

Sensing and Supporting Software Developers' Focus

Manuela Züger
University of Zürich
zueger@ifi.uzh.ch

Thomas Fritz
University of Zürich
fritz@ifi.uzh.ch

ABSTRACT

Software developers regularly have to focus in order to successfully perform their work. At the same time, developers experience many disruptions to their focus, especially in today's highly demanding, collaborative and open office work environments. When these disruptions happen during tasks that require a lot of focus, such as comprehending a difficult piece of source code, they can be very costly, causing a decrease in performance and quality. By sensing how focused a developer is, we might be able to reduce the cost of such disruptions.

In our previous work, we investigated the use of biometric and computer interaction sensors to sense interruptibility—the availability for interruptions—and developed the FlowLight approach—a traffic light like LED indicator of a person's interruptibility—to reduce the cost of external in-person interruptions, a particularly expensive kind of disruption. Our results demonstrate the potential of accurately sensing interruptibility in the field and of reducing external interruption cost to increase focus and productivity of knowledge workers.

CCS CONCEPTS

• **Human-centered computing** → *Human computer interaction (HCI)*;

KEYWORDS

focus, interruptions, sensors, software developers

ACM Reference Format:

Manuela Züger and Thomas Fritz. 2018. Sensing and Supporting Software Developers' Focus. In *ICPC '18: 26th IEEE/ACM International Conference on Program Comprehension, May 27–28, 2018, Gothenburg, Sweden*. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3196321.3196323>

1 INTRODUCTION

Software developers have to perform a broad variety of activities for their work [19] that demand different levels of focus [17]. Especially for tasks such as program comprehension, that require much of a developer's focus and attention to be completed successfully, disruptions that interfere with the main task can have a significant impact on the developer's performance and quality of work. One common source for disruptions from the main task are external interruptions, such as a colleague asking a question in person or an incoming phone call [8]. While many of these interruptions can be

important and beneficial, they can also incur a high cost. Studies have shown that interruptions that happen during highly focused periods of work can significantly increase the error rate, the time to complete the task, and the person's frustration level [3]. Knowing how focused a developer is at any given point in time and fostering focused work, for example, by postponing expensive interruptions, thus has a huge potential to improve the developer's productivity, quality of work, as well as the developer's well-being.

For this paper, we concentrate on research to foster focused developer work by sensing interruptibility—a person's availability for interruptions—and by reducing external in-person interruptions. We present and discuss three studies that we conducted to investigate the accuracy of automatically measuring a person's interruptibility using a variety of biometric and computer interaction sensors in the lab and field [26, 27]. We further present the FlowLight, a traffic light like approach that uses computer interaction sensors to indicate a knowledge worker's interruptibility to co-workers [25]. Our studies on the interruptibility sensing and the FlowLight demonstrate the potential of biometric and computer interaction sensors to determine a knowledge worker's interruptibility and the successful use of such measures to significantly reduce external interruptions and increase productivity in the work place. We believe that the insights can be used to investigate more general measures of focus and foster focus in a developers' work more generally, for instance, by reducing expensive self-interruptions or making sure that developers work on the most cognitively demanding tasks when they are most focused during the day.

2 BACKGROUND AND RELATED WORK

Knowledge workers experience many interruptions a day, each often lasting up to 20 minutes [11]. While many interruptions are necessary in a collaborative work environment, such as software development, they can also disrupt the worker's focus and have multiple negative effects, such as long resumption lags or an increase in errors and frustration [3]. Studies have shown that the moment of interruption thereby plays an important role: the higher the person's cognitive load and focus, the higher the cost of the interruption [1]. Knowing how interruptible a person is at any point in time can thus help to optimize the timing of interruptions.

