Communications of the Association for Information Systems

Volume 39 Article 2

7-2016

Why Forwarded Email Threads are Hard to Read: The Email Format as an Antecedent of Email Overload

Nikolai Sobotta
Goethe University Frankfurt, sobotta@wiwi.uni-frankfurt.de

Communications of the Association for Information Systems

Research Paper ISSN: 1529-3181

Why Forwarded Email Threads are Hard to Read: The Email Format as an Antecedent of Email Overload

Nikolai Sobotta
Faculty of Economics and Business Administration
Goethe University Frankfurt
sobotta@wiwi.uni-frankfurt.de

Abstract:

Research has shown that excessive email use leads to feelings of being overwhelmed and stressed. Existing coping solutions, which mitigate email overload, address the number of emails and, in consequence, the time spent on emails. These approaches are congruent with existing research on antecedents of email overload. Further coping solutions include addressing email threads. However, we lack a theoretical grounding for perceiving email threads as an antecedent of email overload. I suggest cognitive load theory as a means of investigating the format of forwarded email threads in an experiment. I found support for the effects on reading time and performance in terms of correct answers per second, findings that confirm that forwarded email threads are an antecedent of email overload and that we need a new conceptualization of email overload.

Keywords: Email Overload, Forwarded Email Threads, Email, Cognitive Load Theory, Experiment

This manuscript underwent editorial review. It was received 11/19/2015 and was with the authors for 3 months for 1 revision. E. Vance Wilson served as Associate Editor.

1 Introduction

Email is one of the most widely used means of communication. Although email technology is several decades old, it remains well established and offers advantages such as being independent of space and time. However, email use has its disadvantages, such as their high number's resulting in one's feeling overwhelmed (e.g., Dabbish & Kraut, 2006; Whittaker & Sidner, 1996), which can cause consequences such as increased heart rate (Mark, Voida, & Cardello, 2012) and stress (Barley, Meyerson, & Grodal, 2011). Due to the active usage of emails in practice, this phenomenon and its consequences are still present today (Grevet, Choi, Kumar, & Gilbert, 2014).

Existing research investigates this phenomenon as a technology-specific instance of information overload, which researchers (e.g., Gill, 1998; Schroder, Driver, & Streufert, 1967) define as an "excessive supply" (Savolainen, 2007, p. 614) of information. Because researchers have mainly attributed the high volume of information to the emergence of information technology (Bawden, 2001) and Internet use (Beaudoin, 2008), they note that the high number of emails that are sent back and forth in organizations each day represents one of the major causes of information overload (Schultz & Vandenbosch, 1998). Therefore, researchers have used "email overload" to refer to information overload caused by emails (Sumecki, Chipulu, & Ojiako, 2011), a term I use in this paper. Following this definition, researchers assess email overload by the volume of emails one receives and the time one spends on reading and writing emails (e.g., Dabbish & Kraut, 2006; Sumecki et al., 2011). In general, researchers have supposed that the more time people spend on their emails, the higher the degree to which people feel overloaded (Barley et al., 2011). In addition, researchers have investigated email-based interruptions through incoming messages, the resulting impact on time spent on emails, and the optimal timing for reading emails (Gupta, Li, & Sharda, 2013; Gupta, Sharda, & Greve, 2011; Renaud, Ramsay, & Hair, 2006; Vidgen, Sims, & Powell, 2011).

To cope with email overload, computer science research has developed solutions to manage the high number of emails and reduce the time spent on them (e.g., Bälter & Sidner, 2002; Schuff, Turetken, & D'Arcy, 2006). Some researchers have integrated these functionalities into recent state-of-the-art email software (Sumecki et al., 2011). Therefore, one can argue that coping solutions, which address the number of email messages, are well established. Nevertheless, there are other approaches that are not well established. For example, mitigation approaches that try to help with reassembling email threads, which links them together into an email conversation (e.g., Dehghani, Shakery, Asadpour, & Koushkestani, 2013; Sharaff & Nagwani, 2015). Besides solutions for identifying email threads in a set of email messages, some research has suggested solutions to summarize text as a coping mechanism as a means to provide résumé of the content of email threads (Hashem, 2014). In conclusion, research concerned with the development of coping instruments has already acknowledged the problem of reading forwarded email threads by developing coping solutions. However, we lack insights on forwarded email threads as an antecedent of email overload. Established coping software addresses a high number of emails and, in consequence, the time one spends on emails, which research has also investigated as antecedents of email overload. Nevertheless, existing insights do not differentiate between different types of email formats, which may mislead our thinking because existing coping solutions suggest that one needs to approach forward email threads differently. Further, no research has yet developed a theoretical grounding for different types of email formats.

I address this research gap by investigating the format of forwarded email threads. Email threads are emails that comprise a history of emails sent back and forth between two or more recipients who always reply to messages on top or at the bottom of the preexisting chain of messages. I do not look at the emergence of these email threads but at the processing of all the quoted emails at once when such threads are forwarded to new communication participants. I investigate this situation because the increasing division of labor in terms of space and time promotes the forwarding of email threads. I scrutinize cognitive load theory (Sweller, 2010; Sweller, Ayres, & Kalyuga, 2011; Sweller, van Merrienboer, & Paas, 1998; Van Merriënboer & Sweller, 2005, 2010) to assess two measurements. First, I measure the time the recipients need to read forwarded email threads. Second, I record the accuracy in reading related tasks by counting the number of correct answer regarding the email content. We also consider the inherent tradeoff between speed and accuracy that readers face when answering content-based questions, which research refers to as speed-accuracy tradeoff (SAT) (Wickelgren, 1977). Therefore, I suggest focusing on the split-attention effect, which describes the circumstance of the increased time requirement for performing tasks caused by spatially unintegrated presentation formats (Chandler & Sweller, 1991, 1992; Sweller, Chandler, Tierney, & Cooper, 1990; Tarmizi & Sweller, 1988;

Ward & Sweller, 1990), to investigate the impact of forwarded email threads on email overload. As such, I investigate the following research question:

> **RQ:** Is the format of forwarded email threads an antecedent of email overload?

I conducted an independent-measures experiment with 86 students in which the participants had to answer comprehension tasks regarding the content of one text, which appeared either as a consolidated email thread or as a forwarded email thread. While the forwarded email thread appears as it emerges during the conversation, the consolidated email thread presents the same content but can be read from top to bottom and leaves out the meta-information in between the message segments (c.f. Figure 1 for an excerpt).

