Filing, Piling, and Everything In Between: The Dynamics of E-mail Inbox Management

Yoram M. Kalman
Department of Management and Economics, The Open University of Israel, Ra'anana 43537, Israel. E-mail: yoramka@openu.ac.il

Gilad Ravid
Department of Industrial Engineering and Management, Ben-Gurion University of the Negev, Be'er Sheva 84105, Israel. E-mail: gilad@ravid.org

Managing the constant flow of incoming messages is a daily challenge faced by knowledge workers who use technologies such as e-mail and other digital communication tools. This study focuses on the most ubiquitous of these technologies, e-mail, and unobtrusively explores the ongoing inbox-management activities of thousands of users worldwide over a period of 8 months. The study describes the dynamics of these inboxes throughout the day and the week as users strive to handle incoming messages, read them, classify them, respond to them in a timely manner, and archive them for future reference, all while carrying out the daily tasks of knowledge workers. It then tests several hypotheses about the influence of specific inbox-management behaviors in mitigating the causes of e-mail overload, and proposes a continuous index that quantifies one of these inbox-management behaviors. This inbox clearing index (ICI) expands on the widely cited trichotomous classification of users into frequent filers, spring cleaners, and no filers, as suggested by Whittaker and Sidner (1996). We propose that the ICI allows shifting the focus, from classifying users to characterizing a diversity of user behaviors and measuring the relationships between these behaviors and desired outcomes.

Introduction

E-mail is one of the most important tools for online communication. Effective e-mail inbox management is one of the main challenges faced by knowledge workers who are employed in an environment characterized by information overload. In this study, we explore a longitudinal data set that describes the inbox characteristics of thousands of users worldwide whose activities were tracked by an application that assists in managing e-mail overload. We then focus on the dynamics of messages entering and leaving the inboxes over the 24-hr day and throughout the week. These dynamics reflect a diverse repertoire of inbox-management behaviors which is significantly more complex than the often-cited trichotomous classification of e-mail users into frequent filers, spring cleaners, and no-filers (Whittaker & Sidner, 1996), and other classifications (Brogan & Vreugdenburg, 2008; Gwizdka, 2004). We analyze several different inbox-management behaviors in an effort to characterize them as well as to understand their relationship to e-mail load, and their potential value in coping with the sense of e-mail overload.

The Inbox Is the Hub of E-mail Activity

The inbox is where new e-mail usually arrives, and research on the way e-mail is handled focuses on the dynamics of e-mail in the inbox (Dabbish & Kraut, 2006; Dabbish, Kraut, Fussell, & Kiesler, 2005; Szóstek, 2011; Whittaker & Sidner, 1996). When a new message arrives in the inbox, it is added to a list of incoming messages. When users attend to the message, they evaluate whether and when it should be further handled, usually on the basis of attributes such as the name of the sender and the text that appears in the subject line (Szóstek, 2011). When messages are opened and handled, they are usually classified as either requiring a reply (immediate or postponed) or not (Dabbish et al., 2005), and the user decides whether the message needs to be filed outside the inbox, deleted, or left in the inbox. The inbox-management behavior of users is commonly based on classifying users into three categories: "pilers," "filers," and "spring cleaners." This classification, which was proposed almost 2 decades ago (Whittaker & Sidner, 1996) to describe those who never file their e-mail messages, those who file

Received November 20, 2013; revised May 8, 2014; accepted May 8, 2014

 • Published online in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/asi.23337

them on a daily basis, and those who file them only once every 1 to 3 months (respectively) is still used extensively in the literature.

Effective inbox management is one of the main challenges facing knowledge workers who constantly deal with what has been labeled *information overload*, *e-mail overload*, *communication overload*, and so on (Barley, Meyerson, & Grodal, 2011; Dabbish & Kraut, 2006; Eppler & Mengis, 2004; Whittaker & Sidner, 1996). Since e-mail is so central to the work of knowledge workers, effective inbox management is a way to ensure that important information is not missed when it comes in; that information can be retrieved, if necessary, in the future; that messages which require a response are handled as quickly as possible so that senders' expectations are not violated; that assignments which come in through e-mail are carried out efficiently; that interruption to other tasks in which the knowledge worker is engaged is kept to a minimum; and that the inbox remains a pleasant and effective work environment (Barley et al., 2011; Cadiz, Dabbish, Gupta, & Venolia, 2001; Capra, Khanova, & Ramdeen, 2013; Dabbish & Kraut, 2006; Fisher, Brush, Gleave, & Smith, 2006; Kalman & Rafaeli, 2011; Whittaker, Matthews, Cerruti, Badenes, & Tang, 2011).

One consequence of the centrality of e-mail in the daily routines of knowledge workers and of the centrality of the inbox in the ongoing management of e-mail is that the accumulation of messages in the inbox is a key cause of stress and of the sense of overload. Users who succeeded in keeping the inbox small experience less overload (Dabbish & Kraut, 2006), and users have reported that successfully "clearing" their inbox increased their sense that they can cope (Barley et al., 2011). Users use vivid imagery to describe the flow of incoming e-mail messages: a fire hose, a runaway assembly line, and a mountain of messages, to name a few, and describe specific practices focused on keeping the inbox clean: "I try to keep my inbox as clean as possible, although recently I haven't been able to drop it below 200. I normally would like to keep my e-mail inbox below 100" or "I'm usually on [e-mail on Sunday] three or four hours continuously, and that's sort of to sort of clean the plate out" or "if my inbox gets over 100, I start being really frustrated. I feel like I can't prioritize" (Barley et al., 2011). As Barley et al. (2011) described in detail, dealing with e-mail has become a symbol of the stress of the modern knowledge worker. They showed that although there are many sources of overload in contemporary work lives, the unique attributes of e-mail, especially its asynchronicity and textuality, result in users focusing on e-mail as a key source of their overload.

Cadiz et al. (2001) aptly borrowed the medical term *triage* to describe the task of clearing the inbox from incoming messages. Like an emergency room doctor who performs triage by prioritizing incoming cases based on urgency and on the type of required treatment, or like a military medic who needs to allocate limited resources under fire and quickly classify the injured into those who have little or no chance of recovering, those who require immediate attention, and those who can wait, so do the knowledge workers deal with an ongoing flow of messages. Rapid decision making is required to classify the messages and handle those that can or have to be taken care of immediately while carefully marking and/or filing those that require complex and often long-term handling (Bellotti, Ducheneaut, Howard, Smith, & Grinter, 2005). Efforts to improve the effectiveness of e-mail triage are ongoing (e.g., Alberts & Forest, 2012).

In summary, the literature on e-mail management and inbox triage suggests that the sense of e-mail overload is a subjective reaction by users to increases in inbox size, number of unread messages, and response times. As the messages accumulate, users' perception that they can effectively handle the message load is diminished. Users cope with this sense of overload using behaviors that focus on attending to the messages in the inboxes—reading, responding, and removing them from the inbox.

