric effect of text-chart proximity in two-column layout articles or with multiple levels of distance (similar to Bauhoff et al.’s [5]). Another direction is to extend this work to word-scale visualizations [52], which is a tighter integration of the chart within the text. Different reading media—e.g., paper, desktop screen, tablet screen—and their navigational affordances could moderate how the proximity affects reading. We hope that our study will provide a foundation for a better understanding of reading behavior that will ultimately contribute to better layout designs and interactive tools for better reading.

## REFERENCES

- [1] Cengiz Acartürk, Christopher Habel, Kursat Gagitay, and Ozge Alacam. 2008. Multimodal comprehension of language and graphics: graphs with and without annotations. *Journal of Eye Movement Research* 1, 3 (2008).
- [2] Neal Alexander. 2012. Analysis of parasite and other skewed counts. *Tropical Medicine & International Health* 17, 6 (2012), 684–693.
- [3] Nicola Ariasi and Lucia Mason. 2011. Uncovering the effect of text structure in learning from a science text: an eye-tracking study. *Instructional Science* 39, 5 (2011), 581–601.
- [4] American Psychological Association et al. 2009. *Publication manual* (6 ed.). American Psychological Association, Washington, DC.
- [5] Vera Bauhoff, Markus Huff, and Stephan Schwan. 2012. Distance matters: spatial contiguity effects as trade-off between gaze switches and memory load. *Applied Cognitive Psychology* 26, 6 (2012), 863–871.