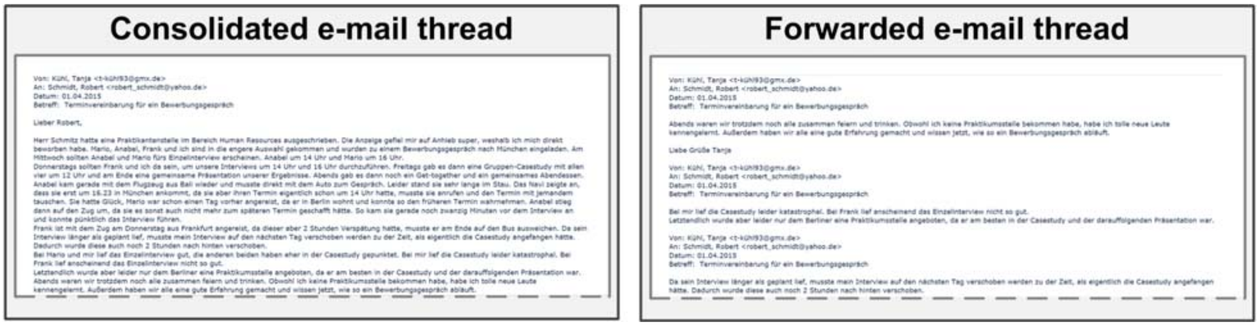

Figure 1. Excerpt of Consolidated Email Thread vs. Forwarded Email Thread

This experiment resulted in a longer reading time and a lower overall performance in terms of correct answers per seconds for forwarded email thread readers.

This paper makes two main contributions. First, I show that the format of forwarded email threads results in taking longer to read emails and, therefore, in email overload. With these results, I confirm that the problem of forwarded email threads exists and that we need to address it using coping solutions, which entails opportunities to further mitigate email overload. Second, I highlight that the email format matters and needs to be integrated into further conceptualizations of email overload to get beyond the high number of emails and time spent on emails.

This paper proceeds as follows. In Section 2, I discuss prior literature concerning email overload, its antecedents, and email threads. I also explain cognitive load theory in Section 3 with the split-attention effect as the theoretical foundation and then develop my hypotheses. In Section 4, I detail the experiment, and, in Section 5, I analyze and present the results. In Section 6, I discuss the results, describe the implications of the findings, and offer an outlook for further research. Finally, in Section 7, I conclude the paper

2 Related Research

According to recent research (Grevet et al., 2014), the term email overload refers to two distinct phenomena that have resulted in two different research streams with different research aims. Researchers observed the phenomenon for the first time when they related it to the high number and different types of emails stored in an inbox. Therein, researchers attributed email overload to the fact that users employed email for purposes other than interpersonal communication such as task management and personal archiving (Whittaker & Sidner, 1996). Therefore, the term email overload emerged as a research stream investigating the number of emails in an inbox. Consequently, researchers focused on technical solutions for coping with the amount of emails in an inbox, which contemporary email software has now partly integrated (Sumecki et al., 2011), such as automatically archiving emails and automatically classifying emails to drop them into specific folders (e.g., Bälter & Sidner, 2002; Schuff et al., 2006). Technical coping approaches try to go beyond reducing the amount of emails in an inbox and address the dependencies between email messages (Dehghani et al., 2013) by integrating emails into email conversations (Lewis & Knowles, 1997; Newman, 2002; Popolov, Callaghan, & Luker, 2000; Rohall, Gruen, Moody, & Kellerman, 2001; Whittaker and Sidner, 1996). An email conversation is a set of email messages mostly discussing a particular topic email users send to each other (Venolia & Neustaedter, 2003). Email threads are a vehicle for email conversations (Black, Levin, Mehan, & Quinn, 1983) and one can use them by applying the reply and append function (Murray, 2000). However, email applications do not summarize, analyze, or synthesize the previous messages, which is a task the applications leave to the user to make sense of the messages. A key characteristic of email threads is the temporal order they appear in (Kleinberg, 2003). Email conversations may span several threads with similar but not exactly the same subject lines (Erera & Carmel, 2008). Also, no standard protocol for the linkage between the email messages comprising an email thread is available (Yeh, 2006), and users inconsistently use the reply functionality in email applications (Minkov, Cohen, & Ng, 2006). Therefore, the email thread relationship, where the email messages are directly connected is not given explicitly; one must derive it from several hints (Klimt & Yang, 2004), which calls for an intelligent approach (Minkov et al., 2006) to reassembling email threads (Yeh, 2006). Coping solutions either reassemble email threads as a linear structure or a tree structure (Dehghani et al., 2013). Recent algorithms go beyond considering the subject and compare the content of the email messages (Sharaff & Nagwani, 2015). Despite this research effort into finding coping solutions, subsequent studies on email overload, have found that email overload has not significantly decreased (Grevet et al., 2014; Szóstek, 2011). These findings may suggest that the technical mechanisms that research has focused on to reduce the number of emails in an inbox do not work.

The development of coping solutions for email threads highlights the fact that email users have to deal with other issues apart from the presence of too many emails in their inbox. To that effect, recent research has conceptualized email overload by deriving it from the more general topic of information overload (e.g., Dabbish & Kraut, 2006; Sumecki et al., 2011) and focused on how humans deal with the high number of incoming emails (Grevet et al., 2014). Researchers have investigated information overload (e.g., Gill, 1998; Sumecki et al., 2011) as an "excessive supply" (Savolainen, 2007, p. 614) of information that overwhelms individuals who deal with it. At first, researchers saw information overload as an inverted "U-curve" with decreasing decision accuracy when too much information was present (e.g., Schroder et al., 1967). After discussing this viewpoint (e.g., Malhotra, Jain, & Lagakos, 1982), researchers have extended this perspective using the antecedent of available time (Schick, Gordon, & Haka, 1990). These two antecedents form the intrinsic factors of information overload (Jackson & Farzaneh, 2004) and are included in the contemporary understanding of information overload (Eppler & Mengis, 2004). Researchers recognized an early stage that computer-mediated communication severely increased the amount of available information in organizations through different technologies such as bulletin board systems (Hiltz & Turoff, 1985) or groupware applications (Schultz & Vandenbosch, 1998). Today, researchers predominantly attribute this increase of available information to the vast number of emails sent in an organization (Bawden, 2001). Regarding email more specifically, studies have consequently looked at the number of emails sent and received, the amount of time spent on emails, and people's perceptions about these numbers when observed (Sumecki et al., 2011). Researchers (e.g., Dabbish & Kraut, 2006; Ducheneaut & Bellotti, 2001; Ingham, 2003; Sumecki et al., 2011) agree that email overload occurs when users are overwhelmed by email (Sumecki et al., 2011), which one can attribute to constraints in the humans' cognitive capacity (Berghel, 1997; Heylighen, 2004; Kirsh, 2000). This stream of research has also identified several aspects that causes strain in humans other than the initial focus on the high number of emails (Dabbish & Kraut, 2006). For instance, studies have examined how often emails interrupt people, how much time they spend reading emails, and the optimal timing decision for checking for new emails (Gupta et al., 2013; Gupta et al., 2011; Renaud et al., 2006; Vidgen et al., 2011). Other research focuses on the perception of email as a business-critical tool (Sumecki et al., 2011) or how their personality may influence their sensitivity to email overload (Reinke & Chamorro-Premuzic, 2014).

