

If Erring is Human, is System Use Divine? Omission Errors During Post-Adoptive System Use

Abstract

Our study contributes to the research on human error during IS use by studying the antecedents of the omission errors that occur during routine instances of computerized work. While attention lapses have been identified as the main mechanism leading to omission errors, we still know little about how such lapses come about during post-adoptive system use. To address this limitation, we draw our theoretical insights from theories of attention and prospective memory to illustrate how the different forms of system use carry the potential to explain patterns of human error. Accordingly, we distinguish between two forms of use history that can consist of features that are either related or unrelated to the execution of a focal task and examine their effects on the frequency of omission errors. We also examine the interaction effects of task variation on the aforementioned relationship. Our hypotheses are tested by analyzing log data associated with the use of a newly introduced mobile application in the context of a sailing sports event. Our results indicate that restricting one's system use on related task features reduces omission errors, whereas a use history based on unrelated task features produces the opposite effects. Further, task diversity positively moderates the relationship between a use history of unrelated features and omission errors, but has no significant moderating effect on the relationship between a use history of related features and omission errors. Our findings hold a number of implications for the literature on human error, and these are discussed alongside with the implications of our study for practitioners and system design.

Keywords: Human error, omission error, attention, use history, task variation, prospective memory, post-adoptive system use, mobile application, field study

1 Introduction

The title of this paper, which paraphrases Alexander Pope's famous quote, "To err is human, to forgive is divine", illustrates the main aim of this paper, viz., an inquiry into the antecedents of human error in the context of routine computerized work. Research on human error has a long-standing tradition, and several works have examined instances and the implications of committing errors during computerized work (Brodbeck et al. 1993; Galletta et al. 1993, 1996; Zapf et al. 1992, 1994). However, we still know little about the antecedents of human error during IS use, and our paper sets out to make a contribution in this direction.

Understanding the conditions under which errors occur during IS use is of paramount importance, because of the high impact that 'small' errors can carry (Carlo et al. 2012). The recent outage in Amazon Web Services (AWS) that originally occurred because one employee incorrectly typed a command caused disruption to thousands of customers¹. Second, 'small' errors carry significant economic and behavioral consequences: users on average spend 10% of their working time correcting their errors (Brodbeck et al. 1993; Zapf et al. 1994). Last, committing errors while using computers has also been associated with negative emotions such as stress and frustration (Frese et al. 1991; de Vries et al. 2003; Zapf et al. 1992).

In this study, we focus on omission errors, which is the most frequent form of error (Love et al. 2009; Reason 2002). Omission errors are typically attributed to some form of attention capture (Reason 1984; Reason 1990). In order to examine how attention failures occur and lead to error during IS use, we are driven by the premise that human error cannot be understood without understanding action, as error is the byproduct of the same cognitive system that produces 'correct' actions (Booth 1991). From this standpoint, we argue that an enquiry into *how patterns of system use are formed* carries the potential to explain why errors occur.

As an indicative measure of the different patterns of IS use, we focus on the concept of use history, which is defined as the accumulated use of a basket of features that are available in a

¹ <https://www.nbr.co.nz/article/xero-instagram-other-services-hit-widespread-aws-outage-ck-200134>

system and are used to accomplish specified tasks (Jasperson et al. 2005; Sun 2012). Most importantly, we extend the concept of use history by distinguishing between use histories that are either related or unrelated to the execution of a focal task. Last, given the acknowledged importance of task characteristics in terms of impacting attention (Randall et al. 2014; Sanjram and Khan 2011), we further examine the moderating effect of task variation between the two forms of use history and the omission errors committed by IS users.

We test our hypotheses by examining the patterns of system use in a newly introduced mobile application in the context of a sailing sports event, namely the 2012 Kiel Week sailing event. Kiel Week involves about 5,000 sailors from 50 different nations and attracts approximately three million visitors every year, and is generally considered to be the world's largest sailing event². The mobile application introduced, named "Race Committee Cockpit" (RCC app), was developed for the purpose of facilitating the work of the event's race officers, who were responsible for monitoring the race conditions, as well as ensuring that race participants comply with the set of rules imposed by the International Sailing Federation (ISAF).

The remainder of the paper is organized as follows. The next section provides a synopsis of work on human error, followed by the presentation of our hypotheses and research model. Section four outlines the research methodology, whereas section five provides the empirical analysis and the results of the study. The final section discusses the broader implications of our enquiry into the antecedents of human error, along with some limitations of the study.

2 Related Research on Human Error

2.1 Defining Error and Understanding its Underlying Causes

The study of human error can be regarded as a discipline in its own right, as this topic can be considered to be as extensive as that covered by the term human performance (Reason 1990). Human error is defined as all the occasions in which a planned sequence of mental or physical

² <http://content.time.com/time/magazine/article/0,9171,2117251,00.html>

activities fails to achieve its intended outcome (Reason 1990). In this respect, errors involve either a departure from an intended course of action (or a path of actions) planned toward a desired goal, or a deviation from an appropriate behavior at work (Reason and Hobbs 2003).

Several taxonomies of human error have been developed in order to classify errors³. Our study sets out to obtain an understanding of the antecedents of error that occur during the execution of routine procedural tasks consisting of well-established, goal-oriented task sequences that are commonly performed during IS use. The errors that occur while executing routine task sequences are known as omission errors (Panko and Aurigemma 2010). An omission error is equivalent to a failure to recall the intention to carry out a task at the right time (e.g. being late to perform an intention), or instances where a necessary item is unwittingly omitted from a task sequence (Reason 2002). Omission errors primarily occur at the rule-based level of behavior and monitoring, where the composition of such a sequence of sub-routines in a familiar work situation is typically controlled by a stored rule or procedure (Rasmussen 1983; Rasmussen and Vicente 1989). Successful task execution at this level involves noticing an environmental cue that is associated in memory with a deferred intention (Dismukes 2006). As such, individuals need to keep the task goals in mind, given that the relevant cues occur while executing a task; moreover, task goals are not uniquely associated with intentions and must compete for retrieval with ongoing task goals (Loft and Remington 2010).

Keeping goal intentions actively maintained is key in terms of avoiding errors (McVay and Kane 2012). A goal intention comprises of a collection of active cognitive schemata, the activation of which depends on the periodic review of the respective intention (Reason 1984). In the absence of such a review process, the activation of cognitive schemata will gradually decline and lead to errors associated with executing a delayed intention (Sanjram 2013). In such cases, internal or external distractors may inappropriately capture ongoing cognition and attention and result in goal neglect errors and action slips (McVay and Kane 2012). Prior to outlining our theoretical foundations on how attention lapses might occur during IS use and

³ For a detailed review on error taxonomies, see Bolton (2017), and Hofmann and Frese (2011)

cause omission errors, we next provide a brief overview of existing IS works on human error.

2.2 Human Error in IS Research

Several works in the IS domain have enquired into the topic of human error. Certain studies have explored the effects of system characteristics on the errors committed by a system's users (Goswami et al. 2008; Lazar et al. 2007), whereas other works have explored human error under the prism of training users to commit fewer errors. An interesting debate that has emerged from this literature concerns the connotations associated with committing errors: a number of training studies have sought to identify mechanisms of reducing errors by preventing them from happening, as errors have been conceived to be frustrating and anxiety provoking, and thus disrupting to individual performance (Brodbeck et al. 1989). Other studies have stressed the positive effects associated with committing errors, especially with respect to exploring the different features of a system and breaking negative habits related to the use of a system (Frese et al. 1991; Keith and Frese 2008).

Instances of human error have also been explored in other domains of IS research: Sein and Santhanam (1999) explored the mechanisms through which committing errors can foster users' learning patterns and the effects of training in this relationship. Recent conceptual works have also attempted to stress the importance of the concept of human error in the IS usage research. For instance, Burton-Jones and Grange (2013) highlighted the role of human error in terms of the relationship between effective IS use and performance, by arguing that effective use can improve the effectiveness and efficiency of a system's use by reducing the errors committed and also by improving error recovery.

3 Theoretical Foundation

3.1 An Attention-Based View of Individual Behavior

Studies on human error converge on the finding that instances of omission errors can be attributed to attention lapses (Norman 1981; Reason 2000; Reason 2002). Attention refers to

“the taking possession by the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought” (James 1890, p. 403–404). Different types of attention have been identified, including selective attention, attentional vigilance, and executive attention (Hoffman and Ocasio 2001; Ocasio 2011). According to Ocasio (2011, p.1287), “selective attention concerns the process by which individuals focus information processing on a specific set of sensory stimuli at a given point in time, whereas attentional vigilance describes the process by which individuals are able to sustain concentration or focus on a particular stimulus (e.g. waiting for a signal to occur). Last, executive attention involves allocating controlled cognitive resources in working memory to information independent from incoming sensory data. It enables individuals to process multiple goals quasi-simultaneously by switching back and forth between different stimuli, including directly observed stimuli and stimuli stored in memory, and bringing them together in working memory”.