Sensing Interruptibility. Techniques to measure interruptibility can broadly be categorized by the sensor types used: biometric, computer interaction, or context sensors. *Biometric* (aka. psychophysiological) sensors are used to measure a body's activities and responses that are linked to cognitive processes. Various studies have shown that biometric data such as heart rate (HR), heart rate variability (HRV), blood volume pulse (BVP), electro-dermal activity (EDA), pupil dilation, skin temperature or electroencephalography (EEG) data can be used to assess mental effort and cognitive load [14], task difficulty [10], emotions [13], or stress [23]. A few

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
ICPC '18, May 27–28, 2018, Gothenburg, Sweden
© 2018 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-5714-2/18/05.
<https://doi.org/10.1145/3196321.3196323>

researchers have also investigated whether such sensors can be used to measure interruptibility, e.g. using EEG data [18], EDA data [12], or eye tracking data [1].

Computer interaction sensors measure a user’s interaction with the computer, in particular, the mouse, keyboard, and applications used [9, 16]. A few studies went a step further, taking into account more context from other sources such as audio and video recordings, calendar or network connection data (e.g. [4]).

In our work, we conducted lab and field studies and demonstrated the feasibility of using biometric sensor data to measure software developer’s interruptibility with high accuracy, and compared biometric with computer interaction sensors in their ability to determine interruptibility [26, 27].

Reducing Interruption Cost. By optimizing the timing of interruptions, one can foster focus at work and reduce interruption cost. Knowledge workers already use various techniques themselves to optimize the timing, for instance, by using instant messaging to negotiate availability beforehand [22], or by using manual and physical indicators, such as headphones or a closed office door to either signal unavailability or tune out distractions [24]. Researchers have primarily identified two ways to optimize the timing of interruptions: deferring interruptions to task boundaries or continuously measuring and indicating interruptibility even during tasks. Since working memory is usually low at task boundaries, the defer-to-boundary policy aims at determining these natural breakpoints during work and delaying interruptions, such as email notifications, to these more opportune moments [2, 16]. Continuously measuring and indicating interruptibility is particularly useful to reduce in-person interruptions at inopportune moments, which have been shown to be particularly costly [24]. Researchers have developed multiple approaches to indicate the current interruptibility state to potential interrupters, for instance in the form of tools installed on the computer [4], wall projectors [15], or physical lights [6, 25]. In our work we developed FlowLight, a physical indicator for interruptibility placed on a knowledge worker’s desk and controlled automatically through sensing computer activity [25].

3 SENSING INTERRUPTIBILITY

With the advances in sensor technology, in particular biometric sensors, researchers have started the use of a variety of sensors to assess a person’s interruptibility. In our research, we investigated (a) whether biometric sensors can be used to accurately measure a person’s interruptibility, and (b) which combination of biometric and computer interaction data is best to predict interruptibility in the field. We conducted three studies: a short lab, a short field and a long field study with a total of 33 participants with a variety of backgrounds. Table 1 provides an overview of the studies.

3.1 Study Design

All studies follow a similar procedure. Participants were asked to either work on predefined coding tasks (lab study) or on their own tasks (field studies) and were interrupted at random intervals to rate their interruptibility. For study 1 and 2, we focused on the use of biometric sensors to predict interruptibility. For study 3, we added computer interaction sensors for a first comparison of the two data sources to predict interruptibility in the field.

Study	Type	Length	Participants	Sensors
1	Lab	1 hour	10 graduate students	Empatica E3, Neurosky Mindband
2	Field	2 hours	10 developers from 4 companies	Empatica E3, Neurosky Mindband
3	Field	2 weeks	13 developers from 3 companies	Computer Monitoring, Polar H7, Fitbit Charge 2

Table 1: Conducted studies

Across all three studies, we used a variety of biometric sensors, including a Neurosky MindBand to collect EEG and eye blink data; an Empatica E3 to collect EDA, skin temperature, and BVP data; a Fitbit Charge 2 to collect HR data with an optical sensor, as well as movement, and sleep data; and a Polar H7 to measure HR data with an ECG-based sensor. To gather computer interaction data, we used our own monitoring tool called WorkAnalytics [21]. After all studies, we conducted short interviews to learn more about the participants’ experience with interruptions and the sensors.

