WHAT DISRUPTS FLOW IN OFFICE WORK? THE IMPACT OF FREQUENCY AND RELEVANCE OF IT-MEDIATED INTERRUPTIONS[1]

Mario Nadj
Business & Information Systems Engineering, Department of Business and Economics,
TU Dortmund University, Dortmund, GERMANY {mario.nadj@tu-dortmund.de}

Raphael Rissler
Human-Centered Systems Lab (h-lab), Department of Economics and Management, Karlsruhe Institute of Technology (KIT), Karlsruhe, and
AI Incubation, SAP SE, Walldorf, GERMANY {raphael.rissler@partner.kit.edu}

Marc T. P. Adam
Discipline of Computing and Information Technology, School of Information and Physical Sciences,
University of Newcastle, Newcastle, AUSTRALIA {marc.adam@newcastle.edu.au}

Michael T. Knierim
Institute for Information Systems and Marketing (IISM), Department of Economics and Management,
Karlsruhe Institute of Technology (KIT), Karlsruhe, GERMANY {michael.knierim@kit.edu}

Maximilian X. Li
Intuitive Robots Lab, Department of Informatics,
Karlsruhe Institute of Technology (KIT), Karlsruhe, GERMANY {maximilian.li@kit.edu}

Alexander Maedche
Human-Centered Systems Lab (h-lab), Department of Economics and Management,
Karlsruhe Institute of Technology (KIT), Karlsruhe, GERMANY {alexander.maedche@kit.edu}

René Riedl
Digital Business Institute, School of Business and Management, University of Applied Sciences Upper Austria, Steyr, and
Institute of Business Informatics – Information Engineering, Johannes Kepler University, Linz, AUSTRIA {rene.riedl@fh-steyr.at}

Flow, the holistic sensation people experience when they act with total involvement, is a known driver for desired work outcomes like task performance. However, the increasing ubiquity of IT at work can disrupt employees' flow. Thus, the impact of IT-mediated interruptions on flow warrants more attention in research and practice. We conducted a NeuroIS laboratory experiment focusing on a typical office work task—an invoice matching task (i.e., matching customer payments to invoices). We manipulated interruption frequency (low, high) and content relevance (irrelevant, relevant) to study the impact of interruptions on self-reported flow, its dimensions, and high-frequency heart rate variability (HF-HRV; calculated from electrocardiography recordings) as a proxy for parasympathetic nervous system (PNS) activation. We found that content relevance moderated the relationship between interruption frequency and self-reported flow and that these results vary along flow dimensions. Content relevance also moderated the relationship between interruption frequency and PNS activation. Furthermore, self-reported flow was positively associated with both perceived and objective task performance, while PNS

activation was not related to either performance measure. Lastly, we found no relationship between PNS activation (measured by HF-HRV) and self-reported flow, contributing to an important debate in the NeuroIS literature on whether physiological evidence constitutes an alternative or a complement to self-reports. Overall, our findings indicate that frequent interruptions are not harmful per se. Rather, considering content relevance is critical for a more comprehensive understanding of the effects on self-reported flow, its dimensions, and the underlying physiology.

Keywords: IT-mediated interruptions, flow, electrocardiography (ECG), heart rate variability (HRV), laboratory experiment, NeuroIS

Introduction ▮▮▮▮▮▮▮▮▮

Despite the many positive impacts, digitalized work also creates negative outcomes for employees, as information technology (IT) can be a source of interruption (e.g., email or instant messaging notifications). Such IT-mediated interruptions are defined as "IT-based external events with a range of content that captures cognitive attention and breaks the continuity of an individual's primary task activities" (Addas & Pinsonneault, 2015, p. 233). In the European Working Conditions Survey, 33% of EU employees reported frequent work interruptions with a disruptive impact on primary tasks (Parent-Thirion et al., 2015).

Information systems (IS) scholars have called for more research into the nomological net (i.e., antecedents, consequences, moderation, and/or mediation effects) of different IT-mediated interruption types (Addas & Pinsonneault, 2015). An important consequence could be that interruptions disrupt humans' flow (Rissler et al., 2017a), which may harm desired work outcomes such as task performance (Fullagar & Kelloway, 2009). Flow is defined as "the holistic sensation that people feel when they act with total involvement" (Csikszentmihalyi, 1975, p. 36). Despite increasing research interest in flow (Mahnke et al., 2014, 2015; Rissler et al., 2017b) and increasing employee concerns about IT-mediated interruptions (Parent-Thirion et al., 2015), there is a *lack of research on how IT-mediated interruptions impact flow* (Rissler et al., 2017b) *and further downstream variables, such as performance* (e.g., Addas & Pinsonneault, 2018a; Chen & Karahanna, 2018). This lack of research is problematic since IT-mediated interruptions, on the one hand, improve task completion and performance (e.g., a user receives task-relevant information via email). Yet, on the other hand, interruptions are not cost-free and may also have a severe downside if they prevent users from effectively returning to their primary task (Addas & Pinsonneault, 2018a; Chen & Karahanna, 2018).

IS scholars have identified frequency and content relevance as key drivers of interruption outcomes (Addas & Pinsonneault, 2015; Galluch et al., 2015).[2] Interruption frequency is critical, as interruptions have negative effects on user states, including annoyance, anxiety (e.g., Bailey et al., 2001; Bailey & Konstan, 2006), overload, and stress (e.g., Galluch et al., 2015; Tams et al., 2015). Also, office work interruptions have increased in recent years and will likely increase even further (Addas & Pinsonneault, 2018b; Parent-Thirion et al., 2015). Moreover, content relevance is critical as it can potentially explain inconsistent findings linking interruptions to both negative and positive outcomes (Addas & Pinsonneault, 2015). Following Addas and Pinsonneault (2015, 2018a), we define content relevance as the degree to which an interruption's content supports humans in resolving a primary task. Importantly, we expect content relevance to act as a moderator in that frequent interruptions may only be harmful for irrelevant but not for relevant content. First, flow theory suggests that *clear and immediate feedback* promotes flow (Csikszentmihalyi, 2014; Nakamura & Csikszentmihalyi, 2009). In fact, relevant interruptions also disclose information about expected and actual primary task performance (Addas & Pinsonneault, 2015). Second, flow theory also suggests *clear goals* as flow facilitators (Csikszentmihalyi, 2014; Jackson & Marsh, 1996). While relevant interruptions keep task goals clear, irrelevant interruptions could introduce unclear feedback and contradictory demands.

In recent years, scholars have made repeated calls for further research into the physiology of flow (Harmat et al., 2015; Knierim et al., 2017). As flow can occur unconsciously, self-report measures alone cannot fully capture this phenomenon (Harmat et al., 2015; Knierim et al., 2017). Referring to Ortiz de Guinea et al. (2013), Léger et al. (2014a) suggested "to enrich current flow measures by capturing implicit (i.e., automatic and unconscious) psychophysiological measures in conjunction with more explicit (i.e., self-report) measures" (p. 275). Because physiological measures may explain different

---

[2] Unless stated otherwise, throughout this article, the terms interruption frequency and content relevance refer to the respective properties of an IT-mediated interruption. Interruption frequency refers to the number of interruptions per time interval, which may either be low or high. Content

relevance refers to the significance of an interruption's content to a task, which may either be irrelevant or relevant. In the following the terms 'irrelevant content' and 'relevant content' will be used for simplification.

variance in outcome variables (e.g., performance), they are valuable complements to self-report measures (Tams et al., 2014). On this basis, and in light of the increasing importance of neurophysiology in IS research (NeuroIS) (e.g., Dimoka et al., 2012; Riedl et al., 2020; Riedl & Léger, 2016), we studied both (1) self-reported flow *after* task execution and (2) individuals' parasympathetic nervous system (PNS) activation by electrocardiography (ECG) *during* task execution. We decided to rely on ECG in our study due to its particular contribution to flow variance explanation, its high robustness, and its low obtrusiveness, which is an important factor for the ongoing study of flow in more ecologically valid scenarios (Riedl et al., 2014).

Overall, there is only limited evidence available for the flow-performance relationship in the office work context (Csikszentmihalyi et al., 2016; Peifer & Wolters, 2021). Existing research often relies on cross-sectional surveys, with most finding a positive direct relationship. Yet, performance has been conceptualized in different ways. For instance, cross-sectional surveys that have conceptualized performance as service quality of service personnel (Kuo & Ho, 2010) or work creativity of software developers (Zubair & Kamal, 2015) have found positive and direct flow-performance links. By contrast, a cross-sectional survey covering different companies did not confirm a direct flow-performance relationship (conceptualized as executing tasks that serve the company's goals); however, conscientiousness emerged as a moderator—that is, flow predicted performance for conscientious employees only (Demerouti, 2006). With this in mind and given calls for further research, particularly research using laboratory experiments (e.g., Peifer & Wolters, 2021), we also explored this relationship for both perceived and objective task performance.

Regarding the PNS-performance relationship, evidence indicates mixed results, namely that high performance is either related to high PNS activation (Chin & Kales, 2019) or moderate PNS activation (Harris et al., 2017a). Therefore, the PNS-performance relationship needs further investigation. Against the background of the described status in research, in this article, we address the following research questions:

RQ1: *Does content relevance moderate the impact of IT-mediated interruption frequency on (**a**) self-reported flow and (**b**) PNS activation?*

RQ2: *Do (**a**) self-reported flow and (**b**) PNS activation positively correlate with perceived and objective task performance?*

To address these questions, we conducted a large-scale NeuroIS laboratory experiment (*N* = 166). Using a 2 (interruption frequency: low, high) by 2 (content relevance:

irrelevant, relevant) design, participants executed an invoice matching task in which they linked customer payments to corresponding invoices, a typical task completed by office staff using enterprise resource planning (ERP) systems.

Our research makes four contributions. First, we extend flow theory by showing that content relevance moderates the relationship between interruption frequency and different self-reported flow dimensions during a typical office task. This is critical, as we find that the frequency-relevance interplay influences the self-reported flow dimensions in different ways. Second, our research sheds light on how interruptions influence flow physiology in terms of PNS activation. Such an understanding is important as flow is often described as a state of optimized physiological activation (Bian et al., 2016; Knierim et al., 2017) and scholars have called for further research into flow physiology (Harmat et al., 2015; Knierim et al., 2017). We thereby also complement a NeuroIS investigation by Ortiz de Guinea et al. (2014) that called for "testing users' neurophysiological reactions to technological interruptions and studying the potential temporal correlations between neurophysiological measures and perceptual ones during IS use" (p. 32). Third, we complement existing research on the relationship of self-reported flow and PNS activation with both perceived and objective task performance. Finally, our work also makes a practical contribution by showing system designers, managers, and employees what impact IT-mediated interruptions have on users' flow during task execution. We make suggestions on how interruptions' negative effects during office work can be avoided using the moderating role of content relevance. Moreover, we discuss how future flow-adaptive systems could help to mitigate the negative effects of interruptions during office work.

Theoretical Background

What is Flow?

Flow was first conceptualized in the 1970s by Mihaly Csikszentmihalyi, who created questionnaires and structured interviews to study various physical and mental tasks in nonprofessional contexts without extrinsic rewards. Later, substantiated by a wealth of empirical research, scientists established that people can experience flow in nearly any activity, regardless of whether its rewards are intrinsic (e.g., leisure activities) or extrinsic (e.g., office work). A fundamental research finding is that a person may enter a flow state if the activity fulfills three preconditions (Csikszentmihalyi, 2014; Nakamura & Csikszentmihalyi, 2009):

Clear goals: Activities with clear goals offer direction and purpose. Their value lies in their ability to structure experience by channeling attention rather than being an end in itself.

Clear and immediate feedback: Activities with clear and immediate feedback from a range of potential sources inform individuals about their progress, giving them the opportunity to adjust their actions and reduce uncertainty about what to do next.

Perceived challenge-skill balance: Flow tends to emerge when the individual conducting an activity perceives a balance between challenges (i.e., opportunities for action or goals) and skills (i.e., the abilities of an individual to achieve these goals).

The flow experience itself is characterized by the following four flow dimensions:

(1) Merging of action and awareness: During flow, individuals are aware of their actions but not of the awareness itself (Csikszentmihalyi, 1975, 2000; Jackson & Marsh, 1996). Hereby, a sense of effortlessness and spontaneity is related to this flow dimension.

(2) Concentration: The second dimension is a human's concentration on the performed activity. Based on interviews, Csikszentmihalyi (1975) found that during flow people shut out the world and are mentally cut off from the world.

(3) Loss of self-consciousness: During flow, an activity's action requirements fully involve the individual, rendering self-considerations about how one is doing in the view of others irrelevant (Jackson & Marsh, 1996). This dimension is also known as "ego loss" or "self-forgetfulness" (Csikszentmihalyi, 1975, 2014).

(4) Sense of control: The feeling of being in control over the activity's demands (whether or not this feeling is objectively justified) ensures that no fear of failure or worry occurs while conducting the activity (Csikszentmihalyi, 1975, 2000; Jackson & Marsh, 1996).

The conceptualization of flow along these four dimensions has proved remarkably robust, confirmed by studies on various activities (Csikszentmihalyi, 2014).[3] Further, flow appears consistent across different cultures, genders, and age groups (Csikszentmihalyi, 2014). Flow has also been studied in IS contexts such as e-commerce (Yi et al., 2015). For a systematic overview of flow in IS contexts, we refer to Mahnke et al. (2014) and Rissler et al. (2017b).

The Physiology of Flow

Based on neurophysiological evidence, flow is a state of optimized cognitive effort and physiological activation (Knierim et al., 2017; Peifer & Tan, 2021), culminating in a perception of effortless action despite challenging task demands. Regarding central nervous system (CNS) activity, research on cognitive effort and attention using electroencephalography (EEG) measures have repeatedly found increased neural activity with increases in cognitive effort (when a task becomes more cognitively or attentionally demanding—see, e.g., Chikhi et al., 2022); researchers have also reported reduced neural activity patterns in expert performers when compared with novices (e.g., Wolf et al., 2015). Scholars have argued that during flow, when high challenge meets high skill, the brain only recruits essential neural resources, leading to neural efficiency that is perceived as effortless, despite some cognitive effort being required to meet the task demands (Peifer & Tan, 2021). Further, during flow, the brain efficiently integrates attentional processes to reduce distractibility (e.g., by down-regulating self-referential attentional processes and attentional alertness), thus facilitating focus on the primary task (Harris et al., 2017b). In simple terms, during flow, attention is highly narrowed on the primary task—attention-competing stimuli from the outer (e.g., task-irrelevant noises) and inner worlds (e.g., thoughts about one's own needs or performances) are discarded (Harris et al., 2017b).

In extension, scholars understand the body as accompanying this neural configuration through the provision of an optimized physiological activation, where flow is considered to be directly related to autonomic nervous system (ANS) activity (Peifer & Tan, 2021). Because CNS and ANS measures appear to be mostly complementary for the study of cognitive effort (Chikhi et al., 2022), they are often used in concert. The ANS regulates the critical functions of the human body, such as heart rate and gland activity (e.g., to release behaviorally relevant hormones). The two major ANS components are the sympathetic nervous system (SNS) and the PNS. SNS activation prepares for action or stressful situations (Kolb & Whishaw, 2009; Léger et al., 2014a; Riedl, 2013). Conversely, PNS activation aims to keep the body in a relaxed state or bring it back to such a state, counterbalancing SNS activation. This function is critical, as it

---

[3] Some literature mentions *transformation of time* as another flow dimension. It denotes a phenomenon where a person is so immersed in an activity that little capacity is left for mental processes that contribute to the experience of activity duration (Csikszentmihalyi, 2014). We decided not to include transformation of time as a flow dimension in our study as it was not initially defined by flow theory. Moreover, there is evidence that this flow dimension is not stable across different tasks (Csikszentmihalyi, 2014; Jackson & Marsh, 1996).

secures homeostasis, a state in which the body is in a stable and constant condition (Kolb & Whishaw, 2009).

SNS activation is related to the release of stress hormones, increases the heart's contraction rate and force (cardiac output), but decreases HRV. Conversely, PNS activation reduces heart rate but increases HRV to restore homeostasis after a state of physiological activation (Berntson et al., 2007; Valentini & Parati, 2009).

Thus, the SNS is stimulatory while the PNS is inhibitory (e.g., Kolb & Whishaw, 2009). SNS and PNS activity correlate with a number of physiological changes in the human body, some of which are briefly summarized in the following. SNS activation correlates with pupil dilation (i.e., increased attention), skin conductance elevation (i.e., increased arousal), airway relaxation, heartbeat acceleration, heart rate variability (HRV) decrease, intense glucose release, and muscle tension. PNS activation results in the opposite effects—namely pupil contraction (i.e., decreased attention), skin conductance reduction, airway constriction, heartbeat slowdown, HRV increase, halted glucose release, and muscle relaxation.

Regarding SNS activity, research has established that some activity is always inherent in flow experiences since a moderate degree of arousal is necessary (e.g., Peifer et al., 2014). By contrast, research has not yet arrived at definitive conclusions about PNS activity. However, initial evidence suggests that PNS activity increases during flow. In a laboratory experiment, researchers found a positive relationship between PNS activation and flow in a cabin air management task ($n = 22$) (Peifer et al., 2014). Two other research teams found an inverted U-shape relationship between PNS activation and flow, that is, moderate PNS activation levels led to higher flow, while lower and higher levels of PNS activation led to lower flow; Chin and Kales (2019) used a Stroop test ($n = 48$) and Bian et al. (2016) used a virtual reality game ($n = 34$).