The ability to devote attention towards the execution of a task is guided by a number of cognitive control capabilities, defined as “the supervisory cognitive mechanisms through which individuals monitor and control their own attention and cognitive processes” (Laureiro-Martinez 2014, p. 1114). What is also important to note is that individuals’ attentional capabilities are bounded because humans have a limited information processing ability and that the number of stimuli that can be attend to are limited (Laureiro-Martinez 2014; Pashler 1999). It goes without saying that attentional deficits have been associated with negative behavioral and performance outcomes, a notable example of which includes human error (Kahneman 1973).

3.2 Sustaining Attention During Task Execution – A Theory of Prospective Memory

Within the context of task execution, the role of attention is critical for two reasons: first, successful task execution requires individuals to dedicate attention towards monitoring environmental stimuli in order to draw memory associations and perform an upcoming task. Second, in order to execute of an intended action, attention needs to be shifted to the execution of an ongoing task (Sanjram 2013). The latter is contingent upon an individual’s ability to maintain attentional control or to recover access to environmental stimuli or stimulus or goal

representations if these are outside an individual's conscious focus (Engle & Kane, 2004).

Individuals tend to combine mechanisms for conscious monitoring for environmental cues and automatic retrieval (Dismukes 2006). The former would involve an attention-demanding executive control system that would encode the association between the external event pertinent to the intended action and the intended action itself (McDaniel and Einstein 2000). In case of the latter, a target event automatically brings to mind the intended action. This system is assumed to support conscious recollection when an external cue automatically interacts with previously encoded actions stored in memory (McDaniel and Einstein 2000). Automatic retrieval mechanisms are increasingly used as individuals gain experience with a task and task execution becomes automatic. As Dismukes (2006, p.910) notes, "It would be uncommon for an experienced pilot to arrive at work thinking "I will lower the landing gear today when I turn onto final approach" (and it would be rather alarming if a pilot found this necessary)".

The extent to which attention-demanding or automatic mechanisms are in operation largely depends on the characteristics of the tasks an individual has to perform, cue quality and strength, as well as the properties of the ongoing activity (McDaniel and Einstein 2000). In sum, the more conscious mechanisms are used for cue monitoring, the higher the attentional demands placed on an individual. Accordingly, lapses can occur when individuals face either external or internal distractions, and their ability to maintain executive control becomes compromised (Casner and Schooler 2015; Randall et al. 2014). Attention lapses can also occur when an individual's limited amount of cognitive resources needs to be devoted to an increased number of sources at the same time (Smallwood and Schooler 2006; Thomson et al. 2014).

Taking stock of these insights, our survey of the literature on human error during IS use revealed that even if the link between attention lapses and omission error is acknowledged (Sanjram and Khan 2011), studies have done little in terms of showing how such lapses come about. Our main proposition is that enquiring into the diverse forms of IS use can help better understand how attention lapses and omission errors occur. This position is strengthened by several studies showing that research into how systems are used can explain performance

outcomes (Burton-Jones and Straub 2006). As Burton-Jones and Grange (2013, p. 640) argue, " ...There is always the potential that users may overinvest, underinvest, or misdirect their efforts in creating and using systems... although the creation and use of information systems can improve over time, the process is likely to be never ending and error prone".

In order to address this topic, and at the same align our study with works arguing that individual experience is a key determinant of human error (Frese and Keith 2015; Sanjram and Khan 2011), we focus on two distinct forms of system use that evolve over time; a use history that consists of related or unrelated features. We propose that these two forms of system use produce diverse attentional requirements, and accordingly differentially impact the frequency of omission errors that users commit during system use. We explain these concepts in the next section, where we outline our proposed model.

4 A Proposed Model on Omission Error During Post-Adoptive System Use

In this section, we outline our research model and hypotheses. In particular, we examine how, 1) the diverse patterns of system use as they evolve over time, and 2) the amount of variation in the tasks that users have to perform, can potentially explain the frequency of omission errors that users commit, given the strain that these two factors place on an individual's attentional ability. By pursuing 1), we join the call by Benlian (2015) for additional research on how the changing patterns of system use over time can influence performance outcomes. Accordingly, our choice to study the effects of task variation on the frequency of omission errors that users perform is grounded upon existing studies showing that task characteristics significantly impact the propensity for human error (Bolton 2017). We explain this argument in more detail in the following sub-sections, where we outline our research hypotheses.

4.1 Related and Unrelated Feature Use History

Our main hypothesis is that the cumulative IS use history (and the diverse patterns that use histories can take) can explain the omission errors that users commit by producing diverse attentional requirements. The concept of use history evolves at the post-adoption stage, where

users actively choose to explore, adopt, use and possibly extend one or more of a system's features (Jasperson et al. 2005). The latter constitutes one's features in use, and essentially refers to "the basket of system features that are ready to be used by a user to accomplish tasks. System features that do not belong to one's features in use include the features of a system that remain unused, such as the features that are unfamiliar or unknown" (Sun 2012, p. 455). Users are in many instances faced with these options, as systems (e.g. word processors and spreadsheets) have many more features than those mandated for work accomplishment (Jasperson et al. 2005). The decision of which features of a system are used can exceed the mandatory use of a system, where users are required to use specific features of an IS in order to execute their tasks (Jasperson et al. 2005).

While users decide which system's features become parts of their basket of features in use, they gain experience with what was initially a novel behavior, and engage less frequently in reflective consideration of this behavior and rely instead on previous patterns of behavior to direct their future behavior (Jasperson et al. 2005). As users routinely apply any IS feature within their work context, the ever-accumulating prior-use experiences will imprint these use behaviors within their cognitive scripts and direct them towards task execution. Accordingly, an individual's past behavior will form his/her use history, which is defined as "... a collective, systematic account of an individual's prior use of an IT application and its features... and ... learned situational-behavior sequences with respect to an IT application and its features that have become automatic" (Jasperson et al. 2005, p. 542).

The concept of use history has been successful in terms of overcoming some of the shortcomings of earlier measures of system use (e.g. the frequency, duration, and the different system functionalities that are used) that have been considered to be too simplistic to capture the relationship between system use and its resulting outcomes (Benbasat and Barki 2007; Sun 2012). This has been achieved through the integration of the diverse features of a system that actively form a user's basket of features in use with the evolution of use over time. In short, it has been argued that examining post-adoptive behavior at a feature level of analysis can

provide insights into why users develop diverse patterns of feature use and, as a result, extract differential value from an IS (Jasperson et al. 2005).

Notwithstanding the promise of the construct of use history in terms of explaining behavioral and performance outcomes (omission errors in our case), the related research on the features in use has yielded inconclusive results with respect to user performance. For instance, it has been argued that expanding one's basket of features in use is a form of exploratory behavior that enhances one's knowledge and mastery of a system's features that enhances individual performance (Sun 2012). An extended number of features in use also leads to other positive outcomes, such as an increased sense of autonomy and cognitive stimulation from the work environment (Gill 1996). In contrast, other studies have shown that a higher number of features in use does not necessarily lead to performance increases; the adoption of additional features can take place in nonproductive ways, or similarly users may be overwhelmed by the presence of too many features, resulting in an inability to choose among feature sets or to apply the features effectively in their work (Jasperson et al. 2005; Silver 1990).

We address these seemingly contradicting results by distinguishing between a related feature use history (RFUH) and an unrelated feature use history (UFUH). The difference between the two constructs is that RFUH includes all those features in use that are necessary for the execution of a focal task, whereas UFUH includes the use of features that are not directly related to the successful completion of this task, but can nevertheless exist in a system and form a user's basket of features in use (and one's use history). To offer an example of UFUH, while typing this paper on a Microsoft Word processor, one of the authors clicked on the 'SmartArt' feature and started experimenting with different options, even though the task of writing this paper did not require the use of this feature in any way. In short, the cumulative use of features that are unrelated to the execution of a task (even though the use of this feature might have been required in the execution of a different task) constitutes UFUH. It can therefore be inferred that what counts as UFUH is dictated by the nature of the task that has to be performed, hence what is as an unrelated feature in use could be considered as a related

feature in use at another point. This distinction can be particularly salient in work environments that involve the execution of routine tasks, where the potential of adapting the different features in use in order to exploit the benefits of explorative behavior makes little sense, given that successful task execution is performed in standardized, pre-specified ways.

