Multiple Object Tracking and the Division of the Attentional Spotlight

in a Realistic Tracking Environment

by

Martin Jewell Lochner

A Thesis
presented to
The University of Guelph

In partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in
Psychology

Guelph, Ontario, Canada

© Martin J. Lochner, January 2012
ABSTRACT

MULTIPLE OBJECT TRACKING AND THE DIVISION OF THE ATTENTIONAL SPOTLIGHT IN A REALISTIC TRACKING ENVIRONMENT

Martin J. Lochner

University of Guelph, 2012

The multiple object tracking task (Pylyshyn and Storm, 1988) has long been a standard tool for use in understanding how we attend to multiple moving points in the visual field. In the current experiments, it is first demonstrated that this classical task can be adapted for use in a simulated driving environment, where it is commonly thought to apply. Standard requirements of driving (steering, maintaining headway) are shown to reduce tracking ability. Subsequent experiments (2a, 2b, 2c) investigate the way in which participants respond to events at target and distractor locations, and have bearing on Pylyshyn’s (1989) “indexing” hypothesis. The final experiment investigates the effect of the colour-composition of the tracking set on performance, and may have implications for our theoretical understanding of how tracking is performed.
ACKNOWLEDGEMENTS:

I would like to thank my family and friends, both those who could stand my pre-occupation with research, and those who couldn’t! I would also like to extend special thanks to my advisor, Dr. Lana Trick for her tireless efforts in clarifying my logic and my syntax. Likewise, this work could not have been completed without the help of DRIVE Lab administrator and technician, Ryan Toxopeus, as well as the on-going support of the many students and interns that have passed through the lab; thank-you also to the Neuroscience and Applied Cognitive Science (NACS) group for their patience and feedback during my intermittent seminar presentations. None of this would have been possible without our Canadian funding organizations, AUTO21, NSERC, and CANDrive, who deserve a large thank-you for their financial support. And last but not least, thank-you, dear reader, for your interest in this research.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>iii</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>viii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>ix</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1. Dividing the attentional spotlight</td>
<td>4</td>
</tr>
<tr>
<td>1.1 Divided attention in the real world</td>
<td>5</td>
</tr>
<tr>
<td>2. FINST theory: Fingers of INSTantiation (a.k.a. Visual Indexing)</td>
<td>6</td>
</tr>
<tr>
<td>2.1 FINSTS and multiple object tracking</td>
<td>7</td>
</tr>
<tr>
<td>2.2 Describing the visual indexing mechanism</td>
<td>8</td>
</tr>
<tr>
<td>2.3 Tracking set composition</td>
<td>11</td>
</tr>
<tr>
<td>3. Object tracking in realistic environments</td>
<td>12</td>
</tr>
<tr>
<td>3.1 Differences in control characteristics between laboratory multiple</td>
<td>14</td>
</tr>
<tr>
<td>Object tracking and real-world equivalent</td>
<td></td>
</tr>
<tr>
<td>3.2 Differences in visual characteristics between laboratory multiple</td>
<td>16</td>
</tr>
<tr>
<td>Object tracking and real-world equivalent</td>
<td></td>
</tr>
<tr>
<td>3.2.1 Composition of the visual display, presence of motion,</td>
<td>17</td>
</tr>
<tr>
<td>and 3d environments</td>
<td></td>
</tr>
<tr>
<td>4. Overall Summary</td>
<td>19</td>
</tr>
<tr>
<td>5. EXPERIMENT 1</td>
<td>20</td>
</tr>
<tr>
<td>5.1 Methods</td>
<td>21</td>
</tr>
</tbody>
</table>
Driving Metrics........................................................................................................... 62

9.3 Discussion............................................................................................................ 63

10. GENERAL DISCUSSION......................................................................................... 66

10.1 Summary............................................................................................................ 67

10.2 Implications....................................................................................................... 70

10.2.1 Preserving the notion of an object oriented, pre-attentive indexing

system (saving the wild FINSTS)........................................................................... 73

10.3 Limitations........................................................................................................ 74

10.4 Future Directions............................................................................................. 76

11. CONCLUSION..................................................................................................... 78

REFERENCES.......................................................................................................... 81

APPENDICES........................................................................................................... 87

Appendix A: Lane Change Data Collection Form.................................................... 87