This literature (both quantitative and qualitative) on the management of e-mail in general and of the inbox in particular often has been limited by the low number and diversity of participants in the studies; usually has been based on interviews, questionnaires, or experiments; and has been rarely longitudinal. An interesting example of the power of a longitudinal analysis of the behavior of a large number of users is Jones, Ravid, and Rafaeli's (2004) article on information overload in 600 Usenet groups. The authors analyzed over 2.65 million Usenet messages posted over a period of 6 months, and demonstrated that the cumulative behavior of large numbers of users can be used to confirm hypotheses about information overload which takes place at the level of the individual user. In this article, we describe a large data set that tracks the inbox behavior of thousands of unrelated users worldwide. Although the data set was not created for research purposes but rather for commercial purposes, it is significantly larger than those used in previous studies—longitudinal, international, and cross-organizational. Moreover, data in this data set were collected unobtrusively. These characteristics enable us to try and replicate findings that were based on relatively small data sets, and to test ideas and concepts suggested by these studies. Specifically, we can quantify specific e-mail inbox-management behaviors in a large, diverse, and ecologically valid sample, and to evaluate the relationship between these activities and measures that are known to be associated with the sense of e-mail overload.

Our first research question (RQ1) seeks to confirm the link between an increase in the number of incoming messages and the increase in inbox size, number of unread messages, and response time. In other words, it asks whether the characteristics of the inboxes of users who receive large amounts of e-mail messages are those characteristics that are associated with a sense of e-mail overload. Our second research question (RQ2) explores the extent to which specific user behaviors are effective in decreasing inbox size, number of unread messages, and response time. In other

DOI: 10.1002/asi

words, we ask what user behaviors are associated with diminishing the sense of e-mail overload.

RQ1 was affirmed, and RQ2 led to the identification of user behaviors that influence the inbox variables that are associated with a sense of overload. Findings suggest an innovative continuous quantitative measure of inbox-clearing behavior that moves beyond the traditional division of users into general, somewhat vaguely defined, discrete categories such as filers, spring cleaners, and no-filers (Whittaker & Sidner, 1996). The characteristics of inboxes of users who display a higher inbox clearing index (ICI) suggest that clearing the inbox is an important way of reducing the sense of e-mail overload.

The study begins with a section that provides a benchmark longitudinal description of the inbox behavior of e-mail users. This is followed by an inferential section in which we test several hypotheses about the influence of increased numbers of incoming messages on inbox characteristics (RQ1), and on the influence of specific inbox-management behaviors on these characteristics (RQ2). Based on this inferential section, we closely examine the ICI and show that it informs us about users with different needs and preferences, suggesting future research using this index to explore the fit between users and their inbox-management strategies.

Hypotheses

The first hypothesis focuses on RQ1—the consequences of increasing message load. More incoming messages require more effort and time to review, sort, respond to, and act upon (Barley et al., 2011; Dabbish et al., 2005; Eppler & Mengis, 2004; Fisher et al., 2006; Szóstek, 2011). We expect this increasing load to increase the variables that are associated with a sense of overload; namely, inbox size at the end of the day, number of unread messages at the end of the day, and response time to messages. Thus:

> H1: There will be a positive linear correlation between the number of incoming messages of users, and their (a) inbox size, (b) number of unread messages, and (c) average response time.

The second and third hypotheses focus on the RQ2—the user behaviors that lead to decreasing the variables associated with a sense of e-mail overload: inbox size, number of unread messages, and response time. Based on the extensive literature on the efforts that users make to clear their inboxes throughout the day (Barley et al., 2011; Cadiz et al., 2001; Dabbish & Kraut, 2006; Fisher et al., 2006; Kalman & Rafaeli, 2011; Whittaker et al., 2011), we hypothesize that specific daily activities on the inbox will lead to a decrease in the three variables associated with overload: inbox size, number of unread messages, and response time to messages. Specifically, we hypothesize that:

> H2: There will be a negative linear correlation between users attending to their inbox more often: (a) removing messages, (b) sending messages (including replying), or (c) performing any activities in the inbox, and their (a) inbox size, (b) number of unread messages, and (c) average response time. Each of the three activities (a–c) will be checked against each of the three measures (a–c).

> H3: There will be a negative linear correlation between users' daily inbox-clearing activity and their (a) inbox size, (b) number of unread messages, and (c) average response time.

Method

Data Set

Data used in this study were contributed by a company that develops, markets, and supports an Outlook add-on designed to assist users who experience information overload in managing their e-mail messages. The add-on provides tools to better organize and prioritize messages, and to more effectively act upon them. In addition, it provides users with tools to better understand their e-mailing behavior. The company will be named "company X" to protect its proprietary information. The initial data set contributed by company X included data about the inbox utilization by its users from January 1, 2010 to August 31, 2010. This included all of the information that the company servers collect from the users to evaluate the users' e-mail load and their coping with it. Hourly records for each user included information such as number of messages in the inbox, number of read messages, number of messages sent, number of responses sent, average response time, and number of unread messages. Company X is based in the United States and has corporate and private users in the United States and abroad. Users' data were supplied in their local time, based on the time provided by the operating system.

The Structured Query Language (SQL) data set was reviewed and purged of records which (a) did not have matching session identifiers, (b) described users whose inbox count was unknown, (c) described users with less than 30 days of activity on the system, (d) occurred during periods during which the inbox size could not be deduced, and (e) included corrupted data. Then, a Perl script was developed to deduce missing data about inbox sizes and about the net number of messages which were cleared (deleted, filed in another folder, etc.) from the inbox. For example, if yesterday Sara had 14 messages in her inbox, and today she received four messages and removed five, it was inferred that at the end of the day there were 13 messages in her inbox. Some reasons for missing data included technical glitches and data missing when the tool developed by company X was unable to record inbox activities. The final data set that was used for the analysis included information about 7,745 users. Removing outliers did not materially influence the results; thus, no outliers were removed, except in the multiple regression (described later). This data set was analyzed using SQL queries. The longitudinal chronemic data (Figures 1 and 2) were based on the nonaggregated data set. The rest of the reports are based on a data set that summarized the averages for each of the 7,745 users.

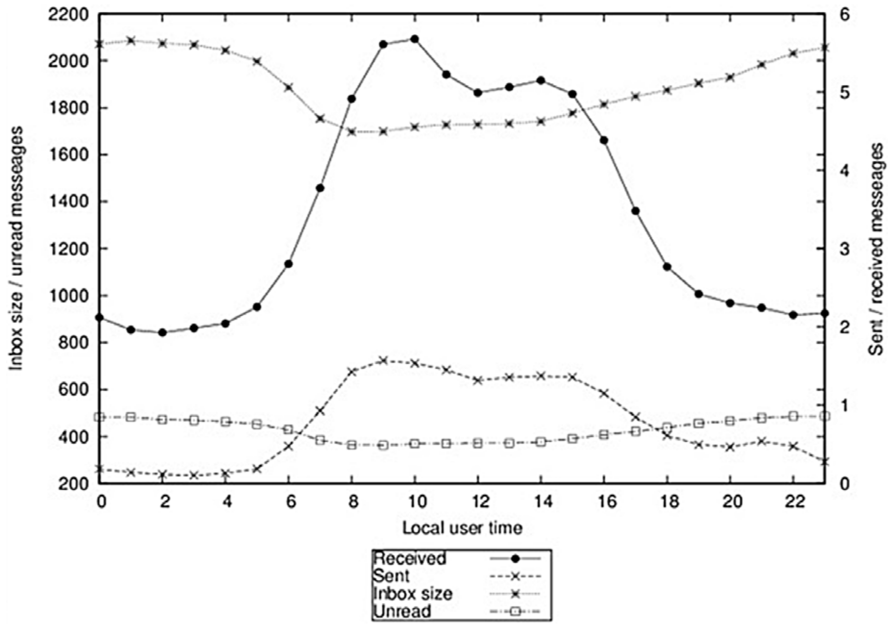

FIG. 1. Average levels of inbox activity, by (local) time of day. •Received per hour; ˣsent per hour; *end of day inbox size; □end of day unread.