Theoretical propositions for this relationship between PNS activation and flow are primarily linked to attention and stress mechanisms (Harris et al., 2017b; Tozman & Peifer, 2016). Research has indicated that during flow attentional mechanisms do not additively increase mental demands—rather they become efficiently aligned (Harris et al., 2017b). A demanding attentional mechanism is conflict monitoring, a process that often unfolds when a person perceives self-evaluative threats when they cannot meet task demands (Tozman & Peifer, 2016). Tozman and Peifer (2016) argued that people may perceive some challenging tasks as more positive if there is no self-evaluative threat, and vice versa. Therefore, on the one hand, a *first theoretical proposition* suggests that increased PNS activation during flow could be linked to an absence of a self-evaluative threat and hence the lack of conflict monitoring. This may be caused passively by efficient resource demands (i.e., an

optimally demanding situation) or actively by a person's voluntary effort to calm down (i.e., a coping mechanism in an overly demanding situation). A study with expert musicians provided support for *this proposition* (Harmat et al., 2011), reporting PNS upregulation in the early minutes of encountering a challenging task, where individuals appeared to purposefully elicit a calming mechanism to cope with the sudden demand. Further evidence was found by Peifer et al. (2015), who studied the isolated effect of the stress hormone cortisol on flow, via controlled administration (i.e., participants received 20 mg oral cortisol). Flow was reduced as cognitive coping strategies were rendered less effective by the hormonal manipulation.

On the other hand, a *second theoretical proposition* for the increased PNS activation builds on foundational physiological observations. The human body is considered to be most effectively functioning when both SNS and PNS are highly active (even during challenging situations), as this provides the body with high flexibility to cope with task demands (Berntson et al., 2007; Shaffer & Ginsberg, 2017; Valentini & Parati, 2009). In contrast, stressful situations typically exhibit SNS dominance and PNS reductions (Peifer & Tan, 2021). Thus, during flow, when both SNS and PNS are elevated, the body can very effectively respond to elevated and possibly varying task demands (Harris et al., 2017b). As a consequence, the individual is effectively coping with the challenging task demands, which is represented by the most functional state of the ANS, a flexible and dynamic interplay of the SNS and PNS (Berntson et al., 2007; Valentini & Parati, 2009).

Related Work

In addition to essential flow theory literature, the present work relies on further research, particularly NeuroIS studies. In a literature review on the biological perspective on technostress, Riedl (2013) reported that interruptions during computer work may elevate the two major stress hormones, adrenaline and cortisol, as well as SNS activity in general (e.g., increased heart rate, blood pressure, electrodermal activity (EDA), and muscle tension, in addition to reduced HRV). Riedl (2013) also indicated that people benefit from PNS activation to avoid a "stage of exhaustion" caused by prolonged SNS activation.

Related work has also emphasized how IT events can trigger positive and negative cognitive-affective phenomena. For instance, Ortiz de Guinea and Webster (2013) conducted two studies (one qualitative and one quantitative) that combined different methods (e.g., open-ended questions, videos, protocol analysis), including a physiological measure (i.e., heart rate data). In doing so, they (1) conceptualized IS use patterns as people's emotions, cognitions, and behaviors while using IT for a work-related task; (2) examined how these

patterns change over time as a result of different IT events (expected and discrepant[4]); and (3) examined how they relate to short-term performance. They thereby identified two distinct IS use patterns: *automatic* (i.e., takes place during expected IT events, characterized by exploitive behaviors as well as cognitions and emotions unrelated to the IT employed; enhanced short-term performance) and *adjusted* (i.e., triggered by discrepant IT events, characterized by adaptive behaviors, negative affect related to the IT employed, as well as computer-related thoughts; no enhanced short-term performance). Physiologically, it was found that heart rate (as a proxy for physiological arousal) significantly *decreased* following discrepant IT events.

Moreover, Léger and colleagues studied cognitive absorption in the context of simulation-based training for ERP software usage (Léger et al., 2014a). Users had to make decisions while running a simulated company. It was found that cognitive absorption was significantly affected by skill, difficulty, and their interaction. Further, five neurophysiological measures were analyzed: EEG alpha, EEG beta, EDA, heart rate, and HRV. Each of those five "measures explained significant unique variance in cognitive absorption over and above skill, difficulty, and their interaction, and collectively more than doubled the explained variance" (p. 273). Importantly, the findings suggest that "cognitive absorption is positively related to a more relaxed, less vigilant state of the learner (after controlling for the difficulty of the task and trainees' beginning level of expertise)" (p. 279). Also, it was argued that it "is possible that the effect of cognitive absorption and training outcome is artificially suppressed if elements of the psychophysiology of the individual, such as HR [heart rate] and HRV, are not controlled for … The positive influence of HR and negative impact of HRV could suggest that some levels of arousal can also contribute to the training outcome" (p. 280). This conclusion seems consistent with flow theory and our theorizing in the previous section on flow physiology.

Finally, Léger et al. (2014b) introduced eye-fixation-related potential (EFRP) as a method to link physiological activity to specific information events on the computer screen. Besides its methodological contribution, we highlight the context where the EFRP method's efficacy was shown. In a natural IS use context, users had to read an industry report while email notifications popped up on their screen. The results revealed three neural processes linked to (1) an attentional response to the pop-up notification, (2) its cognitive processing, and (3) motor planning activity associated with the potential email opening.

Hypotheses Development ▪▪▪▪▪

Impact of Content Relevance and Frequency on Flow (H1)

The content of an IT-mediated interruption can be either supportive (relevant) or not supportive (irrelevant) in resolving a primary task (Addas & Pinsonneault, 2015, 2018a). In the following, we theorize how content relevance may moderate the effect of interruption frequency on flow. Content relevance can be linked to flow theory in two ways.

First, flow theory considers *clear and immediate feedback* on the present task as a flow facilitator, as it provides users with the opportunity to adjust their behavior and resolve a task better (Csikszentmihalyi, 2014; Nakamura & Csikszentmihalyi, 2009). As we outline in more detail below, relevant interruptions may create such opportunities. Conversely, *no or unclear feedback* does not give individuals the opportunity to adjust their behavior (Csikszentmihalyi, 2014; Nakamura & Csikszentmihalyi, 2009). In line with this proposition, researchers noted that relevant interruptions "reveal information about a discrepancy between expected and actual primary task performance" (Addas & Pinsonneault, 2015, p. 237) and thus help users perform their task better. Correspondingly, users reported that such interruptions provide critical information for their tasks (Addas & Pinsonneault, 2015) and yield higher benefits and less frustration than irrelevant interruptions (Gluck et al., 2007).

Second, flow theory considers *clear goals (noncontradictory demands for action)* as a flow facilitator since they offer users a structure for the task (Csikszentmihalyi, 2014) and a sense of knowing exactly what to do (Jackson & Marsh, 1996). In the context of a *relevant* interruption, the task goals remain clear because the interruption helps users achieve their goals. Thus, the interruption and the task work together. Conversely, *irrelevant* interruptions may corrupt goal clarity due to contradictory demands for action that are not part of the task and do not help the users achieve their goals. Thus, a higher frequency of irrelevant interruptions will disrupt flow. Indirect support for this is provided by a study where researchers found that users perceive contradictory demands for action in response to interruptions with irrelevant content (Galluch et al., 2015).

Building on the two aforementioned theoretical considerations, we hypothesize:

H1: *Content relevance moderates the impact of interruption frequency on flow such that an inhibiting effect is present for irrelevant content.*

---

[4] An *expected* IT event refers to a match between the user's expectation and the IT performance. In contrast, a *discrepant* event is an unexpected negative

event that describes a mismatch in this regard, usually due to problems, difficulties, or misunderstandings of the user with the respective IT.

Impact of Content Relevance and Frequency on the Flow Dimensions (H1a–d)

Several theoretical considerations further support a moderating role of content relevance in the relationship between interruption frequency and the flow dimensions. Regarding the *merging of action and awareness*, users need significantly more time to resume their primary tasks for irrelevant than for relevant interruptions (Czerwinski et al., 2000). This shorter "resumption lag" after relevant interruptions indicates that users' merging of action and awareness remains high. Trafton et al. (2003) argued that this can be explained by the fact that the primary task goals remain clear, which is an important condition for flow (Csikszentmihalyi, 2014). Correspondingly, users have reported that they are able to reengage quickly with their primary task without losing flow in response to relevant interruptions compared to irrelevant interruptions (Addas & Pinsonneault, 2015). Thus, an increasing frequency of irrelevant interruptions will impair a merging of action and awareness due to the experience of repeated resumption lags.

Scholars have argued that irrelevant interruptions "capture and divert attention from primary activities," and thus hinder users from *concentrating* on their primary task (Addas & Pinsonneault, 2015, p. 236). In turn, if the interruption is relevant, it does not divert attention but instead reinforces concentration and thinking about how to resolve the task (Addas & Pinsonneault, 2015). Building upon this rationale, we suggest that frequent interruptions with irrelevant content cause repeated diversions of attention that break users' concentration.

Finally, we are not aware of evidence on how content relevance may moderate the influence of interruption frequency on the *loss of self-consciousness* and *sense of control* dimensions. Based on the preceding considerations, we suggest that the unclear feedback and contradictory demands associated with irrelevant interruptions may give users more opportunities to consider how others would evaluate their execution, particularly when these interruptions occur frequently. Conversely, even if relevant interruptions occur frequently, they allow users to mentally stay with the task instead of thinking about how others might evaluate them. A similar line of reasoning applies to *sense of control* in that a higher frequency of unclear feedback and contradictory demands will diminish a user's sense that they are in control of their task execution. Hence, we expect that the more frequently a user receives irrelevant interruptions, the more the user's sense of control is disrupted. In summary, we hypothesize:

H1a: *Content relevance moderates the impact of interruption frequency on the merging of action and awareness such that an inhibiting effect is present for irrelevant content.*

H1b: *Content relevance moderates the impact of interruption frequency on concentration such that an inhibiting effect is present for irrelevant content.*

H1c: *Content relevance moderates the impact of interruption frequency on loss of self-consciousness such that an inhibiting effect is present for irrelevant content.*

H1d: *Content relevance moderates the impact of interruption frequency on sense of control such that an inhibiting effect is present for irrelevant content.*

Impact of Content Relevance and Frequency on PNS Activation (H2)

As outlined before, an interruption's content can be relevant or irrelevant to a primary task (Addas & Pinsonneault, 2015, 2018a). In the following, we first outline what effects interruption frequency may have on PNS activation and then argue why this effect could be moderated by the content relevance of an interruption.

In a field study, employees exhibited lower PNS activation when they had access to email compared to when their access to email was taken away (Mark et al., 2012). In a later study with a larger sample size, a significant reduction in PNS activation was found when employees spent a lot of time on emails during the workday (Mark et al., 2016). Based on these findings, we suggest that higher interruption frequency decreases PNS activation. However, we further propose that this effect is moderated by content relevance. Regarding relevant interruptions, we expect that PNS activation at low interruption frequency does not differ significantly from that at high interruption frequency.

Addas and Pinsonneault (2018a) provided further evidence for a moderating influence of content relevance. The researchers studied the role of interruptions' content relevance based on perceptual and observational measures. Building on action regulation theory, the study found a positive indirect effect of relevant interruptions via mindfulness and a negative indirect effect of irrelevant interruptions via subjective workload (Addas & Pinsonneault, 2018a). These results align well with the previously discussed role of attentional processes for flow.

A high level of irrelevant interruptions triggers attentional processes that are not relevant for processing the main task (Harris et al., 2017b), thus making it less likely that effective coping strategies will emerge and direct attention toward the main task, which in turn inhibits PNS activation. Thus, we propose:

H2: *Content relevance moderates the impact of interruption frequency on PNS activation such that an inhibiting effect is present for irrelevant content.*

Relationship of Flow, PNS Activation, and Task Performance (H3a-d)

Since the beginning of flow research, scholars have theorized a positive relationship between flow and task performance (Csikszentmihalyi, 1975). This flow-performance relationship stems from the flow dimensions that describe a highly functional state (Engeser & Rheinberg, 2008). Specifically, high concentration and a loss of self-consciousness support attentional processes, and the rewarding experience increases task motivation, effort, and perseverance (Engeser & Rheinberg, 2008; Schüler & Brunner, 2009). Research suggests that the lack of self-conscious thought not only explains improved task attention but also the intrinsically rewarding character of the experience (Harris et al., 2017b; Ulrich et al., 2014). This reward, in turn, is frequently discussed as a reason why individuals exert more effort and perseverance during task execution, which additionally explains performance improvements (Engeser & Rheinberg, 2008; Schüler & Brunner, 2009). We hypothesize:

H3a: *Flow is positively correlated with perceived task performance.*

H3b: *Flow is positively correlated with objective task performance.*

Studies often report a positive relationship between PNS activation and task performance (Mathewson et al., 2010; Murray & Russoniello, 2012). However, it is also well documented that task performance can degrade in complex tasks when overly demanding or stressful situations arise and a high level of autonomic arousal is present. Two studies (Chin & Kales, 2019; Harris et al., 2017a) have investigated the relationship between PNS activation and task performance, finding that high or moderate PNS activation is positively linked to high task performance. Such a pattern is explained by more efficient use of attentional processes during high task performance when individuals are coping well with the task demands (Harris et al., 2017b). Thus, we formulate the following hypotheses:

H3c: *PNS activation is positively correlated with perceived task performance.*

H3d: *PNS activation is positively correlated with objective task performance.*

We summarize our research model in Figure 1.

Research Method

Sample

We recruited 179 participants[5] from a public university through its online recruitment system. The homogeneous age group increases the experiment's internal validity. After data screening, we removed five participants due to failed attention checks.[6] Further, we removed eight participants based on outlier detection on the flow dimensions using Tukey's 1.5 interquartile range (IQR) criterion (Tukey, 1977). The final sample consisted of 166 participants (69 women, 97 men; mean age = 23.5, *SD* = 3.02). Each participant received €15.[7]
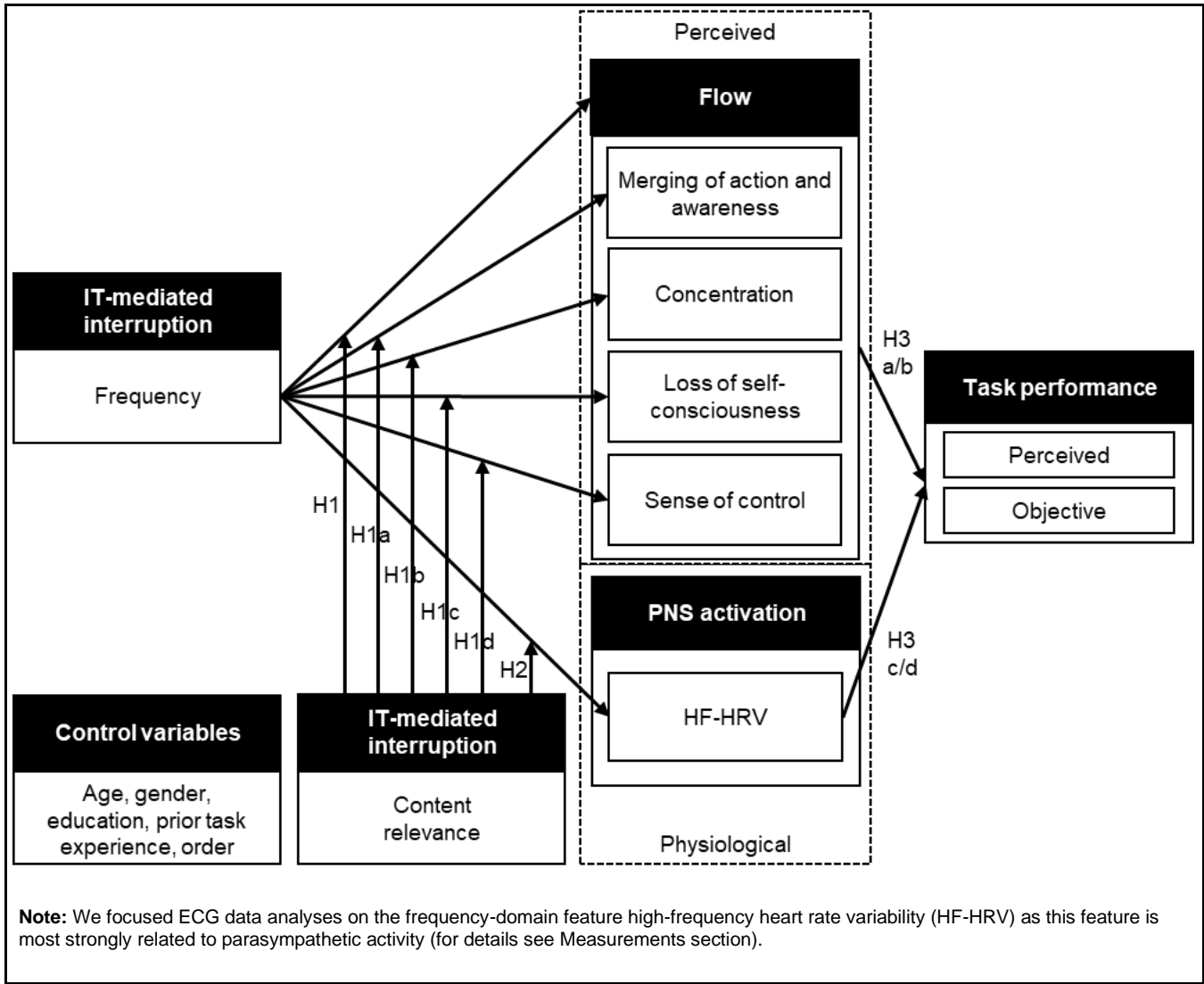

Note: We focused ECG data analyses on the frequency-domain feature high-frequency heart rate variability (HF-HRV) as this feature is most strongly related to parasympathetic activity (for details see Measurements section).