Appendix B: Consent Form....................................................................................... 88

Appendix C: Subject Information Form..................................................................... 90

Appendix D: Simulator Adaptation Syndrome Pre-Screening Form......................... 91

Appendix E: Simulator Adaptation Syndrome Post-Test Form............................... 93

Appendix F: Debriefing Form................................................................................... 94

Appendix G: Graph of D-Prime Means.................................................................... 95

Appendix H: Correlation Charts – Tracking Accuracy vs. Driving Metrics............ 96
LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1</td>
<td>Time course of typical multiple object tracking task</td>
<td>3</td>
</tr>
<tr>
<td>Figure 2</td>
<td>Experiment 1 vehicle tracking stimuli</td>
<td>22</td>
</tr>
<tr>
<td>Figure 3</td>
<td>Calculation of predicted accuracy if guessing 1 or 2 out of 4 targets</td>
<td>26</td>
</tr>
<tr>
<td>Figure 4</td>
<td>Percentage of correctly identified targets for Experiment 1</td>
<td>26</td>
</tr>
<tr>
<td>Figure 5</td>
<td>Experiment 1 driving metrics</td>
<td>29</td>
</tr>
<tr>
<td>Figure 6</td>
<td>Experiment 2a vehicle tracking stimuli</td>
<td>34</td>
</tr>
<tr>
<td>Figure 7</td>
<td>Task-flow diagram of procedure for Experiment 2</td>
<td>35</td>
</tr>
<tr>
<td>Figure 8</td>
<td>Percentage of correctly identified targets for Experiment 2a</td>
<td>36</td>
</tr>
<tr>
<td>Figure 9</td>
<td>Experiment 2a change detection accuracy</td>
<td>38</td>
</tr>
<tr>
<td>Figure 10</td>
<td>Experiment 2a driving metrics</td>
<td>39</td>
</tr>
<tr>
<td>Figure 11</td>
<td>Experiment 2b vehicle tracking stimuli</td>
<td>42</td>
</tr>
<tr>
<td>Figure 12</td>
<td>Task-flow diagram of procedure for Experiment 2b</td>
<td>43</td>
</tr>
<tr>
<td>Figure 13</td>
<td>Percentage of correctly identified targets for Experiment 2b</td>
<td>44</td>
</tr>
<tr>
<td>Figure 14</td>
<td>Experiment 2b change detection accuracy and RT</td>
<td>46</td>
</tr>
<tr>
<td>Figure 15</td>
<td>Experiment 2c vehicle tracking stimuli</td>
<td>49</td>
</tr>
<tr>
<td>Figure 16</td>
<td>Percentage of correctly identified targets for Experiment 2c</td>
<td>51</td>
</tr>
<tr>
<td>Figure 17</td>
<td>Experiment 2c event detection accuracy and RT</td>
<td>52</td>
</tr>
<tr>
<td>Figure 18</td>
<td>Makovski and Jiang (2009) tracking accuracy for coloured stimuli</td>
<td>55</td>
</tr>
<tr>
<td>Figure 19</td>
<td>Example of “Heterogenous” tracking set from Experiment 3</td>
<td>60</td>
</tr>
<tr>
<td>Figure 20</td>
<td>Percentage of correctly identified targets for Experiment 3</td>
<td>62</td>
</tr>
<tr>
<td>Figure 21</td>
<td>Experiment 3 driving metrics</td>
<td>63</td>
</tr>
<tr>
<td>Figure 22</td>
<td>Thomas and Seiffert (2010) Experiment 1</td>
<td>71</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table 1: Experiment 1 predicted scores if guessing................................. 27
Table 2: Experiment 2a tracking set and possible vehicle changes.............. 33
Table 3: Experiment 2a predicted scores if guessing............................... 36
Table 4: Experiment 2b tracking set and possible vehicle-switch combinations ... 43
Table 5: Experiment 2b predicted scores if guessing............................... 46
Table 6: Experiment 2c tracking set.................................................. 50
Table 7: Experiment 2c predicted scores if guessing............................... 53
Table 8: Experiment 3, example of colour conditions.............................. 59
Table 9: Experiment 3 predicted scores if guessing................................. 62
INTRODUCTION

In order to navigate through, and perform actions upon the physical world, it is probable that humans form some kind of mental representation. Modern psychology tends to agree that something called “attention” determines what makes it into the representation, and what does not. The result of this is that thousands upon thousands of research hours have been devoted to determining precisely what attention is: how it operates, and what mediates the ability to apply it to a visual scene. Famously, William James once wrote that:

“Everyone knows what attention is. It is the taking possession by the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought” (James, 1890).