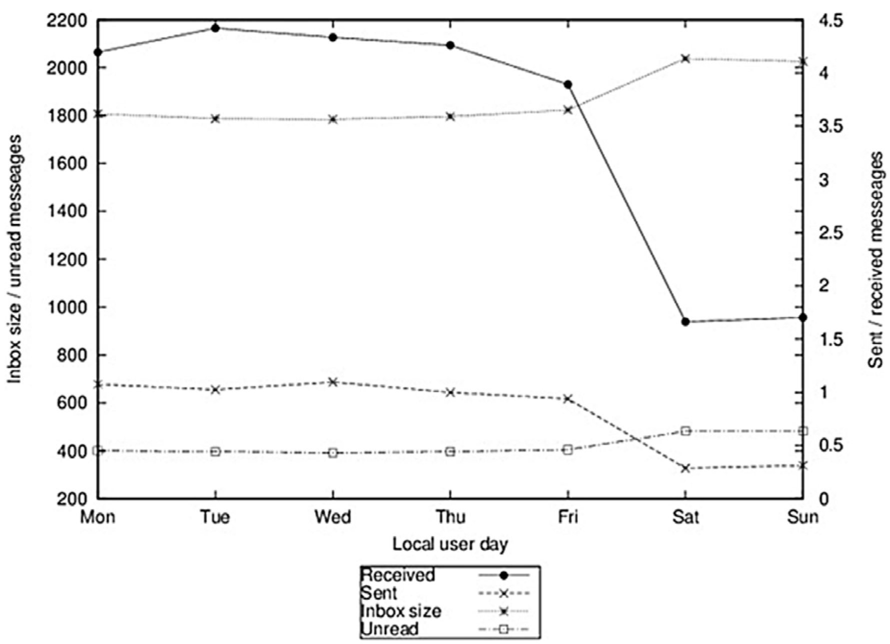

FIG. 2. Average levels of inbox activity, by day of the week. •Received per hour; ˣsent per hour; *end of day inbox size; □end of day unread.

Variables and Analyses

Inbox variables used in this study include the averages of the following:

- **Inbox size:** number of messages in the inbox at the end of the day.
- **Received:** number of messages received by the user during the day.
- **Sent:** number of messages sent by the user during the day. Includes replies to messages.
- **Replied:** number of messages replied by the user during the day. *Replied* is a subset of *Sent*.
- **Unread:** number of unread messages in the inbox at the end of the day.
- **Removed:** number of messages removed from the inbox (e.g. erased, filed) by the user during the day.
- **Response time:** the average of the user's response times during the day.

DOI: 10.1002/asi

TABLE 1. Average characteristics of users' inboxes and activities.

	Units	Average (*SD*)	Range (10th–90th percentile)
Inbox size	Messages by end of day	1,551 (3,980)	14–3,996
Received	Messages/day	57 (103)	10–109
Sent	Messages/day	14 (28)	1–28
Replied	Messages/day	5 (5)	0.3–11.0
Unread	Messages by end of day	345 (1,690)	1–547
Removed	Messages/day	58 (155)	6–119
Response time	Minutes from receipt of incoming message	1,783 (3,505)	403–3,234
Sending frequency	Frequency	3.50 (1.96)	.68–5.98
Removing frequency	Frequency	6.07 (3.87)	1.85–11.03
Activity frequency	Frequency	6.98 (3.77)	2.70–11.72
Inbox clearing index	Correlation coefficient	.50 (.33)	.05–.93

Note. The frequency unit is defined as the number of hours during the day of which the user performed the action at least once.

Variables measured during the day were measured from 00:00 to 23:59. Variables measured at the end of the day were measured at midnight.

Users' frequency of activity variables in this study are used to measure how often during the day users enter their inbox and perform an activity. These variables are based on measuring whether during each hour of the day (e.g., between 00:00–00:59, between 01:00 and 01:59, etc.) the user performed one of the two following activities at least once: sent a message or removed a message from the inbox. The number of hours during which each of these activities took place was summed, resulting in a number between 0 and 24. Three variables were based on these numbers:

- **Sending frequency:** number of hours during the day that the user sent at least one message.
- **Removing frequency:** number of hours during the day that the user removed (e.g., filed, deleted) at least one message.
- **Activity frequency:** number of hours during the day that the user sent or removed at least one message.

Finally, the ICI of each user was derived by calculating the Pearson correlation between the number of messages that the user received each day and the number of messages that the user removed from the inbox on that same day. The ICI is the correlation coefficient between these numbers for each user, and as such, it has values from −1 to 1.

Seven of the variables displayed a highly skewed distribution: inbox size, received, sent, replied, unread, removed, and response time. These variables were normalized using a log transformation. A visual inspection of the distributions confirmed that they were significantly more even and symmetric.

A multivariate correlation was performed on the inbox variables listed earlier, and a multiple regression using backward elimination was performed on the ICI. Both analyses used the ICI and the 10 inbox variables, with log transformation of the seven highly skewed variables.

Results

This section begins with a benchmark longitudinal description of the inbox properties and activities of e-mail users. This description is followed by an inferential part in which we test the three hypotheses.

Description of Inbox Properties and Activities

Inbox characteristics. Analysis of the data set revealed a very diverse set of inbox characteristics or "vitals," as described in Table 1. As described earlier, all variables were calculated at the end of each day: the total number of messages sent, received, replied to, and removed from the inbox during the day, the inbox size and number of unread messages as of the end of the day, and the weighted (by number of messages) average of all the (nonzero) average hourly response times of messages responded to during each of the hours of the day. Then, these data were averaged for each of the users for the period that they were active. Thus, for example, the inbox size measure reported in Table 1 is an average of the average inbox size calculated for each of the users in the study, and the range is the range of the average inbox sizes.

Longitudinal behavior. Figure 1 describes the (average) level of inbox activities throughout the day, and Figure 2 describes the (average) level of inbox activities throughout the week. Note that the left *y*-axis denotes the average number of messages in the inbox by the end of the day and the average number of messages marked as unread by the end of the day. The right *y*-axis denotes the average number of messages sent *per hour* and the average number of messages received *per hour*. All times in Figures 1 and 2 are in the local time of the users' computers.