Given that task execution at the rule-based level of regulation tends to become automatic over time (Dismukes 2006; Sanjram 2013), we essentially propose that a use history based on related features will lead to fewer omission errors, whereas a use history of unrelated features will produce the exact opposite results. RFUH implies higher attentional affordances that can be beneficial in the context of task execution, as the increased ability to maintain focus on task execution will lead to a less error-prone behavior (Aggarwal and Woolley 2013). This is because cues associated with an upcoming task will be able to produce a stronger association with existing memory traces of prior system use that have been based solely on task execution. In this respect, users do not have to ‘consume’ significant attentional resources to switch back and forth between effortful and automatic forms of monitoring for environmental cues. Similar to a routinized behavior that frees up mental resources and allows users to rapidly process information with little effort (Bargh and Chartrand 1999; Cohen et al. 1996; Laureiro-Martinez 2014), RFUH will lead to fewer errors due to the lower attentional demands that it creates. Hence, we hypothesize:

HYPOTHESIS 1 (H1). RFUH is negatively related to the number of omission errors committed by a system’s users.

In contrast, we expect that a use history based on unrelated features will produce the inverse effects. UFUH essentially refers to cumulative system use that occurs during task execution and is unrelated to it. As UFUH is likely to occur because users want to actively experiment with a system or just ‘play around’, it is expected to draw many linkages to the concept of mind wandering, which includes situations where “executive control shifts away from a primary task to the processing of personal goals that are unrelated to the focal task, and occurs without intention or even awareness that one’s mind has drifted” (Smallwood & Schooler, 2006, p.

946). Omission errors are known to occur during such instances of mind wandering (Casner and Schooler 2015), and recent studies have shown that the link between mind wandering and human error becomes even stronger when taking into account the detrimental effects of mind wandering when it accumulates over time (McVay and Kane 2012; Thomson et al. 2014). In the case of UFUH, users will have to resort to more attention-demanding forms of cue monitoring. This is because the link between cues for upcoming tasks and memory traces will be weaker, given that an increased number of unrelated features will form part of users' memory scripts. As users will also have to dedicate attentional resources to ongoing task execution, as well as to recovery from task-unrelated thinking, we expect UFUH to lead to a higher number of omission errors. Accordingly, we hypothesize:

HYPOTHESIS 2 (H2). UFUH is positively related to the number of omission errors committed by a system's users.

4.2 Task Variation

Task characteristics are generally defined as the 'real world' dimensions that relate to the physical nature of a stimulus (Wood 1986). Research into the different aspects of task characteristics (e.g. complexity, variety, autonomy, feedback, identity and significance) has widely exhibited how the different task features can influence individual and group productivity/performance (Fuller and Dennis 2009, Staats and Gino 2012), as well as users' learning patterns (Mukhopadhyay et al. 2011; Narayanan et al. 2009; Schilling et al. 2003).

In the context of human error, task characteristics (e.g. task complexity) have been found to affect both the frequency of errors, as well as the effectiveness of the different error recovery strategies (Chung and Byrne 2008; Galletta et al. 1993; Goswami et al. 2008; Sein and Santhanam 1999). Even if in cases of task execution that take place at the rule-based level of cognitive control and imply a lower level of complexity, the aforementioned findings regarding task complexity are not applicable (Zapf et al. 1992), other task features, in particular the

diversity of the tasks that have to be executed, can shed some light on the errors committed by a system's users at this level of cognitive control.

The concept of task variation, namely the frequency of diverse activities that occur *while performing* a particular task, has gained increased attention in terms of understanding its dynamics with respect to behavioral outcomes (Narayanan et al. 2009, Schilling et al. 2003, Staats and Gino 2012). To offer a better understanding of the concept of task variation, at a general level, while most activities comprising a task can significantly overlap with those of other tasks, workers still engage in high levels of cognitive activity, presumably due to variation in work content across tasks, such as differences between consulting projects, surgeries, or legal cases (Avgerinos and Gokpinar 2018). In line with related works that examined the effects of task characteristics on IS success (e.g., Sharma and Yetton 2007), we hypothesize that task variation will moderate the relationship between use history and omission errors.

Multitasking is widely known to produce additional attentional requirements (Kahneman 1973). In the case of task variation, users will tend to encounter diverse situations, which will require a higher amount of information processing. Moreover, in such cases preplanning tends to be difficult and thus leads to a greater need for acquiring information on an ongoing basis (Karimi et al. 2004). The more absorbing an ongoing activity is, the less likely that resources will be available for attention-demanding approaches to prospective remembering or that subjects will be able to successfully deploy strategic approaches to cue monitoring (McDaniel and Einstein 2000). Accordingly, tasks requiring more focus will leave fewer cognitive resources available for task-unrelated activities (Smallwood & Schooler, 2006). Engaging in the latter under higher task variation should result in larger performance decrements (e.g. errors), because they require more active cognitive control (Kanfer and Ackerman 1989).

Translating these insights into our study, we expect that conditions of higher task variation will place a bigger strain on a user's limited cognitive resources, given that under conditions of higher task variation successful performance will require users to divide attention among competing task activities. Therefore, users who maintain and enhance a use history of

unrelated features are likely to commit more omission errors under task variation. The logic behind this builds on our previous hypothesis: As a use history of unrelated features is in itself attention-consuming, the likelihood of committing omission errors under conditions of higher task variation will be even higher, as users will be less likely to maintain attentional control, given their limited amount of cognitive resources (Smallwood and Schooler 2006).

Following the same logic, a use history of related features is also likely to be negatively affected by an increased level of task variation, as the increased workload will direct users into the more effortful and resource-consuming patterns of cue monitoring. Under such conditions, users' ability to maintain attentional control might also be compromised, even if such users have a higher 'attentional buffer' due to maintaining a use history of related features. In sum, our third and fourth hypotheses are the following:

HYPOTHESIS 3 (H3). Task variation will moderate the relationship between RFUH and the number of omission errors that are committed by a system's users, such that the negative relationship between RFUH and the number of omission errors will be weaker when users have to perform a wider variety of tasks.

HYPOTHESIS 4 (H4). Task variation will moderate the relationship between UFUH and the number of omission errors that are committed by a system's users, such that the positive relationship between UFUH and the number of omission errors will be stronger when users have to perform a wider variety of tasks.

5 Research Methodology

In this section, we first describe the research setting of our study. This is followed by an overview of our sources of empirical data, which include a combination of hand-written protocol data and log data that were derived from the usage of the RCC app that was deployed in the 2012 Kiel Week sailing event that took place between 16-24 June 2012. The last part in this section presents the operationalization of the variables in our model.

5.1 The 2012 Kiel Week Sailing Event and the Deployment of the RCC App

In our attempt to identify omission errors during IS use, we collected data from the field deployment of a tablet application, called the Race Committee Cockpit application (RCC app). The RCC app incorporates the ISAF sailing rules⁴ and, among others, allows the race officers to schedule races, communicate start violations, manage the flags, and end races. To further ensure user acceptance of the new system, the race officers of 2012 Kiel week were involved in the iterative design and development processes of the RCC app. During the design and development processes, screenshots and early prototypes were demonstrated to them on a weekly basis. Prior to the 2012 Kiel Week sailing event, there was a preparatory sailing event where the application developers joined the race committee boat, and operated the RCC app to show how the RCC app should be used. In the evening before the first day of Kiel Week, there was a briefing session with all race officers. Besides communicating instructions on how to use the RCC app, the race officers were asked to practice using the app with 'dummy' races.

During the nine days of Kiel Week 2012, 360 races were conducted in 8 different race courses, with each race course including a number of races taking place in parallel. For each of the race courses, one dedicated team of race officers located on a boat next to the race course was responsible for monitoring and refereeing the races. The head of the team was responsible for refereeing the races and communicating with members at shore via a handheld transceiver (VHF radio). One member was responsible for recording race events to the RCC app, which would then be transmitted via cellular network to an information system on the shore and broadcasted live onto the Internet; and two other members of the team were responsible for independently handwriting the race events into pre-defined forms, the so-called race protocols.

5.2 Data Set: App Log Data and Protocol Data

Our first source of data is the hand-written race protocol data (PD) that served as the official description of events that took place in each race. In case of official complaints by one of the

⁴ <http://www.sailing.org/tools/documents/ISAFRRS20132016Final-%5B13376%5D.pdf>

contestant teams, PD would be examined to respond to the complaints. Hence, we considered PD as our reference data. PD informed us of the actions/tasks that *had to be performed* in the RCC app. Our second source of data included data that was derived from the use of the RCC app. Each action (equal to a click) on the RCC app was logged and stored on the tablet device. Accordingly, the app log data (LD) recorded all the actions that *were performed* during the races. The use of the RCC app and accordingly the recorded LD actions also mirror tasks that had to be performed; RCC users were encouraged to use the app as accurately as possible, given that the use of hand-written protocols would gradually be substituted by the RCC app (in fact, Kiel Week 2015 was the first event where race monitoring was exclusively based on the RCC app). The comparison between PD and LD allowed us to classify the actions recorded in LD to: use of related features, use of unrelated features, and omission errors.

