How E-Mail Threads Contribute to E-Mail Overload: Investigating Intrinsic, Extraneous, and Germene Cognitive Load

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

Abstract

Information overload caused by e-mails is a known issue in practice and academia. Research often investigates the number of e-mails and not the cognitive impact of the e-mail channel. Therefore, it is necessary to know how much the format of a single e-mail contributes to the overall phenomenon. As an example, we examine one of the most promising formats: e-mail threads as a history of appended e-mail quotes that are forwarded to a third person. Cognitive load theory is proposed to investigate the extent to which the format of e-mail threads contributes to intrinsic, extraneous, and germane cognitive load. The resulting hypotheses and measurement scales help to understand the induction of e-mail overload by e-mail threads through different types of cognitive load. The scales are tested for construct validity based on an item-sort task. The results show high values for the substantive validities of the constructs, indicating good construct validity.

1. Introduction

New technologies [1], but also existing technologies such as electronic mails (e-mail), expose humans to stress [2]. One strong predictor for stress is information overload (IO) [3]. IO is an ubiquitous phenomenon [4] and has been characterized by many researchers [e.g., 5, 6] as an “excessive supply” [7] of information. Beside this general perspective, it has been also conceptualized for different technologies like social network sites [8], online interaction spaces [9], and e-mails [10, 11]. Scholars often rely on the sheer amount of information that is generated by a certain technology to predict IO as for example the number of postings in newsgroups [9] or the amount of social information in social network sites [8]. For the e-mail technology, researchers investigate the phenomenon mainly from the perspective of the volume of e-mails and time spent for these e-mails [e.g., 10, 11]. In the remainder of this paper, IO caused by e-mails is called e-mail overload [10].

Previous research on IO and e-mail overload has in common that it looks at the total number as an independent variable and how this is linked to the feeling of being overloaded or at perceived IO as a dependent variable. This approach treats the emergence of e-mail overload as a black box, as it disregards the format of a single e-mail. E-mail is used for many different tasks which results in different e-mail types in terms of formats [12]. In current research, they are not differentiated and its impact on perceived IO is neglected. Therefore, we aim to extend existing work by not looking at the number of e-mails; instead we look at the format of a single mail and its contribution to IO. As an example, we examine one of the most promising e-mail specific formats which annoys many e-mail users: e-mail threads. This is how we call the history of e-mail quotes sent back and forth between two or more recipients. They are characterised by answering messages always on top of the already existing history of messages. We do not look at the emergence of e-mail threads; in fact we look at the processing of all quotes at once when such threads are forwarded to new communication participants. We investigate this e-mail format because the increasing division of labour and the needed information exchange promotes the forwarding of e-mail threads and the high numbers of e-mails that are sent back and forth each day are one of the major causes for IO in organizations [13].

The purpose of this study is to provide a first building block for a novel theoretical perspective on e-mail overload and open the black box of how certain formats of e-mail contribute to the e-mail overload phenomenon. As existing e-mail overload measurements fail in assessing the perception beyond the number of e-mails we develop and test new measurement scales that are unique to determine the impact of the format of e-mail threads on e-mail overload. We propose to scrutinize cognitive load theory [14-18] in order to investigate the extent to which different formats impose cognitive load. The measurement of different types of cognitive load imposed by e-mail threads draws on insights of
cognitive science. We investigate intrinsic cognitive load for the content, extraneous cognitive load for the format of e-mail threads, and germane cognitive load for learning to cope with the format of e-mail threads. These insights aim for optimal presentation of information in order to capture complex problems and situations [15]. Therefore, extraneous cognitive load should be kept low in order to keep residual cognitive capacity high. We therefore ask the following research question: “What are valid scales for measuring the impact of e-mail threads on intrinsic, extraneous, and germane cognitive load?”