Inbox maintenance. A manual examination of many individual inboxes revealed significant diversity between users. A sample of this diversity is demonstrated in Figure 3, which

DOI: 10.1002/asi

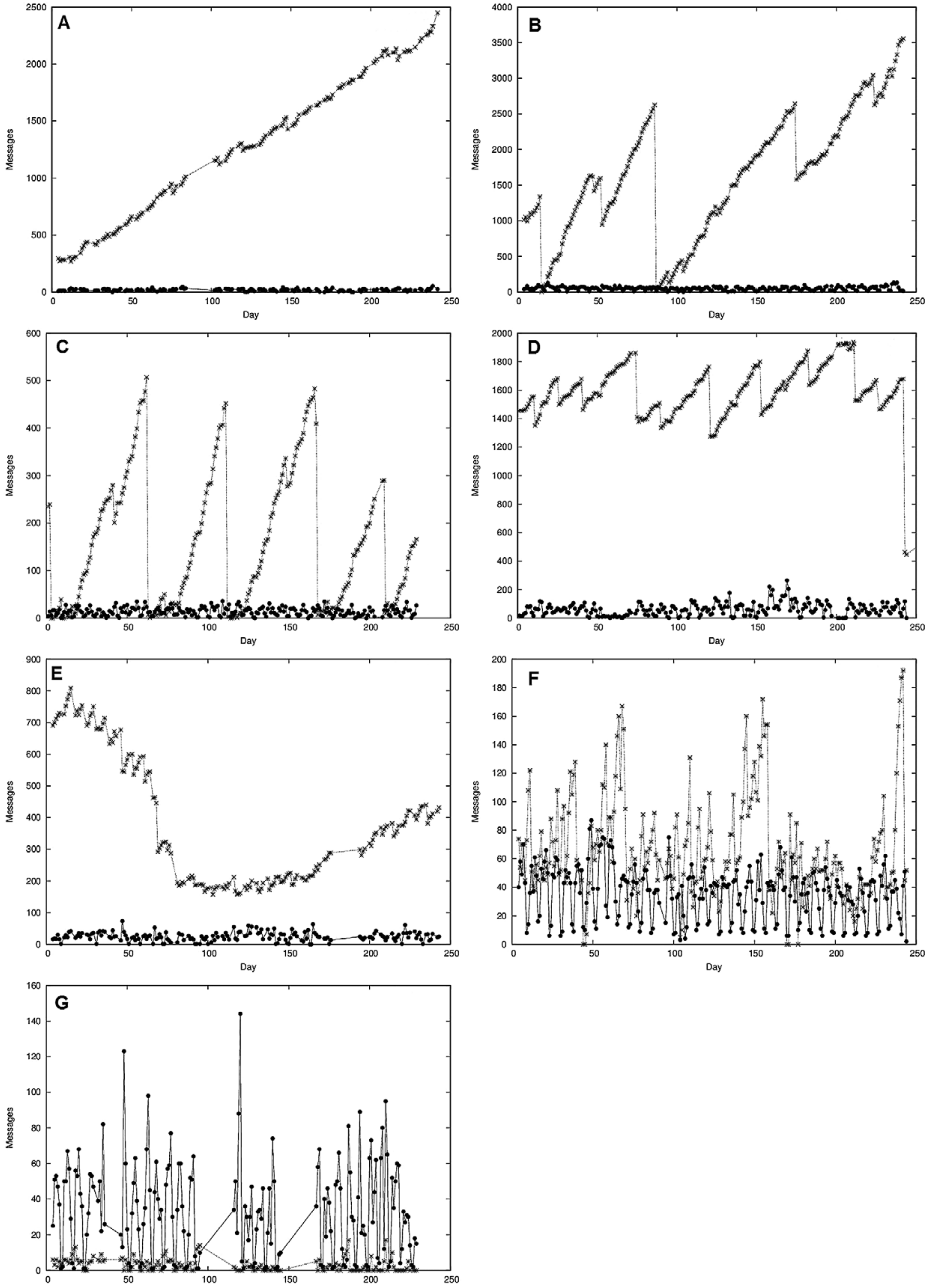

FIG. 3. The inbox activity of seven users included in the study: daily number of incoming messages (light x) and daily inbox size by day's end (filled dark circle). Note that the scale of the *y*-axis varies between users. The ICIs of users A–G are .01, .06, .14, .39, .48, .65, .99, respectively.

DOI: 10.1002/asi

TABLE 2. Correlations among the key study variables.

	Inbox size	Received	Sent	Replied	Unread	Removed	Response time	Sending frequency	Removing frequency	Activity frequency	ICI
Inbox size	1	.22	.03	−.01 n.s.	.64	.11	.02	.02 n.s.	−.09	−.06	−.54
Received		1	.56	.53	.31	.87	−.14	.57	.64	.66	.25
Sent			1	.83	−.00 n.s.	.53	−.16	.81	.35	.44	.29
Replied				1	−.01 n.s.	.51	−.11	.82	.37	.45	.32
Unread					1	.22	.06	−.00 n.s.	.02 n.s.	.02 n.s.	−.36
Removed						1	−.13	.53	.62	.63	.21
Response time							1	−.20	−.12	−.14	−.08
Sending frequency								1	.45	.55	.29
Removing frequency									1	.98	.35
Activity frequency										1	.33
Inbox clearing index											1

Note. n.s. superscript = less than significant ($p > .05$) correlation. All other correlations are statistically significant. The following variables were log transformed: inbox size, received, sent, replied, unread, removed, and response time.

describes the inboxes of seven users from the data set. These users were selected to demonstrate the diversity of strategies as well as to represent a wide range of ICIs from .01 to .99. Note that the scale of the *y*-axis is different for different users, based on their level of activity. The ICI, which describes the correlation between the daily number of incoming messages and the daily number of messages removed from the inbox, is noted in the figure, and is discussed in the inferential section.

Hypotheses Testing

A correlation analysis was performed on the key study variables, and a correlation matrix is presented in Table 2.

H1 predicted that there will be a positive linear correlation between the number of messages users receive and their (a) inbox size, (b) number of unread messages, and (c) average response time. H1(a) was supported ($r = .22$), H1(b) was supported ($r = .31$), and H1(c) was rejected ($r = -.14$).

H2 predicted that there will be a negative linear correlation between users attending to their inbox more often: (a) removing messages, (b) sending messages (including replying), or (c) performing any activities in the inbox, and their (a) inbox size, (b) number of unread messages, and (c) average response time. H2(a) was supported for inbox size ($r = -.09$) and for average response time ($r = -.12$), and H2(b) was supported only for average response time ($r = -.20$). H2c was supported for inbox size ($r = -.06$) and for average response time ($r = -.14$).

H3 predicted a negative linear correlation between users' daily inbox-clearing activity (ICI) and their (a) inbox size, (b) number of unread messages, and (c) average response time, and the three hypotheses were supported ($r = -.54$, $r = -.36$, $r = -.08$, respectively).

Regression

The multiple regression (using backward elimination) was used to test if the inbox variables explained the ICI. As described earlier, all 10 inbox variables were included as effects in the model (seven of which were log transformed). Because of the high correlation between some of the variables, a variance inflation factor (VIF) was calculated; consequently, activity frequency (almost fully correlated to removing frequency) was removed from the model. The analysis resulted in a model (Table 3) comprising seven variables which explain 47% of the variability in the ICI: $F(7, 7359) = 938$, $p < .0001$, $R^2_{\text{adj}} = 0.47$.