Some error types are detected from mismatches between LD and PD (e.g., the necessary actions indicated in PD are not evident in LD). In this case we assume that PD serves as the valid reference point and any mismatches in LD indicate errors. One could question this, in the sense that PD could also be a source of errors. We nevertheless maintain PD as a more objective 'mirror' of the tasks that had to be performed, firstly because PD was used as the reference point for official complaints, and second because all of the errors that we counted depended on actions that had to be performed according to PD; given the nature of tasks that race officers had to perform, we consider it unlikely that false actions were reported in PD.

In total, 360 races were monitored by eight race officers, who were named after the code words of maritime alphabet: Alpha, Charlie, Delta, Echo, Foxtrott, Golf, Juliett and Kilo. Out of the 360 races, 230 races could be linked to complete PD access and be translated into a digital format, and were consequently included in the data set. During the 230 races, a total of 3,439 actions were recorded that were classified as related and unrelated feature uses. During a majority of the races (144 races), each race officer had to administer at least two races in parallel. These 144 parallel races that corresponded to 2,381 actions (related and unrelated feature uses)

were used for the analysis. Our analysis was limited to 144 races, as the computation of some of our variables required the administration of at least one parallel race.

5.3 Operationalization of the Variables

The dependent variable in our study is the number of omission errors committed in each race, which are classified as errors (ERROR). The main variables in our study are: related feature use history (RFUH), unrelated feature use history (UFUH), and task variation (TV).

$ERROR_{ij}$ is the number of omission errors in race j committed by the race officer i . $O_{ERROR_{ij}} \in [0,9]$ consists of the following errors:

$error_1$: The running flag was not set.

$error_2$: The running course was not set.

$error_3$: The XRAY flag was set, but was never unset.

$error_4$: The BLUE flag was set, but was never unset.

$error_5$: According to PD, the race had ended, but the BLUE flag was not set.

$error_6$: The BLUE flag was set and immediately unset. This means the user noticed that he/she did not properly set the BLUE flag, and corrected it by firstly setting the BLUE flag, and then unsetting the BLUE flag again.

$error_7$: According to PD, a race started at time t , but the start of the race in LD was recorded at time $t + n$. To account for the possible processing delay of the RCC app and after consulting with the race committee officers, we only count this error if $|n| > 90$ seconds.

$error_8$: According to PD, the first boat crossed the finish line at time t , but the event in LD was recorded at time $t + n$. To account for the possible processing delays and after consulting with the race committee officers, we only count this error if $|n| > 60$ seconds.

$error_9$: According to PD, the last boat crossed the finish line at time t , but the event in LD was recorded at time $t + n$. To account for the possible processing delays and after consulting with the race committee officers, we only count this error if $|n| > 60$ seconds.

Our pool of omission errors includes both ‘pure’ omissions (*error1-error5*), where a necessary item is unwittingly omitted from a task sequence, as well as an additional set of omission errors (*error6-error9*), where the intention to carry out an action is not recalled at the right time (Reason 2002). For errors 7-9, we had to decide on a decide on an acceptable minimum amount of delay that would count as an omission error beyond reasonable doubt. To achieve this, we consulted the Race Committee Officers, as well as the Head of the Racing Committee whose role was instrumental in the design and rollout of the RCC app. Examples of errors derived from LD and PD are shown in Figure 1. This example shows a correct setting of the PAPA flag (marker 1), but there was an error in scheduling the race (*error2*; the actual starting time according to PD is 17:00, but it was scheduled in LD to start at 17:05) (marker 2), and there was a missing action of unsetting the BLUE flag (*error4*) (marker 3).

meaning	logNR	timestamp	race_identifier	course	
RACE_SELECTED_ELEMENT	5	2012-06-18 16:47:33	16062012-0000-0000-0000-000000000025	Alpha	
RACE_SET_TIME_WITH_TIME	40	2012-06-18 16:47:47	16062012-0000-0000-0000-000000000025	Alpha	
RACE_SET_TIME	16	2012-06-18 16:47:47	16062012-0000-0000-0000-000000000025	Alpha	
RACE_START_PHASE_FLAG_PAPA	206	2012-06-18 16:48:07	16062012-0000-0000-0000-000000000025	Alpha	
RACE_SELECTED_ELEMENT	5	2012-06-18 16:48:07	16062012-0000-0000-0000-000000000025	Alpha	
RACE_START_PHASE_BUTTON_RESET_TIME	15	2012-06-18 16:48:43	16062012-0000-0000-0000-000000000025	Alpha	
RACE_RESET_TIME	17	2012-06-18 16:48:54	16062012-0000-0000-0000-000000000025	Alpha	
RACE_START_PHASE_FLAG_PAPA	1	206	2012-06-18 16:48:54	16062012-0000-0000-0000-000000000025	Alpha
RACE_SELECTED_ELEMENT	5	2012-06-18 16:50:27	16062012-0000-0000-0000-000000000025	Alpha	
RACE_STARTED	2	208	2012-06-18 17:05:00	16062012-0000-0000-0000-000000000025	Alpha
RACE_SELECTED_ELEMENT	5	2012-06-18 17:33:28	16062012-0000-0000-0000-000000000025	Alpha	
RACE_SELECTED_ELEMENT	5	2012-06-18 17:33:36	16062012-0000-0000-0000-000000000025	Alpha	
RACE_RUNNING_FLAG_BLUE_SET	3	10	2012-06-18 17:36:54	16062012-0000-0000-0000-000000000025	Alpha

Figure 1: Examples of errors derived from LD and PD

The set of observed actions in the RCC app that was performed by race officer *i* on race *j* can be classified as related feature use (RFU), unrelated feature use (UFU), and *ERROR*. Let a_{ij} denote an action performed by race officer *i* on race *j*, and a classifier $class(a) \in \{RFU, UFU, ERROR\}$. Previously, we explained the list of actions coded as errors that define our dependent variable. The list of actions coded as RFU and UFU are summarized in Appendix A. We first created this list and split actions into related and unrelated features. In order to avoid any mistakes with the operationalization of actions into related and unrelated actions, we checked and confirmed our operationalization of the two variables with the Head of the Race Committee Officers, who was instrumental in the design and rollout of the RCC

app. Following the consultation process, we did not make the operationalization of the variables, as we had a 100% overlap with the Head of the Race Committee⁵.

Let $A_{ij} = \{a_{ij}: class(a_{ij}) \neq ERROR\}$ denote the set of related and unrelated features in use during race j on course i . The set of prescribed actions consists of a set of 4 necessary actions derivable from PD (see Appendix B). Let pa_{ij} denote a prescribed action performed by race officer i on race j , and let PA_{ij} denote the set of prescribed actions during race j on course i .

Related Feature Use History (RFUH)_{ij} is defined as the cumulative number of related features in use in the RCC app by race officer i , before race j . That is $RFUH_{ij} = \sum_{l=1}^{j-1} |\{a_{il} \in A_{il}: class(a_{il}) = RFU\}|$.

Unrelated Feature Use History (UFUH)_{ij} is defined as the cumulative number of unrelated features in use in the RCC app by race officer i , before race j . That is $UFUH_{ij} = \sum_{l=1}^{j-1} |\{a_{il} \in A_{il}: class(a_{il}) = UFU\}|$.

Task Variation (TV_{ij}) is defined as the amount of variation in the prescribed actions between race j (PA_{ij}) and its corresponding parallel races. A detailed explanation of how we computed task variation can be found in Appendix C.

Control Variables

In addition to the dependent and independent variables, we also measured several control variables that we deemed important. These are: 1) the external environmental conditions in terms of the wind strength, 2) the ‘normality’ of a race, and 3) the individual-specific effects of the different race officers.

Wind strength (WIND) was measured for each race and documented in PD. Measurement was done using a wind measurement device on the race committee boats. The unit of wind is knot.

⁵ Naturally, what counts as a related/unrelated feature has to be seen in the context of the sailing event; using a different artifact in another setting would have a different ‘basket of features in use’, as well as related and unrelated features.

Since the race officers using the RCC app were on a boat offshore, the environmental condition in terms of the wind strength may have an impact on user performance, given that external distractions can potentially cause interruptions and increase the frequency of errors (Speier et al. 1999), and are also known to impact the propensity of experiencing attention lapses and consequently performing omission errors (Casner and Schooler 2015).

Race Flag (FLAG): In 'normal' race situations, race officers start with raising the PAPA flag. In 'special' situations, such as restarting an aborted race, the race officers may start the race with other flags, such as the BLACK flag. We thus include the race flag as a dummy variable for controlling for the 'normality' of the race (i.e., FLAG=0 -> PAPA, FLAG=1 -> other than PAPA). The 'normality' of the race may therefore affect the race officers' tendency to commit errors.

