Electronic Eyes: The Cognitive Effects of Webcam Surveillance

John Kieran-Mc Donagh

Master’s Thesis – 27.5 ECTS

Student Number: 1707440

Programme: Applied Cognitive Psychology MsC

Department of Social and Behavioural Sciences

Supervisor: Dr. Leendert van Maanen

2nd assessor: Dr. Chris Janssen

14/05/21

Universiteit Utrecht

Abstract

A side effect of the recent pandemic is that we now often view one another through the lens of a webcam. Physical communication has been replaced by online technologies. However, relatively little is known about the cognitive, behavioural and psychological effects that they may induce (Claypoole & Szalma, 2019). In this study, we aim to look at webcam surveillance monitoring, part of a pool of literature known as electronic performance monitoring (Ravid, 2020). We hypothesize that webcams elicit similar effects to social presence, and restrict participants' attention capacity, resulting in a decrease in cognitive performance. A total of 87 participants completed an online recognition-memory (RM) task and Stroop task. Half of these participants were told they were being recorded, while the other half were not. Linear mixed effects models revealed that participants were not significantly slower or less accurate in the webcam condition. However, when making errors, participants in the Stroop task were significantly slower in the webcam group, and in the RM task, significantly faster at responding to trials. The bi-directional nature of the effects can be partly explained by aspects of Social Facilitation Theory (Zajonc, 1965). We conclude that when one must elicit cognitive control in the presence of surveillance technology, and errors are made, it seems that this decision-making process is negatively influenced by the presence of a webcam, and the direction of the effect is moderated by the difficulty of the task.

In modern day society there is now an unprecedented amount of surveillance, bringing about the notion of a 'Big Brother' environment where almost all facets of human behaviour can be monitored. This happens continuously, whether it is through our online cookie use, CCTV, or satellites, amongst others. The ever-increasing desire to obtain human behavioural data coincides with an increasing distrust towards technology. Naturally, society's relationship with monitoring technologies is complex. Paradoxically, monitoring can be seen to simultaneously increase and decrease feelings and perceptions of safety (Isnard & Council, 2001). These monitoring technologies are usually put in place to protect us, but may also be perceived as harmful.

In a Covid-19 environment, there has been a seismic shift in the use of technologies. Communicating, working, and educating has become commonplace on websites like Zoom and Microsoft Teams. We now often see our friends, colleagues and family through the lens of a webcam. A surprisingly small amount of literature has been conducted on the psychological effects of video or webcam technologies (Claypoole & Szalma, 2019). Nonetheless, the field suggests that the relationship technological monitoring and webcam monitoring has on human behaviour is multi-dimensional, and can be part explained through social theories like Social Facilitation Theory (SFT) (Zajonc, 1965) and Cue-Utilisation Theory (Baron, 1986). The purpose of this paper is to further explore this relationship, and to investigate the potential factors that impact our behaviour and cognitions when it comes to online monitoring, and more specifically, webcam monitoring, which falls into a category known as Electronic Performance Monitoring (EPM).

EPM is the use of technologies to observe, record, and analyse information on performance (Bhave, 2014). It is a technique that is predominantly used in organizational settings to monitor employee performance. It enables supervisors to track employees discreetly or intrusively, through emails, calls, computers, and video monitoring, with or without consent (Ajunwa, Crawford, & Schultz, 2017; Ravid, 2020). The effects of electronic monitoring on task performance are well documented in literature. Classical studies examined EPM dichotomously (present vs. not present) and generally found that the presence of EPM improved task performance, work-attitudes and productivity (Irving, Higgins, & Safayeni, 1986; Nebeker & Tatum, 1993; Cottrell, 1972). However, the field has progressed over the last 30 years from inspecting task performance on a binary level to reveal a multi-dimensional phenomenon that impacts performance on a variety of levels. These include the level of invasiveness, the perceived rationale for the study, the extent to which subjects are given information about monitoring, the temporal aspects of EPM use, and also individual differences (Ravid, 2020). Each level can be viewed as a potential contributing performative factor.

One of the more relevant factors for this particular study is the perceived rationale for monitoring. Research has shown that the rationale provided to the participant for EPM use effects the perceptions towards EPM, which ultimately alters the participants' reactions (Ravid, 2020). For example, providing clear instructions, like informing the participant that the experiment is for developmental purposes, whilst also motivating and reassuring the subject to perform well on a task will generally yield positive results and attitudes (McNall & Roch, 2009). People tend to react better when there is transparency with regards to the instrumentality of the experiment. Vagueness or bluntness can be read as a distrust or invasion of privacy, resulting in negative responses. Moreover, Hovorka et al (2002) found that strong or even weak justifications for monitoring elicited greater perceptions of injustice than no justification whatsoever. This branch of EPM, whereby researchers attempt to induce arousal due to a lack of rationale, is known as surveillance monitoring.

Surveillance monitoring can be thought of as EPM without a purpose (Ravid, 2020). More often than not, when we are under some form of surveillance, there is little to no explanation or argument provided to us as to why we are being monitored. McNall & Roch (2007) argue that individuals may feel disrespected by surveillance systems as they may feel that the surveillance may be picking up information that is unrelated to the task at hand. Surveillance EPM tries to capture the emotions and cognitions that are representative of these specific settings. These include a reduced perception of fairness (Douthitt & Aiello, 2001), a decrease in satisfaction and mood (Becker & Marique, 2014), an increased sense of privacy invasion (McNall & Roch, 2007), as well as an increase in stress (Davidson, & Henderson, 2000). Moreover, surveillance monitoring has been previously shown to have moderate negative effects on performance levels (Thompson, Sebastianelli & Murray, 2009; Ravid, 2020). The current pool of work tends to argue that negative associations exist between surveillance monitoring and behaviour. However, the behavioural outcomes mentioned consist of values, attitudes, work performance, and stress. There is little to no work conducted on the cognitive, behavioural and psychological effects that bridge these associations. Very little is known whether surveillance EPM impacts cognitions like attention, perception and memory. Understanding the effects of certain cognitive predictors may provide some much needed clarity on the relationship between surveillance monitoring and behaviour. In this study, electronic monitoring will be achieved through the use of a webcam.

Typically, electronic surveillance uses visual stimuli, such as cameras, to observe behaviour and movement (McNall, & Roch, 2007)**.** Webcam monitoring has been theorised to increase self-awareness during an online experimental task (Van Bommel, et al., 2012). Prosocial behaviour increased with the presence of a webcam, which was significantly linked with an increase in self-awareness. Webcams may create the effect that there is someone else watching, which can lead individuals to alter their behaviour positively based on the context they are in. Webcams also have the ability to impair behaviour. For example, Kozar (2016) interviewed language tutors and students found that in the first 1-2 weeks, they used their webcams at the start of lessons for socio-affective reasons but discontinued their use after 2-3 weeks. As the weeks progressed, students generally found webcams as a tiring mode of communication, reporting privacy concerns and feelings of self-consciousness. Focusing on the self and how they are presented through the webcam to others became psychologically taxing. The evidence seems to suggest that feelings of self-consciousness, distrust and self-awareness negatively affect performance (Heaton & Sigall, 1999). These findings are similar to some of the core assumptions of SFT (Zajonc, 1965), a theoretical framework that may provide a conceptual basis for the impact of webcams on behaviour.