3.2 Feature Extraction

To develop a model that allows us to predict a worker’s interruptibility, we first cleaned the data and extracted several features from the sensor data. For instance, from the EEG raw data, we extracted brain wave frequency bands and combinations thereof that have previously been linked to mental states such as alertness or attention [5]. The EDA signal needs to be split into its phasic and tonic components that have previously been linked to arousal and specific emotions [7]. From the computer interaction data, specifically the application usage data, we extracted activity categories through a combination of automatic and manual coding and calculate the time spent per category [20]. All extracted feature groups and samples of some of the features are presented in Table 2.

For study 1 and 2, we collected *baseline* measures for participants while they watched a calming fish tank video for two minutes and used them to normalize the values. This normalization allowed us to develop a model based on multiple participants. For the long study, we normalized the features using a standard scaling technique.

Finally, we had to identify the *time windows* used for extracting and calculating feature values. In the shorter studies (study 1 and 2), we found that a short window of 10s works best for the predictive power biometric data. In the longer study, we also considered longer time windows up to 3 hours for the variety of features, and found that the optimal time window varies widely per feature, e.g. shorter time windows of 20 seconds for HR features, several minutes for application activity and up to 3 hours for calendar features.

3.3 Interruptibility Prediction

To predict participants’ ratings of their interruptibility, we applied a machine learning approach to the extracted feature data. The machine learning classifier for study 1 was trained using a Naïve Bayes classifier. The classifier achieved an accuracy of 91.5% (two states) and 43.9% (five states) to predict interruptibility in the lab, improving significantly over a simple majority classifier, which would always predict the most common class. In study 2 that was conducted in the field, only the classification into two states improved significantly over a majority classifier with 78.6% accuracy. Our results demonstrate the feasibility of using biometric sensors

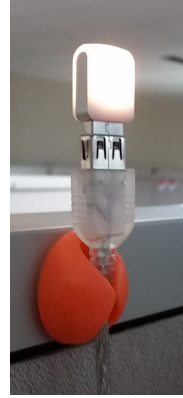
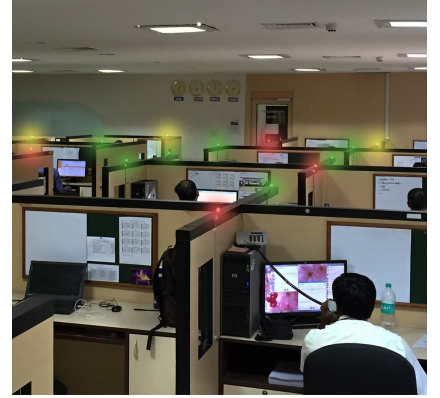
Biometric Features and Examples	
Brain	EEG Frequency bands (e.g. α) and combinations (e.g. α/γ), Attention and Meditation (e.g. max)
Eye	Eye blinks (e.g. # per minute)
Heart	BVP Amplitude (e.g. max peak amplitude), HR (e.g. mean), HRV (e.g. PNN50)
Skin	EDA Phasic signal (e.g. mean peak amplitude), EDA tonic signal (e.g. mean), temperature (e.g. mean)
Movement	Steps (e.g. number per minute)
Sleep	Duration (e.g. total minutes), quality (e.g. restless minutes)
Computer Interaction Features and Examples	
Time	Current time (e.g. hour of day), circadian rhythm (e.g. hour arrived at work)
Calendar	Meetings (e.g. # upcoming meetings)
User Input	Keystrokes (e.g. # delete key presses), mouse clicks (e.g. # left clicks), mouse moves (e.g. moved pixels per minute), mouse scrolls (e.g. time spent scrolling)
Applications	Activity categories (e.g. time spent coding), focus duration (e.g. max. time in one window), activity switches (e.g. # window switches per minute)