Race Officer (OFFICER): To account for the individual differences in cognitive resources (e.g. general mental ability and working memory capacity) that are known to affect the propensity of experiencing attention lapses (Kane and Engle 2003) and conducting omission errors (Randall et al. 2014), we included each race officer as a control variable. Those are: ALPHA, CHARLIE, DELTA, ECHO, FOXTROTT, GOLF, JULIETT, and KILO. The number associated with each officer refers to the frequency of parallel races that he/she had to administer throughout the sailing event. For the analysis, officer ALPHA was used as the baseline.

6 Data Analysis and Results

Before listing the correlation and the descriptive statistics for the variables in our model (Tables 1 and 2), we would like to make a note on the impact and the magnitude of errors in our study: the number of omission errors performed on average per race was 2.43 (St. Dev. = 1.18). To offer a sense of the relative impact of omission errors, the average number of related features in use (RFU) that RCC users performed per race was 10.2 (St. Dev. = 6.03). Also, every error counts; given the criticality of capturing the race accurately in order to avoid and/or address formal complaints from the contestants of the sailing event, no error can be taken lightly.

We also conducted robustness checks for our model using errors 1-5 and errors 6-9 respectively. We obtained similar results as in our main model, with the only difference that the interaction effect between TV and UFUH in the model with errors 1-5 as the dependent variable was not statistically significant. We attribute this to the unbalanced frequency of errors in the two models, given that errors 1-5 did not occur in 52 of the 144 races, whereas errors 6-9 did not occur in only 17 of the 144 races.

Continuous Variables				
	Min	Max	Mean	Std.Dev
ERROR	0	6	2.43	1.18
WIND	4.5	27.0	14.13	4.00
RFUH	0	402	207.45	90.14
UFUH	0	585	191.42	150.63
TV	0	1	0.37	0.26
Categorical Variables				
Variable	Frequency		%	
FLAG-0	96		66.7	
FLAG-1	48		33.3	
COURSE-ALPHA	3		2.1	
COURSE-CHARLIE	25		17.4	
COURSE-DELTA	23		16.0	
COURSE-ECHO	20		13.9	
COURSE-FOXTROTT	16		11.1	
COURSE-GOLF	32		22.2	
COURSE-JULIETT	21		14.6	
COURSE-KILO	4		2.8	

Table 1: Summary of Descriptive Statistics

	ERROR	WIND	FLAG	OFFICER	RFUH	UFUH	TV
ERROR	1						
WIND	-0.091	1					

FLAG	-0.248	0.061	1				
OFFICER	0.191	0.142	0.336	1			
RFUH	0.055	0.001	-0.046	0.133	1		
UFUH	-0.114	0.010	-0.153	0.630	0.362	1	
TV	-0.115	0.212	-0.021	-0.154	0.202	0.033	1

Table 2: Correlation Statistics

Before proceeding with our hypotheses testing, we standardized the coefficients of the continuous variables for ease of comparison. To account for the count data in the dependent variable, we used a Poisson regression model. Allison and Waterman (2002) suggest using the negative binomial model as an alternative to Poisson in the presence of over-dispersion. Since there is no significant indication of over-dispersion (we are not able to reject the null-hypothesis of equi-dispersion at the 5% significance level), we proceed with a Poisson regression model⁶. Model 1 in Table 3 is without the interaction terms (H1 and H2), whereas Model 2 in Table 3 is with the interaction terms (H3 and H4). The values of variance inflation factor (VIF) for our continuous variables were not higher than 5, which means that no multicollinearity problem exists in any of the models (Cohen and Cohen, 1975). The regression results of all models are summarized in Table 4.

The data analysis shows that RFUH reduces omission errors whereas UFUH increases omission errors. Hence, H1 and H2 are supported. The standardized variables make it possible to compare the magnitudes between the coefficient estimates. As shown in Model 1, UFUH has larger effect on omission errors than RFUH. The effect of UFUH on omission errors is magnified by the degree of variation in terms of the parallel tasks that a race officer has to perform. Task variation (TV) however has no significant moderating effect on the relationship

⁶ Over-dispersion in Poisson regression models is present if the conditional variance is larger than the conditional mean (Allison and Waterman 2002). Statistical tests for overdispersion (Cameron and Trivendi 1990) are done against the null hypothesis of equi-dispersion (conditional variance equals conditional mean).

between RFUH and omission errors. Hence, only H4 is supported. In the next section, we discuss these findings in more detail.

	Model 1 DV: ERROR	Model 2 DV: ERROR
Control Variables		
WIND	-0.010 (0.060)***	0.015 (0.062)
FLAG-0 (baseline)		
FLAG-1	-0.166 (0.150)	-0.206 (0.153)
OFFICER-ALPHA (baseline)		
OFFICER-CHARLIE	-0.206 (0.407)	-0.125 (0.430)
OFFICER-DELTA	-3.323 (1.261)**	-3.482 (1.303)**
OFFICER-ECHO	-0.109 (0.395)	-0.064 (0.410)
OFFICER-FOXTROTT	0.546 (0.415)	0.532 (0.433)
OFFICER-GOLF	0.223 (0.430)	0.355 (0.445)
OFFICER-JULIETT	-0.814 (0.474)	-0.743 (0.485)
OFFICER-KILO	-0.511 (0.534)	-0.467 (0.566)
Main Variables		
RFUH	-0.399 (0.172)*	-0.432 (0.175)*
UFUH	1.227 (0.476)**	1.335 (0.489)**
TV		-0.064 (0.062)
RFUH:TV		-0.078 (0.068)
UFUH:TV		0.151 (0.061)*
AIC	474.46	473.71

***Significant at < 0.001; **Significant at < 0.01; *Significant at < 0.05; .Significant at <0.1

Table 3. Analysis Results

Hypotheses	Empirical Support
H1. RFUH is negatively related to the number of omission errors that are committed by a system's users.	Supported
H2. UFUH is positively related to the number of omission errors that are committed by a system's users.	Supported
H3. Task variation will moderate the relationship between RFUH and the number of errors that are committed by a system's users, such that the positive relationship between RFUH and the number of omission errors will be weaker when users have to perform a wider variety of tasks.	Not Supported
H4. Task variation will moderate the relationship between UFUH and the number of errors that are committed by a system's users, such that the negative relationship between UFUH and the number of omission errors will be stronger when users have to perform a wider variety of tasks.	Supported

Table 4. Summary of the Results

7 Discussion

In this section, we discuss the contributions of our study for the research on human error and post-adoptive system use. We also discuss the practical implications of our study in terms of user training and error management. Last, we elaborate on the limitations of our research and the possible extensions that can be made to our study.

7.1 Contributions to Research

Our study firstly contributes to the post-adoptive system use literature by empirically showcasing how different forms of system use can lead to diverse performance outcomes, and in particular into omission errors. We examined the antecedents of omission errors in the context of executing largely standardized, procedural task sequences. Omission errors occurring at this level have typically been attributed to attention lapses. Earlier studies have argued that in such settings, the extent to which such lapses occur can be attributed to the degree of user experience and the characteristics of the tasks that users have to perform (Sanjram and Khan 2011). We extended this argument by showing that omission errors are indeed dependent on individuals' cumulative use of a system, and as such we join other works highlighting the importance of better understanding the patterns of system use over time

(Benlian 2015). Most importantly, however, we show that cumulative system use can unfold in diverse ways and differentially impact the frequency of omission errors.

We drew our insights from the concept of use history (Jasperson et al. 2005; Sun 2012), and we argued that use histories can unfold in diverse ways and produce the inverse effects in terms of committing omission errors; a history of using a system according to features that are related to the execution of specific tasks leads to fewer omissions, whereas a use history that encompasses an extended number of system features leads to the exact opposite results. In this respect, we extend existing studies by arguing that what matters in terms of combatting omission errors during the execution of PM tasks is not the cumulative use of a system per se, but rather cumulative system use in a focused manner.

Additionally, our study shows that using an extended number of a system's features can be a double-edged sword: while a number of studies have highlighted the benefits of IS infusion (Fadel 2012; Kim and Gupta 2014), or that extended, innovative, or emergent use can be a means of fostering individual mindfulness and maintaining performance reliability (Butler and Gray 2006; Li et al. 2013; Sun 2012), we show that in the context of routine, procedural tasks that do not necessarily require devising innovative solutions, an extended use of a system's features can produce adverse effects on user performance. This is not to say that our study downplays any of the acknowledged positive effects of mindful use in terms of combatting error (Butler and Gray 2006). Given that mindfulness is equally about the quality of attention as it is about the conservation of attention (McAvoy et al. 2013), we argue that it is important to emphasize that innovative or extended use (an inherent component of mindful use) is also resource consuming and can be redundant in certain task contexts.

