
Time Estimation Bias in Knowledge Work: Tasks With Fewer Time Constraints Are More Error-Prone

Yoana Ahmetoglu

UCL Interaction Centre
University College London
London, UK
yoana.petrova.15@ucl.ac.uk

Anna L. Cox

UCL Interaction Centre
University College London
London, UK
anna.cox@ucl.ac.uk

Duncan P. Brumby

UCL Interaction Centre
University College London
London, UK
d.brumby@ucl.ac.uk

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).

Copyright held by the owner/author(s).

CHI Conference on Human Factors in Computing Systems Extended Abstracts, April 25–30, 2020, Honolulu, HI, USA.
ACM 978-1-4503-6819-3/20/04.

<https://doi.org/10.1145/3334480.3382917>

Abstract

Previous research has found that people often make time estimation errors in their daily planning at work. However, there is limited insight on the types of estimation errors found in different knowledge work tasks. This one-day diary study with 20 academics compared the tasks people aimed to achieve in the morning with what they actually did during the day. Results showed that participants were good at estimating the duration of time-constrained tasks, such as meetings, however they were biased when estimating the time they would spend on less time-constrained tasks. Particularly, the time needed for email and coding tasks was underestimated, whereas the time needed for writing research and planning activities was overestimated. The findings extend previous research by measuring *in situ* whether some tasks are more prone to time estimation errors than others. Planning and scheduling (AI) tools could incorporate this knowledge to help people overcome these time estimation biases in their work.

Author Keywords

Daily Planning; Time Management; Knowledge Work; Time Estimation Bias; Productivity.

CCS Concepts

•Human-centered computing → *Interaction design theory, concepts and paradigms*;

Introduction

Daily planning and time management are essential skills needed for modern knowledge work. However, many people struggle with planning and time management – many of the tasks that people set for themselves in the morning are still on their to-do list at the end the workday and extra tasks are often added [5]. A failure to complete planned daily tasks can lead to stress, rumination over work during periods of rest and relaxation and sleep problems [12].

Observational studies have found that daily planning is an ongoing challenge for knowledge workers such as academics [7] and software developers [9]. Academics, for instance, need to predict how long different tasks will take to complete in order to meet multiple deadlines [7]. However, such estimates are complex. People may lack the experience or expertise to deal with such complexity and to make accurate time estimates about how long their tasks will take to complete. Time estimates can also be affected by known biases, such as optimism and overconfidence [3].

Only a limited number of studies have examined how accurately people estimate the duration of their daily tasks. Claessens' [5] study had 29 research and development employees keep a diary for three weeks. Participants had to report their plans in the morning and the proportion of each task they managed to complete at the end of each day. These diaries showed that people did not complete 27% of their daily planned work. However, while this study recognised a planning fallacy, it did not consider how accurate people were at estimating how long tasks would take to complete.

Newman [11] conducted a diary study to understand why busy people often did not complete all the tasks that they set out each day to achieve. Participants in the study were asked to make a plan in the morning and to keep a precise

time sheet diary with every tasks they did for a day. Results showed that people underestimated how quickly they could complete information work compared to face-to-face meetings. This study offers evidence that some tasks are more prone to time estimation errors than others. However, the study was conducted before the wide-spread use of online communications and with a sample of people who are not necessarily users of technology at work (e.g. a furniture maker). In contrast, knowledge workers today perform many different types of information work tasks, such as email, collecting and analysing information, writing reports. Those tasks have not been a focus of research on time estimation errors. Examining the potential for biases in these different kinds of tasks may be useful for informing the design of planning and scheduling tools to help people more accurately plan their work tasks.