TABLE 3. Predictors of the inbox clearing index.

Variable	Estimate (*SE*)	[95% confidence level]	
Intercept	.613 (.013)**	.588	.639
Inbox size	−.078 (.002)**	−.082	−.075
Received	.145 (.006)**	.133	.157
Replied	.041 (.004)**	.034	.048
Unread	−.017 (.002)**	−.020	−.014
Removed	−.061 (.004)**	−.070	−.052
Sending frequency	−.008 (.003)*	−.014	−.003
Removing frequency	.009 (.001)**	.007	.011

Note. The following variables were log transformed: inbox size, received, replied, unread, and removed.
$^*p = .001$. $^{**}p < .0001$.

Discussion

The findings of this study paint a detailed picture of the inbox dynamics that manifest as e-mail users manage their loaded inboxes. It shows how thousands of users around the world deal with the constant inflow of hundreds of messages a week and the outflow of dozens of messages. It shows how messages accumulate in the inbox around the clock and throughout the workweek and the weekend, and are then read, responded to, filed away, or erased. In this discussion, we review the findings, compare them to previous findings, discuss the daily and weekly cycles of e-mail, the dynamics of inbox management, the diversity of user inbox-management behaviors, and the power of the

DOI: 10.1002/asi

ICI to capture some of this diversity. Finally, we discuss the limitations of the study and suggest future directions for research.

Participants are Typical Knowledge Workers

Most previous studies of e-mail usage and inbox management have been based on a relatively small sample of well-characterized users. In contrast, this study is based on an anonymous sample of thousands of users who have the add-on developed by company X installed on their machines. Before we proceed to discuss their inbox behavior, we verify that there are no indications that our sampling method biased our findings in unexpected ways. We achieved this by comparing the characteristics of the study's users with the characteristics of users in previous studies of e-mail management, and conclude that our study's users are typical knowledge workers.

Inbox size is the most commonly reported variable in studies of e-mail management and information overload. As can be seen in Table 1, the average inbox size of the users in our study was 1,551 ($SD = 3{,}980$) messages, and the range (10th–90th percentile) was 14 to 3,996. Fisher et al. (2006) compared their findings in 600 mailboxes of employees at a high-tech company to the findings reported by Whittaker and Sidner (1996) on 18 users in another high-tech company. The average inbox size in Whittaker and Sidner's study was 1,624 messages while the average inbox size in the study by Fisher et al. was 1,150. The 124 respondents included in the Dabbish et al. (2005) paper reported an average inbox size of 1,336 ($SD = 2{,}785$) messages. Whittaker et al. (2011) reported that the average inbox size of 345 of their participants (employees in a high-tech company) was 870 ($SD = 1{,}423$). The average inbox size of 484 randomly selected participants from the United States, as reported by Dabbish and Kraut (2006), was 311, with only 10% reporting an inbox larger than 600 messages.

The number of incoming messages per day is another common variable reported in the studies. Our study's participants received 57 messages per day ($SD = 103$), and the range (10th–90th percentile) was 10 to 109. Whittaker and Sidner (1996) reported 49 messages, Fisher et al. (2006) found an average of 87, Dabbish et al. (2005) noted an average of 30 messages *read* per day, and Dabbish and Kraut (2006) reported an average of 41 messages per day.

Overall, we can see that the amounts of e-mail messages handled by the users in our study are similar to those reported in previous studies. Since most of the studies were carried out in the context of knowledge workers, note the exception—Dabbish and Kraut's (2006) study—which studied a more representative sample of the general U.S. population and which reported a somewhat lower level of activity.

Another reference for comparison is the study of 79 employees in a high-tech company reported by Barley et al. (2011). Their study focused on the tasks these employees carried out throughout their workday and included a longitudinal report on media usage (e-mail, telephone, meetings, and teleconferences) throughout the day. The longitudinal chronemic pattern they identified (p. 901, Figure 1) is very similar to the one that we identify in Figure 1 of our study: Activity starts to pick up around 5 a.m., rises until about 9 a.m, remains high until about 4 to 5 p.m. (3 p.m. in our study) with a slight drop around midday (lunch), and drops until late evening.

A final reference point for the characteristic behaviors of the users in our study is average response time. An extensive study by Kalman and Rafaeli (2005) ($N = 14{,}740$) of response times to e-mail messages in a large corporation reported that the average response time was 1,728 min, quite similar to the average of 1,783 min in our study.

Based on the comparison with previous studies, it is possible to say that the inbox characteristics and behavior of the population in our study appear to be similar to those of the populations of knowledge workers studied earlier. This population shows a somewhat higher level of activity than that measured in the Dabbish and Kraut (2006) study of a more representative sample of the general U.S. population, which is a sample that includes a mixture of participants, both typical knowledge employees who often "live in their inboxes" as well as e-mail users who mainly utilize e-mail for personal communication. In addition to confirming that the users in our study are typical knowledge workers, Table 1 can serve as a reliable benchmark for future work on e-mail management.

Inbox chronemics. Figures 1 and 2 provide an interesting glimpse into the chronemic (time-related) dynamics of the inbox. It demonstrates the basal level of incoming messages, which does not drop below 1.9 messages per hour at any time of the day, and which rises to an average of five to six messages an hour during the workday. Note that these averages include both workdays and weekends (and holidays), during which the amount of activity drops. The number of incoming messages peaks at around 9 to 10 a.m., then drops until around lunchtime (noon–1 p.m.), slowly rises again until around 2 to 3 p.m., and starts dropping. The general pattern regarding increases and decreases in average volume is almost identical for incoming and outgoing messages. That is not surprising, given that many of the received messages have been sent by some other user. Nevertheless, note that the ratio of incoming to outgoing messages does not remain constant. At 4 a.m., the ratio of sent to received messages is approximately 1:19, which then rises to a ratio of 1:3 to four messages between 8 a.m. and 6 p.m. One possible explanation for this is that the effort required to send an e-mail message remains relatively constant. It is difficult to rise above a small number of messages per hour. On the other hand, there are no restrictions on incoming messages; thus, messages which are sent to many recipients and/or via automated mechanisms are received by many people.

Another interesting finding described in Figure 1 is that the only time we see a significant drop in the number of messages in the inbox and in the number of unread messages is early in the morning, before the "official" beginning of the office workday, and that starting from 10 a.m., these two indicators of information overload are constantly on the rise until the next morning. These dynamics are quantitative evidence of the stresses expressed by e-mail users who feel swamped by the flood of incoming messages which accumulate, often unread, until the end of the day and beyond (Barley et al., 2011).

Figure 2 complements the picture that Figure 1 provides by describing the weekly chronemic dynamics of the inbox. The most prominent difference is between the level of activity on workdays in most of the Western world (Monday–Friday) and the weekend days (Saturday–Sunday). We see that during the weekend, the average number of incoming messages drops from over four an hour to about 1.7 an hour. Accordingly, the number of sent messages drops from an average of about one per hour to about 3 per hour. On the other hand, despite the relatively low level of activity, the average inbox sizes and the average number of unread messages increase over the weekend. Note that the effects of the weekend begin on Fridays, possibly since it is a traditionally shorter day or due to the effect of countries where Friday is a part of the weekend, such as in some Muslim countries and Israel ("Workweek and weekend," 2013).