In summary, the first research stream focuses on the number of emails in an inbox and on coping mechanisms to reduce the number of emails in an inbox. The second stream of email overload research builds on information overload research dealing with humans, who are not able to cope with the high number of incoming emails and its antecedents. Because research has shown that existing technological solutions, such as email thread reassembly, do not mitigate email overload, I consequently suggest investigating email threads as an antecedent of email overload by following recent research and considering email users.

3 Theoretical Background and Hypothesis Development

Researchers identified the potential link between the use of information technology and the occurrence of cognitive overload at an early stage (Vollmann, 1991). Cognitive load is also caused by the degree of ambiguity that is higher in electronic communication than in standard face-to-face communication (Kock, 2004, 2005, 2009). Encoding emails to compensate for absent stimuli increases cognitive load required for using email (Kock, 2007). Likewise, existing research has found support concerning the same mechanism for decoding (Kock, 2001a, 2001b). Cognitive load theory (Sweller, 2010; Sweller et al., 2011; Sweller et al., 1998; Van Merriënboer & Sweller, 2005, 2010) suggests that the human brain's architecture does not enable it to process an unlimited amount of information; its working memory in particular is limited. The research on information overload propose that the constraints of human processing capability originates from this limit (Berghel, 1997; Heylighen, 2004; Kirsh, 2000). Cognitive load theory distinguishes between working memory and long-term memory. Individuals use working memory to consciously process new information; however, one can process only a limited amount of information (Paas, Renkl, & Sweller, 2004). Researchers have suggested that humans can hold only five to nine (7±2) information items in their working memory simultaneously (Miller, 1956). When humans actively work with these chunks of information compared to simply storing them in their working memory, the limit reduces to four (Cowan, 2001). Individuals also require working memory to create schemas for relating new information to existing information (Sweller, 2010; Sweller et al., 2011). In contrast to working memory, individuals' long-term memory has nearly infinite capacity, and individuals need long-term memory to store new information in schemas (Chi, Glaser, & Rees, 1982) that they can use later on for unconscious automation and, thereby, bypass working memory (Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977).

To measure the effect of forwarded email threads on time and decision accuracy, I compared readers of consolidated email threads with readers of forwarded email threads. Forwarded email threads present emails in the form that they emerge during the conversation. Consolidated email threads present the same content, but one can read them from top to bottom without meta-information in between the message segments. For the theoretical grounding, I draw on cognitive load theory in general and the limit in working memory more specifically. I use the split-attention effect, which researchers have proposed based on cognitive load theory and on empirical findings (Chandler & Sweller, 1991, 1992; Sweller et al., 1990; Tarmizi & Sweller, 1988; Ward & Sweller, 1990). This effect suggests that one should present information in an adequate format so that readers do not need to split their attention between multiple sources of information. Researchers have observed the split-attention effect for presentation formats such as text plus diagrams (Chandler & Sweller, 1992) or when switching between a computer monitor and a written piece of paper (Sweller & Chandler, 1994). As with forwarded email threads, these presentation formats require individuals to search and connect information elements rather than focus on the task at hand (Chandler & Sweller, 1991, 1992; Sweller et al., 1990; Tarmizi & Sweller, 1988; Ward & Sweller, 1990). Researchers have also observed this effect specifically for non-integrated text segments (Chandler & Sweller, 1992) as they occur in forwarded email threads. Thus, when dealing with forwarded email threads, I expect working memory to engage in a searching process to find and extract the required information from the multiple, single-text segments such threads include. Therefore, I generally assume that the split-attention effect will occur and propose that:

> **Proposition:** Individuals who read consolidated email threads perform better in reading-related tasks than individuals who read forwarded email threads.

To observe a split-attention effect, researchers have to rely on indirect measures. Prior research has mainly focused on two established objective measures (e.g., Florax & Ploetzner, 2010; Sweller & Chandler, 1994): time used and task accuracy. Cognitive studies have already used the time needed for the task at hand as a dependent variable (Chandler & Sweller, 1991, 1992). Participants in these experiments needed more time to complete the more difficult tasks at hand than the easier ones. This measurement instrument is well established in existing cognitive studies. In terms of email threads, the time used is the reading time. The presence of a split-attention effect causes longer reading times because of the continuous search process in the working memory and the need to integrate the information. Mentally integrating parts of a forwarded email thread also takes longer because one needs to recall existing information into the working memory by rereading passages. Based on this established measurement instrument, I suggest assessing the time participants need to read a forwarded email thread.

> **Hypothesis 1:** Individuals take more time to read forwarded email threads than consolidated email threads.

Cognitive studies have also found support for lower task accuracy for difficult tasks in terms of test scores and errors (Owen & Sweller, 1985; Sweller et al., 1990; Sweller & Cooper, 1985). More difficult tasks lead to an early excess of working memory limits, which consequently results in more errors (Ayres & Sweller, 1990). When one surpasses working memory error levels earlier, one reaches more error levels due to a lack of capacity of their working memory to deal with the task (Ayres, 2001). For forwarded email threads, the working memory is busy with mentally integrating parts of the forwarded email thread. However, the working memory might be necessary to consider all the information and complete the task accurately with a lower error rate.

> **Hypothesis 2:** Individuals that read consolidated email threads achieve more correct answers in reading-related tasks than those who read forwarded email threads.