- [6] Vivien Beattie and Michael Jones. 1992. The use and abuse of graphs in annual reports: theoretical framework and empirical study. *Accounting and Business Research* 1 (09 1992). <https://doi.org/10.1080/00014788.1992.9729446>
- [7] Tanja Blaschke, Kuno Kurzhals, Michael Raschke, Michael Burch, Daniel Weiskopf, and Thomas Ertl. 2017. Visualization of eye tracking data: A taxonomy and survey. In *Computer Graphics Forum*, Vol. 36. Wiley Online Library, 260–284.
- [8] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101.
- [9] Joy Burrough-Boenisch. 1999. International reading strategies for IMRD articles. *Written Communication* 16, 3 (1999), 296–316.
- [10] Jacob Cohen. 1988. *Statistical power analysis for the behavioral sciences*. 2nd. Hillsdale, NJ: erlbaum.
- [11] Fergus IM Craik and Robert S Lockhart. 1972. Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior* 11, 6 (1972), 671–684.
- [12] Geoff Cumming. 2013. *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. Routledge.
- [13] Geoff Cumming and Robert Calin-Jageman. 2016. *Introduction to the new statistics: estimation, open science, and beyond*. Routledge. [www.routledge.com/cw/cumming](http://www.routledge.com/cw/cumming) (ESCI version 17 November 2016).
- [14] Geoff Cumming and Sue Finch. 2005. Inference by eye: confidence intervals and how to read pictures of data. *American Psychologist* 60, 2 (2005), 170.
- [15] Pierre Dragicevic. 2016. *Fair statistical communication in HCI*. Springer International Publishing, Cham, 291–330. [https://doi.org/10.1007/978-3-319-26633-6\\_13](https://doi.org/10.1007/978-3-319-26633-6_13)
- [16] Sigal Eden and Yoram Eshet-Alkalai. 2012. Print versus digital: the effect of format on performance in editing text. In *Proceedings of the Chais Conference on Instructional Technologies Research*.
- [17] Franz Faul, Edgar Erdfelder, Albert-Georg Lang, and Axel Buchner. 2007. G\*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods* 39, 2 (01 May 2007), 175–191. <https://doi.org/10.3758/BF03193146>
- [18] John Fox and Sanford Weisberg. 2018. *An R companion to applied regression*. Sage Publications.
- [19] Matti Hannus and Jukka Hyönä. 1999. Utilization of illustrations during learning of science textbook passages among low-and-high-ability children. *Contemporary Educational Psychology* 24, 2 (1999), 95–123.
- [20] Mary Hegarty and Marcel-Adam Just. 1993. Constructing mental models of machines from text and diagrams. *Journal of Memory and Language* 32, 6 (1993), 717–742.
- [21] Hsin Ning Jessie Ho, Meng-Jung Tsai, Ching-Yeh Wang, and Chin-Chung Tsai. 2014. Prior knowledge and online inquiry-based science reading: evidence from eye tracking. *International Journal of Science and Mathematics Education* 12, 3 (2014), 525–554.
- [22] Kenneth Holmqvist, Marcus Nyström, Richard Andersson, Richard Dewhurst, Halszka Jarodzka, and Joost Van de Weijer. 2011. *Eye tracking: a comprehensive guide to methods and measures*. OUP Oxford.
- [23] Jana Holsanova, Nils Holmberg, and Kenneth Holmqvist. 2009. Reading information graphics: the role of spatial contiguity and dual attentional guidance. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition* 23, 9 (2009), 1215–1226.
- [24] Anthony J Hornof and Tim Halverson. 2003. Cognitive strategies and eye movements for searching hierarchical computer displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 249–256.
- [25] Jukka Hyönä, Robert F Lorch, and Mike Rinck. 2003. Eye movement measures to study global text processing. In *The Mind's Eye*. Elsevier, 313–334.
- [26] Jukka Hyönä, Robert F Lorch Jr, and Johanna K Kaakinen. 2002. Individual differences in reading to summarize expository text: evidence from eye fixation patterns. *Journal of Educational Psychology* 94, 1 (2002), 44.
- [27] Robert JK Jacob and Keith S Karn. 2003. Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. In *The Mind's Eye*. Elsevier, 573–605.
- [28] Cheryl I Johnson and Richard E Mayer. 2012. An eye movement analysis of the spatial contiguity effect in multimedia learning. *Journal of Experimental Psychology: Applied* 18, 2 (2012), 178.
- [29] Marcel Adam Just and Patricia A Carpenter. 1976. Eye fixations and cognitive processes. *Cognitive Psychology* 8, 4 (1976), 441–480.
- [30] Dae Hyun Kim, Enamul Hoque, Juho Kim, and Maneesh Agrawala. 2018. Facilitating document reading by linking text and tables. In *The 31st Annual ACM Symposium on User Interface Software and Technology*. ACM, 423–434.
- [31] Ha-Kyung Kong, Zhicheng Liu, and Karrie Karahalios. 2018. Frames and slants in titles of visualizations on controversial topics. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 438.
- [32] Nicholas Kong, Marti A Hearst, and Maneesh Agrawala. 2014. Extracting references between text and charts via crowdsourcing. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. ACM, 31–40.
- [33] Jill H Larkin and Herbert A Simon. 1987. Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science* 11, 1 (1987), 65–100.
- [34] Brandon LeBeau. 2018. *pdfsearch: search tools for PDF files*. <https://CRAN.R-project.org/package=pdfsearch> R package version 0.2.3.
- [35] Clayton Lewis. 1982. Using the thinking-aloud method in cognitive interface design. *Research Report RC9265, IBM TJ Watson Research Center* (1982).
- [36] Ziming Liu. 2005. Reading behavior in the digital environment: Changes in reading behavior over the past ten years. *Journal of Documentation* 61, 6 (2005), 700–712.
- [37] Regan Mandryk, Mark Hancock, Mark Perry, and Anna Cox (Eds.). 2018. *CHI '18: proceedings of the 2018 CHI conference on human factors in computing systems* (Montreal QC, Canada). ACM, New York, NY, USA.
- [38] Salvatore S Mangiafico. 2015. An R companion for the handbook of biological statistics. (2015).
- [39] Lucia Mason, Patrik Pluchino, Maria Caterina Tornatora, and Nicola Ariasi. 2013. An eye-tracking study of learning from science text with concrete and abstract illustrations. *The Journal of Experimental Education* 81, 3 (2013), 356–384.
- [40] Richard E Mayer. 2005. Principles for managing essential processing in multimedia learning: segmenting, pretraining, and modality principles. *The Cambridge Handbook of Multimedia Learning* (2005), 169–182.
- [41] Danielle S McNamara, Max M Louwerse, Philip M McCarthy, and Arthur C Graesser. 2010. Coh-Metrix: capturing linguistic features of cohesion. *Discourse Processes* 47, 4 (2010), 292–330.
- [42] Frank Mittelbach. 2014. How to influence the position of float environments like figure and table in LATEX? *TUGboat* 35, 3 (2014), 248–254.
- [43] Adelheid A. M. Nicol and Penny M. Pexman. 2010. *Displaying your findings: a practical guide for presenting figures, posters, and presentations*. American Psychological Association.
- [44] University of Chicago Press. 2017. *The Chicago manual of style* (7 ed.). University of Chicago Press.
- [45] R Development Core Team. 2008. *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org> ISBN 3-900051-07-0.
- [46] Jeff Sauro and James R Lewis. 2010. Average task times in usability tests: what to report?. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2347–2350.
- [47] Florian Schmidt-Weigand, Alfred Kohnert, and Ulrich Glowalla. 2010. A closer look at split visual attention in system-and self-paced instruction in multimedia learning. *Learning and Instruction* 20, 2 (2010), 100–110.
- [48] Wolfgang Schnotz. 2005. An integrated model of text and picture comprehension. *The Cambridge Handbook of Multimedia Learning* 49 (2005), 69.
- [49] Abigail J Sellen and Richard HR Harper. 2003. *The myth of the paperless office*. MIT press.
- [50] Laurence D Smith, Lisa A Best, D Alan Stubbs, Andrea Bastiani Archibald, and Roxann Roberson-Nay. 2002. Constructing knowledge: the role of graphs and tables in hard and soft psychology. *American Psychologist* 57, 10 (2002), 749.
- [51] John Sweller, Jeroen JG Van Merriënboer, and Fred GWC Paas. 1998. Cognitive architecture and instructional design. *Educational Psychology Review* 10, 3 (1998), 251–296.
- [52] Edward R Tufte. 2006. *Beautiful evidence*. Vol. 1. Graphics Press Cheshire, CT.