Social facilitation refers to the enhancement or impairment of performance due to the presence of others (Tricoche et al, 2020). Originally defined by Robert Zajonc (1965), the explanation operates on the basis that task performance, in the presence of others, is moderated by task difficulty. Monitoring technologies have also been found to elicit social facilitation effects (McNall & Roch, 2009; Ravid, 2020). For an easy task, performance will be facilitated due to the social presence and for a difficult task, performance will be inhibited. Zajonc hypothesised that these effects were a result of drive theory (Geen & Gange, 1977), which assumes that if you increase an individual's level of drive and arousal, the presence of others enhances the dominant-response tendency (Huguet et al, 1999). The dominant-response is the tendency to give a correct answer for an easy task, and an incorrect answer for a difficult task, resulting in both social facilitation and inhibition respectively. The existence and recognition of SFT and the link with task performance is widely accepted in academic spheres, and there are a plethora of meta-analyses supporting its effects in a variety of disciplines (Uziel, 2007; Herman, 2015; Bond & Titus, 1983). In a meta-analysis of 241 studies across 24,000 subjects, Bond and Titus (2002) found that the presence of others increases physiological arousal only if the individual is performing a difficult task. However, they found that social presence accounted for between 0.3% to 3% variance in performance, a relatively small effect. SFT provides the most convincing argument for possible performance effects of webcams, with the direction of the effect a result of the dominant-response tendency. However, can the direction of the effect also be explained by our attention and how we allocate it?

An alternative account of the interaction effect between task difficulty and social presence was provided by Baron (1986), known as Cue-Utilization Theory. Baron reasoned that social presence is distracting, which can lead to a cognitive overload and a restriction in the capacity to focus attention. Individuals then need to focus their attention on more specific details within the task. For simple tasks, people are usually better at filtering out irrelevant information and perform better. Whereas for difficult tasks, a restricted focus can result in filtering out relevant information, resulting in performance inhibition (Guguet et al, 1999). Attention is viewed as a finite resource, where we make very quick decisions and filter out what we perceive to be relevant and irrelevant. Theoretically, this decision-making process can be influenced by social presence, and is moderated by task difficulty. For difficult tasks, the level of expertise is what is theorised to ultimately dictate performance. For, example, experts engage in more automatic processing, associated with task familiarity. Whereas non-experts are engaging in controlled processing which is both consciously and cognitively demanding (Schiffrin & Schneider, 1977). These may provide an explanation for potential performance effects. However, the utilisation of cues is a construct that is difficult to evaluate due to the idiosyncratic and nonconscious nature of cues (Wiggins, 2021). One must be sceptical, and if possible, statistically account for individual responses, random effects, and perception of cues.

Whilst there is an abundance of EPM literature investigating the behavioural effects of multiple variants of electronic monitoring, there are no specific studies investigating the cognitive or behavioural effects of webcam monitoring, especially in surveillance literature. EPM is usually operationalised through computer icons or computer monitoring and has even advanced towards including AI and robotic technology to monitor performance (Reither et al, 2012). Surprisingly, there is very little research relating to video-based performance monitoring. Claypoole & Szalma (2019) argue that video monitoring literature is vastly understudied given the prevalence of video-based technologies and called upon future research to analyse further the link between video EPM and performance. To date, there has been no research conducted that also looks at how various cognitions like attention, memory, and perception are impacted by webcam or camera surveillance. There is a clear need for current researchers to narrow the scope of the EPM to webcam surveillance monitoring, and also broadening the scope of performance to examine behaviour through a variety of different cognitions.

Upon identifying a discrepancy between what is known and what is not. It leads us to ask the question, does the presence of a webcam, in a surveillance setting, affect our cognitive performance? We aim to investigate the effect webcams have on an individual's decision-making behaviour. Cognitive performance will be assessed using both a Stroop task and a recognition-memory (RM) task. This means that the independent variables (IV's) are congruence (incongruent vs. congruent) for the Stroop task, and webcam condition (webcam vs. no webcam). The dependent variable (DV) is performance, measured by reaction time for the Stroop task, and accuracy for recognition-memory task. We predict that there will be a significant difference in cognitive performance between conditions, and we expect task performance to be worse in the webcam condition. We theorise that reaction times (RTs) will be significantly higher in the webcam condition than the non-webcam condition. We also expect a decrease in performance on the RM task, and that the presence of a webcam will impair participants' ability to recognise and recall word stimuli. This produces two separate hypotheses for the two tasks;

H_1 – Webcams in an online surveillance setting will result in participants reacting more slowly, and less accurately to trials in the Stroop task.

H_2 – Webcams in an online surveillance setting will result in a significant decrease in performance in the RM task.

The effectiveness of the experiment ultimately hinges upon the illusion that the webcam is turned on. The webcam will not once be recording throughout the duration of the experiment. We assumed that most participants did not have the technological literacy to know that the webcam light and the webcam itself are part of the same circuit. In order for a webcam to record, the light must be on. One paper explored the effectiveness of LED indicators for webcams, and showed that only 45% of subjects noticed that they were being recorded (Portnoff et al, 2015). Even if subjects are aware that the light must be on, there is also a significant probability that they will not check. Therefore, we can assume to a degree of relatively high probability, that if we convincingly tell participants that they are being recorded, then they will believe us.

Method

Participants

87 psychology students, 23 male (26.4%) and 64 female (73.6%) with an average age of 25.49 from Utrecht University participated in an online experiment that lasted approximately 20 minutes. Students signed up through the SONA recruitment system and were awarded one credit upon completion.

Materials

Participants completed both the Stroop task and the RM task on the website Pavlovia.org, the online open source equivalent of the PsychoPy software package for Python (Pierce, 2007). Both surveys were created using Qualtrics, the first survey included basic demographic questions and the second survey was an online debrief form.

Procedure. Once participants had signed up through SONA, they were then required to click on the experiment link. This link brought them to Qualtrics.com, where they initially read an information sheet, followed by a consent form, and a demographic questionnaire. Upon completion, they were automatically directed to Pavlovia.org. Participants were assigned to conditions through Qualtrics, where they were coded into specific groups and conditions based on their unique ID number. Half of the participants in Pavlovia received a disclaimer image, informing them that they were being recorded through their webcam, and that they must not cover it. The other half were simply given instructions for the RM task.