Furthermore, I argue that there is a speed-accuracy tradeoff between reading time and the number of correct answers. The more time one spends reading the text, the more time one's working memory has to engage in searching. If enough time is available to compensate for the split-attention effect, one can use the additional time to find solutions for the task. However, as individuals are able to substitute speed with accuracy, I cannot conclude that the task with a higher reading time is more difficult than the task with the lower reading time. One needs to consider both speed in terms of reading time and accuracy in terms of number of correct answers (Wickelgren, 1977). Consequently, I can conclude only that the task is more difficult if, in the longer lasting task, the accuracy in relation to time is lower. Hence, I draw on the speed-accuracy literature by including and adapting the response-to-stimulus method (Bogacz, Brown, Moehlis, Holmes, & Cohen, 2006). I use an adapted version of the response-to-stimulus method because time is the dependent variable in this method, which is in line with Hypothesis 1. To draw on SAT insights and to establish a standardized SAT measure that allows comparability among the values of the standardized SAT measures, I divide the number of correct answers by reading time to also reflect Hypothesis 2. The resulting measurement shows the correct answers per second and allows one to capture the dependency between accuracy and speed while implicitly comparing the tasks' difficulty.

> **Hypothesis 3:** Individuals who read consolidated email threads achieve more correct answers per second in reading-related tasks than individuals who read forwarded email threads.

4 Research Methodology

I conducted an online experiment to test the hypotheses. During the experiment, I gave participants three questions they had to answer based on a provided text. I employed a between-subjects design with a manipulation of the provided text format. Each of the participants received the same text either as a consolidated email thread or as a forwarded email thread (cf. Figure 2 for details of the manipulation). I measured the reading time and counted the number of correct answers. I conducted the online experiment in the summer of 2014 in a two-week period.

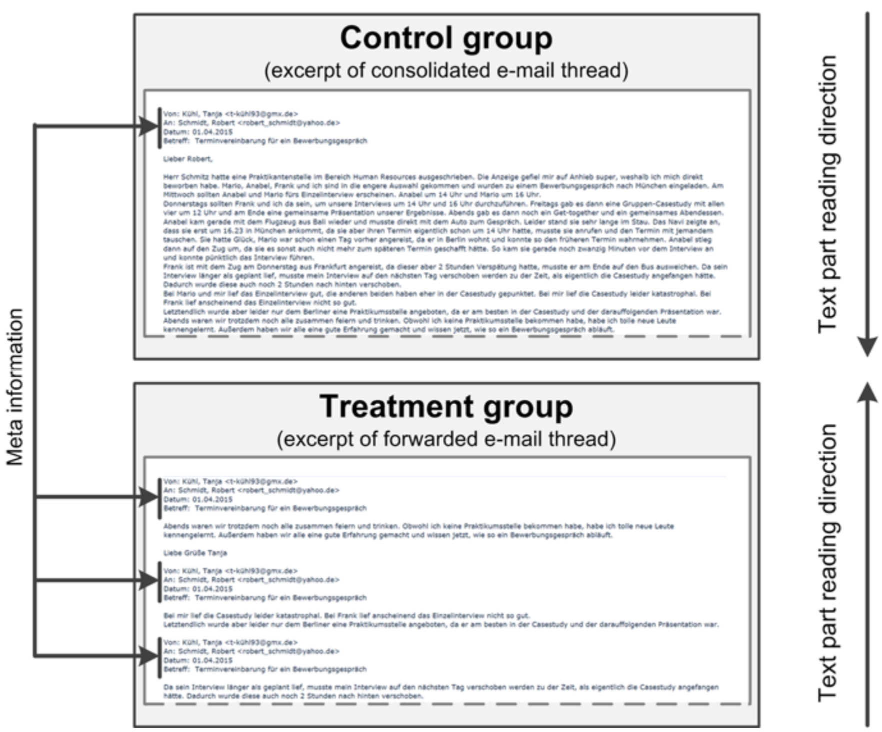

Figure 2. Excerpt of the Manipulation

4.1 Participants

I recruited participants through online platforms such as Facebook or via personal messages such as email. I further encouraged participants to distribute the link for the experiment. In all, 86 students participated in yjr experiment. The participants could participate in a lottery for Amazon vouchers by completing the experiment and answering all the questions correctly. Because only participants who gave correct answers could participate in the lottery, I assume that I sufficiently incentivized and encouraged them to work hard. The participants were invited to a study where they had to answer questions based on a text. I did not mention the terms “email”, “email overload”, or “information overload” in the experiment to mitigate priming effects associated with these terms. The sample included 43 female and 43 male participants. The group who read the consolidated email threads had an average age of 23.91 (SD = 2.07; range = 20-29), and they all checked their emails more than three times a day. The group who read the forwarded email threads had an average age of 23.74 (SD = 2.10; range = 20 - 29), and 92.9 percent of them checked their emails more than three times a day. All participants were either native speakers or fluent in the experiment’s language. Because students do not generally deal extensively with email threads, I consider the student participants as novice email thread readers.

4.2 Experimental Materials

To measure the effect of forwarded email threads, I presented the participants with comprehension tasks that cognitive science studies have used extensively to measure the split-attention effect (e.g., Florax & Ploetzner, 2010; Sweller & Chandler, 1994). To adapt the comprehension tasks to the context of emails threads, I presented the questions to the participants before presenting the text. Prior experiments on the split-attention effect were exam situations; however, in the context of email threads, readers can always reread the email if something is unclear. For the emails’ content, I composed a typical business situation and asked questions concerning the time or the place of a meeting. I asked three questions that one could not answer by light reading the emails because the text implicitly hid the answers. After the participants left the screen with the email text, they could see the posed questions again and 10 possible answers alphabetically ordered in a dropdown. All possible answers were similar to the correct answers but slightly different. I generated the potential answers in such a way that it was not possible to infer the correct answer by looking at how they differed from each other. In my analyses, I included only participants with at least one correct answer because the presence of correct answers was the precondition for assessing correct answers.

I further measured how long the participants took to read the text. The questions appeared before presenting the text so that I could separate the reading time of the text from the reading time of the questions. The reading time started with the appearance of the text and ended when the participants proceeded to the page with the possible answers. I did not inform the participants that I measured how long they took to read the emails, and they had to decide on their own to stop reading. To evaluate the speed-accuracy tradeoff between reading time and number of correct answers, I also assessed the correct answers per second. Therefore, I divided the number of correct answers by the reading time.

