The Effect Of Perceptual Load On Attention Switching

Luca Chech

UCL

MPhil in Cognitive Neuroscience

‘I, Luca Chech, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.'

Abstract

Automation in the car is becoming progressively more sophisticated and we are now approaching a critical junction, where vehicles will be capable of taking care of all aspects of driving but with the expectation that the driver will promptly respond to a request to take-over. With drivers already engaging in a variety of non-driving tasks, it becomes crucial to evaluate the assumption of their readiness to intervene. While simulator studies have partly addressed this expectation by comparing different non-driving tasks, no research has tried to systematically vary the attentional demands of the non-driving tasks and measure their impact on the take-over process. Here, aided by the conceptual framework provided by Perceptual Load Theory, I explore two different scenarios showing how manipulation of attentional load in the non-driving task might hamper drivers' ability to execute different aspects of the take-over process. While testing was performed entirely in the laboratory, each experiment employs tasks designed to be relevant proxies for both the non-driving tasks and the take-over request. In Chapter 2, I present two experiments in which participants are asked to watch a sequence of natural scenes of varied perceptual load – the non-driving task – while monitoring for the occurrence of an auditory stimulus – the take-over request. High perceptual load was associated with reduced detection of the auditory stimulus. The three experiments reported in Chapter 3 instead aim at understanding the extent to which high attentional demands right before a task switch might hamper the ability to correctly process and respond to the motion of other vehicles, assessed with the use of random dot kinematograms. A high level of perceptual load was reliably accompanied by slower responses to the motion stimuli. Finally, in Chapter 4 I describe an fMRI experiment looking at possible neural contributions to the reaction-time delay observed in Chapter 3.

Impact Statement

The following PhD project is part of the TASCC (Towards Autonomy – Smart and Connected Control) programme. This research consortium, jointly funded by Jaguar Land Rover and EPSRC and carried out in over a dozen universities across the UK, aims to tackle several outstanding issues around the introduction of self-driving cars from a multitude of perspectives. These issues range from the legal (e.g., what levels of automation are allowed in a given country; what kind of trials can be conducted; is it mandatory to have a human supervisor in the vehicle at all times) to the technical and even the philosophical. For instance, the re-edition of the trolley problem in the context of self-driving cars asks what choice should be programmed into the car in the case of an unexpected obstacle that would result in certain collision, when each choice that the autonomous vehicle might take carries a toll in terms of people that would be injured or killed. Should the driver be sacrificed to save one or more bystanders? Also, the technical challenges are much broader than just improving the array of sensing technologies (such as radar, lidar, cameras, infrared sensors and so on) that allow the car to make sense of the surrounding environment and carry out the driving task. How should vehicles coordinate when the majority of the cars will be fully autonomous? The need to have a great number of vehicles talking to each other with a very low latency in order to negotiate their respective position in real time carries the additional challenge of a high-speed, pervasive data network to be planned, built and tested. As we can appreciate, these issues are really broad and complex, and the problem space needs to be well thought out long before the actual mass adoption of self-driving vehicles.

There is also another piece to the puzzle, which will be particularly important in the near future: the driver. Despite the rapid progress that we are witnessing, fully automated vehicles will not happen overnight. We are now living a transition phase, the length of which is hard to guess at present, that sees the human driver still in control, albeit aided by an ever-increasing range of automated components. It is therefore of paramount importance that man and machine interact in the most effective way possible. People's trust in the automation, the transparency of the automation itself (i.e., how to have the automation communicate its internal state in an intelligible way to the driver) and the transition of authority between the car and the driver are all problems belonging to this domain. The CogShift (Driver-Cognition-Oriented Optimal Control Authority Shifting for Adaptive Automated Driving) program, of which my PhD is part, engages with the latter. Its rationale is therefore to apply cognitive psychology and neuroscientific expertise to assess how cognitively well-equipped people are to handle authority transition from the car to the driver.

Contents

Chapter 1:
General introduction

The trend toward the introduction of more sophisticated in-vehicle assistive technology (park assist, cruise-control, lane-keeping assist, etc.) has been ongoing for decades, with highly automated driving (HAD) soon to become a reality. An authoritative classification of vehicle automation levels (J3016_202104), formulated by the Society of Automotive Engineers (SAE, 2021) proposes six hierarchical levels of automation, with most of the cars currently on the road belonging to either level 0, 1 or 2. Level 3 vehicles are likely to be the most problematic, because they will be able to take care “of all aspects of the dynamic driving task with the expectation that the human driver will respond appropriately to a request to intervene” (SAE, 2021). In other words, even though the task of monitoring the environment will be left to the car while the automation is active, the driver will still be expected to be able to quickly and efficiently resume control when the need arises (due to, for example, faded lane markings or sensors limitations). The literature has shown that drivers currently engage in a number of non-driving tasks and that these tasks certainly have an impact on their driving performance (Young & Regan, 2007). This trend is bound to increase with level 3 autonomous vehicles, and it is therefore of the utmost importance to ensure that drivers are able to take back control of the vehicle safely and efficiently.

Over the course of the next four chapters I examine two different scenarios, each investigating a critical issue that needs to be addressed in order to provide a satisfactory answer to the fundamental question posed by this project - namely if and to what extent non-driving tasks might hamper the drivers’ readiness and effectiveness at promptly taking back control of the vehicle should the need arise – with the conceptual framework provided by Perceptual Load Theory.

Although testing was performed entirely in the laboratory, each scenario is designed to capture an important aspect of the underlying control authority transition problem, with tasks chosen to be representative proxies of both the take-over request emitted by the vehicle as well as of the non-driving tasks drivers might perform while letting the automation take care of the driving.

Chapter 2 focuses on what is arguably the first step in achieving an effective transition of control: understanding the effect that a non-driving task might have on the driver's ability to perceive the take-over request issued by the car in the first place. Lavie's Perceptual Load Theory of selective attention and cognitive control (Lavie & Tsal, 1994), which ties awareness of task-irrelevant stimuli to the level of perceptual (and cognitive) load posed by each specific task, is inherently applicable to this situation. In Chapter 3 the aim is instead that of quantifying the effects that different levels of perceptual load in the non-driving task right before the take-over request might have on the driver's ability to process motion in the surrounding environment after the switch. Chapter 4 persists on the same question posed in the previous chapter and tries to clarify some of its findings by means of an fMRI experiment.

In this chapter I firstly review the literature on highly automated driving and take-over requests, highlighting two areas that have not received much consideration so far: the lack of experiments in which attentional demands posed by the non-driving task are systematically manipulated, as well as the assumption that the take-over request is always perceived. Given the centrality of Perceptual Load Theory to this thesis, I then move to briefly delineate the debate between early selection and late selection theories that Perceptual Load aimed at resolving and I expand on a few central elements of this theory: how perceptual load has been manipulated in the literature, its effect on processing of distractors stimuli – both behaviourally and through neuroimaging studies – and on awareness of distractors.

Highly automated driving and Take-Over Requests

Take-over requests (TORs) in level 3 vehicles can be issued either as a consequence of an abrupt change in the environment, to which the vehicle is not able to respond appropriately (e.g., faded lane markings, merging of lanes in a motorway), or when the vehicle is about to exit a road segment where automation is supported (e.g., when transitioning from a motorway to a suburban area).

The bulk of the literature on control shifting has therefore sought to determine the optimal lead-time that drivers should be given before automation relinquishes control (Zeeb et al., 2015; Gold et al., 2013), as well as establishing how long it takes for the driver to get the back in the loop if the automation is suddenly disengaged (Mok et al., 2015), while a few studies have instead focused on understanding how other factors, like traffic conditions (Radlmayr et al., 2014), age (Clark et al., 2017) and fatigue (Merat et al., 2012) play a role in determining the drivers' ability to resume control of the car.

None of the aforementioned studies have focused on the drivers' ability to detect the TOR per se or on if (and to what extent) TOR detection can be modulated by the complexity and type of attentional demands posed by the non-driving tasks. These tasks have been thus generically viewed as a mean to prevent drivers from paying attention to the road. Simulator studies often contrasted highly automated driving (HAD) - with or without a concomitant secondary task - with manual driving, generally focusing only on a single non-driving task and without varying task difficulty or load (for a recent review: Eriksson & Stanton, 2017).

Some of the tasks typically employed in simulator studies include the Surrogate Reference Task (SuRT; Gold et al., 2015; Lorenz et al., 2014), in which participants are required to detect a large circle among smaller ones, watching videos (Mok et al., 2015), the n-back task (Radlmayr et al., 2014), the “twenty-question task”, in which participants have to guess the identity of an object by asking yes-or-no questions (Merat et al., 2012), cell-phone texting and voice calls (Neubauer et al., 2012). The dependent variables of interest in these experiments are intended to quantify drivers' reaction times to TORs (i.e., time-to-hands-on-steering-wheel, time-to-first-road-gaze), and to evaluate take-over quality by looking at driving performance directly after the control transition occurs (i.e., deviation from the centre of the lane, longitudinal and lateral acceleration).

Very few studies have attempted to clarify whether different types of task-load (perceptual versus WM load) can lead to different take-over RTs and quality, and even these have not manipulated the load within a task. Radlmayr et al. (2014), for instance, employed both the SuRT and a 2-back task under four different traffic conditions. In the high traffic condition, the SuRT led to a higher number of frontal collisions compared to the 2-back task; no differences between these tasks were found in low-density traffic or among all other dependent variables (RTs, longitudinal acceleration and time-to-collision). They argued therefore that performing a more cognitive task that allows drivers to visually monitor the road does not necessarily translate to a better performance than engaging in a visually distracting task.

Zeeb et al. (2016) compared take-over time and quality in two critical scenarios while participants were either watching a news video, writing an email, reading the news, or simply not engaging in a secondary task. In line with their prediction, the reading and video conditions were accompanied by a larger deviation from the centre of the lane in the 10 seconds following take-over, even though no differences were observed regarding the time taken to physically establish readiness to drive (time-to-hands-on and time-to-eyes-on) among the four conditions. The authors therefore argued that non-driving tasks can exert their influence beyond the speed of the initial motor reaction and question the usefulness of considering RTs in isolation.

Overall, the existing literature on take-over performance and authority transition has shown that highly automated driving, by freeing drivers from the need to constantly monitor their surroundings, can lead to diminished situation awareness, making it more difficult to get back in the loop and resume control of the vehicle, especially if drivers concurrently engage in other tasks. The vast majority of studies have compared HAD with and without a concomitant non-driving task (Eriksson & Stanton 2017), and only a few have assessed how different kinds of load can affect the take-over and driving performance right after the transition (Zeeb et al., 2016; Radlmayr et al., 2014). Critically, none of the studies reviewed have addressed the impact that different levels of load for the same task have on the take-over process. Furthermore, all the studies reviewed share a fundamental assumption: that there is optimal perception of the take-over request regardless of the type of non-driving task drivers engage in and its attentional demands. This assumption is central to the design of take-over systems and has safety-critical implications, thus deserving closer scrutiny.

Selective attention and the early vs late selection debate

It would be hard for anyone to disagree with the idea that we cannot be equally aware of all the information that is available at any given moment to our senses, and that some form of selection is called for. We therefore need to prioritize perception of what will aid us to achieve the task at hand and minimize availability of everything that would instead turn as away from what we are trying to achieve. The usefulness of such a mechanism is readily apparent, and yet, unless the perceiver has total control over the environment, it is also easy to see that this mechanism would quickly become dangerous if it were to operate in a rigid, inflexible way. Tasks are also not all alike. Some have well defined boundaries between what is useful and what is irrelevant (e.g. reading a text whilst trying to avoid distractions) while others are fuzzier by definition (e.g. trying to assess if something is out of place in a certain context). Therefore, a delicate balance must always be struck between the ability to focus primarily on something while remaining to some degree permeable to what is happening around us.

While few researchers nowadays would take issue with this very general description of the role of selective attention – that is, the need to operate a selection of the incoming information and the modulation of such selectivity based on both top-down and bottom-up influences – the ways in which this selection is achieved has been the subject of a long-standing debate. In particular, the processing step at which attention intervenes along the putative processing chain from perception to action, as well as the fate of ignored stimuli, have been the subject of contention for decades between proponents of two opposite accounts: early selection and late selection.

A seminal research in the late fifties (Broadbent, 1958), had given widespread credit to the idea that attention is a capacity limited channel, and that a filter of sorts must be in place in order to only allow processing of task relevant stimuli. Research around this time tried to better characterize how this filter operated, spurring researchers to delineate the properties of the stimuli evaluated by this filter. This kind of explanation of the role of selective attention has been termed early selection, as it states that the filtering of what is selected for further processing happens at an early stage, particularly before semantic processing. Stimuli that do not make it past the filter, in this view, do not receive further processing, and only very generic, coarse features of the unattended stimuli are available for report.

Late selection theorist (Deutsch & Deutsch, 1963) on the other hand, observed that although participants might not be able to report the identity of unattended stimuli, these nonetheless exert an effect on performance through, for instance, negative priming, or on physiological measures, such as galvanic skin response. This seems to indicate that all the stimuli are processed at least until the stage of semantic processing and leads to the (somewhat bizarre) conclusion that initial processing of stimuli does not suffer from any capacity limitations. Perception proceeds in an automatic, parallel fashion, with all the stimuli being always entirely processed. The fact that we are not able to fully report the unattended stimuli has little to do with limited capacity, and it is instead due to the fact that attention gates processes at a later stage, such as entry into working memory and response selection. Kahneman and Treisman (1984) noticed how substantial differences between the paradigms employed by early and late selection theory prevented meaningful generalization of the mechanism proposed by either theory. In particular, support to early selection had come from paradigms that overloaded participants with information and required more complex responses - such as the shadowing task - while findings in line with late selection were brought about by the use of tasks in which participants were presented with a very limited number of stimuli. Yantis and Johnston (1990) went one step further as they not only recognized that methodological differences could be responsible for the different pattern of results, but also introduced an important new element. Their claim was that, instead of exerting its influence at a fixed stage in the processing chain, attention could intervene at different stages depending on task demands.

Perceptual load theory of selective attention and cognitive control

In this background, Lavie (Lavie, 1995; Lavie & Tsal, 1994) put forward her Perceptual Load Theory of selective attention (and cognitive control). According to this theory, the level of perceptual load imposed by the stimuli is what determines the locus of attention. When perceptual load is low, targets and distractors alike will be processed. As perceptual load increases though, the amount of processing received by distractor stimuli is reduced or even eliminated. This is a hybrid account, as it is grounded both in the idea of attention as having limited capacity (coherently with early selection) and the notion that whatever capacity we have is automatically employed and cannot be withheld (in line with late selection). It is not possible to achieve selectivity unless capacity is exhausted, as any left-over capacity would spill over to irrelevant distractors. The perceiver can only set processing priorities and ensure that the target will be processed, but has no control over the amount of processing received instead by distractors.

The theoretical novelty of this line of reasoning lies in the fact that the locus of attention is not seen as an intrinsic property of the cognitive system, but rather as a resultant of the interplay between task demands and the ensuing level of perceptual load imposed on the perceiver. Under this framework, instead of being mutually exclusive, findings in line with early and late selection can be thought of as the kind of results that would be expected when the levels of perceptual load are at the opposite extremes of what can otherwise be thought of as a continuum. Indeed, this is what Lavie and Tsal, (1994) found in her review of the literature on the early versus late selection debate. Tasks favouring late selection were characterized by low perceptual load, and manipulations supporting early selection instead as having high perceptual load.

To get a better understanding of what perceptual load is taken to mean we can look at how this construct has been empirically manipulated. Despite the extremely prolific literature on the topic, most of the studies have adopted variations on one of the following three perceptual load manipulations: set size (Lavie & De Fockert, 2003), similarity between target and distractors (Beck & Lavie, 2005; Lavie & Cox, 1997), and different task demands on the same stimuli (Bahrami et al., 2007; Carmel et al., 2011; Cartwright-Finch & Lavie, 2007).

Manipulations of perceptual load via set size or by altering the similarity between target and distractor can be grouped under the more general label of response competition paradigm. They have often used letters as stimuli for both targets and distractors, and represent a variation of Eriksen's flanker task (Eriksen & Eriksen, 1974). Generally, one of two target letters (X or N) is presented centrally or peripherally. When load is manipulated via set size, the target is flanked by one distractor letter in the low load condition and by multiple distractors in the high load condition (either laterally, vertically or even radially). If, instead, load is manipulated by similarity between target and distractors, the target is always flanked by several distractors, with circular placeholders being used in place of letters in the low load condition. In either case, the underlying rationale is exactly the same. In relation to the target, the distractor can be neutral (e.g., W when the target is N), compatible (e.g., N when the target is N) or incompatible (e.g., X when the target is N). Assuming that the distractors are processed, one would expect reaction times to the target to be faster in the case of a compatible distractors and slower in the case of an incompatible distractor, with reaction times to targets associated with neutral distractors falling in the middle. This is in fact what is commonly found under a low load condition. Under high load, however, the above-mentioned difference, indicative of distractors' interference with target processing, is either reduced or eliminated, and this is taken to indicate that under high load distractors have been processed to a smaller extent than they were under reduced load (Beck & Lavie, 2005; Forster & Lavie, 2008; Lavie & Cox, 1997; Lavie & Tsal, 1994).

The other common manipulation sees the uses of identical stimuli for both perceptual load conditions. Here, the manipulation of perceptual load is achieved by asking participants to perform different tasks in the two conditions. Two such studies have used crosses as stimuli and asked participants to indicate which arm of the cross was of a certain colour in the low load condition, and to indicate instead which arm was slightly longer in the high load condition (Cartwright-Finch & Lavie, 2007; Macdonald & Lavie, 2011). Others instead (Bahrami et al., 2007; Carmel et al., 2011; Lavie, 2006; Schwartz et al., 2005) have similarly used crosses as stimuli in association with a different pair of tasks. The crosses could be displayed in six different colours and two different orientations (upward and downward). In the low load condition participants had to respond to every red cross, regardless of the orientation, whereas in the high load condition the task was a more demanding feature conjunction and required participants to respond only to upward yellow crosses and downward green crosses.

Perceptual load theory also draws a clear distinction between perceptual load and general task difficulty. Both high perceptual load and high task difficulty should result in lower accuracy and increased reaction times. Crucially, though, only high perceptual load should concurrently reduce distractor interference. This is indeed the pattern of results found by Lavie and de Fockert (2003), where they compared a high load display with a display subjected to extreme sensory degradation, which rendered the target very hard to discern.

Of course, one might question whether these results cannot be accommodated by a different account, based on distractor suppression. The possibility that the reduced distractor interference under high load derives from a higher level of active inhibition of the distractor cannot be ruled out a priori. This would be at odds with Perceptual Load Theory, as it would entail that distractors are normally processed but suppressed at a later stage (in line with a late selection account). A perceptual load account would instead predict that the distractors are not processed in the first place, as all the available attentional capacity is exhausted by the loading stimuli. Lavie and Fox (2000) addressed this possibility with a negative priming paradigm. Negative priming is the slowing of responses to a target that has recently been presented as a distractor. On the first presentation, response to the probe has to be actively suppressed to protect performance. When the same item is then presented as a target, this earlier suppression carries over to the new trial, determining a delay in the response. This pattern is clearly indicative of distractor perception. In their study, Lavie and Fox used pairs of letter search displays: a first, “prime” display was followed by a second (“probe”) display. Sometimes, the distractor in the prime display became the target in the probe display. In line with their prediction, a slowing of reaction times, which was indicative of negative priming and therefore distractor processing, was found under low load but disappeared under high load.

Within-modality studies of perceptual load and awareness of task-irrelevant stimuli

For the most part, the first studies on perceptual load were not concerned with directly measuring awareness of the distractors, and instead indirectly addressed the extent to which they are processed based on their ability (or inability) to interfere with processing of the target stimuli and, consequently, to slow down response times. Naturally, one might expect that if high perceptual load severely disrupts distractor processing, awareness of the distractors should also suffer. A few studies have therefore explicitly addressed the result of perceptual load manipulation on awareness of task-irrelevant stimuli. Cartwright-Finch and Lavie (2007) have tested awareness of an unexpected stimulus, whose features were therefore not known in advance. Perceptual load was manipulated with a cross task, using a series of crosses with arms of different length and colour. In the low load condition participants were required to perform a simple colour detection task (i.e., indicate which arm of the cross was of a certain colour) whereas a more demanding length discrimination task was adopted in the high load condition (i.e., indicate which arm was slightly longer). Reports of awareness of the critical stimulus, a grey square appearing in the periphery and presented only once in the very last trial, were drastically reduced in the line-length discrimination. These results have been later replicated by Remington et al. (2014) with the arm-length discrimination previously used in the high load condition and a sample that included both children (7-13 years old) and adults. The difference in line length between the two cross arms was varied to achieve different levels of perceptual load. In line with the predictions, younger children had lower awareness of the critical stimulus than older children and adults, with a smaller increase of perceptual load needed to reduce distractor awareness due to their limited attentional capacity. By raising perceptual load sufficiently, the same pattern was observed in the adults group as well.

Both Remington et al. (2014) and Cartwright-Finch and Lavie (2007) have employed an unexpected critical stimulus that appeared only once. In addition, awareness was assessed retrospectively with a surprise question at the end of the experiment, asking participants if they had noticed the appearance of anything else other than the cross in the last trial. As a result, the possibility of an alternative account in terms of rapid forgetting remained open. Macdonald and Lavie (2008) addressed this issue by adopting an expected detection stimulus that was presented multiple times, thus allowing the assessment of perceptual load effects on detection sensitivity and response criterion. In a series of experiments, they manipulated perceptual load within a modified version of the letter search task (Lavie & Cox, 1997) in which the irrelevant distractor letter was replaced by a meaningless shape. Awareness of the distractor was consistently found to be modulated by perceptual load across experiments, together with reduced detection sensitivity under high perceptual load. At the same time, alternative accounts of the results in terms of memory failures or goal neglect (Duncan et al., 1996) were ruled out respectively by having participants report the presence of the critical stimulus online (i.e., right after its presentation and before providing a response to the letter-search task) and by presenting the critical stimulus in 50% of the trials.

Similarly, the level of perceptual load in a central task has been shown to influence perception of temporal patterns (Carmel et al., 2007). The critical flicker fusion threshold is the threshold at which a flickering light is equally likely to be perceived as flickering or as a steady, continuous light. Carmel et al. (2007) have found that, under high load, a light flickering at around the critical flicker fusion threshold is more likely to be perceived as a steady light than it is under low perceptual load.

Finally, reduced distractor awareness under high perceptual load has been demonstrated in a driving simulator study (Murphy & Greene, 2016) with a gap-task. Participants had to judge whether the gap separating two rows of vehicles was wide enough for the driver to pass through and act accordingly by either passing between the vehicles or driving around them. The perceptual load manipulation was achieved by varying the width of the gap between the vehicles while the distractor was either a person or a large animal appearing in close proximity to the road during two trials. Drivers' awareness of the distractor was significantly reduced when the gap between vehicles was narrower (i.e., the high load condition).

Cross-modal studies of perceptual load effect on awareness of auditory stimuli

While the aforementioned studies present an overall strong case regarding the effect of perceptual load on the awareness of irrelevant visual stimuli, the possibility of a cross-modal load effect has not been researched as extensively so far. Understanding the possibility of a cross-modal load effect is important both on a theoretical level - given the ongoing debate over shared versus modality-specific attentional resources - and for its practical consequences, particularly in the case of a visual-auditory effect. Indeed, it is quite easy to imagine situations where failing to detect a sound while engaging on a visual task would be undesirable or outright dangerous (e.g., missing a car horn while looking at one's mobile phone).

Macdonald and Lavie (2011) were the first to address this possibility. Adapting the visual discrimination task (colour detection versus line-length discrimination) used by Cartwright-Finch and Lavie (2007) by replacing the visual distractor with a short tone presented amidst white noise, they showed reduced awareness of the sound under high perceptual load. The results held even when the white noise was removed, and therefore with a higher signal-to-noise ratio. The authors named this phenomenon "inattentional deafness". Similarly to Cartwright-Finch and Lavie (2007), in Macdonald and Lavie (2011) the critical auditory stimulus was unexpected and presented only during the last trial, hence leaving the possibility open that the effect may have been in part due to a higher rate of memory failure in the more demanding condition.

Raveh and Lavie (2015) sought to replicate the “inattentional deafness” effect with a critical stimulus which was expected by the participants and appeared multiple times. They reported a series of experiments where perceptual load was varied within the letter-search task (Lavie & Cox, 1997), and presented a 1025 Hz pure tone in a subset of trials (17% to 50%) at the onset of the visual-search display. Inattentional deafness, as well as reduced sensitivity with high load, was consistently found across experiments while ruling out alternative accounts in terms of memory failure and goal neglect. These reports of inattentional deafness are also in line with an EEG study (Parks et al., 2009). By using the aforementioned cross task, they found perceptual load to modulate the PAR auditory micro-reflex (although later failing to replicate the results with a very similar design; Parks et al., 2011). In a recent MEG study, Molloy et al. (2015) attempted to clarify the neural mechanisms responsible for inattentional deafness. Their results indicated that load has an influence on both early and late components. Indeed, high perceptual load was characterized by an increased vM100 and a reduction in the aM100, together with a suppression of the P3 related to the auditory stimuli. The source of this modulation was localized in areas of the associative auditory cortex, like the superior temporal sulcus and the posterior middle temporal gyrus. These results speak in favour of the idea of shared, modality-independent attentional resources.