Inbox Management

Figure 3 describes the inboxes of seven users selected from the data set. We chose these individuals to demonstrate the variability of inbox-maintenance behaviors, and that these behaviors are more complex than are the typical trichotomous division of users into pilers, filers, and spring cleaners.

User A accumulates messages in his inbox, which is constantly expanding. We can see some removal of messages from the inbox, but in general, the inbox size rises monotonously. This user is a classic no-filer ("piler") according to the Whittaker and Sidner (1996) classification. User B managed to completely clear his inbox twice during the study period (˜Days 10 and 90) and to clear a significant portion of the messages a few more times during this period. User C is a typical "spring cleaner," who managed to clear his inbox several times a year. Note that such purging events sometimes took more than 1 day (e.g., ˜Day 170), and that this spring cleaning sometimes ended before the inbox was empty. It also appears that for this individual user, an inbox size approaching 500 messages was the trigger for at least some of the purging events. Similarly, User D also tried to maintain an inbox size below 2,000 messages, and tried to achieve this goal by clearing several hundred messages every few weeks. Once, toward the end of the study period, that user cleared over 1,000 messages from the inbox. User E cleared his inbox every few days, and once during the study period went through a "spring cleaning" event when he significantly trimmed the inbox size. Unlike the previous users, the typical inbox size of User F is the same order of magnitude as the number of incoming messages. Note that the user's inbox size fluctuated significantly, but that overall, the inbox was cleared of most messages that had accumulated within a few days. We also can see evidence for three to four spring cleaning events when the inbox was completely or almost completely emptied. Finally, user G constantly cleared his inbox, not allowing the inbox size at the end of the day to be greater than 10 to 20 messages, even on days when many dozens of messages were received.

The ICI

The ICI was developed based on insights derived from a close examination of many individual users such as Users A to G discussed earlier as well as on studies of e-mail maintenance. Even this small sample of seven users demonstrates the limitations of the trichotomous classification of users into frequent filers, spring cleaners, and no-filers (Whittaker & Sidner, 1996) and variations on this classification (Brogan & Vreugdenburg, 2008; Gwizdka, 2004): The same person could be both a piler and a filer (Users D and F). Moreover, users such as D, E, and F demonstrate that spring cleaning could take place more or less often than the 1 to 3 months that was suggested by Whittaker and Sidner (1996), and might be triggered not only by the passage of time but also by inbox size. In summary, many users sometimes pile messages, often remove ("file") them, and occasionally clean ("spring clean") their inbox to some extent. Although these three categories were very useful as a first approximation when they were suggested almost 20 years ago, they are no longer adequate for classifying users and for gaining deeper insight into their behavior (Fisher et al., 2006). We propose that the ICI is a continuous measure that overcomes some of the limitations of the existing classifications.

The ICI is based on several insights about inbox maintenance that were gleaned both from the examination of our study's users' data as well as from conclusions of previous cited literature. One insight was that there is a pervasive perception among e-mail users that the fundamental unit of measure is the *single day*. As described in the Introduction, handling the day's e-mail is a key challenge for users, and the state of the inbox at the end of the day is high on user's minds. Moreover, we know that responding to an e-mail message within 1 day is an acceptable norm (Kalman & Rafaeli, 2011; Kalman, Ravid, Raban, & Rafaeli, 2006; Tyler & Tang, 2003). We also see that when users fail to handle and clear their inbox on a daily basis, many will clear the inboxes at a later time—possibly when some threshold (e.g., number of messages) or another condition (e.g., some free time) is reached.

The ICI is a continuous quantitative measure that looks at the correlation between the number of incoming messages and messages that have been handled and removed from the

DOI: 10.1002/asi

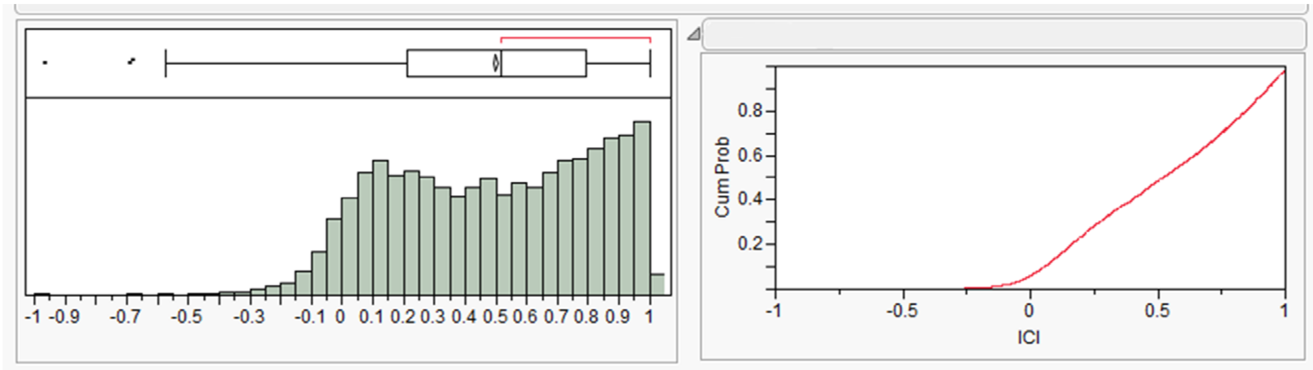

FIG. 4. The distribution of users with different ICIs in the studied population as a histogram (left panel) and cumulative density function (CDF) (right panel). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

inbox during each 24-hr day. Users who are pure pilers (e.g., User A in Figure 3) have a very low correlation: Messages come in, and only a few are removed. Users who are committed filers (e.g., User G in Figure 3) constantly handle and remove messages from the inbox as they come in, and thus the correlation is very high. Given the importance of the daily inbox activities, we propose that the ICI is a useful measure to differentiate between different inbox-management strategies. As can be seen in Figure 4, the positive values of the ICI, which is a correlation coefficient, are distributed relatively evenly, and about 10% of the values are negative.

The multiple regression demonstrates that almost half of the variance in the ICI is explained by a model comprising seven of the inbox variables. The model shows that despite the fact that a high ICI is associated with more received and replied messages, it also is associated with users who have, by the end of the day, a smaller inbox and less unread messages.

To demonstrate the usefulness of the ICI, we divided this continuous measure into five categories (described in Table 4): all ICIs which are either negative or below .2 are classified as ICI Category 1, and the rest of the ICIs are grouped into Categories 2 to 5 in groups of relatively equal size: Category 2, .2 = < ICI < .4; Category 3, .4 = < ICI < .6; Category 4, .6 = < ICI < .8; and Category 5, .8 = < ICI. The resulting Table 4 demonstrates in a more tangible manner the correlations described in Table 2 and the regression described in Table 3.

TABLE 4. Average inbox characteristics of users in each inbox clearing index (ICI) category.