The first part of the task involved memorising a sequence of 14 randomised two and three syllable words that were presented on the computer screen. Each word was shown for 2.5s and followed by a blank screen for 0.5s. Participants were then met with instructions for the second part of the experiment (the Stroop task). The Stroop task consisted of two different conditions, congruent and incongruent. In the congruent condition, the colour words ‘RED’,’GREEN’,’BLUE’, and ‘YELLOW’ were printed in their respective ink colour. For the incongruent condition, their print colour did not match the ink colour, but was assigned one of the other three colours. For example, the word “RED” would be coloured either blue, green, or yellow. Each colour was assigned one of the keys ‘a’,’f’,’k’, or ‘l’. Participants had to respond to the print colour, whilst ignoring the semantic meaning of the word. There were a total of 80 trials, with a congruent to incongruent ratio of 2:3. The Stroop task lasted between 4-7 minutes. The final part of the experiment was the recall part of the RM Task. They were again presented with 14 words, but only half of the words were in the first part of the experiment, and half of the words were not. Participants had to indicate whether or not the words were there previously. Following the RM task, participants were automatically directed back to Qualtrics where they read a debrief form. Once they read and accepted the terms of the debrief form, they were automatically redirected back to SONA where they received one credit for their participation.

Design. For the Stroop task, a 2 x 2 x 2 repeated measures design was used. There were three categorical independent variables (IV's). Congruence was the within-subjects factor and had two levels, congruent and incongruent. The webcam condition was the between-subjects factor and also had two levels, webcam or no-webcam, and the third IV was accuracy with two levels, accurate and inaccurate. The dependent variable (DV) was reaction time (RT) measured in seconds (s). Linear mixed effects models were used in order to carry out the analysis. For the RM task, the statistical analysis used was signal detection theory (SDT)(van Maanen & Miletic, 2020), as responses could easily be categorised into hits, misses, false alarms, and correct rejections. Mixed effects models were also carried out on the RM task data. Performance can be quantified using sensitivity, a measure for the distance between signal and noise distributions (Stanislaw, & Todorov, 1999).

Exclusion Criteria. For the Stroop and RM tasks, trial responses were first graphed per participant to spot any outliers or unusual responses. Participants whose accuracy was less than 55%, or who had too high a proportion of fast (<200ms), or slow responses (>3s), were removed. This led to the exclusion of 53 participants, 61% of the original sample. Therefore, data for 34 participants were included in the final dataset for the Stroop task, and 18 for the RM task.

Results

Stroop Task

The linear mixed effects regression model was fitted using the restricted maximum likelihood (REML) estimation procedure that seeks to find the parameter values that, given the data and our choice of model, make the models predicted values as similar as possible to the observed values (Baayen, Davidson, Bates, 2008). Random effects were also included in the estimation with the colour of the word accounting for 0.46% (SD=.068) of the variance in RT, the response key accounted for 0.98% (SD = 0.099), and the differences between participants accounting for 11.95% (SD=0.35) of the variance in response times. Table 1 displays the fixed effects that were included, together with the estimates, the estimated Satterthwaite degrees of freedom, and the probability that the estimates differ from zero.

Table 1.

Fixed Effects Model with RT as the outcome variable.

Measurement	Estimate	SE	df	t	p
Intercept	0.966	0.099	24.875	9.734	<0.001***
Congruence	0.120	0.031	2647.07	3.875	<0.001***
Acc (Incorrect)	-0.023	0.099	2653.57	-0.237	0.813
Webcam	0.238	0.128	36.057	1.857	0.057
Con*Acc	0.152	0.120	2648.82	1.272	0.204
Con*Web	-.044	0.051	2647.14	-0.869	0.385
Acc*Web	0.393	0.144	2658.19	2.726	<0.01**
Con*Acc*Web	-0.287	0.172	2649.77	-1.666	0.096

N.B. The intercept for the model is the reaction time (s) for correctly answered congruent trials in the no-webcam condition.
*Significance Codes - ***<-0.001, **<-0.01*

On average, participants were 120ms faster at responding to congruent trials than incongruent trials in the webcam condition, relative to the no-webcam condition (β_{con}=.120), resulting in a significant congruence effect. Post Hoc comparisons using the Tukey HSD test indicated that in the webcam condition, there was no significant difference in mean RT's between congruent (M = 1.238, SE = .040) and incongruent trials (M = 1.308, SE = .030). However, there was a congruence effect in the no-webcam condition, where congruent trials (M = .963, SE = .240) were significantly faster (p <.001) than incongruent trials (M = 1.176, SE = .017). These effects are illustrated in Figure 1 below.

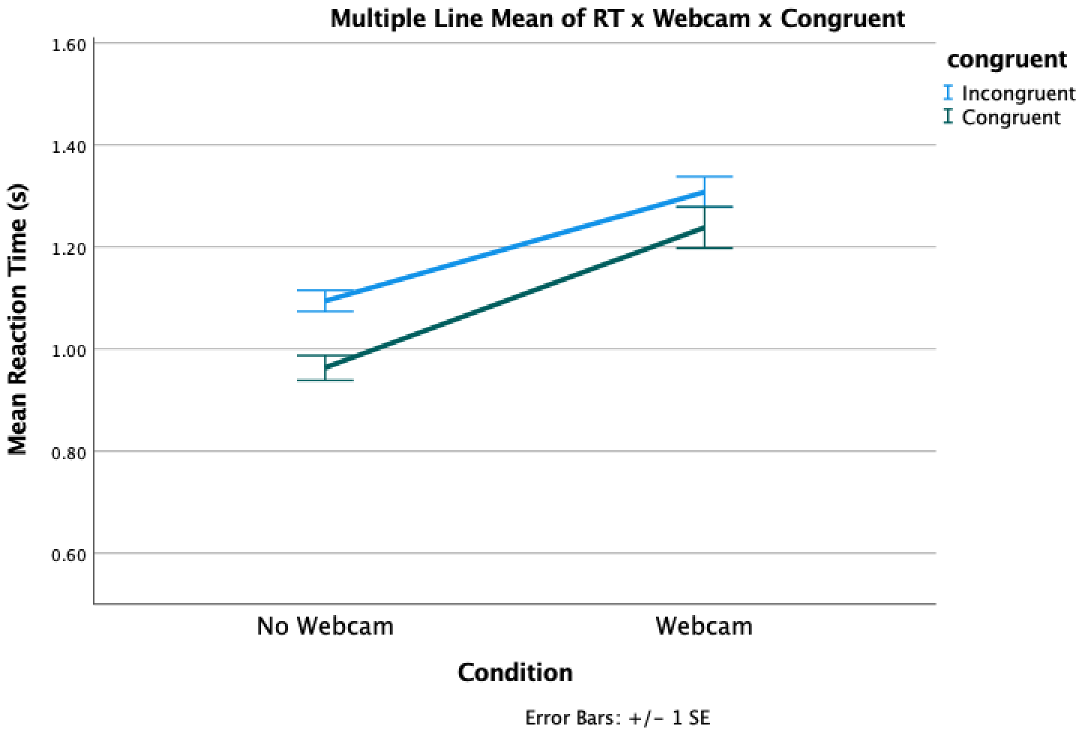

Fig 1 – A graph comparing the average reaction time for the webcam and no webcam conditions across the level of congruency.