Table 2: Features of the interruptibility model

to predict interruptibility, but also point out that external influences and individual differences impact the accuracy of the models. The larger data set collected in study 3 allowed to build a more reliable model, achieving an overall accuracy for individually trained models of 75.3%, which is a 26.6% improvement over a majority classifier. A general model achieved 69.8% accuracy when applied to unseen participants, improving over a majority classifier by 18%. This shows that we can build and apply a general model for interruptibility with reasonable accuracy, which solves the cold-start-problem successfully. When comparing features, we found that the computer interaction features provided more predictive value compared to data from the Fitbit and Polar sensors (74.8% accuracy vs. 68.3%), and that a combination of all works best (75.3% accuracy). Therefore, computer interaction data can serve as a good starting point to sense interruptibility, but especially during tasks without extensive usage of the computer (e.g. reading or thinking), biometric sensors can complement computer interaction data well.

4 REDUCING INTERRUPTION COST

To foster focused work and reduce the number of interruptions at inopportune moments, researchers have developed approaches that postpone interruptions, such as emails, or indicate interruptibility to co-workers on the computer or externally. In our work, we developed the FlowLight, an application and a physical traffic light like LED lamp that indicates interruptibility to co-workers [25]. The FlowLight approach also updates the computer-based instant messaging status, however, with its external indicator it primarily focuses on external and in-person interruptions that have been shown to be one of the most disruptive kind of interruptions due to the immediate nature and often prolonged duration [24].

**(a) FlowLight****(b) Office setup with FlowLights****Figure 1: FlowLights**

4.1 FlowLight Approach

The FlowLight consists of a computer application to automatically determine a user's interruptibility state and a physical LED light to indicate this state to co-workers. The physical LED light is mounted at the desk, cubicle wall or office door of a knowledge worker (see Figure 1). Similar to a traffic light the light uses different colors to indicate a person's interruptibility: *available* as green, *busy* as red, *do not disturb (DnD)* as pulsating red, *idle or away* as yellow. The FlowLight application calculates the user's interruptibility state on the fly based on mouse and keyboard activity. Specifically, the application determines a personalized model of interruptibility for each user based on heuristics for each type of input, the user's historical interaction data and a smoothing function. Based on insights from early pilot studies, the application sets the light to red for approximately 9% of the time spent on the computer (4% for pulsating red). Whenever the user's interruptibility status changes, the FlowLight application updates the color of the LED light as well as the Skype presence status.

4.2 Evaluation and Results

To evaluate the FlowLight's ability to reduce interruption cost, we conducted a large-scale field study with 449 participants of one multi-national company working in 12 different countries. In the study, we asked participants to self-report interruptions for one week before we installed the FlowLight, and again after they familiarized themselves with the new system for a week. We further conducted surveys and interviews, collecting a rich qualitative and quantitative dataset of 183 survey responses, 23 interview transcripts, 36 interruption logs, 47 FlowLight usage data logs and activity logs from all 449 participants.

Our analysis showed that the FlowLight significantly reduced the amount of interruptions by 46%. One of FlowLight's major benefits, as stated by participants, was that the small lights increased the awareness on interruption cost. For instance, one participant mentioned that *"the pilot increased the sensitivity to interruption. Team members think more about whether an interrupt is necessary and try to find a suitable time"* (S45). Furthermore, participants stated that the FlowLight increased their productivity: first, because it reduced the number of expensive interruptions; and second, because

it encouraged them to focus or keep focusing on their task either because they realized that their light was green for a while or because it just turned red. Overall, 85.5% of all users continued using the FlowLight several months after the study period ended.

In terms of the interruptibility status, 71% of the participants perceived the calculated status to be accurate. However, there is potential for improvement, especially in situations in which the participants' focus was high but they did not interact with the computer, e.g. when reading or sketching on a piece of paper. In these cases, biometric sensors worn by the user have a big potential to increase the accuracy of the system.