At the theoretical level, we argued that the key to successful task execution is noticing environmental cues, which in turn trigger intended actions by involving mechanisms that draw associations with existing memory (McDaniel and Einstein 2000). The process of cue monitoring can either involve more conscious, effortful and attention-demanding mechanisms, or it can involve mechanisms where such associations are done automatically. Automatic

retrieval becomes increasingly the norm as task execution becomes repetitive over time. Nonetheless, depending on a number of parameters (e.g. task characteristics, cue quality and strength), attention-demanding mechanisms can complement the more resource-free mechanisms. In sum, given the limited amount of cognitive resources that any individual is equipped with, the more conscious mechanisms are used for cue monitoring, the higher the attentional demands, and consequently the higher the probability of lapses and errors.

While this framework initially provides support to our findings, our study nevertheless provides some nuanced findings and makes some possible extensions: firstly, contrary to what we hypothesized, we did not find support for the moderating effect of task variation on the relationship between RFUH and omission error. While it is likely that users were able to maintain an 'attentional buffer' from this relatively resource-free form of use history, the negative coefficient of the moderating effect (albeit not significant) seems to suggest that the additional burden of task variation starts to consume the attentional 'buffer'. Perhaps more importantly, there is a need for better conceptualization and measurement of the attentional requirements of both ongoing and PM tasks.

Secondly, the literature on human error argues that it is important to examine the characteristics of the *prescribed* tasks that individuals have to perform in order to assess whether attention-demanding or automatic mechanisms of cue monitoring will be in operation. Our study shows that it is equally important to complement the study of ongoing task characteristics with an examination of the *actual activities* during task execution, given that the two forms of use history in our study lead to the exact opposite performance results. In this respect, our study contributes to the ongoing debate around the ways in which task variation impacts task performance (Avgerinos and Gokpinar 2018). In sum, our unique empirical setting and dataset that included a documentation of both prescribed and performed actions, enabled us to conduct such an enquiry and yielded some highly interesting results.

7.2 Implications for Practice

Our study holds several practical implications: First, we observe that providing users with initial training is simply not enough in terms of combating errors. In our study, users received training, not only in terms of obtaining detailed instructions about the procedures they had to follow prior to the real-time use of the system, but also in terms of having ample time to practice with the system on their own, thus resembling the two types of procedural and conceptual model training procedural (Santhanam and Sein 1994). As a number of users tend to disregard training manuals and similar support systems (Lazonder and van der Meji 1994), it is important for organizations to understand the conditions under which errors occur in situ. To this effect, our study shows that in the case of executing procedural tasks, particular attention should be paid on how systems are used over time. Effective error management techniques should therefore not only consist of regular monitoring of the ways in which use histories are formed, but also of targeted interventions in cases of deviance from patterns of related feature use.

It is important for such interventions to occur before a use history of unrelated features turns into routinized action. An additional implication of our study is that suitable mechanisms of sustaining attention also need to be identified, and most importantly including such elements into user training also seems imminent. In this respect, investing in training to enhance users' cognitive abilities (i.e. general mental ability and working memory capacity) seems to be a worthwhile undertaking. While research on the effects of training on working memory abilities has yielded inconclusive results, it has been acknowledged that such efforts make sense in the context of executing routine tasks⁷. Last, it is important to note that in cases where 'playing around' during system use cannot be avoided, it is best if such user behavior takes place under conditions where task variation is the lowest; even if high task variation might make users more alert (Sanjram and Khan 2011), an increasing use of unrelated features will undermine user performance, and make the shift to patterns of related feature use more difficult.

⁷ <http://www.the-scientist.com/?articles.view/articleNo/39768/title/Does-Brain-Training-Work/>

From a system design perspective, it is acknowledged that investing in system-based error correction mechanisms can be problematic, especially in the case of end user computing (Klein et al. 1997). Nevertheless, introducing features that sustain users' attentional abilities (e.g. through arousal techniques and/or regular reminders) appears to be a sensible option. Nonetheless, our results indicate that the most important thing in terms of combating omission errors is to redirect individuals from the use of unnecessary features, and most importantly to avoid that such patterns of usage become routinized over the course of time. As such, restricting system use during the earlier stages of system deployment to only related and mandatory features appears to be eminent in order to combat omission error, as users have limited knowledge of a system's features at this stage (Durcikova et al. 2011).

7.3 Limitations and Suggestions for Future Research

Our study also comes along with a number of limitations: First, we focus only on omission errors that take place at the level of rule-based regulation. While on the one hand we cover a broad spectrum of omission errors at this level of cognitive control, our study does not explore other classes of errors that occur at the knowledge and the skill-based level of cognitive regulation. Future studies can complement ours by examining whether and how the diverse forms of system use also affect the patterns of errors occurring in settings that involve the execution of more complex tasks. It could be the case that under such conditions, an extended basket of features in use that accumulates over time and encapsulates patterns of emergent and integrative use (Saga and Zmud 1993), might produce different results from ours. Second, our hypotheses were tested over a set period of time and in a sailing sport setting; future studies could firstly examine whether our inferences hold over a larger time span, given that patterns of system use are likely to change over the course of time (Burton-Jones and Gallivan (2007). Future studies could also test our hypotheses in more 'typical' settings of computerized work (e.g. in spreadsheet development), as computerized work in the context of sailing sports may produce certain peculiarities that compromise the generalizability of our findings.

Also, we strongly believe that the concept of attention will gain a more prominent position in the literature on post-adoptive system use. Testing the effects of the different aspects of usage (e.g. the duration and the scope of usage, the intention to use a system and the perceived ease of use) on users' attentional abilities and patterns of attention allocation can yield novel insights into how system usage affects both positive and negative aspects of performance.

Last, a large body of research in the area of mind wandering has shown that when individuals experience attention lapses, they tend to lose focus and the mind shifts into internal trains of thought that are unrelated to ongoing task execution, which leads to errors (McVay and Kane 2012; Randall et al. 2014, Smallwood 2013). Future research on post-adoptive system use could also benefit from examining the interplay between such internal trains of thought and the diverse patterns of system use. Such an enquiry can shed further light into how use histories are formed, and how positive and negative aspects of user performance are ultimately affected.

8 Conclusion

In this study we have sought to stress the need to study errors during post-adoptive system use, and we attempted to examine human error under the prism of *how* the information systems are actually used over time, while executing procedural tasks. Our analysis yielded some interesting results: when systems are used repeatedly in a focused manner, users tend to commit fewer errors. In contrast, when unrelated features become embedded in individuals' use history, the effects can be deleterious, even more so under conditions of high task variation. In sum, human error in system usage is a persistent problem in modern organizations, even in times where automation is more pervasive than ever before; given that humans are ultimately the end users of a system and those primarily responsible for error, our study contributes toward the effective management of error by stressing the need to enquire into how systems are used, and in particular by examining the formation of use histories.

List of References

- Aggarwal, I., & Woolley, A. W. (2013). Do you see what I see? The effect of members' cognitive styles on team processes and errors in task execution. *Organizational Behavior and Human Decision Processes*, 122(1), 92-99.
- Allison, P. D., & Waterman, R. P. (2002). Fixed-effects negative binomial regression models. *Sociological Methodology*, 32(1), 247-265.
- Avgerinos, E., & Gokpinar, B. (2018). Task Variety in Professional Service Work: When It Helps and When It Hurts. *Production and Operations Management*, 27(7), 1368-1389.
- Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist*, 54(7), 462.
- Benbasat, I., & Barki, H. (2007). Quo vadis TAM?. *Journal of the Association for Information Systems*, 8(4), 211-218.
- Benlian, A. (2015). IT feature use over time and its impact on individual task performance. *Journal of the Association for Information Systems*, 16(3), 144-173.
- Bolton, M. L. (2017). A task-based taxonomy of erroneous human behavior. *International Journal of Human-Computer Studies*, 108, 105-121.
- Booth, P. A. (1991). Errors and theory in human-computer interaction. *Acta Psychologica*, 78(1), 69-96.
- Brodbeck, F. C., Zapf, D., Prümper, J., & Frese, M. (1993). Error handling in office work with computers: A field study. *Journal of Occupational and Organizational Psychology*, 66(4), 303-317.
- Burton-Jones, A., & Gallivan, M. J. (2007). Toward a deeper understanding of system usage in organizations: a multilevel perspective. *MIS Quarterly*, 657-679.
- Burton-Jones, A., & Grange, C. (2013). From use to effective use: A representation theory perspective. *Information Systems Research*, 24(3), 632-658.
- Burton-Jones, A., & Straub Jr, D. W. (2006). Reconceptualizing system usage: An approach and empirical test. *Information Systems Research*, 17(3), 228-246.
- Butler, B. S., & Gray, P. H. (2006). Reliability, mindfulness, and information systems. *MIS Quarterly*, 30(2), 211-224.
- Cameron, A C., & Pravin K Trivedi, P. K. (1990). Regression-based tests for overdispersion in the poisson model. *Journal of Econometrics*, 46(3), 347-364.
- Carlo, J. L., Lyytinen, K., & Boland Jr, R. J. (2012). Dialectics of Collective Minding: Contradictory Appropriations of Information Technology in a High-Risk Project. *MIS Quarterly*, 36(4), 1081-1108.
- Casner, S. M., & Schooler, J. W. (2015). Vigilance impossible: diligence, distraction, and daydreaming all lead to failures in a practical monitoring task. *Consciousness and Cognition*, 35, 33-41.
- Chung, P. H., & Byrne, M. D. (2008). Cue effectiveness in mitigating postcompletion errors in a routine procedural task. *International Journal of Human-Computer Studies*, 66(4), 217-232.
- Cohen, J., Cohen, P. (1975). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. Lawrence Erlbaum.