There was a significant accuracy x webcam interaction effect for RTs. Participants were significantly slower when making mistakes in congruent trials in the webcam condition than in the no-webcam condition. There was a significantly larger difference in congruent RT's for incorrect trials as opposed to correct trials. Participants were much slower at responding to incorrect congruent trials (1.574ms) than to incorrect incongruent trials

(1.013s) in the webcam condition, as compared to the no-webcam condition, resulting in a 561ms response delay between correct and incorrect trials, which can be seen in Figure 2 below.

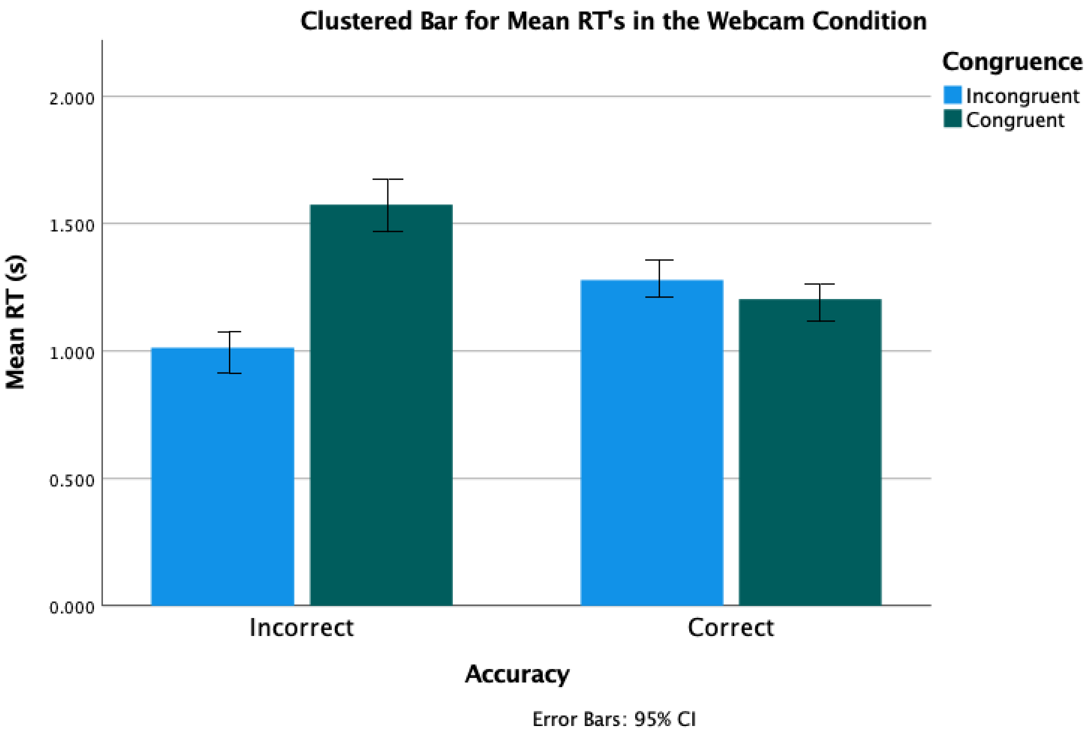

Fig 2 – A bar graph representation comparing the RTs for congruent and incongruent trials according to accuracy in the webcam condition.

A second linear mixed effects model was carried out with accuracy as the outcome variable, and with congruence and webcam as the predictor variables. Random effects revealed that the participant accounted for 1.147% of variance in accuracy, the word accounted for 0.020%, the colour of the word accounted 0.016% and the response key accounting for 0.016% variance. However, the model revealed that there was no significant difference in accuracy between congruent conditions, or between monitoring conditions, resulting in a non-significant interaction effect, which can be observed in Table 2 below.

Table 2

Fixed Effects model with Accuracy as the outcome variable.

Measurement	Estimate	Std. Error	z	p
Intercept	-3.169	0.319	-9.933	<.001 ***
Incongruent	0.329	0.202	1.630	0.103
Webcam	0.382	0.461	0.829	0.407
Con*Web	0.120	0.295	0.677	0.498

N.B. The intercept for the model is the accuracy (%) for congruent trials in the no-webcam condition.
*Significance Codes - ***<-0.001*

Recognition-Memory Task

A second linear mixed effects model was conducted inputting RT as the outcome variable and webcam and accuracy as the fixed effects. The random effects included word type and participant. Word type accounted for 0.51% (SD = .039) variance in RT and participant accounted for 5.18% (SD= .098). Table 3 displays the fixed effects that were included, together with the estimates, the estimated Satterthwaite degrees of freedom, and the probability that the estimates differ from zero.

Table 3

Fixed Effects Model with RT as the outcome variable.

Measurement	Estimate	SE	df	t	p
Intercept	1.326	0.103	44.161	12.853	<.001***
Acc (Correct)	-0.118	.095	406.015	-1.244	0.214
Webcam	-0.408	0.146	69.206	-2.795	.007**
Acc*Web	0.365	0.151	350.987	2.417	.017*

N.B. The intercept for the model is the reaction time (s) for incorrectly answered trials in the no-webcam condition.
*Significance Codes - ***<-0.001, **<-0.01, *<-.05.*

Parameter estimates revealed that there was a significant effect on RT for the webcam condition. Participants were significantly faster at responding to incorrect trials in the webcam condition ($p<.01$) in comparison to the no-webcam . This resulted in a significant accuracy x webcam interaction effect ($p <.05$), whereby trial response time was significantly predicted according to the group and whether or not a participant answered correctly. It was found that RT's were predicted by whether or not they were in the webcam condition, and also whether or not they answered the question incorrectly.

In the webcam condition ($n = 7$), participants were 80.65% ($SE = .028$) accurate and in the no-webcam condition ($n = 11$), participants were 86.90% ($SE = .041$) accurate. An independent samples t-test was carried out in order to test whether or not there was a difference in mean sensitivity (d') and criterion (c) values across webcam conditions. There was no significant difference in d' scores between the webcam ($M = 1.78$, $SE = .223$) and no-webcam ($M = 2.145$, $SE = .420$) conditions, $F(16, 9.420) = 2.788$, $t = .849$, $p = .408$, Cohen's $d = .398$. There was also no significant difference in criterion values between the webcam ($M= .065$, $SE = .271$) and no-webcam ($M=-.058$, $SE =.046$) conditions, $F(16, 8.416) = 2.572$, $t = -1.246$, $p = .231$, Cohens' $d = .583$. The difference in mean d' scores for both the webcam and no-webcam conditions can be seen in Figure 3 below.