Existing research on AI support for daily planning and scheduling of work focuses mainly on tasks with known duration, such as calendar appointments. Intelligent task management systems and AI assistants, such as Microsoft Cortana, attempt to automate the estimation the duration of calendar tasks [15] and predict the time it takes to complete them [14]. Another stream of research works on developing conversational agents such as taskBot [13] and calendar.help [6] to help automatic scheduling of appointments with known duration. However, there is not much work beyond arranging and scheduling well-defined tasks, such as meetings. Few tools exist for tasks that are less clearly constrained in time, such as replying to emails. Those tasks are typical for many knowledge workers and it is often more difficult to estimate how long these tasks will take to complete.

In this paper, we examine how accurate knowledge workers are in their estimates of how long different kinds of

work tasks will take to complete. In doing this we aim to find out whether some tasks are more prone to estimation errors than others. Planning and scheduling (AI) tools could then incorporate this knowledge to help people overcome error in their estimations. We recruited academics and research students in a university as they are able to set their own daily agendas. This group of people also do a mix between solo and team activities over a wide-range of projects, which allows to gain observations about many different tasks. Due to space limitations, we report and discuss part of our quantitative results obtained through the diary method.

Method

Participants

Twenty participants (nine male) took part in the study with a mean age of 29 years ($SD = 4.8$). They were academics and early career researchers at UK universities (3 \times lecturers, 2 \times post docs, 13 \times PhD students and 2 \times internship graduate students). Participation was voluntary. The study was approved by the UCL Ethics Committee.

Design

We were interested in comparing what people aimed to achieve at work with what they actually did. All participants filled in a two-stage report: they indicated their plans in the morning and reported their behaviours continuously in a diary throughout the day.

Procedure

Participants chose one typical day at work to do the study. They listed in the morning all tasks they aimed to achieve during their workday and were asked to estimate the likely duration for each task. They used pen and paper report forms. They were free to consult their calendars or to-do lists to remind themselves of their agendas. The morning

| <i>Describe your current task</i> | <i>Start</i> | <i>End</i> |
|-----------------------------------|--------------|------------|
| Email and communications | 9:40 | 10:37 |
| Writing report | 10:37 | 11:08 |
| Email and communications | 11:08 | 11:16 |
| Socialising | 11:16 | 11:26 |
| Email and communications | 11:26 | 11:45 |
| (Lunch break) | | |
| Writing report | 12:21 | 13:20 |
| Meeting | 13:30 | 14:30 |
| Fixing coding issues | 14:35 | 17:00 |

Table 1: Example of anonymized diary data

list was then handed to the researcher and was not given back to the participants until the end of the workday.

After filling the morning lists, participants were asked to continue their workday as usual while keeping a pen and paper record of the main task they decided to work on. They had to report the start and end times of each task, giving as much detail as possible. At the end of the day, participants were given back their morning lists together with their diaries. They were asked to fill in any missing details and explain the nature of the tasks reported.

Data Analysis

Plans and diaries were transferred to digital tables (see Table 1). Each reported activity was anonymized by deleting information about specific projects and people. Lunch time was omitted from analysis because we wanted participants to feel as minimally observed as possible. A new comparison table was created for each participant with estimated time and actual time spent on all reported tasks (see table 2). This table was used to compare the estimated and actual total workday duration for each participants.

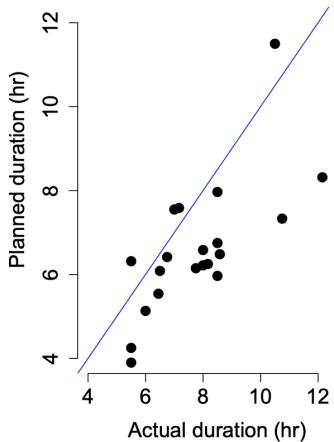


Figure 1: The association between planned (estimated) work duration and actual workday duration for each participant. The blue line shows a perfect relationship with no bias. Data points below the line indicate that the planned duration was longer than the actual one.

| Type of task | Estimated time | Spent time |
|--------------------------|----------------|------------|
| Research report | 180 min | 89 min |
| Meeting | 60 min | 60 min |
| Transcribe | 120 min | 0 min |
| Email and communications | 60 min | 79 min |
| Coding task | 60 min | 145 min |

Table 2: Example of comparison table used for data analysis

All tasks were then thematically analysed and sorted into different categories, for instance, writing research, scheduled meetings and coding. The accuracy of estimates for different categories of tasks was analysed (see table 3). In cases where a participant had more than one task of a given type (i.e., to read two separate works), the averaged duration was used for mean statistics for each type of task (i.e. reading research).