ICI categories	5	4	3	2	1
Inbox size	604^{D}	794^{D}	$1,394^{C}$	$1,855^{B}$	$3,032^{A}$
Received	90^{A}	57^{B}	50^{C}	45^{CD}	39^{D}
Sent	20^{A}	15^{B}	13^{BC}	12^{CD}	11^{D}
Replied	7^{A}	6^{B}	5^{C}	4^{D}	3^{E}
Unread	112^{D}	125^{CD}	245^{C}	389^{B}	798^{A}
Removed	89^{A}	56^{B}	52^{B}	48^{BC}	41^{C}
Response time	$1,306^{C}$	$1,794^{B}$	$1,857^{AB}$	$1,986^{AB}$	$2,066^{A}$

Note. In each of the rows, values not connected by the same superscript letter are significantly different (Student's *t* test).

Dealing with e-mail overload. H1 hypothesized that users with increasing e-mail loads will have larger inboxes and more unread messages, and that their response times will be lower. The first two hypotheses received some support (low, but positive, correlations of about .2–.3) while the third hypothesis was rejected: There is actually a statistically significant low negative correlation between the daily number of incoming messages and response time. It seems that users who receive more messages have to deal with increasingly larger inboxes and with increasingly higher levels of unread messages, but these overloads do not negatively impact their average response times, and the response times of those who are more overloaded are even slightly shorter.

H2 hypothesized about the link between how often users attended to their inboxes, and the same three variables: inbox size, unread messages, and average response time. Results show that the main link is between attending to the inboxes and average response time. Attending more often was consistently correlated with shorter response times. A possible interpretation is that when users attend to their inboxes many times a day, they also respond to messages and thus shorten the average response times, as compared to users who attend less often to their inboxes. The negative link between removing frequency and inbox size was as expected, although it is quite weak.

H3 hypothesized that users whose ICI is higher will present smaller inboxes, less unread messages, and shorter response times, and was supported. The support H3 received suggests that the ICI is a useful variable and should be examined more closely. An examination of Table 2 shows that the ICI is positively correlated with the number of messages received, sent, replied, and removed, and is negatively correlated with inbox size, number of unread messages, and average response time. The ICI also is positively correlated with the three activity measures. Similarly, Table 3 also demonstrates that the ICI is associated with more received

DOI: 10.1002/asi

and replied message, and with smaller inbox sizes and less unread messages at the end of the day. Table 4 demonstrates this by reporting the averages for each of the ICI categories. It shows a monotonous rise or fall in the relevant variable as we move from users with a high ICI to users with a low ICI. It also demonstrates that many of these changes are statistically significant. For example, when we examine Category 5, which comprises about one fourth of the study population with the highest ICI, we can see these users have the smallest average inbox size, receive the highest average number of messages, send the highest average number of messages, reply to the highest average number of messages, remove the highest average number of messages, and their average response time is the lowest. We also can see that these differences are statistically significant from all other ICI categories for the average number of received, sent, replied, and removed messages as well as for the average response time, and that they are statistically different from Categories 1 to 3 for the two remaining variables: average inbox size and number of unread messages. Finally, we can see that these differences are not only statistically significant but also materially significant: Users in Category 5 receive, on average, twice or more messages than do users in Categories 1 and 2; despite that, they manage to respond to and send close to twice the number of messages and to have an average response time which is at least one third faster.

An examination of Tables 2 to 4 also provides an answer to a question raised in the seminal Whittaker and Sidner (1996) paper. They observed (p. 282) that managers were more likely to receive greater volumes of e-mail, and asked whether these managers were less likely to be frequent filers, given their higher volume of received e-mail and greater time spent in meetings. They did not find strong evidence for this relationship, but it was not clear whether this was due to the small sample size. Our findings clarify that the fact that users receive many more messages does not mean that they have less time to organize them. On the contrary, note that users in Category 5 actually receive the *highest* average number of messages and are able to deal with them (respond, remove, etc.) at least as well as their peers from the lower ICI categories.

Limitations and Future Directions. A key limitation of our study is that it is based on a convenience sample of users who installed a specific commercial product and on the information collected by this product. This means that the users are anonymous and that we are only able to learn about a small number of variables in their inboxes. Unlike other studies cited here, we cannot, for example, ascertain whether a message that no longer appears in the inbox was erased or filed away. Similarly, we cannot distinguish between a message that has not yet been read by a user and a message that was read and later marked as unread. Moreover, since the tool is based on an e-mail client, we cannot establish to what extent these conclusions can be generalized to other forms of e-mail, such as web-based interfaces (e.g., Gmail, Yahoo!).

Another important limitation of the study is that most of the analyses focus on average user behaviors and not on their behaviors hour by hour. When all of the activities of a user are summarized into a small number of variables (e.g., average inbox size, average response time), many details are averaged out. This also is the reason why many of the effects that we observed in the study are low to moderate correlations. The size of the data set allowed detecting these effects despite the averaging out, but future research should explore these trends detected at the aggregate level, at the level of smaller groups and of individuals.

An additional limitation that is a consequence of the fact that we only had access to the inbox activities is that it is impossible to directly measure overload of the users. The study is able to infer the sense of overload that the users experienced based on variables identified in previous studies, as well as to demonstrate how specific inbox activities succeed (or do not succeed) in reducing these variables. We infer that this reduction leads to a decrease in the sense of overload.

Finally, as in any study that focuses on the analysis of an unobtrusively collected data set, note that correlation is not causality and that our observations of online human behavior are not manipulations. Although some of our hypotheses that were based on causal mechanisms were supported, this is not conclusive evidence that these causal mechanisms are the explanations of the phenomena that we observed. For example, if H3 was based on the assumption that more diligent clearing of the inboxes will lead to decreased average response time, the correlation we observe has alternative explanations such as that users with a high e-mail load who are overrepresented in the high ICI groups are busier and thus have less time to dedicate to each of the e-mail messages—they simply respond quickly and thoughtlessly, and move on.

These limitations are balanced by the innovativeness of this study, which is based on a significantly larger data set than that of any previous studies. Moreover, the data set was collected unobtrusively (Lee, 2000) over a period of 8 months, was cross-organizational and international, and included data points from almost every hour that the e-mail client was active. Future research could use the quantitative aggregate findings and the insights of this study and explore them in more depth. In particular, it would be interesting to further explore the differences between users with different inbox-management behaviors such as varying ICIs: How do these differences relate to organizational variables such as position, to personal attributes such as personality, and to cultural variables such as a polychronic versus a monochronic orientation (Stephens, Cho, & Ballard, 2012).

Finally, note that our results emphasize the complexity of identifying an inbox-management behavior that is optimal for different users. We did not find that any one approach (e.g., checking the inbox many times during the day or not more than a few times a day) is the right strategy for all users. As is evident from a review of the small sample of seven users described in Figure 2, the preferences and

DOI: 10.1002/asi

circumstances of different users vary significantly, as do the solutions they apply. It would be wrong to suggest, for example, that users with a higher ICI necessarily have a better strategy for handling e-mail than do those with a lower ICI. We can only determine that these high-ICI users seem to be more effective in mitigating the main variables that are the cause of a sense of e-mail overload: a large inbox and a lot of unread messages at the end of the day.

A better understanding of the fit between needs, preferences, and strategies is left for future study. We believe we have demonstrated that the ICI will be a useful variable in such studies. Like any single descriptive construct, the ICI also simplifies a complex human behavior by reducing it to a single number; that is its strength as well as its weakness.