Figure 1. Research Model

Experimental Design

We used a 2 (frequency: low/high) by 2 (content relevance: irrelevant/relevant) mixed design with repeated measures on the first factor (see Figure 2). In line with our research model, we used frequency as the focal variable and content relevance as the moderator variable. This decision was based on the consideration that without an interruption, there is also no content relevance of that interruption.

Experimental Task

During the experiment, participants engaged in an invoice matching task.[8] We selected this task for two reasons. First, invoice matching represents a typical and widely used office task in companies of all sizes. Second, invoice matching has similarities to tasks employed in flow research to induce flow (Ulrich et al., 2014; Ulrich et al., 2016).

---

[8] Before the task periods started, an information period informed participants about how to conduct the task. This information period also entailed the following two aspects to clarify the goals and to motivate the participants: (1) "Your goal is to correctly solve as many invoice matching tasks as possible"

and (2) "Even though this is a difficult task, we want you to really try hard to do your best at it. Try to think of the task as a challenge to be met and overcome. And think of yourself as someone capable of meeting that challenge."

Figure 2. Experimental Design and Conditions
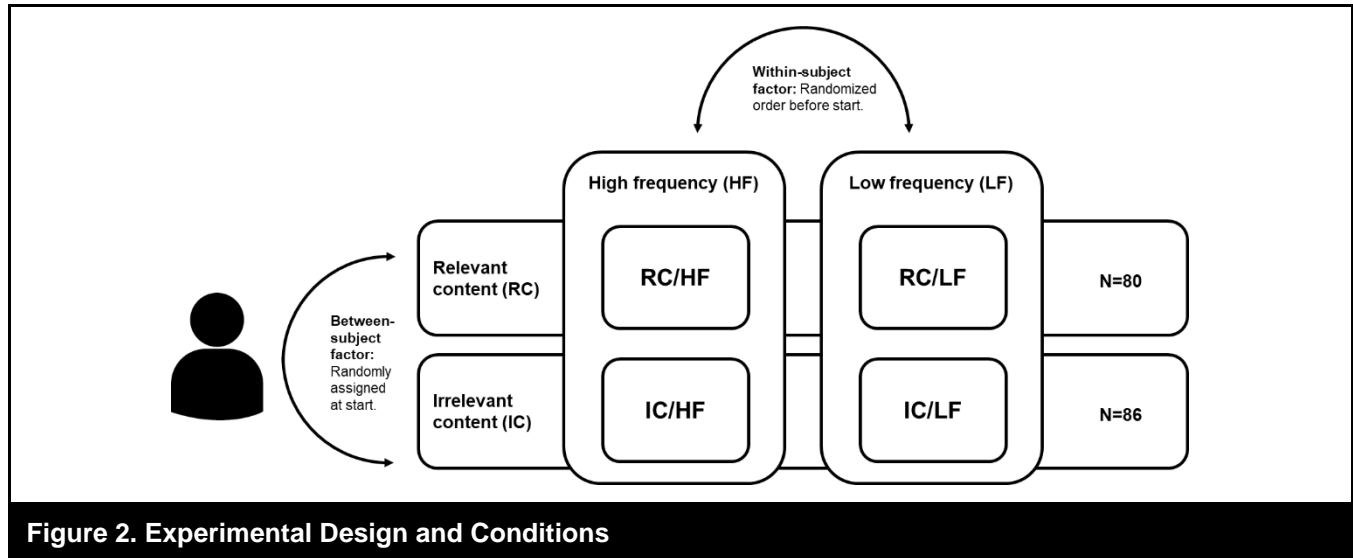
In one task period (which lasted 300 seconds[9]), participants had to match as many customer payments to corresponding invoices as possible. Given the within-subjects factor of interruption frequency, each participant completed the task twice (i.e., HF and LF; see Figure 2). Matching a payment involved two components. First, the left side of the screen displayed a customer's payment amount (e.g., €47). Second, the right side of the screen listed a total of eight invoices that the payment could be matched to. The aim was to match the correct invoices to the payment by using mental arithmetic. For instance, if 3, 44, and six more numbers were displayed on the right side of the screen, the aim of the participants was to select only those invoices for which the sum matched the payment displayed on the left side of the screen (e.g., 47; see Appendix A for a screenshot). For the solution of one invoice matching task, only addition operations (+) were necessary, and there was exactly one correct solution per task. As soon as an answer was chosen, a new payment and a new list of invoices were displayed. This procedure was repeated until the 300 seconds had expired (i.e., the task period was completed). To induce flow, the difficulty of the calculations was continuously adapted to the participant's skill level, a standard approach in flow research (Ulrich et al., 2014; Ulrich et al., 2016). Once a sequence of four invoice matching tasks in a row was solved correctly, the difficulty level increased by one, using two mechanisms: (1) If the last summand had only one digit, the level was increased by changing one solution number to a two-digit number in the next invoice matching task. For instance, if the matching task 88 = 80+**8** (Level 1) was solved correctly and the last three invoice matching tasks were also solved correctly (i.e., 4 correctly solved invoice

matching tasks in a row), the level was increased by changing one solution number to a two-digit number (e.g., 91 = 30+**61**; Level 2). (2) After four correct Level-2 matchings, an additional single-digit number was added for another level-up (e.g., 93 = 30+55+**8**; Level 3). If a sequence of four results was incorrectly solved, the difficulty decreased by one level.

Experimental Manipulations

The goal of our experiment was to establish high internal validity. This included the interruption length (i.e., 13 seconds) and type (i.e., full screen). In particular, the interruptions occurred between two invoice matching tasks (within a 300-second task period) in the form of full-screen notifications. Full-screen notifications are an established method in experimental research to ensure that participants perceive the notification and are not distracted by potential content in the background (Morgan et al., 2009; Trafton et al., 2003). The full-screen notifications were displayed for 13 seconds without providing the participants with the opportunity to close the notifications, ensuring that the interruption time was identical for all participants (Monk et al., 2008; Morgan et al., 2009). The period of 13 seconds was chosen because researchers found that the rate of resumption time (as a measure of disruptiveness of an interruption) rose rapidly in the short duration range (3-13 seconds) before diminishing over the next 45 seconds, approaching asymptote (e.g., a log function) (Monk et al., 2008). Due to the log function, only marginal additional disruptive effects can be achieved for interruptions of more than 13 seconds. As one of

---

[9] To define the task period length, we followed established research that also relied on a difficulty manipulation paradigm (Tozman et al., 2015).

Furthermore, we followed the recommendation that for HF-HRV measurements 300 seconds of data should be collected (Task Force, 1996).

our aims was to disrupt participants' flow in our experimental manipulation, we based our decision on these findings.

In the *high frequency* condition, a notification was shown after every invoice matching task, whereas in the *low frequency* condition, there were only two notifications in the entire 300-second task period. We build upon an operationalization of frequency as outlined by psychologists (Morgan et al., 2009).

In the *irrelevant* interruption condition, the system interrupted participants with notifications in the form of text derived from anonymized business emails unrelated to the task at hand (Galluch et al., 2015). In the *relevant* interruption condition, the system showed notifications containing two distinct cues that helped participants solve the task: (1) the number of invoices required for the summation of the next invoice matching task and (2) all invoice identifiers for the next solution except one (i.e., participants still had to resolve one invoice). Note that interruptions in our manipulation are not necessarily about a secondary task interfering with a primary task but about interrupting the flow state. Participants get into flow, and when they are abruptly taken out of it, it takes time and effort to get back into flow. In our aim to investigate content relevance's moderating role, our manipulation offers a stark contrast between the relevant invoice cues and the irrelevant business emails (see Appendix A for sample screenshots).

Experimental Procedure

After providing their written informed consent, each participant was equipped with the ECG device and was randomly assigned to the treatment levels of the between-subject factor (content relevance; $N = 80$ relevant content, $N = 86$ irrelevant content). To control for possible carryover effects in the within-subjects condition of interruption frequency, the sequence of this factor was also randomized independently for each participant.

The experiment was divided into different periods: (1) An information period, in which the participants were informed on how to conduct the experiment, followed by (2) a quiz period to ensure that they understood how to conduct the experimental task. (3) A practice period (180 seconds) to familiarize the participants with the invoice matching task. (4) A calibration period (300 seconds) based on which an initial difficulty level was independently determined for each participant. Specifically, a participant's average task level in the calibration period served as the starting level for each experimental round. (5) A physiological baseline period (300 seconds) in which participants were asked to breathe deeply, relax, and avoid unnecessary movements while leaving their eyes open and fixating on a black cross in the middle of the screen. And lastly, (6) the task periods accompanied by the described invoice matching task. After each task period, self-reported flow and its

dimensions were captured with a questionnaire, and objective and perceived task performance were assessed (see Appendix B for details). Between the task periods, a resting period of 60 seconds was included to avoid possible effects from the previous invoice matching tasks.
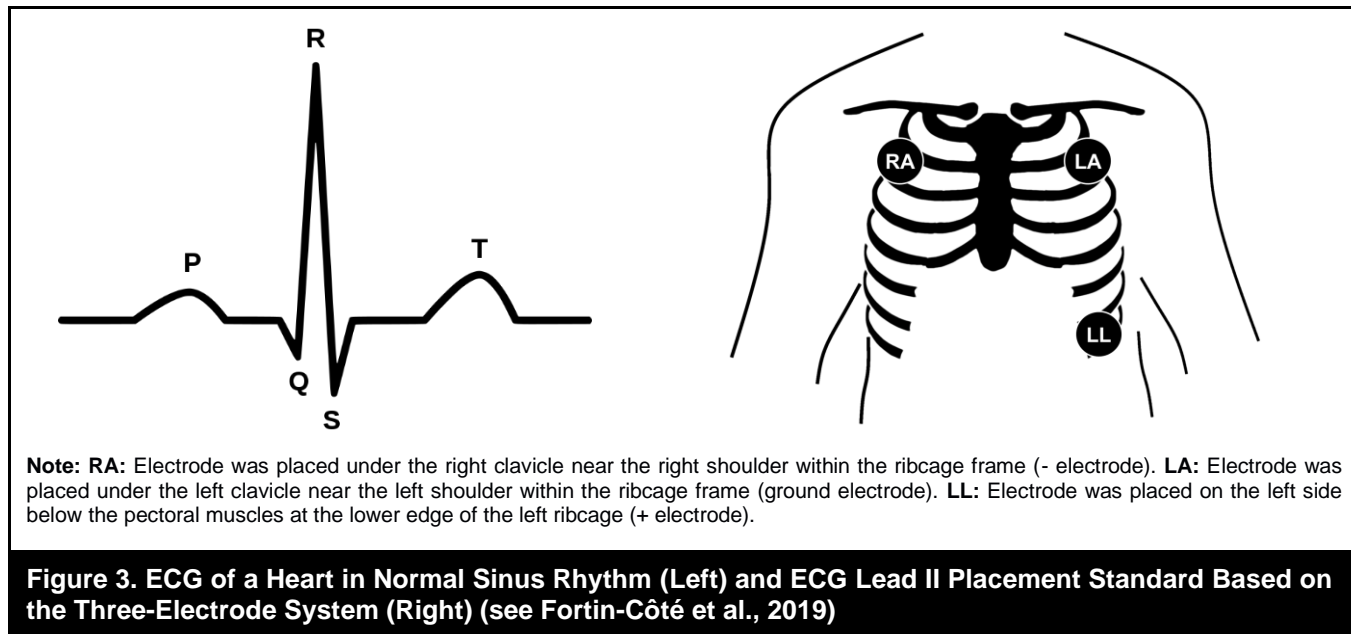
Autonomic Nervous System and Electrocardiography

The human nervous system consists of different parts. We distinguish the CNS (brain and spinal cord) and the peripheral nervous system, which comprises all neural tissue except for the CNS. The ANS is a "component of the peripheral nervous system that regulates involuntary physiologic processes including heart rate, blood pressure, respiration, digestion, and sexual arousal" (Waxenbaum et al., 2022, para. 1).

In the present paper, our focus is on the ANS which consists of three anatomically distinct divisions: the SNS, the PNS, and the enteric nervous system (ENS) (Waxenbaum et al., 2022). The ENS is less relevant for this study, as it governs the function of the gastrointestinal tract. The SNS is responsible for implementing a "fight-or-flight" response in stress situations, while the PNS implements a "rest-and-digest" response. SNS activation is related to the release of stress hormones, increases the heart's contraction rate and force (cardiac output), but decreases HRV. Conversely, PNS activation reduces heart rate, but increases HRV to restore homeostasis after a state of physiological activation (Berntson et al., 2007; Valentini & Parati, 2009). Importantly, this interplay between the sympathetic and parasympathetic division of the ANS makes it possible (in healthy humans) for the heart to instantaneously respond to different situations and needs (Task Force, 1996).

In this context, new sensors enable the capture of a wide range of human signals that can provide information about user states like flow, providing more information than ever before to interpret human communication and interaction (Schultz et al., 2013). Following Schultz et al. (2013), we define such signals as autonomous, energetic-materially measurable physical quantities generated by the living organism. They originate from chemical and physical actions of the human body and serve to control, regulate, and transmit information in the human organism. They are measured in different quantities depending on their origin. An important category is represented by electrical human signals (Loewe & Nadj, 2020) for recording brain activity (electroencephalography, EEG), eye activity (electrooculography, EOG), heart activity (electrocardiography, ECG), muscle activity (electromyography, EMG), and skin conductance (electrodermal activity, EDA, for an overview see also Table 1).

Table 1. Overview of Sensor Devices and Sensor Choice of this Study					
Activity	**Brain**	**Eye**	**Heart**	**Muscle**	**Skin**
Sensor device	Electroencephalography (EEG)	Electrooculography (EOG)	Electrocardiography (ECG)	Electromyography (EMG)	Electrodermal activity (EDA)
Sensor choice	Focus on ECG given its particular contribution to: Flow variance explanation High robustness Low obtrusiveness				

Note: **RA:** Electrode was placed under the right clavicle near the right shoulder within the ribcage frame (- electrode). **LA:** Electrode was placed under the left clavicle near the left shoulder within the ribcage frame (ground electrode). **LL:** Electrode was placed on the left side below the pectoral muscles at the lower edge of the left ribcage (+ electrode).

Figure 3. ECG of a Heart in Normal Sinus Rhythm (Left) and ECG Lead II Placement Standard Based on the Three-Electrode System (Right) (see Fortin-Côté et al., 2019)

In this paper, we explicitly decided to focus on ECG given its particular contribution to flow variance explanation, its high robustness (i.e., less prone to noise, as well as signal processing and data interpretation errors), and its low obtrusiveness, which is an important factor for the ongoing study of flow in more ecologically valid scenarios (Riedl et al., 2014).

As mentioned, ECG measures the electrical activity of the heart. Based on a heartbeat, an electrical impulse travels through the heart. This impulse causes the heart to pump blood. A normal heartbeat on ECG will show the timing of the top and lower chambers of the heart. The right and left atria or upper chambers produce the first wave, referred to as the *P wave*, following a flat line when the electrical impulse goes to the bottom chambers. The right and left bottom chambers, or ventricles, produce the next wave referred to as the *QRS complex*. The final wave is referred to as the *T wave*, representing electrical recovery or return to a resting state for the ventricles. Figure 3

shows an illustration of an ECG of a heart in normal sinus rhythm, as well as the ECG lead II placement standard based on the three-electrode system, as used in the present study. Further information on the physiological foundations of the ECG can be found in Wagner and Strauss (2014).

Measurements

Self-reported flow: After each task period, participants filled in Jackson and Marsh's (1996) flow questionnaire. The questionnaire's reliability and validity have been widely confirmed (Riva et al., 2017).

PNS activation: ECG was used to record heart activity during task execution (1,000 Hz, Bioplux device; PLUX wireless Biosignals S. A., 2018). A variety of features calculated from ECG recordings exist that allow for the assessment of HRV (Task Force, 1996). Most of these features are calculated either

in the time domain or the frequency domain of the signal. We focused our ECG data analyses on the high-frequency heart rate variability (HF-HRV) frequency-domain feature as this feature is most strongly related to parasympathetic activity (Berntson et al., 2007; Valentini & Parati, 2009). In general, by using a fast Fourier transformation, the frequency band components of the heart can be separated, and the frequency-related HRV features can be calculated. In the case of HF-HRV, the power of the signal lies between 0.15-0.4 Hz. HF-HRV was computed with an established Python package (Bartels, 2018).[10] Following established guidelines (Harmat et al., 2015; Tozman et al., 2015), change scores were used to analyze HF-HRV (ΔHF-HRV = HF-HRV$_{Task}$ − HF-HRV$_{Baseline}$).