The group of consolidated email thread readers received a text structured like a conventional email with meta information about who has written to whom, the salutation, the actual text, and a complimentary close. The group with forwarded email threads received the same text; however, it appeared as a forwarded email thread. Therefore, the email appeared in segments of several sentences that make up one part of the conversation. As in contemporary email applications, the top of the existing history of email messages showed the latest contribution of the conversation (i.e., I reversed the content’s temporal order). In consequence, forwarded email thread readers had to read from bottom to top to make sure that they read the email in the correct temporal order. The segments were split by a block of information such as recipients, sender, time, and date as is common in contemporary email applications. Please see Figure 2 for the manipulation details.

4.3 Experimental Procedure

I designed the experiment with inspiration from the original study of the split-attention effect (Sweller & Chandler, 1994) and adapted it to the email context. I used a between-subjects design in which I manipulated the format of the email threads and presented an email thread either as a consolidated email thread or as a forwarded email thread. The experiment was implemented as an online survey using the software LimeSurvey[1]. Figure 3 outlines the procedure of the experiment.

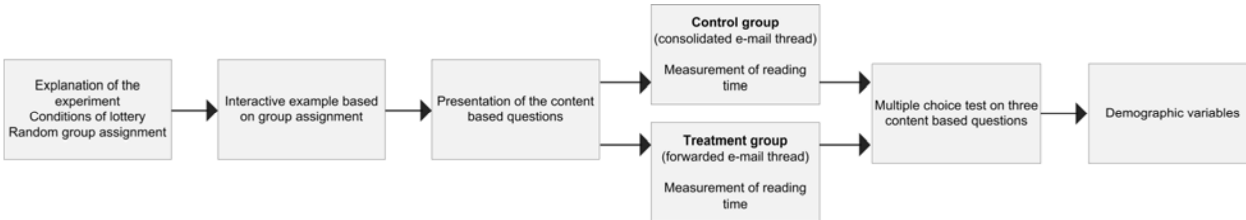

Figure 3. The Applied Procedure to Conduct the Experiment

The experiment started with the instructions for the participants and the conditions of the lottery to incentivize correct answers. The system randomly assigned participants (via random number generation) to a group as the introductory page appeared. After the instructions, an interactive example appeared to help participants learn the experimental task. In this interactive example, one question and a short text appeared with hints regarding the reading directions to ensure that the participants knew the correct reading direction. This example text appeared in the format of a consolidated email thread or a forwarded email thread. The interactive example was compulsory in order to proceed to the actual experiment. In the actual experiment, three questions with one introductory sentence about the context appeared first so that participants could read them beforehand. The next page presented the same questions on top of the page with the email thread underneath. After reading the text, participants had to proceed to the next page that showed the same questions again and also possible solutions. To proceed, the participants had to select one of the solutions for each of the three questions. At the end of the experiment, I collected each participant's demographic information. At no time could the participants go back and see the email thread again.

Before conducting the experiment, I performed a pilot study with students and employees. With the help of the pilot study, I eliminated ambiguous wording, tested the experimental procedures, made instructions clearer and more precise, and determined the experiment's duration.

5 Results

Prior to testing the hypotheses, I examined whether the two experimental groups based on the random assignment belonged to the same population. While the control group held 44 participants, the treatment group had only 42 participants. To rule out alternative explanations of the results, I applied several manipulation checks. Based on the level of measurement, I used the independent samples t-tests for age and information technology knowledge; the Mann-Whitney U test for computer usage duration, Internet usage duration, email knowledge, number of daily email checks, and work experience; and Pearson's chi-square test for gender. I found no significant differences in any of the control variables. Therefore, I do not reject the assumption that my two experimental groups originate from the same population in terms of the tested measures.

Subsequently, I report the results of the three hypotheses[2]. In Section 5.1, I present the results regarding the reading time for different numbers of correct answers. In Section 5.2, I present the results regarding the number of correct answers. In Section 5.3, I present the results regarding correct answers per second.

5.1 Reading Time (H1)

For Hypothesis 1, I investigated the reading time in seconds for different numbers of correct answers. Table 1 highlights the results as the mean and standard deviation (SD) and compares the two groups in terms of p-values and effect sizes (Cohen's d).

[1] https://www.limesurvey.org/

[2] Tests of normality showed a minor violation in one case. The results with parametric and nonparametric tests were essentially similar, and, therefore, we report the results of the parametric tests.

Table 1. Results of Reading Time

Number of correct answers	Consolidated email threads			Forwarded email threads			*p*-value[1]	Effect size[2] (Cohen's d)
	#	Mean	SD	#	Mean	SD		
All (1/2/3)	44	136.91	52.98	42	175.64	52.25	0.001 **	Medium (0.74)
3	28	149.00	47.74	23	181.57	43.76	0.015 *	Medium (0.71)
2	7	151.57	53.61	13	177.46	68.74	0.400	Small (0.42)
1	9	87.89	42.84	6	149.00	39.99	0.016 *	Large (1.47)

[1] P-value: * significant at $p < 0.05$; ** significant at $p \leq 0.01$.
[2] Effect size: small (d = 0.2); medium (d = 0.5); large (d = 0.8).

The results show that participants took longer to perform the comprehension task on forwarded email threads than on consolidated email threads. The overall effect size for all numbers of correct answers was medium, and the effect size for the isolated numbers ranged from small to large. All differences between the two groups were significant except the one with two correct answers, which also exhibited the lowest effect size. The results suggest that individuals take longer to read forwarded email threads than consolidated email threads when performing comprehension tasks. Therefore, the empirical data supports Hypothesis 1.

5.2 Correct Answers (H2)

For Hypothesis 2, I investigated the number of answers that participants answered correctly. Table 2 shows that the means between consolidated email threads (2.43) and forwarded email threads (2.40) differed only by 0.03. Because these results reveal no significant difference in addition to a negligible effect size (0.04), the data does not support Hypothesis 2.

Table 2. Results of Correct Answers

Consolidated email threads			Forwarded email threads			*p*-value[1]	Effect size[2] (Cohen's d)
#	Mean	SD	#	Mean	SD		
44	2.43	0.82	42	2.40	0.73	0.872	< small (0.04)

[1] P-value: * significant at $p < 0.05$; ** significant at $p \leq 0.01$.
[2] Effect size: small (d = 0.2); medium (d = 0.5); large (d = 0.8).

5.3 Correct Answers per Second (H3)

For Hypothesis 3, I investigated the number of correct answers that participants answered per second. Table 3 shows the results. The mean for consolidated email threads (0.0193) was superior to the mean for forwarded email threads (0.0146). This difference is significant at the 0.01 level with a small effect size. The results suggest that the performance in terms of correct answers per second is lower for forwarded email threads than for consolidated email threads and support Hypothesis 3.

