COGNITIVE LOAD IN SMARTPHONE CALENDAR APPLICATIONS
Adrian Hellqvist, Tommy Olsson
Abstract

The use of smartphone applications has increased dramatically in the last decade. With the technology being ubiquitous around us it is important to reduce the required cognitive demand for interacting with the technology. A common use of smartphones is calendar applications. This paper investigates if cognitive load can be lowered by using a calendar application specifically designed with cognitive load in mind compared to the widely used Google Calendar. The hypothesis is that the application will generate a lower cognitive load in comparison. An experiment was conducted where participants were instructed to perform a primary task by following given instructions of what to do in the calendar application while also performing a secondary task. There were no significant differences in performance except for two tasks where the new app performed worse. Participants also reported a higher level of frustration for the new app which could mean that their cognitive load was slightly higher.

Abstrakt

Användandet av smartphone applikationer har ökat dramatiskt under det senaste decenniet, eftersom att tekniken hela tiden är närvarande runt om oss så är det viktigt att minska de kognitiva kraven som behövs vid användandet. Ett vanligt användande av smartphones är kalenderapplikationer, den här studien undersöker om den kognitiva belastningen kan minskas vid användandet av en kalender app genom att använda applikation som designats med kognitiv belastning i åtanke för att sen jämföras mot Google Kalender. Hypotesen är att applikationen kommer att generera en lägre kognitiv belastning i jämförelse. Ett experiment utfördes där deltagarna instruerades att utföra en huvuduppgift som innebar att utföra givna instruktioner med uppdrag i kalenderapplikationen, samtidigt som en sekundär uppgift utfördes där deltagarna fick trycka på en knapp när en signal hördes. Det var ingen signifikant skillnad i prestation förutom i två uppgifter där appen presterade sämre, deltagarna rapporterade också en högre nivå av frustration vilket skulle kunna innebära att den kognitiva belastningen var något högre.
Cognitive load in smartphone calendar applications

Introduction

Smartphone and technology usage in the last decade increased dramatically and the usage has become a part of our everyday life (TechCrunch, 2015). With the increasing amounts of technical interactions, cognitive load in everyday situations is further increased with the pervasive use of smartphone technology. For instance, drivers experience a higher cognitive load when driving while also using cell phones (Block, Hancock & Zakay, 2010). Research has shown that the increasing complexity of technology makes high cognitive load ubiquitous (Hancock & Szalma, 2008). There is therefore a need to reduce the cognitive load. Since our mental resources are limited the decrease of cognitive load makes less mental resources needed which can be helpful in our daily lives.

Smartphones offer a variety of different applications to help with everyday tasks. A common use of smartphone technology is scheduling with the use of calendar applications. This study will investigate if cognitive load can be minimized by using simplified calendar application developed for this project with the purpose of decreasing cognitive load. The application will be compared to Google Calendar to measure differences in cognitive load.

Cognitive load theory

Several competing models of working memory coexist today but all models agree on that working memory as a resource has a limited capacity (Baddeley, 2012). Cognitive load theory (CLT) was introduced by Sweller (1988) where cognitive load is described as the exertion on working memory while interacting with a system or solving a task. Engaging in more complex activities institutes a greater cognitive load on the subject. Cognitive load can be divided into three categories: intrinsic, extraneous and germane.

Intrinsic cognitive load is load that is imposed by the task itself, the level of intrinsic load is dependent only on the complexity of the task and the knowledge level of the person. Germane cognitive load involves the processing of learning, specifically the processing construction and automation of schemas. Extraneous load in return is concerned with presentation of information and the majority of research on CLT concerns techniques involved in trying to reduce extraneous load.

Historically most studies on cognitive load mainly relied on subjective and/or indirect methodologies in the form of self-reported questionnaires (Paas & van Merriënboer, 1993). Since then several objective and direct methods has been brought forward in the form of both physiological and behavioral measurements. These include neuroimaging techniques and the dual-task methodology prominent in much of working memory research (Brunken, Plass & Leutner, 2003).