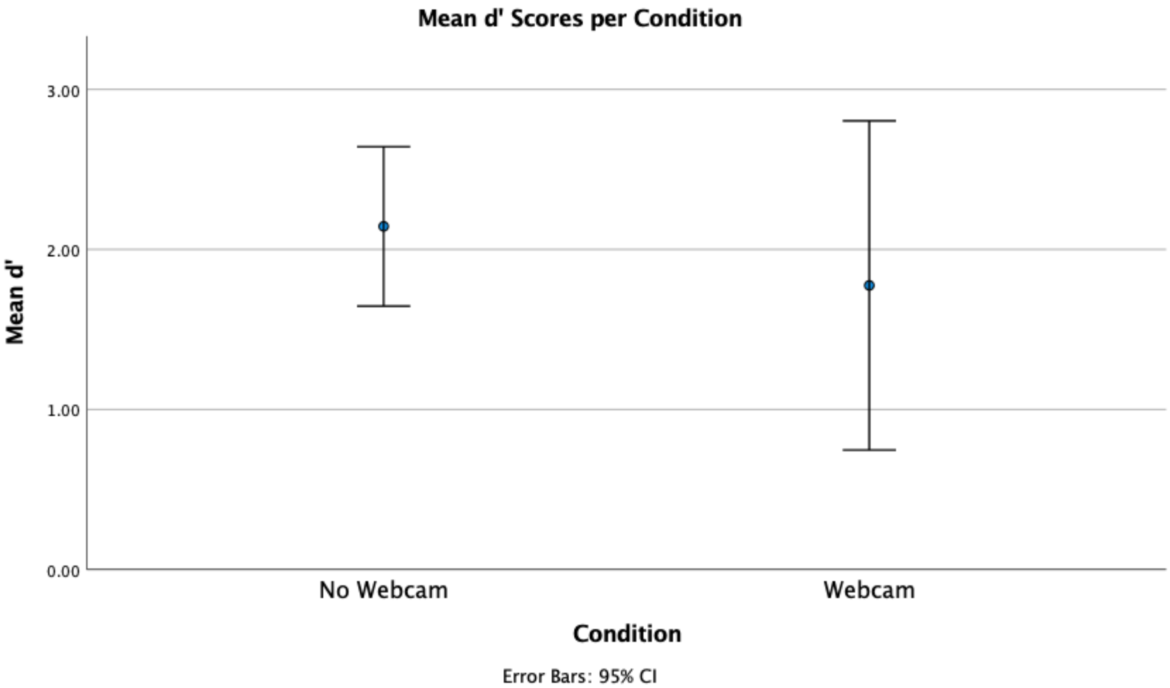

Fig 3 – A box plot representation of the mean d' scores and 95% confidence intervals per condition.

Discussion

Overall, there was not enough evidence to suggest that the presence of a webcam negatively affected Stroop performance. Linear mixed effects models revealed that participants were overall marginally slower and less accurate in the webcam condition, however, these differences between groups were found to be non-significant. Therefore, we can reject the first hypothesis (H_1), which states that webcams in an online surveillance setting will result in participants reacting more slowly, and less accurately to trials in the Stroop task. This is contrary to previous research that found video monitoring technologies negatively affected performance, and elicited social facilitation effects (Claypoole & Szalma, 2019). Although, there was a significant effect on RT when taking errors into account. Individuals were significantly slower when making congruent-trial mistakes in the webcam condition in comparison to the no webcam condition. For the RM task, there was no overall difference in performance between the webcam and no-webcam condition. Whilst participants were effective at distinguishing between noise and signal with high accuracies in both conditions (M_{web} = 80.65%, M_{noweb} = 86.9%), we cannot conclude that they were significantly less accurate in the webcam condition. There was a medium effect size between conditions for d' scores (Cohens d = .398) and a large effect size for criterion values (Cohens d = .583). However, the standard error estimates and spread of means around the population mean were too high for hypothesis confirmation. Therefore, we can reject the second hypothesis (H_2), which hypothesises that webcams in an online surveillance setting will result in a significant decrease in performance in the RM task. Nonetheless, for both the Stroop task and RM task, there was a significant association between errors and reaction time.

The Stroop task was effective at producing a RT effect in the no-webcam condition ($p<.001$), whereby RTs were significantly faster in the congruent condition than the incongruent condition. This parallels previous literature that found congruence effects in groups of participants who were not being monitored (Sharma et al, 2010; Davidson, Zacks, & Williams, 2003). Whilst mean RTs were higher for both congruent and incongruent trials in the webcam condition than in the no-webcam condition, there was no evidence for a RT congruence effect. The absence of an effect in the webcam condition may be a result of inflated RTs in the congruence condition. Congruent trials were significantly higher in the webcam condition, and it can be argued that this was brought about by an increase in errors made, resulting in a highly significant accuracy x webcam interaction effect. Errors in incorrect congruent trials were slower than correct trials, whereas they appear similar in the webcam condition. Congruent trials generally elicit automatic, high confidence, cognitive processing (Bar-Anan et al, 2007). However, this was not the case for errors in the webcam condition. Kinoshita et al (2017), used an evidence accumulation model to look at Stroop responses in the presence of a distractor and found that a distractor did not alter the baseline starting position, but it did alter the overall rate at which evidence for a colour is accumulated. We theorise that for most congruent trials, there was a threshold to make a decision, and once this threshold was passed, the webcam decreased the rate at which the participant collected evidence to make a decision. Which ultimately lead to an error. Juxtaposing incorrect trials with correctly answered trials, we found that responses were marginally faster than errors in both groups, however, the difference between these groups was smaller in the webcam condition. This is similar to the no-webcam group in the RM task, where RTs for errors were slower. However, this was not the case for errors in the webcam condition in RM.

Contrary to the Stroop task, the standout finding from the RM task was that participants in the webcam condition were significantly faster at responding to trials when making errors. Model parameters estimated a 408ms (SE = 0.146) difference between the webcam and no-webcam group when making mistakes. This resulted in a significant difference ($p<.01$). This finding is in agreement with previous social facilitation literature that found vigilance task reaction time decreases in the presence of others (Claypoole & Szalma, 2017). However, previous findings found a decrease in response time for overall task performance, and not for errors made. For recall tasks, the cognitive processes are somewhat automatic. Indicating they are familiar or not familiar with a word previously shown, and therefore can respond without the need for cognitive control. Correct answers should have a lower RT, as they are more familiar and require less deliberation (Huguet et al, 1999). According to Ratcliff & Starns' work on confidence judgements for binary choice tasks (2009), high confidence responses will be fast, and low confidence responses will be slow, leaving very little overlap between distributions. However, after conducting multiple experiments, they found a significant overlap between distributions, proposing a different diffusion model (RTCON model). Results from this study can be assimilated to these findings. When making errors in the webcam condition, high confidence responses were slow in the Stroop task, and low confidence responses were fast in the RM task, producing an overlap in noise and signal distributions. Perhaps part of this overlap can be explained by the presence of a webcam. With it ultimately effecting participants' confidence in responses for errors, and shifting their decision criteria depending on the task. This may have interfered with their ability to recognise and recall relevant cues, therefore the distinction between false alarms and hits may have become less clear. These explanations draw certain comparisons with the main assumptions of SFT.