Task performance: Perceived task performance was adapted from an instrument developed by Novak and Hoffman (2009). Objective task performance refers to a participant's performance score regarding the difficulty and number of correct answers (see Appendix B).

Control variables: We collected demographics (age, gender, education, and prior task experience) to account for interpersonal factors. We also controlled for the within-subject factor order (see Appendix C). Lastly, we kept the lighting, noise, and temperature in the laboratory constant across all sessions.

Manipulation checks: To confirm that our manipulations were successful, we measured participants' perceived interruption frequency and perceived content relevance after each experimental condition (see Appendix B).

Next, we present our results, which were derived with SPSS 25.0 and Python 3.8.5.

Results

MANOVA, ANOVA, and Correlation Analyses

Next, we tested our hypotheses using multivariate analysis of variance (MANOVA), analysis of variance (ANOVA), and Pearson correlation analyses. We confirmed that the assumptions underlying these analyses hold for our data (Appendix C).

Manipulation checks: Participants reported significantly higher perceived frequency in the high interruption frequency condition than in the low frequency condition. This holds for irrelevant content (M_{HF} = 5.888, SD = 1.052; M_{LF} = 3.221, SD = 1.290; $F(1,85)$ = 244.9, $p < 0.001$, η_p^2 = 0.742) and for relevant content (M_{HF} = 4.975, SD = 1.570; M_{LF} = 2.650, SD = 1.043; $F(1,79)$ = 123.7, $p < 0.001$, η_p^2 = 0.610). Further, participants reported significantly higher perceived relevance in the relevant condition than in the irrelevant content condition. This holds for low interruption frequency ($M_{Relevant}$ = 5.308, SD = 1.634; $M_{Irrelevant}$ = 1.512, SD = 1.009; F(1,164) = 329.167, $p < 0.001$, η_p^2 = 0.667) and for high interruption frequency ($M_{Relevant}$ = 6.03, SD = 0.979; $M_{Irrelevant}$ = 1.60, SD = 0.995; $F(1,164)$ = 836.5, $p < 0.001$, η_p^2 = 0.836). Thus, both manipulations were successful.

Control variables: We considered the control variables of age, gender, education, prior task experience, and order as covariates in a MANCOVA (Mangalaraj et al., 2014; Qiu & Benbasat, 2005). None of the control variables significantly affected flow. Hence, none of them were included as covariates in the further analysis (Qiu & Benbasat, 2005) (see Appendix C).

Impact of frequency and content relevance on self-reported flow (H1-H1d): To test our hypotheses, we followed a three-step procedure: In the first step, we conducted a two-way mixed MANOVA to evaluate the hypothesized *moderation effect* of content relevance on the relationship between interruption frequency and the flow dimensions: *merging of action and awareness* (H1a), *concentration* (H1b), *loss of self-consciousness* (H1c), and *sense of control* (H1d). In our MANOVA analysis, we therefore modeled the flow dimensions as dependent variables. The composite dependent variable (i.e., self-reported flow) was created automatically by the MANOVA standard procedure in SPSS (Hair et al., 2014; Verma, 2016). A two-way mixed MANOVA was used, as one of our independent variables was manipulated within-subjects (interruption frequency) and one was manipulated between-subjects (content relevance). The results were interpreted using Wilks' lambda (Λ) (Hair et al., 2014). The analysis confirms a significant moderation effect of content relevance on the relationship between interruption frequency and self-reported flow on the multivariate level ($F(4, 161)$ = 4.96, $p < 0.001$, Wilks' Λ = 0.890, η_p^2 = 0.109, power = 0.957; H1 = supported; see Table 2).

---

[10] In this Python package, HF-HRV is calculated by using the power spectral density of the intervals between R waves (RRi) and by calculating the area under the curve (AUC) of the high frequency bands. AUC is calculated using the trapezoidal method of numpy (Bartels, 2018). The documentation also references to the following two sources of the algorithm: Cokelaer and Hasch (2017) and Task Force (1996).

Table 2. MANOVA Summary Table						
HYP	**IVs**	**DVs**	**_F_-statistic**	**_p_-value**	η_p^2	**Power**
H1	Frequency × relevance	Self-reported flow (dimensions: (1) Merging of action and awareness, (2) concentration, (3) loss of self-consciousness, and (4) sense of control)	$F(4,161) = 4.96$	$p < .001$	.109	.957

Table 3. ANOVA Summary Table; Frequency × Relevance (IVs) on Flow Dimensions						
HYP	**DV**	**_F_-statistic**	**_p_-value**	η_p^2	**Power**	**Low vs. high freq. [Content relevance]**
H1a	MAA	$F(1,164) = 12.849$	$p < .001$	.072	.946	3.70 vs 4.01 [RC] 3.53 vs 3.07 [IC]
H1b	CON	$F(1,164) = .152$	$p = .697$	<.001	.067	5.27 vs 4.44 [RC] 4.94 vs 4.01 [IC]
H1c	LSC	$F(1,164) = 1.974$	$p = .162$	<.05	.287	5.20 vs 4.97 [RC] 4.43 vs 4.56 [IC]
H1d	CTR	$F(1,164) = 4.589$	$p = .034$	.027	.567	5.04 vs 4.97 [RC] 4.75 vs 4.26 [IC]

Note: MAA = Merging of action and awareness, CON = Concentration, CTR = Sense of control, LSC = Loss of self-consciousness; RC = relevant content; IR = irrelevant content

Table 4. ANOVA Summary Table; Interruption Frequency (IV) on Flow Dimensions					
DV	**_F_-statistic**	**_p_-value**	η_p^2	**Power**	**Low vs. high freq.**
Concentration	$F(1,164) = 45.847$	$p < .001$	.218	.999	5.10 vs 4.22
Loss of self-consciousness	$F(1,164) = .135$	$p = .714$	<.001	.065	4.80 vs 4.76

In the second step, as the analysis confirmed a significant moderation effect of content relevance on the relationship between interruption frequency and self-reported flow (multivariate level), we conducted several two-way mixed ANOVAs to assess the moderation effects of content relevance on the relationship between interruption frequency and each flow dimension (univariate level). This confirmed a significant moderation effect of content relevance on the relationship between interruption frequency and two flow dimensions, namely *merging of action and awareness* ($F(1, 164) = 12.849, p < 0.001, \eta_p^2 = 0.072$, power = 0.946; H1a = supported) and *sense of control* ($F(1, 164) = 4.589, p = 0.034, \eta_p^2 = 0.027$, power = 0.567; H1d = supported). There was no support for a moderation effect of content relevance on the relationship between interruption frequency and the other two flow dimensions, *concentration* ($F(1, 164) = 0.152, p = 0.697, \eta_p^2 < 0.001$, power = 0.067; H1b = not supported) and *loss of self-consciousness* ($F(1, 164) = 1.974, p = 0.162, \eta_p^2 < 0.05$, power = 0.287; H1c = not supported; see Table 3).

As we found no significant moderation effect for *concentration* and *loss of self-consciousness*, we reported the main effects of interruption frequency and content relevance on these flow dimensions (Pituch & Stevens, 2016). There was a significant main effect of interruption frequency on *concentration* ($F(1, 164) = 45.847, p < 0.001, \eta_p^2 = 0.218$, power = 0.999). However, there was no significant main effect of interruption frequency on *loss of self-consciousness* ($F(1, 164) = 0.135, p = 0.714, \eta_p^2 < 0.001$, power = 0.065; see Table 4).

Regarding the main effects of content relevance, we found a significant effect on *concentration* ($F(1, 164) = 4.425, p = 0.036, \eta_p^2 = 0.0262$, power = 0.552) and *loss of self-consciousness* ($F(1, 164) = 7.992, p = 0.005, \eta_p^2 = 0.0464$, power = 0.802; see Table 5).

Table 5. ANOVA Summary Table; Content Relevance (IV) on Flow Dimensions					
DV	*F*-statistic	*p*-value	η_p^2	Power	Relevant vs. irrelevant
Concentration	$F(1,164) = 4.425$	$p = .036$	.0262	.552	4.85 vs 4.47
Loss of self-consciousness	$F(1,164) = 7.992$	$p = .005$	.0464	.802	5.08 vs 4.49

In the third step, as the analysis confirmed a significant moderation effect of content relevance on the relationship between frequency and the flow dimensions *merging of action and awareness* and *sense of control* (see Table 3), we calculated the simple main effects of frequency at each content relevance level (relevant and irrelevant) by using one-way repeated measures ANOVAs to assess where the differences occurred (Hair et al., 2014; Verma, 2016). Regarding the merging of action and awareness, for *irrelevant* interruptions, high frequency yielded a lower merging of action and awareness (*Mean (M)* = 3.07, *SD* = 1.40) than low frequency (*M* = 3.53, *SD* = 1.49; *F*(1, 85) = 10.303, *p* = 0.002, $\eta_p^2 = 0.108$, power = 0.887). For *relevant* interruptions, the merging of action and awareness was even slightly *higher* for high frequency (*M* = 4.01, *SD* = 1.47) than for low frequency (*M* = 3.70, *SD* = 1.39). However, this difference was only marginally significant (*F*(1, 79) = 3.723, *p* = 0.057, $\eta_p^2 = 0.045$, power = 0.478). Regarding sense of control, for *irrelevant* interruptions, high frequency yielded a lower sense of control (*M* = 4.26, *SD* = 1.43) than low frequency (*M* = 4.75, *SD* = 1.22; *F*(1, 85) = 12.887, *p* < 0.001, $\eta_p^2 = 0.132$, power = 0.943). By contrast, for *relevant* interruptions, there was no significant difference for sense of control between high frequency (*M* = 4.97, *SD* = 1.28) and low frequency (*M* = 5.04, *SD* = 1.23; *F*(1, 79) = 0.226, *p* = 0.636, $\eta_p^2 = 0.003$, power = 0.075; see Figure 4).
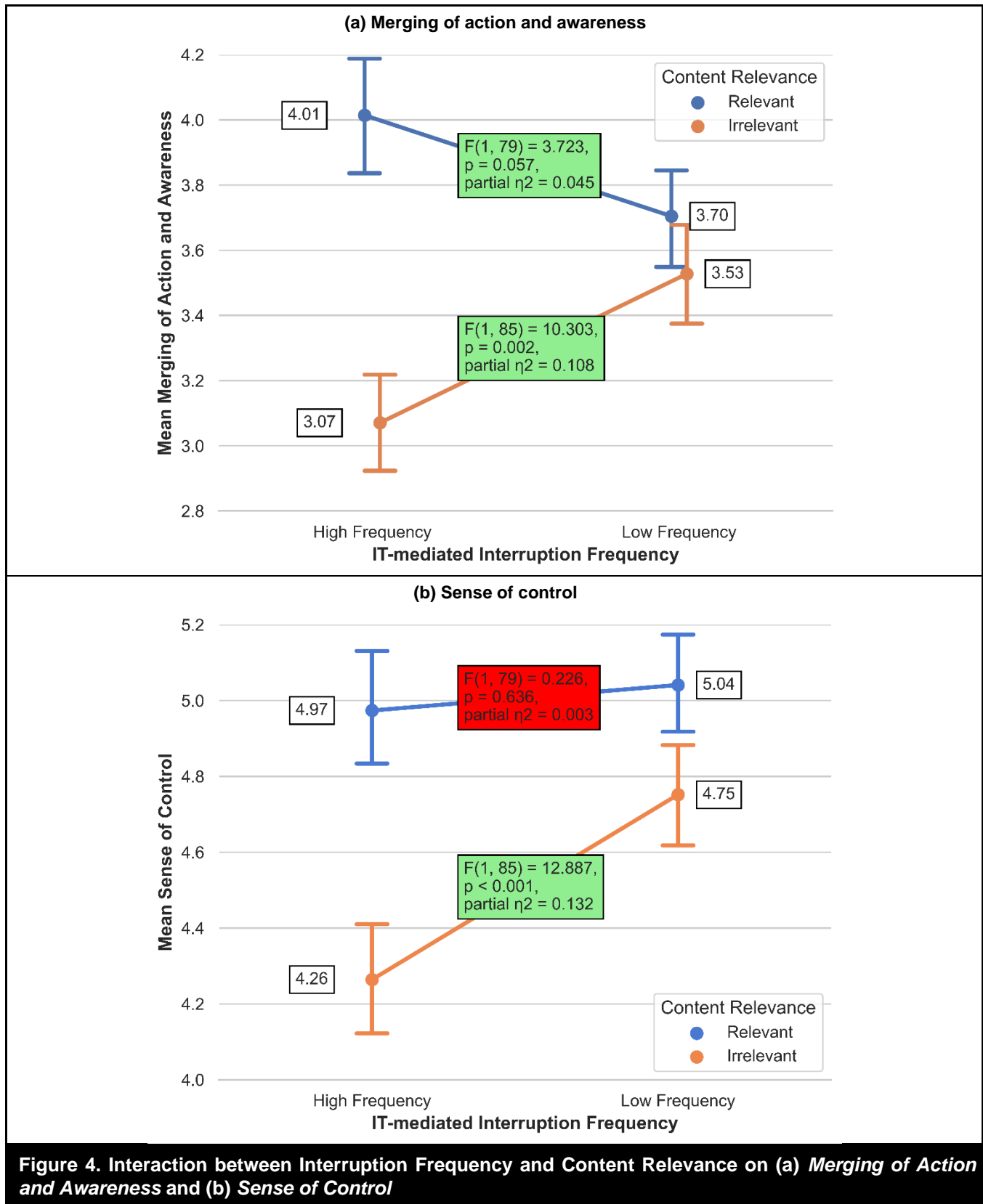
Impact of frequency and content relevance on PNS activation (H2): After data screening, eight participants were removed from the original sample of *N* = 166 due to an ECG recording error. The ECG measurement did not offer any data in this case, meaning that the continuous data stream was null. This indicates that the respective participants (knowingly or unknowingly) removed one or more electrodes or switched off the recording device. Furthermore, outliers were identified and removed based on the 1.5 IQR criterion (Tukey, 1977). This led to a final dataset of 129 participants (53 women, 76 men) with valid ECG recordings. Following recommendations in the cardiac data preparation literature (Berntson et al., 2007) and the applications of those recommendations in empirical research (Harmat et al., 2015), the cardiac features were natural log transformed to normalize their distribution. Further, we

confirmed that the data were in line with the assumptions underlying these analyses (see Appendix C).

To test H2, we applied a two-way mixed ANOVA (see Table 6). The analysis confirmed a marginally significant moderation effect of content relevance on the relationship between interruption frequency and ΔHF-HRV (H2 = partially supported).

Next, we calculated the simple main effects of interruption frequency at each level of content relevance (relevant and irrelevant) using repeated measures ANOVAs for ΔHF-HRV to assess where the differences occurred (see Figure 5). For *irrelevant* interruptions, high frequency yielded a lower ΔHF-HRV (*M* = -0.51, *SD* = 0.558) than low frequency (*M* = -0.29, *SD* = 0.522; *F*(1, 65) = 16.018, *p* < 0.001, $\eta_p^2 = 0.198$, power = 0.976). For relevant interruptions, there was no significant difference between high frequency (*M* = -0.38, *SD* = 0.519) and low frequency (*M* = -0.30, *SD* = 0.506; *F*(1, 62) = 2.491, *p* = 0.120, $\eta_p^2 = 0.039$, power = 0.343).