The main contribution of this paper is twofold. First, we integrate existing literature and propose cognitive load theory to investigate the extent to which e-mail threads contribute to the e-mail overload phenomenon. Second, we report on the development of valid measurement scales of a psychometric instrument which can differentiate intrinsic, extraneous, and germane cognitive load. We design these scales to measure the extent of extraneous cognitive load imposed by e-mail threads contributing to e-mail overload. We test our proposed scales for construct validity in a pre-test assessment that is based on an item-sort task. As we concentrate on the measurement scale development and due to the limited data inherent in an item-sort task, no data is provided to investigate whether the hypotheses can be supported.

In the next section, we discuss related work on e-mail overload. Consequently, we present cognitive load theory as the foundation and develop our hypotheses. Afterwards, the measurement scale development and in Section 5 the pre-test assessment is outlined. Finally we discuss the results and present the implications of our findings, give an outlook on further research, and conclude with a summary.

2. Related research

The term e-mail overload has two meanings that are used in research [19]: First, it deals with the high number and different types of e-mails stored in the inbox, whereas the other meaning is concerned with humans dealing with the high number of incoming e-mails. E-mail overload was observed for the first time when researchers investigated the phenomenon of high number of e-mails in the inbox and the effect that users employ e-mail for other tasks than for conversations [12]. Researchers suggested technological solutions to mitigate these problems, such as classification of e-mails into predefined folders and auto archiving [e.g., 20, 21]. Some of these features are now also evident in popular e-mail programs [10]. Surprisingly, to all appearances, this does not mitigate the problem because ten years later, researchers reinvestigated this phenomenon and realized that basically nothing had changed and e-mail overload still is a problem [22]. This statement is still valid today as recent studies demonstrated [19, 23]. In the meanwhile, the second meaning of e-mail overload emerged which deals with the high volume of incoming messages which is also still an issue today [19].

The second meaning of e-mail overload emerged from the research stream of IO [10, 11]. This meaning is the one that is used for the remainder of this paper because it incorporates the human component and it is more suitable to answer our research question because we investigate different types of cognitive loads. Scholars [e.g., 5, 10] often look on IO by investigating the “excessive supply” [7] of information so that a human cannot cope with it anymore. Computer-mediated communication in organizations has increased the amount of information [24] and especially e-mails sent back and forth lead to a dramatic increase of available information [25]. Based on the limited cognitive capacity of humans [26-28], scholars [e.g., 10, 11, 29, 30] agree that e-mail overload occurs when “email use overwhelms the user” [10]. These studies look at the phenomenon by investigating the number of e-mails sent and received, the amount of time spent with reading or writing e-mails, and humans’ perceptions about these amounts [10]. However, different studies highlight different perspectives, building upon the first work on the high number of incoming e-mails [11]. The perception of e-mail as a business critical tool [10] or the influence of the users’ personality [31] are two extending perspectives. Furthermore, it is also considered how often humans should read their e-mails and how often they are interrupted by e-mails [32-35]. Most research in this stream has focused on perceived IO as well as the effect of the total amount of its antecedents. However, none of them tried to determine the extent to which certain formats of e-mail contribute to the phenomenon of e-mail overload.

3. Theoretical background and hypothesis development

Researchers recognized early that the overuse of information technology leads to cognitive overload [36]. More specifically, electronic communication leads to a higher degree of ambiguity than face-to-face communication and evokes additional cognitive load [37-39]. The needed encoding and decoding of electronic messages in order to compensate for missing stimuli induces additional cognitive load [40-42]. We use a cognitive perspective, drawing on cognitive load theory, in order to investigate the extent to which e-mail threads contribute to e-mail overload by
measuring intrinsic, extraneous, and germane cognitive load of decoding e-mail threads. The insights on limited processing capacity of humans [26-28] are built to a large degree on insights from cognitive load theory [14-18].

The central argument of this theory is the limitation of the human cognitive architecture and especially the working memory. Cognitive load theory distinguishes between two components of the human cognitive architecture: (1) working memory and (2) long-term memory. Working memory is conscious [16] and limited in the number of information chunks it is able to process [43]. A maximum of five to nine chunks of information can be held simultaneously [44]. By dealing with these chunks of information, the number is even reduced to four [45]. Working memory is also used to create new schemas which relate new information to existing information [14, 15]. Long-term memory is used to store schemas and features a virtually unlimited capacity [46]. Acquired schemas can be used for unconscious automation bypassing working memory [47, 48].