Table 3. Results of Correct Answers per Second

Consolidated email threads			Forwarded email threads			*p*-value[1]	Effect size[2] (Cohen's d)
#	Mean	SD	#	Mean	SD		
44	0.0193	0.0078	42	0.0146	0.0057	0.002**	Small (0.42)

[1] P-value: * significant at $p < 0.05$; ** significant at $p \leq 0.01$.
[2] Effect size: small (d = 0.2); medium (d = 0.5); large (d = 0.8).

To recap, I found support for Hypothesis 1 and Hypothesis 3. I did not find support for Hypothesis 2. As such, in summary, the results partially support this paper's proposition.

6 Discussion

I examined the effect of forwarded email threads on email overload. As my theoretical foundation, I used cognitive load theory (Sweller, 2010; Sweller et al., 2011; Sweller et al., 1998; Van Merriënboer & Sweller, 2005, 2010) and drew on the split-attention effect (Chandler & Sweller, 1991, 1992; Sweller et al., 1990; Tarmizi & Sweller, 1988; Ward & Sweller, 1990). I conducted an independent-measures experiment with 86 students in which the participants had to answer comprehension tasks based on an email thread presented either as a forwarded email thread or as a consolidated email thread. I measured reading time, number of correct answers, and the dependency between them with an adapted SAT measure in terms of correct answers per second. I found that forwarded email threads led to longer reading times when participants performed comprehension tasks (i.e., Hypothesis 1 supported). The results highlight that reading time was generally significantly longer, which means that the effect is robust as it concerns the achieved quality of results in terms of the number of correct answers. In contrast, I found that participants did not provide a lower number of correct answers when reading forwarded email threads rather than consolidated email threads (i.e., Hypothesis 2 not supported).

However, I found that performance in terms of correct answers per second was lower with forwarded email threads (i.e., Hypothesis 3 supported). This finding indicates the difficulty of reading forwarded email threads compared with consolidated email threads. The support I found for Hypothesis 1 and Hypothesis 3 reveals the participants' inherent speed-accuracy tradeoff. Since my experiment was a comprehension task in which the participants knew the questions before reading the text and since they had no constraints regarding the reading time, the participants could freely decide their personal level of speed and accuracy. I manipulated neither speed nor accuracy externally, and, therefore, the participants were free to choose between speed and accuracy to achieve the highest reward (Edwards, 1965; Gold & Shadlen, 2002). My lottery supposedly led the participants to substitute speed for accuracy to achieve a high number of correct answers. Therefore, this result supports the findings for Hypothesis 1 and explains the respective results for Hypothesis 2. The results suggests that individuals compensate for the difficulty of the forwarded email thread format by reading longer. Furthermore, the number of correct answers does not reflect this difficulty, which shows that participants substitute speed for accuracy.

With these findings, I contribute to the literature in two ways. First, I empirically investigated and found support for the notion that the format of forwarded email threads is an antecedent of email overload. The format of forwarded email threads leads to longer reading times, and the decreased SAT measure indicates that comprehension tasks are more difficult when information appears as a forwarded email thread. Therefore, I contribute to the list of isolated antecedents such as the number of emails (e.g., Dabbish & Kraut, 2006), time spent (e.g., Sumecki et al., 2011), and email-based interruptions (e.g., Gupta et al., 2013; Renaud et al., 2006), which researchers have also investigated as isolated antecedents. Furthermore, these factors are congruent with the findings from an information-overload perspective in terms of the antecedents amount of information (Hiltz & Turoff, 1985; Nelson, 1994) and available time (Schick et al., 1990). Beyond these already-known antecedents of information overload, such as the quality of information (Slawson, Shaughnessy, & Bennett, 1994), we need to know how the information appears. Although existing research on this is scarce, this study confirms that we need research on coping software that addresses the consolidation of forwarded email threads. We also need research that focuses on presenting emails in a spatially integrated format.

Second, the findings confirm that time spent (Schick et al., 1990; Sumecki et al., 2011) on emails is an important predictor of email overload. However, one needs to keep in mind that an individual can trade time for accuracy depending on the individual's external constraints and the reward situation. As such, time spent solely on emails might not be the strongest predictor for email overload. However, I not only identified findings that are congruent with existing research but also found support for the notion that the number of emails is not an "almighty" predictor for email overload: the same email, including the same content, can lead to a longer reading time. Despite the fact that the number of emails is a strong predictor, it does not reflect my insights into the format of email threads and the format of emails generally and demands further conceptualizations of email overload beyond their number. By counting the number of emails, one neglects their different formats, which may mislead one's thinking, especially when designing mitigation tools and strategies.

I encourage researchers to intensify my experiment on emails' format. In addition, researchers can examine task types other than comprehension tasks because individuals use email for a variety of different tasks (Whittaker & Sidner, 1996). Besides investigating the email thread format with the split-attention effect, I suggest elaborating the impact of further cognitive effects on email such as the redundancy effect (Chandler & Sweller, 1991) to more thoroughly understand the emergence of email overload. Finally, one might consider using further measurement methods of cognitive load to develop a new conceptualization of email overload (Sweller, 2010; Sweller et al., 2011, 1998).

Our study also provides useful implications for email use in practice. First, one should avoid forwarded email threads in general when one needs to quickly understand the content because they lead to longer reading times. When one cannot avoid them, one should summarize the necessary information when forwarding email threads to further parties. This procedure prevents recipients from needing to read the whole email thread again. From a technical perspective, I suggest using features in email applications that allow one to view emails in a spatially integrated format to avoid the emergence of a split-attention effect. These features should also include the possibility of reformatting email threads and disabling the presentation of meta-information in between the message segments. As a last practical implication, I suggest that we need to design and use coping tools that merge email messages into a spatially integrated format to help individuals read forwarded emails threads.

Our paper has several limitations. I used students as participants, who are inexperienced with business email threads, in my experiment. I set up the experiment's incentives up to encourage a high number of correct answers; however, the results might be different when incentivizing a low reading time. In addition, I only investigated one text per person and only one type of task. Therefore, I cannot generalize beyond this setting. However, existing cognitive research has experienced the same limitations. In addition, I conducted my experiment online, so I could not control for a variety of factors, but I am confident that I isolated the essential effect.