Interestingly, this quote highlights an important debate that has been present in the literature on attention. James writes that attention is the taking possession of “one out of ...several” possible foci of attention. This evokes the concept that attention is unitary, or focal in nature, having but one possible focus at a given moment in time. Evidence of the singular attentional focus is readily available to the thoughtful person. When we attempt to listen to more than one audio stream at a time, we tend to process one stream after the other, as opposed to processing the two simultaneously (e.g., Cherry, 1953; Moray, 1959). The same is generally true of visual attention, in that we have one very obvious focus of vision (the fovea), which moves from point to point, and can never be placed at two points simultaneously. Despite this historical tradition (e.g., James, 1890), and the once-generally accepted notion of a unitary “spotlight” of attention (Posner et al., 1980) or “zoom lens” of attention (Eriksen and St. James, 1986), there is accumulating evidence that things are not entirely so simple. Posner
himself described the means by which the focus of attention can be directed at a location that is spatially separate from the focal point of vision, the fovea. This effect has been widely replicated and investigated, and it was only a small step to reach the notion that attention could be allocated simultaneously to multiple non-contiguous locations in physical space.

One popular way of studying the phenomenon of attention to multiple points at once is the multiple object tracking (MOT) task (Pylyshyn and Storm, 1988), in which participants are required to keep track of a certain number of targets (typically between 1 and 5) amongst a larger number of objects (e.g., 10). A schematic representation of the task is available in Figure 1. In this task, participants are shown a number of identical objects (e.g., circles) on a computer monitor. A subset of these items are cued as targets by flashing them off and on a number of times. After the targets have been cued, all of the items begin to move about the computer screen, and the participant is required to mentally keep track of the location of the target items. After a set duration (e.g., 8 seconds) the motion stops, and one of the items is again cued, and the participant is required to report whether or not that item was a target. Using this partial-report methodology, where participants report whether a single cued item was a target or a distractor, the authors found that participants responded correctly on 86.4% of trials when tracking 5 objects amidst a set of 10 items total. Using a full-report methodology where the participant attempts to indicate all of the tracked objects (e.g., Pylyshyn, 2004), accuracy ranges from 85 to 90% for 5 objects when there are a total of 10 items in the display. The task was designed in such a way that, even if the attentional focus was able to shift at an exceptionally high speed (20 ms per degree of visual angle), a serial model of object tracking would only be able to achieve an accuracy of approximately 20%. Instead, participants were achieving 85% to 90% accuracy when required to track 5 independently moving targets when there were 10 items total. Pylyshyn (1989) argues that the high accuracy demonstrated in this task is evidence that the objects were being tracked in parallel.
The ability to attend to multiple objects at the same time is generally assumed to be vital to every-day tasks such as driving an automobile (Fencsik et al., 2007) or playing group sports (Horowitz et al., 2007). However, object tracking research is generally carried out using abstract 2-dimensional visual displays that consist of circles on a monochrome background, as shown in Figure 1. The current experiments will reproduce the multiple object tracking task in a real-world situation (driving an automobile), in which tracking is thought to occur (e.g., Fencsik et al, 2007; Feria, 2008; Horowitz et al., 2007; Oksama and Hyona, 2004), and where item heterogeneity (i.e., heterogeneity of different vehicles / road users) is the norm. In the following sections I will first describe the theoretical context that gave rise to the multiple object tracking task, and then describe a debate within the literature about the processes involved, and the type of information available to the hypothesized tracking mechanism. I propose three main experiments. Experiment 1 will investigate object tracking in a driving environment, and will measure tracking ability, as well as driving performance while performing the tracking task. In Experiments 2a, 2b and 2c, I will investigate the ability to detect changes in tracked vs. non-tracked vehicles on the roadway. In the third experiment I will investigate the effects of homogenous vs. heterogenous display compositions and how this relates to tracking as well as driving performance.