Overall, the FlowLight's success demonstrates that the combination of a physical indicator with an automatic interruptibility measurement that reduces the cost of manually maintaining the interruptibility status, is an effective means to reduce interruption cost and increase worker's focus and productivity.

5 DISCUSSION AND SUMMARY

Software developers often need to focus to successfully perform their work. If the developer's focus is disrupted, it can lead to a lower quality in the outcome and decrease performance significantly, especially for highly demanding tasks and activities, such as program comprehension and bug fixing. In our research, we aimed at reducing the number of external in-person interruptions—one particularly expensive disruption to focus. We explored the use of biometric and computer interaction sensors to measure a developer's interruptibility, and developed the FlowLight approach that uses one such automatic interruptibility measure and combines it with a physical traffic light like LED light to indicate the interruptibility to co-workers. The results of our studies show that biometric and computer interaction sensors can be used in the lab and field to accurately measure a developer's interruptibility and that the FlowLight is able to significantly reduce interruption cost and increase focus, motivation and productivity.

At the same time, external interruptions only make up half of the interruptions that a knowledge worker experiences in a day. The other half stems from self-interruptions [8]. These self-interruptions can also have a big impact on the developer's focus and performance. With a more holistic measure of a developer's focus, we might be able to reduce the cost of self-interruptions and better support developers in their work. For instance, by automatically detecting when a developer's focus is decreasing, we might be able to intervene, e.g. by reducing distracting content on the screen that might cause self-interruptions or by suggesting to take a break. Furthermore, by knowing when a developer is more or less focused during the day, we might be able to optimize the work day by scheduling highly demanding tasks during times of high focus. Overall, the more accurate and less invasive we can automatically measure developers' focus, the better we are able to support them in their work and increase their productivity and well-being.

ACKNOWLEDGMENTS

The authors would like to thank all study participants. This work was funded by SNF under Grant No.: 200021_150050.