Inducing cognitive load

The dual-task paradigm has been used in studies to increase the cognitive load of the participants by introducing a secondary task to induce memory load. The dependent variable of interest in this approach is the performance in the primary task.

In a study by Marcus, Cooper & Sweller (1996) participants with no previous experience in the subject of electricity and connecting electrical resistors were given tasks to connect electrical resistors by being shown what a resistor and connector is, they were also given some background information to read on how these are connected.
Since the participants had no prior experience with the task the execution had higher cognitive demands. To further increase the cognitive load, the participants were given a secondary task to perform in addition to the primary task where they were told that they had to press a foot pedal as quickly as possible when a tone was heard within an interval of every 5 to 20 s.

In another study by Madrid, Van Oostendorp & Melguizo (2009) an experiment was conducted where the effects of number of hyperlinks and how it affects navigation support on cognitive load and learning were analyzed. The experiment included either a higher number of hyperlinks or a fewer number that had link suggestions for navigation support. The participants got to read a text containing 4400 words with multiple hyperlinks to navigate through the text to help with learning the content. The link visibility was designed in a way that the users would not be able to see the links during the reading until they had pressed a button to indicate that they were finished reading. After a link was chosen the selection menu disappeared and a new text was presented. This method was used to test cognitive load during reading and selection phase. In addition to the primary task to further increase the cognitive load the participants were told to press the “Z”-button on the keyboard every time a beep sound was heard, the beeps had an interval between 15-45 s during the reading phase and 4-9 s during the link selection phase.

In the present paper it is investigated if there is a difference in cognitive load and time to complete tasks in two different smartphone calendar applications. The applications are Google Calendar and another application created for this study named Eye that emphasizes on reducing cognitive load by reducing visual clutter and only show relevant information for the calendar activities. Cognitive load is also reduced by separating activities in two different categories depending on whether the activity is a meeting that has to be replied to or not. The hypothesis is that cognitive load will be lower for Eye.

**Method**

**Developing a scheduling app with cognitive load in mind**

A calendar application for the Android platform was created prior to the experiment to be compared to Google Calendar. The purpose of the application is to be a scaled down calendar with limited functionality.

This section describes the development process for the scheduling application with the working title of “Eye”. The application was requested by a local company looking for a way to represent their clients’ schedules through a smartphone app. Requested features included the possibility to view upcoming activities, see available appointments and be able to accept/decline these appointments. The commissioning company uses Outlook for their internal email and scheduling software, Microsoft’s Office 365 and Outlook Calendar was used as the backend platform for that reason.

During development aspects of the design affecting cognitive load were kept in mind and especially extraneous cognitive load aspects were sought to be minimized. According to Whitenton (2013) when a website reduces the needed cognitive load to operate that site its usability increases. The total mental processing power needed to use a website affects how users find content and complete tasks more easily. During our development we have strived to implement some of the ideas that Whitenton brings up such as avoiding to show unnecessary clutter and information for the users, or building on existing mental models of how a calendar works.
Other ways to reduce cognitive load were also taken into consideration, such as segmenting, weeding, and signaling (Mayer & Moreno, 2003). Originally introduced as ways to improve multimedia learning, we sought to implement these guidelines during our development.

Segmenting, similar to the cognitive psychology term chunking, entails the process of breaking down a presentation into bite-size segments. Hence in our application activities and meetings are separated and treated as two different kinds of items.

Weeding, the elimination of interesting but extraneous material, was implemented by only including the current week's calendar events. Furthermore, we only included essential parts of each calendar entry such as the title, date, time, description and location.

Signaling or providing cues so that the user is able to focus on essential information is implemented through the use of making past calendar events gray and semi-transparent.

Use

After logging on to the application the user is provided with three options: Hem (home), Kalender (calendar) and Möten (meeting) to navigate through (see Figure 1).