Cognitive load resides in working memory and according to its causes, there are three types of cognitive load distinguished in the literature [15]: (1) intrinsic cognitive load which is determined by the natural complexity of information, (2) extraneous cognitive load which arises from suboptimal presentation, and (3) germane cognitive load which is invoked from relating new information elements to already existing information and storing these schemas in long-term memory.

Based on the central argument of cognitive load theory, most contemporary cognitive research sees working memory as the limiting resource that leads to a limited information processing capacity [49, 50]. Conversely, we suggest that shortage in working memory is supposedly the cognitive cause for the bottleneck in processing e-mails that is generally called e-mail overload. Therefore, we develop measurement scales to investigate the impact of e-mail threads on intrinsic, extraneous, and germane cognitive load. As an essential step, we first develop our underlying hypotheses regarding these measures to determine the reference points of the cognitive load measures (c.f. Section 4 - “Measurement Scale Development” for further details). The hypotheses and the reference points are necessary to validate the scales appropriately. The limited data collection of our pretest does not allow for hypothesis testing which is not the focus of this article.

Intrinsic cognitive load [14-16] cannot be changed by any intervention and solely relies on the complexity of embedded information elements [15, 51]. The information elements are closely related and may influence each other by interacting [51]. Previous literature found that the more interactive elements have to be stored, the higher is the cognitive load [52]. The intrinsic cognitive load of an e-mail is evoked by the information elements that are encapsulated into the content of the e-mail message. By reading the text and realizing the content of the text that is embedded in the e-mail message, intrinsic cognitive load is evoked. Due to the nature of intrinsic cognitive load, it is not possible to change this type of cognitive load without changing the content of the e-mail message. For our article, we compare e-mail threads to conventional e-mails containing the same text in order to investigate the format of e-mail threads. As for our investigation, e-mail threads have the same embedded text and cover the same information as conventional e-mails, only the presentation format as an e-mail thread is different. The number of information elements associated with the content of the message or the information elements itself are not changed. As the same information elements are present for readers of e-mail threads and for conventional readers, we propose the following manipulation check: Readers of e-mail threads are exposed to the same intrinsic cognitive load than those readers of conventional e-mails.

Extraneous cognitive load [14-16] is solely concerned with the presentation of the information, not with the information itself. Low levels of extraneous cognitive load have been shown to affect understanding positively because an optimal presentation format is used [51, 53]. Amongst others, specifically one relevant effect induces extraneous load: split attention is a presentation format where extraneous load is increased by a not physically integrated presentation. This has been observed for text plus diagrams [54] or computer and paper [51]. These additional interacting elements cause parts of the working memory to engage in searching rather than in dealing with the task [53-57]. Investigating the presentation of e-mail threads, which is a format where each sender replies to the existing history of messages and appends the answer message to the top or bottom, we have a physically not integrated format of text blocks. The split attention effect has also been observed for not integrated text blocks [54]. Therefore, we suppose to observe a split attention effect for e-mail threads because by reading the whole history of messages, the readers’ working memory needs to engage in a search process to integrate the embedded information through several message answers. Therefore we posit the following hypothesis:

Hypothesis 1: Readers of e-mail threads are exposed to a higher extraneous cognitive load than those readers of conventional e-mails.
**Germane cognitive load** is crucial for the construction of schemas and therefore for learning [14, 58]. Humans store new information related to other already existing information in schemas [46]. Appropriate instructional designs such as asking questions about the examples may induce the construction of schemas [16]. Despite the importance of germane cognitive load, cognitive resources are first used for intrinsic and extraneous cognitive load and germane cognitive load is only triggered when there is enough residual cognitive capacity remaining [16-18].