The primary conventions of SFT implies that the direction of task performance is moderated by task difficulty. Essentially, task performance improves for simple tasks, and decreases for complex tasks (Aiello & Douthitt, 2001). Even though these assumptions are tailored towards overall task performance, they may provide an explanation for the bi-directional relationship between errors. For the Stroop task, the RT for errors was significantly slower in the webcam condition. The Stroop task is a hallmark cognitive control task, a 'gold standard' of attentional measures (MacLeod, 1992). Naturally, in incongruent trials, individuals tend to be more cautious when making a decision, and if they make a mistake, it will generally yield a higher RT (Lin et. al., 2020). Therefore, social, or webcam presence, according to SFT, will increase RTs for difficult tasks. Participants may have felt more pressure to perform better under surveillance conditions, and taken marginally more time to make a decision that should have been much faster, inhibiting their automatic cognitive processing ability. This is in accordance with previous research that found a negative association between surveillance monitoring and performance (Thompson, Sebastianelli & Murray, 2009; Ravid, 2020). This may explain the direction of the effects. For example, task performance for simple tasks, which in this case can be operationalised by speed (RT) and accuracy, increases. In the RM task, there was no evidence to suggest that overall task performance improved, as participants weren't more accurate or faster in either group. This was highlighted through the lack of significant difference in sensitivity (d') and criterion (c) scores between groups. However, they were faster when making mistakes in the webcam condition. When uncertain, the presence of the camera may have altered their decision-making threshold, perhaps giving them false confidence that they knew the word, or that the task was relatively easy, encouraging the participant to respond faster.

We assume that the observed effects on performance are a result of a combination of both SFT and Cue-Utilization Theory. Webcams elicit similar responses to human monitoring. If a person is being observed, a natural tendency is to focus more on the self, and to become more self-conscious of one’s own behaviour. We assume that webcams also increase the likelihood that an individual will adjust their behaviour, and focus more on themselves and their actions. This effect is enhanced when the observer is of a surveillance-like nature (Lyon, 2001). More focus may be given to a task, and participants may feel a sense of added pressure to perform, resulting in an increase in task performance. However, the presence of a webcam may also reduce the amount of attention given to a task, resulting in the poorer selection of specific cues based on perceived relevance (Baron, 1986). It is also worth noting that for the Stroop task, individual participants accounted for 11.95% of variance, and 5.18% of variance in the RM task. Consequently, it is important to observe that a large proportion of variance in performance can be explained by the individual. This finding is logical, as individual differences may account for privacy concerns, comfort in front of a camera, or neurotic individuals. A previous study found that extraversion was positively correlated, and neuroticism negatively correlated, with task performance (Ravid, 2020). Therefore, it is important to look at both the differences between individuals and the experimental context. Nonetheless, the idea that there is a symbiotic relationship between the two theories could form the basis for future cognitive models. Investigating attention to cues, self-presentation, decision thresholds, and evidence accumulation at specific time intervals could paint a clearer picture of the true cognitive processes elicited by webcams. It would also tell us whether or not cue-utilisation theory and SFT are valid predictors of webcam behaviour. As no previous literature has examined this link, for this paper, we assume that they are.

Inevitable, there are a number of limitations to the current study. Firstly, the lack of studies in this area means there is no reference point for which we can compare and assess the results to. It is difficult to confidently draw conclusions about the true cognitive effects of webcam surveillance, as this is the first study of its kind. Secondly, it is within the realms of possibility that participants were aware that in order for the webcam to record, the webcam light must be on. Failing to control for this, by not assessing the extent to which they noticed the webcam recording, may have been a misjudgement. This would ensure that we could remove participants who were aware that the webcam was not on when they were told it is. Gaining access to participants' webcams would increase the probability that participants were aware they are being monitored. Thirdly, there was a lack of participants for the signal detection analysis (n=18). Whilst mixed-effects models can accommodate small sample sizes due to their clustered orientation (Baayen, Davidson, & Bates, 2008), the standard error estimates for SDT models were too high. The final limitation was the absence of evidence accumulation models to pinpoint a decision threshold. This would have allowed us to see how the rate of evidence was accumulated according to teach task, and also to investigate what occurs after a threshold has passed. This may have provided a stronger explanation for the effects we observed in this study.

Future research would benefit from testing samples of different demographics. In this study, a large proportion of the sample consisted of young, educated, female psychology students. By including more participants from different ages, socio-economic backgrounds and cultures, this would give a stronger representation of the general public. There may also be some fascinating between-group effects according to age, as some studies have shown that older individuals are more sceptical of technology and online privacy than younger (Willis, 2006). Collecting physical data in a laboratory setting would also be recommended, over 58% of participants had to be removed from the sample size due to inconsistent, inaccurate and strange responses. We would expect a much smaller percentage of removed participants and trials if the experiment was conducted in person, as in a more professional environment, participants would have felt a stronger obligation to complete the experiment correctly. One could look at the effects webcams have on cognitive performance with more refined detail and under various scopes. Including different types of cognitive, perceptual or attentional tasks may also uncover a variety of different effects. For example, vigilance tasks could reveal how these effects develop over a relatively long period of time. We call upon future researchers to use evidence accumulation model in order look at how the rate of evidence accumulation is influenced before and after the threshold, specifically for errors. Manipulating rationale as an independent variable could also be beneficial for future research. One would imagine that performance could be influenced by how positive, negative, or neutral the explanation for the presence of a camera is. One could also manipulate task difficulty as an IV, to investigate social facilitation and cue utilisation theories in more detail. These findings could also be extrapolated into the effects of cameras. For example, one may find somewhat replicable results in CCTV settings which could be beneficial for businesses and government policy. We would also recommend including physiological measures for arousal, to test for both physiological and neurological responses to webcam surveillance.

The knowledge that the error reaction time is moderated by task difficulty be applied in a variety of pedagogical, EPM, social-psychological and informational-technological settings. For instance, this information may be of value in settings where the cost of a false alarm is high. Take airport security as an example. It would be worth noting that although monitoring employees may not increase the rate of errors, it may impact the response. Given the nature of the environment, errors must be notified at a fast rate, it would be worth taking into consideration the deployment of surveillance cameras. Also, given the shift from physical to online assessment due to Covid-19, this could be important knowledge for a variety of educational and academic sectors. By allowing students to turn off their camera during assessment, this may remove any social facilitation effects. These results could also be extrapolated to surveillance monitoring, or CCTV domains. A surveillance like atmosphere can produce delays in cognitive processing, which could be of use for organisations. However, based on the results of this experiment, it may not affect memory and recall, which again could be of use to educators, or people in commercial spheres.