Conclusion

This study provides a broad quantitative description of the inbox activities of thousands of e-mail users worldwide, over a period of 8 months in 2010. The data for this study were collected unobtrusively on an hourly basis by an information-overload management tool that the users installed for their e-mail client, and the activities are thus highly detailed and representative of the behavior of knowledge workers across many organizations and contexts. Our findings demonstrate the high variability in behaviors between different e-mail users, and move research beyond the oft-cited division of users into general categories such as filers, spring cleaners, and no-filers. The finding that the ICI of most users is somewhere between 0 (pure pilers) and 1 (pure filers) suggests that most users engage in "piling"—accumulating messages in the inbox, "filing"—daily removing of messages from the inbox into the "trash" or other folders, and "cleaning"—a periodic removal of a larger number of messages from the inbox. We can assume that the way each of these is carried out is determined by each user's circumstances and preferences. Apparently, we should move beyond classifying *users* to classifying user *behaviors*. Users combine behaviors and switch between them. Moreover, we can assume that when the same user manages his Outlook inbox, their behavior might differ from their behavior when they manage a web-based e-mail inbox (e.g., Gmail). Based on these insights, we studied different inbox-management behaviors such as continuous attending to the inbox during the day as well as continuous daily clearing of the inbox. We identified a quantitative index to measure this inbox maintenance behavior, the ICI. This index is based on the correlation between the daily number of messages received in the inbox and the daily number of messages removed from the inbox by the user. We showed that this index is effective in distinguishing between different users without placing them into rigid categories, as previous classifications have done. Specifically, we showed that users with a high ICI (.8 or above) are characterized by a large average number of incoming messages, sent messages, and replied messages, and by a shorter response time to messages. Despite the large number of incoming messages, the average number of messages and unread messages in their inboxes by the end of the day is smaller. We suggest that this high-ICI behavior is exhibited by users who are trying to mitigate the sense of e-mail overload which is caused by a high number of incoming messages, unread messages, and messages in the inbox.

Acknowledgment

We wish to thank the executives at "company X" for their assistance and support.

References

Alberts, I., & Forest, D. (2012). Email pragmatics and automatic classification: A study in the organizational context. Journal of the American Society for Information Science and Technology, 63(5), 904–922. doi:10.1002/asi.21702

Barley, S.R., Meyerson, D.E., & Grodal, S. (2011). E-mail as a source and symbol of stress. Organization Science, 22(4), 887–906. doi:10.1287/orsc.1100.0573

Bellotti, V., Ducheneaut, N., Howard, M., Smith, I., & Grinter, R.E. (2005). Quality versus quantity: E-mail-centric task management and its relation with overload. Human–Computer Interaction, 20(1–2), 89–138. doi:10.1207/s15327051hci2001&2_4

Brogan, M., & Vreugdenburg, S. (2008). You've got mail: Accountability and end user attitudes to email management. In Proceedings of International Conference on e-Government (pp. 63–70). Melbourne, Australia: Academic Publishing Limited.

Cadiz, J.J., Dabbish, L., Gupta, A., & Venolia, G.D. (2001). Supporting email workflow—Microsoft Research. Retrieved from http://research.microsoft.com/apps/pubs/default.aspx?id=69881

Capra, R., Khanova, J., & Ramdeen, S. (2013). Work and personal e-mail use by university employees: PIM practices across domain boundaries. Journal of the American Society for Information Science and Technology, 64(5), 1029–1044. doi:10.1002/asi.22815

Dabbish, L.A., & Kraut, R.E. (2006). Email overload at work: An analysis of factors associated with email strain. In Proceedings of the 20th Anniversary Conference on Computer Supported Cooperative Work (pp. 431–440). Banff, Alberta, Canada: ACM. doi:10.1.1.83.9750

Dabbish, L.A., Kraut, R.E., Fussell, S., & Kiesler, S. (2005). Understanding email use: predicting action on a message. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 691–700). Portland, Oregon: ACM. doi:10.1145/1054972.1055068

Eppler, M.J., & Mengis, J. (2004). The concept of information overload: A review of literature from organization science, accounting, marketing, MIS, and related disciplines. The information society, 20(5), 325–344. doi:10.1080/01972240490507974

Fisher, D., Brush, A.J., Gleave, E., & Smith, M.A. (2006). Revisiting Whittaker & Sidner's "email overload" ten years later. In Proceedings of the 20th Anniversary Conference on Computer Supported Cooperative Work (pp. 309–312). Banff, Alberta, Canada: ACM. doi:10.1145/1180875.1180922

Gwizdka, J. (2004). Email task management styles: The cleaners and the keepers. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 1235–1238). Vienna, Austria: ACM. doi:10.1145/985921.986032

Jones, Q., Ravid, G., & Rafaeli, S. (2004). Information overload and the message dynamics of online interaction spaces: A theoretical model and empirical exploration. Information Systems Research, 15(2), 194–210.

Kalman, Y.M., & Rafaeli, S. (2005). Email chronemics: unobtrusive profiling of response times. In 38th Hawaii International Conference on System Sciences (pp. 108b). Big Island, HI: IEEE. doi:10.1109/HICSS.2005.231

Kalman, Y.M., & Rafaeli, S. (2011). Online pauses and silence: Chronemic expectancy violations in written computer-mediated communication. Communication Research, 38(1), 54–69. doi:10.1177/0093650210378229

Kalman, Y.M., Ravid, G., Raban, D.R., & Rafaeli, S. (2006). Pauses and response latencies: A chronemic analysis of asynchronous CMC. Journal of Computer Mediated Communication, 12, 1–23.

Lee, R.M. (2000). Unobtrusive methods in social research. Buckingham, United Kingdom: Open University Press.

Stephens, K.K., Cho, J.K., & Ballard, D.I. (2012). Simultaneity, sequentiality, and speed: Organizational messages about multiple-task completion. Human Communication Research, 38(1), 23–47. doi:10.1111/j.1468-2958.2011.01420.x

Szóstek, A.M. (2011). "Dealing with my emails": Latent user needs in email management. Web 2.0 in Travel and Tourism: Empowering and Changing the Role of Travelers, 27(2), 723–729. doi:10.1016/j.chb.2010.09.019

Tyler, J.R., & Tang, J.C. (2003). When can I expect an email response? A study of rhythms in email usage. In ECSCW 2003. Retrieved from http://www.hpl.hp.com/research/idl/papers/rhythms/index.html

Whittaker, S., & Sidner, C. (1996). Email overload: Exploring personal information management of email. In Proceedings of the Conference on Human Factors in Computing Systems (pp. 276–283). doi:10.1145/238386.238530

Whittaker, S., Matthews, T., Cerruti, J., Badenes, H., & Tang, J. (2011). Am I wasting my time organizing email?: A study of email refinding. In Proceedings of the annual Conference on Human Factors in Computing Systems (pp. 3449–3458). Vancouver, British Columbia, Canada: ACM. doi:10.1145/1978942.1979457

Workweek and weekend. (2013). *Wikipedia, the Free Encyclopedia*. Retrieved from http://en.wikipedia.org/w/index.php?title=Workweek_and_weekend&oldid=582010582

DOI: 10.1002/asi