7 Conclusion

Existing research focuses on number of emails as an antecedent of email overload and neglects to investigate the format of each email. I conducted an experiment exposing participants to the same content either as a consolidated email thread or as a forwarded email thread. The experiment resulted in an increased reading time and lower correct answers per second in comprehension tasks for email thread readers. Based on cognitive load theory and the split-attention effect, these results suggest that the format of forwarded email threads also contributes significantly to email overload. With these findings, I contribute to identifying email threads as a potential antecedent of email overload and the fact that emails' format matters in addition to their number. These results serve as a theoretical foundation to investigate further coping instruments.

References

Ayres, P., & Sweller, J. (1990). Locus of difficulty in multistage mathematics problems. *The American Journal of Psychology,* 103(2), 167-193.

Ayres, P. L. (2001). Systematic mathematical errors and cognitive load. *Contemporary Educational Psychology, 26*(2), 227-248.

Bälter, O., & Sidner, C. (2002). Bifrost inbox organizer: Giving users control over the inbox. In *Proceedings of 2nd Nordic Conference on Human-Computer Interaction* (pp. 111-118).

Barley, S. R., Meyerson, D. E., & Grodal, S. (2011). Email as a source and symbol of stress. *Organization Science, 22*(4), 887-906.

Bawden, D. (2001). Information overload. *Library & Information Briefings, 92*, 1-15.

Beaudoin, C. E. (2008). Explaining the relationship between Internet use and interpersonal trust: Taking into account motivation and information overload. *Journal of Computer-Mediated Communication, 13*(3), 550-568.

Berghel, H. (1997). Cyberspace 2000: Dealing with information overload. *Communications of the ACM, 40*(2), 19-24.

Black, S. D., Levin, J. A., Mehan, H., & Quinn, C. N. (1983). Real and non real time interaction: Unraveling multiple threads of discourse. *Discourse Processes, 6*(1), 59-75.

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review, 113*(4), 700-765.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction, 8*(4), 293-332.

Chandler, P., & Sweller, J. (1992). The split attention effect as a factor in the design of instruction. *British Journal of Educational Psychology, 62*(2), 233-246.

Chi, M. T. H., Glaser, R., & Rees, R. (1982). Expertise in problem solving. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (vol. 1, pp. 7-76). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*(1), 87-185.

Dabbish, L., & Kraut, R. (2006). Email overload at work: An analysis of factors associated with email strain. In *Proceedings of the Conference on Computer Supported Cooperative Work* (pp. 431-440).

Dehghani, M., Shakery, A., Asadpour, M., & Koushkestani, A. (2013). A learning approach for email conversation thread reconstruction. *Journal of Information Science, 39*(6), 846-863.

Ducheneaut, N., & Bellotti, V. (2001). Email as habitat: An exploration of embedded personal information management. *Interactions, 8*(5), 30-38.

Edwards, W. (1965). Optimal strategies for seeking information: Models for statistics, choice reaction times, and human information processing. *Journal of Mathematical Psychology, 2*(2), 312-329.

Eppler, M. J., & Mengis, J. (2004). The concept of information overload: A review of literature from organization science, accounting, marketing, MIS, and related disciplines. *Information Society, 20*(5), 325-344.

Erera, S., & Carmel, D. (2008). Conversation detection in email systems. In *Proceedings of the European Conference on Advances in Information Retrieval* (pp. 498-505). Berlin, Heidelberg: Springer.

Fisher, D., Brush, A. J., Gleave, E., & Smith, M. A. (2006). Revisiting Whittaker & Sidner's "email overload" ten years later. In *Proceedings of the Conference on Computer Supported Cooperative Work* (pp. 309-312).

Florax, M., & Ploetzner, R. (2010). What contributes to the split-attention effect? The role of text segmentation, picture labelling, and spatial proximity. *Learning and Instruction, 20*(3), 216-224.

Gill, S. (1998). UK business information on the Internet. *New Library World, 90*(6), 238-242.

Gold, J. I., & Shadlen, M. N. (2002). Banburismus and the brain: Decoding the relationship between sensory stimuli, decisions, and reward. *Neuron, 36*(2), 299-308.

Grevet, C., Choi, D., Kumar, D., & Gilbert, E. (2014). Overload is overloaded: Email in the age of Gmail. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 793-802).

Gupta, A., Li, H., & Sharda, R. (2013). Should I send this message? Understanding the impact of interruptions, social hierarchy and perceived task complexity on user performance and perceived workload. *Decision Support Systems, 55*(1), 135-145.

Gupta, A., Sharda, R., & Greve, R. A. (2011). You've got email! Does it really matter to process emails now or later? *Information Systems Frontiers, 13*(5), 637-653.

Hashem, M. I. (2014). Improvement of email summarization using statistical based method. *International Journal of Computer Science and Mobile Computing, 3*(2), 382-388.

Heylighen, F. (2004). Complexity and information overload in society: Why increasing efficiency leads to decreasing control. *The Information Society, 87*, 1-19.

Hiltz, S. R., & Turoff, M. (1985). Structuring computer-mediated communication systems to avoid information overload. *Communications of the ACM, 28*(7), 680-689.

Ingham, J. (2003). E-mail overload in the UK workplace. *Aslib Proceedings, 55*(3), 166-180.

Jackson, T. W., & Farzaneh, P. (2004). Theory-based model of factors affecting information overload. *International Journal of Information Management, 32*(6), 523-532.

Kirsh, D. (2000). A few thoughts on cognitive overload. *Intellectica, 30*(1), 19-51.

Kleinberg, J. (2003). Bursty and hierarchical structure in streams. *Data Mining and Knowledge Discovery, 7*(4), 373-397.

Klimt, B., & Yang, Y. (2004). The Enron corpus: A new dataset for email classification research. *Machine Learning: ECML 2004 Proceedings* (pp. 217-226).

Kock, N. (2001a). Asynchronous and distributed process improvement: The role of collaborative technologies. *Information Systems Journal for East European Management Studies, 11*(2), 87-110.

Kock, N. (2001b). Compensatory adaptation to a lean medium: An action research investigation of electronic communication in process improvement groups. *IEEE Transactions on Professional Communication, 44*(4), 267-285.

Kock, N. (2004). The psychobiological model: Towards a new theory of computer-mediated communication based on Darwinian evolution. *Organization Science, 15*(3), 327-348.

Kock, N. (2005). Media richness or media naturalness? The evolution of our biological communication apparatus and its influence on our behavior toward e-communication tools. *IEEE Transactions on Professional Communication, 48*(2), 117-130.

