

# Barriers to Engagement with a Personal Informatics Productivity Tool

Emily I M Collins

UCL Interaction Centre  
Gower Street  
London, W1CE 6BT  
e.collins@ucl.ac.uk

Anna L Cox

UCL Interaction Centre  
Gower Street  
London, W1CE 6BT  
Anna.cox@ucl.ac.uk

Jon Bird

City University London  
School of Engineering &  
Mathematical Sciences  
London, EC1V 0HB  
Jon.bird@city.ac.uk

Cassie Cornish-Tresstail

UCL Interaction Centre  
Gower Street  
London, W1CE 6BT  
cassandra.cornish-tresstail.10@ucl.ac.uk

## ABSTRACT

Technology helps us get work done but also provides many distractions. As a result, seemingly unproductive activities such as social networking sites (SNS) cause considerable stress. This paper reports a series of studies into whether personal informatics (PI) tools for productivity can make people more aware of their SNS usage and encourage behaviour change. The first two studies took an in-the-wild approach, encouraging students to use a PI tool, RescueTime, to improve their estimations of how much time they spent using SNS, in line with research that had used this technique to reduce participants' stress. However, participants simply did not engage with RescueTime in the studies. A further interview study found that there are four barriers that inhibit engagement with this PI tool and reduce its potential to facilitate behaviour change. In particular, the way it presents data lacks: *salience*; *contextual information*; *credibility*; and *action advice*.

## Author Keywords

Personal informatics tools; productivity; barriers; social networking; engagement; stress

## ACM Classification Keywords

H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

## INTRODUCTION

The ubiquity of modern technologies has revolutionised how people complete daily tasks. Individuals are more reliant on computers than ever before, and while this has made a number of activities easier and quicker, it has also increased the opportunities for interruption. Not only can individuals be actively notified of incoming information in the form of e-mails and instant messages, but they also have easy access to a number of other sources of distraction. While in some workplaces, internet access may be restricted or monitored, thus removing the temptation to access distractions such as social networking sites (SNS), this is not the case for students. Like knowledge workers, their productivity relies heavily on their own time management and therefore distractions by anything deemed to be unproductive may be especially

Permission to make digital or hard copies of all or part 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. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

OZCHI '14, Dec 2-5, 2014, Sydney, Australia

Copyright © 2014 ACM 978-1-4503-0653-9... \$15.00

<http://dx.doi.org/10.1145/2686612.2686668>

stressful.

Students are the most prolific users of SNS [11], and despite being able to use SNS in various productive ways from coordinating meetings to working on shared assignments, they do report SNS use to be a considerable source of stress [9]. This stress may be, at least in part, reliant on an inaccurate perception of how much time individuals spend on SNS. Junco [10] found that self-reported SNS use was almost six times higher than their actual use, as measured by logging software. Efforts to make individuals more aware of their actual level of use may have a positive impact on how stressed they feel, and personal informatics (PI) tools seem particularly well-suited to achieving this.

Consequently, a number of PI tools have been developed with the aim of supporting users in becoming more 'productive'. These tools allow individuals to view automatically collected data on how many hours they work, how long they spend on particular documents and when they have been distracted by non-work related applications, such as SNS. While this information should be highly valuable to anyone seeking to improve their productivity (even if just to confirm that their SNS use is not impacting negatively on their work), it is currently unclear whether PI tools are actually successful in this aim. Research in other domains indicates that PI tools are often ineffective at facilitating behaviour change. For instance, despite users' intentions to utilize health apps to increase physical activity, a recent study found that 26% were found to have only been used once and 74% were used fewer than ten times [5]. Similarly, over 50% of wearable activity trackers have been found to be discarded after 6 months [7], indicating that simply owning these devices or applications does not mean that they are engaged with or that the information is applied to behaviour. Previous work assessing the use of other tools that provide feedback to change behaviour, namely a household energy monitor, reported that once the initial novelty of owning the device wore off, users disengaged and remained unmotivated to change their energy use behaviour [14]. It is therefore still unclear whether PI tools for productivity would be effective in helping students monitor SNS use in a way that would be beneficial in terms of changing behaviour or reducing the stress related to these sites.

However, there is research to suggest that people respond well to visualizations of their computer usage, and believe

that it will help them make better use of their time [1]. Li and colleagues [12] found that people tend to reflect on automatically collected information as it is readily available and avoids the typical difficulties experienced in the manual collection and integration of personal data, such as having to organize the information or remember to record it.

Additional evidence that PI tools could be beneficial in this context comes from Zhou et al [16]. In that study, participants completed questionnaire measures of stress and time management behaviours, and having installed RescueTime (a tool that tracks productivity), were then asked to provide daily estimations of how much time they thought they had spent on SNS so far that day. This task motivated participants to check their RescueTime data to assess how accurate their estimate had been, and provided a measure of the discrepancy between perceived and actual SNS use. It was found that participants' estimation accuracy significantly improved over the two weeks of estimations. Interestingly, participants also experienced a reduction in their perceived stress and an increase in their satisfaction with their levels of SNS use. This suggested that a lack of awareness of the actual time spent on SNS could be an important factor in the relationship between SNS use and subjective stress, and that RescueTime could help bring about this awareness. However, there was no change in the actual amount of time spent on SNS. This provides further evidence for the hypothesis that although PI tools are effective in raising awareness of behaviour, there are situations in which they may be ineffective at facilitating behaviour change.