For e-mail threads, we investigate the induction of germane cognitive load that is induced by learning to cope with the format of e-mail threads. Readers of e-mail threads are exposed to the structural outline of e-mail threads and cognitive load is imposed to their working memory. In comparison to readers of conventional e-mails who have no e-mail threads, only readers of e-mail threads experience the structural outline of e-mail threads. Only by learning to cope with e-mail threads, schemas are constructed and germane cognitive load is imposed:

**Hypothesis 2:** Readers of e-mail threads are exposed to a higher germane cognitive load than those readers of conventional e-mails.

In order to find support for our hypothesis, we adapted the cognitive load types to match our research domain. The reference point of the cognitive load definitions was adapted from the general learning context to the specific e-mail thread context presented in Table 1.

**Table 1. Constructs and their descriptions**

<table>
<thead>
<tr>
<th>Construct</th>
<th>Definition and description</th>
</tr>
</thead>
<tbody>
<tr>
<td>E-mail intrinsic cognitive load</td>
<td>Intrinsic cognitive load that is determined by the complexity of the content of the e-mail text.</td>
</tr>
<tr>
<td>E-mail extraneous cognitive load</td>
<td>Extraneous cognitive load arising from the suboptimal presentation of the content of an e-mail in terms of the structural outline.</td>
</tr>
<tr>
<td>E-mail germane cognitive load</td>
<td>Germane cognitive load that is invoked by learning to cope with the structural outline of an e-mail.</td>
</tr>
</tbody>
</table>

**4. Measurement scale development**

For our next step, we developed measurement scales for our constructs [59] based on our hypotheses. To find appropriate measurement scales, we conducted a content validity assessment on the basis of the construct definitions. We set out to find and adopt existing, established measurement scales from prior literature, if possible [60]. Unfortunately, no adequate scales that measure our constructs were found, although cognitive load in general can be measured by different methods [61]. Subjective rating scales have been used quite intensively in existing cognitive research to measure cognitive load [see 61, 62 for an extensive discussion]. Based on early works [63], rating scales for mental effort are unidimensional scales [e.g., nine-point unidimensional scale, 64], where learners introspect the amount of mental effort they invested during learning [64]. The same is also valid for unidimensional scales assessing how difficult or easy a learning task was perceived [65]. These unidimensional scales have in common that they cannot assess intrinsic, extraneous, and germane cognitive load, only cognitive load in general. In terms of multidimensional scale, the NASA-TLX scale [66] is assessing the related concept of workload. But the exact relation to cognitive load is not investigated. Another multidimensional scale is based on three levels (high, medium, and low) of three subscales resulting into 21 gradations [67, 68]. Also none of them can distinguish intrinsic, extraneous, and germane cognitive load.

Several studies attempted to develop measurement instruments to differentiate intrinsic, extraneous, and germane cognitive load [69-73]. However, all of them measure at least one of the cognitive load types only with one item, but several items are desirable [74]. A recent measurement instrument overcomes this lack by distinguishing three different types of cognitive load as latent variables: intrinsic, extraneous, and germane cognitive load [75, 76]. However, these measurement instruments are limited to the context of statistics and language in terms of learning instructions. These scales are designed with the purpose of optimizing learning and the needed instructions. The measures aim to capture the three cognitive load types in a sense to include all imposed cognitive loads through learning instructions.

We built on these recent measurement scales and developed new scales meeting our requirements specified by the hypotheses in terms of the domain, specifically to the used reference points: the content and the structural outline of e-mails. As our constructs are related but distinct and close to each other, new items have been developed keeping the basic idea of the existing items in mind in order to improve content validity. Also, one of the existing items on mental effort encountered an issue with internal consistency and as we had to redevelop all of the items, we conducted a pre-test assessment in terms of an item-sort task. Table 2 contains all developed measurements scales and their items.
We performed a two stage process in order to develop our measurement items. First, we adapted the existing scales to our domain and for our purpose [75, 76]. As our hypotheses and the measurement of different types of cognitive load require topic specific constructs, we had to adapt the reference point of the items. For intrinsic cognitive load, we adapted the reference point to the content of the texts in order to measure how difficult the participants perceive the complexity of the text. Therefore, we dropped one of the items because it was based on the complexity of the terms and not on the content of the text. Then we adapted the items for extraneous cognitive load in order to measure only the structural outline in terms of e-mail threads. We also included the explicit note in the items that the structure is not related to the content of the text in order to avoid interferences and distinguish between intrinsic and extraneous cognitive load. Also, for germane cognitive load, we adapted the items in terms of how much load is imposed by learning to cope with the structural outline and added the explicit note that the structural outline is not content related. Extraneous and germane cognitive load refer both to the structural outline. Therefore we sharpened the wording of both of the constructs in order to make the concepts clearer, more precise and avoid overlapping.