Finally, a 20% threshold was used to typcategories of tasks into correctly predicted (less than 20% bias), underestimated (took 20% longer) or overestimated (took 20% shorter). The threshold follows previous research on estimation of effort for knowledge work (e.g., in software developers [10]).

Results

The average duration of workday tasks was estimated to be 7 hours 44 min ($SD = 102$ min) whereas the actual duration of workday tasks reported was 6 hours 40 min ($SD = 101$ min), including breaks but excluding lunch time. A paired sample t -test suggested that participants planned to work for significantly longer than they actually did, $t(19) = 4.01$, $p = .001$.

| Type of tasks | N | Estimated | Actual |
|------------------------------|----|-----------|-----------|
| Correct | | | |
| Scheduled meetings | 16 | 78 (43) | 64 (42) |
| Scheduled seminars | 6 | 130 (113) | 110 (95) |
| Reading research | 11 | 68 (56) | 66 (31) |
| Creating presentation | 5 | 91 (27) | 77 (52) |
| Underestimation bias | | | |
| Email and communications | 12 | 51 (35) | 68 (55) |
| Coding and data analysis | 5 | 90 (30) | 182 (240) |
| Overestimation bias | | | |
| Writing research | 12 | 119 (78) | 78 (54) |
| Planning and decision-making | 5 | 63 (50) | 19 (28) |

Table 3: Summary of time spent on different types of tasks. N = number of participants. Time values are *Mean (SD)* in minutes.

Figure 1 shows a scatter-plot of actual and planned workday duration for each participant in the study. The diagonal line indicates a perfect prediction. Data points below the diagonal line indicate optimistic prediction where participants planned to work for longer than they did. The plot shows that a majority of participants (16 out of 20) were optimistic in their estimates and expected to work for longer than they actually did.

Table 3 summarises the estimation errors in predicting different types of tasks. The next paragraphs discuss each group of tasks in turn, starting with correctly predicted, followed by underestimated and, finally, overestimated.

Correctly predicted

Sixteen participants attended scheduled meetings. On average they planned to spend 78 min and they actually spent 64 min in meetings. Examples of the type of meetings participants attended are meetings booked by a student, meet-

ing one's supervisor, one's research group or team. Apart from scheduled meetings, six participants attended scheduled seminars which were planned as 130 min and lasted 110 min. Examples include weekly departmental seminars. Further, eleven participants planned to spend 68 min on reading tasks. Later, they reported spending 66 min on those tasks. Examples of such tasks include reading an article prior to supervisory meeting, or reading a draft of students' work. Finally, three participants planned to spend 91 min and spent 77 min on creating presentations, such as preparing slides for their Viva examinations or for their lectures at university. Almost all of those tasks were related to an important commitment (e.g. lecture or a meeting) on the same or on the next day.

Underestimation bias

Twelve participants planned to spend 51 min and reported spending 68 min on average on email and communication activities. Coding and data analysis was the other group of tasks which took longer than expected. Five participants planned to spend 90 min but actually spent 182 minutes on those tasks. The most frequent reasons for delays were either an unexpected code issue that needed urgent attention or because they forgot to preprocess the data.