In order to further investigate the effects of RescueTime on awareness of SNS use, the perceived stress of this activity and behaviour change, we adopted an in-the-wild design. This paper therefore describes two attempts at replicating and extending of Zhou et al's original study [16]. Despite specifically recruiting participants who were intrinsically motivated to change their SNS behaviour, we did not find any evidence of behaviour change as a result of using the PI tool, and only limited evidence for an increase in awareness of actual behaviour. One explanation for this may be the relatively low levels of engagement with RescueTime. We therefore carried out an interview study that identified several barriers that inhibit engagement with RescueTime that may have impacted on the participants involved in the previous studies. In particular, the way RescueTime presents user data lacks: salience; contextual information; credibility; and action advice. We discuss the implications of our findings for in-the-wild experimental studies into the use of PI tools and for the design of such tools.

## STUDY 1

### Introduction

As stated above, Zhou et al [16] identified that providing daily retrospective estimations by phone (that is, estimating how much time they had spent on SNS from when they woke up until they received the phone call) encouraged participants to refer to RescueTime and consequently improve the accuracy of their estimations. Participants also reported a reduction in stress, although

interestingly, participants did not reduce the amount of time they spent on SNS.

Zhou et al [16] had used retrospective estimations, and so participants had to wait until they had made their estimation before checking RescueTime to see how accurate their estimate had been. However, PI tools are not just effective at recording past behaviour; they can also be used to help set goals for future behaviour, a tactic well established in the behaviour change literature [15]. We decided to investigate whether making prospective estimates about future SNS usage would not only have the same benefits as retrospective estimates but also encourage a change in SNS behaviour.

Very little research exists in relation to the differential effect of prospective and retrospective estimations. Despite arguments for a dissociation between the two in terms of the required processes [3] (and consequently, possible outcomes), there is general agreement that both rely heavily on memory [2]. Therefore, outcomes of prospective estimations, like those of retrospective estimations, may increase awareness of SNS use. This has implications for attempts to decrease SNS related stress, but also more broadly for the effective design of personal informatics tools and other behaviour change interventions.

It was therefore hypothesized that estimations of SNS use would significantly improve over the course of the study as participants gained an awareness of their actual levels of usage. It was also predicted that with improved estimations would come reduced stress and improved perception of time management skills, as participants realize their SNS usage is substantially less than expected, or that it does not impact on productivity [C]. Once again, this is in line with the findings of Zhou et al [16] and the theoretical importance of prospective estimations in goal setting. A final hypothesis was that there would be a reduction in SNS usage because prospective estimations would encourage participants to set goals.

## Method

### Participants

A total of 16 taught postgraduate students took part (seven of whom were female), aged between 22-28 ( $M=23.75$ ,  $S.D.=1.76$ ). The study was conducted during the summer period, during which time postgraduates were still studying and attending classes but did not have exams. They were recruited by posters in university buildings and by advertisements on the university participation pool website. The advert for the study specifically targeted those who were interested in improving their productivity in order to ensure a minimum level of intrinsic motivation. Additional extrinsic motivation to adhere to the study protocol was provided by informing the participants that for every estimation they made, they were awarded an extra entry into a prize draw for Amazon vouchers.

## Materials

As in Zhou et al [16], time management skills were measured by the Time Management Behavior Scale [13] (TMB;  $\alpha = .482-.601$ ), which contained the sub-scales of setting goals ( $\alpha = .691-.801$ ), mechanics of time management ( $\alpha = .746-.808$ ), preference for scheduling/organisation ( $\alpha = .265-.337$ ) and perceived control of time ( $\alpha = .528-.683$ ). Participants were asked to indicate how true 34 statements were to them, which were answerable along a seven point Likert scale.

Stress was measured by the 10 item Perceived Stress Scale [4] (PSS;  $\alpha = .891-.957$ ). This was also answered on a seven point Likert scale.

The questionnaire also contained a question enquiring as to how stressed they felt about their SNS use and how satisfied they were with their current level of use, both answerable on a seven point Likert scale, ranging from 1 (not at all) to 7 (very much so).

RescueTime was used to measure the amount of time spent accessing SNS, in particular, Facebook. Participants were provided with a premium account which allowed them to view data on their “dashboard” (Figure 1), which outlined the amount of time spent on each application and other information, including the number of hours worked and their “productivity pulse”. This is calculated automatically according to the category an application belongs to and its assumed purpose. Applications can be re-categorised by users to better reflect their use, although this was not observed in the present study.



Figure 1. Example display of RescueTime data

## Procedure

Participants were first asked to install RescueTime on all their personal computers and smartphones. Information on their SNS usage was collected for a two week period while the application was in restricted mode (meaning participants were unable to access their own data), in order to provide an indication of baseline SNS use.

Participants completed the above questionnaires, were granted access to their RescueTime data and then began a two week period during which they received daily phone calls. These phone calls asked them to estimate how long they expected to spend on SNS from the phone call to when they went to sleep. They were made at random times between noon and midnight on weekdays only. If they did not answer the call, the same question was sent to them by text message, and if they did not respond within five minutes, this was recorded as a missing data

point. At the end of the two weeks, participants then completed the questionnaires once more.

## Results

### Questionnaire measures

Paired t-tests with Bonferroni correction indicated that there was a significant increase in the TMB total score following the estimation period ( $t = -4.38$ ,  $p < .01$ ). There were no changes in any of the other measures (see Tables 1 and 2).

	Before	After
SNS Stress	4.33 (1.83)	3.92 (1.73)
SNS Satisfaction	4.42 (1.62)	4.42 (1.62)

Table 1. Means (and standard deviations) for ratings before and after the estimation period

	Before	After
PSSS	29.0 (6.88)	26.83 (8.78)
TMB Goals	34.8 (4.09)	36.8 (4.84)
TMB Mechanics	32.9 (6.02)	33.9 (6.19)
TMB Scheduling	26.4 (3.22)	26.8 (2.49)
TMB Perceived control of time	16.2 (3.28)	15.67 (3.06)

Table 2. Means (and standard deviations) of questionnaire measures before and after the estimation period

### SNS use

Actual SNS usage did not change over the 10 day period (Figure 2). The amount of time spent using SNS ranged from 0 to 4993 seconds (83 minutes 13 seconds), and the average amount of time spent on them over the 10 days was 1272 seconds (21 minutes 12 seconds).