Second, to increase content validity, we developed new items based on the wording and the sentence structure from existing items [75, 76]. Two new items were developed respectively for intrinsic and germane cognitive load and three new items for extraneous cognitive load.

We evaluated our constructs in terms of whether they are formative or reflective on the basis of established guidelines [77, 78]. Based on our construct definitions, we came to the conclusion that the constructs are reflective because the constructs are intended to measure one type of cognitive load each. Each type of cognitive load contains exactly one reference point, content for intrinsic and the structural outline for extraneous and germane cognitive load. These reference points are reflected in all the items of the respective constructs, thus we modelled them as reflective. A formative measurement model could not be observed. Based on our hypotheses and our theoretical argument, no distinguishable subcomponents could be inferred forming the respective constructs.

### 5. Pretest

The development of the scales to our research domain and also to the different reference points in terms of the content and structural outline of the text

---

**Table 2. Overview of constructs and their items**

<table>
<thead>
<tr>
<th>Construct</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>E-mail intrinsic cognitive load (reflective)</td>
<td>The content of this text was complex.</td>
</tr>
<tr>
<td></td>
<td>The storyline covered in this text was complex.</td>
</tr>
<tr>
<td></td>
<td>I invested a high mental effort in the complexity of this text's content.</td>
</tr>
<tr>
<td></td>
<td>The content of this text was hard to understand.</td>
</tr>
<tr>
<td></td>
<td>I had difficulties to keep up with the text's content.</td>
</tr>
<tr>
<td>E-mail extraneous cognitive load (reflective)</td>
<td>The text's structural outline (not content-related) was unclear.</td>
</tr>
<tr>
<td></td>
<td>The text's structural outline (not content-related) was confusing.</td>
</tr>
<tr>
<td></td>
<td>The text's structural outline (not content-related) was inefficient for grasping the content.</td>
</tr>
<tr>
<td></td>
<td>I invested a high mental effort in the text's inefficient structural outline (not content-related).</td>
</tr>
<tr>
<td></td>
<td>The text's structural outline (not content-related) impeded my understanding of the text.</td>
</tr>
<tr>
<td></td>
<td>The text's structural outline (not content-related) was burdensome to me.</td>
</tr>
<tr>
<td>E-mail germane cognitive load (reflective)</td>
<td>Reading the text enhanced my understanding of its structural outline (not content-related).</td>
</tr>
<tr>
<td></td>
<td>Reading the text enhanced my ability to cope with this type of structural outline (not content-related).</td>
</tr>
<tr>
<td></td>
<td>Reading the text enhanced my experience regarding this type of structural outline (not content-related).</td>
</tr>
<tr>
<td></td>
<td>Reading the text helped me to deal with texts with a similar structural outline (not content-related).</td>
</tr>
<tr>
<td></td>
<td>Reading this text enhanced my knowledge regarding this type of structural outline (not content-related).</td>
</tr>
<tr>
<td></td>
<td>I feel that it will be easier for me to deal with a text with a similar structural outline (not content-related) in the future.</td>
</tr>
</tbody>
</table>
demanded for a preliminary assessment of construct validity in order to ensure that our items measure what they are supposed to measure. As our measurement scales are based on reflective items, construct validity is enabled by ensuring unidimensionality, reliability and validity [59]. In order to investigate those three components, we conducted an item-sort task which evaluates substantive validity, which is a major contributor towards construct validity [60, 79].