- Cohen, M. D., Burkhart, R., Dosi, G., Egidi, M., Marengo, L., Warglien, M., & Winter, S. (1996). Routines and other recurring action patterns of organizations: contemporary research issues. *Industrial and Corporate Change*, 5(3), 653-698.
- De Vries, P., Midden, C., & Bouwhuis, D. (2003). The effects of errors on system trust, self-confidence, and the allocation of control in route planning. *International Journal of Human-Computer Studies*, 58(6), 719-735.
- Dismukes, K. (2006). Concurrent task management and prospective memory: pilot error as a model for the vulnerability of experts. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50(9), 909-913.
- Durcikova, A., Fadel, K. J., Butler, B. S., & Galletta, D. F. (2011). Research Note-Knowledge Exploration and Exploitation: The Impacts of Psychological Climate and Knowledge Management System Access. *Information Systems Research*, 22(4), 855-866.
- Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. Ross (Ed.), *The psychology of learning and motivation* (pp. 145–199). New York, NY: Academic Press.
- Fadel, K.J. (2012). "User adaptation and infusion of information systems," *Journal of Computer Information Systems*, 52(3), 1-10.
- Frese, M., Brodbeck, F., Heinbokel, T., Mooser, C., Schleiffenbaum, E., & Thiemann, P. (1991). Errors in training computer skills: On the positive function of errors. *Human-Computer Interaction*, 6(1), 77-93.
- Frese, M., & Keith, N. (2015). Action Errors, Error Management, and Learning in Organizations. *Annual Review of Psychology*, 66, 661-687.
- Fuller, R. M., & Dennis, A. R. (2009). Does fit matter? The impact of task-technology fit and appropriation on team performance in repeated tasks. *Information Systems Research*, 20(1), 2-17.
- Galletta, D. F., Abraham, D., El Louadi, M., Lekse, W., Pollalis, Y. A., & Sampler, J. L. (1993). An empirical study of spreadsheet error-finding performance. *Accounting, Management and Information Technologies*, 3(2), 79-95.
- Galletta, D. F., Hartzel, K. S., Johnson, S. E., Joseph, J. L., & Rustagi, S. (1996). Spreadsheet presentation and error detection: An experimental study. *Journal of Management Information Systems*, 45-63.
- Gill, T. G. (1996). Expert systems usage: task change and intrinsic motivation. *MIS Quarterly*, 20(3), 301-329.
- Goswami, S., Chan, H. C., & Kim, H. W. (2008). The role of visualization tools in spreadsheet error correction from a cognitive fit perspective. *Journal of the Association for Information Systems*, 9(6), 321-343.
- Hofmann, D., & Frese, M. (2011). Errors, error taxonomies, error prevention, and error management: Laying the groundwork for discussing errors in organizations. In D. Hofmann and M. Frese (Eds.), *Errors in Organizations*, (pp. 1-44). New York: Routledge, Taylor & Francis.
- Hoffman, A. J., & Ocasio, W. (2001). Not all events are attended equally: Toward a middle-range theory of industry attention to external events. *Organization Science*, 12(4), 414-434.
- James, W. (1890). *The Principles of Psychology*. Henry Holt, New York. Volume 1.
- Jaspersen, J. S., Carter, P. E., & Zmud, R. W. (2005). A comprehensive conceptualization of post-adoptive behaviors associated with information technology enabled work systems. *MIS Quarterly*, 29(3), 525-557.

- Kahneman, D. (1973). *Attention and Effort*. Englewood Cliffs, NJ: Prentice-Hall.
- Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: the contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General*, 132(1), 47-70.
- Kanfer, R., & Ackerman, P. L. (1989). Motivation and cognitive abilities: An integrative/aptitude-treatment interaction approach to skill acquisition. *Journal of Applied Psychology*, 74, 657–690.
- Karimi, J., Somers, T. M., & Gupta, Y. P. (2004). Impact of environmental uncertainty and task characteristics on user satisfaction with data. *Information Systems Research*, 15(2), 175-193.
- Keith N., & Frese M. (2008). Effectiveness of error management training: a meta-analysis. *Journal of Applied Psychology* 93(1), 59.
- Kim, S.S., & Malhotra, N. (2005). "A user empowerment approach to information systems infusion," *IEEE Transactions on Engineering Management*, 61(4), 656-668.
- Klein, B. D., Goodhue, D. L., & Davis, G. B. (1997). Can humans detect errors in data? Impact of base rates, incentives, and goals. *MIS Quarterly*, 21(2), 169-190.
- Knudsen, E. (2007). Fundamental components of attention. *Annual Review of Neuroscience*, 30, 57-78.
- Laureiro-Martinez, D. (2014). Cognitive Control Capabilities, Routinization Propensity, and Decision-Making Performance. *Organization Science*, 25(4), 1111-1133.
- Lazar J., Allen A., Kleinman J., and Malarkey C. (2007). What frustrates screen reader users on the web: A study of 100 blind users. *International Journal of human-computer interaction* 22(3), 247–269.
- Lazonder, A. W., & van der Meij, H. (1994). Effect of error information in tutorial documentation. *Interacting with Computers*, 6(1), 23-40.
- Li, X., Hsieh, J. P. A., & Rai, A. (2013). Motivational differences across post-acceptance information system usage behaviors: An investigation in the business intelligence systems context. *Information Systems Research*, 24(3), 659-682.
- Loft, S., & Remington, R. W. (2010). Prospective memory and task interference in a continuous monitoring dynamic display task. *Journal of Experimental Psychology: Applied*, 16(2), 145-157.
- Love, P. E., Edwards, D. J., Irani, Z., & Walker, D. H. (2009). Project pathogens: The anatomy of omission errors in construction and resource engineering project. *IEEE Transactions on Engineering Management*, 56(3), 425-435.
- McAvoy, J., Nagle, T., & Sammon, D. (2013). Using mindfulness to examine ISD agility. *Information Systems Journal*, 23(2), 155-172.
- McDaniel, M. A., & Einstein, G. O. (2000). Strategic and automatic processes in prospective memory retrieval: A multiprocess framework. *Applied Cognitive Psychology*, 14(7), S127-S144.
- McVay, J. C., & Kane, M. J. (2012). Drifting from slow to "d'oh!": Working memory capacity and mind wandering predict extreme reaction times and executive control errors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38(3), 525-549.
- Mukhopadhyay, T., Singh, P., & Kim, S. H. (2011). Learning curves of agents with diverse skills in information technology-enabled physician referral systems. *Information Systems Research*, 22(3), 586-605.