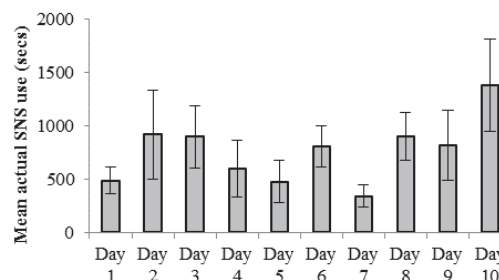


Figure 2. Mean SNS use across the duration of the investigation

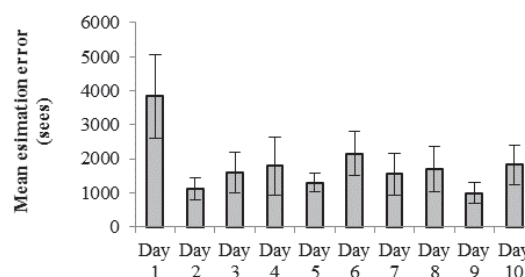


Figure 3. Mean SNS use estimation error across the duration of the investigation



However, estimation accuracy (Figure 3) significantly improved ( $F_{9,99}=2.07, p<.05$ )

### **RescueTime**

The mean number of days RescueTime was accessed during the 10 days was 4.13 ( $S.D.=3.34$ ), ranging from 0 to 10, with the average visit lasting for 78.91 seconds. The overall average amount of time spent on RescueTime per day was 38.85 seconds ( $S.D.=52.51$ ).

The difference between the TMB score before and after the estimation period was positively correlated with the amount of time spent on RescueTime ( $r=.659, p<.05$ ) and the frequency with which it was accessed ( $r=.787, p<.01$ ). Both the amount of time and the frequency were also positively correlated with the TMB Mechanics subscale ( $r=.729, p<.01$  and  $r=.668, p<.05$ , respectively). However, there were no differences in any of the questionnaire scores when comparing those who did engage with software (identified by accessing it 3 or more times during the experimental period) with those who did not.

### **Discussion**

This study aimed to investigate the effect of using RescueTime, a PI productivity tool, and to explore whether it could be used in conjunction with daily prospective estimations of SNS usage. In particular, we hypothesised that getting participants to think about their future behaviour might encourage goal setting and lead to changes in their SNS use or increased awareness and acceptance of current SNS use. However, there were no significant changes in the participants' SNS behaviour, indicating that prospective estimations may not encourage goal setting, or if they do, this may not result in behaviour change. There is evidence that participants' awareness of their behaviour increased as their prospective estimations were originally highly inaccurate but improved over the course of the study. This accompanied an increase in perceived time management behaviours, but contrary to predictions, there were no differences in perceived stress before and after the estimation period. These findings are partly in keeping with existing literature, which suggests that individuals overestimate their SNS usage [10] and that tasks encouraging greater reflection on SNS use can improve these estimations [16]. However, these results also suggest that improving estimations in itself does not guarantee a reduction in stress.

Several possibilities exist as to why this study did not replicate the findings of Zhou et al [16]. The most significant difference between the present investigation and Zhou et al [16] is the kind of estimation being investigated. It is therefore possible that this discovery reflects the nature of prospective estimations, and indicates that conceptualizing SNS use in this manner is not effective in reducing stress. This is perhaps due to the approach of goal setting not being as valuable when the main aim is not behaviour change per se, but more increasing awareness of existing behaviours.

Conversely, the different results may also reflect differences in the amount of SNS use in the two groups of participants. The sample in this study consisted of

postgraduate rather than undergraduate students, and perhaps due to having greater workloads, the level of SNS use was much lower (the average daily use was 12.82 minutes, compared to 51.14 minutes in Zhou et al [16]). This could mean that participants felt that the tool was simply not needed. However, in the present study, satisfaction with SNS use was not substantially higher (4.42 out of 7, compared to Zhou et al's 4.06), suggesting this may not entirely be responsible.

One of the most interesting, and potentially most influential, differences between the present and previous investigations is the variation in the frequency of accessing RescueTime. On average, participants accessed RescueTime 4.13 times during the 10 day experimental period, compared to Zhou et al [16], in which the average number of days was 6.94. A second interesting aspect of this is *why* engagement was lower than has been previously observed. Once again, this may be due to the nature of prospective compared to retrospective estimations. The added difficulty in accessing or remembering the relevant information for prospective estimation may therefore be a substantial barrier in participants using RescueTime in this context. Although there are options within the application to set goals, as no participants in the present experiment chose to do this, it is possible that this process needs to be made easier or quicker in order to be useful.

Alternatively, it is possible that checking RescueTime once would be sufficient to realise the first estimation had been incorrect and adjust expectations accordingly, removing the need to continue to engage with RescueTime. That said, for participants who only accessed RescueTime once or twice, this tended to occur on day two or three and then towards the end on days nine or ten. This therefore suggests that even for those without substantial interest in engaging with RescueTime, it did not appear to be a case of checking once at the beginning and feeling satisfied that this information was sufficient.

It is consequently clear that despite these interesting findings, several questions remain. Without a comparison group, it is impossible to ascertain whether the different results gained from the present study and Zhou et al [16] are in fact due to the type of estimation or the other factors that distinguish the two studies. Lower overall SNS use and less frequent RescueTime access cannot be ruled out as influential in this study, and it was therefore necessary to further explore these possibilities with a second experiment.