For the item-sort task, 25 participants were recruited both from an academic as well as from an industry background. Such a small sample sizes is appropriate when evaluating substantive validity [79]. We have chosen this sample in order to include participants that suffer from e-mail overload. 60% of the test persons were male and 40% female respectively. Each of the test persons was instructed by the researcher about the design of the item-sort task before the definitions of the constructs were explained in order to avoid misunderstandings. Also, written definitions of each of the construct were provided to the participants. They were encouraged to read the definitions carefully before and during the item-sort task. Each participant had to decide for each questionnaire item which construct fit the respective item best. The items were randomly sorted. Participants were encouraged to report unclear and ambiguous wordings.

In order to evaluate substantive validity, two established indices were used for data analysis [see 79 for an extensive discussion]. The $P_{SA}$ index reflects how well the appropriate items fit to the construct, whereas the $C_{SV}$ index considers also the loadings on other constructs. The $P_{SA}$ index ranges from 0.0 to 1.0, whereas the $C_{SV}$ index ranges from -1.0 to 1.0. The higher the better is valid for both of the indices. However, the $C_{SV}$ index should have a minimum of at least 0.5. The overall values for both of the indices are presented in Table 3.

**Table 3. Results on substantive validity**

<table>
<thead>
<tr>
<th>Construct</th>
<th>Number of items</th>
<th>$P_{SA}$</th>
<th>$C_{SV}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>E-mail intrinsic cognitive load</td>
<td>5</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>E-mail extraneous cognitive load</td>
<td>6</td>
<td>0.97</td>
<td>0.93</td>
</tr>
<tr>
<td>E-mail germane cognitive load</td>
<td>6</td>
<td>0.97</td>
<td>0.94</td>
</tr>
</tbody>
</table>

The results of the three constructs show high values for both of the indices. The values of the $C_{SV}$ index are higher than the recommended threshold of 0.5 for all constructs, indicating high substantive validity and therefore good construct validity. For intrinsic cognitive load, all corresponding items of each of the participants were assigned to the correct construct. For the two other constructs, values less than 1.0 indicate that at least one person assigned one of the items to an inappropriate construct. The most wrong items were assigned to the germane cognitive load construct. All of them would have been extraneous cognitive load items. However, the construct definitions of extraneous and germane cognitive load share the structural outline as the same reference point. In contrast, the wrong assigned germane cognitive load items are allocated nearly equally to both intrinsic and extraneous cognitive load. The extraneous and germane cognitive load constructs have the same number of incorrect assigned items. In sum, there are only a few wrong assignments, so it can be concluded that the items measure the corresponding constructs appropriately.

6. Discussion

We set out to investigate and develop valid scales for measuring the impact of e-mail threads on intrinsic, extraneous, and germane cognitive load. Existing work is focusing on the total number of e-mails and how the increasing volume of e-mails leads to perceived cognitive load. This perspective neglects how formats like e-mail threads in terms of quoted conversations is impacting cognitive load. We proposed cognitive load theory [14-18] to distinguish between intrinsic, extraneous, and germane cognitive load [75, 76]. We conceptualize intrinsic cognitive load based on the content of the e-mail, which should not differ regardless of the presentation of the information. Consequently there is nothing to optimize in order to cope with e-mail overload. Extraneous cognitive load is devoted to deal with the structural outline of the e-mail threads. In order to theoretically explain e-mail threads we draw on the split attention effect which evokes extraneous cognitive load by physically not integrated information [53-57]. Therefore, extraneous cognitive load and e-mail overload may be reduced by choosing a more suitable presentation format. Furthermore, there is also germane cognitive load in order to build up knowledge for dealing with the structural outline of the e-mail threads.