![Figure 1: Time course of typical multiple object tracking task. Stage 1: Cue targets (target items flash off and on for 5 sec). Stage 2: tracking (items move randomly around the computer screen); Stage 3: Report (items stop, and the participant uses the computer mouse to indicate the target objects)](image-url)
1. Dividing the attentional spotlight

Does the fact that participants can track the locations of multiple independently moving objects (e.g., Pylyshyn and Storm, 1988) indicate that the attentional spotlight can be divided? Support for this notion is mixed. As Jans et al. (2010) point out, there has been a shift in the way that visual attention is understood over the past two decades. A 1998 review concluded that “most of the evidence favors the unified model” of attention (McCormick et al., 1998, pp. 350). Historically, the notion that visual-spatial attention has but a single focus has enjoyed wide acceptance among researchers (Eriksen and Yeh, 1985; Heinze et al., 1994; McCormick and Klein, 1990; Pan and Eriksen, 1993; Posner et al., 1980). Visual attention has been generally understood to operate in a “spotlight” (Posner, 1980) or a “zoom lens” (Eriksen and St. James, 1986) which covers a single contiguous region in space, inside of which processing is facilitated (as evidenced by generally faster reaction times and greater response accuracy). However, recent investigations into the nature of visual focal attention have given rise to the notion that the attentional spotlight may actually be divisible, as demonstrated by the fact that we can process information from multiple non-contiguous locations at once.

Two papers in particular from the recent literature are commonly cited. Awh and Pashler (2000) tested an observer's ability to report the identity of briefly-displayed numbers on a 5 x 5 grid. Two target positions, at non-contiguous locations on the 5 x 5 grid, were cued. While participants were able to successfully report both of the simultaneously-cued numbers with high accuracy, performance at reporting numbers in-between the cued locations was poor generally poor. Given this finding, the authors concluded that attention was distributed evenly between the two cued locations, and not to the intermediate space between them, thereby calling into question the notion of a single contiguous region of attentional facilitation. In a similar manipulation, Kramer and Hahn (1994) demonstrated that subjects could accurately determine when a pair of non-adjacent letters matched,
regardless of the presence or absence of potentially interfering stimuli that occurred between the two target locations. Because the distracting letter fell in between the target letters, and did not interfere with performance at the matching task, it was concluded that attention was divided between the two locations, and not distributed to the intermediate space that contained the interfering letters. The general conclusion of these two studies, and others that are similar (e.g., Hahn and Kramer, 1998), is that there is potentially evidence that the attentional spotlight can be divided at will.

1.1 Divided attention in the real world?

A vital test for any theoretical advancement is whether or not it makes sense outside of the laboratory. The notion of a unitary spotlight resonates well with our everyday visual experience. Can the same be said of multiple attentional foci? Consider the following scenario. You are playing a friendly game of ultimate frisbee, and you currently have the disc. You are standing just short of the opponents' goal, and most of the other players (your team and the other team) are in front of you, waiting to be passed the disc, or intercept that pass. You need to make a decision as to whom to pass the disc to. To make this decision, you must wait until you receive a signal from one of your teammates, letting you know to pass him or her the disc. In this situation, you will be visually monitoring a number of your teammates for a signal, for example, a wave. At the same time, you want to inhibit that same signal if it originates from one of your opponents, lest you pass the disc to the other team. In this situation, it could be expected that you will respond quickly to a signal from any of your teammates (the target stimuli) but not at all to a signal from an opponent (the distractor stimuli), even if he or she is standing immediately in between two of your teammates. This requirement, to monitor a number of moving non-contiguous stimuli (your teammates) while ignoring information from others that are interposed amongst them (the opponents), generally mirrors tasks that purport to demonstrate a division of the attentional spotlight, above (note that the ultimate frisbee example was chosen because
of the commonly informal nature of the game, in which distinguishing uniforms would not be worn).

A critic of the divided-attention hypothesis could interject here and point out that this scenario does not require simultaneous non-contiguous distribution of the attentional spotlight. Rather, the observer can continue to make discrete, serial visual shifts from location to location. However, the fact remains that some sort of representation has been formed, and that a wave from any of the team-mate players can trigger a response (i.e., passing the disc), regardless of whether they are the focus of the attentional spotlight or not, and that a signal from a non-team-mate will not trigger a response, even if he or she is immediately in between two team-mates. Thus at one level the unitary spotlight of attention can be maintained, while at another, inputs from multiple non-contiguous stimuli can trigger a response, even though responses from the space between them is inhibited. As such, perhaps studies such as Awe and Pashler (2000) and Kramer and Hahn (1994) are not demonstrating a division of the attentional spotlight, per se, but rather a division of some other, related mechanism.

2. FINST theory: Fingers of INSTantiation (a.k.a. Visual Indexing)

In this section I discuss a theoretical framework used to understand the phenomenon of multiple non-contiguous points of attention, and its relation to the multiple object tracking task. In order to reconcile the single spotlight of attention with the findings that we can track multiple independently- moving stimuli in the visual field, categorize some but not others as targets, and rapidly (without attentional, serial scanning) foveate to the items in question, Pylyshyn and his collaborators (Pylyshyn and Storm, 1988; Pylyshyn, Burkell, Fisher, Sears, Schmidt and Trick, 1994; Sears and Pylyshyn, 2000) have posited a limited capacity object indexing mechanism that operates after the feature detection stage of visual processing, but prior to the operation of the attentional spotlight. The term FINST (Fingers of INSTantiation) was originally used by Pylyshyn (1989), and makes reference
to the similarity with using a finger to point to 'this' or 'that' item.

According to a 2-stage theory of vision, (e.g., Jonides, 1983) visual processing occurs first rapidly and in parallel across the entire visual field, and second, slowly and in serial, moving under conscious control. The differences between these stages of processing have been implicated in a number of common psychological findings, such as the difference between feature and conjunction searches (Treisman and Gelade, 1980), guided search (Wolfe et al., 1989) and subitizing vs. counting (Trick and Pylyshyn, 1993). However, Pylyshyn et al. (1994) point out that if the spotlight can be rapidly and precisely moved to a particular location, at which a moving object currently resides, then there must be a mechanism to instruct attention where to move, as for example if the frisbee player wanted to attend to particular team-mate. One way in which this could be done would be by attending to a stimulus onset, such as a hand wave (i.e., a “bottom-up”, or stimulus-driven process). However, the observer would have to wait for such an onset in order to attend to the desired target. Alternately, the observer could direct the attentional spotlight to scan in a random, serial manner, picking a direction in relation to the current focus. However, this would prove to be a very inefficient strategy, as the observer would be constantly searching in a slow and effortful manner. As Pylyshyn et al. (1994) conclude, the ability to track a moving stimulus or group of stimuli, whose only identifying factor is temporal continuity, requires a mechanism that can individuate those items using something akin to a demonstrative reference in natural language (‘this’, or ‘that’, e.g., Clark 2004; Pylyshyn, 2009), and rapidly direct visual attention to them for further processing.

2.1 FINSTS and Multiple Object Tracking

Visual Indexing (a.k.a. FINST theory) has been used to explain the ability to successfully track multiple moving objects. As described, participants are able to track multiply independent moving objects at a rate which suggests a parallel rather than serial function (e.g., Pylyshyn and Storm, 1988).
To reconcile high performance at this task with the common notion of a unified attentional spotlight, an object indexing system is suggested. In this conception, the indexes (sic) are assigned during the target acquisition stage, and are maintained as the objects move around the computer monitor. Because the object indexes stay with the object or objects as they travel around the screen, the observer has rapid access to the objects’ physical locations, as is demonstrated by the high accuracy rates. It is important to note that while the index can direct focal attention to the object, the index itself is not focal attention. This is supported by studies demonstrating that observers have very limited access to the identities of the tracked objects, (e.g., Pylyshyn 2004). In this study, Pylyshyn assigned unique identities to each of the target objects during the encoding stage (each target circle had, displayed within it, a unique number, 1 through 5). The numbers were removed during the tracking stage, and participants were quizzed regarding the objects’ identities at the end of each trial. While memory for object location was high (80%), memory for object identity was significantly poorer (~30%), which precludes the presence of focal attention at the object level. That is, if full focused attention were present at each object index, reporting the object identity would not be a problem. As such, the notion of a single unified spotlight of attention can be maintained, even though we have access to multiple non-contiguous locations within that focus.

2.2 Describing the visual indexing (FINST) mechanism

Despite fairly consistent findings regarding the number of objects that can be tracked (approximately 4 or 5) and the factors that affect tracking ability (speed, number of objects), there is still some disagreement in the literature regarding the mechanism that allows us to perform such complex tasks as tracking multiple independently moving objects. The notion that there is a limited capacity, yet parallel mechanism that can index a certain number of objects and follow them as they
move through the visual field, is a powerful way of explaining such phenomena as the ability to rapidly foveate to a peripheral object, or number of objects that bear no individuating features. Besides this great explanatory power, however, it is necessary to have some independent validation of this mechanism's existence; otherwise it is a risky venture to add an intermediate stage within a well-established concept such as the 2-stage theory of vision. Scholl (2009) makes precisely this point, stating that if the presence of multiple pre-attentive indexes cannot be empirically differentiated from simply positing the multiple division of the attentional spotlight, then we should go with the simpler, accepted theoretical framework. However, as we have seen, by positing just such a mechanism we can maintain the sanctity of a contiguous visual spotlight (which does make a lot of experiential / phenomenological sense), while explaining the complex visual behaviour described above. Importantly, then, we must look for predictions which would support the existence of an object indexing mechanism.

As pointed out by Pylyshyn et al. (2004), the most obvious prediction that can be made from visual indexing theory is that, once indexed, a small number of items in the visual field will have facilitated access to conscious attention, over and above that experienced by non-indexed entities. This is shown in the multiple object tracking paradigm, (e.g., Pylyshyn and Storm, 1988) simply by the fact that subjects are accurately able to track up to 5 independently moving objects amongst as many, or more, distractors. This demonstrates that positional information regarding the tracked objects is available to the indexing mechanism. The next obvious step is to determine whether, and moreover, how much, information is available to conscious processes once the indexes are assigned.

We know from studies such as Pylyshyn (2004) that the availability of identity information regarding the indexed objects is limited, in that participants perform poorly when asked to link unique identities to the objects being tracked. Likewise, Scholl, Pylyshyn, and Franconeri (1999) have
demonstrated that while participants have access to information regarding the tracked objects' location and direction, information regarding their shape and colour is lacking. This in itself is evidence that full conscious attention cannot be present at each location, otherwise such information would be readily accessible to conscious awareness. Nevertheless, objects that have been indexed (as determined by the fact that they are successfully tracked in a multiple object tracking scenario) have been shown to enjoy some benefits that untracked (and thus unindexed) items do not. Specifically, Sears and Pylyshyn (2000) demonstrated that when the form, or physical identity, of an object changes mid-task, these changes are identified more rapidly for target items than for distractor items.

In this experiment, participants were required to identify the nature of a form change, which could be one of two possibilities (a digital-type figure “8” which could turn into either an “H” or an “E”, also in digital form). Because the change from a digital figure “8” to an “H” or “E” did not involve an attention-grabbing visual onset (also called a “transient”), but rather used a visual offset, which does not automatically attract attention (e.g., Kramer and Hahn, 1995; Yantis, 1993), the task was considered to require the application of focused attention. That is, attention was not allocated automatically based on an external signal, but rather had to be consciously applied to determine whether a shift to an “H” or “E” had taken place. Because the task demanded the application of focused attention, and because indexed items enjoy preferential access to focused attention, changes in tracked items were identified faster than changes in non-tracked items. Given this interpretation, we can predict that attention-demanding tasks will be performed faster for tracked vs. non-tracked objects – a prediction that will have implications when we consider that the tracked objects may be real world entities, and that efficient and accurate processing of their information may have immediate implications for the successful and safe performance of a given task such as driving.
2.3 Tracking set composition

Another controversial issue in the object tracking literature is the importance of the composition of the tracking set, particularly, whether heterogeneity in the tracking set affects object tracking performance. Because tracking tasks have generally excluded individual item detail for a number of reasons (e.g., item detail would simplify the task because participants could just remember the target identity and locate it that way, rather than by tracking), and because tracking can be carried out very well without such detail, it has generally been assumed that the object indexing system does not store individual item detail (e.g., Pylyshyn 2009). This is reflected in FINST theory, in that the indexes are thought to operate based on early-visual signals that provide (and update) the position of a moving object, but are unaware of what it is. This notion is again supported by the finding that while positional information is usually very accurate (80% or higher), memory for individual object identity, shape or colour is generally much poorer (e.g., Pylyshyn 2004; Scholl, Pylyshyn, and Franconeri, 1999).