## **STUDY 2**

### **Introduction**

The findings of Study 1 indicated that asking participants to make daily prospective estimations of their SNS use and providing them with RescueTime data could improve their awareness of SNS and potentially, improve their perception of time management skills. However, the aforementioned questions that still remained regarding the generalizability of these findings and the issues with

very low levels of RescueTime use required further investigation.

This study therefore aimed to combine the approaches of Study 1 and Zhou et al [16] and directly compare prospective and retrospective estimations, while also testing these against a control group who provided no estimations. Based on the previous findings, it was hypothesised that across both estimation conditions there would be an increase in accuracy across the experimental period. Moreover, it was predicted that retrospective estimations would be accompanied by a reduction in perceived stress, and that prospective estimations would be accompanied by an improvement in time management behaviour. The latter prediction was based on the findings of Study 1, and was therefore only a tentative hypothesis. We were also interested in whether evidence of behaviour change would be present, particularly in the prospective condition considering the aforementioned importance of goal setting in behaviour change theories.

## Method

### Participants

A total of 30 participants were recruited (18 of whom were female), aged between 18 and 25 ( $M=20.33$ ,  $S.D.=1.54$ ). All were undergraduate students who mostly used their own personal laptops and computers. They were recruited towards the end of the first term and so did not have any impending exams. Once again, the advert targeted those intrinsically motivated to monitor and take interest in their SNS data, and extrinsic motivation was provided by entry into a prize draw reliant on timely estimations.

### Materials

As in the previous study, the TMB scale ( $\alpha=.819-.855$ ) was used to measure self-perceived time management skills [13], and the PSS ( $\alpha=.892-.894$ ) measured subjective stress [4]. Participants were also asked about their attitudes towards SNS, their perception of SNS as a stressor and how satisfied they were with their current level of use. *RescueTime* was once again used to measure SNS usage.

### Procedure

The same procedure was followed as in Study 1, with the exception that instead of a daily phone-call, participants received a text message, sent at random times during working hours every week day for a total of two weeks. This was intended to keep response methods constant across participants, irrespective of whether they were able to answer the phone, and to allow for responses to be collected in situations in which a phone-call may not be possible, for instance, during lectures. Participants were

instructed that they must respond within five minutes in order to remain in the prize draw for extra Amazon vouchers.

This text asked them to estimate how much time they spent on SNS yesterday (for the retrospective group) or how much they thought they would spend on SNS tomorrow (the prospective group). This was slightly different to the previous study, in which estimations were made in relation to that day from when they had woken up. The control group did not receive any texts, but their SNS usage was still monitored during this period. The final set of questionnaires was then completed.

## Results

### Questionnaire measures

There were no significant changes in any of the questionnaire measures before or after the experimental period, and this did not differ with condition (see Tables 3 and 4)

Estimation	SNS Stress		SNS Satisfaction	
	Before	After	Before	After
Prospective	2.20 (1.398)	2.30 (1.34)	3.90 (1.29)	4.00 (1.41)
	3.00	3.40	3.80	4.00
Retrospective	(1.70)	(1.84)	(0.92)	(1.12)
	2.20	2.75	4.40	4.67
No estimation	(1.23)	(1.39)	(1.35)	(1.00)

**Table 3. Means (and standard deviations) of SNS stress and satisfaction ratings across conditions**

### SNS use

Actual SNS behavior did not change (Figure 4). This figure shows data for the ten days estimations were elicited, as well as the days they were elicited for, which included the days before in the case of the retrospective condition and the days after in the case of the prospective condition. This resulted in a total of 14 days. The number of seconds SNS was used for per day ranged from 0 to 17340 (4 hours and 49 minutes). The average amount of time spent on SNS per day was 40.66 minutes, and this was similar across the groups.

The accuracy of estimations also did not change during the course of the experiment (Figure 5). The estimation error ranged from 0 to 357.58 seconds (5 minutes and 37.58 seconds).

### RescueTime

The mean number of days RescueTime was accessed during the 13 days (including the days before after the estimation period to account for the retrospective and prospective estimations) was 2.93 ( $S.D.=12.03$ ). This ranged from 0 to 12. The average amount of time spent on RescueTime per day was 4.68 seconds ( $S.D.=12.03$ ),

	PSSS		TMB Goals		TMB Mechanics		TMB Scheduling		TMB Perceived Control of Time	
	Before	After	Before	After	Before	After	Before	After	Before	After
Prospective	27.4 (6.93)	28.5 (7.84)	32.7 (8.59)	32.2 (8.16)	35.2 (6.68)	33.2 (7.80)	26.9 (5.17)	26.2 (4.49)	12.2 (3.46)	12.6 (3.44)
	32.5 (7.26)	30.7 (5.31)	31.3 (6.58)	31.1 (4.38)	34.5 (5.52)	33.5 (6.70)	25.6 (4.38)	24.9 (2.47)	11.1 (2.96)	11.7 (3.43)
Retrospective	26.4 (7.03)	24.0 (6.16)	33.22 (6.24)	33.25 (6.67)	34.8 (6.84)	33.7 (6.21)	26.0 (4.64)	26.4 (4.84)	14.7 (2.98)	14.6 (3.16)

Note. PSSS = Perceived Stress Scale, TMB = Time Management Behavior scale.

**Table 4. Means (and standard deviations) of questionnaire measures before and after the estimation period across conditions**

which translated to an average of 12.99 seconds per visit. The mean total of RescueTime use across the 13 days was 60.9 (*S.D.*=156.43). There were no significant differences in RescueTime use across the different conditions. There were also no differences in any of the questionnaire scores when comparing those who did engage with software (identified by accessing it 3 or more times during the experimental period) with those who did not.