In order to investigate this research gap and answer our research question, we developed appropriate measurement scales for each of our constructs by redeveloping items or creating new items. We based our work on a recently developed psychometric instrument which can differentiate intrinsic, extraneous, and germane cognitive load [75, 76]. On this basis we developed a measurement instrument to the domain of e-mails and also to its reference points for the different types of cognitive load in order to
reflect e-mail threads. The reference point for intrinsic cognitive load is the complexity of the text in terms of content, whereas for extraneous and germane cognitive load, it is the structural outline. Due to new items and the new domain, we wanted to gain first indications of construct validity before conducting the experiment on a large scale. Therefore, we conducted an item-sort task with 25 participants. The results of our analysis show that our scales measure the respective constructs appropriately and substantive validity is present. Therefore this paper provides several contributions. Foremost, it reports the rigorous development of valid measurement scales to measure the impact of e-mail threads on intrinsic, extraneous, and germane cognitive load. The described procedure, we followed, ensures high levels of confidence in content and substantive validity as a strong indicator for construct validity. Furthermore, this paper presents a crucial step towards the measurement of different types of cognitive load to determine the extent to which e-mail threads contribute to e-mail overload and can now be used in further studies. Our scales are not only useful for investigating e-mail threads. They can be also used for investigating the cognitive impact of any text based communication such as chats or social networks, the structural outline of any text and also for assessing computer interfaces presenting information in terms of text.

Once validated, the new constructs with its created items as well as the developed hypotheses are an important contribution for practice. In order to cope with e-mail overload, it is important to know the extent to which e-mail formats induce cognitive load and which type of cognitive load. This offers a valid starting point in order to increase efficiency of e-mail communication which reduces e-mail overload. Successful coping tools and mechanisms should be designed in such a way that they diminish extraneous cognitive load. In order to achieve this, an optimal presentation of information and the use of features allowing the reduction of extraneous cognitive load are needed. Our study also aims towards the suggestion that e-mail threads should be avoided because the overall cognitive load of this format is high due to the high extraneous cognitive load. The use of e-mail threads also consumes more cognitive load which reduces the ability to deal with complex problems and situations in terms of many information chunks. Information sharing, as it is needed through the division of labour in companies, should not be implemented by forwarding existing e-mail threads in terms of quoted conversations. It seems as if these practices should be avoided as the processing of these e-mail threads induces extraneous cognitive load.

For future studies, we encourage researchers to test our hypotheses with our measurement scales in an experimental setting, as it was never our intention to test our hypotheses with the item-sort task. We will adapt and sharpen the items again in order to cope with false assignment of items to the appropriate construct. Furthermore, we suggest investigating existing antecedents of IO and e-mail induced IO by distinguishing between intrinsic, extraneous, and germane cognitive load. It is also promising to identify further formats of e-mail by applying exploratory research methods. Afterwards, the impact of these formats of e-mail on e-mail overload can be determined by using our existing scales, measuring the different types of cognitive load.

There are also limitations inherent in our study. From a content perspective, it needs to be remarked that e-mail threads are not supposed to be the sole source of IO. Existing research demonstrates that the number of e-mails is an indicator for IO [10, 11] and stress [2]. Also, various technologies are a potential source for IO [25] as well as any information in general [6]. Furthermore, we scrutinized a cognitive lens on IO, but there might by further possible perspectives (e.g., affective perspective). From a methodological perspective, we could only evaluate substantive validity which gives indications of construct validity due to use of an item-sort task. For the final assessment of the constructs in terms of construct validity, it is necessary to test the scales using a larger sample size. Based on the data, it is also not possible to draw any conclusions in terms of our hypotheses as the model needs to be tested on a larger scale.

To sum up, existing research of e-mail induced IO relies solely on the total number of e-mails and its impact on perceived IO caused by e-mails. This approach implies that there is a clear need for considering the format of a single e-mail. Our central argument is that the high number of received e-mails is not the core of the problem; it is rather the format of a single e-mail that burdens humans. We proposed cognitive load theory to evaluate different e-mail formats and developed experimental measurement scales to assess e-mail threads as a history of appended e-mail quotes that are forwarded to a third person in terms of intrinsic, extraneous, and germane load. The scales were tested for construct validity based on an item-sort task. These scales are unique in assessing e-mail induced IO in terms of cognitive load which contributes to the research streams of e-mail overload and IO.

7. References