Nonetheless, in real world tracking events, the presence of unique item information is generally unavoidable. As such, there is growing interest in how the presence of unique item identity affects tracking ability. There is recent evidence (Makovski and Jiang 2009; Oksama and Hyona, 2007) that heterogenous item displays have a positive impact on our ability to track multiple objects. Makovski and Jiang (2009) used a coloured-disc version of the multiple object tracking task with 4 conditions, in which they varied the composition of the tracking stimuli. The stimuli maintained their colour during the tracking stage, and reverted to black during the response stage of the task. The 5 colour conditions were: homogenous, heterogenous, “4-unique”, “paired 4” and “paired 2”.

The authors found that performance was better when the tracking set was heterogenous (all different) as compared to homogenous (all the same), and particularly good in the “4-unique”
condition, in which the targets were always 4 specific colours, and the distractors 4 other colours. Performance in the “paired” conditions, where targets and distractors shared colours, was generally poorer, and comparable with the homogenous condition. In order to counter the possibility that participants were able to simply use a verbal rehearsal technique to remember the item colours, the authors included an articulatory suppression condition in which the participants repeated “a three-letter word” (actual word not provided) as rapidly as possible, thus preventing the articulatory repetition of the target colours. Interestingly, this manipulation produced no effect, indicating that participants were not using verbal encoding to perform the tracking task, or that this strategy had no effect. The authors concluded that individual item identity can be used to improve tracking ability. This advantage is attributed to the operation of visual working memory for the tracked items, in parallel with the positional information provided by the object indexing system. This notion is supported by Experiment 4 of the Makovski and Jiang study. In this experiment, participants were given a centrally-presented ‘one-back’ colour memory task in which they were required to report whether a change occurred between presentations of 1, 2 or 4 coloured dots. This was performed concurrently with the tracking task. The fact that the colour memory task negated the advantage previously demonstrated for tracking uniquely coloured items, was taken as evidence that visual working memory played a role in producing that advantage. Further interest in how visual working memory for surface features and the object indexes interact is indicated, with specific mention of the notion that parallel systems for tracking and object identity may exist.

3. Object Tracking in Realistic Environments

So far I have discussed the possibility of dividing the attentional spotlight (e.g., Kramer and Hahn, 1995), as well as a theoretical framework which supports the findings that lead to this division,
but maintains the possibility of a unified attentional field. Additionally, I have discussed one major
difference between object tracking in realistic vs. standard laboratory conditions, which is object
heterogeneity. The following section investigates further the differences between tracking in a realistic
vs. standard laboratory environment.

Since Pylyshyn and Storm's (1988) introduction of the multiple object tracking task, it has been
widely applied in investigations into the nature of visual attention. It is commonly stated that the task
is useful because it taps resources that are used in every-day tasks. Assertions that the multiple object
tracking task can be directly generalized to real-life situations include, among others, Horowitz et al.,
2007: air traffic control, driving and sports; Fencsik et al, 2007: driving, air traffic control, children on
a playground; Boot et al., 2007: air traffic control; Fehd and Seiffert, 2008: walking on sidewalk, team
sports; Oksama and Hyona, 2004: air traffic control, sports, driving; Feria, 2008: driving, playing
sports; Tombu et al, 2008: tracking wild gazelles. Interestingly, most of these examples are never
actually demonstrated or otherwise verified in the literature. Because the goal of the current research
is to better understand the nature of single vs. multiple foci of attention, and because attention is by
necessity based in the real world that exists around us, I will be attempting to model the multiple
object tracking task in one of the places which it is commonly assumed to apply: the driving
environment.