The Home screen provides the user with today’s date and the three forthcoming activities. The calendar screen shows all the coming activities in a list, or all the activities for the current week displayed as a week calendar, the meeting screen displays activities that the user have been invited to. The meeting activities can be either accepted or declined by tapping the meeting item for the additional information screen, where a button can be clicked to accept or decline (see Figure 2).

Figure 1. Screenshots of different views in the Eye application. Displayed is the home screen and the two different ways to view activities.
Figure 2. Screenshots of different views in the Eye application. The screenshots show the screen where you are able to view meetings, and the additional information screens when selecting an activity or meeting.

Development tools

The application uses Microsoft Graph (2018), a developer platform and API for connecting services to Office 365. A snippet code provided by Microsoft on Github was implemented into the application in order to integrate with Outlook Calendar. Java and Android Studio were used for development with focus on following the design guidelines provided in Google's Material Design (2018).
The empirical study

Participants

20 persons participated in the experiment, 12 were men and 8 women ranging from the age of 21 to 34 (avg. 25.58). Participation was voluntary and no form of compensation was offered to any of the participants.

Materials

A python script played a beeping sound at a random time with an interval of 2 to 10s and a keyboard was provided for the participants to click a button after the sound was heard. A stopwatch was used to measure time of completion. NASA-TLX questionnaires (Hart, 1988), the most common tool to measure subjective cognitive load, and another form was provided to collect information on previous experience with smartphone use and Google’s calendar application.

Google calendar was used in comparison to Eye. After logging on to Google calendar (see Figure 2), the user is provided with a list of activities starting from the current day. When pressing the menu button, a navigation drawer is displayed with the options of listing the activities by showing them in a list, by day, by every three days, by every week or by every month. There is also an option to search for activities or to create a new activity. As well as options for settings and for help and feedback.

Figure 3. Screenshots of Google Calendar
Procedure

Prior to the testing of the experiment, a pilot study was conducted to investigate if there was anything in the design of the test that did not work. After the pilot study some changes were made to the experiment to further increase the induction of cognitive load, and some additional changes were made to the application to reduce confusion of the participants. Another pilot study was conducted after the changes were made to test if the experiment worked as intended.

The experiment was carried out in an experimental environment in a study room at Umeå University and consisted of three segments. In the first segment the users answered a questionnaire regarding their previous experience of using smartphones and the Google calendar application. The questionnaire included two questions to be answered using a 1-5 Likert scale, scaling from “Never” to “Very often” with the questions “Do you use smartphone applications in your daily life?” and “Do you use Google Calendar?”

In the second segment the participants got to perform two different tasks, one primary and one secondary that the participants were asked to perform at the same time. In the primary task the users received tasks to complete in two different calendar applications, using the calendar app created for this study and Google calendar. The participants received the tasks on a piece of paper one at a time with instructions on what to do, when they felt ready to perform the task the participant handed over the paper to the study director and received a mobile phone to perform the task on.

The tasks were the following:

1. Check what activity occurs on Friday the 11th at 13:00.
2. Check what the three next upcoming activities are.
3. Look up the activity “Handledarmöte” and accept the invite.
4. Look up the activity “Möte med chefen” and decline the invite.
5. Look up the activity “Fika” and check what location the activity takes place at.
6. Look up the activity “Lektion” and check for how long time the activity lasts.

While performing the primary task the participants were also asked to perform a secondary task which consisted of having to press a keyboard button when a beep sound was heard. The beep sound arose at a random time with an interval of 2-10s between beeps. The third segment occurred after the second segment was finished. The participants got to answer a NASA-TLX questionnaire to measure cognitive load regarding the experience of the second segment.
Results

Measuring prior knowledge showed that all participants use smartphone applications in their daily life (mean of 4.6). Familiarity with Google Calendar is much lower, however (mean 2.25).

When analyzing differences in performance, mean time (see figures 4,5) and standard deviation were first calculated to control if the task data were normally distributed. None of the tests were normally distributed and as such the data was analysed with a Mann-Whitney U-test. Two of the tasks (task 1 and task 5) showed a significant difference between the apps ($p < 0.05$), in both cases Eye being significantly slower. The four other tasks did not show a significant difference between the apps although Eye was a little bit faster in general.