Overestimation bias

Twelve participants expected to work on at least one research writing task with an estimated mean duration of 119 min and an actual duration of 78 min. Some participants defined a clear objective in their writing tasks, such as *fix introduction comments* whereas others aimed to spend a certain amount of time on their research papers or chapters, such as *write research report for 2 hours*. Next, five participants planned to do a planning or decision-making task with a mean duration of 63 min and actual time spent of 19 min. For instance, they wanted to plan their month ahead,

plan how to approach a writing task or decide on a research direction. Overestimated tasks tended to be left unfinished unless they were related to a very pressing deadline.

Discussion

The results of the study show that participants were optimistic in their time estimates for work activities, planning to work one hour longer than they actually did. Participants were good at estimating the completion time for tasks that had very clear time constraints and were already scheduled in their calendars, such as those related to urgent deadlines, meetings or preparation for meetings. However, participants showed an estimation bias for tasks which did not have clear time constraints and were not scheduled in their calendars: they spent more time than was estimated on coding, data analysis, and email tasks. Although participants did complete these tasks, it often took far longer than expected. In contrast, participants reported spending less time on writing, planning work and decision-making tasks than they had estimated. They tended to leave these tasks incomplete, conceivably because they had less time left than they had expected, due to other tasks taking longer than initially thought.

Previous literature has documented that people experience challenges in managing the time spent on email [8]. A qualitative study revealed that employees felt a loss of control in managing email, psychological pressure to respond quickly to email messages, and fear of missing important information communicated by email [1]. A possible explanation for our finding is that people lose track of time when working on emails due to the psychological pressure they experience. In addition, previous research also suggests that different email checking behaviours can impact the total amount of time spent on emails, with once per day strategy being the most efficient one [2].

People spent less time on writing, planning work and decision-making tasks than they had estimated. Arguably, this kind of work is the most important one in research-oriented universities. Previous research has found that writing requires high levels of focus and is often negatively impacted by interruptions, such as meetings and emails, which are omnipresent in many knowledge work environments. To support completion of writing tasks, previous research has suggested the use of micro-tasking strategies and tools where long writing tasks are broken down into small components and done whenever workers get a chance to do them [4, 6]. Given that people had difficulties in estimating the completion time of these tasks, encouraging the use of micro-tasking strategies is in line with the current findings. However, more research is needed to investigate how well knowledge workers can estimate the duration of micro-tasks, and whether this strategy will lead to less estimation errors.

An interesting research question for future work would be to explore the impact of interventions for improving the identified time estimation bias in daily planning. Would people correct their estimates without changing the way they allocate time to tasks or would they change the way they allocate time to tasks during the day in order to match their estimates? A longitudinal study might give an answer to this question where participants are given feedback on their time estimation errors. Similar work on improving estimation skills for daily time spent on social media found that people felt more in control of their time even though they did not change the amount of time they spent on social media [16].

The current study has limitations. Even though the morning plan was given back to the researcher in the morning and participants were instructed to behave as usual during the day, they might have had a recollection of their plans.

Hence, they might have put more effort than usual to follow their plans. Nonetheless, the current study still found time estimation errors. It is likely that those errors would be higher when people are not explicitly asked to make a morning plan. Further, the sample of participants included PhD students, postdocs and lecturers. Future research can extend this work with other groups of knowledge workers. In addition, there might be important differences in the type of work required at different levels of seniority for the same job role. In the future, we aim to explore how those differences can impact time estimation errors and whether different planning support tools work better for different job levels in knowledge work.

In conclusion, this paper has explored how accurate people are estimating how long different kinds of daily work tasks will take to complete. The results show that participants found it difficult to estimate how long they would work on open-ended tasks that were poorly constrained. In particular, participants underestimated the time that they needed to work on email and coding tasks and overestimated the time that they needed to work on writing and planning activities. These results suggest that there is a need for more support to improve accuracy in estimation of tasks which are not typically scheduled as events in the calendar, such as email, coding, data analysis, writing and planning tasks.

Acknowledgements

This work was supported by a studentship from the Division of Psychology and Language Sciences, University College London. We'd like to thank the reviewers for their constructive feedback for improving this paper. We also wish to thank all the participants who were willing to take the time to keep a daily diary and to share their experiences with us.