Effect of perceptual load on neural processing

If an increased level of perceptual load reduces distractor processing and awareness, it stands to reason that this should also be accompanied by a clear reduction in neural markers of distractor processing. A good body of fMRI results now goes in this very direction. In the first fMRI paper to address this issue, Rees et al. (1997) contrasted peripheral presentation of an optic flow – achieved with dots moving radially outward – with static dots presentation. This allowed them to have an index of motion induced activation in V5. At the same time, participants were engaging with a word processing task at fixation that consisted in responding either to uppercase letters (low load) or to disyllabic words (high load). The results showed the predicted interaction between perceptual load and the type of peripheral stimulation. This means that the increased V5 activity in response to the moving dots (compared to when they appeared static) was only observed under low load. Under high load, V5 activation when viewing moving dots was no greater than when the dots were static, which indicates reduced motion processing. This reduction of activity as a consequence of high perceptual load is not specific to V5 either. Schwartz et al. (2005) manipulated perceptual load in a centrally presented task while irrelevant checkerboards were displayed peripherally. Response to the checkerboards under high load were found to be reduced all the way from V1 to V4.

Bishop et al. (2007) found indications of reduced processing of emotionally salient stimuli, with decreased processing of fearful (versus neutral) faces in the amygdala under high load on the usual letter flanking task. Similar results were reported by Pessoa et al. (2002), who analysed the activation of the amygdala in response to faces of different emotional valence while performing tasks with different attentional demands. Both these studies lent support to Perceptual Load Theory and contradicted the belief that processing of faces and of emotionally salient stimuli might deserve a special status because they appeared to be processed automatically – that is, without the need of attention.

Even the well documented phenomenon of repetition suppression – a progressively minor neural response to a stimulus that is presented multiple times (Henson & Rugg, 2003) - seems to be modulated by task demands. Yi et al. (2004) presented composite visual stimuli consisting of smaller face stimuli at fixation surrounded by larger pictures of outdoor scenes stimuli in the background. Subjects were instructed to ignore the outdoor peripheral scenes and to monitor instead the faces for repetitions. They focused the analysis on the parahippocampal place area (PPA) as it responds strongly to places and only minimally to faces. They found reduced activity when scenes were shown repeatedly. Under high load instead – which was achieved by adding salt and pepper noise to the faces – they found not only a smaller response to scenes, but also no sign of repetition suppression.

Other studies instead have tried to establish what is the earliest point in the visual processing chain where perceptual load exerts an effect. The effect of load could arise even sooner than early visual cortices, and it can be found at a subcortical level according to O'Connor et al. (2002). The lateral geniculate nucleus (LGN), with its afferent connections coming straight from the retina, acts as the main relay of visual input to the cortex and it constitutes the first station that can be subjected to top-down signals affecting visual processing. In their fMRI study, Connor et al. presented peripheral checkerboards while participants were engaged in central tasks of different load (counting infrequent colour changes of the fixation cross under low load, versus counting letters in a RSVP under high load). On top of the load effect in V1, they found significantly greater checkerboard-related signal reduction under high load in the lateral geniculate nucleus.

Criticisms to Perceptual Load Theory

The Perceptual Load Theory of Selective Attention and Cognitive Control was first introduced in 1994 (Lavie & Tsal, 1994). During these almost thirty years, this theory has garnered a wealth of supporting evidence, numerous new applications and, of course, some criticisms. In this section I will discuss in detail two criticisms in particular. The first points to an alternative account, at least in part, of the results obtained in experiments where perceptual load is manipulated with some version of the flanker task. The predictable spatial layout of the stimuli raises the possibility that participants direct their attention differently in the high and the low load conditions and that this different attentional set is in turn responsible for the presence or absence of the flanker compatibility effect commonly reported. The second criticism pertains to the lack of a clear definition of the concept of perceptual load and the risks this creates.

The role of attentional set in flanker tasks

As we have discussed, a good portion of perceptual load experiments has manipulated perceptual load with one variant or another of the flanker task (Eriksen & Eriksen, 1974). Participants search for one of two possible target letters, usually among a group of five non-targets, while trying to ignore a peripheral distractor. The distractor can be congruent, neutral or incongruent with the response afforded by the target and thus respectively facilitate or hinder the response to it. This flanker compatibility effect manifests as a decrease in reaction times in case of compatible distractors and in increase instead when incompatible distractors are shown. It is taken as an index of distractor processing, and the fact that it is reduced or suppressed under high perceptual load (but not with low perceptual load) is explained by perceptual load theory in terms of exhaustion of processing resources or capacity leading to the exclusion of the distractor from further processing.

However, such an arrangement of stimuli as is typical in perceptual load experiments entails a clear, fixed and predictable spatial separation between target and distractors. It therefore suggests the possibility of a contrasting account of the results which could be attributed to the adoption of alternative strategies with different allocation of spatial attention between the two perceptual load conditions. The reduced processing and awareness of distractors could arise not due to the exhaustion of perceptual resources or capacity *per se* but as a consequence of switching from a broader, less focused spread of spatial attention in the less demanding condition to a narrower, more focused allocation in the high perceptual load condition. Under this view, during high perceptual load trials the distractors are processed less because they fall outside the scope of the narrower spread of spatial attention. A few studies have investigated whether the “attentional set” or “zoom” can be responsible for some of the results attributed to perceptual load (Chen & Cave, 2016; Johnson et al., 2002; Theeuwes et al., 2004).

Johnson et al. (2002) devised an experiment which used a slightly modified version of the flanker task, with the addition of a cue factor. As in other experiments that made use of this task (e.g., Lavie & Cox, 1997), participants were looking for one target among six letters arranged in a circle, with an additional distractor in the periphery. The twist here is that trials started without the search stimuli and with a fixation cross lasting either 1000 ms (no-cue trials) or 800 ms followed by a 200 ms cue in the form of an arrow pointing to the location of the upcoming target (cue trials). The cue was either absent or present with equal probability. When present, the cue was always informative (100% validity). The goal was to find evidence of efficient (i.e., early) selection even in a condition of low perceptual load. In fact, another important tenet of perceptual load is that attentional capacity is automatically employed until exhausted, and cannot be voluntarily withheld. This is the reason why, in a low perceptual load scenario, attention necessarily spills over to the distractor, ensuring that it is processed. It follows that high attentional selectivity, which equates to the filtering out of the distractor, can only be achieved when perceptual load is high enough to exhaust the available attentional capacity. Finding that high attentional selectivity can be achieved under low perceptual load would be at odds with perceptual load theory. If perceptual load is low and resources are available, distractors should be processed regardless of the presence of a cue. On top of a main effect of both load and flanker type (congruent, neutral or incongruent) on target reaction times, the results showed an interaction between flanker type and load: the flanker effect was larger under low load than under high load. This is exactly what is expected according to perceptual load theory and the results closely match those reported by Lavie and Cox (1997). Additionally, though, a three-way interaction between flanker type, load and cue (present vs absent) was also found. This means that the presence of a cue caused reduced flanker effect in the low load condition, and this resulted in the significant interaction. Remembering that the strength of the flanker effect indexes the level of processing of distractors, a reduction in this effect under low load when a cue is given means that participants were able to benefit from the cue to perform a more efficient search without processing the distractors as much, even when resources were not likely to be an issue. It is likely that providing a cue and therefore declaring the position of target beforehand creates a high incentive to switch to a more localized attentional deployment. When instead there is no cue and no prior expectations about the location of the upcoming target a less efficient search is performed, taking into account information coming from a broader area.

The plausibility of an account in terms of differential allocation of spatial attention in different perceptual load conditions deserves careful consideration, particularly when other aspects of the experimental design contribute to make this strategy even more advantageous. One such example is keeping the level of perceptual load constant within a block of trials, which is the case in the vast majority of perceptual load experiments. Even though it is not possible to exclude *a priori* that participants decide which attentional set to employ on a trial-by-trial basis as soon as the stimuli are displayed, the short presentation times usually employed make this possibility less of a concern. When, on the other hand, the level of perceptual load is certain to be constant throughout a block of trials, there is suddenly a much greater incentive to set the spread of spatial attention accordingly (i.e., broader with low perceptual load and narrower with high perceptual load). Theeuwes et al. (2004) devised two experiments to test this possibility. Their first experiment was basically a replica of the stimuli and method described by Lavie and Cox (1997), and perceptual load was manipulated between blocks. In their second experiment instead, perceptual load was allowed to vary on a trial-by-trial basis. Additionally, the exact location of the distractor kept changing between a more central and a more peripheral placement, to ensure uncertainty about the position of the distractor, thus making it impossible to filter out a particular location. While the first experiment showed the expected pattern of results, the second experiment failed to find a significant interaction between flanker type (congruent vs incongruent) and perceptual load. In other words, a flanker compatibility effect was found both under low as well as under high perceptual load. They also analysed the flanker congruence effect in relation to the level of perceptual load in the preceding trial. When perceptual load was high both in the preceding and in the following trials there was no congruence effect, in line with the prediction of perceptual load theory. When instead a high load trial was preceded by a low perceptual load trial, selection was inefficient and a flanker effect was indeed found. This pattern, however, did not apply to low load trials. In this case a flanker compatibility effect was always found, regardless of the level of perceptual load in the preceding trial. The authors concluded that perceptual load seems to be the prevailing factor in low perceptual load trials. Instead, they concluded, high load trials are also influenced by the expectations about the upcoming level of perceptual load. Participants tend to expect the same level of load in successive trials, and when the attention is set according to the demands of a low perceptual load trial (i.e., with a broader spread), the same set is then carried over to next trial, even though this might turn out to be a high load trial. The authors remarked that this expectation need not necessarily be a conscious choice, and in fact might very well be an involuntary carry over effect from the previous trial. This phenomenon is referred to as task-set inertia and it is a well-known phenomenon in the study of task-switching (Allport et al., 1994; Wylie & Allport, 2000). It must be noted, however, that other studies have allowed perceptual load to vary between trials and have still observed reduced distractor processing under high load, albeit to a lower extent than is usually found (Macdonald & Lavie, 2011).

Additionally, Linnell et al. (2013) provided an interesting insight derived from the study of the Himba, a secluded population inhabiting the northern part of Namibia. In flanker tasks not too dissimilar from those employed in perceptual load experiments, Himba that were born and raised in Namibia were able to display focused spatial attention with good selectivity at the lowest level of load. On the other hand, both British participants living in London as well as Himba that relocated to western countries showed the usual low selectivity under low load. The authors therefore went on to propose a default attentional state which is not universal and instead differs among various cultures. They reasoned that in fast-paced urban environments it might be advantageous to have a broader and inefficient attentional set, which takes in more contextual information, given that distractors might frequently and unpredictably become targets.

The literature presented here shows that, under the proper conditions, it is possible to observe both efficient selection under low load (Johnson et al., 2002) as well as inefficient selection under high perceptual load (Theeuwes et al., 2004). This entails that factors other than just perceptual load might sometimes contribute to the results observed in flanker tasks, and that care should be taken to minimize the effect of expectations or attentional set when designing perceptual load experiments.

Defining perceptual load

Perhaps the most poignant – and not entirely resolved – criticism pertaining the concept of perceptual load is that it lacks a precise definition. That is not to say that a definition of it has not been attempted by one author or another, but rather that a definition was neither proffered alongside the first experimental findings of Perceptual Load Theory nor one has emerged later on that is largely agreed upon. A lot has been said about perceptual load and what it is not. It does not coincide with a general increase in task difficulty (Lavie & De Fockert, 2003). Its effect should not be confused with negative priming (Lavie & Fox, 2000). It is different from the concept of cognitive load (Lavie, 2005). Yet, while undisputedly useful, all these clarifications do not substitute for nor do they collectively amount to a more formal definition.

The lack of a precise definition raises concerns when comparing the results achieved under different perceptual load manipulations. More importantly, though, this absence creates a risk of circularity. Other than an intuitive understanding of what a high perceptual load condition looks like, perceptual load is usually implicitly explained by illustrating its effect: when high enough, it prevents or reduces the processing of distractors. That is, the definition of what perceptual load is has been implicitely accomplished via the definition of what it does. When the definition of a concept or a variable is tied to its effect on another variable, as it is the case here, the interpretation of the results becomes precarious. For instance, if an experiment fails to find the expect results, the doubt arises of whether this can be taken as a true result or just a consequence of an unsuccessful load manipulation. Clearly, a more precise definition of the concept of perceptual load is needed, as well as an explanation of the mechanism by which it results in reduced processing of distractors.

Torralbo and colleagues (Lavie & Torralbo, 2010; Scalf et al., 2013; Torralbo & Beck, 2008) proposed a neurally plausible explanation of how high perceptual load generates its effect, which also takes care of what might be argued to be another downside of perceptual load theory, namely its reliance on the concept of limited resources. On the surface, it might seem that referring to the idea of resources is a useful heuristic, but their definition is also problematic. Resources have been likened to regulatory juice (Mozer & Sitton, 1998) or a power supply (Kahneman, 1973). None of these metaphors helps in arriving at a clear understanding of what these resources might be, or the extent to which they are shared between different processes.

Torralbo and colleagues start with the observation that often the neural representation of a stimulus is stronger when that stimulus is presented alone as opposed to when it is presented alongside other stimuli in close proximity (Kastner et al., 2001; Reynolds et al., 1999). The magnitude of this difference changes accordingly to the distance between the stimuli: the greater the separation the smaller the difference in activation between sequential and simultaneous stimuli presentation (Kastner et al., 2001). These observations are in line with biased competition models (Desimone & Duncan, 1995) and are thought to be due to local suppressive interactions between populations of neurons in the visual cortex. The result of these interactions is a global weakening of the representations of all competing stimuli. When a clearer stimulus representation is needed in order to guide behaviour and emit a response, attention intervenes and biases the competition in favour of the target at the expense of all other stimuli. However, when attention is spread among multiple stimuli, the resolution of this conflict is more problematic, and requires a top-down bias. In a series of experiments, Torralbo and Beck (2008) showed that manipulations that increase these mutually suppressive interactions, such having both target and distractors in the same visual hemifield, or reducing the separation between target and distractors, both resulted in reduced distractor processing and interference, akin to a high perceptual load scenario.

Therefore, the model they proposed envisions two steps, both of which are needed to explain the effect of high perceptual load. First of all, attention is directed towards multiple stimuli and this creates mutually suppressive interactions between their neural representation in visual cortex. These mutual interactions do not always occur. Instead, their presence depends on the spatial arrangement of the competing stimuli, and are stronger when stimuli are in close proximity. Secondly, this competition is resolved via a top-down bias which has two effects: it enhances the representation of the target and, due to the suppressive interactions with other populations of neurons, results in the suppression of the representations of the competing stimuli, including the distractor responsible for the congruence effect commonly found in perceptual load experiments (Scalf et al., 2013; Torralbo & Beck, 2008).

This can be seen as a refinement of perceptual load theory, as it explains in a plausible fashion just how and why perceptual load might bring upon its effect. It does require a modification of the tenets of perceptual load, as they don't allow for different predictions depending on how the stimuli are spatially arranged – i.e., on whether their arrangement would cause the respective neural representation to overlap and conflict with each other or not. Another consequence of explaining perceptual load in terms of the strength of the top-down bias needed to resolve local competitive interactions is that the exhaustion of resources cannot be seen as the drive underpinning perceptual load effects and thus becomes redundant. There is, in a sense, a limited resource, but this is given by the difficulty in having clear simultaneous representations of multiple stimuli in visual cortex.

Torralbo and Beck (2008) recognize that this account was not intended to explain how the concept of reciprocally suppressive interactions might apply to the other common manipulation of perceptual load, which varies only the task to be performed between conditions while keeping the stimuli exactly the same. These studies often adopt a feature detection task for the low load condition and a feature conjunction task in the high load condition. One possibility, as the authors suggest, is that a mechanism similar to that needed to bias representation in favour of the target is needed when it is necessary to bind features (such as colour and orientation) and not in the case of a simpler feature detection task (Scalf et al., 2013). Of course, another possibility is that we need to attribute the effect of perceptual load when the latter manipulation is employed to a different, albeit unspecified, process.

Roper et al. (2013) tried instead to arrive at a definition of perceptual load via another route. They explored the extent to which a parallel can be drawn between factors that influence search efficiency in canonical visual search tasks and those that determine the level of perceptual load in canonical perceptual load studies. Two factors in particular have been known to influence search efficiency, namely T-D (i.e., target-distractors) similarity and D-D similarity (i.e., distractors-distractors) similarity (Duncan & Humphreys, 1989) . They found that most perceptual load experiments confound these two factors, when considering as distractors also the non-targets which are not competing for response selection. Low perceptual load conditions usually have low T-D similarity and high D-D similarity. On the other hand, high perceptual load conditions have high T-D similarity and low D-D similarity (when compared to low load conditions). Therefore, in order to properly assess the potentially separate contributions of these two factors, they should be manipulated independently of each other. The results of a set of three experiments demonstrated that the same stimulus set which generated an inefficient visual search, as indexed by a steeper search slope in the visual search task, generated high perceptual load in the perceptual load task with the ensuing high attentional selectivity, as indexed by the lack of flanker congruence effect. Conversely, the stimuli that afforded a more efficient visual search also showed the mark of low perceptual load – low attentional selectivity with the presence of flanker congruence effects. While D-D similarity played a minor but quantifiable role, the level of T-D similarity was shown to be by far more important (about 4.5 times) in predicting the level of perceptual load.

Clearly, attempts have been made at defining the concept of perceptual load independently of its effect on distractor processing. More effort will certainly be required in the future to arrive at a precise and shared understanding of what perceptual load is. This definition should assign equal importance to all perceptual load manipulations, not just those based on variations of the flanker task but also those accomplished by performing different tasks on identical stimuli.

To summarize, Perceptual Load Theory has been widely applied over the last two decades to explain the extent to which task-irrelevant stimuli are processed. Conditions of high perceptual load have been shown to reduce processing and awareness of distractors, both behaviourally and in neuroimaging studies.

Over the following three chapters I consider how the conceptual framework afforded by Perceptual Load Theory can help evaluate the impact of non-driving tasks on the drivers' ability to take back control of the vehicle. A series of laboratory experiments will be presented, with tasks chosen to be appropriate proxies for both the take-over request and the non-driving tasks. Two different scenarios will be evaluated in turn. In Chapter 2 I consider how the effect of high perceptual load generated by real-world pictures (such as those that would be encountered while, for instance, on social media) might affect the driver’s ability to perceive the take-over request emitted by the vehicle. In Chapter 3 I instead focus on the effect that modulation of attentional demands before a task-switch has on response times to a task that requires to discriminate motion, as the driver would be expected to when returning to the driving task. Chapter 4 aims to clarify how the behavioural results found in Chapter 3 can be related to the processing of motion in V5 by means of an fMRI experiment.

Chapter 2:
Inattentional deafness and perceptual load in natural scenes perception

Introduction

The first step to ensure a prompt response to the take-over request emitted by the vehicle is of course to make sure that the driver perceives the TOR in the first place. The hypothetical scenario investigated in this chapter sees the driver engaging in a fairly common non-driving task (i.e., looking at pictures) while having to detect an infrequent auditory TOR presented amidst a noisy environment. As it will be detailed below, there is a strong practical and theoretical rationale for the selection of this specific non-driving task. On one hand, tasks that include looking at pictures (or, more, formally, natural scenes) were the top picks of a poll realized for this very purpose (Sullman et al., 2017). On the other hand, natural scenes have been the centre of a heated debate between proponents of the idea that their perception does not require attention and others that see no grounds to grant them this special status.

For this reason, I start by recounting the debate surrounding natural scenes perception and then move on to analyse studies exploring cross-modal effects of perceptual load - i.e., the possibility that visual perceptual load modulates perception of auditory stimuli. Two experiments are then presented that explore the aforementioned scenario while varying the temporal relation between the onset of the natural scenes and that of the auditory TOR.

Attentional demands of natural scene processing

Several findings have accrued over the past decades to indicate that natural scenes might deserve a special status inasmuch as their processing seemingly does not require attention (e.g., Mack & Rock, 1998; Potter, 1975). Thorpe et al. (1996) employed a go/no-go categorization task in which participants had to detect the presence of an animal in outdoor natural scenes after a viewing time of only 20 ms. Despite the short presentation time, participants performed the task with almost perfect accuracy. By comparing ERPs to correct go trials (animal present) with ERPs to correct no-go trials (animal absent) it was observed that they start to diverge at around 150 ms, concluding that by that time, too short for attention to play an important role, natural scenes are processed to such an extent as to reliably assess the presence or absence of a target. VanRullen & Thorpe (2001) replicated these behavioural results with natural scenes depicting animals as well as vehicles (presented for 20 ms) and found detection accuracy close to 95% for both categories. Their results seemingly indicated that the high efficiency in processing natural scene is not confined to categories of inherent biological significance (such as animals) but could in fact be a more general mechanism.

Fabre-Thorpe et al. (2001) consequently went on to propose the existence of a ultra-rapid visual categorization mechanism for natural scenes, which cannot be sped up even after extensive training (Fabre-Thorpe et al., 2001) and does not require foveal vision (Thorpe et al., 2002), again calling into question the involvement of attention. Similarly, Kirchner & Thorpe (2006) presented two scenes at a time (for 20 ms), one of which depicting an animal, and instructed participants to make a saccade toward the target scene. While the median RT was 228 ms, correct responses significantly outnumbered incorrect responses after as little as 120 ms.

The claim for attention-free natural scene processing seems to be strengthened by the finding that animal detection in natural scene does not suffer from the addition of a concurrent task. In a series of experiments, Li et al. (2002), had participants perform a central task and three different peripheral tasks, each under single and dual-task conditions. As a central task, participants had to detect whether five randomly rotated letters (Ts and Ls) were identical or if one differed. One of the peripheral tasks was the go/no-go animal detection task with natural scenes (displayed for 67ms). The remaining two tasks were both discrimination task either on either Ls and Ts or vertically bisected disks. Performance on the scenes task did not differ between single and dual-task condition, and neither did performance on the central task (indicating that participants were not diverting attention from the central task). This finding, in conjunction with the fact that performance on the other two peripheral tasks dropped to chance levels under dual task condition, led the authors to conclude that natural scene perception can be performed without attentional involvement.

Rousselet et al. (2002) have also argued against the need of attention in the processing of natural scenes on the grounds that the speed of processing is unaltered when viewing one versus two images simultaneously. They modified the previously mentioned go/no-go task with animals (Thorpe et al., 1996) by presenting either one or two images. No differences were found in median reaction times and latencies of early responses. ERPs results replicated the enhanced negativity on occipito-temporal sites for target-present trials found by Thorpe et al. (1996), but this difference was not found to be modulated by the number of images presented. Indeed, in both conditions the difference in ERP waves between target and distractor trials became significant at identical times (152 and 150 ms with, respectively, one and two images) and subsequently developed with the same slope and reached analogous amplitudes. However, a later study with one, two and four simultaneously presented images, reached a different conclusion (Rousselet et al., 2004). Behaviourally, RTs, accuracy and speed of processing worsened linearly with increasing number of scenes. Differential ERPs between distractor and target trial at occipital sites did not differ between one and two images in terms of peak time and amplitude, replicating previous findings (Rousselet et al., 2002). With four images, instead, the amplitude of this differential ERP activity reduced significantly, leading the authors to conclude that animal detection in natural scene might not be entirely performed in parallel.

Two alternative accounts have since been proposed, challenging the idea of pre-attentive processing for natural scenes. The first, proposed by Evans and Treisman (2005), is that detection of a target belonging to a certain category (i.e. animals, vehicles) can actually be mediated by the perception of sets of disjunctive features that, in turn, activate all the high-level nodes they are compatible with. For instance, a beak is indicative of the presence of a bird; similarly, detecting features compatible with the shape of legs is sufficient to infer the presence of a wide range of animals. Coherently with Feature Integration Theory (Treisman & Gelade, 1980), they reasoned that detection of such features could indeed be performed pre-attentively, but that feature binding, necessary to fully identify the target, would not be possible without attention. Indeed, in a series of experiments they displayed natural scenes using rapid serial visual presentation (RSVP) for 75 ms and asked participants not only to report the presence of a target (which, depending on the condition, could either be an animal or a vehicle), but also to identify the subordinate category to which it belonged (e.g., bird, fish, mammal are all subordinates of the superordinate animal category) and to localize it. Results showed that while correct detection rates for animals and vehicles were respectively 73% and 74%, targets were poorly identified and localized (53% for animals and 56% for vehicles). What's more, when people (who share more features with the superordinate animal category than with vehicles) were included as distractors, detection was significantly reduced when participants were looking for animals but not when the targets were vehicles. Hence, the possibility that a simple go/no-go detection could be performed via the detection of simple features and without full identification of the target seems very plausible and speaks against claims of pre-attentive processing of natural scenes.

Furthermore, Walker et al. (2008) noticed how the vast majority of images employed in previous ultra-rapid categorization studies had the target animal as the only (or one of two) foreground object. This high regularity in the structure of the images makes it easy for saliency-based mechanisms to rapidly orient attention to the only relevant part of the scene, substantially decreasing the complexity of the stimulus. They reasoned that if all that was required in URC experiments was to scan one or two items, these findings would hardly pose a challenge to conventional theories of high-level vision. Instead, they employed scenes with four superimposed objects, equally distributed between the foreground and the background. The usual animal detection task (Thorpe et al., 1996) was performed under single and dual-task conditions with much longer presentation times (170 to 500 ms) compared to Thorpe et al. (1996). Two different tasks were employed alongside the animal detection task. In a VSTM (visual short-term memory) task four letters were superimposed on the centre of the scene for 125ms, disappeared for 250 and reappeared for the last 125ms during which the scene was presented, and participants had to detect if one of the letters had changed position. In the other task the letters were displayed only once for 120ms and participants instead indicated if one of them was a vowel or not. Under dual-task conditions, both the VSTM task and the selective attention task with no memory involvement led to significant reduction in accuracy for the animal detection task. What's more, the focal attention task (a vowel detection task) significantly reduced animal detection accuracy compared to a single-task condition even for scenes with only one foreground object. The authors explain this discrepancy with Li and colleagues' findings by arguing that the vowel-detection task was more demanding than the letter task chosen by Li et al. (2002).

This alternative account of the attentional demands posed by natural scenes is also in line with Cohen et al. (2011), who argue that previous results can be explained without postulating any kind of pre-attentive processing, by simply conceding that natural scene processing is highly efficient, thus requiring very little attention. It follows that impairment of scene processing should be observed in a dual-task setting provided that the attentional demands of the concomitant task are high enough. In a series of experiments, they employed RSVPs of coloured checkerboards, where the second to last item of each RSVP was unexpectedly replaced by an image during one critical trial. All the while, participants were either engaged in a superimposed multiple object tracking (MOT) task or asked to count the occurrence of numbers in a stream of letters that were time-locked to the change of the checkerboards in the background. Under dual task conditions, the majority of participants showed inattentional blindness for the image under the MOT task and half of them also did not perceive the scene while performing the number task. On the contrary, when instructed to attend only the background stream of checkerboards, participants performed close to ceiling. In a similar experiment, with images presented on 50% of trials, they manipulated the speed of the dots for the MOT task. As opposed to the single task image classification, no decrement in accuracy was observed under dual- task condition at lower dot speed (which, they argued, could be mistakenly interpreted as evidence of pre-attentive scene processing) but accuracy indeed deteriorated when the dots moved faster. Natural scenes processing seems therefore to be susceptible to interference from other tasks under sufficiently taxing conditions.

Additional support to the role of attention in natural scenes processing, though from a different perspective, comes from Peelen et al. (2009). In a fMRI study, they presented four scenes at a time for 130 ms, two on each axis, and instructed participants to look for a target category (people or cars) in two scenes (by cuing which axis to attend). Response patterns to each category in a high-level visual area, the object selective visual cortex (OSC), were correlated with response patterns evoked in the same region by a separate set of images of people and cars in isolation. Two findings are of interest. Firstly, there was a significant difference between within-category and between-category correlations of activation pattern, indicating that patterns of activity in OSC conveyed information about the attended category. In other words, when participants were looking for, e.g., people (and people were present in the scenes), the activity pattern in the OSC correlated more with the pattern evoked by presenting people in isolation than it did with the car-related activity pattern. Centrally, attention plays a critical role given that OSC carried only information related to the target category. Secondly, this attentional mechanism favouring the target category operates globally in the visual field, biasing the processing of both attended and unattended images. The role of attention and object-selective cortex in visual search was also addressed in another fMRI study (Peelen & Kastner, 2011). Again, participants were looking for people or cars (in a single image), with cues given before each trial. On one third of the trials, no scene followed the cue, allowing the investigation of neural activity while participants prepared for the upcoming search. The degree of category-specific activation pattern in OSC during the pre-search window was found to positively correlate with search accuracy. Peelen and Kastner (2014) argued in favour of a top-down biasing signal that results in enhanced processing of stimuli matching a search template. Such a template, evolved via constant exposure to the target category, and stored in the OSC, would have to consist of intermediate level features (i.e., a wheel when looking for a car) in order to be general enough as to be useful when looking for a target whose precise appearance cannot be anticipated (due to different possible points of view).

In short, while detecting the presence of an animal or vehicle in a scene can indeed be performed with remarkable speed and accuracy, it does not necessarily follow that it can be done without any attentional involvement (Evans & Treisman, 2005; Peelen et al., 2009). Furthermore, studies supporting an attention-independent mechanism for processing natural scenes have typically employed outdoor images where the animal is often presented in isolation and in the foreground, therefore becoming highly salient and free of the competition which would derive from having other objects in the scene. More recent studies have established that when natural scenes with multiple objects are employed (Walker et al., 2008), or under sufficiently taxing dual-task conditions (Cohen et al., 2011), scene processing performance rapidly deteriorates, findings that appear quite hard to reconcile with a view of natural scenes processing that proceeds in the absence of attention.

Cross-modal studies of perceptual load effect on awareness of auditory stimuli

As discussed in the previous chapter, high perceptual load has been found to reduce processing and awareness of distractor stimuli. Its effect on processing of distractors has been observed indirectly, by measuring differences in response times in flanker-like experiments between displays including coherent and uncoherent distractors at different levels of load (Beck & Lavie, 2005; Lavie & Cox, 1997; Lavie & De Fockert, 2003). Such a difference in reaction times is typically observed under low perceptual load (suggesting that distractors have been processed) but not under high load. A more direct effect of perceptual load on awareness of distractors has also been shown in numerous studies (Carmel et al., 2007; Cartwright-Finch & Lavie, 2007; Macdonald & Lavie, 2008; Remington et al., 2014). Crucially the effect of perceptual load on awareness is not only observed within the visual modality. Macdonald and Lavie (2011) were the first to show the phenomenon referred to as “inattentional deafness”: high perceptual load decreases the probability of detecting an auditory stimulus. This result was later replicated by Raveh and Lavie (2015) with an auditory stimulus that, unlike in Macdonald and Lavie (2011), was both expected and presented multiple times, even as often as in 50% of the trials.

Of course, the fact that increased perceptual load might affect processing of auditory stimuli carries not only relevant theoretical implications for the debate between modality-specific versus shared attentional capacity, but also – and more importantly for the present discussion – practical ones in terms of safety in a number of scenarios.

Two experiments are reported in this chapter. Experiment 1 tries to establish if processing natural scenes of different complexity and perceptual load can lead to reduced perception of a pure tone presented amidst white noise – intended as a proxy for the take-over request issued by the car in a noisy environment. Importantly, the onset of the tone is fixed relatively to the onset of the images and always occurs after 200 ms. Indeed, reduced tone perception is observed when viewing images of high perceptual load. Given the fixed and close succession between the onset of the tone and that of the images in the previous experiment, Experiment 2 asks if such a strict temporal relation is needed to attain reduced tone perception under high perceptual load. In this experiment the onset of the tone is varied continuously over a wider time frame, while replicating the results of Experiment 1.

Experiment 1

This experiment aims to extend previous reports of the effect of perceptual load on awareness of task-irrelevant auditory stimuli (i.e., inattentional deafness) to a novel and more ecological task involving a different class of stimuli: natural scenes. The rationale for the choice of natural scenes is twofold. Firstly, it relates to the overall aim of the research project, which is to better characterize the effect of different levels of perceptual and cognitive load of the non-driving tasks on the driver's ability to respond to take-over requests. An online survey (with 277 respondents) was conducted by Sullman et al. (2017) to understand what kind of tasks people believe they might engage in while riding in a self-driving vehicle. The survey contained the following question: "if you were to drive in a vehicle that can drive itself for some part of the journey (in autonomous vehicle mode), which of the following activities would you like to engage in instead of driving?". Respondents were asked to choose as many activities as they wished from a wide range of options and to indicate how many minutes they would spend on each of the selected activities. A *count* measure and an *interest* measure (compounding both count and minutes spent on each activity) were then derived. According to both the *count* and the *interest* measure, the three non-driving tasks that came out on top were: browsing the internet, reading and keeping up with the news (Figure 1). Images are constantly encountered during internet browsing (particularly on social media and news websites), and constitute therefore an important part of activities that respondent pictured themselves doing during a journey in a self-driving vehicle.

Secondly, previous studies on inattentional deafness have typically varied load within either a letter-based visual search (Raveh & Lavie, 2015) or a cross-based feature detection and length discrimination task (Macdonald & Lavie, 2011). Consequently, it would be of interest to see if inattentional deafness can be generalized to load manipulations involving more complex and ecological stimuli that are more likely to be of relevance in our everyday life.

In the present experiment, participants attended a sequence of natural scenes of either low load or high load. The natural scenes were assigned to their respective perceptual load condition according to a complexity rating which was calculated in a separate study (Nagle & Lavie, 2020; see *dataset generation* for further details). On 30% of trials the images were followed by a single word probe referring to one of the object categories depicted in the full dataset. Participants had to indicate whether the probed category was present or absent in the image immediately preceding the appearance of the probe. White noise was played constantly throughout the experiment in order to make the experimental setup more akin to a driving scenario, in which the driver is constantly exposed to a multitude of auditory stimuli. Additionally, 10% of the images were accompanied by a pure tone, which participants had to detect and report right after its presentation.

This 'post-cue visual search task' required participants to fully process each image, given that the identity of the target was not revealed until after image offset. If the complexity rating indeed reflected perceptual load (and if natural scenes processing required attention), we would expect participants to perform worse, with lower accuracy in the high load condition. While performing the visual search task on high load images, less attentional resources should be available to process the auditory tone compared with the low load condition. Hence, our main hypothesis, in line with previous findings on inattentional deafness and Load Theory, is that the tone detection rates should depend on load – with lower detection rates in the high load condition.

Design

The experiment used a within-participants design with a single factor. The independent variable was the level of perceptual load (low, high) assigned to the images in the main task. The dependent variables related to image task were the accuracy in judging whether the image contained the probed category as well as the reaction times. The dependent variables for the tone detection task were a series of measures related to the detection of the infrequent auditory tone that was sometime presented while performing the image task. These measures were: correct detections, correct rejections, false alarms, misses and reaction times. These were in turn used to compute detection sensitivity and response bias.

Dataset generation

The experiments described in this chapter draw upon a common dataset, generated as part of a separate study (Nagle & Lavie, 2020): a collection of varied, real-world images with human-generated complexity ratings. All the images were taken from the PASCAL VOC (visual object challenge) object recognition dataset, which consists of natural scene images containing one or more categories from the following 20 classes, which are constructed around the *person*, *animals*, *vehicles* and *indoor* macro classes:

- person: person;
- animals: bird, cat, cow, dog, horse, sheep;
- vehicles: aeroplane, bicycle, boat, bus, car, motorbike, train;
- indoor: bottle, chair, dining table, potted plant, sofa, tv/monitor.

Data Collection. Participants (N=100) were shown a pair of images and instructed to choose, from each pair, the image they thought was most visually complex. The experimenters did not answer any questions concerning which features should be taken to indicate complexity, telling observers that they should make the decision themselves.

Post-processing. These comparisons between image pairs were used as input to the TrueSkill algorithm, which uses a Bayesian framework to assign a score to each image based on its performance in each competition (i.e., each time that image was evaluated against another image). For the purpose of these studies, images that had obtained a complexity rating ranging from 8 to 21.3 (corresponding to the 25th

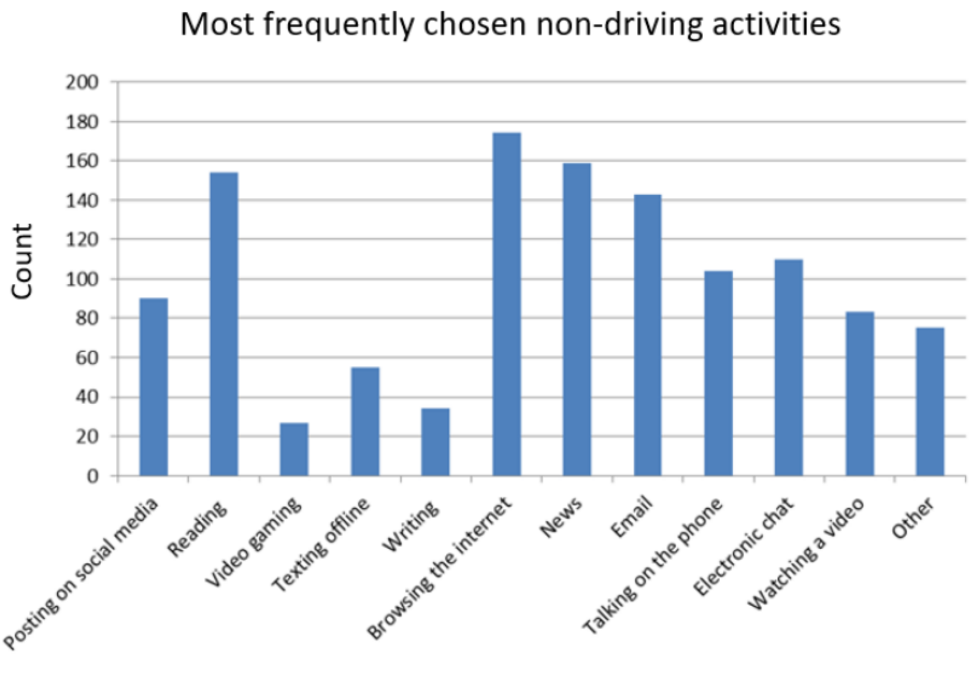

Figure 1. Number of choices for each non-driving task in Sullman et al. (2017). Browsing the internet, keeping up with the news and reading are the most frequently chosen activities.

percentile of the complexity rating distribution) were assigned to the low load condition, while images with a complexity rating ranging from 28.7 (corresponding to the 75th percentile) to 42 were assigned to the high load condition. A subset of 480 images was hence selected so as to have 20 images per level of load for each of the following 12 categories: person, bird, cat, dog, horse, aeroplane, car, boat, potted plant, chair, bottle, sofa.

Participants

Twenty-four participants were recruited for this experiment (mean age: 20.7 years, range 18-24; 18 females). Recruitment was carried out through the UCL SONA platform. All participants reported normal or corrected-to-normal vision and normal hearing. Participation was compensated with £5, which increased to £7 if the accuracy to the cross task was 90% or higher. In compliance with the declaration of Helsinki, participants gave their written informed consent prior to the beginning of the experiment.

Apparatus and stimuli

The experiment was created with MATLAB (Mathworks, Inc.) and run on a PC using windows 7 Enterprise with a 24” monitor. Audio stimuli were created and presented with a pair of Sennheiser 240 headphones. A total of 480 unique natural scenes, selected from the PASCAL VOC dataset, were employed in this experiment. The images comprised both animate and inanimate items and they were set both indoor and outdoor. Images were equally divided between low load and high load. Each image could depict one or more exemplars from a range of twelve categories. On 30% of the trials (n=144), the image was followed by a single-word probe (with the exception of the potted plant category), corresponding to one of the aforementioned categories. The probed category was present in the preceding image on 50% of these trials. In the remaining trials followed by a probe, a random image was provisionally selected for each probe. If none of the categories present in the image matched the probed category, the image was accepted as a target-absent trial for that category, otherwise a new random image was selected and evaluated. The probes were counterbalanced over load level and category, so that each category was probed six times for each load condition. Each image was presented once, for a total of 480 trials.

The size of the images varied from a minimum of 115 by 400 pixels to a maximum of 500 by 500 pixels.

In addition to the probe, 10% of images (24 per load condition) were accompanied by the critical stimulus, a 500 Hz pure tone played at 28 dB (over ambient noise) with a 50 ms duration. The onset of the tone was fixed at 200 ms after the onset of the image. Tones and probes positions within a block were randomized, with the following constraints. Firstly, a two-trials gap was enforced between successive critical trials (i.e., trials containing a tone). Similarly, a two-trials gap was introduced between critical trials and probe trials. Images followed by a probe were not barred a priori from also being paired with a tone, and this was randomly determined. Lastly, a tone, was never played during the first and last trial of a block. Each low load block was paired with a high load block, which received the same randomization in terms of tone and probe position as well as the same ordering of absent and present probes.

Procedure

Each block consisted of a sequence of eighty images, with each image presented at the centre of the screen for 1000 ms. White noise was played constantly throughout the block via the headphones at 48 dB (over ambient noise). A 300 ms gap, filled by a uniform grey background, separated one image from the next. The stream of images was interrupted at random intervals by a single-word probe (appearing 100 ms after image offset). Participants were required to provide a present/absent response by considering only the last image before the probe was displayed. Participants pressed ‘q’ if they believed the probed category to be present in the previous image or ‘p’ to indicate that the probed category was absent. The probe remained visible until participants made a response. Feedback was provided in the form of a one-second buzzer sound if the response provided was incorrect. The image sequence resumed 1.5 seconds after the present/absent response was made. In addition to the visual search task, participants were asked to detect and report the presence of a 50 ms pure tone (played on 10% of trials) by pressing the spacebar as soon as they heard the tone. The experiments comprised 6 blocks, with 3 blocks per load condition. Perceptual load was kept constant in a block and changed at the beginning of each new block. The block order was counterbalanced across participants by having odd-numbered participants starting with a high load block and even numbered participants with a low load block instead. Task instructions emphasized accuracy over response speed. Additionally, to make sure that participants fully engaged with the visual search task, they were warned prior to the beginning of the experiment that points would be assigned for each correct response (to the visual search task), and that they had the chance to earn an additional monetary reward if the number of points collected at the end of the experiment exceeded the threshold specified in the instructions, which corresponded to 90% accuracy. At the end of each block, participants were presented with a message indicating the number of points earned during that block along with a reminder of the average number of points they needed to earn in each block to receive the additional monetary reward. Before the main experiment, participants were played the tone twice without the white noise and then went on to complete twelve practice trials, with a total of four probes (two present and two absent) as well as four tones. The whole experimental session took approximately 30 minutes.

Results and discussion

For both the visual search and the tone detection task, reaction times falling below -3 SDs or above +3 SDs from the mean were discarded. This led to an average rejection of 3% of data points for the visual search task and 1.6% of data points for the tone detection task. One-tailed paired-samples t tests were conducted on the mean accuracy rates and reaction times for both tasks.

Visual search task. Trials could potentially contain both a tone and a probe, and in order to assess the effectiveness of the load manipulation, only trials without tones (or false alarms in the tone detection task) were considered in the accuracy analysis.

A significant difference was found between the accuracy in the low load condition ($M= 93.4\%$, $SD= 4.8\%$) and the accuracy in the high load condition ($M= 82.9\%$, $SD= 5\%$): $t(23) = 10.75$, $p< 0.001$, $d= 2.19$ (one-tailed). Similarly, a one-tailed t test on correct reaction times showed that participants were faster in judging the presence or absence of the target in the low load condition ($M= 1004$ ms, $SD= 274$ ms) as opposed to the high load condition ($M= 1078$ ms, $SD= 352$ ms): $t(23) = 2.82$, $p< 0.01$, $d= 0.57$. These results confirmed that the visual search task was effective in manipulating the level of perceptual load while ruling out a speed-accuracy trade-off.

Tone detection task. For the auditory detection task, paired-samples t-tests were carried on the mean detection rates and RTs. With regard to the detection rates, a significant difference was found between the low load condition (M=76.1%, SD=29.7%) and the high load condition (M = 71.9%, SD = 33.8%): t (23) = 1.9, p = .03, d = 0.3. No difference was found for the reaction times between low load (M = 739 ms, SD = 109 ms) and high load (M = 735 ms, SD = 143 ms): t (22) = 0.1, p = .9. Detection sensitivity for the tone was assessed with A given the very low false alarm rates (less than 1%; Zhang & Mueller, 2005). Detection sensitivity was reduced under high load as opposed to low load: t (22) = 1.7, p = .05. No effect of load on response bias (β) was observed (low load: M = 2.22, SD = 2.3; high load: M = 2.9, SD = 3.9; t (22) = 1.68, p = .1).

These first results provide support to the validity of our load manipulation via complexity rating and to the claim that processing natural scenes does indeed require attention, given the lower accuracy and slower reaction times under high load. More importantly, the results are in line with the hypothesis that perceptual load is capable of modulating auditory detection of a pure tone – specifically, that high perceptual load leads to lower tone detection accuracy. Additionally, reduced detection sensitivity was found under high load, with no difference in response bias (β). The absence of a significant difference in the RTs to the tone detection task between high and low load is not contrary to our hypothesis, given that task instructions emphasized accuracy over speed on both tasks. Nonetheless, in this study tone onset always occurred at 200 ms from image onset, which precludes any conclusions on the time-course of inattentional deafness. Responses to tone made before responses to visual search rules out alternative account in terms of memory failure and rapid forgetting (Moore, 2001; Wolfe, 1999).

Experiment 2

The previous experiment employed a fixed onset time for the auditory tone relative to the onset of the images and did not therefore allow an assessment of the time course of the load effect on tone detection accuracy. In their EEG investigation of inattentional deafness, Molloy et al. (2015) argued for a time-sensitive effect of perceptual load on tone awareness. Therefore, in this experiment, the tone onset time was randomly scattered throughout the image presentation time in order to understand whether, with this type of visual search task, the effect of perceptual load is time-sensitive and confined to a short time following image onset or whether it can exert its influence over a wider timeframe.

Design

The experiment employed a 2X4 within-participants design. The factors were perceptual load (low, high) and tone onset time. The latter was delivered as a continuous variable and grouped into four discrete bins during the analysis. In other words, tones could have any onset time between 50ms and 950ms (after image onset), and the onset time was then discretized into four bins, grouping together tone onsets in the 50-249ms, 250-499ms, 500-749ms and 750-950ms range respectively. The dependent variables were identical to those reported in the previous experiment for both the image task and the tone detection task.

Participants

Sixteen participants were recruited for this experiment (mean age: 24.3 years, range 19-34; 10 females). All participants reported normal or corrected-to-normal vision and normal hearing. In compliance with the declaration of Helsinki, participants gave their written informed consent prior to the beginning of the experiment.

Apparatus, stimuli and procedure

The apparatus, stimuli and procedure were identical to those of the previous experiment, except for the following. While the proportion of critical trials was maintained at 10%, each image was presented twice, making for a total of 960 trials with 48 tone-detection trials per load condition. Furthermore, the percentage of images followed by the category probe was reduced to 15%. The experiment was divided into 16 blocks of 60 images each. Images started repeating after the whole image dataset had been presented once, meaning that no image was repeated within the first half of the experiment nor within the second half. Randomization of image sequence as well as probe position within a block was conducted separately for the first and the second half of the experiment. Odd-numbered participants received a block sequence following a LHHL HLLH pattern while even-numbered participants received the opposite block order. The experiment was preceded by a practice block on a set of 160 images, which were different from those used in the experiment.

Additionally, given that each block contained 9 probes, the number of present probes (i.e., probes for which the correct answer was “present”) and absent probes was counterbalanced across blocks: the first low load block received 4 present probes and 5 absent probes while the second low load block had instead 5 present probes and 4 absent probes (the same applies for each pair of high load blocks). The overall number of absent and present probes was counterbalanced across the first and second half of the experiment: during the first randomization cycle odd numbered categories (from the list of all 12 categories) received 2 present probes and 1 absent probe, while even numbered categories received 1 present probe and 2 absent probes. The opposite criterion was adopted while randomizing the second half of the experiment. Ten percent of images (48 per load condition) were accompanied by the critical stimuli, which were with equal probability either a 1500 or 2000 Hz pure tone with a 50 ms duration. The onset of the tone relative to the onset of the image was randomly determined in a 0-950 ms range. Tone onset, tone frequency and position of tones within a block were randomized independently for the first and the second half of the experiment. To prevent learning effects, images that were paired with a tone in the first half of the experiment were never paired with a tone in the second half, but were allowed to be paired with a category probe. Conversely, images that were paired with a probe in the first half of the experiment could not be paired with another probe in the second half, but could instead be selected as critical trials.

Results and discussion

For the visual search task, paired-samples t tests were conducted on the mean accuracy rates and RTs as a function of the level of perceptual load (two levels: low, high). Again, only trials without tones (and without false positives concerning the tone) were included in the analyses of visual search accuracy and RTs. Furthermore, only trials in which the correct response was provided to the visual search were considered in the RT analysis. A significant difference was found between the accuracy in the low load condition (M = 91.4%, SD = 4.9%) and in the high load condition (M = 77.9%, SD = 4.4%): t (15) = 12.22, $p < .001$, d =3. Similarly, a paired-samples t test on RTs shows that participants were faster in judging the presence or absence of the target in the low load condition (M = 1132 ms, SD = 308 ms) as opposed to the high load condition (M =1176 ms, SD = 285): t (1,15) = 2.38, p = .03, d = .59. These results confirm that the visual- search task was effective in manipulating the level of perceptual load.

As for the tone detection task, in order to assess the time-course of the potential load effect, a time factor was computed so that each level corresponded to the tone detection responses provided during 250 ms bins (level 1: 0-250 ms; level 2: 251-500 ms; level 3: 501-750 ms; level 4: 751-950 ms). A two-way ANOVA was performed with load as the first factor (low, high) and time as the second factor (on four levels; Figure 2). A main effect of load was found (F (1, 15) = 20.2, $p < .001$, $\eta^2 = .57$), with lower detection rates in the high load condition (M = 71.8%, SD = 17.2%) compared to low load (M = 78.5%, SD = 17.2%) as well as a main effect of time (F (3, 45) = 5.1, $p < .01$, $\eta^2 = .25$). Post hoc comparison (with Bonferroni correction) indicated that time 2 was significantly different from time 3 ($p <.001$) and that time 3 differed from time 4 ($p < .01$). The interaction between the factors Load and Time was not significant (F (3, 45) = 0.55, p = 0.6). Detection sensitivity, assessed with A, was significantly reduced during high load as opposed to low load: t (15) = 4.9, $p < .001$. Response bias (β) was found to be significantly higher for high load (M = 2.11, SD = .68) compared to low load (M =1.82, SD = .66): t (15) = 5.3, $p <.001$.

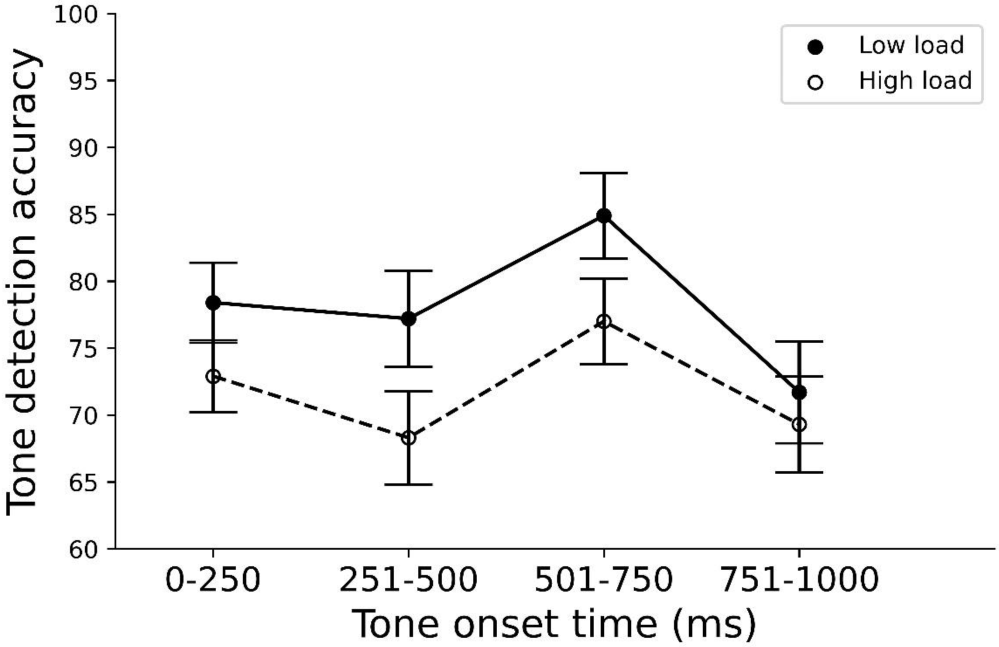

Figure 2. Accuracy to the motion detection task at each level of load (low, high) and discretized tone onset times (0-250, 251-500, 501-750, 751-950 ms).

Two main conclusions can be derived from these results. First of all, they serve as a replication of the results of Experiment 1, given that a perceptual load effect on tone awareness is consistently observed across both studies. Secondly, they shed some light on the temporal dynamics of inattentional deafness. Specifically, the lack of an interaction between load and time indicates that, under high perceptual load, perception of the tone is impaired throughout the whole image presentation time. Again, reduced detection sensitivity was found under high load, this time in concomitance with an increased response bias (β). This is consistent with participants adopting a more stringent response criterion under high load. Given that A (detection sensitivity) and β (response bias) are independent measures, the results point in the direction of both an effect of perceptual load on tone detection, as assessed by A, and some degree of de-prioritization for the tone detection task under the more demanding condition, driving the response bias.

Chapter 3: Task switching and perceptual load

In the previous chapter I have asked the question of whether an increase in perceptual load can cause participants not to notice – and respond to – an auditory stimulus. Relating this question back to the driving scenario underpinning this thesis, it is easy to see how this question is important: the auditory stimulus is a proxy for the take-over request issued by the self-driving vehicle, and a failure to respond to it can have catastrophic consequences. Nonetheless, failure to notice a take-over request because of the demands posed by a non-driving task is unlikely to be the norm and more likely to be a worst-case scenario for a number of reasons. Firstly, the signal-to-noise ratio in the car might be better and favour take-over-request detection more than in our setup. Secondly, the load imposed by the non-driving task of choice might not be as high or as consistent in time. Lastly, some drivers might simply be unable or unwilling to focus as much on the non-driving task in the real word as they would in a much safer lab environment.

Here instead we move from the premise of a successful detection of the take-over request, and focus on other possible sources on interference when switching from the non-driving to the driving task. More precisely, we ask whether the level of perceptual load in a primary task can modulate response times to a less frequent secondary task, when the transition between the two is sudden, with no time given in advance to prepare for the switch.

In the introduction, I reviewed how perceptual load can be manipulated by having identical stimuli in the different load conditions and instead varying only the task to be performed. The primary task on this set of studies does just that. The stimuli are always crosses, and they are defined by two main parameters: colour and orientation. Crosses can have one of six possible colours and be oriented either upward or downward. In the low perceptual load condition, the task requires only a trivial feature detection: respond to all the red crosses. A more demanding feature conjunction is needed in the high perceptual load condition, which requires participants to consider both colour and orientation. Here the instructions are to respond only to upward yellow crosses and to downward green crosses. This task has been used in numerous studies (Bahrami et al., 2008; Carmel et al., 2011; Schwartz et al., 2005) and is known to reliably decrease accuracy and slow response times in the high load condition compared to the low load condition. From time to time, an auditory tone warns the participants to abruptly change task and focus instead on discriminating the prevalent motion direction in a random dot kinematogram (RDK).

This alternation of two tasks in rapid succession might, at first glance, place this line of work within the well-established task-switching literature. While there are some general similarities, there are also some important differences, most notably in the rationale and the measures of interest, which make the experiments contained in this chapter stand apart from the rest of the task-switching literature. I now turn to a brief overview of the task-switching paradigm for two reasons: clarifying the relation between task-switching and the present set of studies while at the same time overviewing the relevant topic of the role of attention during a task-switch.

As a field of research, task-switching (for reviews, see Kiesel et al., 2010; Monsell, 2003) is mostly concerned with understanding the nature of (and interaction between) the processes that allow a quick transition from one task to another. Such a transition is likely to involve several processes, such as retrieval of the new goals and of the associated stimulus-response mapping, shifting attention to new location, stimuli, or stimulus properties, selection of the appropriate response and possibly even an active inhibition of the previous task. This suite of processes is termed task-set and the update of task-set parameters during a task shift is known as task-set reconfiguration (TSR). According to the majority of models, when the task remains the same for two or more successive trials (i.e., repeat trials) there is usually no reason to engage in TSR, which might instead be needed when the task changes between one trial and the next (i.e., switch trials). The alternation between tasks can be planned for in numerous ways. Some studies have employed a predictable task sequence, where the task changes every second trial (i.e. AABBAA..; Rogers & Monsell, 1995). Others instead have opted for an unpredictable sequence, relying on cues to warn the participants about the identity of the upcoming task. With an unpredictable sequence, cues can either be presented on every trial (cuing paradigm; Meiran, 1996), or only when the task is about to change (intermittent instructions paradigm; Altmann & Gray, 2008; Gopher et al., 2000). Others yet have even left it up to the participants to decide when to switch task (voluntary task selection; Arrington & Logan, 2005).

By comparing average RTs and error rates of repeat versus switch trials it is commonly found that performance is considerably worse during switch trials on both measures. This is particularly true for RT, where the delay on switch trials can even amount to up to 40% compared to repeat RTs (Monsell, 2003). This is traditionally referred to as the switch cost, and its most widely accepted interpretation explains it as a result of the time consuming need to reconfigure the task-set. After establishing the phenomenon of the switch-cost, the literature has turned to assess if it is at all possible to engage in TSR ahead of the actual switch by warning participants in advance that a task switch is imminent. This possibility can be explored by manipulating the cue-stimulus-interval (CSI) between the cue that indicates which task is to be performed next and the onset of the new stimuli. An abundance of studies has shown that by allowing sufficient time to prepare there is indeed a reduction in switch cost (RISC effect). Often, the switch cost is only reduced, not entirely eliminated. With a long enough CSI, this reduction reaches an asymptote, at which point allowing even more time to prepare does not further reduce the switch cost. This persisting portion of the switch cost is called residual switch cost, and its interpretation is still open to debate. There is a general consensus around the idea that some components of the task set can be reconfigured endogenously (i.e., in advance) while others can only be updated exogenously, meaning that they require the actual onset of the new stimuli[1]. Which processes belong to the former category and which to the latter, though, is much less clear. Response selection has often been indicated as the primary generator of the residual switch cost, on the account of the congruence effect. Reaction times to a task are slower if the alternative task maps the same stimulus to a different response, and while the overall switch cost is reduced with a longer preparation time, the congruence effect does not seem to benefit by preparation.

Of course, attention also plays an important role when switching task. If the new stimuli are presented at a different location, one must attend the newly relevant location and disregard the previous one. Similarly, if the upcoming task requires classification of a different feature, such as shape versus colour, this desired property needs to be selected and attended to. The role of attention has been recognized by

[1] An entirely different account of the residual switch cost has been put forward by De Jong 2000). Its justification is derived from the analysis not only of the average response times but also of the response times distribution for switch trials. He noticed a subset of responses that were as fast as the average repeat trial RT, while others were much slower. He adopted a probabilistic explanation for this discrepancy and proposed that subjects attempt to perform TSR ahead of time, and each attempt is associated with a probability of succeeding and a probability of failing. If the attempt succeeds, TSR is performed entirely ahead of time and subject will be as ready to perform the new task as they would be for a repeat trial. If the attempt fails instead, TSR will have to wait for stimuli onset to be initiated.

various models, such as the Meiran's CARIS framework (Meiran, 2000a, 2000b; Meiran et al., 2008) or Logan and Gordon's ECTVA (Logan & Gordon, 2001). Without getting into the specifics of each model, it is sufficient to note here that the role assigned to attention differs drastically between models. Some postulate that attention can efficiently be reconfigured prior to stimuli onset (Lien et al., 2010), pointing to a limited contribution of attention to the residual switch cost, while others do not. It would not be far-fetched to imagine that there could be some sort of inertia in resetting attentional parameters and redirecting attention in line with the demands of the upcoming task. In fact, the possibility of a general task-set inertia for non-attentional components of the TSR, by which task-set parameters carry over from one trial to the next and require active suppression, has been proposed by Allport et al. (1994) and is often discussed as a contributor to the switch cost. Nonetheless, only a minority of studies have investigated attention as a possible limiting factor during a task switch. Among these, some have pursued this line of enquiry with the aid of eye-tracker measurements. Traditional task-switching studies rely on the difference between switch and repeat RTs to assess the switch cost. These overall measures make it difficult to disentangle different components of the TSR and to pick apart the possible contribution of attentional factors. The use of eye trackers instead grants a more precise understanding of the role of attention via the close (albeit indirect) relation between gaze position and spatial attention. By sampling the gaze location with very high temporal resolution (up to 1000 Hz), and with an opportune arrangement of the stimuli on the screen it is possible to determine the focus of spatial attention at any point in time.

Mayr et al. (2013) for instance, presented participants with three vertical bars arranged as the vertices of an imaginary equilateral triangle. One of the bars had a slightly different color from the other two, and a second bar had a small horizontal gap close to either the bottom or the top of the bar. The first task was to indicate if the color of the differently coloured bar was of a darker or lighter shade. The second task was to indicate if the gap in the bar was positioned high or low with respect to that bar. Fixations were analyzed from the appearance of the stimuli until a response was made, and they were classified as relevant if they occurred in the vicinity of the target or irrelevant if they ended on one of the two distractor bars. Analysis of average RT showed the usual switch cost as well as its reduction with a long (1000 ms) compared to a short (300 ms) CSI. What is more interesting is the insight that the analysis of the gaze probability distribution allowed regarding the role of attention. In order to perform the task, one must of course start by attending to – and therefore fixate - the correct stimulus. The average probability of fixating relevant and irrelevant areas was computed for every 25 ms period after stimulus onset and then the difference between the two was calculated. A positive difference means that at that particular time participants were more likely to fixate the target than either of the distractors, and this can be considered a timestamp for the onset of top-down attentional deployment. By computing this parameter separately for repeat and switch trials at both the short and long CSI it becomes possible to assess how long after stimulus onset attention is directed according to the task demands. When there were only 300 ms to prepare to the upcoming task, this point in time was delayed by about 100 ms on switch trials as opposed to repeat trials. To put it differently, there was a window of 100 ms during which participants were already more likely to be fixating the relevant stimulus during repeat trials, while on switch trials they were still equally likely to be fixating a relevant as they were an irrelevant stimulus. On top of that, switch trials seemed to have a separate, longer lasting effect on attention. At the longer CSI (1000 ms) there was no difference between repeat and switch trials regarding the timepoint at which participants were more likely to fixate the target, which indicates that attentional settings can be – at least to some extent – reconfigured proactively before stimulus onset. Nonetheless, even at 400-700 ms after stimulus onset participants had higher probability of fixating the relevant target in repeat as opposed to switch trials. This indicates a more efficient selectivity for repeat trials and might imply a persisting interference of the previous task on attention.

Similar results were observed by Longman et al. (2013). Their stimuli consisted of four different faces with one of four letters superimposed over the forehead, and the tasks consisted in categorizing either the letter or the face. With a 200 ms CSI, participants tended to fixate the irrelevant region (the forehead in the face identification task and the face in the letter categorization task) markedly more on switch trials for up to 500 ms from stimuli onset. This tendency was reduced but not eliminated even with a 800 ms CSI. Additionally, the magnitude of this attentional inertia phenomenon was positively correlated with the magnitude of the switch cost.

Both Longman et al. (2013) and Mayr et al. (2013) only analysed eye movements from the onset of the stimuli, not from the cue that identified the upcoming task. As a result, it was not possible to observe the gaze dynamics during the preparation interval. This makes it more difficult to unequivocally attribute the results to the process of reconfiguring attentional parameters. Longman et al. (2014) addressed this shortcoming. Participants performed one of three tasks, each associated with a fixed position throughout the experiment (a vertex of an equilateral triangle), while a centrally presented cue indicated the next task to be performed. Interestingly, they also introduced a control condition in which only one task had to be performed, and the cue in this case only indicated the location of the relevant stimulus (not the task to be performed) to contrast a switch in both task and location (main condition) with a change in location only. This study confirmed the previously established finding at attentional inertia and added new insights. Attention was not simply captured by any irrelevant location, but only by the previously relevant bar, rendering an account in terms of general distractibility less likely. The control condition demonstrated that this effect is not induced by a simple change in task location, but only by the combination of location and task change. On top of that, the fact that attentional inertia was found at all CSIs, even above one second, strongly indicated that attention is a contributor to the residual switch cost.

The literature reviewed so far points to a likely contribution of attention to the switch cost. Regardless of whether attention parameters can be fully or only partly reconfigured in advance given a long enough CSI, with shorter preparation times attention seems to play a role and to introduce a delay, most likely alongside contributions from other processes of the TSR. Attention is therefore affected differently in repeat vs switch trials, and this speaks generally to the importance of considering attention in a task switch setting. However, the question asked in this chapter is only partially related to the previous literature. Specifically, we are not interested in comparing performance on repeat versus switch trials. What we ask instead is if, on switch trials, the level of pre-switch attentional demands influence post-switch performance, above and beyond the general attentional inertia so far described. None of the aforementioned studies attempted to manipulate pre-switch attentional demands. To the author's best knowledge, Chan and colleagues (Chan et al., 2017; Chan & Desouza, 2013) conducted the only two studies attempting to investigate this specific question. In their first experiments participants had to either make a saccade toward a peripherally presented dot (prosaccade task) or away from it (antisaccade task). Before making the saccade, fixation was maintained on a central grey box, the outline of which changed colour to indicate the type of task to be performed next (prosaccade or antisaccade). After a CSI of 200-1500 ms the box reverted back to grey and the peripheral stimulus was presented. This experiment served as a baseline characterization of pro and antisaccade performance and the associated switch cost without additional attentional demands. In a second experiment, an attentional load was imposed by having participants perform an additional task during the CSI interval. For as long as the colour cue persisted, participants engaged in a rapid serial visual presentation (RSVP) inside the central box. Ten letters per seconds were presented along with an occasional number. Participants had to respond to the appearance of the number singleton. All other aspects of the experiment remained constant. The addition of the RSVP had an effect on both error rates and saccadic response times. For both the prosaccade and the antisaccade task, a greater attentional load was associated with higher error rates as well as increased saccadic reaction times. These results were replicated by Chan et al. (2017) in a series of fMRI experiments with the same paradigm.

While in principle Chan and Desouza (2013) and Chan et al. (2017) addressed the question posed in this chapter, their studies present an important limitation. They contrasted an experiment in which only the prosaccade and antisaccade tasks were performed with a second experiment that required the execution of an additional task. A similar methodological concerned was raised in the early days of task-switching research, with regards to studies that compared blocks with one task to blocks in which two tasks were alternated. The criticism, and the reason why this methodology has since been abandoned, is that the different blocks had different memory demands. Maintaining two task-sets in a state of activation is surely more taxing for working memory than having to deal with only one task-set. Therefore, any switching cost that emerges by this comparison cannot be unequivocally attributed to the need of reconfiguring the task-set. The cost in response times and accuracy observed under this paradigm has been termed mixing cost and is thought to reflect a mixture of TSR and additional working memory demands. Similarly, in the studies by Chan and colleagues the RSVP does create attentional demands that were not present in Experiment 1, but its presence also adds to the overall memory demands of the experiment.

There were also other differences with the scenario inspiring this chapter. One is that participants could prepare in advance. The RSVP and the cue coexisted for 200-1500 ms. During this period participants could use the foreknowledge of the upcoming task to prepare for the switch. Lastly (and tightly related to previous point) the knowledge offered by the cue made the correct action to be taken after the switch totally unequivocal. Given that our primary interest is toward totally unexpected control authority transitions between the self-driving vehicle and the hypothetical human driver, allowing time to prepare for the transition would decrease the similarity between the experiment and the problem domain. What's more, in a sudden transition, the best course of action to take after control of the vehicle is handed back to the operator is far from obvious and instead requires careful consideration of the environment. Finally, in the experiments by Chan and colleagues the antisaccade or prosaccade task always followed the RSVP. This of course creates a very high expectation – better, a certainty - that either of these tasks would have to be performed frequently and with a very predictable regularity. Emergency transition with no prior warning are a last resort and imply that something totally unexpected has occurred. They are therefore likely to be very rare and their timing unpredictable.

For all these reasons, the three experiments reported here all follow a newly designed paradigm that addresses all of these points. Different levels of attentional demands are manipulated in the primary cross task, which takes up the vast majority of the time spent on the experiment. A task switch only occurs in 25% of the trials, making expectations of a switch reasonably low. The task switch, signalled by a clearly audible tone, is sudden, with no time at all to prepare for it. Additionally, the correct response is not known before the switch, as it rests on detecting the prevalent motion direction in a random dot kinematogram. Lastly, given that the hypothesis tested in this chapter only concerns the effect of the level of perceptual load in the cross task on performance in the motion task - and does not rely on computing the switch cost - we can do away with the necessity of having repeat and switch trials for both tasks.

Experiment 1

Design

The experiment followed a within-participants design with perceptual load (low, high) as the only factor. In each condition participants were required to perform the cross task under either low or high perceptual load until an auditory cue signalled the need to switch to the motion task. Dependent measures for both the cross task and the motion task were accuracy and reaction times.

Participants

Twenty-three participants took part in this experiment (age range 18-33). Participants were recruited via the UCL SONA database and were paid £7.5 to take part in the study. Exclusion criteria were as follows: age (outside the 18-35 range), colour blindness, impaired hearing, astigmatism. All participants signed a written informed consent form according to the declaration of Helsinki, and the experiment received approval from UCL ethics committee. Given that, to the author's best knowledge, no prior study has used this particular paradigm, the sample size was estimated based on a conceptually similar experiment, reported in Carmel et al. (2011; Experiment 2). Desired power was set at 90%, with $\alpha = 0.05$.

Apparatus and stimuli

The stimuli were presented on a 27'' screen which was viewed at a distance of approximately 57 cm, so that one cm on the screen corresponded roughly to one degree of visual angle. The screen resolution was 1920X1080 pixels with a refresh rate of 60 Hz. All stimuli were displayed against a uniform black background.

Cross task. The stimuli employed in the cross task were upward and downward crosses, with the long arm measuring 80 pixels (corresponding to 2.49 degrees of visual angle) and the short arm measuring 40 pixels (corresponding to 1.24 degrees of visual angle). The crosses were positioned at the centre of the screen. There were six possible colours for the crosses: red (rgb: 255 0 0), green (rgb: 0 255 0), yellow (rgb: 255 255 0), blue (rgb: 0 0 255), brown (rgb: 156 102 31) and purple (rgb: 160 32 240). Each cross was presented for 250 ms and was followed by a 500 ms gap before a new cross appeared.

Random dot kinematogram (RDK) Task. White dots were presented within an invisible square field at the centre of the screen, subtending 16 X 16 degrees of visual angle. Each dot had a diameter of 8 pixels, corresponding approximately to 0.25 degrees of visual angle, and a random "life" of 1 to 12 frames (16-200 ms). Whenever the life parameter of a given dot reached 0, that dot would disappear and be replaced by a new one at a random position. Dots were also eliminated and replaced by new ones whenever their updated position, calculated at each frame, would touch any of the borders of the surrounding square field. At any given moment the dot field contained 200 dots, which moved at a constant speed equal to 5 cm/sec. At the moment of their creation, dots were assigned a random direction vector, which remained constant throughout the dot life. During the time windows of partially coherent motion, a random subset of 80 dots, corresponding to 40%, started moving either upward or downward depending on the trial. Dots positions were re-initialised at the beginning of each trial. Crosses were in the foreground and dots in the background, so that the crosses remained visible whenever dots overlapped with their position.

Procedure

Participants attended an RSVP with crosses shown at fixation every 750 ms. Each cross was displayed for 250 ms and followed by a 500 ms gap, during which only the moving dots were visible. A trial was defined as a sequence of 12 crosses, where there could be 0, 1 or 2 target crosses[2]. If no mistakes were made (i.e., if all target crosses were correctly detected) a trial would continue seamlessly into the next so that participants experienced a continuous RSVP encompassing several trials. The RSVP was paused every 6 trials in order to display a reminder message showing the

---

[2] The precise number of trials containing 0, 1 or 2 target crosses was determined randomly for each participant in the following way. First, we computed all the possible combinations of 0, 1 and 2, of length equal to the number of trials in each perceptual load condition (i.e., 144) that sum up to the number of cross targets per load condition (i.e., 288). With a more formal definition, we calculated the subset of the partition of 288, of length 144, containing only 0, 1 and 2 as addends. One of these lists was randomly selected for each participant and shuffled once per each perceptual load condition. Therefore, a participant received the same number of trials containing 0, 1, and 2 cross targets in both perceptual load conditions, albeit in a different order.

upcoming target crosses for 4 seconds. Secondly, when no response was recorded for a target cross, at the end of the trial a feedback message lasting 5 seconds warned the participants that a target cross had been missed.

Responses to targets crosses were provided by pressing the spacebar with the left hand. Participant could make their response in the 1500 ms following the onset of a target cross (i.e., until the end of the gap after the cross following the target one). In order to achieve this response window, the position of the targets had to be constrained so that no two targets could appear consecutively or occupy the last (and first) position in a trial. In the low load condition targets were red crosses (regardless of their orientation); upward yellow crosses and downward green crosses were instead the targets in the high load condition. The frequency of target crosses was approximately 7.6%.

On 25% of trials participants had to stop performing the cross task and engage instead in a motion direction discrimination based on the prevalent direction of motion of the dots. The task change was signalled by a pure tone with a frequency of 1500 Hz and a 50 ms duration, the onset of which coincided with the onset of a cross in a random position. Again, to ensure a response window of at least 1500 ms, the onset of the coherent motion (signalled by the tone) could coincide with the onset of any cross except the first and the last. From the onset of the tone, 40% of the dots started moving either upward or downward with equal probability, while the rest of the dots kept moving randomly. Motion responses were made by pressing either the “up” or “down” key with the right hand. The coherent motion ended either when participants made their response or at the end of the trial (whichever came first). Crosses kept appearing even while coherent motion was displayed but participants were instructed to disregard any cross appearing after the tone and to focus solely on the dots. To discourage them for paying attention to the dots before the occurrence of the tone, they were also told that this would not have facilitated performance on the motion task in any way given that the dots were all moving randomly up until the onset of the tone. Trials containing a coherent motion segment ended as soon as a response to the motion was provided. The subsequent trial started after a 1.5 seconds delay, during which no stimuli were displayed.

Perceptual load was maintained constant within a block, and alternated between blocks in an ABBA fashion, with the perceptual load of the first block counterbalanced between odd and even numbered participants. Each block comprised 24 trials and lasted approximately 6 minutes. Each perceptual load condition received the same number of cross and motion targets, upward and downward coherent motion events, but no constant number of events for each of these parameters was enforced at the level of each block. Participants were free to take breaks between blocks as often as needed.

Before the beginning of the experiment, participants had the chance to familiarize themselves with the task by performing 4 practice trials, which were presented in a fixed order. While both crosses and dots were presented in all four of the trials, participants performed only the cross task for the first two trials, with the first being a low perceptual load trial and the second a high perceptual load trial. The last two trials instead, (again, one per perceptual load condition) contained both cross targets and a task switch to the motion task, with upward and downward prevalent motion respectively in the third and fourth trial.

Results

Data for one participant was discarded as they achieved 0% accuracy on one condition of the motion task. Additionally, the majority of data for two participants was lost due to a technical error, therefore those participants were discarded as well, leaving a total sample size of 20 participants.

Cross Task. Reaction times to cross targets further than 3 SD from the mean RT in either direction were discarded, which lead to the rejection, on average, of 0.5% of cross responses. A two-tailed paired samples t test revealed that participants were less accurate in detecting cross targets in the high perceptual load condition (M= 85.6%, SD= 7.4%) compared to the low perceptual load condition (M= 91.7%, SD= 2.5%): $t(19) = 4.03$, $p < .001$, $d = 0.91$. A separate two-tailed paired t test was conducted on reaction times for correct responses. This showed that participants also took longer to respond in the high perceptual load condition (M= 558 ms, SD= 64 ms) compared to the low perceptual load condition (M= 440 ms, SD= 58 ms): $t(19) = 11.09$, $p < .001$, $d = 2.48$. The perceptual load manipulation achieved the expected results. Accuracy and reaction times were both affected, which rules out a speed-accuracy trade-off.

Motion task. Similarly to the cross task, separate two-tailed paired t tests were conducted on accuracy and reaction times. In a limited number of trials (on average

1.8% of motion responses), participants did not respond to the coherent motion task and these trials were excluded from accuracy calculations. Also, an average of 1.5% of motion responses were removed as they were further than 3 SD from the mean.

No difference in accuracy to the motion task was observed between the high (M= 73%, SD= 22.8%) and the low (M= 71.1%, SD= 27.5%) perceptual load condition: t (19) = 0.89, p= 0.38. On the other hand, there was a significant difference in reaction times, with faster responses in the low perceptual load condition (M= 971 ms, SD= 295 ms) compared to high load condition (M= 1054 ms, SD= 365 ms): t (19) = 3.53, p = 0.002, d=0.79 (Figure 3).

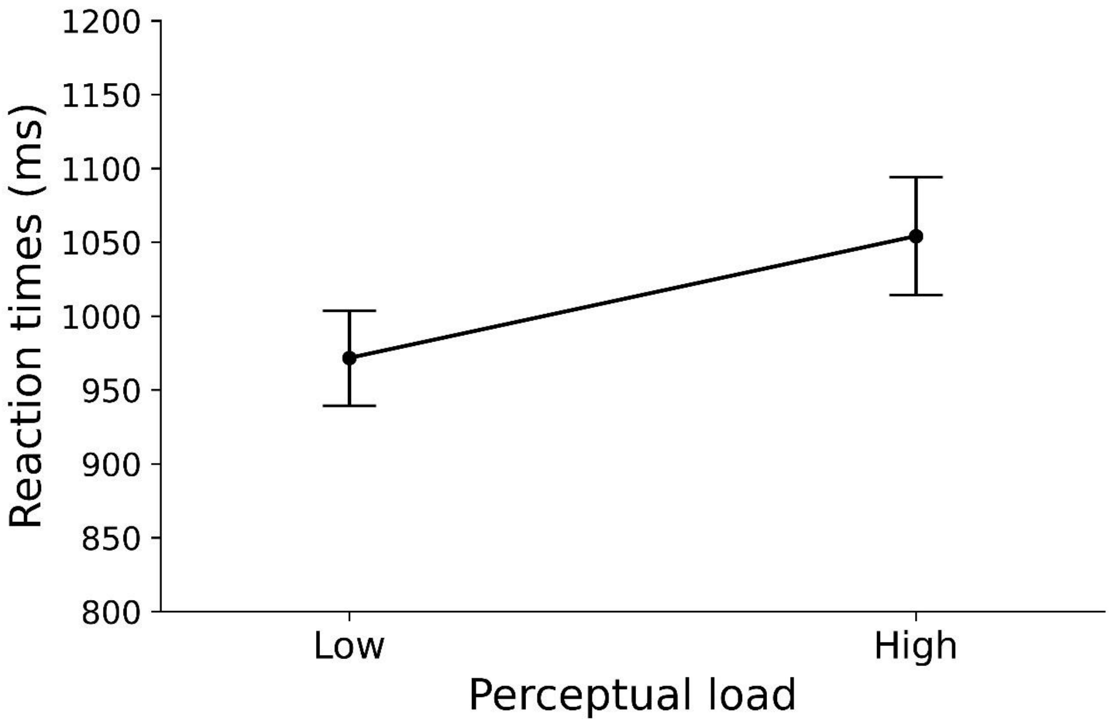

Figure 3. Effect of perceptual load on reaction times to the motion task. Bars represent the standard error of the mean.

Discussion

In this study we have examined the role of attentional demands on a task switch with no preparation time. Results showed that reaction times to the RDK task were modulated by the level of attentional demands posed by the cross task. Responses to the dots motion were faster when the preceding cross task consisted of a trivial colour detection and were instead delayed when the cross task required integration of two different features: colour and orientation. To the author's best knowledge, this is the first study to assess the effect of increased attentional demands on switch trials in a task-switch. Chan & Desouza, (2013) and Chan et al. (2017) posed the same question but went about it by contrasting experiments in which a RSVP-based task was either present or absent prior to performing either a prosaccade or an antisaccade task. As mentioned, this manipulation has the downside of changing memory demands alongside with attentional demands. In this study instead, memory demands were similar between conditions.

The observed delay differs from the classic phenomenon of the switch cost conceptually, operationally and quantitatively. Conceptually, insofar as the switch cost likely reflects multiple contributing sources other than attention, such as stimulus-response selection, inhibition of the previous task set and possibly others. Operationally, as the computation of the switch cost requires that both repeat and switch trials are performed for both tasks in the same block. In the present study instead, only switch trials were present for our primary task of interest (i.e., the RDK) and participants never performed the RDK task for two or more trials consecutively. As for the magnitude, the difference between average reaction times between the high perceptual load (M= 1054 ms) and low perceptual load condition (M= 971 ms) amounted to 83 ms, or approximately 8.6%, while the switch cost can be equal to 40% (Monsell, 2003). This is not surprising however, as the switch cost captures multiple sources, as discussed.

This attentional modulation in instead absent when accuracy is taken into account, with analogous accuracy rates in the two conditions. The fact that the two tasks were performed serially instead of in parallel (as the experiments reported in the previous chapter) renders alternative accounts in terms of rapid forgetting or goal neglect inapplicable.

Experiment 2

The aim of this experiment was twofold. Firstly, given the novelty of the paradigm employed in Experiment 1, the first goal was to replicate the main finding of slower responses to the post-switch task with increased perceptual load to the pre-switch task. Secondly, we wanted to assess the magnitude of this delay in comparison to a single task condition. Therefore, this experiment closely resembled Experiment 1 except for the presence of an additional *motion only* condition during which participants only performed the motion task without the preceding cross task.

Design

The experiment followed a within-participants design with perceptual load (low, high) as the only factor. In these two conditions participants were required to perform the cross task under either low or high perceptual load until an auditory cue signalled the need to switch to the motion task. In a third condition there was no cross task and participants only performed the motion task. Dependent measures for both the cross task and the motion task were accuracy and reaction times.

Participants

A new sample of 23 participants (age range 18-24) was recruited for this study. Participants were again recruited via the UCL SONA database and were paid £9 to take part in the study. Exclusion criteria were the same as in Experiment 1, with the additional constraint that participants who had taken part in the first experiment could not participate to this experiment. All participants signed a written informed consent form according to the declaration of Helsinki, and the experiment received approval from UCL ethics committee.

Procedure

Apparatus, stimuli and procedure were identical to those of Experiment 1, except for the following two differences. Firstly, on top of the 12 blocks included in Experiment 1, participants performed four additional blocks which constituted the *motion only* condition. The underlying structure and timing in these blocks were preserved. Each trial was still structured as a sequence of 12 crosses, but the crosses themselves were made invisible and the participants only saw the random dot kinematogram. As in experiment 1, the coherent motion could not start sooner than 750 ms from the beginning of a trial or later than 1500 ms from the end of a trial. These additional blocks had 18 trials each, for a total of 72 observations. The experiment started with a motion-only block, with subsequent ones separated by four regular blocks (i.e., blocks 6, 11 and 16).

Additionally, a change was made in the way in which feedback about the cross task was communicated. In Experiment 1, a feedback message appeared at the end of each trial in which a cross target was not detected. This created a number of interruptions and unnecessarily increased the duration of the experiment. Additionally, accuracy in the high load condition was lower and this resulted in more frequent interruptions in this condition. These shortcomings were addressed in Experiment 2 by presenting a feedback message every six trials (e.g., “You have missed 2 targets out of 10”)[3].

Results

Removing outliers further than 3 SD from the mean resulted on average in a loss of 1% of responses to the cross task and 1.4% of responses to the motion task. On average, 1.7% motion responses per participant were excluded because no response to the motion task was provided.

Cross task. Again, participants were more accurate in the low load (M= 95.18%, SD= 3.36%) as opposed to the high load condition (M= 91.36%, SD= 4.98%): $t(22) = 4.3$, $p < 0.001$, d= 0.9. Similarly, responses were faster in the low load condition (M= 439 ms, SD= 51 ms) than in the high load condition (M= 576 ms, SD= 39 ms): $t(22) = 18.9$, $p < 0.001$, d= 3.9. The results on both accuracy and reaction times confirms that the load manipulation had the desired effect.

Motion task. A repeated measures ANOVA on accuracy scores revealed no significant difference between the motion-only *condition* (M= 91%, SD= 18%.7%), the low load condition (M= 90.6%, SD= 19.2%) and the high load condition (M= 90.3%, SD = 18.3%): $F(2, 21) = 0.1$, $p= 0.9$. Instead, the repeated measures ANOVA on

[3] This feedback structure was also employed in Experiment 3 and in the fMRI experiment reported in Chapter 4.

correct response times highlighted a significant main effect of condition F (1, 21) = 17.7, $p < 0.001$, $\eta_p^2 = 0.44$. All post hoc t-tests were significant (Figure 4): not only were RT in motion-only condition faster than in both the low load (t= 3.77, p= 0.001, d= 0.79), and the high load conditions (t= 4.53, p< 0.001, d= 0.95), but the low load condition also had faster RT compared to the high load condition (t= 4.13, p< 0.001, d= 0.86). Average response times were 793 ms (SD= 225 ms), 981 ms (SD= 340 ms) and 1050 ms (SD= 377 ms) respectively in the motion-only, low perceptual load and high perceptual load conditions.

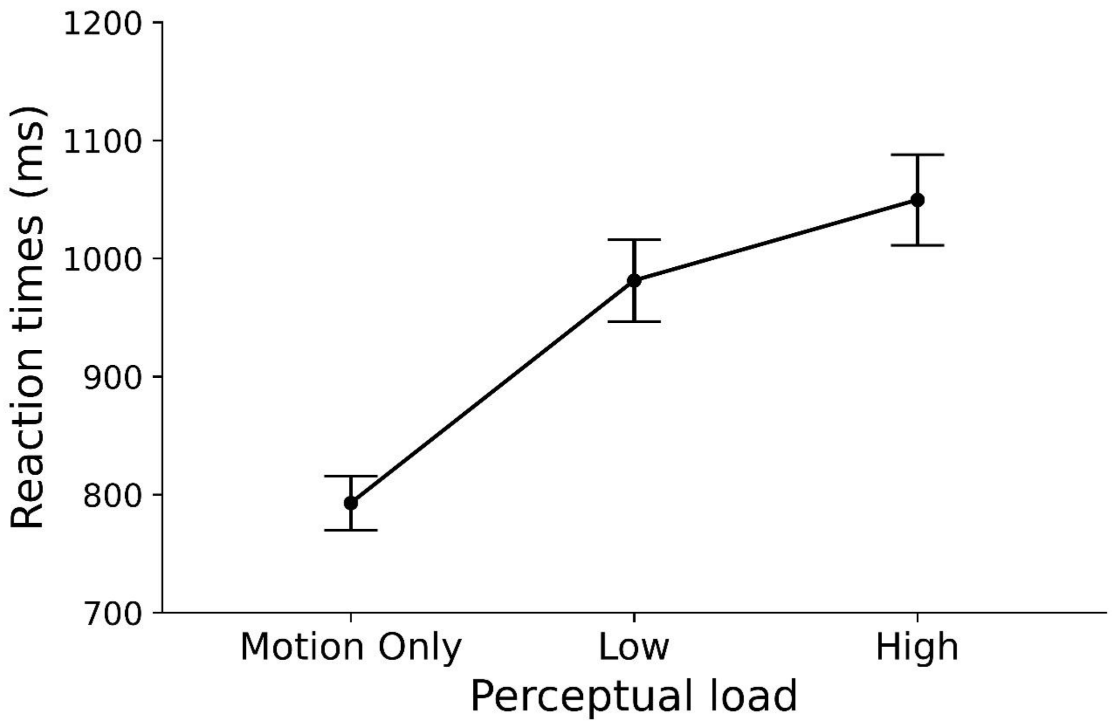

Figure 4. Effect of perceptual load on reaction times to the motion task. Bars represent the standard error of the mean.

Discussion

Experiment 2 successfully replicated the main finding of Experiment 1: increased attentional demands prior to a task switch led to slower response times to a subsequent task. Performing the motion task following a high perceptual load cross trial resulted in a 7% increase in reaction times, which is consistent with the 8.5% reaction time increase found in Experiment 1. Given the novelty of the paradigm, the replication adds to the reliability of this result.

A further similarity was found with regard to accuracy scores for the motion task, which were not in the least affected by the attentional manipulation. Interestingly, accuracy for both low and high perceptual load are considerably higher compared to experiment 1: 17.3% and 19.5% higher accuracy to the motion task in Experiment 2 for the low load and the high load condition respectively. This difference can at least partly be explained by two factors. Firstly, and most importantly, Experiment 2 had 72 more coherent motion trials compared to Experiment 1, and this presumably led to a higher degree of familiarity and experience with this task, making it easier to some extent. Secondly, Experiment 1 included frequent interruptions because each cross target that was not detected triggered a feedback message at the end of the trial. Conversely, in Experiment 2, only one feedback message was displayed every six trials. It is conceivable that these frequent interruptions led to a lower engagement with the task in Experiment 1 compared to Experiment 2, and the lower accuracy observed in the cross task as well in Experiment 1 might be in support of this interpretation.

Experiment 3

Experiment 1 and 2 have shown that increasing the attentional demands of the cross task results in slower discrimination of the dots direction in the motion task. Both experiments shared the same temporal relation between the onset of the crosses and the onset of the partially coherent motion (with the exception of the *motion-only* condition in Experiment 2). In fact, the tone that signalled the need for a switch, as well as the beginning of the coherent motion, always coincided with the onset of a cross. It is therefore unclear whether this RT delay under high perceptual load would still be manifest should the onset of the cross and the onset not coincide with of the start of the coherent motion – more precisely, if the onset of the coherent motion occurred after the cross stimuli have disappeared. In a similar context, Carmel et al. (2011) have asked whether the reduced perception of distractors is due to the actual presence of the loading stimuli (such as the crosses used both in the experiments described here and in Carmel et al., 2011) or simply to their ongoing processing, which does not necessarily coincide – or at least not entirely – with their physical presence. They had participants perform the very same cross task employed here and additionally required them to detect the occasional occurrence of a meaningless shape presented peripherally. This shape, referred to as the critical stimulus, could either be presented synchronously with a cross or 250 ms after its offset. The authors argued that the perceptual load should exert its effect as long as the stimuli generating it are still being processed. Their results supported this view, as they did not observe an interaction between perceptual load and the onset time of the critical stimulus. This means that perceptual load had the same effect on perception of the critical stimulus as long as the crosses were being processed, regardless of whether they were actually present or not. On the other hand, a follow-up experiment showed that if the critical stimulus was presented once the processing of the crosses was likely to have ended, perceptual load did not - and could not have – affect detection of the critical stimulus. Despite this, it does not necessarily follow that a similar pattern of results should be expected under the present paradigm. Though the cross task was identical, there are in fact several noteworthy differences between the experiments by Carmel et al. (2011) and the ones reported in this chapter. First and foremost, participants in Carmel et al. (2011) were constantly performing the cross task at fixation while monitoring for the appearance of the critical stimulus in the periphery. The tasks were therefore performed in parallel. Here instead, tasks were performed serially and participants were instructed to completely disregard the crosses and to focus entirely on the RDK upon hearing the auditory tone. This entails that participants did not have to rely on their already taxed visual perception to spot the occurrence of the relevant stimulus for the secondary task, but were instead explicitly made aware that a relevant change had occurred. Secondly, contrary to Carmel et al. (2011), the dots composing the RDK were presented centrally in the same region occupied by the crosses. Lastly, while Carmel et al. (2011) focused entirely on detection sensitivity, computing this measure would not have been meaningful under the current paradigm. For any detection sensitivity measure to be calculated false alarms, even if infrequent, need to at least be possible. Here instead there was by design no uncertainty regarding the presence of the coherent motion: it never occurred prior to the tone (which was conveyed explicitly in the instructions) and it always occurred after the tone. We focused instead on accuracy and reaction times and found only the latter to be affected by the attentional modulation prior to the task switch in Experiment 1 and 2.

In Experiment 3, therefore, we aimed to test the effect of a different temporal layout on reaction times to the motion task with a between participants design. Half of the participants had the coherent motion onset matching the cross onset, while for the other half motion onset coincided with cross offset, with a delay of 250 ms compared to the other group.

Design

The experiment followed a 2X2 mixed design. The independent variables were perceptual load (low, high) and onset time of the motion task (same onset, delayed onset). Perceptual load was manipulated within participants, while onset time was manipulated between participants. In each condition participants were required to perform the cross task under either low or high perceptual load until an auditory cue signalled the need to switch to the motion task. Dependent measures for both the cross task and the motion task were accuracy and reaction times.

Participants

Forty participants took part in this study (age range 18-35). Participants were again recruited via the UCL SONA database and were paid £7.5 to take part in the study. The same exclusion criteria as in the previous experiments were applied. Participants in this study could not have participated in Experiment 1 or Experiment 2. All participants signed a written informed consent form according to the declaration of Helsinki, and the experiments received approval from UCL ethics committee.

Procedure

Apparatus, stimuli and procedure were identical to those of Experiment 1, with the only exception that, for even numbered participants, tone and coherent motion onsets always coincided with the offset of a cross. For odd numbered participants, instead, the procedure was identical to that of Experiment 1, with tone and coherent motion onsets always occurring at the onset of a cross.

Results and discussion

Data points further than 3 SD from the mean were removed; this resulted in a loss on average of 0.7% of responses to the cross task and 1.4% of responses to the motion task. Given that the onset of the tone and of the coherent motion was manipulated between participants while load was varied within participants, a series of mixed ANOVAs was carried out to analyse accuracy and reaction times on both the cross and the motion task. Onset was considered a between factor on two levels while perceptual load was a within factor, also on two levels.

Cross task. The mixed ANOVA on accuracy scores showed a main effect of load: (F (1, 38) = 34.8, p< 0.001, η_p^2= 0.18) but no interaction between load and onset (p= 0.9). A post hoc t test between high and low load (averaging over onset levels) indicated that participants were more accurate when performing the cross task under the low load condition (M= 95.8%, SD= 3.3%) than in the high load condition (M= 91.7%, SD= 5.4%): t (39= 5.9, p< 0.001, d= 0.9). Similar results were found on the mixed ANOVA on reaction times to the cross task. Again, there was a significant main effect of load (F (1, 38) = 413.1, p< 0.001, η_p^2= 0.6) with no interaction between load and onset time (p= 0.57). A post hoc t test (averaged across onset levels) revealed that participants responded faster in the low load condition (M= 455 ms, SD= 48 ms) compared to the high load condition (M= 588 ms, SD= 62 ms): t (39) = 20.3, p< 0.001, d= 3.2.

Motion task. Low load (M= 86.2%, SD= 16.4%) and high load (M= 86.7%, SD= 16.7%) accuracy scores did not differ (F (1, 38) = 0.2, p= 0.6), nor was there an interaction between load and onset (F (1, 38) = 0.02, p= 0.89). A different scenario emerged when considering reaction times instead. Here we observed a main effect of load (F (1, 38) = 5.6, p= 0.024, η_p^2= 0.003), meaning that, when collapsing across motion onset times, participants responded faster to the coherent motion in the low load condition (M= 999 ms, SD= 344 ms) compared to the high load condition (M= 1040 ms, SD= 396 ms). The interaction between load and onset time was not significant: F (1,38) = 3.8, p= 0.058, η_p^2= 0.002. When coherent motion started at cross offset, response times were almost identical across load conditions (low load: M= 959 ms, SD= 286 ms; high load: M= 966 ms, SD= 278 ms). Instead, when the coherent motion and cross onset coincided, reaction times varied to some extent as a function of perceptual load (low load: M= 1037 ms, SD= 398 ms; high load: M= 1113 ms, SD= 483 ms), though this difference was not sufficiently ample to cause a significant interaction between perceptual load and motion onset (Figure 5).

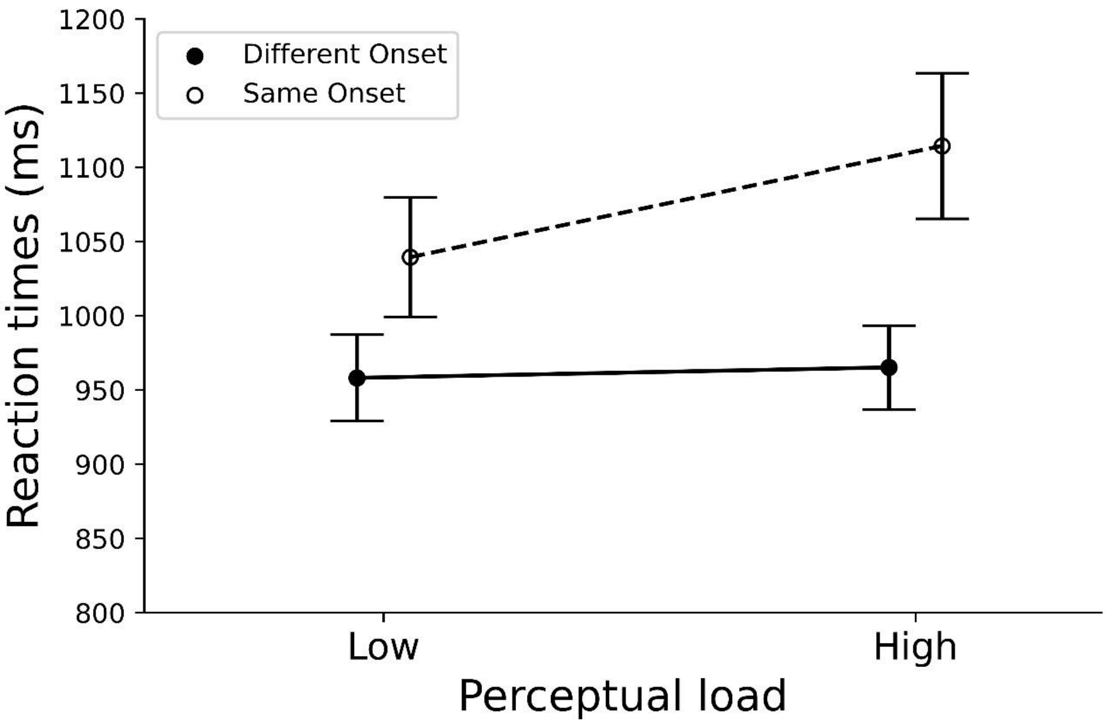

Figure 5. Effect of onset time on reaction times to the motion task at different levels of perceptual load. Bars represent the standard error of the mean.

Discussion

The three experiments described above present a rather stable picture. Our perceptual load manipulation worked as expected and consistently resulted in lower accuracy scores and longer reaction times in the high load condition compared to the low load condition. This double effect on both accuracy and reaction times rules out an alternative account in terms of speed accuracy trade-off. As for the motion task, the level of perceptual load on the cross task had no influence at all on accuracy in the motion task.

The main goal of these experiment was to investigate the possibility that higher attentional demands right before a task switch could impair post-switch performance. Indeed, this was reliably observed in all three experiments. The stimuli employed in the high and low perceptual load conditions were exactly the same. The only difference was in the difficulty and attentional demands posed by the different kinds of detection required by the low and the high perceptual load condition: respectively, a simple feature detection or a more demanding feature conjunction. Experiment 2 showed that this attentionally induced RT cost is a fraction of the cost incurred when comparing performance to the motion task alone with performance to the motion task when preceded by an additional task, regardless of the level of perceptual load.

Experiment 3 attempted to clarify the conditions necessary for this delay to occur. In light of the opposing points of view derived by the literature (Muggleton & al., 2008; Torralbo & Beck, 2008; Carmel et al., 2011), we tested whether the very presence of the loading stimuli at the moment of the take-over request is mandatory or simply their ongoing processing. Unfortunately, the results proved less conclusive on this front. On one hand, we did not observe a clearly significant interaction between the level of perceptual load and the onset time of the coherent motion. This is in line with a very similar finding reported by Carmel et al. (2011) albeit with a partially different paradigm which required the parallel execution of two tasks. On the other hand, the result of the interaction, which was just shy of significance, could be due to low power for the interaction test, which is a common issue in between participants designs. Clearly, further investigation is required on this point.

What could be the mechanism responsible for the observed difference in reaction times? A straightforward explanation cannot be easily derived from load theory. Under this framework, the exhaustion of attentional resources under high load leaves no other to process additional stimuli. This in turn can result in reduced interference from or lack of awareness of distractor stimuli, and it is accompanied by evidence of reduced processing of such stimuli in neuroimaging studies. Two other characteristics are commonly found in most perceptual load experiments. Firstly, both the loading stimuli and the distractors are delivered at the same time. Secondly, participants are constantly engaged in the task afforded by the loading stimuli. Carmel et al. (2011) have shown that the physical presence of the loading stimuli is not mandatory. Still, this does not change or challenge the rationale for the observed effect. Present or not, if the stimuli are being processed, the reduction of available attentional resources perdures and leads to the same effect that we would observe if the stimuli were present. The same, however, cannot be said for the second point. If participants stop performing the task by means of which perceptual load is manipulated, the attentional resources invested in that task should become available again. How, then, can perceptual load exert an effect over performance on a subsequent task? There are at least two plausible reasons for this. The first possibility is that the effect of high perceptual load does not compromise processing of the motion stimuli per se but instead delays the onset motion processing. High perceptual load could hence compromise attentional disengagement from the crosses, delaying it.

An alternative account instead could reside in the carry-over of the previous task set. The task switching literature has shown on numerous studies that that previous task set activation can carry over in the next trial and gradually dissipate over time. For this very reason it is now common in task switching studies to have a sufficiently long interval occurring from the response to a trial to the appearance of the next set of stimuli. This interval is called response stimulus interval (RSI). Common values for the RSI are around 1 to 1.5 seconds. A long RSI interval allows for the dissipation of the previous task set before the onset of the new stimuli occurs, therefore disentangling effects of preparation from effects due to the carryover of the previous task set. In the three experiments reported instead, the RSI is less than 750 ms[4]. Consequently, there might be some residual activation of the task set related to the cross task even after the onset of the motion task. This could be enough to explain the

[4] In task switching experiments a response is often required every time new stimuli are presented. In our case instead, there could be 0-2 target crosses every trial, which comprised 12 crosses. The use of the term RSI is nonetheless adequate because the task set clearly needs to be active throughout a trial despite the fact that not all stimuli require a response.

results even without assuming that such an interval would differ in length between the low and the high perceptual load conditions. During this transient time window both task set would be active, as if the tasks were being performed in parallel. Until the cross-related task set has a chance to dissipate, the motion task would suffer the effect of high perceptual load, with lower attentional resources available to process the motion direction, thereby explaining the observed delay in reaction times.

While mostly tangential to the goal of the experiments, it was surprising to see a substantial variation in accuracy between experiments. This was confirmed by a mixed ANOVA on the accuracy scores, with *load* as a within participants factor (on two levels) and *experiment* as a between participants factor (on three levels). Indeed, there was a main effect of *experiment* ($F(2, 85) = 4.2$, $p= 0.018$, $\eta_p^2= 0.09$). Post-hoc tests showed that the only significant difference occurred between Experiment 1 and Experiment 2 ($t(86) = 2.86$, $p= 0.016$, $d= 0.31$). A partial explanation for this discrepancy could lie in the different number of coherent motion trials between these two experiments. Experiment 2 had twice the number of motion trials compared to experiment 1 (144 versus 72) due to the four extra blocks constituting the motion-only condition. This extra practice might have proven particularly useful given that participants only performed four practice trials before the beginning of the experiment. This explanation, though, can only be a partial one. In fact, Experiment 3 also had half of the motion trials present in Experiment 2, and while accuracy was indeed lower in both the low load (-4.4%) and the high load condition (-3.6%), this difference did not approach significance. Another – admittedly posthoc – explanation could be that the lower accuracy is attributable to the different feedback system adopted for the cross task in Experiment 1. This experiment suffered frequent interruptions due to the feedback message shown for every trial in which a target cross was not detected. The resulting numerous interruptions might have lowered focus on the task and participants' motivation to perform well.

Chapter 4
Exploring the effect of pre-switch perceptual load on motion processing in V5: a fMRI study

Introduction

The experiments exposed in the previous chapter have shown that, after a task switch, reaction times to the upcoming task are not only slowed when compared to single task performance, but also, crucially, that they are influenced by the attentional demands of the pre-switch task. The aim of the present chapter is to clarify this result by addressing the possibility that the observed difference in RT between the high and low load condition arises as a consequence of different levels of processing in motion-related brain areas.

Perceptual load theory posits that, as the attentional demands of a task increase, less attentional capacity is left to concurrently process other stimuli. Rees et al. (1997) have addressed the issue of how peripheral motion processing is modulated by central task demands in a fMRI study. They had participants perform in turn an easier linguistic task (respond to uppercase words) and a more demanding task (respond only to disyllabic words) at fixation. All the while dots were presented in the periphery, which were either static or moving radially towards the edge of the screen. What they found, coherently with their hypothesis, was that the motion of the dots produced greater evoked activity in V5 (compared to the static dots) when perceptual load was low and thus arguably leaving sufficient resources to process the motion. Under high perceptual load, instead, there was no increase in motion-related activity between the motion and the no-motion conditions. While Rees' et al. study is of clear relevance to the hypothesis addressed in this chapter, a few key differences make it unclear whether their findings would be easily extended to this new setup: the fact that there was no task switch, that the dots were always irrelevant, and, lastly, that their processing was assessed while participants engaged with another task.

In the experiment described below, participants performed the same task described in Chapter 3 in the fMRI scanner. This allowed to us to test the hypothesis that the longer RT found when switching away from a more demanding attentional task are accompanied by reduced motion processing in V5 and possibly other areas involved in the processing of motion. In order to assess the level of motion-related processing, a second level of motion coherence was added to the paradigm. An easier condition where 70% of the dots move coherently was added alongside the previously employed 40% coherence condition.

The validity of this motion manipulation rests on the assumption that a higher level of coherent motion in one direction would elicit a greater activation compared to an identical level of motion but with no dominant motion direction. This has in fact shown to be the case in both animals and humans. Of the many areas involved in motion processing the middle temporal area (MT/V5) is certainly the most studied, as the majority of its neurons respond more to moving than to static stimuli and have a preferred motion direction. Britten et al. (1993) trained monkeys to perform a direction discrimination task involving random dot kinematograms (as in the studies reported here) at different levels of motion coherence. More than 50% of recorded neurons increased the frequency of their firing linearly as motion coherence increased, while a minority exhibited a non-linear relationship. Regardless of the particular shape of the response as a function of motion coherence, the responses of all recorded neurons were clearly dependent of the level of motion coherence. Rees et al. (2000) did something very similar with humans using fMRI. Again, several levels of motion coherence were tested in RDKs, and the results confirmed both the relation between motion coherence and activation as well as the shape of this response. In addition to the V5 complex, other areas were found to have a similar relation with motion coherence, such as the kinetic occipital area, V2 and V3a. What's more, two frontal areas, namely the anterior cingulate gyrus and the left insula were found to have their activation linked to the level of motion coherence, albeit in the opposite direction (i.e., bold contrast decreased in these areas at progressively higher levels of motion coherence).

A second, relevant area of enquiry for the current study is whether attention is capable of modulating how motion appears. The answer is in the affirmative. This issue has been tackled in a number of ways. One line of research exploited the motion aftereffect (MAE): after prolonged exposure to motion in a certain direction, the observer experiences an illusory motion, after motion has ceased, in the direction opposite to that of the previously moving stimulus. Chaudhuri (1990) employed a textured background continuously moving in the same direction, a condition that, on its own, generated a robust aftereffect. The duration of this illusory motion was reduced when participants were asked to perform a task on an alphanumeric rapid-serial-visual-presentation at fixation while the background kept moving. While Chaudhuri relied on the presence or absence of a central task to modulate the MAE, Shulman (1993) and Lankheet and Verstraten (1995) took a different approach, utilizing stimuli comprising two opposite motion direction with, respectively, rotatory and linear motion. Lankheet and Verstraten, for instance, showed stimuli containing both leftward and rightward motion. Without explicit instructions to attend one motion direction or the other, this kind of stimulus did not elicit a MAE, because the adaptions to either direction (that would, on their own, cause MAEs) cancelled out. Instructions to attend to either direction caused a MAE in the opposite direction. Therefore, the direction of the aftereffect was shown to depend and be reversed by the motion component that selective attention was applied to. Yet another approach to this topic was taken by Rees et al. (1997), and it tied the duration of the MAE not just to selective attention in general, therefore to the presence or absence of a concurrent task while watching motion stimuli, but to the level of attentional demands demanded by such a task. Perceptual load sees the fate of irrelevant stimuli as dependent on the availability of attentional resources. As mentioned, Rees et al. had participants perform either a low load task or a high load task at fixation, all the while ignoring an outward radial optic flow of dots. The motion related activation elicited in V5 was significantly lower in the high load condition, as was the duration of the MAE. Motion processing, and perception, can be suppressed to a degree if attention demands posed by another concurrent task are sufficiently high.

The literature on the relation between motion perception and attention is not limited to the study of illusory motion, such as the MAE. Other confirmations of this attentional modulation come from the adoption of the pre-cueing paradigm (Dobkins & Bosworth, 2001). The general finding here is that, when the location in which the motion stimulus will appear is cued shortly before the stimulus presentation, performance on the motion related task increases. The rationale is that the cue reflexively captures attention, and hence the improved performance is a result of attentional deployment to the cued area. Dobkins and Bosworth (2001) found that pre-cueing the location of the RDK (as opposed to a no-cue condition) helped with the motion direction decision both when the RDK was presented alongside three other motion distractors (RDKs with no coherent motion) and, to a lesser extent, when the

RDK was presented in isolation. Similar findings were reported by Liu et al. (2006). They presented two RDKs to the left and right of fixation, and asked participants to report the motion direction (out of four possible choices) of the RDK with the more coherent pattern. One of the two stimuli had a fixed 50% coherence (the standard stimulus), while the second one could have a range of coherence values from 10% all the way up to 90% (the test stimulus). The cues were totally uninformative and could either be valid, invalid or neutral (the central fixation point was cued). Looking at the coherence judgments, they found that at all values of coherence, participants were more likely to believe that the test stimulus had more coherent motion when it was cued than when it was not, with intermediate results in the neutral cue condition. Also, motion direction judgment on the standard stimulus with 50% coherence were more accurate when it was cued compared as opposed to when it was not.

Given that some areas have proven to have a stronger response with increased motion coherence, and that attention is capable of modulating motion perception and processing, it is conceivable that varying the pre-switch attentional demands could have a differential impact on post-switch motion processing when an abrupt task-switch is carried out.

Design

The experiment followed a within-participants design with the following factors: perceptual load (low, high), motion coherence level (40%, 70%) and tone volume (low, high). The manipulation of the volume of the tone signalling a task switch is tangential to the purpose of the present discussion. In each condition participants were required to perform the cross task under either low or high perceptual load until an auditory cue signalled the need to switch to the motion task with a level of dot coherence equal to either 40% or 70%. Dependent measures for both the cross task and the motion task were accuracy, reaction times and the BOLD signal level.

Participants

A new sample of 24 participants were recruited for this study through the UCL SONA platform and paid £25 pounds for their participation. Like in the previous experiments, exclusion criteria were age above 35, impaired hearing, colour blindness and astigmatism. Participants could not have taken part in any of the experiments reported in Chapter 3.

Apparatus and stimuli

All stimuli were identical to those presented in chapter 3. They were viewed through a mirror mounted inside the scanner and responses were provided by pressing the corresponding keys on a custom response box which the subjects held on their chest. The auditory tone was delivered through earbuds.

Behavioural Procedure

Overall, the procedure was quite similar to that of the previously reported experiments, with a few exceptions, both in terms of trial structure and of counterbalancing of key parameters. The first notable difference concerns the percentage of coherently moving dots. Contrarily to previous experiment, here we wanted to test the possibility that perceptual load might modulate motion representation in V5 and other areas of the brain. This made it necessary to have two distinct levels of motion coherence, at each level of perceptual load. By comparing the brain activity associated with these two levels of coherence, the neural representation of motion could be detected. This, in turn, could be compared between the two perceptual load conditions. Therefore, while half of the motion coherent trials retained the previously employed 40% coherence, the motion coherence level in the remaining half was set to 70%. Consequently, in order to retain the same number of observations for each level of motion coherence, the percentage of trial containing a motion coherent sequence was doubled, from 25% to 50%.

Additionally, the volume of the tone signalling the task switch was now varied between a louder and a quieter tone, with both tones being clearly audible, as verbally confirmed by each participant after they were positioned and set up inside the scanner. No formal tone volume assessment was employed for this study. The difference in sound intensity between the two tones was determined within Matlab (2020A) by initializing the audio channels with volume values respectively of 0.7 for the louder tone and 0.2 for the quieter tone (with 1 being the highest volume possible and 0 the minimum) and keeping PC volume controls at 50%. In previous experiments, the coherent motion remained visible until one of two conditions was satisfied: either the participant made a response or the trial ended, whichever happened first. In this experiment instead, coherent motion was only visible for 400 ms. Both the dots and the crosses were removed after 400 ms had passed. The experiment resumed with the next trial once the response windows of 2 seconds from motion onset had elapsed. This change was intended to prevent possible confounds in the interpretation of the fMRI results by equalizing the motion presentation times across different motion coherence and perceptual load conditions. Blocks comprised 48 trials (double the length of blocks employed in earlier experiments), 24 of which contained a coherent motion segment. Each combination of motion coherence (40%, 70%), motion direction (up, down) and tone volume (quiet, loud) were presented three times in a block (2 X 2 X 2 X 3 = 24). At three points within each block (between the $12^{th}/13^{th}$ trial, $24^{th}/25^{th}$ trial, and $36^{th}/37^{th}$ trial) there was a 7.5 pause while participants received feedback on their performance on the cross task so far. The perceptual load condition remained constant within each block and alternated from one block to the next, with the starting condition counterbalanced between participants . Lastly, the length of the practice block, which was performed outside the scanner, increased from 4 to 16 trials, so as to give participants a chance to better familiarize with all the possible task combinations.

fMRI Procedure

A 3T scanner (Siemens PRISMA) was used for the acquisition of both structural (T1) images and T2∗-weighted echoplanar (EPI) images (106 ×106; 2 × 2 mm pixels; echo time (TE), 40 ms) with blood oxygen level-dependent (BOLD) contrast. The functional scans were acquired during 8 sessions. Each session comprised 272 volumes each and lasted approximately 6.5 minutes. Each volume was made up of 60 axial slices, oriented approximately to the anterior commissure-posterior commissure plane and with a thickness of 2 mm. Volumes were acquired continuously by means of a multi-band sequence, with an acceleration factor of 4 and an effective repetition time of 1.62s per volume. The first nine volumes in each session were discarded to allow for T1 equilibration effects. A T1-weighted functional scan lasting approximately 7 minutes was administered midway through the task, between the fourth and the fifth functional scan, to allow participants to rest. Additionally, a field-map sequence lasting approximately 40 seconds was acquired before the first functional scan and used to correct for inhomogeneities in the magnetic field. At the end of the session, a brief additional task was administered in a ninth functional scan lasting approximately 3.5 minutes, however this was not used for the analyses below.

Data analysis

FMRI data were analysed with SPM12 and Matlab 2020a, alongside with the Matlab extension MarsBar (for the ROI definition). Data from one participant were discarded due to excessive movements, while a second dataset was discarded because of an acquisition error, leaving 22 complete datasets. The volumes were realigned and un-warped, then normalized into 2 mm cubic voxels using the Montreal Neurological Institute (MNI) reference brain via the co-registered structural scan, using 4th-degree B-spline interpolation. Finally, volumes were smoothed with an isotropic 8-mm full-width half-maximum Gaussian kernel. The volumes acquired during the ninth session were treated as separate time series. For each series, the variance in the BOLD signal was decomposed with a set of regressors in a general linear model. Within each block, there were four regressors. Two stick regressors indexed the onset of a) high- and b) low-coherence motion displays. Two boxcar regressors indexed a) the duration of the cross task, prior to the coherent motion display, and b) the three feedback displays. These regressors, together with the regressors representing residual movement-related artifacts and the mean over scans, comprised the full model for each session. The data and model were high-pass filtered to a cut-off of 1/128 Hz. Parameter estimates for each regressor were calculated from least mean squares fit of the model to the data. Effects of interest were assessed in a random effects analysis as follows. To assess the main effect of motion coherence, collapsing over perceptual load condition, a contrast image was calculated at the first level for each participant and the resulting images were entered into a one-sample t test at the second level. This contrast was thresholded at a height threshold of $p < .001$, combined with an extent threshold calculated by SPM to achieve a whole-brain family-wise error corrected threshold of $p < .05$. We also performed analogous one-sample t tests looking separately at each perceptual load condition. In addition, mean beta images corresponding to each of the four conditions of interest (2 perceptual load levels x 2 motion coherence levels) were extracted for each participant. These were used to perform region of interest analyses, using the MarsBar SPM toolbox to extract the data. Our initial analysis plan was to use the results of the one-sample t test collapsed over perceptual load to define a motion-sensitive ROI which could then be used to interrogate the orthogonal Load x Coherence interaction effect. This would test the main hypothesis that the effect of coherence should be greater at the low load than the high load.

Results

Behavioural results

Responses faster than -3 SD or slower than +3 SD from the mean response times were excluded from analysis. Outliers were infrequent though, with only 0.5% of responses removed on average for the cross task and 0.9% of responses removed for the motion task. Only correct responses were entered into the analysis of reaction times.

Cross Task. For the cross task, responses to the low motion coherence and high motion coherence trials were analysed together, as the cross task always preceded the motion task. Participants were informed that dots would always move randomly before the tone was played and were also explicitly instructed to disregard the dots until then. A two-tailed paired t test showed that participants responded with greater accuracy in the low load condition (M= 89.2%, SD= 8%) compared to the high load condition (M= 85.3%, SD= 7.8%): $t(21)$: 2.4, $p = .025$, Cohen's d = 0.51. Similarly, participants were faster in the low load condition (M= 486 ms, SD= 39 ms) compared to the high load condition (M= 618 ms, SD= 45 ms): $t(21) = 15.1$, $p < .001$, Cohen's d = 3.2 (two-tailed t test).

Motion task. Data for the motion task were analysed by performing a repeated-measures ANOVA on the accuracy scores and on the reaction times. Perceptual load and motion coherence were treated as within participants factors, each with two levels: low and high. The only significant results concerning the accuracy was a main effect of motion coherence: $F(1, 21) = 68.2$, $p < .001$, $\eta^2 = 0.68$. Accuracy decreases from 81.1% (SD= 16.3%) in the high coherence condition to 60.3% (SD= 14.4%) in the low coherence condition (Figure 6). Neither the main effect of perceptual load ($F(1,21) = 2.9$, $p = 0.11$, $\eta^2 = 0.01$) nor the interaction between perceptual load and motion coherence ($F(1,21) = 0.26$, $p = 0.26$, $\eta^2 = 0$) were close to reaching the significance threshold.

The repeated-measures ANOVA on reaction times to the motion task again shows a main effect of motion coherence (F (1, 21) = 60.6, $p < .001$, $\eta^2 = .46$), with faster responses in the high motion coherence condition (M= 901 ms, SD= 103 ms) compared to the low motion coherence condition (M= 992 ms, SD= 121 ms; Figure 7). Additionally, though, response speed was also significantly influenced by perceptual load: F (1, 21) = 14.4, $p = .001$, $\eta^2 = .12$. Shorter response times were recorded in the low load condition (M= 924 ms, SD= 119 ms) as opposed to the high load condition (M= 970 ms, SD= 105 ms; Figure 8). The interaction between motion coherence and perceptual load was not significant (F (1,21) = 0.003, $p = 0.958$, $\eta^2 = 0$; Figure 9).

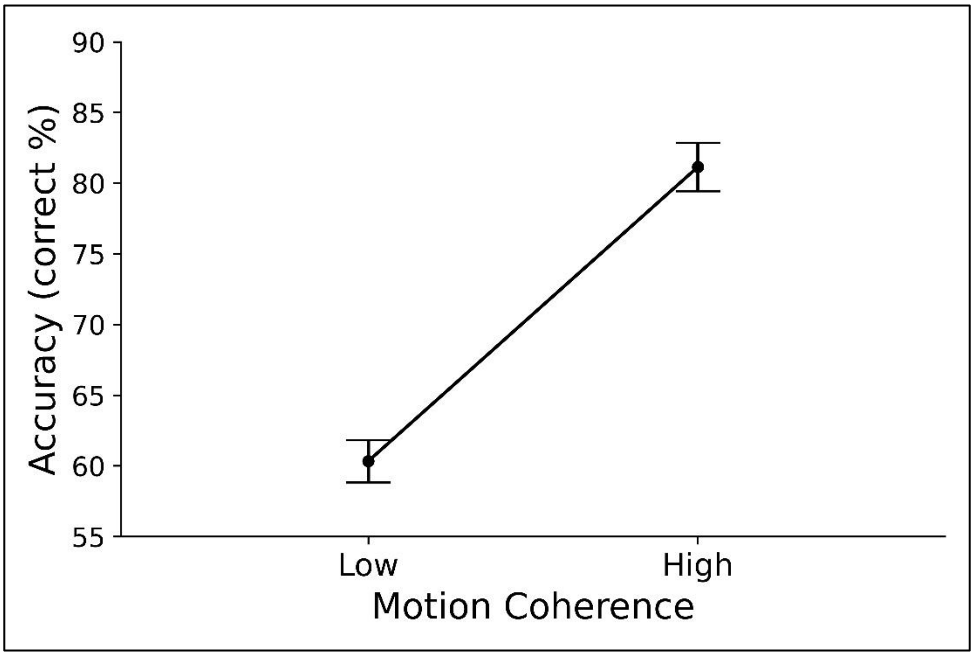

Figure 6. Effect of motion coherence on accuracy scores on the motion task. Bars represent the standard error of the mean.

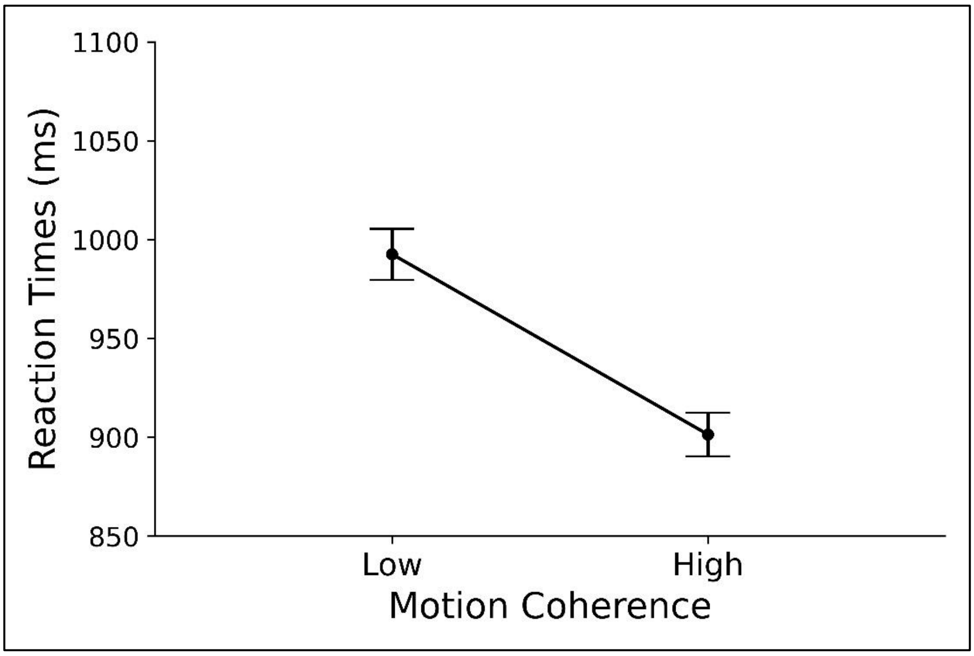

Figure 7. Effect of motion coherence on response times to the motion task. Bars represent the standard error of the mean.

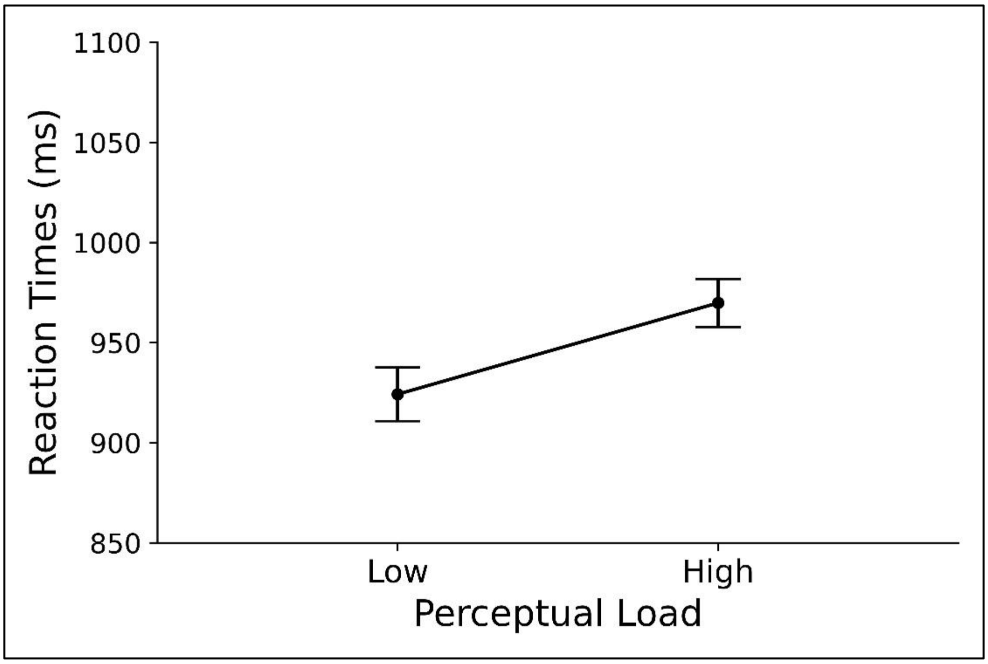

Figure 8. Effect of perceptual load on response times to the motion task. Bars represent the standard error of the mean.

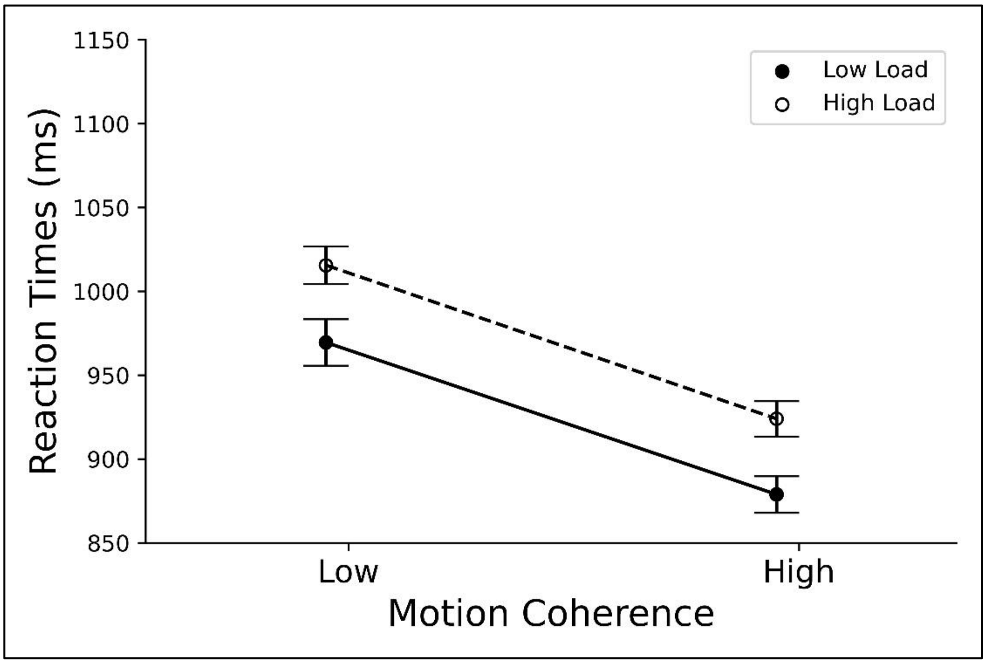

Figure 9. Effect of motion coherence on reaction times to the motion task at different levels of perceptual load. Bars represent the standard error of the mean.

fMRI results

A whole brain analyses was conducted in SPM looking at areas that showed greater activation in response to high-coherence motion compared to low-coherence motion. This analysis was performed both collapsing over perceptual load as well as separately for each perceptual load condition. None of these analyses produced any significant regions of activation at our whole-brain corrected threshold. We therefore needed to find an alternative method for defining a region of interest that could be used to interrogate the Load x Coherence interaction effect. Given the a priori interest in the V5 area, we defined our region of interest based on the V5 region as specified by the AAL mask images supplied with the WFU_PickAtlas toolbox version 3.0.5. Mean signal was extracted from this region of interest, separately in the four possible combinations of motion coherence and perceptual load. These data were entered in a repeated-measures 2x2 ANOVA. This analysis showed a significant main effect of motion coherence: $F(1, 21) = 4.5$, $p = .045$ (Figure 10). Therefore, the independently-defined V5 ROI showed significantly greater signal at high than low motion coherence. The interaction between motion coherence and perceptual load, contrarily to our hypothesis, was not significant ($F(1,21) = 0.4$, $p = 0.54$). Two follow-up two-tailed paired t tests looked at the difference in activation between the low motion coherence and the high motion coherence conditions at each level of perceptual load. This difference in the average level of activation of V5 was found to be significant when perceptual load was low ($t(21) = 2.23$, $p = .037$, Cohen's $d = 0.48$), with greater activation under high motion coherence compared to low motion coherence, but was not significant when perceptual load was high ($t(21) = 0.9$, $p = 0.37$, Cohen's $d = 0.19$; Figure 11).

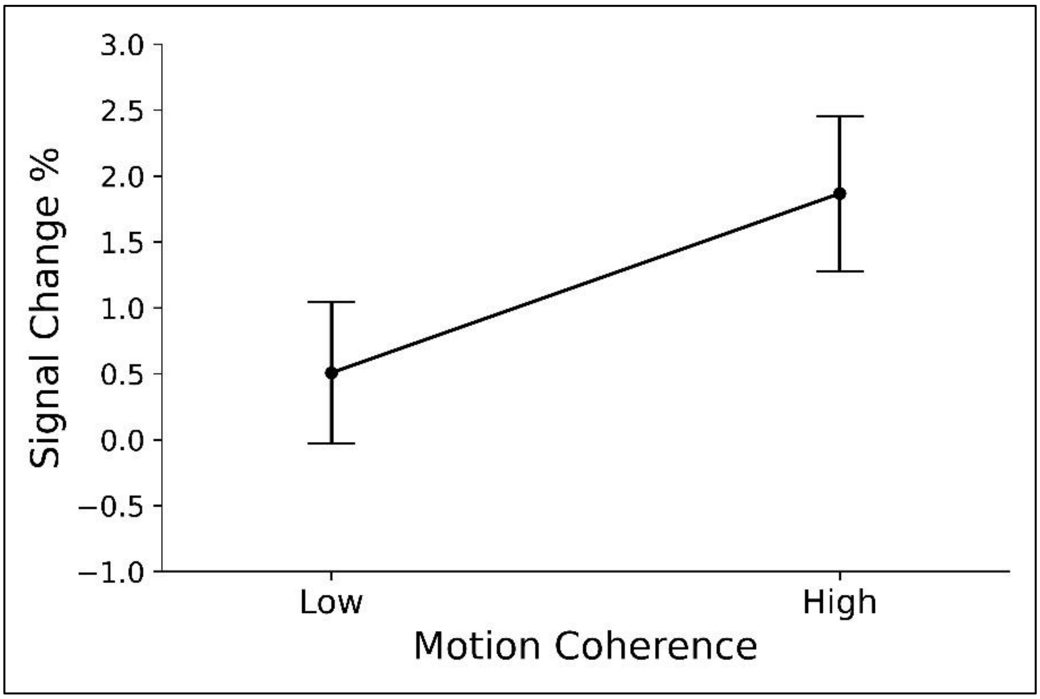

Figure 10. Effect of motion coherence on BOLD signal change in V5. Bars represent the standard error of the mean.

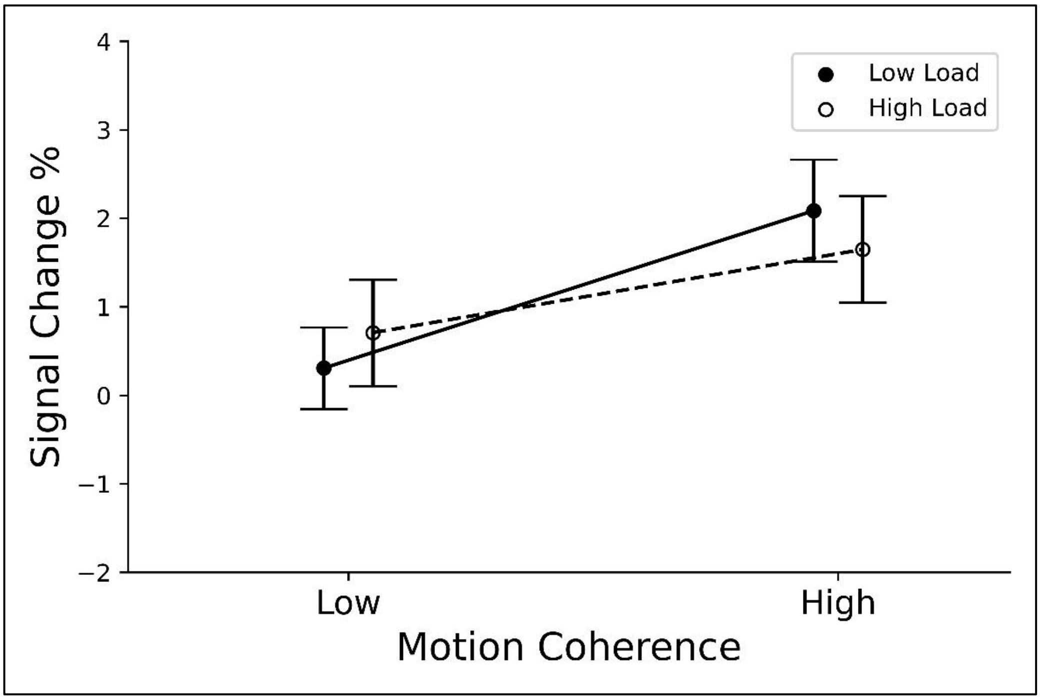

Figure 11. Effect of motion coherence on BOLD signal change in V5 at different levels of perceptual load. Bars represent the standard error of the mean.

Discussion

The aim of this experiment was to clarify the previously reported behavioural results, confirmed once again here, that saw reaction times to the motion task slower when following a high perceptual load cross trials as opposed to a low perceptual load trial. More in particular, the goal was to test an hypothesis derived from the perceptual load theory framework that links these results to differential level of motion processing. According to this theory, when perceptual load is low, stimuli that are not related to the task at hand are nonetheless processed. On the other hand, when perceptual load is high enough that the task exhausts all attentional capacity, processing of distractors is reduced or inhibited as a consequence. Rees et al. (1997) showed that concurrent processing of unrelated background motion is reduced in V5 while a high perceptual load task is performed, relative to a low perceptual task. Here instead, participants had to judge motion direction after a cross task of varied perceptual load. By performing the motion task at different levels of motion coherence (i.e., 40% and 70% coherence) it was possible to compute the difference in BOLD signal and observe how the expected increase in motion processing is affected by different levels of perceptual load.

The V5 area was the main focus of the analysis due to the a priori interest in it and also in light of the fact that a first whole brain analysis failed to show areas where activation was greater in response to high coherence motion versus low coherence motion. The key test was the interaction between load and motion coherence, which, unfortunately, was not significant. However, individual follow-up t tests were in line with the predicted pattern of results. When the perceptual load of the cross task was low, the BOLD signal was significantly higher with high motion coherence as opposed to low motion coherence, in line with the hypothesis. Also in line with the hypothesis is that the same t test on high perceptual load trials showed no significant difference in BOLD contrast between high and low motion coherence. This difference, though, was not strong enough to drive a significant interaction. Overall then, although our results are consistent with the initial hypothesis, they should be interpreted with caution, and more studies are needed to reach a stronger conclusion regarding the effect of perceptual load on task-switching.

One possibility, of course, is that there is indeed no difference in the level of processing received by the RDK as consequence of different levels of perceptual load, and that the difference in reaction times arises as a consequence of mechanisms acting at a later, post-perceptual stage.

Another possibility is that some aspect of the methods employed have contributed to the lack of a significant Load x Coherence interaction. One potential candidate is the proportion of trials containing a task-switch. In this experiment this proportion was set to 50%, double that of previous experiments. Increasing the probability of a task switch has been found to increase RT in repetition trials and reduce the switch cost (Monsell & Mizon, 2006). This happens because participants are likely to either engage proactively in task switch before the cue is given as they believe a task switch to be impending, or in light of the fact that they might try to maintain both task sets in a higher activation state. Also, the RDK was always visible, even while participants were only engaging with the cross task. These two factors might have contributed to generate the belief that maintaining a high readiness to switch to the motion task would be advantageous. A possible replication could focus on reducing the frequency of the task switch and on making the motion stimuli less conspicuous when the cross task is being performed.

All the experiments reported both in this chapter and the previous chapter share the same stimuli layout: the cross stimuli and the RDK are all displayed at the centre of the screen, and they overlap with each other. Future work could manipulate the location of the stimuli for the different tasks so that the RDK and crosses would occupy non-overlapping areas. For instance, the cross task could be placed just above fixation and the RDK slightly below (or vice-versa). Future studies might also explore the effect of different tone volumes on RTs and motion processing. Another useful addition, in conjunction with a different spatial arrangement of the stimuli, would be the adoption of an eye-tracker. A few studies (Longman et al., 2013, 2014; Mayr et al., 2013) have shown that one of the hallmarks of the attentional component of the task switch is the so-called attentional inertia: a lingering tendency to fixate the irrelevant (but previously relevant) area of the screen right after the switch. A more pronounced attentional inertia following a switch away from a high perceptual load trial would strengthen the case for a modulating effect of pre-switch attentional demands on post-switch performance.

General discussion

The material presented in the past four chapters has been inspired by the question of if and to what extent the attentional demands posed by non-driving secondary tasks can negatively impact the drivers' ability to resume control of the vehicle when needed.

This question is justified by the trend toward self-driving cars, which has now arrived to a stage where the most sophisticated vehicles are capable of taking care "of all aspects of the dynamic driving task with the expectation that the human driver will respond appropriately to a request to intervene" (SAE, 2021). This expectation regarding the driver's readiness to intervene cannot be left unchecked. It is well known that drivers already engage in a variety of non-driving tasks, and it is easy to predict that this tendency will increase in parallel with the increase in the automation of the driving task.

It is important to note, however, that no testing has been conducted in either a real or a simulated driving environment. Several studies conducted in driving simulators – reviewed in Chapter 1 – have addressed drivers' readiness to take back control of the vehicle while engaging in other tasks. This has often been achieved by comparing the impact of different non-driving tasks (such as reading the news, performing specific games, etc.) on a certain set of dependent measures (such as time-to-eyes-on-the-road or time-to-hands-on-the-steering-wheel). Nonetheless, a more systematic manipulation of pre-switch attentional demands and the assessment of such a manipulation on different components of the act of resuming control of the vehicle has not been carried out.

I have presented two different scenarios to elucidate the potential impact of a high level of perceptual load on two important aspects of the take-over process: perceiving the TOR (in Chapter 2) and correctly judge and react to the motion of stimuli when resuming the driving task (in Chapter 3 and Chapter 4).

Chapter 2 has extended the phenomenon of "inattentional deafness", which is the reduced detection of an auditory stimulus when performing a visual task of sufficiently high perceptual load, to a more ecological class of stimuli: natural scenes. This class of stimuli is commonly found in tasks that were most frequently chosen in a survey conducted with the purpose of better understanding what sort of activities people thought they would like to engage in while riding in a self-driving vehicle

(Sullman et al., 2017). Therefore, observing “inattentional deafness” to an auditory tone in a task that employs this kind of stimuli is of particular significance. In a second experiment, the main result of reduced tone perception under high load was replicated even though the tone was presented several hundreds of milliseconds (as long as 950 ms) after the onset of the natural scenes. A possible takeaway is that a unimodal, auditory take-over request should be avoided in favour of a multimodal TOR, perhaps combining haptic, auditory and visual signals all together.

In Chapter 3 and Chapter 4 instead, I have explored the effects of different levels of perceptual load on participants’ ability to correctly process the prevalent motion direction in a random dot kinematogram. Of course, judging the motion of vehicles and other stimuli in the environment correctly and quickly is vital upon regaining control of the vehicle in a sudden control authority transition. In three different experiments, participants reliably took longer to evaluate the prevalent motion direction when perceptual load right before the task-switch was high. Finally, in Chapter 4, I have tried to relay this slowing of reaction times to the level of motion processing in V5. While some results pointed to reduced motion processing in V5 under high load, the main hypothesis of an interaction between the level of motion coherence and perceptual load was not significant.

There are, however, some limitations to the experiments presented here that deserve to be addressed. For instance, all the work exposed in this thesis speaks of the negative effects of tasks characterized by a high level of perceptual load. These negative consequences manifest either on the probability of detecting a tone (intended as a proxy for a take-over request emitted by the car) or on the ability to quickly judge the prevalent motion direction in a field of moving dots (which is taken as a general index of the drivers’ capability to assess the motion of vehicles in a driving environment). However, given that the only manipulation has been that of perceptual load, how is it possible to exclude that any task, not just those that entail a high level of perceptual load, would have not yielded the same results? In other words, the effects observed on tone and motion detection could be due not to perceptual load but to a generic increase in task difficulty in any task, perceptual or not. Future studies could build on the present work by exploring different types of tasks, such as tasks requiring different levels of working memory in different conditions. At present, a partial response to this question can be provided, at least for the problem of tone detection reported in Chapter 2, considering the different predictions made by perceptual load theory regarding perceptual and non-perceptual manipulations. While high perceptual load should result in efficient selection and exclusion of distractors from processing, a high load on working memory functions should have the opposite effect. The argument is that processing priorities to perform a task are maintained in working memory in order to ensure that the relevant information is processed. When working memory capacity is taxed, the ability to exert control over the processing priorities is impaired, with the result that distractors are accidently processed. Increased identifications of distractors under high cognitive load has been shown in numerous studies (De Fockert et al., 2001; Konstantinou et al., 2014; Konstantinou & Lavie, 2013; Lavie et al., 2004; Lavie & De Fockert, 2005). This literature, however, has so far only focused on the different effects that loading perceptual capacity or working memory control processes have on a distractor presented in the visual modality. A future experiment could test the hypothesis that loading working memory control processes results in a higher incidental detection of an auditory stimulus.

Another general shortcoming has to do with a potentially low statistical power. Statistical power is inversely related to the probability of committing a type II error, which is the probability of not rejecting the null hypothesis when the null hypothesis is actually false. In other words, when statistical power is not sufficiently high there is a risk of not finding a statistically significant difference between conditions when a significant difference is actually present. This could have posed an issue not simply in all statistical test reported in the present thesis, but rather only in those that did not find a significant difference. There are in particular three interactions that deserve closer scrutiny.

Experiment 2 of Chapter 2 followed a 2 X 4 within-participants design. The factors were perceptual load (low, high) and tone onset time (0-250, 251-500, 501-750, 751-950 ms). The rationale here was to see how long the reduced tone detection under high perceptual load reported in Experiment 1 would perdure after image onset. In Experiment 1 the tone onset was fixed at 200 ms after image onset and could not inform on this point. The result of the interaction between perceptual load and tone onset time was not significant ($F(3, 45) = 0.55$, $p = 0.6$). If low statistical power is the reason why this test failed to find a significant difference, this would not call into question the main effect of reduced tone detection under high perceptual load. This finding was in fact not only assessed separately in Experiment 1 but also confirmed in

Experiment 2 by the main effect of perceptual load on tone detection. Rather, it would call into question the exact temporal dynamics of this effect. Given that the tone onset time was set at 200 ms after image onset in Experiment 1, the uncertainty would likely be circumscribed to later tone onset times (i.e., between 200 and 950 ms). The second interaction was reported in Experiment 3 of Chapter 3. Given that the experiments reported in this chapter used what was, to the authors' best knowledge, a novel design, the decision about the sample size was complicated by the lack of knowledge of the hypothesized effect size. For this reason, sample size for main tests was chosen according to the most conceptually similar study found in the literature (Carmel et al., 2011). This choice should have afforded a 90% power with an α of 0.05. Both Experiment 1 and Experiment 2 found a significant difference in the reaction times to the motion task as a function of the level of perceptual load in the preceding cross task. This test was the rationale for performing both experiments and the fact that a significant difference was indeed found suggests that the estimate of the required sample size was at least reasonably accurate. In Experiment 3 instead we wanted to test whether the presence of the loading stimuli (i.e., the crosses) at the moment of the task switch request was necessary to observe the results found in the previous experiments. A previous study (Carmel et al., 2011) suggested that as long as the stimuli were still being processed, their physical presence was not a strict requirement. For this reason, a mixed design experiment was conducted with a sample size of 40 participants. However, interactions in mixed and between participants designs require a very high sample size in order to achieve a power similar to that of a main effects, as much as 16 times the sample size needed to find a main effect (Gelman, 2018). The result of the interaction between perceptual load and tone onset time was $F(1,38) = 3.8$, $p= 0.058$, $\eta p2= 0.002$. It is therefore entirely possible, if not very likely, that this interaction test suffered from low statistical power and that with a more congruous sample size a significant difference would have emerged. This would have indicated that it is just as easy to disengage from a low perceptual load task as it is from a high perceptual load task. Given the novelty of this design, a replication of Experiment 3 with sample size tailored to properly tackle such an interaction could be the subject of a future study. Finally, there is the matter of the interaction reported in the fMRI experiment in Chapter 4. Here, we explored a possible interaction between perceptual load and motion coherence. The result of the interaction was: $F(1,21) = 0.4$, $p = 0.54$. The two follow-up t tests that looked at the different activation in V5 between the low and high motion coherence conditions at each level of perceptual load showed indeed results in line with the initial hypothesis. Finding a significant interaction would have therefore more strongly supported the conclusion that high perceptual load levels in the prior cross task reduced motion processing in V5 during the following motion task.

In the introduction I have analysed some of the criticisms that have been moved against Perceptual Load Theory. According to this theory a high level of selectivity, which is evidenced by the fact that distractors are not able to influence response times in the flanker task, is only possible when resources are exhausted and none can be spared for the processing of the distractors. This reliance on the concept of limited resources has been called into question and some studies have recast the role of perceptual load, grounding it instead on the idea of a competition for representation in visual cortex by multiple stimuli, which is then biased through a top-down signal (Desimone & Duncan, 1995; Lavie & Torralbo, 2010; Scalf et al., 2013; Torralbo & Beck, 2008). Furthermore, it has been shown that the attentional set, or the expectations about the upcoming level of perceptual load can also play a role. Theeuwes et al. (2004) have shown that when perceptual load is allowed to change between trials (as opposed to between blocks of trials) a flanker compatibility effect can also happen in high load trials when they are preceded by a low load trial. They therefore hypothesized that participants tend to expect the same level of perceptual load and this attentional set is carried over to the next trial. In other words, high selectivity in high load conditions would not be a necessity brought upon by the exhaustion of resources, but rather a consequence of the expectations about the level of perceptual load in the upcoming trial. Specularly, Johnson et al. (2002) have shown that when a 100% valid cue about the location of the target is given in a flanker task, participants are able to ignore the distractor even under a low level of perceptual load. Taken together, these evidences indicate that, under the proper conditions, participants' expectations could potentially set the spatial spread of their attention according to the task demands: broader when load is low and narrower when load is high. Could this mechanism be at play in some of the experiments reported here? In Chapter 2 the goal was to assess the effect of visual perceptual load on perception of an auditory stimulus. It seems therefore unlikely that a broadening or narrowing of spatial attention could have any bearing on the detection of a stimulus presented in a different modality. This possibility is instead present in the experiments reported in

Chapter 3 and 4, where the main finding is that it reliably takes longer to evaluate the prevalent motion direction of a set of dots when the perceptual load in the previous task (i.e., the cross task) is high. As I previously touched on, there is the concrete possibility of task set inertia occurring between the cross task and the motion task. This is made more likely by the fact that the task switch occurs abruptly, with no time for the previous task set to dissipate before the onset of the motion task (Allport et al., 1994; Wylie & Allport, 2000). As Theeuwes et al. (2004) have shown, one aspect of the task set than can be carried forward is the spread of spatial attention. The dots employed in the experiments in Chapter 3 and 4 were shown over a square field encompassing 16 X16 degrees of visual angle, with the crosses measuring 2.49 X 1.24 degrees at its centre. If attention is indeed spread more broadly in the low perceptual load condition, at the onset of the motion task participants would be able to gather information about the motion of the dots from a much greater portion of the field of dots, compared to the smaller area afforded by a narrower spatial distribution of attention in the high perceptual load condition. Therefore, an alternative account of the results on the basis of different attentional sets cannot be ruled out.

A final, general concern is related to the way in which the interaction between the driver and the self-driving vehicle has been modelled in this thesis. While all the experiments were done in a laboratory setting, the choices about the stimuli and the tasks were informed by the scenario underpinning this work. This scenario is predicated upon the belief that once in a self-driving vehicle drivers will be willing to engage in all sorts of non-driving tasks. This is not just an assumption. In fact, the choice of the stimuli employed in the experiments reported in Chapter 2 was performed according to the results of a survey conducted for this very purpose (Sullman et al., 2017). The respondents indicated a high level of interest in performing various non-driving tasks even when imagining themselves in a moderately brief car ride (60 minutes). Their top picks were: browsing the internet, reading and keeping up with the news. Given that self-driving cars have not yet reached widespread adoption, the validity of this belief is to some extent unclear. First of all, drivers might not be inclined to be totally immersed in a non-driving task, considering the severe negative repercussions of not being able to resume control of the vehicle in a timely and effective fashion. At least for a good percentage of the population, a profound trust in the capability of the automation will have to be built before drivers might feel comfortable to perform non-driving tasks without at least trying to simultaneously pay attention to the road. In addition to this, some literature has found that the kind of activities that have been thought to be of interest to prospective drivers of automated vehicles (e.g., browsing the internet, looking up social media, watching videos, etc.) might significantly increase the risk of motion sickness (Diels & Bos, 2016). If this is indeed the case, participants might decide to give up performing these tasks when driving. That, of course, is not to say that drivers will be attentive at all times. Common experience and a wealth of evidence tell us that drivers do perform non-driving tasks even in cars that do not possess any level of automation (for a review, Young & Regan, 2007). Should this new scenario be more in line with what is to come this would to some extent limit the applicability of the current findings, as they would come to reflect more of a worst-case scenario than a common occurrence. Nonetheless, it remains important to have a sense of how profound a deficit drivers would suffer in their capacity to promptly take back control of the vehicle should they choose to pursue these tasks while driving. While preparing for and understanding the worst-case scenario is of great importance, totally unstructured authority transitions will hopefully be only a last resort. Future work might also focus on gaining a better understanding of how to optimally support drivers' situation awareness while negotiating between monitoring the environment and performing non-driving tasks. Similarly, another venue of research could investigate how to structure the take-over request and when to deliver it in relation to different phases of the non-driving tasks in order to maximize the drivers' ability to regain control of the car.

References

Allport, A., Styles, E. A., & Hsieh, S. (1994). Shifting Intentional Set: Exploring the Dynamic Control of Tasks. In *Attention and Performance XV*. Cambridge, MA: MIT Press. https://doi.org/10.7551/mitpress/1478.003.0025

Altmann, E. M., & Gray, W. D. (2008). An Integrated Model of Cognitive Control in Task Switching. *Psychological Review*, *115*(3), 602–639. https://doi.org/10.1037/0033-295X.115.3.602

Arrington, C. M., & Logan, G. D. (2005). Voluntary task switching: Chasing the elusive homunculus. *Journal of Experimental Psychology: Learning Memory and Cognition*, *31*(4), 683–702. https://doi.org/10.1037/0278-7393.31.4.683

Bahrami, B., Carmel, D., Walsh, V., Rees, G., & Lavie, N. (2008). Unconscious orientation processing depends on perceptual load. *Journal of Vision*, *8*(3), 1–10. https://doi.org/10.1167/8.3.12

Bahrami, B., Lavie, N., & Rees, G. (2007). Attentional Load Modulates Responses of Human Primary Visual Cortex to Invisible Stimuli. *Current Biology*, *17*(6), 509–513. https://doi.org/10.1016/j.cub.2007.01.070

Beck, D. M., & Lavie, N. (2005). Look here but ignore what you see: Effects of distractors at fixation. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(3), 592–607. https://doi.org/10.1037/0096-1523.31.3.592

Bishop, S. J., Jenkins, R., & Lawrence, A. D. (2007). Neural processing of fearful faces: Effects of anxiety are gated by perceptual capacity limitations. *Cerebral Cortex*, *17*(7), 1595–1603. https://doi.org/10.1093/cercor/bhl070

Britten, K. H., Shadlen, M. N., Newsome, W. T., & Movshon, J. A. (1993). Responses of neurons in macaque MT to stochastic motion signals. *Visual Neuroscience*, *10*(6), 1157–1169. https://doi.org/10.1017/S0952523800010269

Broadbent, D. E. (1958). *Perception and communication. London: Pergamon*. Pergamon Press.

Carmel, D., Saker, P., Rees, G., & Lavie, N. (2007a). Perceptual load modulates conscious fl icker perception. *Journal of Vision*, *7*(14), 1–13. https://doi.org/10.1167/7.14.14.Introduction

Carmel, D., Saker, P., Rees, G., & Lavie, N. (2007b). Perceptual load modulates conscious flicker perception. *Journal of Vision*, *7*(14), 1–13.

https://doi.org/10.1167/7.14.14

Carmel, D., Thorne, J. D., Rees, G., & Lavie, N. (2011). Perceptual Load Alters Visual Excitability. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(5), 1350–1360. https://doi.org/10.1037/a0024320

Cartwright-Finch, U., & Lavie, N. (2007). The role of perceptual load in inattentional blindness. *Cognition*, *102*(3), 321–340. https://doi.org/10.1016/j.cognition.2006.01.002

Chan, J. L., & Desouza, J. F. X. (2013). The effects of attentional load on saccadic task switching. *Experimental Brain Research*, *227*(3), 301–309. https://doi.org/10.1007/s00221-013-3452-1

Chan, J. L., Kucy, A., & Desouza, J. F. X. (2017). Stable Task Representations under Attentional Load Revealed with Multivariate Pattern Analysis of Human Brain Activity. *Washington, DC: APA, Guideline Development Panel for the Treatment of Posttraumatic Stress Disorder in Adults.*, *December*, 139. https://doi.org/10.1162/jocn

Chaudhuri, A. (1990). Modulation of the motion aftereffect. *Nature*, *344*(March), 60–62.

Chen, Z., & Cave, K. R. (2016). Zooming in on the cause of the perceptual load effect in the Go/No-Go paradigm. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(8), 1072–1087. https://doi.org/10.1037/xhp0000168

Clark, H., Mclaughlin, A. C., Williams, B., & Feng, J. (2017). *Performance in Takeover and Characteristics of Non-driving Related Tasks during Highly Automated Driving in Younger and Older Drivers*. 37–41.

Cohen, M. A., Alvarez, G. A., & Nakayama, K. (2011). Natural-Scene Perception Requires Attention. *Psychological Science*, *22*(9), 1165–1172. https://doi.org/10.1177/0956797611419168

De Fockert, J. W., Rees, G., Frith, C. D., & Lavie, N. (2001). *The Role of Working Memory in Visual Selective Attention Published by : American Association for the Advancement of Science Stable URL : http://www.jstor.com/stable/3082574 Linked references are available on JSTOR for this article : You may need to log in*. *291*(5509), 1803–1806.

De Jong, R. (2000). *Control of Cognitive Processes: Attention and Performance XVIII - International Symposium on Attention and Performance - Google Books* (Control of). https://books.google.co.uk/books?hl=en&lr=&id=kO_baYlSVbwC&oi=fnd&pg=PA175&dq=lavie+2000&ots=pt1HGIimTZ&sig=kN6Kv34lp8cq4BnwVLR-eq4OxRY#v=onepage&q&f=false

Desimone, R., & Duncan, J. (1995). Neural Mechanisms of Selective Visual Attention. *Annual Review of Neuroscience*, *18*(1), 193–222. https://doi.org/10.1146/annurev-psych-122414-033400

Deutsch, J. A., & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological Review*, *70*(1), 51–60. https://doi.org/10.1037/h0042712

Diels, C., & Bos, J. E. (2016). Self-driving carsickness. *Applied Ergonomics*, *53*, 374–382. https://doi.org/10.1016/j.apergo.2015.09.009

Dobkins, K. R., & Bosworth, R. G. (2001). Effects of set-size and selective spatial attention on motion processing. *Vision Research*, *41*(12), 1501–1517. https://doi.org/10.1016/S0042-6989(01)00038-4

Duncan, J., Emslie, H., Williams, P., Johnson, R., & Freer, C. (1996). Intelligence and the frontal lobe: The organization of goal-directed behavior. *Cognitive Psychology*, *30*(3), 257–303. https://doi.org/10.1006/cogp.1996.0008

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity (1989). *Attention, Perception and Action: Selected Works of Glyn Humphreys*, *96*(3), 70–123. https://doi.org/10.4324/9781315712819-8

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*, 143–149. https://doi.org/10.1007/978-3-319-57111-9_9085

Eriksson, A., & Stanton, N. A. (2017). Takeover Time in Highly Automated Vehicles: Noncritical Transitions to and From Manual Control. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *59*(4), 689–705. https://doi.org/10.1177/0018720816685832

Evans, K. K., & Treisman, A. (2005). Perception of Objects in Natural Scenes: Is It Really Attention Free? *Journal of Experimental Psychology: Human Perception and Performance*, *31*(6), 1476–1492. https://doi.org/10.1037/0096-1523.31.6.1476

Fabre-Thorpe, M., Delorme, A., Marlot, C., & Thorpe, S. (2001). A Limit to the Speed of Processing in Ultra-Rapid Visual Categorization of Novel Natural Scenes. *Journal of Cognitive Neuroscience*, *13*(2), 171–180.

https://doi.org/10.1162/089892901564234

Forster, S., & Lavie, N. (2008). Attentional capture by entirely irrelevant distractors. *Visual Cognition*, *16*(2–3), 200–214. https://doi.org/10.1080/13506280701465049

Gelman, A. (2018). *You need 16 times the sample size to estimate an interaction than to estimate a main effect*. https://statmodeling.stat.columbia.edu/2018/03/15/need-16-times-sample-size-estimate-interaction-estimate-main-effect/

Gold, C., Berisha, I., & Bengler, K. (2015). *Utilization of Drivetime – Performing Non-Driving Related Tasks While Driving Highly Automated*. *1*(2013), 1666–1670.

Gold, C., Damböck, D., Lorenz, L., & Bengler, K. (2013). “Take over!” How long does it take to get the driver back into the loop? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *57*(1), 1938–1942. https://doi.org/10.1177/1541931213571433

Gopher, D., Armony, L., & Greenshpan, Y. (2000). Switching tasks and attention policies. *Journal of Experimental Psychology: General*, *129*(3), 308–339. https://doi.org/10.1037//0096-3445.129.3.308

Henson, R. N. A., & Rugg, M. D. (2003). Neural response suppression, haemodynamic repetition effects, and behavioural priming. *Neuropsychologia*, *41*(3), 263–270. https://doi.org/10.1016/S0028-3932(02)00159-8

Johnson, D. N., McGrath, A., & McNeil, C. (2002). Cuing interacts with perceptual load in visual search. *Psychological Science*, *13*(3), 284–287. https://doi.org/10.1111/1467-9280.00452

Kahneman, D. (1973). *Attention and effort* (Englewood Cliffs (ed.)). Prentice-Hall.

Kahneman, D., & Treisman, A. (1984). Changing views of attention and automaticity. In *Varieties of Attention* (pp. 29–61).

Kastner, S., De Weerd, P., Pinsk, M. A., Elizondo, M. I., Desimone, R., & Ungerleider, L. G. (2001). Modulation of sensory suppression: Implications for receptive field sizes in the human visual cortex. *Journal of Neurophysiology*, *86*(3), 1398–1411. https://doi.org/10.1152/jn.2001.86.3.1398

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., & Koch, I. (2010). Control and interference in task switching-a review. *Psychological Bulletin*, *136*(5), 849–874. https://doi.org/10.1037/a0019842

Kirchner, H., & Thorpe, S. J. (2006). Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited. *Vision Research*, *46*(11), 1762–1776. https://doi.org/10.1016/j.visres.2005.10.002

Konstantinou, N., Beal, E., King, J.-R., & Lavie, N. (2014). Working memory load and distraction: dissociable effects of visual maintenance and cognitive control. *Attention, Perception, & Psychophysics*, *76*(7), 1985–1997. https://doi.org/10.3758/s13414-014-0742-z

Konstantinou, N., & Lavie, N. (2013). Dissociable roles of different types of working memory load in visual detection. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(4), 919–924. https://doi.org/10.1037/a0033037

Lankheet, M. J. M., & Verstraten, F. A. J. (1995). Attentional modulation of adaptation to two-component transparent motion. *Vision Research*, *35*(10), 1401–1412. https://doi.org/10.1016/0042-6989(95)98720-T

Lavie, N. (1995). Perceptual Load as a Necessary Condition for Selective Attention. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(3), 451–468. https://doi.org/10.1037/0096-1523.21.3.451

Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in Cognitive Sciences*, *9*(2), 75–82. https://doi.org/10.1016/j.tics.2004.12.004

Lavie, N. (2006). The role of perceptual load in visual awareness. *Brain Research*, *1080*(1), 91–100. https://doi.org/10.1016/j.brainres.2005.10.023

Lavie, N., & Cox, S. (1997). On the Efficiency of Visual Selective Attention: Efficient Visual Search Leads to Inefficient Distractor Rejection. *Psychological Science*, *8*(5), 395–396. https://doi.org/10.1111/j.1467-9280.1997.tb00432.x

Lavie, N., & De Fockert, J. W. (2003). Contrasting effects of sensory limits and capacity limits in visual selective attention. *Perception and Psychophysics*, *65*(2), 202–212. https://doi.org/10.3758/BF03194795

Lavie, N., & De Fockert, J. W. (2005). *The role of working memory in attentional capture*. *12*(4), 669–674.

Lavie, N., & Fox, E. (2000). The role of perceptual load in negative priming. *Brain Research*, *1080*(1), 91–100. https://doi.org/10.1016/j.brainres.2005.10.023

Lavie, N., Hirst, A., De Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General*, *133*(3), 339–354. https://doi.org/10.1037/0096-3445.133.3.339

Lavie, N., & Torralbo, A. (2010). Dilution: A Theoretical Burden or Just Load? A Reply to Tsal and Benoni (2010). *Journal of Experimental Psychology: Human*

Perception and Performance, *36*(6), 1657–1664. https://doi.org/10.1037/a0020733

Lavie, N., & Tsal, Y. (1994a). Perceptual load as a major determinant of. *Perception & Psychophysics*, *56*(2), 183–197.

Lavie, N., & Tsal, Y. (1994b). Perceptual load as a major determinant of the locus of selection in visual attention. *Perception & Psychophysics*, *56*(2), 183–197. https://doi.org/10.3758/BF03213897

Li, F. F., VanRullen, R., Koch, C., & Perona, P. (2002). Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Sciences*, *99*(14), 9596–9601. https://doi.org/10.1073/pnas.092277599

Lien, M. C., Ruthruff, E., & Johnston, J. C. (2010). Attentional Capture With Rapidly Changing Attentional Control Settings. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(1), 1–16. https://doi.org/10.1037/a0015875

Linnell, K. J., Caparos, S., de Fockert, J. W., & Davidoff, J. (2013). Urbanization decreases attentional engagement. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(5), 1232–1247. https://doi.org/10.1037/a0031139

Liu, T., Fuller, S., & Carrasco, M. (2006). \. *Psychonomic Bulletin and Review*, *13*(6), 1091–1096. https://doi.org/10.3758/BF03213931

Logan, G. D., & Gordon, R. D. (2001). Executive control of visual attention in dual-task situations. *Psychological Review*, *108*(2), 393–434.

Longman, C. S., Lavric, A., & Monsell, S. (2013). More attention to attention? An eye-tracking investigation of selection of perceptual attributes during a task switch. *Journal of Experimental Psychology: Learning Memory and Cognition*, *39*(4), 1142–1151. https://doi.org/10.1037/a0030409

Longman, C. S., Lavric, A., Munteanu, C., & Monsell, S. (2014). Attentional inertia and delayed orienting of spatial attention in task-switching. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(4), 1580–1602. https://doi.org/10.1037/a0036552

Lorenz, L., Kerschbaum, P., & Schumann, J. (2014). Designing take over scenarios for automated driving : How does augmented reality support the driver to get back into the loop? *Proceedings of the Human Factors and Ergonomics Society 58th*

Annual Meeting - 2014, *1*, 1681–1685. https://doi.org/10.1177/1541931214581351

Macdonald, J. S. P., & Lavie, N. (2008). Load induced blindness. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(5), 1078–1091. https://doi.org/10.1037/0096-1523.34.5.1078

Macdonald, J. S. P., & Lavie, N. (2011). Visual perceptual load induces inattentional deafness. *Attention, Perception, & Psychophysics*, *73*(6), 1780–1789. https://doi.org/10.3758/s13414-011-0144-4

Mack, A., & Rock, I. (1998). Inattentional Blindness. *Cambridge MA MIT Press Malik J Perona P*, *7*(1998), 287. https://doi.org/10.1016/j.aorn.2010.03.011

Mayr, U., Kuhns, D., & Rieter, M. (2013). Eye movements reveal dynamics of task control. *Journal of Experimental Psychology: General*, *142*(2), 489–509. https://doi.org/10.1037/a0029353

Meiran, N. (1996). Reconfiguration of processing mode prior to task performance. *Journal of Experimental Psychology: Learning Memory and Cognition*, *22*(6), 1423–1442. https://doi.org/10.1037/0278-7393.22.6.1423

Meiran, N. (2000a). Modeling cognitive control in task-switching. *Psychological Research*, *63*(3–4), 234–249. https://doi.org/10.1007/s004269900004

Meiran, N. (2000b). Reconfiguration of stimulus task sets and response task sets during task switching. In *Attention and Performance* (Vol. 18, Issue May).

Meiran, N., Kessler, Y., & Adi-Japha, E. (2008). Control by action representation and input selection (CARIS): A theoretical framework for task switching. *Psychological Research*, *72*(5), 473–500. https://doi.org/10.1007/s00426-008-0136-8

Merat, N., Jamson, A. H., Lai, F. C. H., & Carsten, O. (2012). Highly Automated Driving, Secondary Task Performance, and Driver State. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *54*(5), 762–771. https://doi.org/10.1177/0018720812442087

Mok, B., Johns, M., Lee, K. J., Miller, D., Sirkin, D., Ive, P., & Ju, W. (2015). Emergency, Automation Off: Unstructured Transition Timing for Distracted Drivers of Automated Vehicles. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, *2015-Octob*, 2458–2464. https://doi.org/10.1109/ITSC.2015.396

Molloy, K., Griffiths, T. D., Chait, M., & Lavie, N. (2015). Inattentional Deafness: Visual Load Leads to Time-Specific Suppression of Auditory Evoked Responses.

Journal of Neuroscience, *35*(49), 16046–16054. https://doi.org/10.1523/JNEUROSCI.2931-15.2015

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, *7*(3), 134–140. https://doi.org/10.1016/S1364-6613(03)00028-7

Monsell, S., & Mizon, G. A. (2006). Can the task-cuing paradigm measure an endogenous task-set reconfiguration process? *Journal of Experimental Psychology: Human Perception and Performance*, *32*(3), 493–516. https://doi.org/10.1037/0096-1523.32.3.493

Moore, C. M. (2001). Inattentional blindness: Perception or memory and what does it matter. *Psyche*, *7*(2), 178–194.

Mozer, M., & Sitton, M. (1998). Computational modeling of spatial attention. *Attention*, *9*, 341–393.

Murphy, G., & Greene, C. M. (2016). Perceptual Load Induces Inattentional Blindness in Drivers. *Applied Cognitive Psychology*, *30*(3), 479–483. https://doi.org/10.1002/acp.3216

Nagle, F., & Lavie, N. (2020). Predicting human complexity perception of real-world scenes. *Royal Society Open Science*, *7*(5). https://doi.org/10.1098/rsos.191487

Neubauer, C., Matthews, G., & Saxby, D. (2012). The Effects of Cell Phone Use and Automation on Driver Performance and Subjective State in Simulated Driving. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *56*(1), 1987–1991. https://doi.org/10.1177/1071181312561415

O'Connor, D. H., Fukui, M. M., Pinsk, M. A., & Kastner, S. (2002). Attention modulates responses in the human lateral geniculate nucleus. *Nature Neuroscience*, *5*(11), 1203–1209. https://doi.org/10.1038/nn957

Parks, N. A., Hilimire, M. R., & Corballis, P. M. (2009). Visual perceptual load modulates an auditory microreflex. *Psychophysiology*, *46*(3), 498–501. https://doi.org/10.1111/j.1469-8986.2009.00801.x

Parks, N. A., Hilimire, M. R., & Corballis, P. M. (2011). Steady-state signatures of visual perceptual load, multimodal distractor filtering, and neural competition. *Journal of Cognitive Neuroscience*, *23*(5), 1113–1124. https://doi.org/10.1162/jocn.2010.21460

Peelen, M. V., Fei-Fei, L., & Kastner, S. (2009). Neural mechanisms of rapid natural scene categorization in human visual cortex. *Nature*, *460*(7251), 94–97. https://doi.org/10.1038/nature08103

Peelen, M. V., & Kastner, S. (2011). A neural basis for real-world visual search in human occipitotemporal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(29), 12125–12130. https://doi.org/10.1073/pnas.1101042108

Peelen, M. V., & Kastner, S. (2014). Attention in the real world: Toward understanding its neural basis. *Trends in Cognitive Sciences*, *18*(5), 242–250. https://doi.org/10.1016/j.tics.2014.02.004

Pessoa, L., McKenna, M., Gutierrez, E., & Ungerleider, L. G. (2002). Neural processing of emotional faces requires attention. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(17), 11458–11463. https://doi.org/10.1073/pnas.172403899

Potter. (1975). Meaning in visual search. *Science*, *March*, 965–966.

Radlmayr, J., Gold, C., Lorenz, L., Farid, M., & Bengler, K. (2014). How Traffic Situations and Non-Driving Related Tasks Affect the Take-Over Quality in Highly Automated Driving. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *58*(1), 2063–2067. https://doi.org/10.1177/1541931214581434

Raveh, D., & Lavie, N. (2015). Load-induced inattentional deafness. *Attention, Perception, & Psychophysics*, *77*(2), 483–492. https://doi.org/10.3758/s13414-014-0776-2

Rees, G., Friston, K., & Koch, C. (2000). A direct quantitative relationship between the functional properties of human and macaque V5. *Nature Neuroscience*, *3*(7), 716–723. https://www.nature.com/articles/nn0700_716.pdf

Rees, G., Frith, C., & Lavie, N. (1997). Modulating Irrelevant Motion Perception by Varying Attentional Load in an Unrelated Task. *Science*, *278*(5343), 1616–1619. https://doi.org/10.1126/science.278.5343.1616

Rees, Geraint, Frith, C. D., & Lavie, N. (1997). Modulating irrelevant motion perception by varying attentional load in an unrelated task. *Science*, *278*(5343), 1616–1619. https://doi.org/10.1126/science.278.5343.1616

Remington, A., Cartwright-Finch, U., & Lavie, N. (2014). I can see clearly now: the effects of age and perceptual load on inattentional blindness. *Frontiers in Human Neuroscience*, *8*(April), 1–11. https://doi.org/10.3389/fnhum.2014.00229

Reynolds, J. H., Chelazzi, L., & Desimone, R. (1999). Competitive mechanisms subserve attention in macaque areas V2 and V4. *Journal of Neuroscience*, *19*(5), 1736–1753. https://doi.org/10.1523/jneurosci.19-05-01736.1999

Rogers, R. D., & Monsell, S. (1995). Costs of a Predictable Switch Between Simple Cognitive Tasks. *Journal of Experimental Psychology: General*, *124*(2), 207–231. https://doi.org/10.1037/0096-3445.124.2.207

Roper, Z. J. J., Cosman, J. D., & Vecera, S. P. (2013). Perceptual load corresponds with factors known to influence visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(5), 1340–1351. https://doi.org/10.1037/a0031616

Rousselet, G. A., Fabre-Thorpe, M., & Thorpe, S. J. (2002). Parallel processing in high-level categorization of natural images. *Nature Neuroscience*, *5*(7), 629–630. https://doi.org/10.1038/nn866

Rousselet, G. A., Thorpe, S. J., & Fabre-Thorpe, M. (2004). How parallel is visual processing in the ventral pathway? *Trends in Cognitive Sciences*, *8*(8), 363–370. https://doi.org/10.1016/j.tics.2004.06.003

SAE. (2021). *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*. https://www.sae.org/standards/content/j3016_202104/

Scalf, P. E., Torralbo, A., Tapia, E., & Beck, D. M. (2013). Competition explains limited attention and perceptual resources: Implications for perceptual load and dilution theories. *Frontiers in Psychology*, *4*(MAY), 1–9. https://doi.org/10.3389/fpsyg.2013.00243

Schwartz, S., Vuilleumier, P., Hutton, C., Maravita, A., Dolan, R. J., & Driver, J. (2005). Attentional load and sensory competition in human vision: Modulation of fMRI responses by load at fixation during task-irrelevant stimulation in the peripheral visual field. *Cerebral Cortex*, *15*(6), 770–786. https://doi.org/10.1093/cercor/bhh178

Shulman, G. L. (1993). Attentional effects of adaptation of rotary motion in the plane. *Perception*, *22*(8), 947–961. https://doi.org/10.1068/p220947

Sullman, M., Chen, L., Dongpu, C., Wildegger, T., Auger, D., Zhao, Y., & Lavie, N. (2017). *Internal TASCC report*.

Theeuwes, J., Kramer, A. F., & Belopolsky, A. V. (2004). Attentional set interacts with perceptual load in visual search. *Psychonomic Bulletin and Review*, *11*(4), 697–702. https://doi.org/10.3758/BF03196622

Thorpe, S. J., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. In *Nature* (Vol. 381, Issue 6582, pp. 520–522).

https://doi.org/10.1038/381520a0

Thorpe, S. J., Gegenfurtner, K. R., Fabre-Thorpe, M., & Bülthoff, H. H. (2002). Detection of animals in natural images using far peripheral vision. *European Journal of Neuroscience*, *14*(5), 869–876. https://doi.org/10.1046/j.0953-816X.2001.01717.x

Torralbo, A., & Beck, D. M. (2008). Perceptual-load-induced selection as a result of local competitive interactions in visual cortex. *Psychological Science*, *19*(10), 1045–1050. https://doi.org/10.1111/j.1467-9280.2008.02197.x

Treisman, A. M., & Gelade, G. (1980). A Feature-Integration Theory of Attention. *Cognitive Psychology*, *12*, 97–136.

VanRullen, R., & Thorpe, S. J. (2001). Is it a bird? Is it a plane? Ultra-rapid visual categorisation of natural and artifactual objects. *Perception*, *30*(6), 655–668. https://doi.org/10.1068/p3029

Walker, S., Stafford, P., & Davis, G. (2008). Ultra-rapid categorization requires visual attention: Scenes with multiple foreground objects. *Journal of Vision*, *8*(4), 21. https://doi.org/10.1167/8.4.21

Wolfe, J. (1999). Inattentional Amnesia. In V. Coltheart (Ed.), *Fleeting Memories* (pp. 71–94). MIT Press.

Wylie, G., & Allport, A. (2000). Task switching and the measurement of “switch costs.” *Psychological Research*, *63*(3–4), 212–233. https://doi.org/10.1007/s004269900003

Yantis, S., & Johnston, J. C. (1990). On the locus of visual selection: evidence from focused attention tasks. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(1), 135–149.

Yi, D. J., Woodman, G. F., Widders, D., Marois, P., & Chun, M. M. (2004). Neural fate of ignored stimuli: Dissociable effects of perceptual and working memory load. *Nature Neuroscience*, *7*(9), 992–996. https://doi.org/10.1038/nn1294

Young, K., & Regan, M. (2007). Driver distraction: A review of the current state-of-knowledge. *Distracted Driving*, *September*, 1–22.

Zeeb, K., Buchner, A., & Schrauf, M. (2015). What determines the take-over time? An integrated model approach of driver take-over after automated driving. *Accident Analysis and Prevention*, *78*, 212–221. https://doi.org/10.1016/j.aap.2015.02.023

Zeeb, K., Buchner, A., & Schrauf, M. (2016). Is take-over time all that matters? the impact of visual-cognitive load on driver take-over quality after conditionally automated driving. *Accident Analysis and Prevention*, *92*, 230–239. https://doi.org/10.1016/j.aap.2016.04.002

Zhang, J., & Mueller, S. T. (2005). A note on ROC analysis and non-parametric estimate of sensitivity. *Psychometrika*, *70*(1), 203–212. https://doi.org/10.1007/s11336-003-1119-8