- Narayanan, S., Balasubramanian, S., & Swaminathan, J. M. (2009). A matter of balance: Specialization, task variety, and individual learning in a software maintenance environment. *Management Science*, 55(11), 1861-1876.
- Norman, D. A. (1981). Categorization of action slips. *Psychological Review*, 88(1), 1-15.
- Ocasio, W. (2011). Attention to attention. *Organization Science*, 22(5), 1286-1296.
- Panko, R. R., & Aurigemma, S. (2010). Revising the Panko–Halverson taxonomy of spreadsheet errors. *Decision Support Systems*, 49(2), 235-244.
- Pashler, H. E. (1999). *The psychology of attention*. Cambridge, MA: MIT Press.
- Powell, S. G., Baker, K. R., & Lawson, B. (2008). A critical review of the literature on spreadsheet errors. *Decision Support Systems*, 46(1), 128-138.
- Randall, J. G., Oswald, F. L., & Beier, M. E. (2014). Mind-wandering, cognition, and performance: A theory-driven meta-analysis of attention regulation. *Psychological Bulletin*, 140(6), 1411-1431.
- Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man and Cybernetics*, (3), 257-266.
- Rasmussen, J., & Vicente, K. (1989). Coping with human errors through system design: implications for ecological interface design. *International Journal of Man-Machine Studies*, 31(5), 517-534.
- Reason, J. (1984). Lapses of attention in everyday life. In: R. Parasuraman & D. Davis (Eds.), *Varieties of Attention* (pp. 179–189). New York: Academic Press.
- Reason, J. (1990). *Human Error*. Cambridge University Press.
- Reason, J. (2000). Human error: models and management. *BMJ*, 320 (7237), 768-770.
- Reason, J. (2002). Combating omission errors through task analysis and good reminders. *Quality and Safety in Health Care*, 11(1), 40-44.
- Reason, J., & Hobbs, A. (2003). *Managing maintenance error: A practical guide*. Ashgate: Aldershot.
- Saga, V. L., and Zmud, R. W. (1993). The nature and determinants of IT acceptance, routinization, and infusion. *Proceedings of the IFIP TC8 working conference on diffusion, transfer and implementation of information technology* (pp. 67-86).
- Sanjram, P. K. (2013). Attention and intended action in multitasking: An understanding of cognitive workload. *Displays*, 34(4), 283-291.
- Sanjram, P. K., & Khan, A. (2011). Attention, polychronicity, and expertise in prospective memory performance: Programmers' vulnerability to habit intrusion error in multitasking. *International Journal of Human-Computer Studies*, 69(6), 428-439.
- Santhanam, R., & Sein, M. K. (1994). Improving end-user proficiency: Effects of conceptual training and nature of interaction. *Information Systems Research*, 5(4), 378-399.
- Schilling, M. A., Vidal, P., Ployhart, R. E., & Marangoni, A. (2003). Learning by doing something else: Variation, relatedness, and the learning curve. *Management Science*, 49(1), 39-56.
- Sein, M. K., & Santhanam, R. (1999). Research report. Learning from goal-directed error recovery strategy. *Information Systems Research*, 10(3), 276-285.

- Sharma, R., & Yetton, P. (2007). The contingent effects of training, technical complexity, and task interdependence on successful information systems implementation. *MIS Quarterly*, 31(2), 219-238.
- Silver, M. S. (1990). Decision support systems: directed and nondirected change. *Information Systems Research*, 1(1), 47-70.
- Smallwood, J. (2013). Distinguishing how from why the mind wanders: a process–occurrence framework for self-generated mental activity. *Psychological Bulletin*, 139(3), 519-535.
- Smallwood, J., & Schooler, J. W. (2006). The restless mind. *Psychological Bulletin*, 132, 946–958.
- Speier, C., Valacich, J.S., and Vessey, I. (1999). “The influence of task interruption on individual decision making: An information overload perspective,” *Decision Sciences*, 30(2), 337-360.
- Staats, B. R., & Gino, F. (2012). Specialization and variety in repetitive tasks: Evidence from a Japanese bank. *Management Science*, 58(6), 1141-1159.
- Sun, H. (2012). Understanding user revisions when using information system features: adaptive system use and triggers. *MIS Quarterly*, 36(2), 453-478.
- Thomson, D. R., Seli, P., Besner, D., & Smilek, D. (2014). On the link between mind wandering and task performance over time. *Consciousness and Cognition*, 27, 14-26.
- Wood, R. E. (1986). Task complexity: Definition of the construct. *Organizational Behavior and Human Decision Processes*, 37(1), 60-82.
- Zapf, D., Brodbeck, F. C., Frese, M., Peters, H., & Prümper, J. (1992). Errors in working with office computers: A first validation of a taxonomy for observed errors in a field setting. *International Journal of Human-Computer Interaction*, 4(4), 311-339.
- Zapf, D., Maier, G. W., Rappensperger, G., & Irmer, C. (1994). Error detection, task characteristics, and some consequences for software design. *Applied Psychology*, 43(4), 499-520.

Appendix A: List of Actions

Identifier	Action	Related Feature	Unrelated Feature	Explanation and Comments
id1	select race	x	x	Selecting a race to officiate. When this action was performed one time per each parallel race, we coded it as related feature use. When it was performed more than once per each parallel race, we coded it as unrelated feature use.
id2	set start time	x		Setting the starting time of a race.
id3	reset start time	x		Reschedule the starting time of a race.
id4	reset time button	x		Pressing the button reset time after the start time was rescheduled
id5	set running flag/PAPA	x	x	Setting running flag. When each of these actions was performed one time per race, we coded it as related feature use. When they were performed more than once per race, it was coded as unrelated feature use, as only one out of the four flags can be raised during the course of a race.
id6	set running flag/ZULU	x	x	
id7	set running flag/BLACK	x	x	
id8	set running flag/INDIA	x	x	
id9	postpone race	x		Postpone a race which is not yet started

id10	postpone race/ALPHA	x		Postpone a race which is not yet started with a remark that there is no more racing today
id11	postpone race/HOTEL	x		Postpone a race which is not yet started with a remark that further information is available offshore
id12	abortSt	x		Abandon and postpone a race during the start phase
id13	abortSt/ALPHA	x		Abandon race during the start phase with a remark that there is no more racing today
id14	abortSt/HOTEL	x		Abandon a race during start phase with a remark that further information is available offshore
id15	set individual recall /XRAY	x		Starting an individual recall
id16	unset individual recall /XRAY	x		Ending of an individual recall
id17	general recall/yes	x		Recall the whole race (answering Ok in confirmation dialog)
id18	general recall/no		x	Canceling recall of the whole race
id19	set running course/yes	x		Setting the running course
id20	set running course/no		x	Canceling setting the running course
id21	abortAp	x		Abandon and postpone a running race
id22	abortNovember	x		Abandon a running race with a remark that more information is available at the start line
id23	set blue flag/yes	x		Signaling that the first boat has crossed the finish line
id24	set blue flag/no		x	Canceling the signal that the first boat has crossed the finish line
id25	unset blue flag/yes	x		Signaling that the last boat has crossed the finish line
id26	unset blue flag/no		x	Canceling the signal that the last boat has crossed the finish line
id27	login		x	Login to another race course administered by another user
id28	cancel login		x	Canceling login
id29	select event		x	Select sailing event (after login)
id30	select course		x	Select course (after selecting an event)
id31	select label		x	Click on the label of the RCC app that has no functionality
id32	reset race/no		x	Canceling reset race

Appendix B: List of prescribed actions

Identifier	Action	Explanation and Comments
pa1	Start time	The start time of the race, as documented on the hand-written protocols.
pa2	Running flag	The running flag that was raised (PAPA, ZULU, BLACK or INDIA) as documented on the hand-written protocols.
pa3	Set blue flag	Signaling that the first boat has crossed the finish line (as documented on the hand-written protocols)
pa4	Unset blue flag	Signaling that the last boat has crossed the finish line (as documented on the hand-written protocols).

Appendix C: Operationalization of Task Variation

Task Variation (TV_{ij}) is defined as the variability of the prescribed actions between race j (PA_{ij}) and its corresponding parallel races. We measure TV_{ij} as $(1 - \text{task similarity})$. Task similarity is quantified in a vector space model, a technique frequently used for information retrieval tasks such as comparing similarities of documents or websites (Salton et al. 1975). Objects can be represented as a vector. Similarity between the vectors is calculated as the absolute value of the cosine of the angle between the vectors. Consider the set PA_{ij} , for each parallel race k , we extract the overlapping actions (OA) in *time* for race k as follows:

$OA_k = \{pa_{ik} \in PA_{ik} : \text{time}(pa_{ik}) \in [t_{min}^j, t_{max}^j]\}$, namely all actions on race k between the first action (at time t_{min}^j) and the last action (at time t_{max}^j) on race j respectively.

Let t_{min}^k and t_{max}^k denote the time of the first and last related feature use in OA_k respectively.

We extract the overlapping actions in the race under investigation (race j) as follows:

$OA_j = \{pa_{ij} \in PA_{ij} : \text{time}(pa_{ij}) \in [t_{min}^k, t_{max}^k]\}$, namely all actions on race j between the first action (at time t_{min}^k) and the last action (at time t_{max}^k) on race k respectively.

We then map the overlapping actions OA_k and OA_j to a vector space. As the actions correspond to 4 related features in use that are listed in Appendix C, we map OA_k and OA_j to

4-dimensional vectors $\overrightarrow{OA_k}$ and $\overrightarrow{OA_j}$ which contain a value of zero or one. The similarity

between $\overrightarrow{OA_k}$ and $\overrightarrow{OA_j}$ is defined as follows: $(\overrightarrow{OA_k}, \overrightarrow{OA_j}) = \left| \frac{\overrightarrow{OA_k} \cdot \overrightarrow{OA_j}}{\|\overrightarrow{OA_k}\| \|\overrightarrow{OA_j}\|} \right|$, whereas $\overrightarrow{OA_k}$ is the

vector with the overlapping actions of race k , and $\overrightarrow{OA_j}$ is the vector with the overlapping actions

of race j . The absolute value of cosine yields a value between 0 and 1, ranging from minimum

to maximum similarity. In case of multiple parallel races, we take the average of the similarity

indices to obtain the TS measure.