In conclusion, webcams in a surveillance setting seem to both inhibit and enhance cognitive control in the specific context of error-making, and the direction of the effect is dependent on the task difficulty. Results from the Stroop task indicate that when mistakes are made, webcams delay cognitive control. In the RM task, the presence of a webcam seems to speed up the decision-making process when making errors in high confidence trials. Juxtaposing both tasks, the RT effects are bi-directional. We assume that these results can be explained by a combination of SFT and Cue Utilisation Theory. Whereby the electronic presence of a webcam produces similar reactions to social facilitation, improving simple task performance and impairing difficult task performance. However, these conclusions will need further exploration and interrogation, backed up by a larger sample size. There was no evidence to suggest that webcams effected our memory decision-making capability, but it did increase the decision rate for incorrect responses. We call on future researchers to replicate or build upon this study to investigate the underlying mechanisms further. Nevertheless, there has been a seismic shift from physical to online communication, and the knowledge that when making errors, webcams have the possibility to both speed up and delay our cognitive processing ability, depending on the type of task, may lay the foundation for future discoveries in this area.

References

Aiello, J. R., & Douthitt, E. A. (2001). Social facilitation from Triplett to electronic performance monitoring. *Group Dynamics: Theory, Research, and Practice*, *5*(3), 163.https://doi.org/10.1037/1089-2699.5.3.163

Ajunwa, I., Crawford, K., & Schultz, J. (2017). Limitless worker surveillance. *Calif. L. Rev.*, *105*, 735. https://dx.doi.org/10.15779/Z38BR8MF94

Augustinova, M., & Ferrand, L. (2012). The influence of mere social presence on Stroop interference: New evidence from the semantically-based Stroop task. *Journal of Experimental Social Psychology*, *48*(5), 1213-1216. https://doi.org/10.1016/j.jesp.2012.04.014

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modelling with crossed random effects for subjects and items. *Journal of memory and language*, *59*(4), 390-412. https://doi.org/10.1016/j.jml.2007.12.005

Bar-Anan, Y., Liberman, N., Trope, Y., & Algom, D. (2007). Automatic processing of psychological distance: evidence from a Stroop task. *Journal of Experimental Psychology: General*, *136*(4), 610. https://doi.org/10.1037/0096-3445.136.4.610

Becker, T. E., & Marique, G. (2014). Observer effects without demand characteristics: an inductive investigation of video monitoring and performance. *Journal of Business and Psychology*, *29*(4), 541-553. https://doi.org/10.1007/s10869-013-9338-1

Bond, C. F., & Titus, L. J. (1983). Social facilitation: A meta-analysis of 241 studies. *Psychological bulletin*, *94*(2), 265. https://doi.org/10.1037/0033-2909.94.2.265

Bhave, D. P. (2014). The invisible eye? Electronic performance monitoring and employee job performance. *Personnel psychology*, *67*(3), 605-635. https://doi.org/10.1111/peps.12046

Brunet, P. M., & Schmidt, L. A. (2007). Is shyness context specific? Relation between shyness and online self-disclosure with and without a live webcam in young adults. *Journal of Research in Personality*, *41*(4), 938-945. https://doi.org/10.1016/j.jrp.2006.09.001

Brunet, P. M., & Schmidt, L. A. (2008). Are shy adults really bolder online? It depends on the context. *CyberPsychology & Behavior*, *11*(6), 707-709. https://doi.org/10.1089/cpb.2007.0259

Byrne, J. H. (2017). *Learning and memory: a comprehensive reference*. Academic Press. https://doi.org/10.1016/b978-0-12-805159-7.05001-4

Claypoole, V. L., & Szalma, J. L. (2019). Electronic Performance Monitoring and sustained attention: Social facilitation for modern applications. *Computers in Human Behavior*, *94*, 25-34. https://doi.org/10.1016/j.chb.2019.01.001

Claypoole, V. L., & Szalma, J. L. (2017). Examining social facilitation in vigilance: a hit and a miss. *Ergonomics*, *60*(11), 1485-1499. https://doi.org/10.1080/00140139.2017.1308563

Davidson, R., & Henderson, R. (2000). Electronic performance monitoring: A laboratory investigation of the influence of monitoring and difficulty on task performance, mood state, and self-reported stress levels. *Journal of Applied Social Psychology*, *30*(5), 906-920. https://doi.org/10.1111/j.1559-1816.2000.tb02502.x

Davidson, D. J., Zacks, R. T., & Williams, C. C. (2003). Stroop interference, practice, and aging. *Aging, Neuropsychology, and Cognition*, *10*(2), 85-98. https://doi.org/10.1076/anec.10.2.85.14463

Douthitt, E. A., & Aiello, J. R. (2001). The role of participation and control in the effects of computer monitoring on fairness perceptions, task satisfaction, and performance. *Journal of Applied Psychology*, *86*(5), 867. https://doi.org/10.1037/0021-9010.86.5.867

Geen, R. G., & Gange, J. J. (1977). Drive theory of social facilitation: Twelve years of theory and research. *Psychological Bulletin*, *84*(6), 1267. https://doi.org/10.1037/0033-2909.84.6.1267

Heaton, A. W., & Sigall, H. (1991). Self-Consciousness, Self-presentation, and Performance Under Pressure: Who Chokes, and When?. *Journal of applied social psychology*, *21*(3), 175-188. https://doi.org/10.1111/j.1559-1816.1991.tb02721.x

Herman, C. P. (2015). The social facilitation of eating. A review. *Appetite*, *86*, 61-73. https://doi.org/10.1016/j.appet.2014.09.016

Hovorka-Mead, A. D., ROSS JR, W. H., Whipple, T., & Renchin, M. B. (2002). Watching the detectives: Seasonal student employee reactions to electronic monitoring with and without advance notification. *Personnel Psychology*, *55*(2), 329-362. https://doi.org/10.1111/j.1744-6570.2002.tb00113.x

Huguet, P., Galvaing, M. P., Monteil, J. M., & Dumas, F. (1999). Social presence effects in the Stroop task: further evidence for an attentional view of social facilitation. *Journal of personality and social psychology*, *77*(5), 1011. https://doi.org/10.1037/0022-3514.77.5.1011

Irving, R. H., Higgins, C. A., & Safayeni, F. R. (1986). Computerized performance monitoring systems: Use and abuse. *Communications of the ACM*, *29*(8), 794-801. https://doi.org/10.1145/6424.6430

Isnard, A., & Council, T. C. (2001, August). Can surveillance cameras be successful in preventing crime and controlling anti-social behaviours. In *Character, Impact and Prevention of Crime* *https://doi.org/10.1057/9781137316059.0014*

Kinoshita, S., de Wit, B., Aji, M., & Norris, D. (2017). Evidence accumulation in the integrated and primed Stroop tasks. *Memory & cognition*, *45*(5), 824-836. https://doi.org/10.3758/s13421-017-0701-8

Klauer, K. C., Herfordt, J., & Voss, A. (2008). Social presence effects on the Stroop task: Boundary conditions and an alternative account. *Journal of Experimental Social Psychology*, *44*(2), 469-476. https://doi.org/10.1016/j.jesp.2007.02.009