## Discussion

This study aimed to directly compare the impact of prospective and retrospective SNS use estimations and RescueTime use on perceived stress and time management behaviours. It was predicted that retrospective estimations would reduce perceived stress and that prospective estimations would improve time management behaviours, in line with previous research. Moreover, both estimation types were expected to result in improved estimation accuracy. However, neither estimation type was found to have any effect on the questionnaire measures, the accuracy of the estimations, nor actual SNS use, despite participants being intrinsically and extrinsically motivated.

This result was unexpected, as potentially influential factors had been kept consistent with Zhou et al [16], including the measures employed, the frequency of estimation requests and incentives for participation. Furthermore, the questions that had been raised following Study 1 had been addressed and potential issues with this study were corrected in this instance, for example the focus on undergraduate students. The discrepancies in results are therefore unclear.

One possible explanation is the fact that participants did not engage with RescueTime. The amount of time spent on RescueTime in Study 1 was substantially lower than that observed in Zhou et al [16] and yet the accuracy of participants' estimations still improved as a result. In the present study, the level of engagement with RescueTime

was much lower still. Some participants in the present study did not access RescueTime at all beyond the initial check to ensure that there were no technical issues, and the average number of days on which RescueTime was visited was only 2.93, compared to 6.94 and 4.13 times observed in Zhou et al [16] and Study 1 respectively. This is despite participants in the present study reporting similar or lower levels of satisfaction with their SNS use (4.03 out of 7, overall) compared to Zhou et al [16] (4.06) and Study 1 (4.42).

It had been assumed that the estimation task would encourage participants to become interested in their actual SNS use and determine the accuracy of their estimates by using RescueTime. The relatively infrequent use of RescueTime in Study 1 was hypothesized to be due to the lower level of SNS use in the sample, motivating our use of undergraduate, rather than postgraduate, students in the current study. However, even with similar samples who are enthusiastic users of SNS, engagement with RescueTime cannot be guaranteed. This has important implications; if completion of a task requiring the use of RescueTime cannot motivate the use of the software in purportedly interested participants, it raises the question of what might be preventing people using this PI tool.

We therefore decided to carry out a more in depth investigation into the experience of RescueTime users. We hoped this would help our interpretation of the first two studies, and more broadly, establish whether design or usability issues may prevent the information provided being useful.

## STUDY 3

### Introduction

The findings of Study 1 and Study 2 indicated that asking participants to estimate their daily SNS usage was not sufficient to ensure participants actually accessed RescueTime to a substantial degree. This was the case across contact methods (phone or text) and student groups

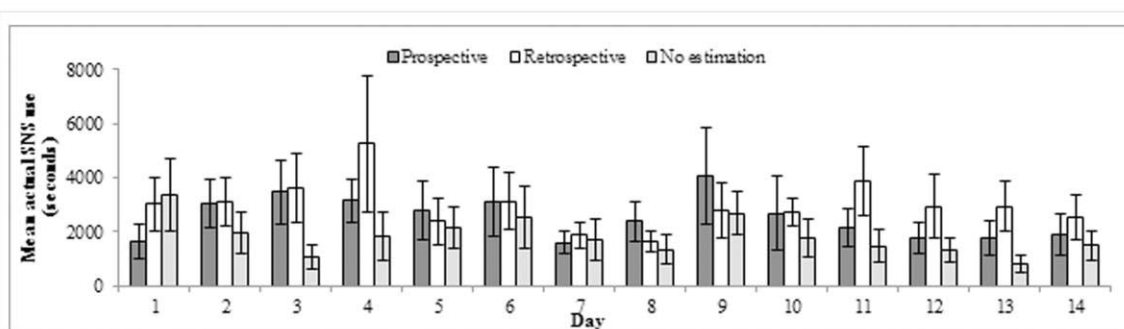


Figure 4. Actual SNS use in each condition across the full experiment period.

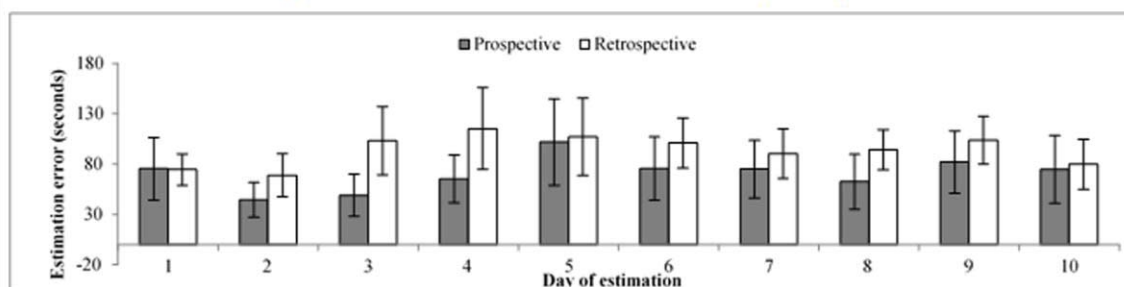


Figure 5. Estimation error in the estimation conditions across the 10 day estimation period



(postgraduates and undergraduates). The data collected as part of these studies alone could not provide information on the underlying reasons for the apparent lack of interest in using RescueTime. This is not only an issue with the present line of research but is widespread in the area of PI tools and therefore has broader implications for the use of PI tools in real life situations; there can be a lack of engagement with PI tools [11], even when there is intrinsic motivation to use information that could be provided by the tools. Consequently it was important to supplement these studies with a more in depth exploration of how RescueTime is used, interpreted and applied. This is in order to both understand the results of the previous studies and also gain insight into what could be done to improve the experience of using RescueTime.

We conducted an interview study, focusing on those who had installed RescueTime for the purpose of the study or previous studies and longstanding users of the application. The interviews aimed to identify barriers to engagement and explore if and how RescueTime leads to reflection and greater productivity.