Correlation analyses (H3a-d): Our analysis confirms positive correlations for self-reported flow with perceived task performance (*r* = 0.534, *p* < 0.001, large effect size; H3a = supported) and objective task performance (*r* = 0.153, *p* = 0.003, small effect size; H3b = supported; see Table 7). Here, self-reported flow was calculated as the average of the four dimensions. Positive correlations were also confirmed for all four dimensions, both for perceived (MAA: *r* = 0.351, CON: *r* = 0.348, CTR: *r* = 0.413, LSC: *r* = 0.373, all *p* < 0.001, medium effect size) and objective task performance (MAA: *r* = 0.103, p = 0.030, CON: *r* = 0.109, *p* = 0.024, CTR: *r* = 0.114, *p* = 0.019, LSC: *r* = 0.096, *p* = 0.040, small effect size). For ΔHF-HRV, we observed a significant correlation with CTR (*r* = -0.128, *p* = 0.020) but not with the other flow dimensions. Neither perceived (*r* = -0.026, *p* = 0.339; H3c = not supported) nor objective task performance (*r* = -0.053, *p* = 0.200; H3d = not supported) exhibited significant correlations with ΔHF-HRV. In summary, Table 8 provides an overview of the hypotheses and the empirical support for each hypothesis.
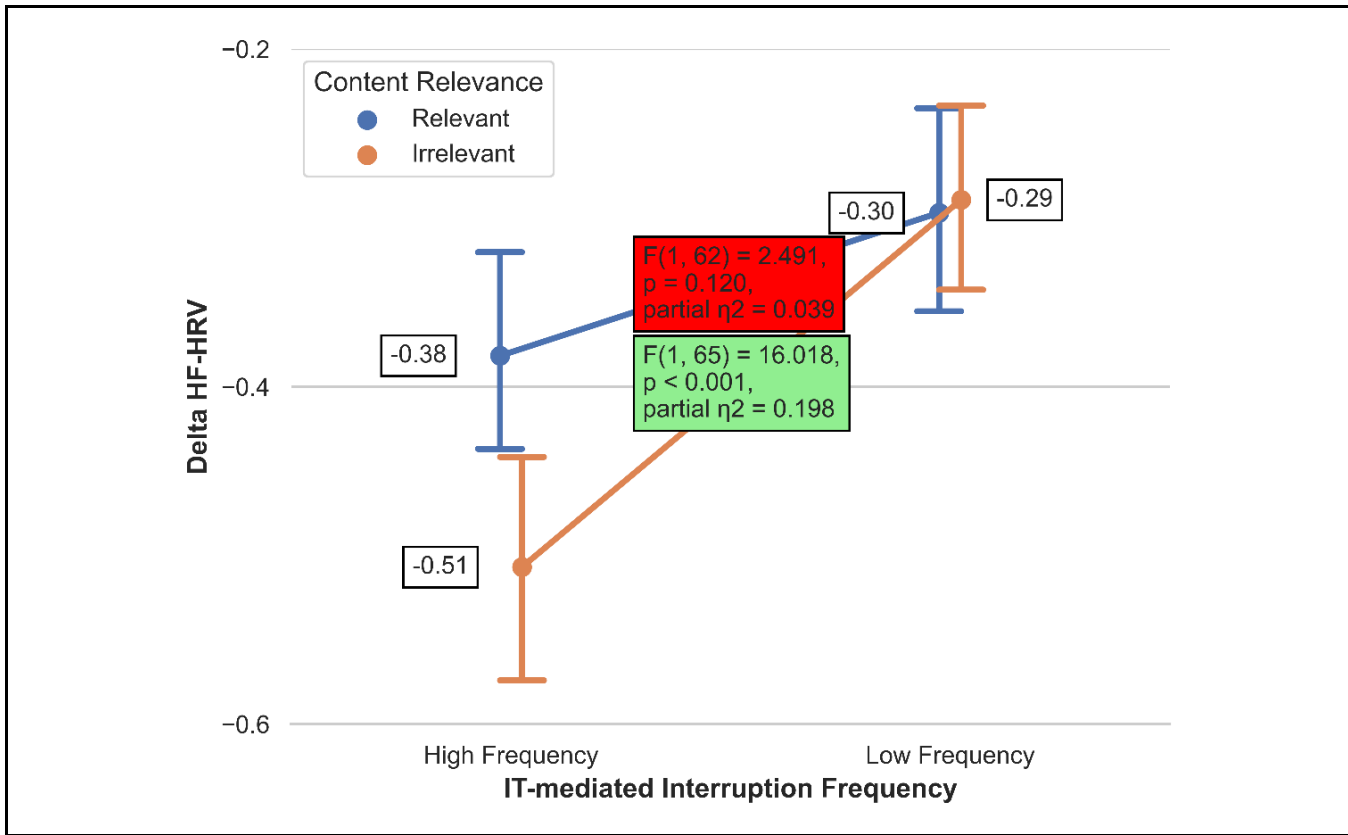
Figure 4. Interaction between Interruption Frequency and Content Relevance on (a) *Merging of Action and Awareness* **and (b)** *Sense of Control*

Table 6. ANOVA Summary Table; Frequency × Relevance on PNS Activation							
HYP	DV	IVs	*F*-statistic	*p*-value	η_p^2	Power	Low vs. high freq. (Content relevance)
H2	ΔHF-HRV	Frequency × relevance	$F(1, 127) = 3.017$	$p = .085$	.023	.407	-.30 vs -.38 [RC] -.29 vs -.51 [IC]

Note: RC = relevant content; IR = irrelevant content

Figure 5. Interaction between Interruption Frequency and Content Relevance on ΔHF-HRV

Table 7. Correlation Table; Results for H3a-d Are Shaded in Gray								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
PPE	1							
OPE	.147***	1						
MAA	.351***	.103*	1					
CON	.348***	.109*	.350***	1				
CTR	.413***	.114*	.376***	.595***	1			
LSC	.373***	.096*	.223***	.182***	.285***	1		
FLW	.534***	.153**	.670***	.721***	.757***	.604***	1	
ΔHF-HRV	-.026	-.053	-.071	-.072	-.128*	.004	-.068	1

Note: Pearson correlations (one-tailed). * = $p < 0.05$; ** = $p < 0.01$; *** = $p < 0.001$; MAA = merging of action and awareness, CON = concentration, CTR = sense of control, LSC = loss of self-consciousness, FLW = self-reported flow, PPE = perceived task performance, OPE = objective task performance; ΔHF-HRV = delta high-frequency heart rate variability; Sample of $N = 166$ (except for correlations involving ΔHF-HRV where all involved constructs refer to the Sample of $N = 129$)

Table 8. Results of the Hypotheses Testing		
HYP	**Description**	**Support**
H1	Content relevance moderates the impact of interruption frequency on flow, such that an inhibiting effect is present for irrelevant content.	Supported
H1a	Content relevance moderates the impact of interruption frequency on the merging of action and awareness, such that an inhibiting effect is present for irrelevant content.	Supported
H1b	Content relevance moderates the impact of interruption frequency on concentration, such that an inhibiting effect is present for irrelevant content.	Not supported
H1c	Content relevance moderates the impact of interruption frequency on loss of self-consciousness, such that an inhibiting effect is present for irrelevant content.	Not supported
H1d	Content relevance moderates the impact of interruption frequency on sense of control, such that an inhibiting effect is present for irrelevant content.	Supported
H2	Content relevance moderates the impact of interruption frequency on PNS activation, such that an inhibiting effect is present for irrelevant content.	Partially supported
H3a	Flow is positively correlated with perceived task performance.	Supported
H3b	Flow is positively correlated with objective task performance.	Supported
H3c	PNS activation is positively correlated with perceived task performance.	Not supported
H3d	PNS activation is positively correlated with objective task performance.	Not supported

Moderated Mediation Analyses

As our research model implies moderated mediation, we conducted several moderated mediation analyses (Hayes, 2021) in order to test whether self-reported flow or ΔHF-HRV mediates perceived interruption frequency's effect on (objective or perceived) task performance with perceived content relevance as a moderator. Using Model 7 (see Figure 6) of the PROCESS macro (Hayes, 2021), we found a significant moderated mediation effect of perceived interruption frequency (as the independent variable) on *perceived task performance* (as the dependent variable) through perceived content relevance (as moderator) and self-reported flow (as mediator; measured as the average of the four dimensions). Specifically, the lower and upper confidence intervals (LLCI and ULCI) of the calculated moderated mediation index did not include zero (LLCI = 0.000041; ULCI = 0.033882).[11] Note that there was no significant direct effect of perceived interruption frequency on perceived task performance ($p = 0.177$). Using Models 14, 8, and 58 of the PROCESS macro (Hayes, 2021), we found no evidence for a significant moderated mediation effect.

Using Model 7 (see Figure 6) of the PROCESS macro (Hayes, 2021), we also found a significant moderated mediation effect of perceived interruption frequency (as the independent variable) on *objective task performance* (as the dependent variable) through perceived content relevance (as moderator) and self-reported flow (as mediator) (LLCI = 0.000186; ULCI = 0.012255). Note that there was no significant direct effect of perceived interruption frequency on objective task performance ($p = 0.216$). Using Model 14, 8, and 58 of the PROCESS macro (Hayes, 2021), we found no evidence for a significant moderated mediation effect.

In contrast, we found no evidence for a moderated mediation effect using various models (Model 7, 14, 8, 58) of the PROCESS macro (Hayes, 2021) when using ΔHF-HRV as the mediator.

## Discussion ▮▮▮▮▮▮

Theoretical Implications

Our research has several implications for theory. We systematically extend flow theory by showing that interruption frequency and content relevance are important flow antecedents, as we confirmed that content relevance moderates the impact of interruption frequency on self-reported flow (H1). This contribution is valuable, as scholars have suggested an interruption-flow link that has not yet been empirically confirmed (Jett & George, 2003; Rissler et al., 2017a).

---

[11] Note that for the moderated mediation analyses, we used the perceived measures of interruption frequency and content relevance (see Appendix B, Table B2 for their operationalization), whereas for the (M)ANOVA analyses, we relied on the experimental manipulations of interruption frequency and content relevance.
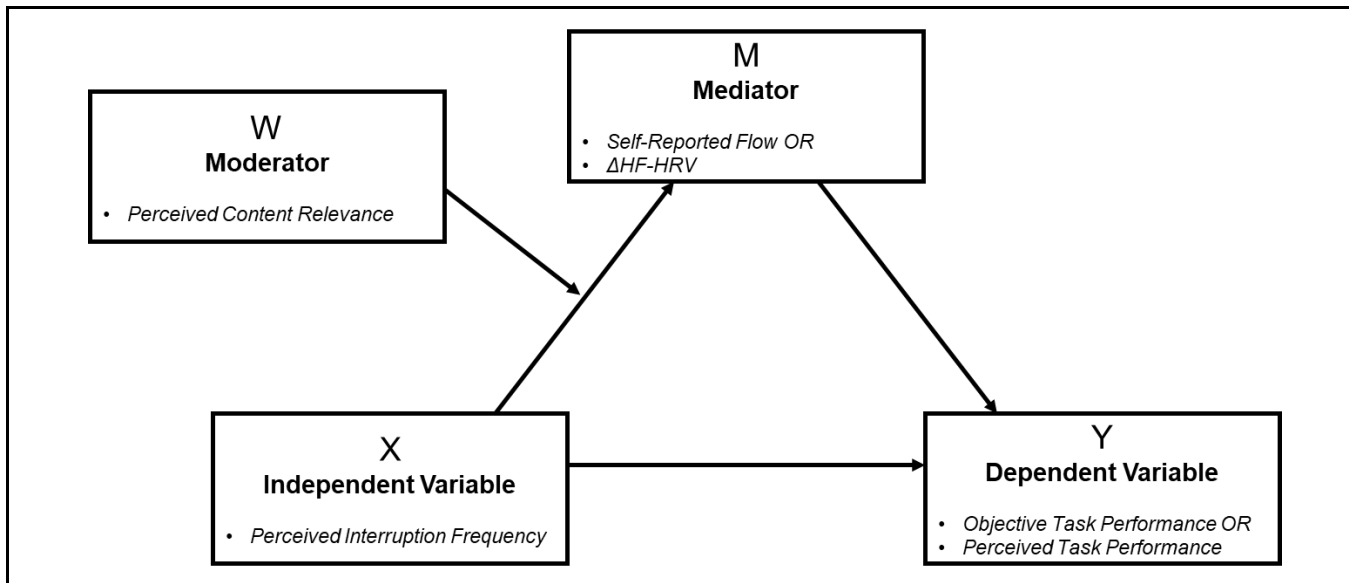
Figure 6. Model 7 of Hayes's PROCESS Macro

Thus far, no study, including major NeuroIS studies (see Related Work section), has empirically examined how interruption frequency and content relevance impact different dimensions of self-reported flow. This contribution is important because complex patterns became visible that would not have been revealed without this focus. Our data reveal that content relevance moderates the relationship of interruption frequency on the *merging of action and awareness* (H1a) and *sense of control* (H1d). The follow-up analysis of the simple main effects revealed that if the interruption content is irrelevant to the task at hand, higher interruption frequency decreases the merging of action and awareness and sense of control. Conversely, if the interruption content is relevant to the task, higher interruption frequency neither impairs the merging of action and awareness nor the person's sense of control. Notably, for relevant interruptions, we even found marginally significant evidence indicating that interruption frequency has a positive effect on the merging of action and awareness. A possible explanation for this finding could be that as long as the new stimulus conveys task-relevant content, it will hardly be perceived as an interruption and the person will remain merged with the task. For sense of control, there was *no* significant difference between high and low interruption frequency for relevant content. This finding suggests that notifications conveying cues to solve the task increase perceived control over the task. Content relevance's moderating influence on the relationship between interruption frequency and the merging of *action and awareness*, as well as *sense of control*, was thereby so prominent that it also had a significant moderating effect on the relationship between *interruption frequency* and *self-reported flow* at the multivariate level (H1).

Further, we contribute to understanding how interruptions influence human physiology by addressing the flow phenomenon from the biological view. Riedl (2013, p. 47) noted:

> *I believe that biology offers a valuable knowledge base for the investigation of IS phenomena, as demonstrated in the present article, which is based on the example of technostress. If technology is the users' foe, corresponding stress perceptions can be objectively measured. However, making technology the users' friend must be a major goal of IS research. Whether or not a specific technology is user-friendly can also be investigated by means of biological approaches.*

In our article, we study flow, which—unlike stress— is a state that makes users feel like technology is a "friend" rather than a "foe." Moreover, unlike Ortiz de Guinea and Webster (2013), who examined IS use patterns and included a biological view by using heart rate as a proxy for physiological arousal, we used an HRV measure similar to Léger et al. (2014a). Importantly, a recent literature review revealed that NeuroIS research has used HRV much less frequently than heart rate (Stangl & Riedl, 2022). Considering that HRV captures important information for psychological constructs in general and IS constructs in particular (e.g., Berntson et al., 2007; Valentini & Parati, 2009), this constitutes a shortcoming in the current NeuroIS literature, which the present study contributes to overcoming.

Specifically, we found marginally significant evidence for a moderating effect of content relevance on the relationship between interruption frequency and HF-HRV as a proxy for PNS activation (H2). Analyzing the simple main effects showed that higher interruption frequency significantly reduced PNS activation for irrelevant content. Note that this finding contrasts with Ortiz de Guinea and Webster (2013). Contrary to their hypothesis of an increased heart rate in response to negative stimuli (i.e., discrepant IT events), Ortiz de Guinea and Webster (2013) found the opposite, a decrease in heart rate. As an explanation for this *unexpected* result, the authors noted: "discrepant IT events may trigger participants to automatically pay attention to the stimuli, reducing bodily activity and thus, heart rate" (p. 1176). In our case, negative stimuli (i.e., highly frequent irrelevant interruptions) decreased HF-HRV, which would correspond to elevated arousal. Following the theoretical considerations from our Physiology of Flow section, our finding could be explained by a frequent occurrence of physiologically demanding conflict monitoring processes that arise (more often) when processing task-irrelevant information (Botvinick et al., 1999; Botvinick et al., 2004). In contrast, both with low interruption frequencies and with relevant information, individuals can more easily integrate the presented additional information into the primary task, which results in a more efficient allocation of attentional resources—a pattern that is considered to be a dominant characteristic of the flow experience (Harris et al., 2017b; Tozman & Peifer, 2016). This comes along with increased PNS activation (Mark et al., 2012; Mark et al., 2016). These observations offer a valuable basis for the development of future research designs that evaluate the effects of different interruption types on physiological and experiential factors. Léger et al.'s (2014b) innovative EFRP method merits future consideration by scholars interested in (1) studying direct and unmediated IT effects upon humans' cognitive processing while using IS, (2) measuring individuals' cognitive processes in a relatively unobtrusive manner at the time of their occurrence, and/or (3) measuring automatic cognitive processes that may occur outside human awareness (see also Limitations and Future Research section).

We also contribute to flow theory by also studying task performance as a major flow outcome. Our results confirm that self-reported flow is positively correlated with both perceived and objective task performance in an invoice matching task (H3a and H3b), a previously unexamined office work context. Our findings position flow as a proxy for desirable outcomes (i.e., higher task performance) in such contexts. This result extends the results of prior studies where performance was conceptualized as service quality (Kuo & Ho, 2010) or work creativity (Zubair & Kamal, 2015). Further, we contribute to the literature by being among the first to develop an objective performance measure in the office work context based on a difficulty manipulation paradigm (for details, see Appendix B).

We further contribute to the understanding of the relationship between PNS activation and task performance. Two studies (Chin & Kales, 2019; Harris et al., 2017a) have shown that high or moderate PNS activation is positively linked to high task performance. However, our results showed that PNS activation (measured by HF-HRV) correlates neither with perceived nor objective task performance (H3c and H3d). Following the Yerkes-Dodson law (Yerkes & Dodson, 1908) scholars have established that different tasks require varying arousal levels for optimal performance. For instance, difficult tasks demand moderate arousal levels, while simpler tasks can be better accomplished at higher arousal levels. However, it is possible that some participants perceived our experimental task as rather easy to perform, whereas others did not, leading to different arousal levels, which, in turn, might have affected the PNS-performance relationship (Yerkes & Dodson, 1908). Future research could further study the *interplay* between PNS and SNS activation because the latter necessarily accompanies physiological arousal (e.g., Gunnar & Quevedo, 2007).

