Learning and Instruction xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Learning and Instruction

journal homepage: www.elsevier.com/locate/learninstruc

Task-irrelevant data impair processing of graph reading tasks: An eye tracking study

Benjamin Strobel*, Marlit Annalena Lindner, Steffani Saß, Olaf Köller

Leibniz-Institute for Science and Mathematics Education (IPN), Kiel, Germany

ARTICLE INFO

Keywords:
Graph comprehension
Task processing
Cognitive load
Eye tracking
Linear mixed-effects models

ABSTRACT

For Instruction, teachers often rely on prefabricated material that may include irrelevant information. However, graphs can place a heavy burden on the cognitive system if their complexity is not suitable for a given task. In this study, we compared bar graphs showing task-irrelevant data points or task-irrelevant data series with a control condition using a within-subject design and eye tracking methodology. Data were analyzed using linear mixed-effects models. Results show that task-irrelevant data significantly elevated processing time, error rate and cognitive load. Even though perceptual grouping by color was expected to aid the process when a task-irrelevant data series was included in a graph, effects were strongest in this condition. Analyses of attention distribution using eye tracking measures revealed that task processing differed qualitatively between the conditions, yielding important implications for instruction.

1. Introduction

Visualizations such as graphs and diagrams play an important role in everyday life and can be found in newspapers, science, engineering, and education (e.g., Glazer, 2011; Mayer, 2009; Pereira-Mendoza, Goh, & Bay, 2004; Purchase, 2014). They are especially important in the context of problem solving (Baker, Corbett, & Koedinger, 2001), for teaching and learning mathematics (Cucuo & Curcio, 2001) and for understanding scientific data (Shah & Hoeffner, 2002).

In learning, graphs provide cognitive support by offering computational advantages, such as perceptual grouping (see Wertheimer, 1938) and serving as external memory (Larkin & Simon, 1987; Tory & Möller, 2004). In contrast, inferences from non-spatial representations (e.g., text) are often more demanding, because some information must be computed at great cognitive expense (e.g., comparing numbers in a continuous text). However, graphs do not always make comprehension of information more effective and less demanding. In many learning situations graphs display more information than is relevant to readers, because teachers often rely on prefabricated material from textbooks and other sources. For example, in a graph that displays population sizes over several years, only a subset of years might be of interest to a learner (e.g., in a comparison task where a learner has to compare a number of specific data points). In this example, *task-irrelevant data* contribute to the overall *complexity* of the graphs, resulting in a higher complexity that is unnecessary for a learner to complete the given tasks.

Even though data complexity has rarely been the focus of graph comprehension research, it has been investigated as a background variable in several studies (e.g., Casali & Gaylin, 1988; Meyer, Shinar, & Leiser, 1997; Schutz, 1961a, 1961b; Spence & Lewandowsky, 1991). Still, more recently published studies provide some evidence that a high complexity negatively impacts performance in regular graph tasks (e.g., Kim & Lombardino, 2015; Kumar & Benbasat, 2004). Understanding a graph can be especially challenging if its complexity is not suitable for the task, placing a heavy burden on the cognitive system (Demetriadis & Cadoz, 2005; Huang, Eades, & Hong, 2009; Sedig & Parsons, 2013). Because human memory is a limited capacity information processing system (Baddeley & Hitch, 1974), this burden may result in a cognitive overload for the reader (Sweller, 1994). But does complexity affect graph processing and related task performance even if the additional data are completely irrelevant to the task at hand?

In the present study we focus on task-irrelevant data as a source of complexity in graphs. We compare different situations of task-irrelevant data: (1) When a given series of data points includes more data points than necessary to complete a task (*task-irrelevant data points*) and (2) when a second, task-irrelevant series of data points is presented next to a relevant data series (*task-irrelevant data series*). We investigate effects of task-irrelevant data on error rates, processing time and cognitive load. Additionally, (3) we explore underlying processes by applying eye tracking methodology, which has proven useful in previous graph comprehension studies (e.g. Kim & Lombardino, 2015; Strobel, Saß, Lindner, & Köller, 2016).

* Corresponding author. Leibniz Institute for Science and Mathematics Education, Olshausenstraße 62, 24118 Kiel, Germany.
E-mail addresses: strobel@ipn.uni-kiel.de (B. Strobel), mlindner@ipn.uni-kiel.de (M.A. Lindner), sass@ipn.uni-kiel.de (S. Saß), koeller@ipn.uni-kiel.de (O. Köller).

http://dx.doi.org/10.1016/j.learninstruc.2017.10.003
Received 30 March 2017; Received in revised form 29 August 2017; Accepted 22 October 2017

1.1. Data complexity

Data complexity is for the most part reflected in the number of data points and variables displayed in a graph. According to Meyer et al. (1997), data complexity can be understood as a result of three factors: (a) the number of data points in a graph, (b) the configuration of the data points (i.e., the organization of data points into data series), and (c) the regularity of the data.[1]

Schutz (1961a, 1961b) was one of the first researchers to investigate the number of data points. Looking at tasks that require the identification of trends, he found that processing time increased with the number of irrelevant data points regardless of the graph type (bar and line graphs), but they did not exhibit differences in accuracy. In line with these findings, Kumar and Benbasat (2004) found elevated processing times when additional data points were included in a graph. In a series of three experiments, Spence and Lewandowsky (1991) compared multiple types of representations including bar and line graphs. In contrast to Schutz, they found that accuracy in regular comparison tasks (i.e., the comparison of two or more data points) was significantly lower when more data points were present.