REFERENCES

- [1] Brian P Bailey and Shamsi T Iqbal. 2008. Understanding changes in mental workload during execution of goal-directed tasks and its application for interruption management. *ACM Transactions on Computer-Human Interaction (TOCHI)* 14, 4 (2008), 21.
- [2] Brian P Bailey and Joseph A Konstan. 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 (2006), 685–708.
- [3] Brian P Bailey, Joseph A Konstan, and John V Carlis. 2001. The effects of interruptions on task performance, annoyance, and anxiety in the user interface. In *Proceedings of INTERACT*, Vol. 1. 593–601.
- [4] James Bo Begole, Nicholas E Matsakis, and John C Tang. 2004. Lilsys: sensing unavailability. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work*. ACM, 511–514.
- [5] Hans Berger. 1933. Über das elektroencephalogramm des menschen. *Archiv für Psychiatrie und Nervenkrankheiten* 98, 1 (1933), 231–254.
- [6] Milan Z Bjelica, Bojan Mrazovac, Istvan Papp, and Nikola Teslic. 2011. Busy flag just got better: Application of lighting effects in mediating social interruptions. In *MIPRO, 2011 Proceedings of the 34th International Convention*. IEEE, 975–980.
- [7] Wolfram Boucsein. 2012. *Electrodermal activity*. Springer.
- [8] Mary Czerwinski, Eric Horvitz, and Susan Wilhite. 2004. A diary study of task switching and interruptions. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 175–182.
- [9] James Fogarty, Andrew J Ko, Htet Htet Aung, Elspeth Golden, Karen P Tang, and Scott E Hudson. 2005. Examining task engagement in sensor-based statistical models of human interruptibility. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. ACM, 331–340.
- [10] Thomas Fritz, Andrew Begel, Sebastian C Müller, Serap Yigit-Elliott, and Manuela Züger. 2014. Using psycho-physiological measures to assess task difficulty in software development. In *Proceedings of the 36th International Conference on Software Engineering*. ACM, 402–413.
- [11] Victor M González and Gloria Mark. 2004. Constant, constant, multi-tasking craziness: managing multiple working spheres. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 113–120.
- [12] Nitesh Goyal and Susan R Fussell. 2017. Intelligent Interruption Management using Electro Dermal Activity based Physiological Sensor for Collaborative Sense-making. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (2017), 52.
- [13] Andreas Haag, Silke Goronzy, Peter Schaich, and Jason Williams. 2004. Emotion recognition using bio-sensors: First steps towards an automatic system. In *Tutorial and research workshop on affective dialogue systems*. Springer, 36–48.
- [14] Eija Haapalainen, SeungJun Kim, Jodi F Forlizzi, and Anind K Dey. 2010. Psycho-physiological measures for assessing cognitive load. In *Proceedings of the 12th ACM international conference on Ubiquitous computing*. ACM, 301–310.
- [15] Juan David Hincapié-Ramos, Stephen Volda, and Gloria Mark. 2011. Sharing availability information with InterruptMe. In *Proceedings of the 13th international conference on Ubiquitous computing*. ACM, 477–478.
- [16] Shamsi T Iqbal and Brian P Bailey. 2008. Effects of intelligent notification management on users and their tasks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 93–102.
- [17] Gloria Mark, Shamsi T Iqbal, Mary Czerwinski, and Paul Johns. 2014. Bored Mondays and focused afternoons: the rhythm of attention and online activity in the workplace. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 3025–3034.
- [18] Santosh Mathan, Stephen Whitlow, Michael Dorneich, Patricia Ververs, and Gene Davis. 2007. Neurophysiological estimation of interruptibility: Demonstrating feasibility in a field context. In *Proceedings of the 4th International Conference of the Augmented Cognition Society*.
- [19] Andre N Meyer, Laura E Barton, Gail C Murphy, Thomas Zimmermann, and Thomas Fritz. 2017. The work life of developers: Activities, switches and perceived productivity. *IEEE Transactions on Software Engineering* 43, 12 (2017), 1178–1193.
- [20] André N Meyer, Thomas Fritz, Gail C Murphy, and Thomas Zimmermann. 2014. Software developers' perceptions of productivity. In *Proceedings of the 22nd ACM SIGSOFT International Symposium on Foundations of Software Engineering*. ACM, 19–29.
- [21] Andre N. Meyer, Gail C. Murphy, Thomas Zimmermann, and Thomas Fritz. 2017. Retrospecting on Work and Productivity: A Study on Self-Monitoring Software Developers' Work. *Proc. ACM Hum.-Comput. Interact.* CSCW (2017), 79:1–79:24.
- [22] Bonnie A Nardi, Steve Whittaker, and Erin Bradner. 2000. Interaction and outercation: instant messaging in action. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work*. ACM, 79–88.
- [23] Akane Sano and Rosalind W Picard. 2013. Stress recognition using wearable sensors and mobile phones. In *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, 671–676.
- [24] Edward R Sykes. 2011. Interruptions in the workplace: A case study to reduce their effects. *International Journal of Information Management* 31, 4 (2011), 385–394.

- [25] Manuela Züger, Christopher Corley, André N Meyer, Boyang Li, Thomas Fritz, David Shepherd, Vinay Augustine, Patrick Francis, Nicholas Kraft, and Will Snipes. 2017. Reducing Interruptions at Work: A Large-Scale Field Study of Flow-Light. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 61–72.
- [26] Manuela Züger and Thomas Fritz. 2015. Interruptibility of software developers and its prediction using psycho-physiological sensors. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 2981–2990.
- [27] Manuela Züger, Sebastian C Müller, André N Meyer, and Thomas Fritz. 2018. Sensing Interruptibility in the Office: A Field Study on the Use of Biometric and Computer Interaction Sensors. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM.