Habituation and learning in a steering task

Tuisku Tammi
Master's Thesis
Cognitive Science
Department of Digital Humanities
University of Helsinki
April 2019

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Tiedekunta/Osasto Fakultet/Sektion – Faculty Faculty of Arts	Laitos/Institution– Department Department of Digital Humanities

Tekijä/Författare – Author
Tuisku Tammi

Työn nimi / Arbetets titel – Title
Habituation and learning in a steering task

Oppiaine /Läroämne – Subject
Cognitive Science

Työn laji/Arbetets art – Level Master’s Thesis	Aika/Datum – Month and year April 2019	Sivumäärä/ Sidoantal – Number of pages 44+6

Tiivistelmä/Referat – Abstract

Objectives. This thesis aims to explore temporal changes in task-related physiological arousal and their connection to performance in repeated trials of a steering task. Moderate physiological arousal is believed to direct attention towards task-relevant stimuli, leading to performance improvements, while too high or low arousal is detrimental (the Yerkes-Dodson law). However, this approach does not explicitly account for changes in arousal over time. In this study, temporal changes in task-related sympathetic arousal are modelled as habituation, which has traditionally been used to describe changes in orienting responses to repeated presentations of non-target stimuli. Habituation during task performance is interpreted in terms of predictability and significance, aiming to describe changes in attentional processing during learning in an evolutionarily plausible manner. Furthermore, connections between performance and individual differences in habituation rate and spontaneous (task-unrelated) sympathetic activity are examined. Finally, habituation is compared to deviations from predicted performance.

Methods. Participants (N = 9) played a total of 40 trials of a high-speed steering task in eight sessions over a period of 2-3 weeks. Electrodermal activity during baseline and task performance was recorded in five sessions. Change in task-related skin conductance response (SCR) frequency over trials 1-5 within sessions was used to determine individual rates of habituation whereas SCR frequency during baseline indicated individual spontaneous activity. Trial-level difference scores were used to explore habituation and deviations from predicted performance (a power-law learning curve) within participants.

Results and conclusions. Task-related arousal was found to decrease with repeated trials for all participants in nearly all sessions, indicating that a habituation model was successful in capturing changes in arousal in a task situation. Furthermore, sustained task-related arousal (slow habituation) was connected to better performance both between and within participants. High spontaneous activity, on the other hand, was associated with performance decrements. Taken together, these results suggest that temporal changes in task-related arousal during learning are related to the processing of task-relevant cues and may reflect motivational states that direct selective attention, while high spontaneous activity is related to performance decrements, perhaps due to interference from task-unrelated stress.

Avainsanat – Nyckelord – Keywords
task-related arousal, learning, attention, habituation, electrodermal activity

Säilytyspaikka – Förvaringställe – Where deposited
E-Thesis

Muita tietoja – Övriga uppgifter – Additional information

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Tiedekunta/Osasto Fakultet/Sektion – Faculty Humanistinen tiedekunta	Laitos/Institution– Department Digitaalisten ihmistieteiden osasto

Tekijä/Författare – Author
Tuisku Tammi

Työn nimi / Arbetets titel – Title
Habituaatio ja oppiminen ohjaustehtävässä

Oppiaine /Läroämne – Subject
Kognitiotiede

Työn laji/Arbetets art – Level Pro gradu	Aika/Datum – Month and year Huhtikuu 2019	Sivumäärä/ Sidoantal – Number of pages 44+6

Tiivistelmä/Referat – Abstract

Tavoitteet. Tämän tutkielman tavoitteena on tarkastella tehtäväsidonnaisten fysiologisten vasteiden ajallisia muutoksia ja näiden yhteyttä suorituskykyyn toistetussa ohjaustehtävässä. Kohtalaisen fysiologisen virittymisen uskotaan suuntaavan tarkkaavaisuutta kohti tehtäväsuorituksen kannalta olennaisia ärsykkeitä ja siten parantavan suoriutumista, kun taas liian matala tai korkea virittyminen on haitallista tehtäväsuorituksen kannalta (Yerkes-Dodson-laki). Tämä lähestymistapa ei kuitenkaan selkeästi huomioi ajallisia muutoksia fysiologisessa virittymisessä. Tässä tutkielmassa tehtäväsidonnaisen sympaattisen virittymisen ajallisia muutoksia mallinnetaan habituaationa, jota on perinteisesti käytetty kuvaamaan orientaatiovasteiden muutoksia toistettujen ärsykkeiden passiivisessa asetelmassa. Tehtäväsuorituksen aikaista habituaatiota tulkitaan ennustettavuuden ja merkittävyyden näkökulmasta, tavoitteena kuvata oppimisen aikaisia tarkkaavaisuuden muutoksia evolutiivisesti uskottavalla tavalla. Lisäksi tarkastellaan tehtäväsuoriutumisen yhteyksiä yksilöllisiin eroihin habituaatiotahdissa ja spontaanissa (ei-tehtäväsidonnaisessa) sympaattisessa aktivaatiossa. Lopuksi habituaatiota verrataan poikkeamiin ennustetusta suoritustasosta.

Menetelmät. Koehenkilöt (N = 9) pelasivat nopeatempoista ohjaustehtävää yhteensä 40 kierroksen ajan kahdeksassa eri sessiossa 2-3 viikon jakson aikana. Ihon sähkönjohtavuuden vasteita (skin conductance responses, SCR) mitattiin viidessä sessiossa perustason ja tehtäväsuorituksen aikana. Yksilöllinen habituaatiotahti määritettiin tehtäväsidonnaisten SCR-frekvenssien muutoksista sessioiden aikana. Perustasomittausten SCR-frekvenssi puolestaan ilmaisi yksilöllistä spontaania aktiivisuutta. Kierrostason SCR-frekvenssien erojen avulla tutkittiin habituaation ja oppimiskäyrämallin avulla ennustetun suoriutumisen välisiä yhteyksiä.

Tulokset ja johtopäätökset. Tehtäväsidonnaisen virittymisen havaittiin vähenevän toistuvien kierrosten mittaan kaikilla koehenkilöillä lähes kaikissa sessioissa, mikä osoitti habituaatiomallin sopivan tehtäväsidonnaisen virittymisen ajallisten muutosten kuvaamiseen. Lisäksi pitkäaikainen tehtäväsidonnainen virittyminen (hidas habituaatio) yhdistyi parempaan suorituskykyyn sekä yksilö- että kierrostasolla. Korkea spontaani aktiivisuus sen sijaan yhdistyi huonompaan suoriutumiseen. Kokonaisuudessaan tulokset viittaavat siihen, että tehtäväsidonnaisen virittymisen ajalliset muutokset oppimisen aikana heijastavat tehtävään liittyvien vihjeiden prosessointia ja otaksuttavasti ilmentävät motivaatiota, joka suuntaa tarkkaavaisuutta, kun taas korkea spontaani aktiivisuus on yhteydessä huonompaan suoritukseen mahdollisesti tehtävään liittymättömän stressin häiritsevän vaikutuksen vuoksi.

Avainsanat – Nyckelord – Keywords
tehtäväsidonnainen virittyminen, oppiminen, tarkkaavaisuus, habituaatio, ihon sähkönjohtavuus

Säilytyspaikka – Förvaringställe – Where deposited
E-Thesis

Muita tietoja – Övriga uppgifter – Additional information

Acknowledgements

I want to thank Benjamin Cowley, Otto Lappi and Jussi Palomäki for persistent guidance and encouragement throughout this project. The experiments were conducted at the Traffic Research Unit at the University of Helsinki as part of a laboratory course in cognitive science. I'm grateful to fellow students in the course group for useful discussions, peer support and teamwork in implementing this complex design. Advice offered by Jami Pekkanen, Paavo Rinkkala and Samuel Tuhkanen with setting up the lab and getting the data processing going was invaluable. Finally, I'd like to thank my family and friends for support and patience.

TABLE OF CONTENTS

Appendices

1 Introduction

Arousal is believed to influence behaviour through its effect on attentional processing (Campbell, Wood, & McBride, 1997; Näätänen, 1992), making its role fundamental in learning and performance. While arousal has been widely studied in connection to performance, temporal changes in arousal and their role in learning have often been overlooked. This thesis aims to explore temporal patterns in task-related arousal, assessing the dynamic relationship of arousal and performance in repeated practice.

Perhaps the most well-known formulation of the arousal-performance relationship is the inverted U-shape curve known as the Yerkes-Dodson law (Broadhurst, 1957; Hebb, 1955; Yerkes & Dodson, 1908). The law implies that peak performance requires an optimal level of arousal while both under-arousal and over-arousal lead to performance decrements. It has been argued that with increasing arousal, less task-irrelevant (nonsignificant) cues are attended to, up to an optimal level, beyond which increasing arousal impairs the processing of task-relevant (significant) cues as well (Easterbrook, 1959; Hanoch & Vitouch, 2004). From an evolutionary perspective, arousal is believed to underlie ecologically rational behaviour by directing attention to stimuli that are perceived as significant, and preparing for action (Hanoch & Vitouch, 2004; Ursin & Eriksen, 2004).

Temporal patterns in arousal as an indicator of stimulus processing have been studied in habituation paradigms in humans and animals, in which a decrease in arousal responses to repeated stimulus presentations is interpreted as a sign of increased predictability – or decreased significance – of the stimulus (Bradley, Lang, & Cuthbert, 1993; Rankin et al., 2009; Sokolov, 1963; Thompson & Spencer, 1966). More recently, habituation models have been extended to include not only separate stimulus presentations but arousal and stress in other situations (Grissom & Bhatnagar, 2009; Eckman & Shean, 1997; Hamer, Gibson, Vuononvirta, Williams & Steptoe, 2006), such as psychosocial stress (Gerra et al., 2001) or parachute jumps (Deinzer, Kirschbaum, Gresele, & Hellhammer, 1997). This suggests that habituation paradigms might be useful in studying temporal aspects of arousal-performance relationships. Furthermore, the habituation framework can help distinguish between task-related arousal and general (task-unrelated) arousal.

This thesis proposes that temporal changes in task-related physiological arousal over repeated trials, measured by electrodermal activity (EDA), can be characterised as habituation, and that these changes are related to performance in a relatively complex visuomotor task. More specifically, maintained task-related arousal, i.e. slow habituation, is shown to be associated with enhanced performance – both between and within individuals – possibly reflecting higher motivation to succeed in the task. Furthermore, it is suggested that high general (task-unrelated) arousal, as opposed to task-related arousal, relates to performance decrements through interfering stress effects.

The structure of the thesis is built around two core themes: habituation and the arousal-performance relationship. First, the habituation framework is reviewed, considering the use of electrodermal activity as a measure of arousal, as well as the concepts of prediction and significance in attentional processing. The second theme focuses on arousal and performance, bringing together the Yerkes-Dodson law and temporal aspects of arousal. Individual differences in electrodermal reactivity in connection to performance are explored, as well as the relationship between deviations from predicted performance and task-related arousal.

The task used in this study was a high-speed steering task (Figure 1), which required visual spatial attention and planning of consecutive motor actions. Each trial lasted for approximately 2-3 minutes, and forty trials were divided into multiple sessions over a period of 2-3 weeks. Learning in the task has previously been found to fit well with a power law learning curve (Cowley et al., 2019). Moreover, deviations from predicted performance were associated with flow experience self-reports so that better-than-expected performance – rather than an objective performance level – was related to higher flow (Cowley et al., 2019). Based on these findings, it is interesting to investigate arousal and learning in a longitudinal design, which allows for a critical evaluation of the Yerkes-Dodson law, and consideration of the dynamic relationships underlying attention, learning and performance.

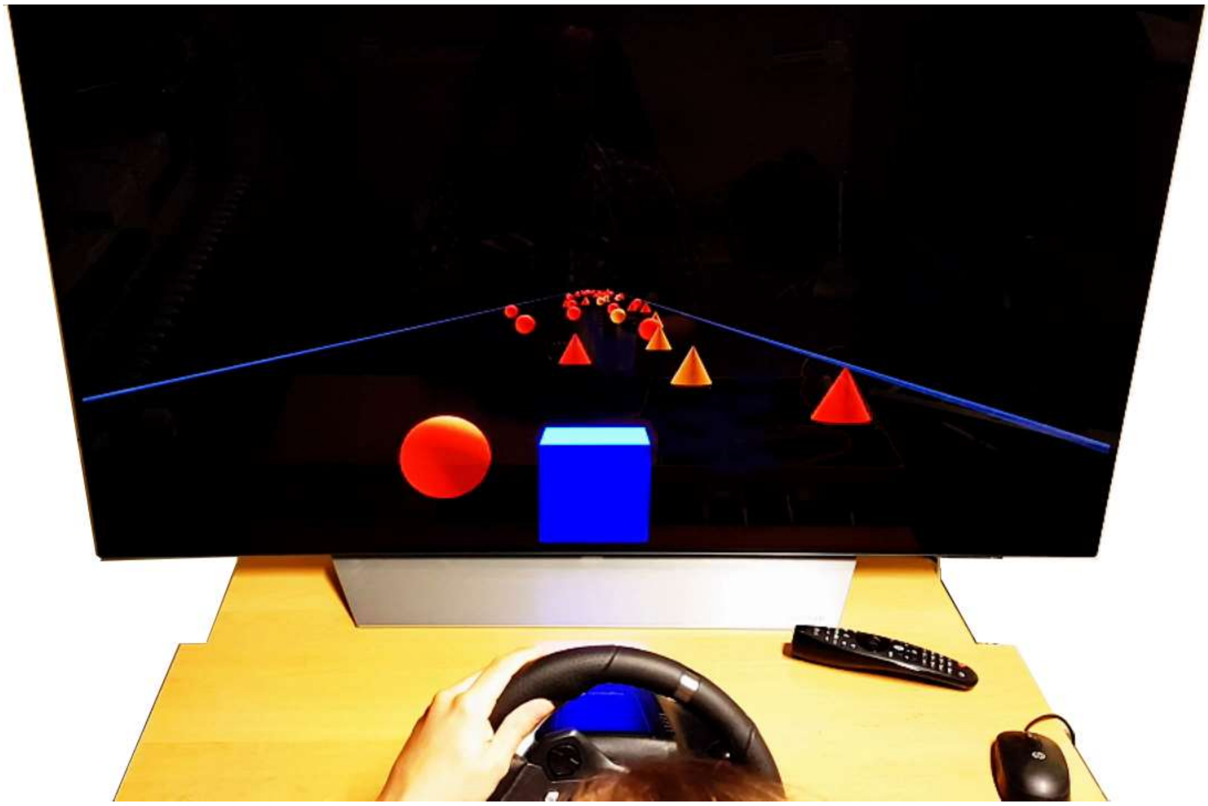


Figure 1. The steering task. The participant had to steer the forward-moving blue cube and avoid collisions to stationary obstacles.

The term *arousal* is used variably by different authors, and it can be defined in psychological terms, referring to a state of increased alertness or vigilance (Oken, Salinsky, & Elsas, 2006), or in terms of physiological responses, such as activation of the autonomic nervous system and increased noradrenaline and adrenaline levels (Bradley, Miccoli, Escrig, & Lang, 2008). Sometimes, the terms arousal and stress are used interchangeably; however, arousal typically refers to more moderate (or shorter) activation compared to stress, which is generally characterised by a feeling of unpleasantness and reduced control (Sapolsky, 2015). In this thesis, arousal is defined as an increase in physiological activation, measured by electrodermal activity. The term *task-related arousal* is used to refer to arousal during task performance. Baseline (resting) arousal, on the other hand, is referred to as *spontaneous activity*, emphasising that it is not directly related to the task situation. Arousal is used as a superordinate term that covers both task-related arousal and spontaneous activity. *Habituation* is used simply to refer to an observed decline in arousal responses over time and does not imply any interpretation of underlying processes.

1.1 Habituation

1.1.1 Electrodermal activity as a measure of arousal

In psychophysiological research, electrodermal activity (EDA) refers to changes in the electrical properties (conductance/resistance) of the skin, resulting from sweat secretion on the epidermis of the skin by eccrine glands (Dawson, Schell & Filion, 2007). These glands are innervated by cholinergic sudomotor fibres of the sympathetic nervous system. Conductance between two sites on the skin is increased (and resistance decreased) when the amount of sweat increases, and EDA can be used as a marker of changes in physiological arousal (Boucsein, 2012; Dawson et al., 2007). Electrodermal activity is affected predominantly by the sympathetic nervous system and not the parasympathetic one, which makes it a fairly reliable measure of sympathetic arousal.

EDA can be broken down into tonic (skin conductance level) and phasic (skin conductance responses) components (Boucsein, 2012). Stimulus-specific phasic skin conductance responses (SCRs), as opposed to spontaneous or nonspecific SCRs, have been extensively studied, and are principally associated with the orienting response, which directs attention towards novel or significant stimuli (Bradley, 2009; Dawson et al., 2007; Öhman, Flykt, & Esteves, 2001; Sokolov, 1963). Sokolov (1978) describes the *predictability* of the stimulus, i.e. the discrepancy between the expected and the observed stimulus, as the main constituent of the orienting response. He states that expectations are formed by ‘traces left in the nervous system’ by previous stimuli, which form a selective attentional filter (Sokolov, 1978). This view closely resembles some more recent accounts of stimulus processing (Bradley, 2009; Ursin & Eriksen, 2004), and the notion of predictability has gained a substantial role in psychophysiological research, such as event-related potentials (Sokolov, Spinks, Näätänen, & Lyytinen, 2002), as well as predictive coding theories of cognition (Clark, 2013).

1.1.2 Characteristics of habituation

Habituation refers to the decline of electrodermal responses with repeated presentations of the eliciting stimulus (Rankin et al., 2009; Sokolov, 1963; Thompson & Spencer, 1966). It has been suggested to result from growing predictability of the repeated stimulus. In the Sokolovian perspective, the orienting response acts as a comparator and habituates as prediction accuracy increases, signalling the *information* carried by the stimulus rather than its absolute intensity (Sokolov, 1963). By the same token, unexpected omission of a stimulus in a repeated sequence would cause a phasic response. In this framework, habituation can be interpreted as nonassociative learning that depends primarily on stimulus predictability but is affected by other parameters of stimulus presentation, such as stimulus strength, presentation frequency, and presence of other stimuli (Bradley, 2009; Thompson & Spencer, 1966; Rankin et al., 2009).

Besides familiarity, habituation - or its absence - can signal the significance or emotional content of a stimulus (Bradley, 2009; Dawson et al., 2007; Grissom & Bhatnagar, 2009). First, it can be argued that predictability is inherently related to significance: less predictable stimuli are also perceived as more significant because of their ‘surprise value’ (Koolhaas et al., 2011). Second, the pattern of responses to a repeated stimulus depends on its significance, and habituation to significant (task-relevant or emotionally salient) stimuli has been shown to be slower than to nonsignificant stimuli (Barry, 2004; Bradley, 2009). This is in line with the evolutionary function of arousal: it may be adaptive to habituate to other stimuli, but not significant, possibly even life-threatening stressors. This is the case even if the stimuli were highly predictable.

In their review, Grissom and Bhatnagar (2009) provide a comprehensive account for habituation in stress-related arousal and conclude that stress-related HPA axis activity displays similar habituation patterns to phasic orienting responses. In their cognitive activation theory of stress, Ursin & Eriksen (2004) frame the concept of arousal/stress responses as reactions to ‘something that is missing’, meaning a discrepancy between the expected and the observed. In this view, the authors incorporate both predictability and significance, covering individual stimuli as well as broader situations, such as a homeostatic imbalance or threat (Ursin & Eriksen, 2004). By viewing stressors as unpredictable or uncontrollable stimuli or events, as proposed by Koolhaas et al. (2011), stress habituation conforms to the information theoretic view of habituation: arousal decreases as the stressor becomes more predictable and possibly more controllable. This suggests that habituation could be used in studying task-related arousal measured over several trials of a task, even though this is not a similar stimulus presentation paradigm to the ones in traditional orienting response studies.

Detailed criteria for habituation have been presented by Thompson and Spencer (1966) and revised by Rankin et al. (2009). Grissom and Bhatnagar (2009) classified these criteria into four themes: first, habituation is seen as a decline in responses to repeated stimuli. Second, it is reversible, meaning that a response can re-occur if stimulation is withheld (spontaneous recovery). Third, it is affected by parameters such as frequency of stimulation: the more frequent the stimulation, the more rapid the habituation rate (potentiation of habituation). Fourth, habituation can progress beyond resting (baseline) levels. For the purposes of this thesis, it is useful to adhere to these four themes in investigating patterns of arousal during repeated task performance and learning.

1.2 Arousal and performance

1.2.1 The Yerkes-Dodson law

The Yerkes-Dodson law presents an inverse U-shaped relationship between arousal and performance, with optimal performance at moderate levels of arousal (Broadhurst, 1957; Hebb, 1955; Yerkes & Dodson, 1908). It is based on the notion that moderate arousal is beneficial for cognitive and physical functioning whereas amplified or prolonged arousal may have detrimental effects. This idea is supported by inverse U-shaped relationships in many neurobiological stress effects on neurocognitive function and health (McEwen & Gianaros, 2010; Sapolsky, 2015). The Yerkes-Dodson law was originally based on experiments on aversive reinforcement stimulus strength (electric shocks administered upon errors) and habit formation/discrimination learning in the dancing mouse (Yerkes & Dodson, 1908). In its current formulation, however, stimulus strength has been paralleled to arousal and habit formation to task performance.

Easterbrook's cue utilisation theory (1959), which has become a well-known interpretation of the Yerkes-Dodson law, suggests that the capacity to process cues decreases with increasing arousal and that first cues to be filtered out are task-irrelevant. Too high arousal therefore prevents the processing of task-relevant cues, whereas moderate arousal prevents processing of task-irrelevant cues while preserving task-relevant ones (Easterbrook, 1959). Peak performance occurs at optimal arousal, when all task-irrelevant cues have been filtered out, but all task-relevant cues are being attended to.

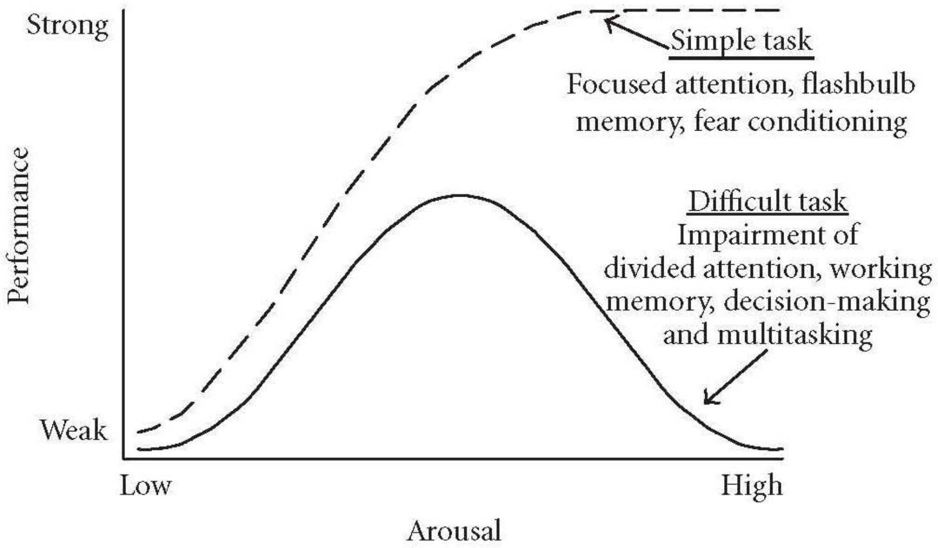


Figure 2. The arousal-performance relationship for simple and difficult tasks as described in the Yerkes-Dodson law. Adapted from Diamond et al. (2007).

Task complexity or difficulty is suggested to moderate this relationship so that performance in a difficult task is more vulnerable to the adverse effects of arousal (Diamond, Campbell, Park, Halonen, & Zoladz, 2007; Easterbrook, 1959); optimal level of arousal in easier tasks might be higher than in difficult tasks due to a more monotonic relationship between arousal and performance (Figure 2)[1]. Easterbrook (1959) argues that complex tasks require allocation of attention to a wider range of stimuli, making performance in them more susceptible to interference from high arousal than in simple tasks. Diamond et al. (2007) differentiate between tasks based on the degree of prefrontal cortex (PFC) processing needed. They argue that even though arousal enhances some aspects of performance, the more PFC processing is required, the more impaired performance will be under high arousal (i.e. inverted U-shape). This is in line with research showing a tendency from goal-directed to habitual processing under extraneous stress (McEwen & Gianaros, 2010; Plessow, Kiesel, & Kirschbaum, 2012): in stressful environments, it is energy-conserving to rely on pre-existing

---

[1] The Hebbian version of the relationship does not make a similar distinction, although it is stated that optimal arousal can be higher for simple tasks (Hebb, 1955).

associations rather than responding to, or even generating, new ones. In contrast, task performance with less PFC processing demand should display a more linear relationship to arousal (Diamond et al., 2007). This is in agreement with Easterbrook's cue utilization hypothesis and the observation that stress narrows attention according to some personal relevance estimate or utility function.

The Yerkes-Dodson law has been heavily criticised for its unitary concept of arousal (Hanoch & Vitouch, 2004, Sanders, 1983). In terms of task performance, it cannot be trivial whether arousal is related to the task at hand or something else, such as psychosocial stress related to performing the task under supervision, or an upcoming exam. From an evolutionary perspective, it is not only arousal itself but *task-relevance* of the stressor that determines how performance is affected (Hanoch & Vitouch, 2004). In fact, Easterbrook (1959) touched upon the subject in his cue-utilisation hypothesis in distinguishing significant from nonsignificant stimuli – however, the underlying assumption of the Yerkes-Dodson law seems to be that arousal is always related to the task, making task-relevant stimuli perceived as significant. However, if this is not the case – if the stressor is task-irrelevant – performance is likely to be impaired due to the allocation of attention elsewhere (Öhman, Flykt, & Esteves, 2001; Plessow et al., 2012): an upcoming exam could distract attention from the actual task.

Furthermore, as arousal is regarded as a marker of attentional processing, a distinction should be made between stimulus-driven and goal-directed attention (Corbetta & Shulman, 2002), or passive and active orienting (Bradley, 2009; Frith & Allen, 1983). This bears some resemblance to the differentiation of arousal (general energetic state) and activation (task-related change in arousal) proposed by Barry, Clarke, McCarthy, Selikowitz, and Rushby (2005). All these dichotomies are inherently connected to motivation: in fact, motivation has been described as consisting of an arousal component and a goal-directed component (Hebb, 1955; Simpson & Balsam, 2016). The effect of arousal on task performance would therefore be influenced by the source of stress and perceived relevance of task stimuli, as well as general arousal or alertness. Perceived relevance can refer to both biological relevance, such as immediate threats to survival, or motivational relevance, such as task-relevant cues (Boonstra, 2013; Bradley, 2009).

While the Yerkes-Dodson law does not explicitly consider learning effects over time, it has been suggested that learning in a task would result in a higher optimal level of arousal, similar to that of a simple task (Watters, Martin, & Schreter, 1997; see Figure 2). For example, it has been suggested that the presence of an audience (social stressor) enhances performance in well-learned tasks but impairs it in less learned tasks (Baumeister & Showers, 1986), implying a monotonous or right-shifted curve and higher optimal arousal for well-learned tasks. However, these views do not suggest what kind of mechanisms might be related to the shifts or transformations in the arousal-performance relationship.

In sum, task-related arousal may enhance learning and performance while task-irrelevant stressors may disrupt the learning process, but the Yerkes-Dodson law does not explicitly make a distinction between task-relevant and task-irrelevant arousal. At the physiological level, both could be observed as elevations in peripheral sympathetic responses. Another fundamental caveat is that the theory does not consider learning effects over time, further indicating that its explanatory power in repeated performance is limited; some authors have proposed a change in the shape of the curve as learning occurs, but these explanations remain superficial. For the theory to overcome these caveats, it should be ecologically reframed in terms of task-relevance (Hanoch & Vitouch, 2004) and include a more detailed account of learning effects.

1.2.2 Individual differences: electrodermal lability

Inter-individual variation in electrodermal activity is observed both in resting (spontaneous) and stimulus presentation conditions (Boucsein, 2012; Crider, 1993; Dawson et al., 2007).

Individuals with high frequency or amplitude of non-specific electrodermal responses, and/or slow habituation, are characterised as *electrodermal labiles*, in contrast to *stabiles,* who have low responsivity and/or high habituation rate (Schell, Dawson, & Filion, 1988). Among other things, electrodermal lability has been linked to attentional and information processing abilities: electrodermal labiles seem to perform better than stabiles in tasks that require sustained attention or fast reaction times (Dawson et al., 2007; Munro, Dawson, Schell, & Sakai, 1987; Sakai, Baker, & Dawson, 1992). This is believed to reflect their higher ability to allocate and maintain attention to task-relevant stimuli.

It should be noted that electrodermal lability can be assessed by two distinct criteria: first, non-specific responses (spontaneous lability), and second, habituation of electrodermal responses during stimulus presentation (habituation lability). These measures have historically been highly correlated (Boucsein, 2012; Schell, Dawson, & Filion, 1988) but might reflect separate phenomena. Some authors have proposed a diverging relationship between task performance and electrodermal lability depending on the criterion used: for example, both Sostek (1978) and Vossel & Rossman (1984) found decrements in vigilance performance to be associated more strongly with habituation lability than spontaneous lability. On the other hand, Crider (1993, 2008) distinguishes between types of cognitive demands imposed by tasks, suggesting that labiles actually perform worse in tasks requiring short term memory and non-distractibility while succeeding in rapid response execution and sustained attention to external stimuli. Crider (2008) proposes an effortful control hypothesis of electrodermal lability, concluding that lability, indexed by high spontaneous activity, marks a tendency for cognitive preoccupation, which impairs performance in certain types of tasks.

It is interesting to study whether habituation of task-related arousal – in contrast to habituation to non-target stimuli – is a measure of electrodermal lability that corresponds to lability measured by spontaneous (baseline) activity. High spontaneous activity might relate to better performance if it were connected to better attentional capability (Dawson et al., 2007). However, baseline arousal can also indicate situational general arousal, in contrast to a lability trait, in which case its effect could be different. Similarly, maintained task-related arousal could be a signal of motivational processes, reflecting higher perceived importance of the task, rather than trait-like stimulus-processing capabilities.

1.2.3 Deviation from predicted performance

Arousal has been suggested to signal discrepancy between the expected and the observed (Sokolov, 1963; Ursin & Eriksen, 2004), and it has been linked to performance monitoring and the occurrence of unexpected errors – *and* the non-occurrence of anticipated errors (Braem, Coenen, Bombeke, Van Bochove, & Notebaert, 2015). Learning in our steering task has previously been found to fit well with a power law learning curve (Cowley et al., 2019), indicating that performance is linearly connected to the logarithm of the number of trials (Newell & Rosenbloom, 1981). In this study, within-subject deviations from predicted learning curves are compared to changes in task-related arousal. Based on the results by Braem et al. (2015), it would be expected that larger absolute deviations from the learning curve are manifested as higher arousal responses, resulting in slower trial-level habituation. On the other hand, maintained arousal (slower habituation) could reflect motivational processes which would lead to enhanced performance.

1.3 Research questions and hypotheses

Research questions and hypotheses are divided into two themes: habituation of task-related arousal, and relationship between arousal and performance. While habituation of phasic orienting responses to repeated stimuli is a rather well-established phenomenon - although not mechanistically fully explained - habituation in broader arousal paradigms is less well studied. It is therefore necessary to first investigate whether the pattern of arousal during repeated task performance follows the criteria for habituation put forward by Thompson and Spencer (1966) and revised by Rankin et al. (2009). This approach is motivated by the review by Grissom and Bhatnagar (2009) and the criteria are divided into four themes similarly to that review. Only the criteria applicable in this context are reviewed: for example, dishabituation by another stimulus, as well as the strength of habituating stimulus, have been excluded due to indistinguishability of stressors or their intensity in our task.

Regarding performance, individual differences in physiological responsiveness (electrodermal lability) and habituation have previously been shown to contribute to performance. It is interesting whether these differences (spontaneous activity or habituation) are manifested in individual learning rates or performance level in the task. Furthermore, the relationship between task-related arousal and deviations from predicted performance is explored. The concept of prediction is extended to whole trials of continuous task performance, rather than separate stimuli.

The research questions and hypotheses are as follows:

RQ1. Habituation: Do the changes in task-related arousal during multiple trials of a steering task follow the criteria for habituation (Rankin et al., 2009)?

H1.1 SCR frequency decreases with repeated trials (1-5) within each session.

H1.2 There is spontaneous recovery in SCR frequency between sessions: the change in SCR frequency between the last and first trials of consecutive sessions is greater than zero.

H1.3 The within-session decrease in SCR frequency is amplified in later sessions (potentiation of habituation).

H1.4 SCR frequency during trials can progress below baseline.

RQ2. Arousal and performance: Are between-participant differences in electrodermal reactivity, or within-participant changes in habituation, related to task performance?

H2.1 Participants with high spontaneous activity (SCR frequency during baseline) and/or maintained arousal (slow habituation) perform better than participants with low spontaneous activity and/or fast habituation.

H2.2 Higher perceived importance is connected to high spontaneous activity and maintained arousal.

H2.3 Maintained arousal (slow habituation) between trials is connected to larger absolute deviations from predicted power-law learning curve within participants.

2 Methods

2.1 Participants

There were nine participants (6 male, 3 female) aged between 22 and 38 (M = 27, SD = 3), recruited from university mailing lists as well as personal contacts. The participants had normal or corrected-to-normal visual acuity and did not report any neurological or psychiatric disease. Two of the participants had no or very little previous gaming experience, two stated playing 1-3 hours a month, and five stated playing weekly.

Participants were given nine cultural vouchers (each worth 5 euros) for participating. In addition, two extra vouchers were promised to participants who improved performance during the study; the criteria for performance improvement were not stated explicitly, and all participants received the two extra vouchers in the end. Before the experiment, participants were told that the study is related to player experience and learning.

2.2 Design

The experiment consisted of eight sessions on eight different days, over a period of 2-3 weeks (Figure 3). In each session, participants played five trials of a simple steering task; each round lasted 2-4 minutes depending on their performance. Total playtime therefore ranged from 80 to 160 minutes. After each of the five trials, participants were shown the duration (seconds) and number of collisions, along with their own top-10 trial times, after which they filled in a short self-report questionnaire, Flow Short Scale (FSS; Engeser & Rheinberg, 2008).

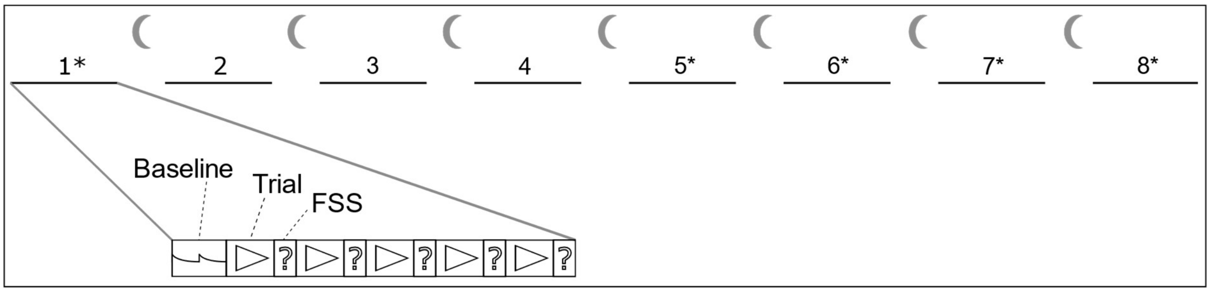


Figure 3. The game was played in eight sessions, each consisting of five trials and Flow Short Scale (FSS) self-reports. Sessions 2-4 were training sessions whereas physiological measurements during baseline and trials were taken in sessions 1 and 5-8 (marked with an asterisk).

Physiological (electrodermal activity and heart rate) and eye-tracking measurements were taken in five sessions (1 and 5-8); these lasted approximately one hour. ‘Training’ sessions 2-4 were intended to provide additional playtime for all participants to learn the task to a reasonably high level of skill, and they lasted 20-30 minutes each.

2.3 Procedure

In the first session, participants filled out their background information (driving and gaming experience, health information) and gave signed informed consent. In every session, participants filled out a form about contact lenses, restedness, and medication, caffeine, and nicotine intake. In the measurement sessions (1 and 5 to 8), physiological sensors and an eye tracking headset were then placed, and a baseline measurement of five minutes was recorded, during which the participant was asked to sit still while looking at a dark blue screen. The participant then played five trials of the task and filled the FSS after each trial. Physiological signals and eye movements were recorded during playing. In the last session, participants were debriefed and given a chance to give feedback on the game and experiment.

The sessions took place between 8 am and 7 pm at the Traffic Research Unit at the University of Helsinki, in a quiet and dimly-lit room. The experiment was carried out by two experimenters at a time, who remained behind a partition wall during playing, out of the participant's sight. The participants started the trials by pressing a button when they were ready. During the measurement, the experimenter took notes about possible confounding factors and problems within the session.

2.4 Materials

2.4.1 Steering task

The task used in the experiment was a simple steering game (CogCarSim, originally designed by Otto Lappi & Juha Vepsäläinen, Python code available at https://doi.org/10.6084/m9.figshare.7269467). In the game, the participant had to steer a forward-moving cube and try to avoid obstacles on the lane (Video available at https://doi.org/10.6084/m9.figshare.7269395). The cube's side length was 2 units, and the track was 25 units wide. The obstacles were red or yellow spheres and red cones with the same height or diameter as the cube (2 units). The track was straight, approximately 24,198 units long, bounded on both sides by walls, and included 2,000 stationary obstacles that were randomly placed. The horizontal placement of the obstacles on the track was fixed at every second unit (-11, -9, …, 9, 11), and longitudinal placement was constrained such that there were always large enough gaps between obstacles for a possible path through. The field of view angle of the virtual camera was 60 degrees horizontal and 32 degrees vertical. The camera was positioned behind the cube at 4 units height, pointing forward along the track.

Participants had no direct control of velocity along the longitudinal axis. The initial velocity of the cube was 1.6 units per step (corresponding to 96 units per second) and increased at a constant rate of 0.0012 units per step if there were no collisions. In case of collision, the screen flashed, and speed decreased by 0.102 units. There was an immunity period of a hundred steps, meaning that if two collisions occurred within that period, there was no more than one speed drop. The participant was instructed to avoid as many obstacles as possible in order to reach the finish line fast.

Data collected from the game included trial-level performance data (trial duration, number of collisions and speed drops, average velocity) and within-trial behavioural data (steering wheel position, cube and obstacle coordinates, collisions and speed drops) sampled at 60 Hz.

The game was played on Windows 10, using a Corsair Anne Bonny computer with an Intel i7 7700k processor, 55’’ screen (LG 55UF85, resolution 1920 x 1080 pixels) and an Nvidia GTX 1080 graphics card. Participants used a Logitech G920 Driving Force steering wheel with 100 percent sensitivity, 4 percent centering spring strength and 900 degrees wheel operating range in Logitech Gaming Software 8.96.88. The participants sat on a Playseat Evolution Alcantara driving seat (Playseats B.V., The Netherlands), the position of which was adjusted for each participant so that they could place their hands on the steering wheel comfortably, resulting in the distance from the eye to the screen ranging approximately between 90 and 120 cm. The seat was aligned to the horizontal midpoint of the screen.

2.4.2 Physiological signals

Electrodermal activity (EDA) and blood volume pulse (BVP, not reported here) were recorded at 128 Hz sampling rate using NeXus-10 (Mind Media B.V, Roermond-Herten, The

Netherlands) connected via Bluetooth to an Asus UX303L laptop running Debian GNU/Linux 9 OS. The data was collected using Trusas software (https://github.com/jampekka/trusas-nexus).

For EDA, silver-silver chloride (Ag-AgCl) electrodes with 0.5% saline paste were attached to the medial-plantar surface of the left foot with adhesive skin tape and gauze. The plantar site was used instead of the palmar site to minimise artefacts resulting from the use of the steering wheel, as per guidelines by Boucsein (2012). The blood volume pulse (heart rate) sensor, measuring relative change in blood flow, was attached to the left index toe of the participant. Eye tracking (not reported here) was recorded with a Pupil Labs Binocular 120 Hz headset with a custom-built headband, using Pupil Capture software to collect data on the same Asus laptop as above.

2.4.3 Flow Short Scale

A brief questionnaire, the Flow Short Scale (FSS, Engeser & Rheinberg, 2008), was used to measure *flow experience* (ten items) and *perceived importance* (three items) after each trial. Additional three items on *perceived fit of demands and skills* were asked at the end of each session. The scale was translated into Finnish and modified to the context of the task by Lehtonen, Tammi, Pölönen, Frantsi, Inkilä, and Palomäki (see Appendix 5 for the FSS in English and Finnish).

Only the perceived importance scale is reported here, as the other measures do not fall within the scope of the research questions; see Cowley et al. (2019) for a report on flow experience and performance. The perceived importance scale consisted of items ‘*Something important to me is at stake here*’, *‘I must not make any mistakes here’*, and *‘I am worried about failing’*, and participants responded on a 7-point Likert scale (1 = *not at all*, 4 = *partly*, 7 = *very much*). Cronbach’s alpha for the importance scale was .73.

2.5 Data processing

The EDA signal was inspected visually for artefacts, and data from 13 trials (5.8 %) was excluded due to missing or low-quality data. Trials with more than 10 % of missing segments were excluded. There was no systematicity in missing/excluded data, apart from five trials that were all from the same session.

Because the physiological signals and game data could not be synchronised during recording, it was done afterwards by obtaining the starting point of each trial and interpolating the physiological data to the timestamps of trial data. The starting point was determined by looking at videos recorded by the eye tracker’s front camera, defined as the time at which the text ‘Press any button to start’ disappeared. For session baselines, the first and last minute from each five-minute recording were omitted due to a large number of artefacts in those periods, resulting in three-minute baselines.

EDA signal processing was performed using MATLAB (MathWorks, Natick, MA, US) with Ledalab 3.4.9 toolbox (http://www.ledalab.de). The signal was downsampled to 10 Hz and smoothed (Gaussian with a window width of 20 samples), then decomposed into tonic and phasic components using Continuous Decomposition Analysis (CDA, see Appendix 1; Benedek & Kaernbach, 2010a). For this, two-step optimisation of the parameters of the impulse response function was done for each recording by a gradient descent method. Skin conductance responses (SCRs) were detected using a threshold of 0.05 μS, and they were further processed in R (version 3.5.1).

For each trial and baseline, the number of SCRs was scaled by the length of the recording, giving SCR frequency per minute. Because EDA can vary considerably between sessions due to, for example, differences in electrode contact from session to session, SCR frequency during baseline was subtracted from SCR frequency during trial.

2.6 Statistical methods

Most of the hypotheses were tested with linear regression or linear mixed models (i.e. hierarchical linear regression); a summary of models corresponding to each hypothesis, where applicable, is outlined in Table 1. Linear mixed models are described in more detail below. Models with group effects, i.e. H2.1-H2.2, were fitted without random effects for simplicity. All *p* values were adjusted for multiple comparisons using Bonferroni-Holm.

To look at differences between groups based on electrodermal reactivity (H2.1-H2.2), two groupings of participants were made, into low/high spontaneous activity and fast/slow habituation rate. Median SCR frequency of baseline measurements was used to measure spontaneous activity. Habituation rates were participant-level random slope coefficients obtained from the habituation model (H1.1). These groupings were also compared to background variables (gender, driving experience, gaming frequency); however, no independence tests were performed due to a small sample size.

Table 1

Linear regression models and linear mixed models, and corresponding hypotheses

Hypothesis	DV	IV(s)	Random effect(s)
H1.1	SCR frequency	Log(trial)	Participant, session
H1.3	SCR frequency	Log(trial) * session	Participant
(learning)	Log(duration)	Log(CT)	Participant
H2.1	Log(duration)	Log(CT) + spontaneous activity group	
H2.1	Log(duration)	Log(CT) + habituation group	
H2.2	Perceived importance	CT + spontaneous activity group	
H2.2	Perceived importance	CT+ habituation group	
H2.3	Habituation score	Trial + deviation score	Participant

Note. Random effects specified for linear mixed models.
DV = dependent variable, IV = independent variable , CT = cumulative trial
+ = main effects only, * = both main and interaction effects

Performance differences (H2.1) were explored with separate regression models for each grouping (baseline and habituation), due to a small sample and overlap in these groups. In a similar manner for hypothesis H2.2, separate regression models were used for each grouping in predicting perceived importance (mean of the three items in the FSS). By fitting separate models, some exploration of main effects was possible, even though sophisticated interaction analysis – or controlling for confounding factors, such as gender, driving experience, or gaming frequency – was not possible with this data.

2.6.1 Linear mixed models

Linear mixed models were fitted with the lme4 R package (Bates, Maechler, Bolker & Walker, 2015) with a maximum likelihood method. The lmerTest package (version 3.0.1; Kuznetsova, Brockhoff, & Christensen, 2017) was used to obtain *p* values; degrees of freedom were approximated with Satterthwaite's method.

For hypothesis H1.1, habituation was modelled with a linear mixed model. SCR frequency (SCR count per minute during trial - SCR count per minute during baseline) over trials 1-5 of each session was predicted with log-transformed trial number (1-5) as a fixed effect predictor. Participant and session were included as random effect predictors in a nested structure, i.e. participant and participant-session interaction. Changes in both intercept and slope of log(trial) were allowed. This was done to account for variation between participants as well as sessions, as suggested by habituation theory (Rankin et al., 2009). The model was compared to a null model without the predictor log(trial) to estimate variance explained by log(trial).

Individual learning curves were modelled using cumulative trial number (log-transformed) to predict trial duration (log-transformed) in a similar manner to Cowley et al. (2019). Instead of fitting separate models, participant was used as a random factor with both random intercept and slope. Participant intercepts indicated starting levels and slopes indicated learning rates (lower duration indicated better performance).

To have a closer look at habituation and performance (H2.3), a 'trial-level' measure of habituation (in contrast to slope over trials 1-5) was determined by calculating the difference in SCR frequency from the previous trial within the same session. Trial-level habituation scores were therefore obtained for trials 2-5 of each session, more negative scores indicating faster habituation from one trial to the next. These were compared to deviations from predicted performance, which were residuals of the learning curve model outlined above. Habituation scores were predicted with a linear mixed model by trial number and deviation score, and participant as a random effect (random intercept and slope for trial number).

3 Results

The frequency of SCRs ranged from 0 to 15 per minute (M = 3.9, SD = 4.1) during baseline, and from 0 to 23.62 per minute (M = 11.2, SD = 6.1) during trials (before baseline was removed). There were no differences in SCR frequencies related to background information (gender, age, gaming experience) of participants, or time of day.

3.1 Habituation

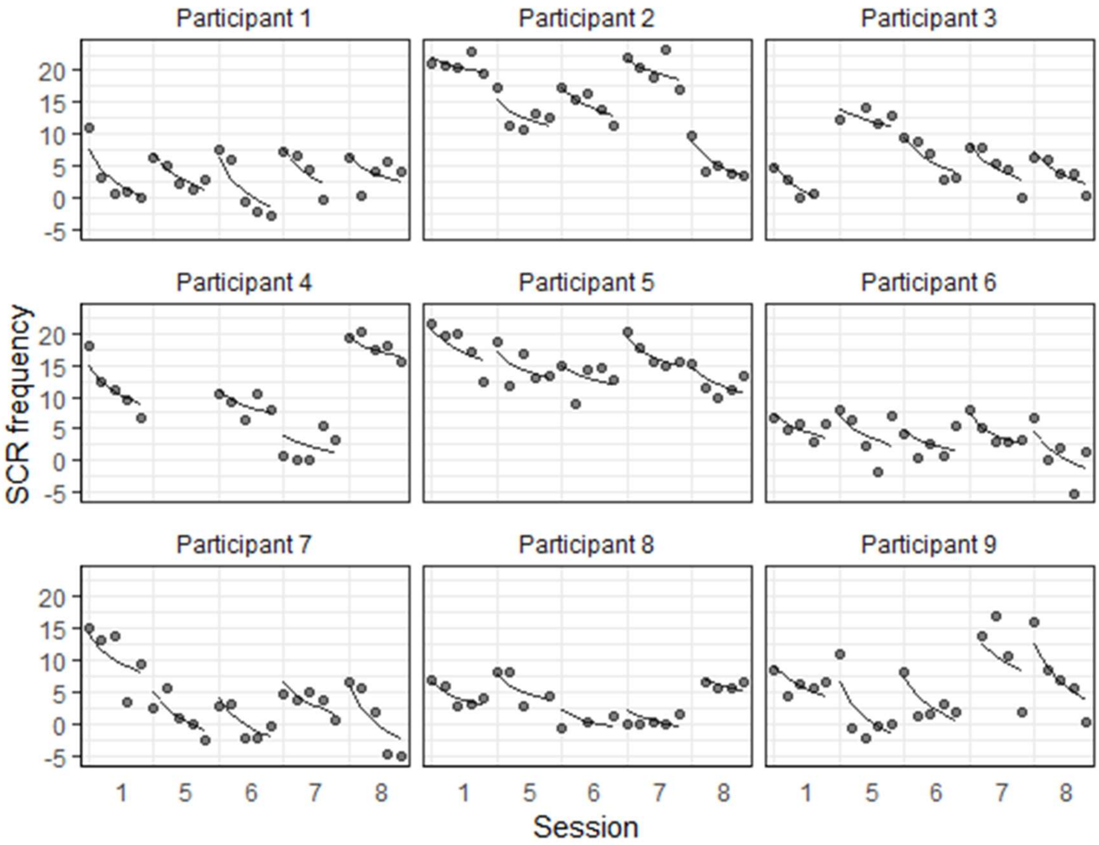


Figure 4. SCR frequencies for trials 1-5 (back-transformed from logarithmic scale) of physiological measurement sessions 1 and 5-8. Lines represent the linear mixed model fit.

Trial number affected SCR frequency negatively on a log-linear scale (*b* = -3.03, SE = 0.44, *p* < .001), meaning that physiological responses habituated with increasing trials (Figure 4). For example, from trial 1 to 2, the predicted decrease in SCR frequency was 2.1 units[2]. Negative random slopes for all nine participants and 44 sessions show that habituation occurred in every session, supporting hypothesis H1.1. Comparison to null model implied that variance explained by the two models was different ($\chi2 = 16.4$, *p* < .001) and log(trial) improved the explanatory power of the model. Akaike information criterion (AIC) for the full model was 1127 compared to 1142 of the null model.

---

[2] For an increase of X % in trial number, change in SCR frequency would be -3.03*log(1 + X/100).

There was spontaneous recovery in SCR frequencies between sessions. Change in SCR frequency between the last and first trials of consecutive sessions (5-8) was mostly positive (M = 4.08, SD = 6.31) indicating recovery of habituation between sessions (t(29) = 3.54, *p* < .001), and supporting hypothesis H1.2. However, time between sessions (1-14 days, M = 3, SD = 2.44) had no effect on the magnitude of the change (Figure 5).

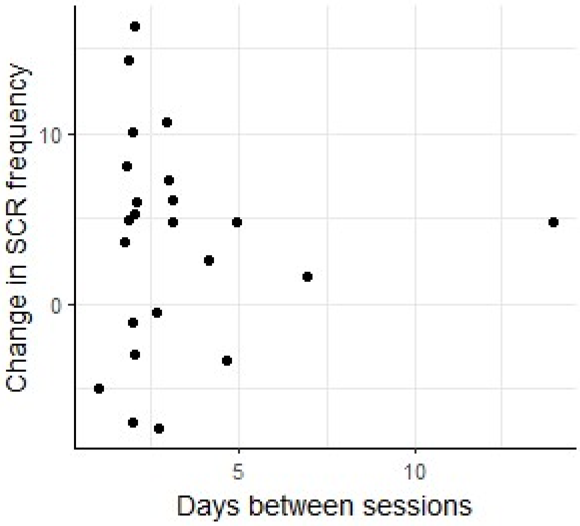


Figure 5. Change in SCR frequency between consecutive sessions 5-8 and time (in days) between sessions.

Potentiation of habituation (H1.3) could not be confirmed: when SCR frequency was predicted with linear regression by session number and log(trial), there was some indication of a main effect of session (*b* = -0.38, SE = 0.18, *p* = .08), but no interaction effect between log(trial) and session was found, i.e. there was no clear pattern in the rate of habituation within sessions. Time between sessions had no effect on SCR frequency or habituation rate.

There were 20 trials (9 %) where SCR frequency during trial was smaller than during baseline. A majority of them, 18 trials, were in sessions 5-8, and 14 were in trials 4 or 5. This supported hypothesis H1.4.

3.2 Arousal and performance

3.2.1 Learning curve

Trial duration was lower with increasing trial number (*b* = -0.07, SE = 0.006, $p < .001$), indicating that all participants improved their performance over cumulative trials (individual learning curves depicted in Appendix 3). The slopes and intercepts of learning curves were very strongly correlated (*r*(7) = -0.99, $p < .001$).

3.2.2 Electrodermal lability, perceived importance, and performance

Participants were grouped by habituation slopes into fast (n = 4) and slow (n = 5) habituators, and by baseline SCR frequency into low (n = 5) and high (n= 4) spontaneous activity (Figure 6). Most slow habituators - 4 out of 5 - had low spontaneous activity, and 3 out of 4 fast habituators had high spontaneous activity.

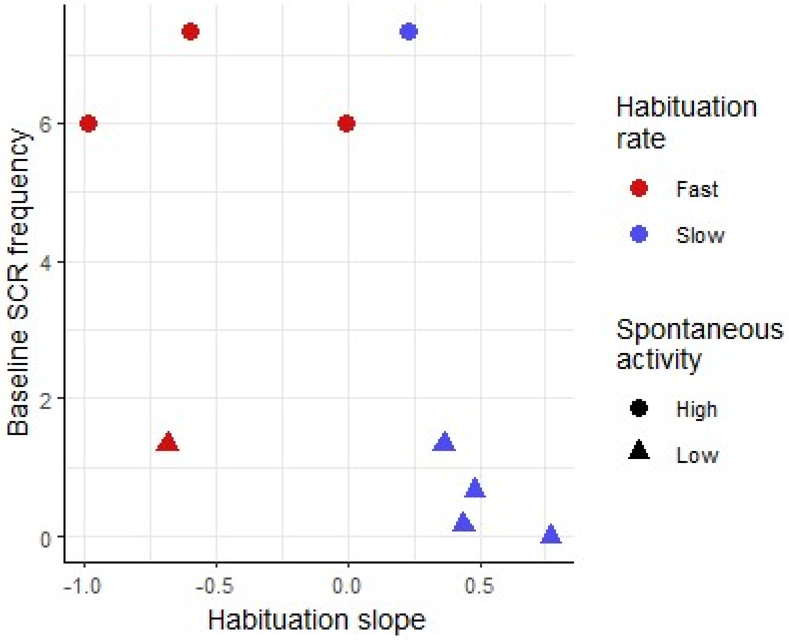


Figure 6. Groupings of participants based on spontaneous activity (baseline SCR frequency) and habituation rate (model slope coefficient).

The relationships between the habituation and spontaneous activity groups and background variables (gender, driving and gaming experience) is shown in Table 2. Participants' habituation model coefficients, baseline SCR frequencies, learning measures, and background information can be found in Appendix 2.

Table 2

Number of participants in habituation and spontaneous activity groups, cross-tabulated with background variables

	Habituation rate		Spontaneous activity		
	Fast	Slow	High	Low	Total
Gender					
Female	2	1	2	1	3
Male	2	4	2	4	6
Driving experience					
< 10,000 km	2	1	1	2	3
> 10,000 km	2	4	3	3	6
Gaming frequency					
Less than weekly	2	2	3	1	4
Weekly	2	3	1	4	5
Total	4	5	4	5	9

Having low spontaneous activity corresponded to better performance (b = -0.06, SE = 0.007, $p < .001$), when added as a predictor in a log-log model, where duration was the dependent variable and cumulative trial was the independent variable. The model explained 52 % of variance in performance. Slow habituation showed a similar association (b = -0.03, SE = 0.007, $p < .001$), with $R^2 = .43$. Figure 7 shows the learning curves for both groupings. Hypothesis H2.1 was therefore only partially supported.

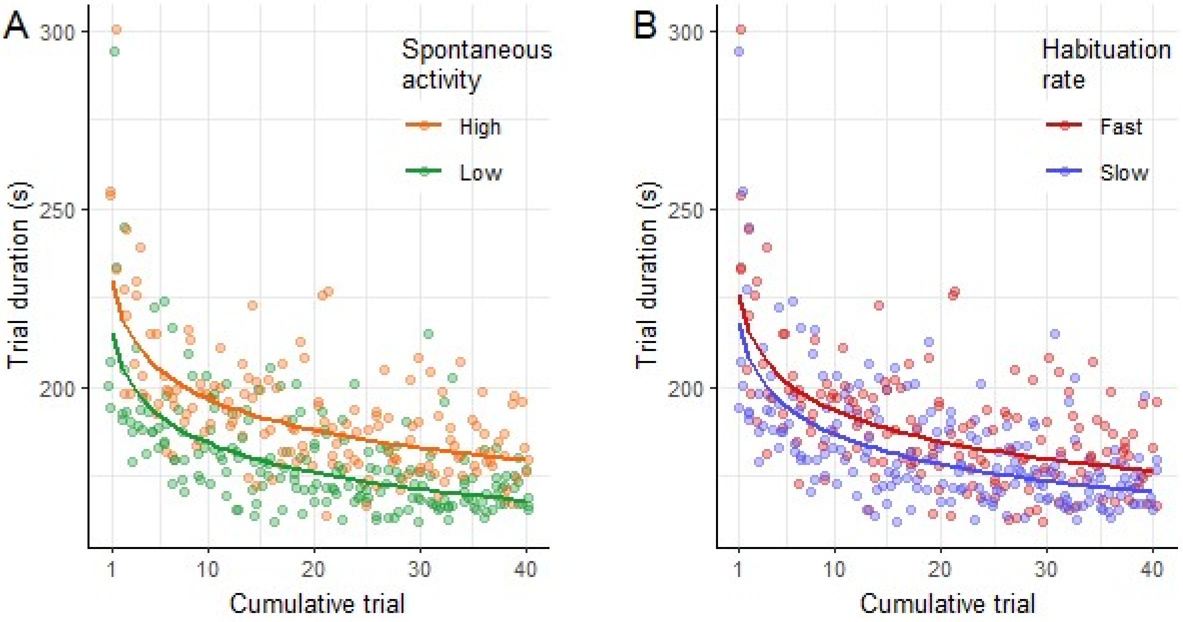


Figure 7. Linear regression model fits of duration predicted by cumulative trial and group (A: spontaneous activity, B: habituation rate). Both axes back-transformed from logarithmic scale.

Perceived importance ranged between 1.33 and 6.00 (M = 3.77, SD = 1.14). There were differences between both spontaneous activity and habituation groups, when analysed with separate linear regression models, as outlined in Table 1 (both models had cumulative trial as a control variable, the effect of which was not significant). Average perceived importance was 4.15 (SD = 0.81) for the low spontaneous activity group and 3.30 (SD = 1.30) for the high spontaneous activity group (t(357) = 11.82, $p < .001$, model $R^2 = .28$). These values were 4.31 (SD = 0.75) and 3.10 (SD = 1.19) for slow and fast habituation groups, respectively (t(357) = 7.63, $p < .001$, model $R^2 = .14$). See Appendix 4 for boxplots of perceived importance items by groups, as well as perceived importance scores over cumulative runs for each participant.

3.2.3 Trial-level habituation and deviation from predicted performance

Average trial-level habituation scores (SCR frequency during trial – SCR frequency during previous trial) were -1.82 (SD = 3.22) for the fast habituation group and -0.90 (SD = 3.09) for the slow habituation group. Deviation scores were residuals of the learning curve model above. *Absolute* values of the deviation scores were not correlated with trial-level habituation scores (*r*(162) = .02, *p* = .84), and hypothesis H2.3 was not supported. However, deviation scores were weakly correlated with trial-level habituation scores (*r*(162) = -0.25, *p* = .004), meaning that negative deviation scores (better-than-predicted performance) was connected to slower habituation. A similar relationship was seen in a linear mixed model with trial-level habituation as dependent variable, and trial (1-5) and deviation score as independent variables (Table 3). Comparison to a null model indicated that deviation score explained significantly more variance than the null model without that predictor ($\chi 2 = 11.1$, *p* = .001); AIC of the full model was 836 while AIC of the null model was 845.

Table 3

Results of linear mixed model with habituation score predicted by trial and deviation score, and a random participant effect for trial (intercept and slope)

	b	SE	t
(Intercept)	-3.75***	0.81	-4.62
Trial	0.67*	0.23	2.96
Deviation score	-20.35**	6.00	-3.39

*** *p* < .001, ** *p* < .01, * *p* < .05

4 Discussion

The purpose of this study was to explore whether temporal changes in task-related arousal can be characterised as habituation, and to utilise the concept of habituation in critically examining the relationship between arousal and performance. It was explored whether individual differences in habituation rate or spontaneous activity are associated with differences in performance. Furthermore, deviations from predicted performance were compared to habituation between trials. Based on the results, habituation occurred: task-related arousal decreased with repeated trials. This was the case for all participants and nearly all sessions. Individual differences in electrodermal reactivity – spontaneous activity and habituation rate – were associated with differences in performance, and maintained arousal at trial level was connected to better-than-predicted performance. These results are considered in light of two approaches – the cue utilisation theory and the orienting response – that could be seen as implying somewhat different causal mechanisms but are bound together by the notion of significance.

4.1 Habituation

A habituation model was successful in depicting changes in electrodermal activity over repeated trials of the steering task. Therefore, the concept of habituation could be extended to cover trials of 2-3 minutes in duration as aggregated data points, in contrast to individual SCRs to stimulus presentations. There was also some evidence of spontaneous recovery in responses between sessions, meaning that task-related arousal recovered when there was a delay between trials. This is consistent with the concept of habituation in terms of prediction (Bradley, 2009; Ursin & Eriksen, 2004): if the 'stimulation' – in this case, trial – is withheld, its next presentation (first trial of next session) carries more information than the previous (last trial of previous session). This is seen as recovery of the habituated response. However, time was not found to affect the magnitude of recovery. As Thompson (2009) points out, no prior estimate of how much time would be expected for spontaneous recovery to occur in this context can be given; but these results suggest that it might be less than one day.

While task-related arousal depended on the frequency of trials (i.e. withholding stimulation resulted in spontaneous recovery), less can be inferred about the rate of habituation, which was expected to amplify in later sessions. However, it is not certain that habituation slope alone could capture habituation in later sessions; on the contrary, low starting SCR frequency (intercept) in later sessions could indicate faster habituation already during the first trial compared to earlier sessions, since the habituation measure is an aggregate over several minutes. This would correspond to the 'potentiation of habituation' effect described by Rankin et al. (2009), and it is supported by the observation that habituation progressed below baseline in some sessions, mostly during later trials. However, based on these results, it remains inconclusive whether the rate of habituation can be said to change with repeated sessions, especially in the absence of EDA data in sessions 2-4.

4.2 Arousal and performance

While high spontaneous activity was connected to performance decrements, individuals with slow habituation performed better. A similar connection between habituation and performance was observed at trial level using two relative measures: better-than-expected performance (deviation from the learning curve) was linked to slow trial-level habituation (change in task-related arousal between trials).

Individual differences in electrodermal reactivity were explored in terms of spontaneous activity and habituation rate. Interestingly, the two measures resulted in a slightly different grouping of participants even though they have historically been considered analogous (Boucsein, 2012; Dawson et al., 2007, Crider et al., 2004). Most fast habituators had high spontaneous activity and vice versa. Unfortunately, the small sample size did not allow for the analysis of interaction effects of these groups. This discrepancy in classification may result from differences in experimental design: most habituation paradigms have consisted of repeated, usually non-target, stimuli, and habituation has been defined in terms of changes in individual SCRs (Dawson et al., 2007; Crider et al., 2004). In our study, habituation rate was determined from aggregated measures of electrodermal activity, and it was measured during continuous task performance. Therefore, a different result is not surprising given the different interpretation of passive (non-target) versus active (target) habituation in relation to significance.

4.2.1 Spontaneous activity

Low spontaneous activity was linked to enhanced performance. While high spontaneous activity has previously been linked to enhanced performance especially in sustained attention and reaction-time tasks due to enhanced stimulus processing (Sakai et al., 1992; Sostek, 1978), it has been proposed to be detrimental in tasks that require a large information processing capacity, in contrast to simple tasks (Crider, 2008). Therefore, it is important to consider the nature of our task in interpreting these results: it was a high-speed steering task where multiple stimuli were provided simultaneously, and it required skilful motor control, visual spatial attention and planning of consecutive actions. It cannot be paralleled to tasks where electrodermal reactivity has traditionally been studied, such as vigilance/monitoring tasks, or simple reaction-time tasks. It can be argued that the task is fundamentally different from simple reaction time or sustained attention tasks.

It might be the case that the higher spontaneous activity observed in our study reflects increased arousal related to anticipation of the task, or anxiety about wearing measurement equipment, or the experiment in general (MacIntosh, Mraz, McIlroy, & Graham, 2007), rather than an electrodermal lability trait. This would correspond to the idea of general, perhaps task-unrelated arousal. Arousal not related to the task would be expected to show as impaired performance, as attention would be directed to the source of arousal, not to the task (Hanoch & Vitouch, 2004). One explanation might be that participants with less gaming background, most of whom had higher spontaneous activity, were more anxious about the upcoming task – especially in the presence of experimenters – than those with more gaming experience and possibly higher perceived skill in games in general. This interpretation would fit the diverging effects of social stressors on performance depending on skill level, observed by, for example, Baumeister & Showers (1986). The way arousal is appraised would play an important role here: for instance, Hong (1999) argued that perceived difficulty affected test performance through the mediation of worry, and Jamieson, Mendes, Blackstock and Schmader (2010) found that reappraisal of arousal improved performance on an exam. Of course, based on our data it cannot be concluded whether worse performance observed in the high spontaneous activity group resulted simply from their less experience with games, rather than higher anxiety/worry about the situation, or something else.

4.2.2 Habituation rate

Slow habituation (maintained arousal) was connected to better performance in the steering task, both between participants and at trial level within participants. The cue utilisation theory would suggest that the maintained high arousal observed in slow habituators would be somewhere close to the ‘optimal level of arousal’ of the Yerkes-Dodson law, facilitating processing of task-relevant stimuli while filtering out irrelevant ones. However, even though decreases in arousal were connected to decrements in performance, there was no inverse U-shaped association between arousal level and performance, and no single ‘optimal level’ was found. In fact, it was the *change* in arousal that was reflected in learning and performance. Furthermore, all participants improved their performance to some extent, even though all of them displayed habituation. These results are better understood when habituation is interpreted as a signal of increased predictability. In this framework, events in the task became more predictable as participants learned to steer through the track. This increased predictability would then be manifested as habituation of task-related arousal (Bradley, 2009; Sokolov, 1963).

A distinction between goal-directed and stimulus-driven attention, or top-down and bottom-up processing (Corbetta & Shulman, 2002; Näätänen, 1992), is useful in interpreting temporal patterns in task-related arousal. While responses to non-target stimuli are elicited primarily by bottom-up signals – though influenced by expectations – task-related arousal stems from goal-directed allocation of attention and can be regarded as an indicator of motivation (Bradley, 2009; Näätänen, 1992). Bradley (2009) argues that habituation in active contexts, such as task performance, differs markedly from habituation in passive contexts due to its connection to motivational and emotional systems. Habituation has indeed been shown to be slower to target stimuli than to non-target stimuli (Barry, 2004; Bradley, 2009; Bradley, Lang, & Cuthbert, 1993). Higher perceived significance, perhaps resulting from different motivational states (Bradley, 2009), would be linked to slower habituation and better performance due to enhanced attentional processing. This is supported by the observation that slow habituators had higher perceived importance.

Slow habituation rate was connected to better overall performance at individual level, and a similar relationship was observed within individuals: maintained arousal (slow habituation) between trials was associated with better-than-expected performance. This is an intriguing result that highlights the need to conceptualise the relationship between performance and arousal in relative terms. Interestingly, changes in arousal have been linked to performance monitoring and deviations from predicted performance in a repeated task with two difficulty conditions (Braem et al., 2015). In their study on a flanker task, Braem et al. (2015) found that errors in an easy version of the task - which were presumably less predictable than errors in the difficult version - elicited phasic arousal responses, and vice versa. The authors concluded that this signalled the surprise associated with a deviation from the predicted number of errors. However, in our results, slower habituation was only linked to *better*-than-predicted performance, not worse, as would be expected if it were to signal the 'absolute' surprise value. Because the measures used here were aggregates over a period of 2-3 minutes, it is less clear what 'deviation from prediction' entails. Together, these results support the interpretation that slow habituation is an indicator of higher motivation, rather than simply (un)predictability.

In their study on fMRI and EDA features during motor skill acquisition, MacIntosh et al. (2007) argue that a temporal decrease in EDA can be related to short-term learning effects, decrease in overall arousal, or decrease in sustained attention. While learning effects would explain some of the concomitant EDA decrease and performance enhancement in our task, spontaneous recovery of task-related arousal between sessions suggests that there were additional factors which were not reflected in performance (there was no ‘spontaneous recovery’ effect in performance between sessions). These may include a decrease in general arousal within sessions (which could be measured by including another baseline period at the end), or re-orientation to the task after a break. Furthermore, faster decrease in task-related arousal was not linked to faster learning, but to a better performance level. On the whole, because some level of habituation was observed in almost all sessions, some of it can probably be connected to temporal changes in arousal related to being exposed to the task (similarly to traditional habituation paradigms). However, differences in the rate of habituation perhaps reflect motivational processes both at individual and trial level. There were slightly more experienced players in the slow habituation group; this might explain their higher perceived importance.

It is noteworthy that in the task, the speed of the cube constantly increased unless obstacles were hit, adapting difficulty to the skill level of the participant. Consequently, the participants who performed better also completed more difficult trials. Furthermore, while the same number of obstacles were included in all trials, they were placed randomly for each trial and the exact track arrangement was slightly different, making it possible for some trials to be easier by coincidence. This emphasises the fact that no causal connections can be made between arousal and performance: it might be that the mere experience of playing an ‘easier’ trial resulted in lower physiological arousal.

4.3 Limitations and future research

The clearest limitations of this study are related to the number of participants and the arrangement of trials within sessions. While the number of longitudinal measurement points was adequate for analysing general trends in electrodermal habituation and learning, random effects related to participants and sessions call for a larger sample of participants. In group analyses, the number of participants was a clear limiting factor, and even though both spontaneous activity and habituation rate were found to have a significant effect on performance on their own, no solid inferences could be made due to the variance in participants with respect to gender, as well as gaming or driving experience. A substantially larger sample size would be needed to make these comparisons.

The current design did not allow for sophisticated comparisons about potentiation of habituation or spontaneous recovery due to random session effects that could not be sufficiently controlled by baseline measurements. These phenomena, and associated time effects, could be better studied by including varying time intervals between trials within one measurement session. Alternatively, skin temperature could be recorded to facilitate the comparison of skin conductance responses from different sessions (Boucsein, 2012).

Baseline measurements could be done between trials or at the end of the session to separate baseline arousal and task-related arousal more clearly. Introducing multiple novel tasks in the same session would also help distinguish between learning effects and general arousal (MacIntosh et al., 2007). Although movement artefacts were minimised by not attaching sensors to the hands, moving the steering wheel is a potentially substantial source of error. Therefore, the baseline period should include movement of the wheel.

There is controversy on how to quantify habituation, and while a commonly used alternative, the *trials-to-habituation* measure (Crider et al., 2004; Dawson et al., 2007) - the number of stimulus presentations until zero response - is not applicable in this design as such, a similar assessment could be made by calculating trials to baseline (or a predefined number of consecutive values at the same level). Based on our data, this would require more trials per session.

The features of skin conductance responses used in analysis are another topic of concern. Isen, Iacono & Malone (2013) argue that habituation measures should include both the frequency and amplitude of responses, and they used latent class analysis to categorise participants based on these features, finding four classes of habituators. While only SCR frequency was reported in this study for simplicity, the analyses were also conducted with an amplitude sum measure yielding similar results. Again, a more thorough investigation about these features would necessitate a larger sample.

The analyses presented here were based on aggregated data points confined at trial level. It would also be possible to study electrodermal activity and its habituation *within* trials - or its connection to trial events or features of steering. For example, the task gave clear feedback with a flash of the screen when an error (collision) was made. Physiological responses to these errors could provide more insight into the relationship between arousal and performance. On the other hand, if analysis remained at trial level, trials could be made shorter to achieve denser data points.

In results reported in Cowley et al. (2019), performance deviation scores in this task were related to subjective reports on flow experience, so that better-than-predicted performance was connected to higher reported flow. This suggests that physiological habituation, learning and flow may be interconnected and should be studied in more detail. In addition to exploring individual differences, different experimental conditions could be implemented by varying task difficulty (starting velocity, object arrangement) or presenting external, task-irrelevant stimuli (such as distractor sounds).

4.4 Conclusion

In this study, it was shown that temporal changes in task-related arousal in a steering task, measured by frequency of skin conductance responses, can be characterised as habituation, according to criteria presented by Spencer and Thompson (1966) and revised by Rankin et al. (2009). A trial-level connection was found between better-than-predicted performance and maintained arousal (slow habituation). The relationship between predicted performance and arousal responses is an intriguing concept and hopefully studied more in the future. Moreover, individual differences in electrodermal reactivity were found to be reflected in performance, perhaps relating to perceived importance. However, individual differences and possible effects of background variables should be studied in more detail in future studies.

Interpreting the results about arousal and performance in light of the Yerkes-Dodson law or the cue utilisation theory alone is challenging as no direct relationship between arousal and performance levels was found. However, by using the concept of habituation, a temporal dimension could be added to the Yerkes-Dodson law. In the context of this task, maintained arousal or, in other words, slower habituation, could be said to reflect an optimal level of *relative* arousal and attentional processing for stimuli that carry information - which can change over time.

The study was founded on a relatively novel approach to habituation, and the results provided valuable insights into the relationship between arousal and performance, suggesting that changes in task-related arousal over time should be studied more to deepen the understanding of attentional processes and arousal in learning and performance. Taken together, these results call for a critical evaluation of existing theories of arousal-performance relationships, as well as formulation of new theories that consider the dynamic nature of these phenomena.

References

Barry, R. J. (2004). Stimulus significance effects in habituation of the phasic and tonic orienting reflex. *Integrative Physiological & Behavioral Science, 39*(3), 166-179.

Barry, R. J., Clarke, A. R., McCarthy, R., Selikowitz, M., & Rushby, J. A. (2005). Arousal and activation in a continuous performance task. *Journal of Psychophysiology*, *19*(2), 91-99.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48.

Baumeister, R. F., & Showers, C. J. (1986). A review of paradoxical performance effects: Choking under pressure in sports and mental tests. *European Journal of Social Psychology, 16*(4), 361-383.

Benedek, M., & Kaernbach, C. (2010). A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods, 190*(1), 80-91.

Benedek, M., & Kaernbach, C. (2010). Decomposition of skin conductance data by means of nonnegative deconvolution. *Psychophysiology, 47*(4), 647-658.

Boonstra, R. (2013). Reality as the leading cause of stress: rethinking the impact of chronic stress in nature. *Functional Ecology, 27*(1), 11-23.

Boucsein, W. (2012). *Electrodermal activity* (2nd ed.). New York, NY, US: Springer Science & Business Media.

Bradley, M. M. (2009). Natural selective attention: orienting and emotion. *Psychophysiology, 46*(1), 1-11.

Bradley, M. M., Miccoli, L., Escrig, M. A., & Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology, 45*(4), 602-607.

Bradley, M. M., Lang, P. J., & Cuthbert, B. N. (1993). Emotion, novelty, and the startle reflex: habituation in humans. *Behavioral Neuroscience, 107*(6), 970.

Braem, S., Coenen, E., Bombeke, K., Van Bochove, M. E., & Notebaert, W. (2015). Open your eyes for prediction errors. *Cognitive, Affective, & Behavioral Neuroscience*, *15*(2), 374-380.

Broadhurst, P. L. (1957). Emotionality and the Yerkes-Dodson law. *Journal of Experimental Psychology*, *54*(5), 345.

Campbell, B. A., Wood, G., & McBride, T. (1997). Origins of orienting and defensive responses: an evolutionary perspective. In P. J. Lang, R. F. Simons, M. Balaban, & R. Simons (Eds.), *Attention and orienting: Sensory and motivational processes* (pp. 41-67). New York, NY, US: Psychology Press.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36*(3), 181-204.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience, 3*(3), 201-15.

Cowley, B. U., Palomäki, J., Tammi, T., Frantsi, R., Inkilä, V. P., Lehtonen, N., ... & Lappi, O. Flow experiences during visuomotor skill acquisition reflect deviation from a power-law learning curve, but not overall level of skill. Manuscript submitted for publication.

Crider, A. (1993). Electrodermal response lability-stability: individual difference correlates. In *Progress in Electrodermal Research* (pp. 173-186). Springer, Boston, MA.

Crider, A. (2008). Personality and electrodermal response lability: an interpretation. *Applied Psychophysiology and Biofeedback*, *33*(3), 141.

Crider, A., Kremen, W. S., Xian, H., Jacobson, K. C., Waterman, B., Eisen, S. A., ... & Lyons, M. J. (2004). Stability, consistency, and heritability of electrodermal response lability in middle-aged male twins. *Psychophysiology, 41*(4), 501-509.

Dawson, M. E., Schell, A. M., & Filion, D. L. (2007). The electrodermal system. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (3rd ed., pp. 159-181). New York, NY, US: Cambridge University Press.

Deinzer, R., Kirschbaum, C., Gresele, C., & Hellhammer, D. H. (1997). Adrenocortical responses to repeated parachute jumping and subsequent h-CRH challenge in inexperienced healthy subjects. *Physiology & Behavior, 61*(4), 507-511.

Diamond, D. M., Campbell, A. M., Park, C. R., Halonen, J., & Zoladz, P. R. (2007). The temporal dynamics model of emotional memory processing: a synthesis on the neurobiological basis of stress-induced amnesia, flashbulb and traumatic memories, and the Yerkes-Dodson law. *Neural Plasticity*, *2007*.

Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review*, *66*(3), 183.

Eckman, P. S., & Shean, G. D. (1997). Habituation of cognitive and physiological arousal and social anxiety. *Behaviour Research and Therapy*, *35*(12), 1113-1121.

Engeser, S., & Rheinberg, F. (2008). Flow, performance and moderators of challenge-skill balance. *Motivation and Emotion, 32*(3), 158-172.

Frith, C. D., & Allen, H. A. (1983). The skin conductance orienting response as an index of attention. *Biological Psychology, 17*(1), 27-39.

Gerra, G., Zaimovic, A., Mascetti, G. G., Gardini, S., Zambelli, U., Timpano, M., ... & Brambilla, F. (2001). Neuroendocrine responses to experimentally-induced psychological stress in healthy humans. *Psychoneuroendocrinology, 26*(1), 91-107.

Grissom, N., & Bhatnagar, S. (2009). Habituation to repeated stress: get used to it. *Neurobiology of Learning and Memory, 92*(2), 215-224.

Hamer, M., Gibson, E. L., Vuononvirta, R., Williams, E., & Steptoe, A. (2006). Inflammatory and hemostatic responses to repeated mental stress: individual stability and habituation over time. *Brain, Behavior, and Immunity*, *20*(5), 456-459.

Hanoch, Y., & Vitouch, O. (2004). When less is more: Information, emotional arousal and the ecological reframing of the Yerkes-Dodson law. *Theory & Psychology*, *14*(4), 427-452.

Hebb, D. O. (1955). Drives and the CNS (conceptual nervous system). *Psychological review, 62*(4), 243.

Hong, E. (1999). Test anxiety, perceived test difficulty, and test performance: Temporal patterns of their effects. *Learning and Individual Differences, 11*(4), 431-447.

Isen, J. D., Iacono, W. G., & Malone, S. M. (2013). Characterizing electrodermal response habituation: A latent class approach with application to psychopathology. *Psychophysiology, 50*(10), 954-962.

Jamieson, J. P., Mendes, W. B., Blackstock, E., & Schmader, T. (2010). Turning the knots in your stomach into bows: Reappraising arousal improves performance on the GRE. *Journal of Experimental Social Psychology, 46*(1), 208-212.

Koolhaas, J. M., Bartolomucci, A., Buwalda, B. D., De Boer, S. F., Flügge, G., Korte, S. M., ... & Richter-Levin, G. (2011). Stress revisited: a critical evaluation of the stress concept. *Neuroscience & Biobehavioral Reviews*, *35*(5), 1291-1301.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26.

Lang, P. J. (2010). Emotion and motivation: Toward consensus definitions and a common research purpose. *Emotion Review, 2*(3), 229-233.

MacIntosh, B. J., Mraz, R., McIlroy, W. E., & Graham, S. J. (2007). Brain activity during a motor learning task: an fMRI and skin conductance study. *Human Brain Mapping, 28*(12), 1359-1367.

McEwen, B. S., & Gianaros, P. J. (2010). Central role of the brain in stress and adaptation: links to socioeconomic status, health, and disease. *Annals of the New York Academy of Sciences, 1186*(1), 190-222.

Munro, L. L., Dawson, M. E., Schell, A. M., & Sakai, L. M. (1987). Electrodermal lability and rapid vigilance decrement in a degraded stimulus continuous performance task. *Journal of Psychophysiology, 1*(3), 249-257.

Näätänen, R. (1992). *Attention and brain function* (6th ed.). Hillsdale, N.J: L. Erlbaum.

Newell, A., & Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the law of practice. *Cognitive Skills and their Acquisition, 1*(1981), 1-55.

Öhman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: detecting the snake in the grass. *Journal of Experimental Psychology: General*, *130*(3), 466.

Oken, B. S., Salinsky, M. C., & Elsas, S. M. (2006). Vigilance, alertness, or sustained attention: physiological basis and measurement. *Clinical Neurophysiology, 117*(9), 1885-1901.

Plessow, F., Kiesel, A., & Kirschbaum, C. (2012). The stressed prefrontal cortex and goal-directed behaviour: Acute psychosocial stress impairs the flexible implementation of task goals. *Experimental Brain Research, 216*(3), 397-408.

Rankin, C. H., Abrams, T., Barry, R. J., Bhatnagar, S., Clayton, D. F., Colombo, J., ... & McSweeney, F. K. (2009). Habituation revisited: An updated and revised description of the behavioral characteristics of habituation. *Neurobiology of Learning and Memory*, *92*(2), 135-138.

Sakai, L. M., Baker, L. A., & Dawson, M. E. (1992). Electrodermal lability: Individual differences affecting perceptual speed and vigilance performance in 9 to 16 year-old children. *Psychophysiology*, *29*(2), 207-217.

Sanders, A. F. (1983). Towards a model of stress and human performance. *Acta Psychologica, 53*(1), 61-97.

Sapolsky, R. M. (2015). Stress and the brain: individual variability and the inverted-U. *Nature Neuroscience*, *18*(10), 1344.

Schell, A. M., Dawson, M. E., & Filion, D. L. (1988). Psychophysiological correlates of electrodermal lability. *Psychophysiology, 25*(6), 619-632.

Simpson, E. H., & Balsam, P. D. (2016). The behavioral neuroscience of motivation: an overview of concepts, measures, and translational applications. *Current Topics in Behavioral Neurosciences, 27*(1), 1–12.

Sokolov, E. N. (1963). Higher nervous functions: the orienting reflex. *Annual Review of Physiology*, *25*(1), 545-580.

Sokolov, E. N. (1978). A neuronal model of the stimulus in a reflex arc. *Soviet Psychology, 17*(2), 3-22.

Sokolov, E. N., Spinks, J. A., Näätänen, R., & Lyytinen, H. (2002). *The orienting response in information processing.* Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Sostek, A. J. (1978). Effects of electrodermal lability and payoff instructions on vigilance performance. *Psychophysiology*, *15*(6), 561-568.

Thompson, R. F. (2009). Habituation: a history. *Neurobiology of Learning and Memory, 92*(2), 127.

Thompson, R. F., & Spencer, W. A. (1966). Habituation: a model phenomenon for the study of neuronal substrates of behavior. *Psychological Review*, *73*(1), 16.

Ursin, H., & Eriksen, H. R. (2004). The cognitive activation theory of stress. *Psychoneuroendocrinology, 29*(5), 567-592.

Vossel, G., & Rossmann, R. (1984). Electrodermal habituation speed and visual monitoring performance. *Psychophysiology, 21*(1), 97-100.

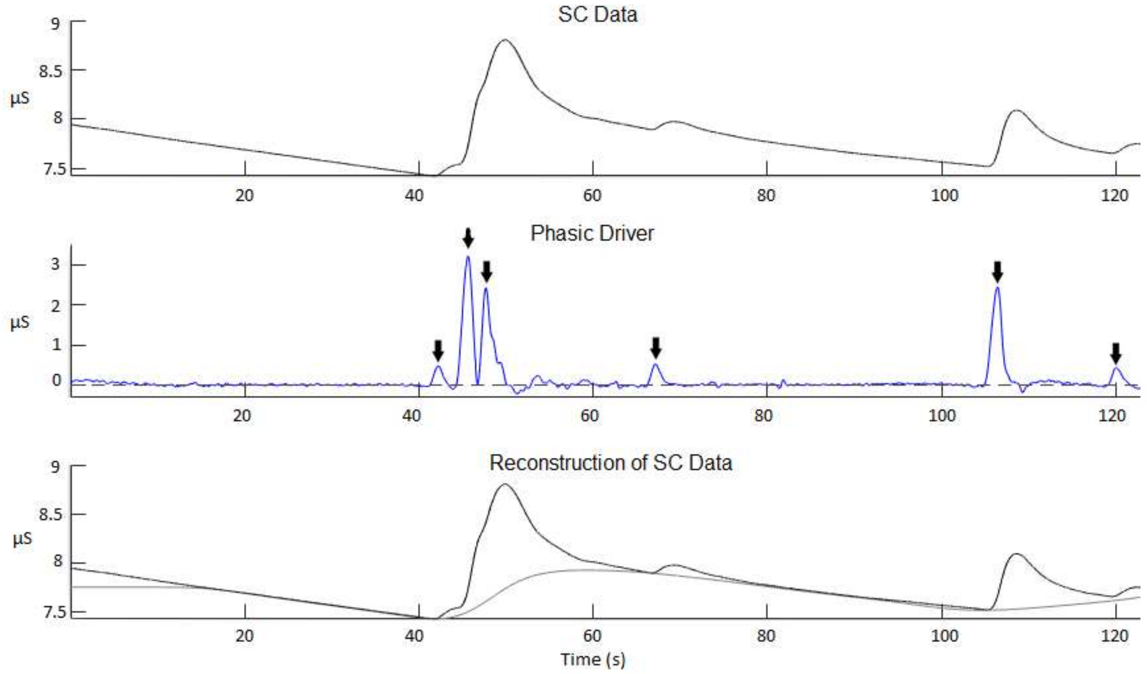
Watters, P. A., Martin, F., & Schreter, Z. (1997). Caffeine and cognitive performance: The nonlinear Yerkes–Dodson law. *Human Psychopharmacology: Clinical and Experimental, 12*(3), 249-257.

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

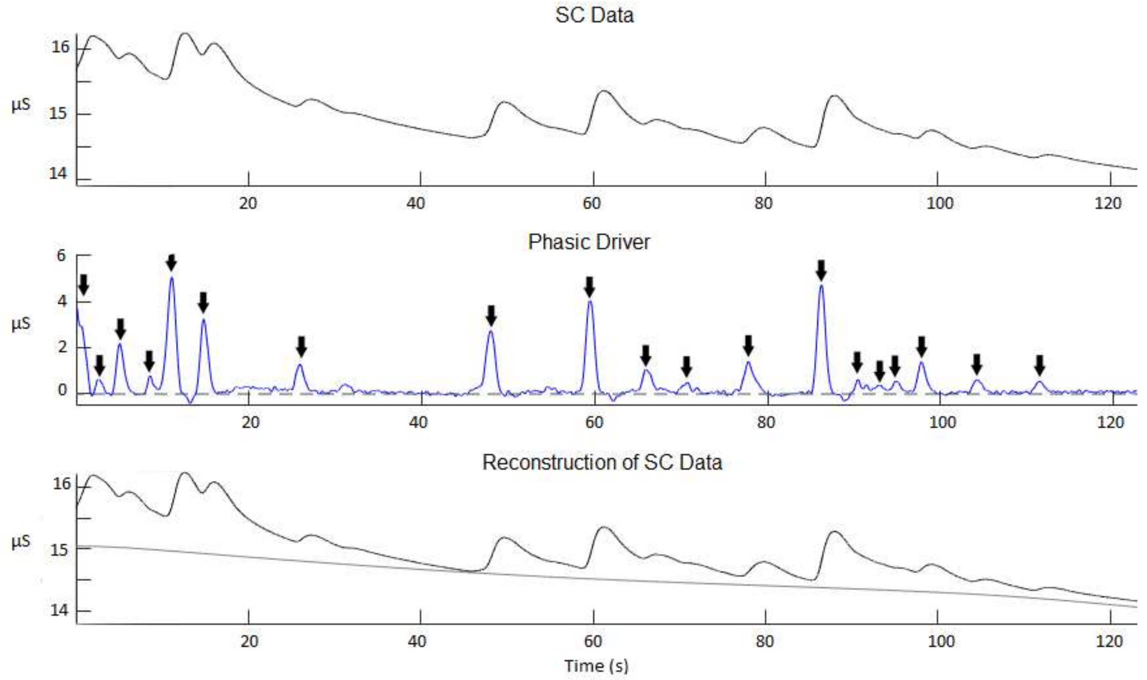
APPENDIX 1: SCR extraction with CDA

Example 120-second periods of baseline (A) and trial (B) EDA. The EDA signal (SC data, top panel) is decomposed by Continuous Decomposition Analysis (Benedek & Kaernbach, 2010a) into a phasic driver (middle panel) and a tonic component (not shown). Phasic responses over 0.05 μS are marked in the middle panel. Note that the amplitude criterion refers to SCRs reconvolved from corresponding phasic driver peaks, not the phasic driver signal itself (Benedek & Kaernbach, 2010b). The bottom panel shows a reconstructed signal from tonic and phasic components.

A. Baseline



B. Trial



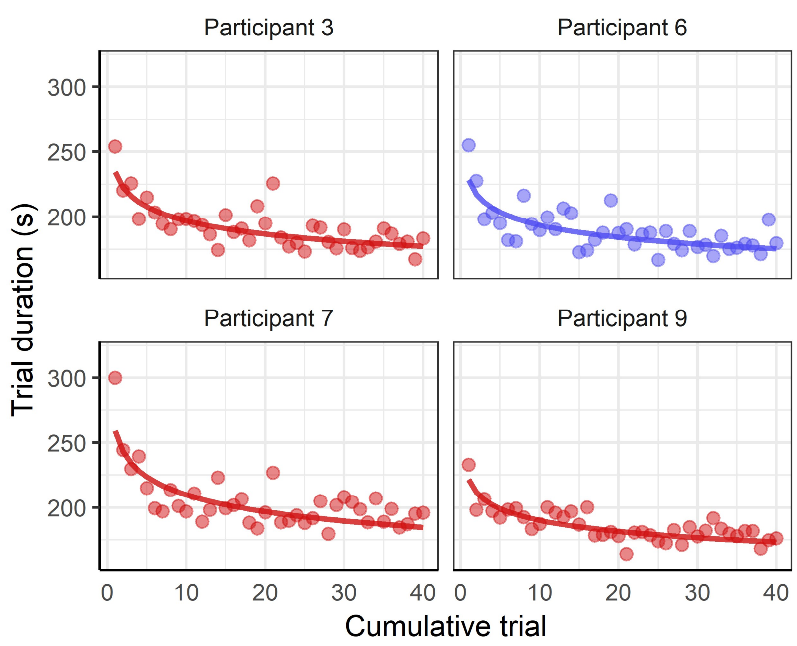
APPENDIX 2: Participant information

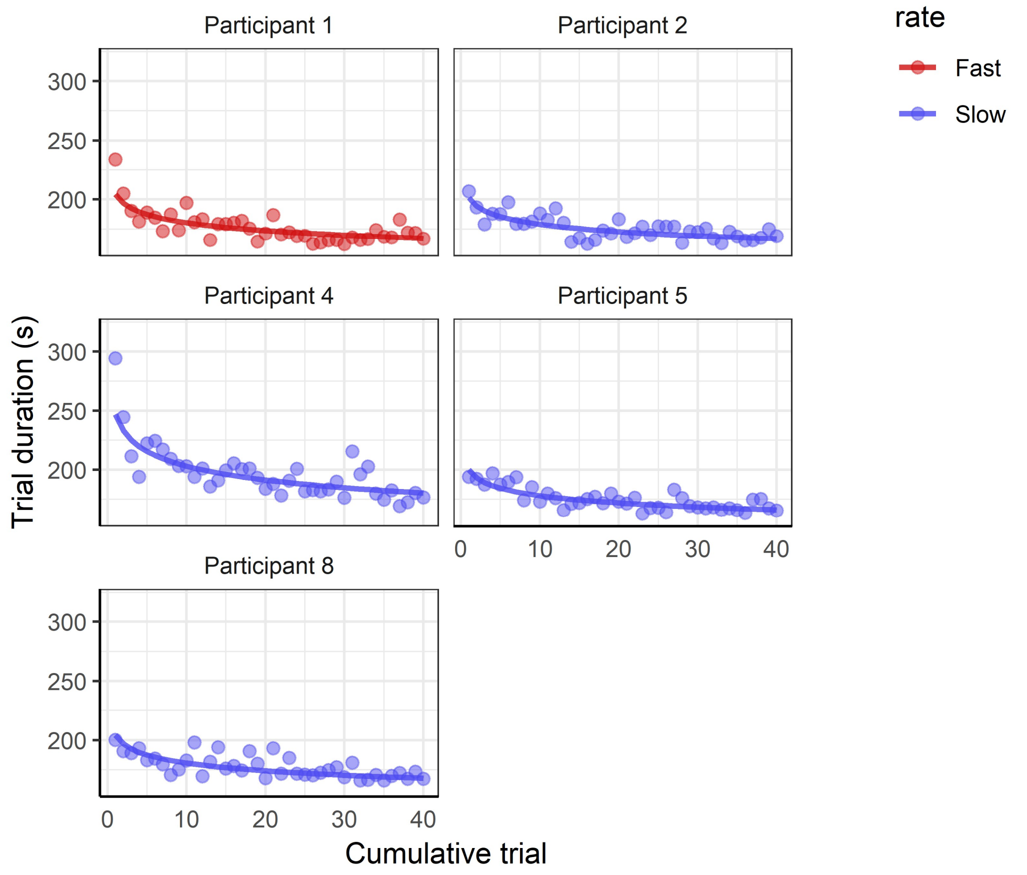
Participant background information and individual measures of learning, habituation, spontaneous activity, and perceived importance. Inter-quartile range in parentheses for baseline median, standard deviation for perceived importance mean.

Participant	1	2	3	4	5	6	7	8	9
Learning curve									
Slope	-0.05	-0.05	-0.08	-0.08	-0.05	-0.07	-0.09	-0.05	-0.07
Intercept	5.33	5.31	5.46	5.51	5.3	5.44	5.56	5.32	5.41
Habituation rate	Fast	Slow	Fast	Slow	Slow	Slow	Fast	Slow	Fast
Slope	-3.61	-2.6	-3.04	-2.61	-2.7	-2.85	-3.59	-2.28	-3.95
Intercept	7.5	15.9	9.05	11.91	16.14	6.88	7.6	6.13	10.17
Spontaneous activity	Low	Low	High	Low	Low	High	High	Low	High
Baseline SCR frequency	1.3 (1.7)	0.7 (1.0)	6.0 (2.3)	0.2 (0.5)	1.3 (1.0)	7.3 (4.9)	7.3 (3.7)	0.0 (1.1)	6.0 (3.3)
Perceived importance	3.53 (0.57)	4.29 (0.47)	2.03 (0.50)	4.12 (0.46)	5.16 (0.64)	4.33 (0.56)	2.22 (0.53)	3.67 (0.70)	4.63 (0.55)
Gender	Male	Male	Female	Female	Male	Male	Female	Male	Male
Driving experience (km)	≤10,000	>10,000	≤10,000	≤10,000	>10,000	>10,000	>10,000	>10,000	>10,000
Gaming frequency	Weekly	Weekly	Less than weekly	Weekly	Weekly	Less than weekly	Less than weekly	Less than weekly	Weekly

APPENDIX 3: Individual learning curves by groups

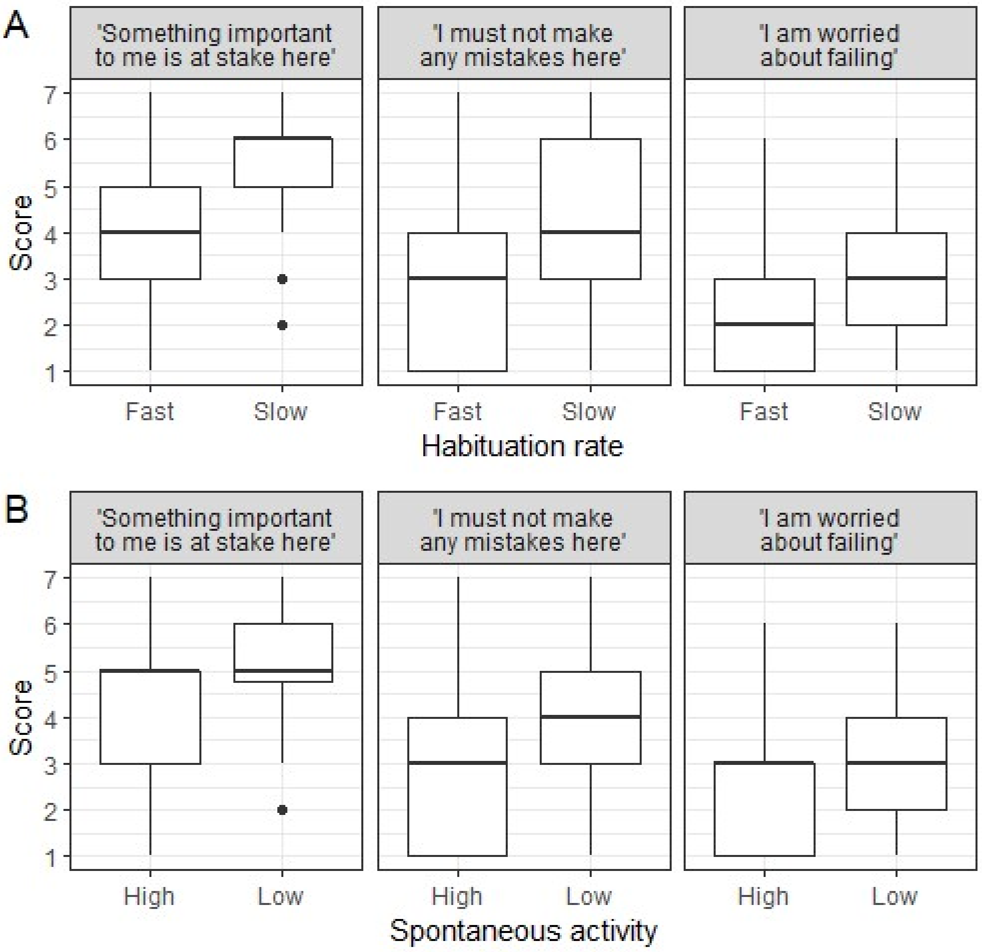
Individual learning curves for high (A) and low (B) spontaneous activity groups.


A. High spontaneous activity


B. Low spontaneous activity

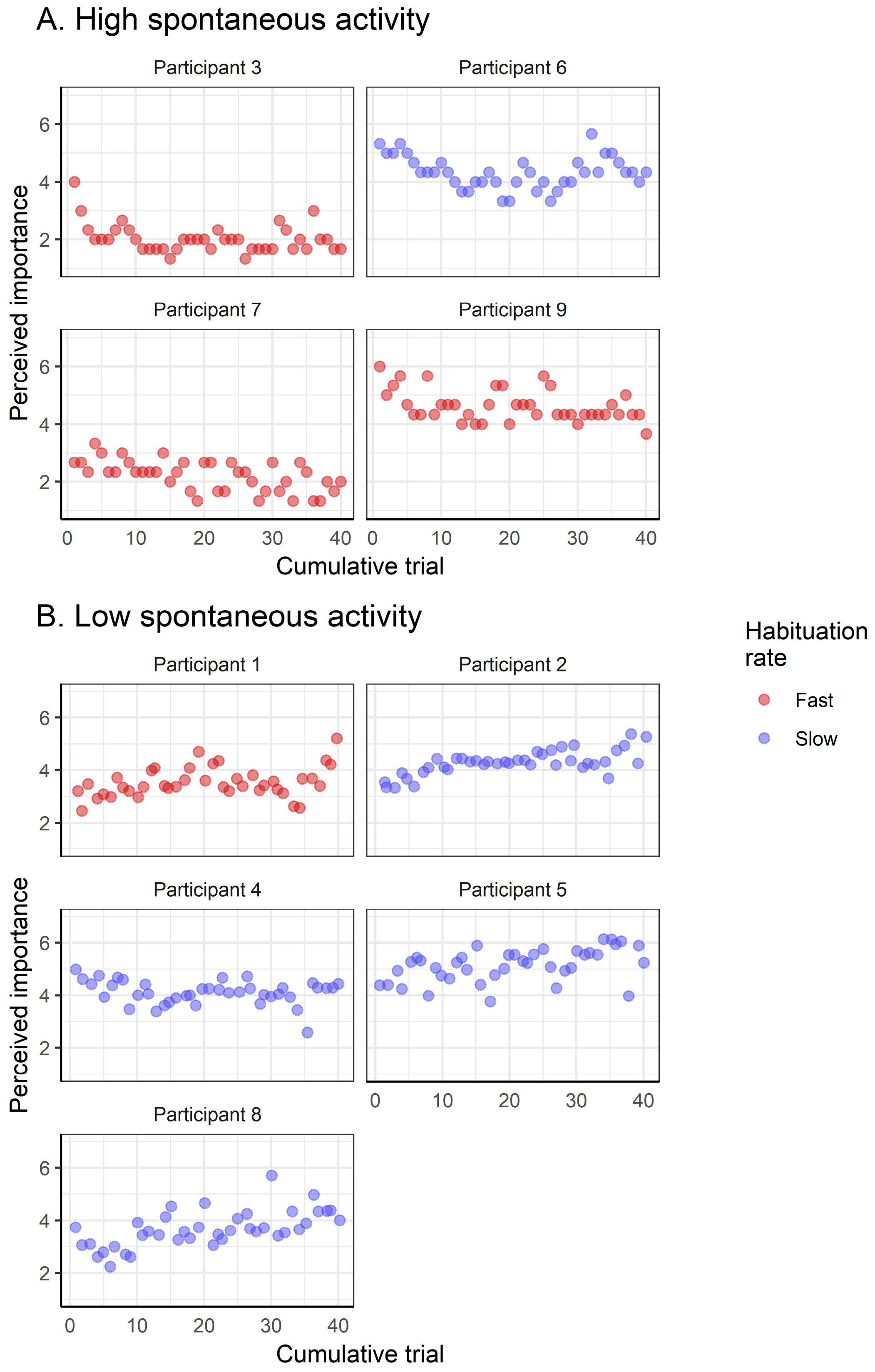
APPENDIX 4: Perceived importance

Perceived importance item distributions for habituation (A) and spontaneous activity (B) groups.



APPENDIX 4: Perceived importance

Perceived importance for cumulative trials for high (A) and low (B) spontaneous activity groups.



APPENDIX 5: Flow Short Scale

Flow Short Scale by Rheinberg & Engeser (2008), and a Finnish translation by Noora Lehtonen, Tuisku Tammi, Pasi Pölönen, Roosa Frantsi, Ville-Pekka Inkilä and Jussi Palomäki.
The *flow* scale consists of subscales *fluency* (6 items) and *absorption* (4 items).
Additional scales: *perceived importance* (3 items) and *perceived fit of demands and skills* (3 items).

English	Finnish
Fluency	
2. My thoughts/activities run fluidly and smoothly	Pelasin sujuvasti
4. I have no difficulty concentrating	Pystyin hyvin keskittymään
5. My mind is completely clear	Mieleni oli selkeä
7. The right thoughts/movements occur of their own accord	Löysin oikeat liikkeet kuin itsestään
8. I know what I have to do each step of the way	Olin koko ajan tilanteen tasalla
9. I feel that I have everything under control	Tunsin hallitsevani tilannetta
Absorption	
1. I feel just the right amount of challenge	Peli tuntui juuri sopivan haastavalta
3. I do not notice time passing	En huomannut ajankulkua
6. I am totally absorbed in what I am doing	Uppouduin täysin pelaamiseen
10. I am completely lost in thought	Syvennyin peliin täysin
Perceived importance	
11. Something important to me is at stake here	Koin pelissä onnistumisen tärkeäksi
12. I must not make any mistakes here	Minusta tuntui siltä, etten saisi tehdä yhtäkään virhettä
13. I am worried about failing	Pelkäsin epäonnistuvani
Perceived fit of demands and skills	
Compared to all other activities which I partake in, this one is … (easy - difficult)	Verrattuna muihin tekemiini asioihin, tämä on ... (helppoa - vaikeaa)
I think that my competence in this area is … (low - high)	Osaamiseni taso on … (matala - korkea)
For me personal, the current demands are … (too low - just right - too high)	Pelin vaativuus on tällä hetkellä minulle ... (liian matala - sopiva - liian korkea)