REFERENCES

- [1] Inge Alberts. 2013. Challenges of information system use by knowledge workers: The email productivity paradox. *Proceedings of the American Society for Information Science and Technology* 50, 1 (2013), 1–10. DOI:
<http://dx.doi.org/10.1002/meet.14505001089>
- [2] Adam Bradley, Duncan P. Brumby, Anna L. Cox, and Jon Bird. 2013. How to Manage Your Inbox: Is a Once a Day Strategy Best?. In *Proceedings of the 27th International BCS Human Computer Interaction Conference (BCS-HCI '13)*. BCS Learning Development Ltd., Swindon, GBR, Article Article 20, 6 pages.
- [3] Roger Buehler, Dale Griffin, and Johanna Peetz. 2010. The planning fallacy: Cognitive, motivational, and social origins. In *Advances in experimental social psychology*. Vol. 43. Elsevier, 1–62.
- [4] Justin Cheng, Jaime Teevan, Shamsi T Iqbal, and Michael S Bernstein. 2015. Break it down: A comparison of macro-and microtasks. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 4061–4064.
- [5] Brigitte JC Claessens, Wendelien Van Eerde, Christel G Rutte, and Robert A Roe. 2010. Things to do today...: A daily diary study on task completion at work. *Applied Psychology* 59, 2 (2010), 273–295.
- [6] Shamsi T. Iqbal, Jaime Teevan, Dan Liebling, and Anne Loomis Thompson. 2018. Multitasking with Play Write, a Mobile Microproductivity Writing Tool. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology (UIST '18)*. Association for Computing Machinery, New York, NY, USA, 411–422. DOI:
<http://dx.doi.org/10.1145/3242587.3242611>
- [7] Amirrudin Kamsin, Ann Blandford, and Anna L. Cox. 2012. Personal Task Management: My Tools Fall Apart When I'm Very Busy!. In *CHI '12 Extended Abstracts on Human Factors in Computing Systems (CHI EA '12)*. Association for Computing Machinery, New York, NY, USA, 1369–1374. DOI:
<http://dx.doi.org/10.1145/2212776.2212457>
- [8] Kim McMurtry. 2014. Managing email overload in the workplace. *Performance Improvement* 53, 7 (2014), 31–37.
- [9] A. Meyer, E. T. Barr, C. Bird, and T. Zimmermann. 2019. Today was a Good Day: The Daily Life of Software Developers. *IEEE Transactions on Software Engineering* (2019), 1–1. DOI:
<http://dx.doi.org/10.1109/TSE.2019.2904957>
- [10] Kjetil Moløkken and Magne Jørgensen. 2003. A review of surveys on software effort estimation. (2003).
- [11] William M Newman. 2004. Busy days: exposing temporal metrics, problems and elasticities through diary studies. In *CHI 2004 Workshop on Temporal Issues in Work*. Citeseer.
- [12] Christine J Syrek, Oliver Weigelt, Corinna Peifer, and Conny H Antoni. 2017. Zeigarnik's sleepless nights: How unfinished tasks at the end of the week impair employee sleep on the weekend through rumination. *Journal of Occupational Health Psychology* 22, 2 (2017), 225.

- [13] Carlos Toxtli, Andrés Monroy-Hernández, and Justin Cranshaw. 2018. Understanding chatbot-mediated task management. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 58.
- [14] Ryen W White, Ahmed Hassan Awadallah, and Robert Sim. 2019. Task Completion Detection. (2019).
- [15] Ryen W White and Ahmed Hassan Awadallah. 2019. Task duration estimation. In *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining*. ACM, 636–644.
- [16] Yan Zhou, Jon Bird, Anna L Cox, and Duncan Brumby. 2013. Estimating usage can reduce the stress of social networking. (2013).