Due to some outliers, additional analyses were performed. When excluding outliers (2.5 standard deviations higher than mean time), task 3 time results were normally distributed, although the t-test showed no significant differences.

Furthermore, some participants reported a high level of experience using Google Calendar (4 and above), which may have given them an advantage in using that application due to prior training. After excluding these participants task 1 and 3 had a normal distribution. In both cases Eye was slower but t-test showed no significant differences.

![Figure 4. Mean time for individual tasks.](image-url)
The reported cognitive load from the NASA-TLX scores showed a normal distribution for the frustration-parameter for both applications, a paired sample t-test showed a significant difference ($p < 0.05$) where Eye scored a higher degree of frustration. The four other parameters were analysed with Mann-Whitney U-tests but provided no significant differences. However, for each parameter the mean score for Eye was higher than for Google Calendar (see Figure 6).

The NASA-TLX scores were further investigated with the same exclusion criterias used in the performance analysis and although Eye’s test scores improved, reducing the difference between the applications, no significant differences were found and Google calendar still scored lower means for each parameter.

![Figure 5. Mean time for performing all tasks combined.](image)

![Figure 6. Mean NASA-TLX scores.](image)
Discussion

In this paper two smartphone calendar applications were compared to investigate differences in cognitive load by measuring completion time for different tasks, such as finding out information about a specific calendar entry. The results showed a difference in time of completion in two out of six tasks where Google Calendar proved to be significantly faster compared to Eye. Results also showed that users experienced Eye to be more frustrating to use, which indicates a higher cognitive load. Even though the Eye application did not perform as well as Google Calendar the application still performed as good in four out of six of the tasks and five out of six parameters in the cognitive load measurement. Based on this it is hard to draw any comprehensive conclusions.

One issue that may have affected the results of this study is the low number of participants in the experiments, which resulted in that there were only ten participants performing each task in each application. When excluding outliers and participants with high experience of using Google Calendar the remaining data points decreased in number even more and for a more reliable analysis this should be taken into consideration.

Another issue that may have affected both the completion time and the level of frustration is that Google Calendar already is a well-known and established application with a design widely used for different Google applications. According to Whitenton (2013), when applications are built on existing mental models on how websites work based on previous experience then cognitive load will be reduced since the amount of learning to use the applications is reduced.

Additionally some parts of the Eye application could not be finalized as intended due to limitations in resources and time, including the “timeline” and the “weekly” screens. The layout of these screens was meant to be revamped but since these screens were not finalized a majority of participants missed the weekly screen feature and only used the timeline layout. With more time available a “Tab layout” would have been implemented to further indicate the existence of the weekly screen which could have influenced Eye’s performance.

Before every task all running applications were killed to standardize for the next task. A problem with this approach is that the Eye-application does not save any scheduling data locally on the device between sessions. Instead the application fetches this data from Microsoft’s API on startup, resulting in a 3 to 4 second delay before the user is able to start navigating the interface. This has obviously impacted the Eye-applications time scores and may also have contributed to the significant difference in NASA-T LX frustration ratings between the two applications. The issue with the delay could have been prevented by first loading the data into the application before handing over the phone, unfortunately this issue was only discovered after a few participants had already done the experiment. Due to limited time it was decided to continue the testing with the same procedure. Performing more pilot studies prior to the experiment may have prevented this.

Another issue that may have affected the outcome of the experiment is that all the activities in the tasks were scheduled chronologically during the same week and therefore there was not a great need of navigation through the calendar. The Eye application was developed for a client to be used in their daily work as a calendar application for their participants and has as such been designed after their needs. It would be interesting to investigate if there would be a difference in the results if there was a greater range of non-meeting activities and meeting activities over the time span of weeks or months to increase the ecological validity for real life use. It would also be interesting to test the application over a longer period of time to see differences in learnability and cognitive load differences in experienced users.
Reference list