Finally, in an effort to provide a comprehensive analysis, we also explored the possibility of a linear relationship between PNS activation and self-reported flow with a correlation analysis (see Table 7). We found no linear relationship between PNS activation (measured by HF-HRV) and self-reported flow. This result is consistent with Léger et al. (2014a), who also found no significant correlation between HRV and cognitive absorption ($r = 0.16$, ns.). Interestingly, no significant correlations with cognitive absorption were found for EEG alpha ($r = 0.09$, ns.) or EDA ($r = -0.15$; ns.) either. Only two of the five neurophysiological measures, heart rate ($r = -0.24$; $p < 0.05$) and EEG beta ($r = -0.28$; $p < 0.05$), correlated with cognitive absorption. However, in their regression analysis, Léger et al. showed that the inclusion of these five neurophysiological measures over and above skill, difficulty, and their interaction "increased the amount of variance explained from 16% (skills and difficulty alone) to 34% for the comprehensive CA [cognitive absorption] construct" (p. 279) and that each of these five measures explained unique variance. We believe that these and our own results contribute to an important discussion in the NeuroIS literature. In essence, a fundamental question is whether physiological evidence constitutes either an *alternative* or a *complement* to self-reports. As an alternative, the implication would be that both approaches (physiological and self-report) capture the same dimension of an underlying construct; as a complement, the implication would be that both approaches capture different dimensions of a construct. In a study involving an interaction task in which participants were asked to acquaint themselves with another person (computer-mediated vs. face-to-face), subjective anxiety and arousal were higher in the face-to-face condition than in the computer-mediated condition, but physiological arousal (measured by heart rate and skin conductance) showed no significant differences between conditions (Shalom et al., 2015).

Reviewing the technostress literature, Riedl (2013) concluded:

> *Biological measurement is indispensable. Empirical evidence shows that **conscious** stress perceptions of humans, measured by means of self-report instruments ... often do not correlate with the usually **unconscious** elevations of stress hormones ... This finding, reported in the literature on general stress research, has been replicated in several technostress studies that combine biological and self-report measures* (p. 46, emphasis in the original in italics, here in bold font).

Moreover, we refer the reader to a pathbreaking paper by Tams et al. (2014). For the context of technostress, they found that both kinds of data (physiological and self-report) tap into different aspects of this construct "and that, together, they can yield a more complete or holistic understanding of the impact of technostress on a theoretically-related outcome, rendering them complements" (p. 723). More specifically, Tams et al. analyzed the correlation between a self-report and a physiological (salivary alpha-amylase) measure of technostress to assess their incremental validity in explaining objective performance in a computer-based task. The physiological stress measure predicted variance in performance over and above the stress self-report. Thus, the study constitutes direct evidence that physiology is a complement to self-reports, not an alternative. Tams et al. concluded with a call for more research on the role of physiological measures in IS: "we conducted this study in the context of technostress so that it is currently unclear to what extent the results can generalize to other IS phenomena … Future research could investigate this generalization" (p. 742). Following this call in the context of flow, we confirm Tams et al.'s original finding that physiological measures constitute a critical complement to self-reports. We acknowledge here that our goal was not to triangulate variance components of flow experience measures but to complement the present understanding of flow observation by the ANS perspective, which we deem particularly useful due to the robustness and pervasiveness of ANS measures (in particular for HR and HRV).

An alternative reason for the lack of correlation could be that the form of analysis might be decisive in establishing empirical support for this relationship. Currently, there is no conclusive evidence on which models can explain flow based on physiological measures. We believe that this finding suggests a complex and hence nonlinear underlying relationship between self-reported flow and PNS activation. Initial evidence in this direction is offered by Chin and Kales

(2019) and Bian et al. (2016), as both research teams found an inverted U-shaped relationship between HF-HRV and flow. Based on this finding and considering recent developments in artificial intelligence, we argue that future research should apply more complex mathematical models, such as those created with machine learning (ML), to better uncover the often complex relationships between physiological and self-report measures. Initial promising results have been published recently. Building on a laboratory and a field study, Rissler et al. (2023) trained different ML models to distinguish between low and high flow based on HRV features alone. In each study, the model trained with random forest showed superior results compared to the models trained with the other ML algorithms. An accuracy of 68.5% was achieved in the laboratory and 70.6% in the field. Thus, both studies showed that it is generally possible to differentiate between low and high flow based on HRV features alone. Moreover, Maier et al. (2019) used a deep learning architecture and achieved 67.5% accuracy for the binary classification task (i.e., distinguishing between low and high flow) based on HRV data alone, 58.54% accuracy based on EDA data alone, and 63.03% accuracy based on a combination of EDA and HRV data. Interestingly, in the binary classification task, the model based on HRV data alone performed better than the other two models (note that this was not the case for the three-class classification task).

Practical Implications

Our research has several implications for system designers, managers, and employees trying to mitigate the negative effects of IT-mediated interruptions on flow. A key factor in this regard is content relevance. If the interruption provides information that helps resolve the task, the interruption does not reduce flow. Even more interesting, our results indicate that even a high number of interruptions, if relevant, does not have a negative impact on flow. In other words, as long as the interruption is relevant to the task at hand, flow is not impaired. Therefore, we recommend that individuals try to anticipate whether their message is relevant to the ongoing task of the message receiver or not. However, it can be difficult for the sender of a message to ascertain whether the provided information is relevant to the task that the recipient is currently engaged in. However, we emphasize that several contexts are conceivable in which users have an idea of their colleagues' current work. For example, software engineers increasingly use agile practices such as Scrum. Based on the information that developers exchange in daily standup meetings, it is easier than it was in the past to become aware of the work tasks of colleagues. Moreover, various agencies such as those in the marketing or social media industries also frequently apply agile work practices, including frequent

morning meetings.[12] It follows that employees today are more likely to accurately predict colleagues' tasks than they were in the past.

Even more promising seems to be the development of a nonobtrusive real-time measurement method that can automatically detect flow, as it does not rely on employees' assessment of their colleagues' work. As mentioned, recent research has already contributed to this challenge by relying on ML algorithms to train models that predict or classify flow based on physiological signals (e.g., Maier et al., 2019; Rissler et al., 2023). With these findings, the development of a flow-adaptive system that could improve employees' performance and well-being is becoming attainable (Rissler et al., 2023). On this basis, we see two promising scenarios for such systems: (1) Concerning asynchronous communication, emails could be blocked during task execution by hiding them in the background or not forwarding them as long as the flow-adaptive system "senses" flow. The sender would not know when exactly the message would arrive; however, since asynchronous communication, by definition, creates a time delay before recipients can absorb the information and respond, this should not present a problem. Thus, a flow-adaptive system could help solve the difficulties related to the timing of delivering messages to employees. (2) In synchronous communication, employees rely on instant messaging to communicate with colleagues, customers, and third parties. In this case, either the messaging service itself changes the status of the employee to "away" (e.g., when the person's device is idle) or "on a call" or the employee manually manages their status. In the future, employees could also rely on a flow-adaptive system. Such a research stream would be a direct response to NeuroIS calls for the development of adaptive systems (Adam et al., 2017; Demazure et al., 2021; Riedl & Léger, 2016; vom Brocke et al., 2013; vom Brocke et al., 2020). Such research is also motivated by papers promoting the application of emotion-sensing technology in the workplace to "help employees make better decisions, improve concentration, and adopt healthier and more productive work styles" (Whelan et al., 2018, p. 7).

Limitations and Future Research

Our research has some limitations that offer avenues for future research. First, we focused on two characteristics of IT-mediated interruptions (i.e., frequency and content relevance). Although these characteristics are considered important and have been systematically derived from suggestions in the literature (Addas & Pinsonneault, 2015, 2018a; Galluch et al., 2015), there may be opportunities for future research to also investigate other characteristics of IT-mediated interruptions and their impact on flow, such as presentation mode (e.g., whole

screen versus smaller information boxes, or colors) and interruption channel (e.g., email or instant messenger) (Rissler et al., 2017a). Second, another avenue for future research would be to create an experimental design in which three levels of content relevance are manipulated—for instance, irrelevant, moderately relevant, and highly relevant content. This would help us understand the effect of interruptions that fall between the extremes of irrelevant and highly relevant. Third, future research could also consider other physiological indicators, particularly brain activation, to capture specific physiological patterns of flow (e.g., based on fMRI, EEG, or fNIRS). The use of brain imaging tools and other ANS activity measures (e.g., skin conductance) in future research might also result in the establishment of a relationship between neurophysiological activation and self-reported flow (as we could not find a linear relationship between HF-HRV and self-reported flow). In this context, we also highlight the EFRP method by Léger et al. (2014b), which, among others, can be used "to identify the neural activity associated with the experience of IT discrepant events … [that] can break automatic use patterns by altering users' cognitive, emotional, and behavioral processes, which, in turn, influence performance in a given task" (p. 670). Importantly, the EFRP method was developed based on the experimental task of email pop-up notifications, resembling the type of interruption that we studied in the present article. Fourth, there is the possibility of common method variance (CMV) for the self-report measures since self-reported flow and perceived task performance were measured at the same time and in the same instrument after each within-subjects task period. This is mitigated by our recording of objective task performance; we retained this measure for comparison. If CMV was present, there is the possibility that self-reported flow and perceived task performance might be more highly correlated.

Conclusion

In summary, this study addresses an important research gap in terms of extending flow theory by understanding the role of IT-mediated interruptions as antecedents of flow by investigating the influence of two interruption characteristics (i.e., frequency and content relevance). Our work on flow in the context of interruptions offers the groundwork for scholars to advance the understanding of flow in the context of office work. We hope that future research—including research featuring behavioral, perceptual, and physiological levels of analysis—will continue to build upon this work by exploring the impact of other interruption characteristics on flow, further investigating the physiology of flow, and extending the knowledge of the specific impact of flow on user behavior and metrics like task performance.

---

[12] For details, please see https://www.agilealliance.org/ and https://stateofagile.com/

References

Adam, M. T. P., Gimpel, H., Maedche, A., & Riedl, R. (2017). Design blueprint for stress-sensitive adaptive enterprise systems. *Business & Information Systems Engineering, 59*(4), 277-291. https://doi.org/10.1007/s12599-016-0451-3

Addas, S., & Pinsonneault, A. (2015). The many faces of information technology interruptions: A taxonomy and preliminary investigation of their performance effects. *Information Systems Journal, 25*(3), 231-273. https://doi.org/10.1111/isj.12064

Addas, S., & Pinsonneault, A. (2018a). E-mail interruptions and individual performance: Is there a silver lining? *MIS Quarterly, 42*(2), 381-405. https://doi.org/10.25300/MISQ/2018/13157

Addas, S., & Pinsonneault, A. (2018b). Theorizing the multilevel effects of interruptions and the role of communication technology. *Journal of the Association for Information Systems, 19*(11), 1097-1129. https://doi.org/10.17705/1jais.00521

Ateş, C., Kaymaz, Ö., Kale, H. E., & Tekindal, M. A. (2019). Comparison of test statistics of nonnormal and unbalanced samples for multivariate analysis of variance in terms of type-I error rates. *Computational and Mathematical Methods in Medicine*, Article 2173638. https://doi.org/10.1155/2019/2173638

Bailey, B. P., & Konstan, J. A. (2006). On the need for attention-aware systems: Measuring effects of interruption on task performance, error rate, and affective state. *Computers in Human Behavior, 22*(4), 685-708. https://doi.org/10.1016/j.chb.2005.12.009

Bailey, B. P., Konstan, J. A., & Carlis, J. V. (2001). The effects of interruptions on task performance, annoyance, and anxiety in the user interface. In *INTERACT Proceedings 2001* (pp. 593-601).

Bartels, R. (2018). *Pythonic package for heart rate variability analysis*. https://github.com/rhenanbartels/hrv

Berntson, G. G., Quigley, K. S., & Lozano, D. (2007). Cardiovascular psychophysiology. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of Psychophysiology* (pp. 182-210). Cambridge University Press.

Bian, Y., Yang, C., Gao, F., Li, H., Zhou, S., Li, H., Sun, X., & Meng, X. (2016). A framework for physiological indicators of flow in VR games: Construction and preliminary evaluation. *Personal and Ubiquitous Computing, 20*(5), 821-832. https://doi.org/10.1007/s00779-016-0953-5

Botvinick, M., Nystrom, L. E., Fissell, K., Carter, C. S., & Cohen, J. D. (1999). Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature, 402*(6758), 179-181. https://doi.org/10.1038/46035

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: An update. *Trends in Cognitive Sciences, 8*(12), 539-546. https://doi.org/10.1016/j.tics.2004.10.003

Chen, A., & Karahanna, E. (2018). Life interrupted: The effects of technology-mediated work interruptions on work and nonwork outcomes. *MIS Quarterly, 42*(4), 1023-1042. https://doi.org/10.25300/MISQ/2018/13631

Chikhi, S., Matton, N., & Blanchet, S. (2022). EEG power spectral measures of cognitive workload: A meta-analysis. *Psychophysiology, 59*(6), Article e14009. https://doi.org/10.1111/psyp.14009

Chin, M. S., & Kales, S. N. (2019). Is there an optimal autonomic state for enhanced flow and executive task performance? *Frontiers in Psychology, 10*, Article 1716. https://doi.org/10.3389/fpsyg.2019.01716

Cokelaer, T., & Hasch, J. (2017). "Spectrum": Spectral analysis in Python. *The Journal of Open Source Software, 2*(18), 348-358. https://doi.org/10.21105/joss.00348

Csikszentmihalyi, M. (1975). *Beyond boredom and anxiety*. Jossey-Bass.

Csikszentmihalyi, M. (2000). *Beyond boredom and anxiety: Experiencing flow in work and play*. Jossey-Bass.

Csikszentmihalyi, M. (2014). *Flow and the foundations of positive psychology: The collected works of Mihaly Csikszentmihalyi*. Springer. https://doi.org/10.1007/978-94-017-9088-8

Csikszentmihalyi, M., Khosla, S., & Nakamura, J. (2016). Flow at work. In L. G. Oades, M. F. Steger, A. D. Fave, & J. Passmore (Eds.), *The Wiley Blackwell handbook of the psychology of positivity and strengths-based approaches at work* (pp. 99-109). Chichester: Wiley. https://doi.org/10.1002/9781118977620.ch7

Czerwinski, M., Cutrell, E., & Horvitz, E. (2000). Instant messaging: Effects of relevance and timing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 71-76).

Demazure, T., Karran, A., Léger, P.-M., Labonté-LeMoyne, É., Sénécal, S., Fredette, M., & Babin, G. (2021). Enhancing sustained attention. *Business & Information Systems Engineering, 63*(6), 653-668. https://doi.org/10.1007/s12599-021-00701-3

Demerouti, E. (2006). Job characteristics, flow, and performance: the moderating role of conscientiousness. *Journal of Occupational Health Psychology, 11*(3), 266-280. https://doi.org/10.1037/1076-8998.11.3.266

Dimoka, A., Banker, R. D., Benbasat, I., Davis, F. D., Dennis, A. R., Gefen, D., Gupta, A., Ischebeck, A., Kenning, P., Pavlou, P. A., Muller-Putz, G., Riedl, R., vom Brocke, J., & Weber, B. (2012). On the use of neuropyhsiological tools in IS research: Developing a research agenda for NeuroIS. *MIS Quarterly, 36*(3), 679-702. https://doi.org/10.2307/41703475

Engeser, S., & Rheinberg, F. (2008). Flow, performance and moderators of challenge-skill balance. *Motivation and Emotion, 32*(3), 158-172. https://doi.org/10.1007/s11031-008-9102-4

Fortin-Côté, A., Beaudin-Gagnon, N., Campeau-Lecours, A., Tremblay, S., & Jackson, P. L. (2019). Affective computing out-of-the-lab: The cost of low cost. In *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics* (pp. 4137-4142). https://doi.org/10.1109/SMC.2019.8914646

Fullagar, C. J., & Kelloway, E. K. (2009). Flow at work: An experience sampling approach. *Journal of Occupational and Organizational Psychology, 82*(3), 595-615. https://doi.org/10.1348/096317908X357903

Galluch, P. S., Grover, V., & Thatcher, J. B. (2015). Interrupting the workplace: Examining stressors in an information technology context. *Journal of the Association for Information Systems, 16*(1), 1-47. https://doi.org/10.17705/1jais.00387

Gluck, J., Bunt, A., & McGrenere, J. (2007). Matching attentional draw with utility in interruption. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 41-50). https://doi.org/10.1145/1240624.1240631

Gunnar, M., & Quevedo, K. (2007). The neurobiology of stress and development. *Annual Review of Psychology, 58*, 145-173. https://doi.org/10.1146/annurev.psych.58.110405.085605

Gupta, A., Li, H., & Sharda, R. (2013). Should I send this message? Understanding the impact of interruptions, social hierarchy and perceived task complexity on user performance and perceived

workload. *Decision Support Systems, 55*(1), 135-145. https://doi.org/10.1016/j.dss.2012.12.035

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). *Multivariate data analysis*. Pearson Education.