## Method

### Participants

A total of seven participants were recruited (four of whom were female), aged 25-40. All were in positions that required working on computers and juggling multiple tasks, and all worked within a university setting. This included academics, researchers, professional services staff and PhD students. Three participants had previously installed RescueTime (one of whom as part of a previous experiment) and had been using the software for between three months and a year. Four others responded to an advertisement which called for people who had an interest in monitoring their productivity.

### Procedure

Participants were asked to install RescueTime on their devices for a minimum of two weeks prior to the interviews. However, those that had previously used RescueTime had this installed for a longer period, some for up to a year. Some of these therefore used the free version of the software. The semi-structured interviews took approximately 30-60 minutes.

## Results and Discussion

Despite the participant group consisting of both long term and new users of RescueTime, no substantial differences were found between these groups in the data.

A thematic analysis was used to analyse the interviews. Participants reported a lack engagement with the data provided by the tool. This corresponds strongly with the finding that failure to maintain interest arose from ineffective feedback outlined in Snow et al [14]. We found that participants did find some utility in using RescueTime. However, there were four barriers that inhibited engagement. In particular, the way this PI tool presents data lacks: salience; contextual information; credibility; and action advice. As a result the tool fails to initiate behaviour change.

### Utility

Participants felt that RescueTime failed to provide them with new information about their behaviour reaffirmed what they already knew: *"It really did just confirm what I already knew. I'm very productive in the morning and then have a slump in the middle of the day and it picks up again towards the end, which I kind of knew already."* – P5

This seemed to mostly be the case for information relating to periods of productivity or the specific hours worked: *"It showed me that I mainly work on Monday, Tuesday and Wednesday and then sometimes in the evening during the rest of the week... but I knew that!"* – P4

However, when it came to more detailed information on tasks that may not be as readily monitored, participants did gain new insights. This was particularly the case for e-mail use: *"I realized that I spent less time on email than I thought. Frequent short glances feels like it adds up to a lot of time when it actually doesn't."* – P6

### Barrier 1: Data Salience

As in Study 1 and Study 2, participants admitted to very rarely accessing the website to view their data, and instead would rely on their weekly summary e-mail to provide them with an overview of their digital behaviours: *"That is the thing that reminds me I have RT on my computer and then I might go and look at my dashboard and have a look at my report and think 'hm, that's interesting'"* *"I pretty much forget I have RT installed"* – P4

This appeared to be partly due to the ease with which this information was accessed, and not necessarily because they intended to seek out their data: *"I looked at it because it came to me and I didn't have to proactively go looking for RT data. It was just there."* – P7

One aspect that prevents engagement is not having immediate access to the data, and having to stop any activity currently being performed in order to seek this information out on the website. This is similar to the findings of Snow et al [14], in which energy meters were found to encourage greater engagement when the display was readily visible rather than when effort was required to check the information. This indicates that one way engagement could be improved is to make the data more readily available, and not only accessible through the website.

### Barrier 2: Contextual information

The lack of context provided by the tool was also problematic; participants found it difficult to gain a true understanding of their work patterns from the data without contextual information relating to factors such as deadlines or working hours: *"The numbers themselves don't really make much sense. You need activity of what's going on side by side with it, like tasks and deadlines etc."* – P6

This also applied to goals, as participants felt that aiming for specific amounts of times on applications rather than the completion of tasks was not helpful in motivating behaviour: *"You don't go back home and say, 'I should*

*do one more hour of Microsoft Word', you go back and say, 'I should finish that job by tomorrow.' It may take half the time, it may take twice – you just don't know.*" – P7

There was also a need for a broader context in terms of how others are working. In particular, there was a limited understanding of what the productivity pulse actually meant, particularly in relation to what a "good" score is: *"I'm not even sure what a productivity score really is. It doesn't make any sense to me, I'm not sure what I should be aiming for."* – P6

*"The most interesting would be seeing other people with the same job as me in [the university] to see what their data looks like. Just to see if they do it better."* – P5

The lack of opportunities for participants to reflect on their data or access to some kind of social context for this information meant that they were not sure whether they should be happy with their score or be striving to improve it. This is in line with a number of behaviour change theories, which argue that social norms and social influence are useful for creating the impetus to alter behaviour [12].

#### **Barrier 3: Data Credibility**

Participants tended to mistrust the data or believe that it was not a true representation of their actual activities: *"It's assuming that you're doing all of your work at one computer and that's all you do, but that's not the case."* – P7

Even when completing tasks on the computer, it is possible that this issue may also be present when tasks involve multiple applications. For instance, for situations in which dealing with e-mails involves updating spreadsheets, RescueTime would categorise the time spent on the e-mail application as separate to the time spent on the spreadsheet, despite the two applications being used for the same purpose. This lack of trust may be a further barrier to engagement, as viewing data that indicates much lower productivity than what the participant believes to be true can be demotivating: *"The numbers feel depressingly low. You'll be working hard all week and then get really low numbers because it's not logging meetings and travel time."* – P6

Even when faced with a score they believed could be lower than ideal, the lack of trust in the validity of the data meant that they were unlikely to use this as motivation to change their behaviour.

#### **Barrier 4: Action advice**

One possible barrier to RescueTime being useful in encouraging behaviour change is that even when new insights were gained from the data provided, participants were not sure what to do with this information: *"Even though it's good at telling you whether you are completing your time goal, it's not quite allowing me to reflect on why I'm not meeting that goal."* – P2

This seemed to be at least partly due to a lack of certainty in what kind of goals to be setting: *"What should I be spending my time doing?"* – P4

The difficulty in interpreting this information led to participants not actively accessing their data, instead relying on the weekly summary e-mail: *"That is the thing that reminds me I have RT on my computer and then I might go and look at my dashboard and have a look at my report and think 'hm, that's interesting'. I pretty much forget I have RT installed"* – P4

#### **Failure to promote behaviour change**

For both confirmations of existing behaviour and new insights, participants used this information more to accept their current behaviour than to motivate efforts to change. This could be described as an "acceptance epiphany" [6], which is when an individual becomes aware of their behaviour but is satisfied with it and therefore not motivated to try to change it. This appeared to apply to both productive and unproductive activities: *"It made me accept that the unproductive period just is what it is."* – P5.