Kock, N. (2007). Media naturalness and compensatory encoding: The burden of electronic media obstacles is on senders. *Decision Support Systems, 44*(1), 175-187.

Kock, N. (2009). Information systems theorizing based on evolutionary psychology: An interdisciplinary review and theory integration framework. *MIS Quarterly, 33*(2), 395-418.

Lewis, D. D., & Knowles, K. A. (1997). Threading electronic mail: A preliminary study. *Information Processing & Management, 33*(2), 209-217.

Malhotra, N. K., Jain, A. K., & Lagakos, S. W. (1982). The information overload controversy: An alternative viewpoint. *Journal of Marketing, 46*(2), 27-37.

Mark, G. J., Voida, S., & Cardello, A. v. (2012). “A pace not dictated by electrons”: An empirical study of work without email. In *CHI 2012 Proceedings* (pp. 555-564).

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *The Psychological Review, 63*(2), 81-97.

Minkov, E., Cohen, W. W., & Ng, A. Y. (2006). Contextual search and name disambiguation in email using graphs. In *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development on Information Retrieval*.

Murray, D. E. (2000). Protean communication: The language of computer-mediated communication. *TESOL Quarterly, 34*(3), 397-421.

Nelson, M. R. (1994). We have the information you want, but getting it will cost you!: held hostage by information overload. *Crossroads, 1*(1), 11-15.

Newman, P. S. (2002). Exploring discussion lists: Steps and directions. In *Proceedings of the 2nd ACM/IEEE-CS Joint Conference on Digital Libraries*.

Owen, E., & Sweller, J. (1985). What do students learn while solving mathematics problems? *Journal of Educational Psychology, 77*(3), 272-284.

Paas, F., Renkl, A., & Sweller, J. (2004). Cognitive load theory: Instructional implications of the interaction between information structures and cognitive architecture. *Instructional Science, 32*(1-2), 1-8.

Popolov, D., Callaghan, M., & Luker, P. (2000). Conversation space: Visualising multi-threaded conversation. In *Proceedings of the Working Conference on Advanced Visual Interfaces*.

Reinke, K., & Chamorro-Premuzic, T. (2014). When email use gets out of control: Understanding the relationship between personality and email overload and their impact on burnout and work engagement. *Computers in Human Behavior, 36*, 502-509.

Renaud, K., Ramsay, J., & Hair, M. (2006). “You've got e-mail!” …Shall I deal with it now? Electronic mail from the recipient's perspective. *International Journal of Human-Computer Interaction, 21*(3), 313-332.

Rohall, S. L., Gruen, D., Moody, P., & Kellerman, S. (2001). Email visualizations to aid communications. In *Proceedings of the IEEE Symposium on Information Visualization* (p. 12-15).

Savolainen, R. (2007). Filtering and withdrawing: Strategies for coping with information overload in everyday contexts. *Journal of Information Science, 33*(5), 611-621.

Schick, A. G., Gordon, L. A., & Haka, S. (1990). Information overload: A temporal approach. *Accounting, Organizations & Society, 15*(3), 199-220.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84*(1), 1-66.

Schroder, H. M., Driver, M. J., & Streufert, S. (1967). *Human information processing: Individuals and groups functioning in complex social situations*. New York: Holt, Rinehart and Winston.

Schuff, D., Turetken, O., & D’Arcy, J. (2006). A multi-attribute, multi-weight clustering approach to managing “e-mail overload”. *Decision Support Systems, 42*(3), 1350-1365.

Schultz, U., & Vandenbosch, B. (1998). Information overload in a groupware environment: Now you see it, now you don’t. *Journal of Organizational Computing & Electronic Commerce, 8*(2), 127-148.

Sharaff, A., & Nagwani, N. K. (2015). Email thread identification using latent Dirichlet allocation and non-negative matrix factorization based clustering techniques. *Journal of Information Science*.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review, 84*(2), 127-190.

Slawson, D., Shaughnessy, A., & Bennett, J. (1994). Becoming a medical information master: Feeling good about not knowing everything. *The Journal of Family Practice, 38*(5), 505-513.

Sumecki, D., Chipulu, M., & Ojiako, U. (2011). Email overload: Exploring the moderating role of the perception of email as a “business critical” tool. *International Journal of Information Management, 31*(5), 407-414.

Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review, 22*(2), 123-138.

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. New York, NY: Springer.

Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction, 12*(3), 185-233.

Sweller, J., Chandler, P., Tierney, P., & Cooper, M. (1990). Cognitive load as a factor in the structuring of technical material. *Journal of Experimental Psychology: General, 119*(2), 176-192.

Sweller, J., & Cooper, G. A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction, 2*(1), 59-89.

Sweller, J., van Merrienboer, J. G., & Paas, F. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*(3), 251-296.

Szóstek, A. M. (2011). "Dealing with my emails": Latent user needs in email management. *Computers in Human Behavior, 27*(2), 723-729.

Tarmizi, R. A., & Sweller, J. (1988). Guidance during mathematical problem solving. *Journal of Educational Psychology, 80*(4), 424-436.

Van Merriënboer, J. J. G., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review, 17*(2), 147-177.

Van Merriënboer, J. J. G., & Sweller, J. (2010). Cognitive load theory in health professional education: Design principles and strategies. *Medical Education, 44*(1), 85-93.

Venolia, G. D., & Neustaedter, C. (2003). Understanding sequence and reply relationships within email conversations: A mixed-model visualization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*.

Vidgen, R., Sims, J., & Powell, P. (2011). Understanding e-mail overload. *Journal of Communication Management, 15*(1), 84-98.

Vollmann, T. E. (1991). Cutting the Gordian knot of misguided performance measurement. *Industrial Management & Data Systems, 91*(1), 24-26.

Ward, M., & Sweller, J. (1990). Structuring effective worked examples. *Cognition and Instruction, 7*(1), 1-39.

Whittaker, S., & Sidner, C. (1996). Email overload: Exploring personal information management of email. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 276-283).

Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica, 41*(1), 67-85.

Yeh, J.-Y. (2006). Email thread reassembly using similarity matching. In *Proceedings of the Conference on Email and Anti-Spam.*

About the Author

Nikolai Sobotta is a doctoral candidate at the Department of Information Systems Engineering, Goethe-University of Frankfurt, Germany. Nikolai earned a Master of Science degree from the University of Hohenheim, Germany. His research focuses on cognitive perspectives of email, information overload, and computer based communication.