Harmat, L., Manzano, Ö. de, Theorell, T., Högman, L., Fischer, H., & Ullén, F. (2015). Physiological correlates of the flow experience during computer game playing. *International Journal of Psychophysiology, 97*(1), 1-7. https://doi.org/10.1016/j.ijpsycho.2015.05.001

Harmat, L., Ullén, F., Manzano, Ö. de, Ölsson, E., Elofsson, U., Schéele, B. von, & Theorell, T. (2011). Heart rate variability during piano playing: A case study of three professional solo pianists playing a self-selected and a difficult prima vista piece. *Music and Medicine, 3*(2), 102-107. https://doi.org/10.1177/1943862110387158

Harris, D. J., Vine, S. J., & Wilson, M. R. (2017a). Is flow really effortless? The complex role of effortful attention. *Sport, Exercise, and Performance Psychology, 6*(1), 103-114. https://doi.org/10.1037/spy0000083

Harris, D. J., Vine, S. J., & Wilson, M. R. (2017b). Neurocognitive mechanisms of the flow state. In M. R. Wilson, V. Walsh, & B. Parkin (Eds.), *Sport and the brain: The science of preparing, enduring and winning: Part B* (pp. 221-243). Elsevier. https://doi.org/10.1016/bs.pbr.2017.06.012

Hayes, A. F. (2021). *Introduction to mediation, moderation, and conditional process analysis*: *A regression-based approach*. Guilford.

Jackson, S., & Marsh, H. (1996). Development and validation of a scale to measure optimal experience: The flow state scale. *Journal of Sport & Exercise Psychology, 18*(1), 17-35. https://doi.org/10.1123/jsep.18.1.17

Jett, Q. R., & George, J. M. (2003). Work interrupted: A closer look at the role of interruptions in organizational life. *Academy of Management Review, 28*(3), 494-507. https://doi.org/10.2307/30040736

Knierim, M. T., Rissler, R., Dorner, V., Maedche, A., & Weinhardt, C. (2017). The psychophysiology of flow: A systematic review of peripheral nervous system features. In *Information Systems and Neuroscience*: *Gmunden Retreat on NeuroIS 2017 Proceedings* (pp. 109-120). https://doi.org/10.1007/978-3-319-67431-5_13

Kolb, B., & Whishaw, I. Q. (2009). *Fundamentals of human neuropsychology*. Palgrave Macmillan.

Kuo, T.-H., & Ho, L.-A. (2010). Individual difference and job performance: the relationships among personal factors, job characteristics, flow experience, and service quality. *Social Behavior and Personality, 38*(4), 531-552. https://doi.org/10.2224/sbp.2010.38.4.531

Léger, P.-M., Davis, F. D., Cronan, T. P., & Perret, J. (2014a). Neurophysiological correlates of cognitive absorption in an enactive training context. *Computers in Human Behavior, 34*, 273-283. https://doi.org/10.1016/j.chb.2014.02.011

Léger, P.-M., Sénecal, S., Courtemanche, F., Guinea, A., Titah, R., Fredette, M., & Labonte-LeMoyne, É. (2014b). Precision is in the eye of the beholder: Application of eye fixation-related potentials to information systems research. *Journal of the Association for Information Systems, 15*(10), 651-678. https://doi.org/10.17705/1jais.00376

Loewe, N., & Nadj, M. (2020). Physio-adaptive systems: A state-of-the-art review and future research directions. In *Proceedings of the European Conference on Information Systems*.

Mahnke, R., Benlian, A., & Hess, T. (2014). Flow experience in information systems research: Revisiting its conceptualization, conditions, and effects. In *Proceedings of the International Conference on Information Systems*. https://doi.org/10.13140/2.1.4852.0486

Mahnke, R., Benlian, A., & Hess, T. (2015). A grounded theory of online shopping flow. *International Journal of Electronic Commerce 19*(3), 54-89. https://doi.org/10.1080/10864415.2015.1000222

Maier, M., Elsner, D., Marouane, C., Zehnle, M., & Fuchs, C. (2019). DeepFlow: Detecting optimal user experience from physiological data using deep neural networks. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence* (pp. 1415-1421). https://doi.org/10.24963/ijcai.2019/196

Mangalaraj, G., Nerur, S., Mahapatra, R., & Price, K. H. (2014). Distributed cognition in software design: An experimental investigation of the role of design patterns and collaboration. *MIS Quarterly, 38*(1), 249-274. https://doi.org/10.25300/MISQ/2014/38.1.12

Mark, G., Iqbal, S. T., Czerwinski, M., Johns, P., Sano, A., & Lutchyn, Y. (2016). Email duration, batching and self-interruption. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1717-1728). https://doi.org/10.1145/2858036.2858262

Mark, G., Voida, S., & Cardello, A. (2012). A pace not dictated by electrons: An empirical study of work without email. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 555-564). https://doi.org/10.1145/2207676.2207754

Mathewson, K. J., Jetha, M. K., Drmic, I. E., Bryson, S. E., Goldberg, J. O., Hall, G. B., Santesso, D. L., Segalowitz, S. J., & Schmidt, L. A. (2010). Autonomic predictors of Stroop performance in young and middle-aged adults. *International Journal of Psychophysiology 76*(3), 123-129. https://doi.org/10.1016/j.ijpsycho.2010.02.007

Maynard, D. C., & Hakel, M. D. (1997). Effects of objective and subjective task complexity on performance. *Human Performance, 10*(4), 303-330. https://doi.org/10.1207/s15327043hup1004_1

Moller, A. C., Meier, B. P., & Wall, R. D. (2010). Developing an experimental induction of flow: Effortless action in the lab. In B. Bruya (Ed.), *Effortless attention: A new perspective in the cognitive science of attention and action* (pp. 191-204). MIT Press. https://doi.org/10.7551/mitpress/9780262013840.003.0010

Monk, C. A., Trafton, G., & Boehm-Davis, D. (2008). The effect of interruption duration and demand on resuming suspended goals. *Journal of Experimental Psychology: Applied, 14*(4), 299-313. https://psycnet.apa.org/doi/10.1037/a0014402

Morgan, P. L., Patrick, J., Waldron, S. M., King, S. L., & Patrick, T. (2009). Improving memory after interruption: Exploiting soft constraints and manipulating information access cost. *Journal of Experimental Psychology: Applied, 15*(4), 291-306. https://doi.org/10.1037/a0018008

Murray, N. P., & Russoniello, C. (2012). Acute physical activity on cognitive function: A heart rate variability examination. *Applied Psychophysiology and Biofeedback, 37*(4), 219-227. https://doi.org/10.1007/s10484-012-9196-z

Nakamura, J., & Csikszentmihalyi, M. (2009). Flow theory and research. In S. J. Lopez and C. Snyder (Eds.), *The Oxford handbook of positive psychology* (pp. 195-206). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195187243.013.0018

Novak, T. P., & Hoffman, D. L. (2009). The fit of thinking style and situation: New measures of situation-specific experiential and rational cognition. *Journal of Consumer Research, 36*(1), 56-72. https://doi.org/10.1086/596026

Ortiz de Guinea, A., Titah, R., & Léger, P.-M. (2013). Measure for measure: A two study multi-trait multi-method investigation of construct validity in IS research. *Computers in Human Behavior, 29*(3), 833-844. https://doi.org/10.1016/j.chb.2012.12.009

Ortiz de Guinea, A., Titah, R., & Léger, P.-M. (2014). Explicit and implicit antecedents of users' behavioral beliefs in information systems: A neuropsychological investigation. *Journal of Management Information Systems, 30*(4), 179-210. https://doi.org/10.2753/MIS0742-1222300407

Ortiz de Guinea, A., & Webster, J. (2013). An investigation of information systems use patterns: Technological events as triggers, the effect of time, and consequences for performance. *MIS Quarterly, 37*(4), 1165-1188. https://doi.org/10.25300/MISQ/2013/37.4.08

Parent-Thirion, A., Vermeylen, G., van Houten, G., Wilkens, M., & Wilczynska, A. (2015). *First findings: Sixth European working conditions survey.* Publications office of the European Union. https://doi.org/10.2806/59106

Peifer, C., Schächinger, H., Engeser, S., & Antoni, C. H. (2015). Cortisol effects on flow-experience. *Psychopharmacology, 232*(6), 1165-1173. https://doi.org/10.1007/s00213-014-3753-5

Peifer, C., Schulz, A., Schächinger, H., Baumann, N., & Antoni, C. H. (2014). The relation of flow-experience and physiological arousal under stress: Can u shape it? *Journal of Experimental Social Psychology, 53*, 62-69. https://doi.org/10.1016/j.jesp.2014.01.009

Peifer, C., & Tan, J. (2021). The psychophysiology of flow experience. In C. Peifer & S. Engeser (Eds.), *Advances in flow research* (pp. 191-230). Springer. https://doi.org/10.1007/978-3-030-53468-4_8

Peifer, C., & Wolters, G. (2021). Flow in the context of work. In C. Peifer & S. Engeser (Eds.), *Advances in flow research* (pp. 287-321). Springer. https://doi.org/10.1007/978-3-030-53468-4_11

Pituch, K. A., & Stevens, J. (2016). *Applied multivariate statistics for the social sciences: Analyses with SAS and IBM's SPSS.* Routledge.

PLUX Wireless Biosignals S. A. (2018). *Biosignalsplux.* http://biosignalsplux.com/en/

Qiu, L., & Benbasat, I. (2005). An investigation into the effects of text-to-speech voice and 3D avatars on the perception of presence and flow of live help in electronic commerce. *ACM Transactions on Computer-Human Interaction, 12*(4), 329-355. https://doi.org/10.1145/1121112.1121113

Riedl, R., & Léger, P.-M. (2016). *Fundamentals of NeuroIS.* Springer. https://doi.org/10.1007/978-3-662-45091-8

Riedl, R. (2013). On the biology of technostress: Literature review and research agenda. *The Data Base for Advances in Information Systems, 44*(1), 18-55. https://doi.org/10.1145/2436239.2436242

Riedl, R., Davis, F. D., & Hevner, A. R. (2014). Towards a NeuroIS research methodology: Intensifying the discussion on methods, tools, and measurement. *Journal of the Association for Information Systems, 15*(10), i-xxxv. https://doi.org/10.17705/1jais.00377

Riedl, R., Fischer, T., Léger, P.-M., & Davis, F. D. (2020). A decade of NeuroIS research: Progress, challenges, and future directions. *The Data Base for Advances in Information Systems, 51*(3), 13-54. https://doi.org/10.1145/3410977.3410980

Rissler, R., Nadj, M., Adam, M. T., & Maedche, A. (2017a). Towards an integrative theoretical framework of IT-mediated interruptions. In *Proceedings of the European Conference on Information Systems* (pp. 1950-1967).

Rissler, R., Nadj, M., & Adam, M. T. P. (2017b). Flow in information systems research: Review, integrative theoretical framework, and future directions. In *Wirtschaftsinformatik 2017 Proceedings* (pp. 1051-1065).

Rissler, R., Nadj, M., Li, M. X., Loewe, N., Knierim, M. T., & Maedche, A. (2023). To be or not to be in flow at work: Physiological classification of flow using machine learning. *IEEE Transactions on Affective Computing*, *14*(1), 463-474. https://doi.org/10.1109/TAFFC.2020.3045269

Riva, E. F. M., Riva, G., Talò, C., Boffi, M., Rainisio, N., Pola, L., Diana, B., Villani, D., Argenton, L., & Inghilleri, P. (2017). Measuring dispositional flow: Validity and reliability of the Dispositional Flow State Scale 2. *PLOS ONE, 12*(9), 1-15. https://doi.org/10.1371/journal.pone.0182201

Schüler, J., & Brunner, S. (2009). The rewarding effect of flow experience on performance in a marathon race. *Psychology of Sport and Exercise, 10*(1), 168-174. https://doi.org/10.1016/j.psychsport.2008.07.001

Schultz, T., Amma, C., Wand, M., Heger, D., & Putze, F. (2013). Biosignale-basierte Mensch-Maschine Schnittstellen. *at – Automatisierungstechnik, 61*(11), 760-769. https://doi.org/10.1524/auto.2013.1061

Shaffer, F., & Ginsberg, J. P. (2017). An overview of heart rate variability metrics and norms. *Frontiers in Public Health, 5*, 1-17. https://doi.org/10.3389/fpubh.2017.00258

Shalom, J. G., Israeli, H., Markovitzky, O., & Lipsitz, J. D. (2015). Social anxiety and physiological arousal during computer mediated vs. face to face communication. *Computers in Human Behavior, 44*, 202-208. https://doi.org/10.1016/j.chb.2014.11.056

Stangl, F. J., & Riedl, R. (2022). Measurement of heart rate and heart rate variability with wearable devices: A systematic review. In *Wirtschaftsinformatik 2022 Proceedings*.

Tams, S., Hill, K., Guinea, A. O. d., Thatcher, J., & Grover, V. (2014). NeuroIS: Alternative or complement to existing methods? Illustrating the holistic effects of neuroscience and self-reported data in the context of technostress research. *Journal of the Association for Information Systems, 15*(10), 723-753. https://doi.org/10.17705/1jais.00374

Tams, S., Thatcher, J., Grover, V., & Pak, R. (2015). Selective attention as a protagonist in contemporary workplace stress: Implications for the interruption age. *Anxiety, Stress, & Coping 28*(6), 663-686. https://doi.org/10.1080/10615806.2015.1011141

Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. (1996). Heart rate variability: Standards of measurement, physiological interpretation and clinical use. *Circulation 93*(5), 1043-1065. https://doi.org/10.1161/01.CIR.93.5.1043

Tozman, T., Magdas, E. S., MacDougall, H. G., & Vollmeyer, R. (2015). Understanding the psychophysiology of flow: A driving simulator experiment to investigate the relationship between flow and heart rate variability. *Computers in Human Behavior 52*, 408-418. https://doi.org/10.1016/j.chb.2015.06.023.

Tozman, T., & Peifer, C. (2016). Experimental paradigms to investigate flow-experience and its psychophysiology: Inspired from stress theory and research. In L. Harmat, F. Ørsted Andersen, F. Ullén, J. Wright, & G. Sadlo (Eds.), *Flow experience: Empirical research and applications* (pp. 329-350). Springer. https://doi.org/10.1007/978-3-319-28634-1_20

Trafton, J. G., Altmann, E. M., Brock, D. P., & Mintz, F. E. (2003). Preparing to resume an interrupted task: Effects of prospective goal encoding and retrospective rehearsal. *International Journal of Human Computer Studies, 58*(5), 583-603. https://doi.org/10.1016/S1071-5819(03)00023-5

Tukey, J. W. (1977). *Exploratory data analysis*. Addison-Wesley.

Ulrich, M., Keller, J., & Grön, G. (2016). Neural signatures of experimentally induced flow experiences identified in a typical fMRI block design with BOLD imaging. *Social Cognitive and Affective Neuroscience, 11*(3), 496-507. https://doi.org/10.1093/scan/nsv133

Ulrich, M., Keller, J., Hoenig, K., Waller, C., & Grön, G. (2014). Neural correlates of experimentally induced flow experiences. *NeuroImage, 86*, 194-202. https://doi.org/10.1016/j.neuroimage.2013.08.019

Valentini, M., & Parati, G. (2009). Variables influencing heart rate. *Progress in Cardiovascular Diseases 52*(1), 11-19. https://doi.org/10.1016/j.pcad.2009.05.004

Verma, J. P. (2016). *Repeated measures design for empirical researchers*. Wiley.

Vom Brocke, J., Hevner, A., Léger, P.-M., Walla, P., & Riedl, R. (2020). Advancing a NeuroIS research agenda with four areas of societal contributions. *European Journal of Information Systems 29*(1), 9-24. https://doi.org/10.1080/0960085X.2019.1708218

Vom Brocke, J., Riedl, R., & Léger, P.-M. (2013). Application strategies for neuroscience in information systems design science research. *Journal of Computer Information Systems 53*(3), 1-13. https://doi.org/10.1080/08874417.2013.11645627

Wagner, G. S., & Strauss, D. G. (2014). *Marriott's practical electrocardiography*. Wolters Kluwer Health / Lippincott Williams & Wilkins.

Waxenbaum, J. A., Reddy, V., & Varacallo, M. (2022). *Anatomy, autonomic nervous system*. StatPearls Publishing.

Whelan, E., McDuff, D., Gleasure, R., & Vom Brocke, J. (2018). How emotion-sensing technology can reshape the workplace. *MIT Sloan Management Review, 59*(3), 7-10.