Kozar, O. (2016). Perceptions of webcam use by experienced online teachers and learners: A seeming disconnect between research and practice. *Computer Assisted Language Learning*, *29*(4), 779-789. https://doi.org/10.1080/09588221.2015.1061021

Lin, H., Saunders, B., Friese, M., Evans, N. J., & Inzlicht, M. (2020). Strong effort manipulations reduce response caution: A preregistered reinvention of the ego-depletion paradigm. *Psychological science*, *31*(5), 531-547.https://doi.org/10.1177/0956797620904990

Lynn, S. K., & Barrett, L. F. (2014). "Utilizing" signal detection theory. *Psychological science*, *25*(9), 1663-1673. https://doi.org/10.1177/0956797614541991

Lyon, D. (2001). *Surveillance society: Monitoring everyday life*. McGraw-Hill Education (UK). https://doi.org/10.1108/itp.2001.14.4.406.5

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: an integrative review. *Psychological bulletin*, *109*(2), 163. https://doi.org/10.1037/0033-2909.109.2.163

MacLeod, C. M. (1992). The Stroop task: The" gold standard" of attentional measures. *Journal of Experimental Psychology: General*, *121*(1), 12. https://doi.org/10.1037/0096-3445.121.1.12

Macmillan, N. A. (2002). Signal detection theory. *Stevens' handbook of experimental psychology*. https://doi.org/10.1002/0471214426.pas0402

Macmillan, N. A., & Creelman, C. D. (2004). *Detection theory: A user's guide*. Psychology press. https://doi.org/10.4324/9781410611147

McNall, L. A., & Roch, S. G. (2007). Effects of Electronic Monitoring Types on Perceptions of Procedural Justice, Interpersonal Justice, and Privacy 1. *Journal of Applied Social Psychology*, *37*(3), 658-682. https://doi.org/10.1111/j.1559-1816.2007.00179.x

McNall, L. A., & Roch, S. G. (2009). A social exchange model of employee reactions to electronic performance monitoring. *Human Performance*, *22*(3), 204-224. https://doi.org/10.1080/08959280902970385

Moors, A., Spruyt, A., & De Houwer, J. (2010). In search of a measure that qualifies as implicit. *Handbook of implicit social cognition: Measurement, theory, and applications*, 19-37. https://doi.org/10.1037/a0014211

Nebeker, D. M., & Tatum, B. C. (1993). The Effects of Computer Monitoring, Standards, and Rewards on Work Performance, Job Satisfaction, and Stress 1. *Journal of Applied Social Psychology*, *23*(7), 508-536. https://doi.org/10.1111/j.1559-1816.1993.tb01101.x

Peirce, J. W. (2007). PsychoPy—psychophysics software in Python. *Journal of neuroscience methods*, *162*(1-2), 8-13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Portnoff, R. S., Lee, L. N., Egelman, S., Mishra, P., Leung, D., & Wagner, D. (2015, April). Somebody's watching me? assessing the effectiveness of webcam indicator lights. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 1649-1658). https://doi.org/10.1145/2702123.2702164

Ratcliff, R., & Starns, J. J. (2009). Modeling confidence and response time in recognition memory. *Psychological review*, *116*(1), 59. https://doi.apa.org/doi/10.1037/a0014086

Ravid, D. M., Tomczak, D. L., White, J. C., & Behrend, T. S. (2020). EPM 20/20: A review, framework, and research agenda for electronic performance monitoring. *Journal of Management*, *46*(1), 100-126. https://doi.org/10.1177/0149206319869435

Riether, N., Hegel, F., Wrede, B., & Horstmann, G. (2012, March). Social facilitation with social robots?. In *2012 7th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 41-47). IEEE. https://doi.org/10.1145/2157689.2157697

Salovaara, A., Helfenstein, S., & Oulasvirta, A. (2011). Everyday appropriations of information technology: A study of creative uses of digital cameras. *Journal of the American Society for Information Science and Technology*, *62*(12), 2347-2363. https://doi.org/10.1002/asi.21643

Scarpina, F., & Tagini, S. (2017). The Stroop colour and word test. *Frontiers in psychology*, *8*, 557. https://doi.org/10.3389/fpsyg.2017.00557

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological review*, *84*(2), 127. https://doi.org/10.1037/0033-295x.84.2.127

Sharma, D., Booth, R., Brown, R., & Huguet, P. (2010). Exploring the temporal dynamics of social facilitation in the Stroop task. *Psychonomic Bulletin & Review*, *17*(1), 52-58. https://doi.org/10.3758/pbr.17.1.52

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior research methods, instruments, & computers*, *31*(1), 137-149. https://doi.org/10.3758/bf03207704

Thompson, L. F., Sebastianelli, J. D., & Murray, N. P. (2009). Monitoring Online Training Behaviours: Awareness of Electronic Surveillance Hinders E-Learners 1. *Journal of Applied Social Psychology*, *39*(9), 2191-2212. https://doi.org/10.1111/j.1559-1816.2009.00521.x

Tricoche, L., Ferrand-Verdejo, J., Pélisson, D., & Meunier, M. (2020). Peer Presence Effects on Eye Movements and Attentional Performance. *Frontiers in Behavioural Neuroscience*, *13*, 280. https://doi.org/10.3389/fnbeh.2019.00280

Uziel, L. (2007). Individual differences in the social facilitation effect: A review and meta-analysis. *Journal of Research in Personality*, *41*(3), 579-601. https://doi.org/10.1016/j.jrp.2006.06.008

van Maanen, L., & Miletić, S. (2020). The interpretation of behaviour-model correlations in unidentified cognitive models. *Psychonomic Bulletin & Review*, 1-10.https://doi.org/10.3758/s13423-020-01783-y

van Maanen, L., van Rijn, H., & Borst, J. P. (2009). Stroop and picture—word interference are two sides of the same coin. *Psychonomic bulletin & review*, *16*(6), 987-999. https://doi.org/10.3758/pbr.16.6.987

Van Bommel, M., van Prooijen, J. W., Elffers, H., & Van Lange, P. A. (2012). Be aware to care: Public self-awareness leads to a reversal of the bystander effect. *Journal of Experimental Social Psychology*, *48*(4), 926-930. https://doi.org/10.1016/j.jesp.2012.02.011

Mark W. Wiggins (2021) A behaviour-based approach to the assessment of cue utilisation: implications for situation assessment and performance, Theoretical Issues in Ergonomics Science, 22:1, 46-62, DOI: 10.1080/1463922X.2020.1758828

Willis, S. (2006). Technology and learning in current and future generations of elders. *Generations*, *30*(2), 44-48. http://anitacrawley.net/Resources/Articles/Technology%20and%20Learning%20in%20Current.pdf

Zajonc, R. B. (1965). Social facilitation. *Science*, *149*(3681), 269-274. https://doi.org/10.1126/science.149.3681.269