Many participants consequently did not feel the need to use RescueTime's goal feature in order to set themselves targets. Even when this feature was utilised, it did not seem to promote actual changes in behaviour but once again, led to acceptance epiphanies. This was irrespective of how achievable the goals themselves were: *"From using RescueTime I discovered that I actually tended to spend less than an hour a day [on e-mails]. I considered changing the goal to spending less than 1 hour on emails, but then thought I'll just keep it at 2 and then I'll always reach that goal."* – P2

This consequently highlights how the barriers identified result in a lack of behaviour change.

#### **Limitations**

This study did, however, have certain limitations. For instance, some of the participants had only used RescueTime for a two week period, which may not have been adequate to assess its usefulness. Participant 3, for instance, had reported to be interested in the points during the day in which he got distracted, but that several more weeks of data would have been useful in trying to establish these patterns. Moreover, one of the participants only had RescueTime installed due to their involvement in previous studies the research group had conducted. Although this is unlikely to have impacted their view of RescueTime's usefulness, the motivation for engaging with the software in the first place is markedly different to those seeking it out independently.

However, despite these issues, the data confirm that those interviewed were not readily able to perceive what their data meant, nor were they able to instantly apply this to their behaviour. Although this is interesting in its own right, this is particularly relevant to the findings of Study 1 and Study 2, in which a lack of engagement was a substantial problem. Addressing these factors in the design of the tool, how the information is displayed and what accompanying advice is provided may assist developers in creating PI tools that facilitate engagement.



## GENERAL DISCUSSION

The series of studies outlined in this paper have demonstrated that engagement with RescueTime across a range of studies is low. Despite both intrinsic and extrinsic motivation to engage with the tool in order to monitor digital behaviours, specifically time spent on SNS, participants experienced a number of barriers to engagement. This was evident in the first two studies in which participants spent less time using the PI tool and checked it much less frequently than in previous investigations, despite the same tasks, incentives and similar levels of satisfaction with their current SNS use. Especially in the case of Study 2, this prevented any benefits of using RescueTime being evident, as SNS usage, perceived stress, perception of time management behaviours and estimation accuracy did not show any change during the course of the experiment. Study 3 was able to explore this further and following interviews of new, current and long standing users, identified four barriers to engagement when discussing their use of RescueTime. The way this PI tool presents data lacks: salience; contextual information; credibility; and action advice. Taken together, these findings indicate that there can be substantial individual differences in the level of engagement with PI systems, beyond those that would be immediately obvious (for instance, age or student status, which were comparable across the studies). Moreover, simply providing participants with a goal setting task in the hope that this will be sufficient motivation to engage with PI systems may not guarantee they will actually be used. This appears to be true regardless of intrinsic and extrinsic motivation to change and observe SNS use. This is due to factors within the individual, but also the design of the tool itself, and how it presents the data.

Our work was able to make a number of recommendations in relation to the design of PI tools, and how they can be adapted to improve engagement. Study 3 identified that one of the main issues with RescueTime is a lack of direction; information is presented, often without explanation or context, and no advice, points for comparison or opportunities for reflection are provided. Although many of these factors could be combatted by altering settings or actively soliciting information from other users, the onus is on the individual to seek out this information and invest more time in tailoring the tools to their preferences. However, especially in the case of productivity PI tools, the most enthusiastic adopters will be those that already consider themselves to be time poor and will not be motivated to make these kinds of changes. Not having these capabilities readily available consequently limits their usefulness and could undermine the possible benefits of using it at all.

The four barriers that inhibit engagement with RescueTime are in agreement with other research and appear to be an issue across a number of PI tools. Li, Dey and Forlizio [12] report that the reflection stage of PI tool use can be disrupted by a lack of context and an inability to interpret the data. Moreover, the next stage, action, can be prevented by a lack of direction in what to do next. While some PI tools do have this in-built capability, for those that do not, users are required to seek

this kind of advice elsewhere, for instance, from a doctor for guidance on the output of fitness related apps. However, in the case of SNS use, there is no clear alternative source of help, and it is therefore likely that RescueTime and similar PI tools leave many users unable to act on their information.

The very short amount of time spent on the RescueTime website evident in Study 1 and Study 2 suggest a lack of time or motivation for seeking out this information. Therefore, one design recommendation would be to lower the barrier for access to this information by improving the salience of the data. This approach has been supported by Snow et al's work [14], which identified that when an energy monitor was hidden away from view, this resulted in greater disengagement as participants were required to invest more effort into remembering to check the meter and to locate it. Replicating this in manner similar to the wearable activity tracker, Fitbit, may consequently be a positive influence on ease of use, as some models are able to display activity measures. This means that users are able to view basic data without having to even open a smartphone application, but can view more complex breakdowns on the website should they require them. Although wearable devices are not necessarily a solution for a PI tool such as RescueTime, this could be emulated in desktop displays, browser windows or smartphone widgets. However, care should be taken to insure this information is not made too visible to those whom the user may not wish to see it, for instance, co-workers as this may have the opposite effect; when investigating water meters in the home, Froehlich et al [8] highlighted the need for some visibility, but also the knowledge that personal information would not be made available to guests or other third parties.