Wolf, S., Brölz, E., Keune, P. M., Wesa, B., Hautzinger, M., Birbaumer, N., & Strehl, U. (2015). Motor skill failure or flow-experience? Functional brain asymmetry and brain connectivity in elite and amateur table tennis players. *Biological Psychology 105*, 95-105. https://doi.org/10.1016/j.biopsycho.2015.01.007

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology 18*(5), 459-482. https://doi.org/10.1002/cne.920180503

Yi, C., Jiang, Z., & Benbasat, I. (2015). Enticing and engaging consumers via online product presentations: The effects of restricted interaction design. *Journal of Management Information Systems 31*(4), 213-242. https://doi.org/10.1080/07421222.2014.1001270

Zubair, A., & Kamal, A. (2015). Work related flow, psychological capital, and creativity among employees of software houses. *Psychological Studies 60*(3), 321-331. https://doi.org/10.1007/s12646-015-0330-x

About the Authors

Mario Nadj is an assistant professor at the TU Dortmund University in Germany and heads the Business & Information Systems Engineering group. He is dedicated to investigating, designing, and evaluating digital technologies to actively shape the future of work. For this purpose, he uses design-oriented and empirical research methods. His work is positioned at the intersection of information systems and human-computer interaction. Specifically, he investigates novel information systems in the fields of business intelligence & analytics, physiological computing, and machine learning. His research has been published in leading international journals and conference proceedings, such as *Journal of Cognitive Engineering & Decision Making*, *ACM Conference on Human Factors in Computing Systems*, *IEEE Transactions on Affective Computing*, and *Decision Support Systems*. ORCID ID: https://orcid.org/0000-0002-6901-9450

Raphael Rissler received his Ph.D. from the Karlsruhe Institute of Technology (KIT) in 2019 after he successfully conducted a doctoral program in close cooperation between the KIT and SAP SE. His research interests include artificial intelligence/machine learning, flow, physiological computing, and interruptions. He publishes in leading outlets, like *IEEE Transactions on Affective Computing*, and the proceedings of the *European Conference on Information Systems* and the *ACM Conference on Human Factors in Computing Systems*. Since 2019 he has worked as a senior data scientist at SAP SE, where he leads innovative projects on artificial intelligence/machine learning to enable the intelligent enterprise. He also published a book called *Data Science Process Model (DASC-PM),* which developed a process model for data science projects in companies to make real-world artificial intelligence/machine learning projects a success. ORCID ID: https://orcid.org/0000-0002-6206-0733

Marc T. P. Adam is an associate professor in computing and information technology at the University of Newcastle, Australia. He received an undergraduate degree in computer science from the University of Applied Sciences Würzburg, Germany, and a Ph.D. in business information systems from the Karlsruhe Institute of Technology (KIT), Germany. His research investigates the interplay of cognition and affect in human-computer interaction. He is a founding member of the NeuroIS Society. ORCID ID: https://orcid.org/0000-0002-6036-4282

Michael T. Knierim, Ph.D. is a post-doctoral researcher at the Institute of Information Systems and Marketing (IISM) at the Karlsruhe Institute of Technology (KIT) in Germany. His research interests lie in the use of neurophysiological measurement methods and feedback systems to facilitate positive psychological experiences such as flow or relaxation states. He has published in international outlets in the domain of information systems and human-computer interaction—for example, in the proceedings of the *International Conference on Information Systems* and the *ACM Conference on Human Factors in Computing Systems*—and has received a publication award for a biofeedback prototype published in the proceedings of the *Conference on Design Science Research in Information Systems and Technology*. ORCID ID: https://orcid.org/0000-0001-7148-5138

Maximilian X. Li is a research scientist and Ph.D. candidate in the Intuitive Robots Lab at the Karlsruhe Institute of Technology (KIT). His research interests lie at the intersection of robotics and machine learning, with a focus on human-robot interaction. Maximilian investigates how explainable artificial intelligence techniques can be leveraged in interactive learning settings. He obtained his M.Sc. in computer science at KIT in 2022. From 2018-2020 he was a junior researcher at the Human-Centered Systems Lab at KIT. ORCID ID: https://orcid.org/0000-0002-5974-3322
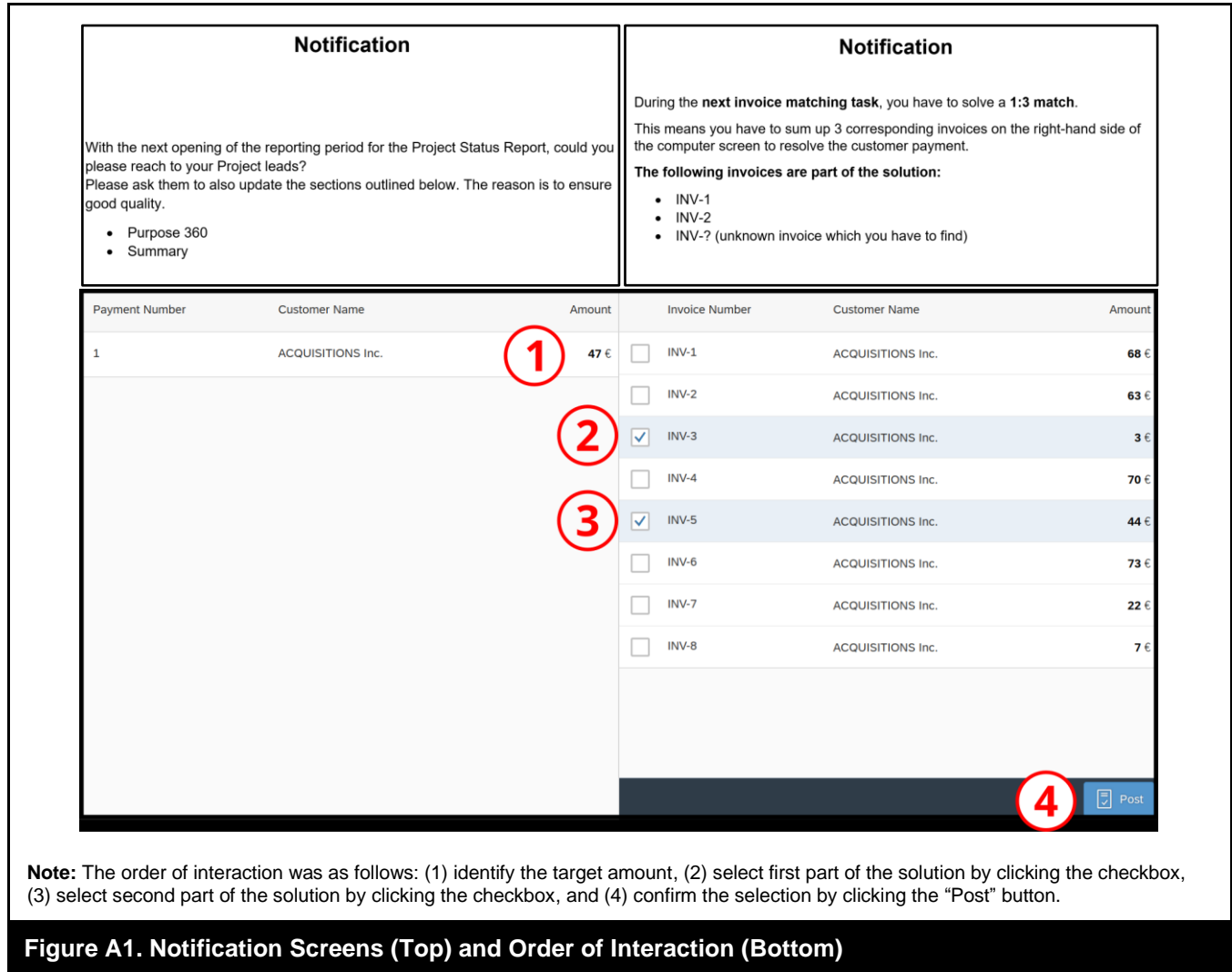
Alexander Maedche has been a professor of information systems at the Karlsruhe Institute of Technology (KIT), Germany since 2015. Before that, he was a professor at the University of Mannheim (2009-2015) and worked between 2003 and 2009 outside academia at Bosch and SAP in different managerial roles. He currently heads the Human-Centered Systems Lab (h-lab), which researches the design of human-centered systems for better work and life at the intersection of information systems and human-computer interaction. His work has been published in leading journals such as *MIS Quarterly, Journal of the Association for Information Systems, Computers & Human Behavior, International Journal of Human-Computer Studies,* and *IEEE Transactions on Software Engineering*. He serves as an editorial board member for renowned journals such as *MIS Quarterly, Journal of the Association for Information Systems,* and *Business & Information Systems Engineering*. He emphasizes knowledge transfer by actively engaging in company cooperations and acting as a founding member of the nonprofit associations Usability & UX in Germany (UIG) e.V. and Die Wirtschaftsinformatik e.V. ORCID ID: https://orcid.org/0000-0001-6546-4816

René Riedl is a professor of digital business and innovation at the University of Applied Sciences Upper Austria and served as vice dean for R&D in the School of Business and Management from 2018-2021. He is also an associate professor of business informatics at Johannes Kepler University Linz, Austria. He serves as the scientific director of the NeuroIS Society (Vienna, Austria) and as a conference co-chair of the NeuroIS Retreat (www.NeuroIS.org). He has published in many scientific journals such as *Addictive Behaviors Reports, Behavior Research Methods, Business & Information Systems Engineering, European Journal of Information Systems, Frontiers in Neuroscience, Frontiers in Psychology, Journal of Information Technology, Journal of Management Information Systems, Journal of Systems & Software, Journal of the Association for Information Systems, MIS Quarterly, PLOS ONE,* and *Scientific Reports.* He also serves as a senior editor or an associate editor for several scientific journals. Finally, he advises companies on questions about digital transformation, organizational measures against technostress, and several other questions about the use of digital technologies in organizations. ORCID ID: https://orcid.org/0000-0003-2468-5860

Appendix A

Experimental Screens

Note: The order of interaction was as follows: (1) identify the target amount, (2) select first part of the solution by clicking the checkbox, (3) select second part of the solution by clicking the checkbox, and (4) confirm the selection by clicking the "Post" button.

Figure A1. Notification Screens (Top) and Order of Interaction (Bottom)

Appendix B

Experimental Design and Measures

| (1) Information period | (2) Quiz period | (3) Practice period | (4) Calibration period | (5) Physio baseline | (6) Task periods |
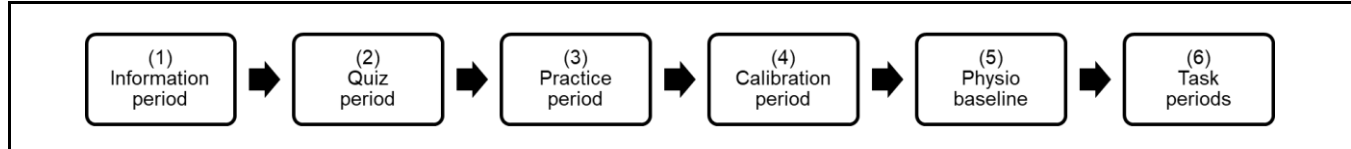
Figure B1. Steps in the Experimental Design

Table B1. Measurement Items

Construct	Item	Item description	Source
Merging of action and awareness (MAA)	MAA1	During the invoice matching task, things just seemed to be happening automatically	Adapted from (Jackson & Marsh, 1996)
	MAA2	During the invoice matching task, I performed automatically	
	MAA3	During the invoice matching task, I did things spontaneously and automatically without having to think	
Concentration (CON)	CON1	During the invoice matching task, my attention was focused entirely on what I was doing	
	CON2	During the invoice matching task, it was no effort to keep my mind on what was happening	
	CON3	During the invoice matching task, I had total concentration	
	CON4	During the invoice matching task, I was completely focused on the task at hand	
Sense of control (CTR)	CTR1	During the invoice matching task, I felt in total control of what I was doing	
	CTR2	During the invoice matching task, I felt like I could control what I was doing	
	CTR3	During the invoice matching task, I had a feeling of total control	
	CTR4	I felt in total control of my actions during the invoice matching task	
Loss of self-consciousness (LSC)	LSC1	During the invoice matching task, I was not concerned with what others may have been thinking of me	
	LSC2	I was not worried about my performance during the invoice matching task	
	LSC3	During the invoice matching task, I was not concerned with how I was presenting myself	
	LSC4	During the invoice matching task, I was not worried about what others may have been thinking of me	
Perceived task performance (PPE)	PPE1	How would you rate your performance on the invoice matching task? *[(1) extremely poor to (7) extremely good]*	Adapted from (Novak & Hoffman, 2009)
	PPE2	If you were to grade your performance on the invoice matching task, what grade would you give yourself? *[(1) extremely poor to (7) extremely good]*	
	PPE3	Compared to other people, how well do you think you did on the invoice matching task? *[(1) well below average to (7) well above average]*	

Note: Unless stated otherwise, items were anchored on a 7-point Likert scale from (1) *strongly disagree* to (7) *strongly agree.*

Table B2. Manipulation Checks			
Construct	**Item**	**Item description**	**Source**
Perceived interruption frequency (PIF)	PIF1	During the invoice matching task, I received too many interruptions	Adapted from (Galluch et al., 2015)
	PIF2	During the invoice matching task, I experienced many distractions	
	PIF3	During the invoice matching task, the interruptions came frequently	
Perceived content relevance (PCR)	PCR1	During the invoice matching task, the interruptions helped me accomplish my task	Adapted from (Galluch et al., 2015)
	PCR2	During the invoice matching task, the interruptions helped me think about my task	
	PCR3	During the invoice matching task, the interruptions were relevant to my task	

Note: Items were anchored on a 7-point Likert scale from (1) *strongly disagree* to (7) *strongly agree.*

Table B3. Control Variables			
Construct	**Item**	**Item description**	**Source**
Age	AGE1	How old are you? *[measured with an integer input box]*	Adapted from (Galluch et al., 2015)
Gender	GEN1	What is your gender? *[checkbox: (1) male (2) female (3) other]*	Adapted from (Gupta et al., 2013)
Education	CST1	My Class status is currently (e.g., Bachelor-Student, Master-Student, ...) *[Measured with an input box.]*	Adapted from (Galluch et al., 2015)
Prior task experience	TEX1	How much experience have you had in the past with tasks similar to those (invoice matching tasks) that you have just worked on?*[checkbox: (1) none (2) a little (3) some, (4) a lot]*	(Maynard & Hakel, 1997)
Order of the within-subject factor	ORD1	Measured by the experimental apparatus	-

Table B4. Objective Task Performance Examples	$level_{end}$	$level_{mean}$	$direction$	**OPE**
Example				
Starts weak; improves towards the end	2	1	1	4
Average participant; performs worse to the end	2	3	-1	4
Average participant; improves to the end	3	2	1	6
Strong participant, but cannot level up in a high difficulty, but also does not lose a level	5	5	0	10
Starts below final level, but improves greatly	5	4	1	10
Starts very strong, but loses a level in the end	4	5	-1	8

The objective task performance measure was computed with the following formula:

$$OPE = level_{end} + level_{mean} + direction$$

$level_{end}$ is the difficulty level the participant has reached when the treatment round ends. As higher levels are more difficult, reaching such a level is rewarded. $level_{mean}$ is the individual's average level for the treatment. This metric acknowledges the difficulty to further level up once a higher difficulty is reached. $direction$ is an integer value from $[-1; 1]$, or the output of the *signum* function for the difference between the end and start level. If the last level is higher than the start level, the subject is rewarded with a direction score of +1, if it is lower, they are penalized with a direction score of -1. If the last level is the same than the start level, the subject gets a direction score of 0.

Appendix C

Assumptions Testing for MANOVA & ANOVA

Table C1a. Levene's Test of Homogeneity of Variance

Frequency condition – Construct	*F*	df1	df2	Sig.
High frequency – MAA	0.021	1	164	0.886
Low frequency – MAA	0.266	1	164	0.606
High frequency – CON	2.966	1	164	0.087
Low frequency – CON	0.185	1	164	0.668
High frequency – CTR	3.981	1	164	0.048
Low frequency – CTR	0.003	1	164	0.958
High frequency – LSC	0.133	1	164	0.716
Low frequency – LSC	6.671	1	164	0.011
High frequency – HF-HRV	0.208	1	127	0.649
Low frequency – HF-HRV	0.036	1	127	0.850

Note: The homogeneity of variances assumption is violated in two out of ten conditions: (1) High frequency – CTR; $p = 0.048$, and Low frequency – LSC; $p = 0.011$. Against this backdrop, we followed literature guidelines and also used Roy's largest root as a test statistic (Ateş et al., 2019). The findings remained exactly the same.

Table C1b. Box's Test of Homogeneity of Covariance Matrices

Construct	Box's *M*	*F*	df1	df2	Sig.
Flow	55.872	1.472	36	89519.221	0.034
HF-HRV	0.356	0.116	3	3121746.069	0.950

Table C2. Skewness and Kurtosis

Construct	Item	Skewness (Kurtosis)			
		High frequency		Low frequency	
		Relevant	Irrelevant	Relevant	Irrelevant
Merging of action and awareness	MAA1	-0.062 (-0.687)	0.344 (-0.687)	0.114 (-0.901)	0.014 (-0.670)
	MAA2				
	MAA3				
Concentration	CON1	-0.542 (-0.591)	0.000 (-1.192)	-0.739 (0.180)	-0.504 (-0.074)
	CON2				
	CON3				
	CON4				
Sense of control	CTR1	-0.716 (0.274)	0.022 (-1.071)	-0.425 (-0.366)	-0.325 (-0.337)
	CTR2				
	CTR3				
	CTR4				
Loss of self-consciousness	LSC1	-0.434 (-0.781)	-0.330 (-0.848)	-0.464 (-0.307)	-0.294 (-0.848)
	LSC2				
	LSC3				
	LSC4				
Perceived task performance	PPE1	-0.009 (-0.284)	0.084 (-0.636)	-0.290 (-0.509)	0.006 (-0.302)
	PPE2				
	PPE3				
HF-HRV	HFH1	-0.12 (-0.43)	-0.62 (0.03)	-0.20 (-0.13)	-0.27 (0.25)

Table C3. Checks for Potential Covariates on Flow			
Control variable	**Sig.**	η^2_p	**Power**
Age	*p* = .061	.054	.659
Gender	*p* = .569	.018	.233
Education	*p* = .335	.028	.355
Prior task experience	*p* = .570	.018	.233
Order of the within-subject factor	*p* = .599	.017	.221