Regarding the organization of data into data series, Schutz (1961b) found that processing time increased with the number of data series in line graphs, but only if multiple graphs were used (i.e., one graph per data series). Single graphs displaying multiple lines were unaffected by the data series number. Recently, Kim and Lombardino (2015) conducted an eye tracking study and varied the number of data series and the task type (i.e., single point location vs. point comparison). They found that processing time in both task types was significantly higher when an additional data series was present. This was also reflected in longer fixation times on the graph regions.

Data complexity has consistently shown effects on task performance. In addition to performance measures, Huang et al. (2009) exhibited cognitive load with a subjective rating scale for mental effort and found that reported mental effort increased with the level of data complexity. Along with processing times and task performance, cognitive load measures are important to reveal effects on the burden of working memory.

1.2. Cognitive Load Theory

Cognitive Load Theory (CLT; Sweller, 1988) describes how cognitive processing can be facilitated or inhibited under the constraints of a limited working memory capacity. Three types of cognitive load are distinguished in CLT: *Intrinsic load* refers to the inherent complexity of the information, especially the number of interactive elements that must be understood in relation to each other. *Extraneous load* originates from the instructional material und involves unnecessary processing of irrelevant or unrelated information. Finally, *germane load* describes the mental effort invested by the learner to comprehend the material and involves processes such as interpreting and organizing (De Jong, 2010; Sweller, 1994).

A high data complexity in graphs may result in cognitive overload if the complexity is not suitable for the task (Demetriadis & Cadoz, 2005; Huang et al., 2009; Sedig & Parsons, 2013). However, in many of the presented studies the additional data were relevant to the task. In the context of CLT, there are two reasons to distinguish data that are required to solve a given task from task-irrelevant data. First, data necessary to complete a given task contribute to intrinsic cognitive load, because the information that must be extracted is inherently more complex. Task-irrelevant data on the other hand are expected to induce extraneous cognitive load because they give rise to unnecessary processing of irrelevant information. Second, while a higher burden on working memory is to be expected when a bigger amount of data *must* be processed (i.e., when additional data is required to complete the task), it is unclear if and to what extent task-irrelevant data are processed during task-oriented graph reading. For example, if the oil price of two specific years is of interest to a learner, giving the prices for a wide range of years (i.e., showing task-irrelevant data points) might induce additional load and make the task more difficult. When data points are inserted into an existing data series like in this example, readers have to refer to the corresponding labels (i.e., the specific years) in order to distinguish relevant from irrelevant data. Given an additional data series on the other hand, a color-coded legend allows for perceptual grouping of information in the graph (e.g., Freedman & Shah, 2002; Pinker, 1990). Here, task-relevant and task-irrelevant data points are easily distinguishable by color, eliminating the need to check multiple labels and thus facilitating task processing.

In summary, task-irrelevant data can be expected to induce extraneous cognitive load, to extend the processing time of the task and to make this process less accurate. Based on the principles of perceptual grouping, these effects may be mitigated when task-irrelevant data can be identified by color.

1.3. Using eye tracking to gain insight into processing of graph reading tasks

The use of eye tracking (for an introduction, see Duchowski, 2007; Holmqvist et al., 2011) in cognitive psychology is based on the assumption that the location of eye-fixations represents the focus of attention. In other words, it is assumed that we process the visual information we are currently looking at. This idea is called the eye-mind hypothesis (Just & Carpenter, 1980). Even though shortcomings of the eye-mind hypothesis have been discussed (Hyönä, 2010; Wright & Ward, 2008), researchers have shown that eye-fixation measures and cognitive performance are closely related (e.g. Canham & Hegarty, 2010; Jarodzka, Scheiter, Gerjets, & van Gog, 2010; Lindner, Eitel, Strobel, & Köller, 2017). In contrast to other process tracing methods (e.g., verbal protocols; Ericsson & Simon, 1980; Van Gog, Kester, Nievelstein, Giesbers, & Paas, 2009), eye tracking does not place additional load on participants' working memory.

Mayer (2010) suggested that eye tracking technology can contribute to the study of learning with graphics by providing information relevant to the instructional design of graphics. In traditional experiments, we may draw conclusions about if and when a manipulation of the material has an effect on performance measures (e.g., accuracy, processing time). However, we often are clueless on *how* processing of the material changes. In the context of graph reading, eye tracking allows us to allocate processing time to the spatiotemporal attention of the graph reader. In a number of recent graph studies, researchers were able to successfully attribute processing time to important subregions of graph tasks (i.e., x-axis, y-axis, legend, pattern, question, answers; Kim & Lombardino, 2015; Peebles & Cheng, 2003; Strobel et al., 2016). For the present study, eye tracking enables us to estimate how much time can be attributed to the processing of specific regions in the display, such as task-irrelevant data points in graphs.

1.4. Research hypotheses

In this study, we compared two different types of data complexity: (1) task-irrelevant data points (DP) that are included in an existing data series and (2) a task-irrelevant data series (DS) presented next to an existing data series in comparison to a control condition with no task-irrelevant data. Additionally, (3) we explored differences in task processing by using eye tracking methodology. In summary, we expected the following:

(1) *Task-Irrelevant Data Points*: If task-irrelevant DP are inserted into an existing DS, graph readers need to examine their labels in order to distinguish task-irrelevant from relevant data. Accordingly, we

[1] Data regularity is not investigated in the present study, because task-irrelevant data is not the source of complexity in this factor (for findings on data regularity see Kumar & Benbasat, 2004; Schutz, 1961b).