## CONCLUSION

One might expect that a lack of engagement would be more of an issue with tools that aggregate measurements from a number of people, because the presented data are not solely a consequence of one individual's behaviour. However, our studies show that lack of engagement is as much of an issue for tools that display personal information, such as RescueTime, as for those that display aggregated data, such as household electricity use measured by energy monitors. Our studies therefore support the findings of Snow et al [13] but in a different domain, that of PI tools for productivity. One difference to Snow et al is that some of our participants did not even show an initial short term engagement with RescueTime, even though we provided intrinsic and extrinsic motivation for using the tool.

We have identified four barriers to engagement that are a consequence of the way RescueTime presents information and can be addressed by design interventions:

*Salience:* Productivity PI tools should have low barriers to data access; particularly in this domain, users have limited time to spend on accessing and interpreting their data, and this should therefore be as quick to access as possible. In line with the previously noted need for visibility of displays [14], while maintaining privacy [8].

*Contextual information:* The data is difficult to interpret without some indication as to what average values for similar individuals may be and such tools should therefore support social comparison. Moreover, greater support for integration of personal contexts, such as schedules and deadlines, would make the data more holistic.

*Credibility:* Users need to feel that the data is accurate and believable for any action to be considered. This could be improved through the inclusion of contextual information, so the values displayed also reflect time away from the computer. However, this work also highlights that users have inaccurate perceptions of how much time they spend on digital tasks, and the lack of belief in the data may stem from this. Consequently, the design needs to address this as well, perhaps through providing verification tasks throughout the initial tutorial to prompt users to confirm or deny the collected data so far. This would highlight the data to be accurate and build a trusting relationship with the technology.

*Action advice:* The tool needs to help users identify what further actions they may take. Study 3 highlighted that a sense of uncertainty in how to proceed was a dominant factor in the participants' use of RescueTime. This is very much in line with more general research on PI tools [12], which suggests that even basic information on alternative sources of information to support changes in behaviour may be sufficient. For example, the activity application Runkeeper includes access to external training programmes that users can follow. Therefore, in a similar vein, RescueTime could suggest achievable and appropriate goals or could implement other behaviour change techniques into the tool. This could, for example, include guidelines in reducing time spent using SNS to three 5 minutes blocks per day.

There are clear benefits from being more aware of how we spend our time including a reduction in perceived stress, an increase in the perception of time management skills and, more generally, an awareness of our most time consuming activities. However, our study shows that awareness alone is not sufficient to facilitate behaviour change and that there are a number of barriers that inhibit people from engaging with RescueTime (measured by both frequency and duration of usage). These barriers are all a consequence of how information is presented to users and we have proposed how better design could remove them. Although our study focuses on RescueTime, these findings are applicable to other PI productivity tools and more generally to systems that provide users with information to facilitate behavioural change.

## REFERENCES

1. Barata, G., Nicolau, H. and Goncalves, D. ApplInsight: what have I been doing? In Proc. International Working Conference on Advanced Visual Interfaces, 2012.
2. Block, R. A., & Zakay, D. (2006). Prospective remembering involves time estimation and memory processes. In J. Glicksohn & M. S. Myslobodsky (Eds.), *Timing the future: The case for a time-based prospective memory* (pp. 25– 49). Singapore: World Scientific.
3. Block, R.A., Hancock, P.A. and Zakay, D. How cognitive load affects duration judgments: A meta-analytic review. *Acta Psychologica*, 134, (2010), 330–343
4. Cohen, S., Kamarck, T. and Mermelstein, R. A global measure of perceived stress. *Journal of Health and Social Behavior* 24, (1983), 386-396.
5. Consumer Health Information Corporation (2011). *Motivating Patients to Use Smartphone Health Apps*. McLean, VA. Retrieved from <http://www.prweb.com/releases/2011/04/prweb5268884.htm>
6. Cox, A L., Bird, J. and Fleck, R. Digital epiphanies: How self-knowledge can change habits and our attitudes towards them. Workshop paper presented at: The 27th International British Computer Society Human Computer Interaction Conference (2013).
7. Endeavour Partners. *Inside wearables: How the science of human behaviour change offers the secret to long-term engagement* (White Paper), 2014.
8. Froehlich, J., Findlater, L., Ostergren, M., Ramanathan, S., Peterson, K., Peterson, J. et al. The design and evaluation of prototype eco-feedback displays for fixture-level water usage data. In *Proc. CHI'12*, ACM Press (2012).
9. Gabre, H. and Kumar, G. The effects of perceived stress and facebook on accounting students' academic performance. *Accounting and Finance Research* 2, (2012), 87-100.
10. Junco, R. Comparing actual and self-reported measures of Facebook use. *Computers in Human Behavior* 29, (2013), 626-631.
11. Junco, R. and Cotton, S. R. Perceived academic effects of instant messaging use. *Computers & Education* 56, (2011), 370-378.
12. Li, I., Dey, A. and Forlizzi, J. A stage-based model of personal informatics systems. In *Proc. SIGCHI Conference of Human Factors in Computing Systems 2010*, ACM (2010), 557-566.
13. Macan, T. H. Time management: test of a process model. *Journal of Applied Psychology* 79, (1994), 381–391
14. Snow, S., Buys, L., Roe, P. and Bereton, M. Curiosity to cupboard: Self reported disengagement with energy use feedback over time. In *Proc. OzCHI'13*, ACM Press (2013), 245-254.
15. Webb, T. L. and Sheeran, P. Mechanisms of implementation intention effects: The role of goal intentions, self-efficacy, and accessibility of plan components. *British Journal of Social Psychology* 47, (2008), 373-395.
16. Zhou, Y., Bird, J., Cox, A. and Brumby, D. Estimating usage can reduce the stress of social networking. Workshop paper presented at CHI'13 (2013).