The present experiments will not only make a contribution to the research on multiple object
tracking (and demonstrate its role in a practical task), but will make a contribution to the driving
literature insofar as they relate to other aspects of driving performance; specifically, how the
requirement to track the positions of multiple independently-moving non-contiguous objects (vehicles)
affects standard metrics of driving. It may be that tracking significantly influences steering, driving
speed maintenance and braking. The way in which tracking influences these metrics is important as it
is argued that driving a vehicle requires two different attentional systems. Wickens (2002) describes how visual processing can be divided into different streams of processing. According to this model there is a *focal* stream which deals with foveal requirements such as reading, form identification, and obstacle (hazard) detection, as well as an *ambient* stream that is related to peripheral visual information such as balance, position, and “ego motion” (Wickens, 2002). With these details in mind, the following section will describe some important differences in both the control requirements and the visual display when tracking objects in a realistic environment.

3.1 Differences in control characteristics between laboratory multiple object tracking and real-world equivalent

An important factor that must be considered when comparing traditional laboratory object-tracking studies to object-tracking in the real world, is the way in which attention needs to distributed across a series of sub-tasks in order to perform a given activity. Michon (1985) breaks down the task of driving an automobile into 3 discrete levels: the Strategic, or Planning level, which involves such issues as route choices, trip goals, and evaluations of costs and risks; the Tactical, or Maneuvering level, which involves dealing with situation-relevant action decisions (e.g., lane change, passing, emergency stop), and the Operational or Control level, which involves the millisecond time-frame operations of maintaining the desired heading and acceleration values. It is thus understood that driving involves a multi-level hierarchy of attention, and that even in the most simplistic scenario drivers are required to perform, at the very minimum, control level actions such as monitoring their environment, steering the vehicle, and applying braking and acceleration as necessary. As such, it is evident that driving a motor vehicle is a complex task, one that involves co-ordinating many activities, even when the task involves nothing more than maintaining a straight course at an even speed.
In comparison, the original multiple object tracking paradigm requires no secondary task, allowing participants to devote their undivided attention to tracking. Breaking down the multiple object tracking task into its component parts, one finds that it ultimately involves three discrete, sequential activities. First, items are cued, during which the spatial location for each item to be tracked is encoded, or indexed. Second, there is period of object tracking, which involves visually keeping track of the cued items as they move around the display, while maintaining a fixed central stare (Pylyshyn and Storm, 1988) or active, unconstrained viewing (Doran et al., 2009; Zelinsky and Neider, 2008). Third, there is a report stage, where the items stop moving, and the participant is required to indicate which objects are targets (in the full-report methodology) or whether a re-cued item is a target or a distractor (in the partial report methodology). As such, on its own, multiple object tracking task is a 3-part serial task.

Studies investigating dual task manipulations in multiple object tracking indicate that the inclusion of a secondary task is detrimental to performance on both the tracking task and the secondary task. Tombu and Seiffert (2008) tested participants on a tone discrimination task concurrently with a multiple object tracking task. They found that increases in object tracking difficulty (e.g., increased object speed, proximity) interfered with speed and accuracy of tone discrimination, and that object tracking performance was impaired by 15.2% when tracking 4 targets out of a total of 8 objects, using a full-report methodology. It is likely that this estimate of interference is low, given that tone discrimination is an auditory task, while object tracking is a visual / spatial task. Multiple Resource Theory (Wickens, 1984; 2002) describes how tasks which use different processing resources, such as auditory and visual, interfere less with each-other than if they use the same resource (e.g., visual/visual). Because a tone detection task would not be expected to interfere extensively with a spatial task such as object tracking, it is therefore likely that the requirements of maintaining heading
and speed (‘ambient' tasks: Wickens 2002), which require spatial processing, may cause an even greater decrement to object tracking performance, which is also a spatial task. This being the case, any attempt to describe object tracking in a realistic environment will have to account for the way in which processing resources are spread across all task components.

3.2 Differences in visual characteristics between laboratory multiple object tracking and real-world equivalent

Although it is commonly assumed in the literature that driving is exactly the type of task where multiple object tracking is required, there are substantial differences between tracking in the standard laboratory task, and tracking in the real world. The classical multiple object tracking task involves participants tracking 2-dimensional objects (e.g., circles, in the original task) moving randomly across a blank 2-dimensional background; this background rarely subtends more than 25º of visual angle. (Pylyshyn and Storm's 1988 version used a 21.5º x 21.5º background.) There are a fixed number of distractors in any given trial (i.e., distractors do not enter or leave the scenario). The standard method of cueing the targets via flashing is very robust, and leaves no ambiguity during the target acquisition stage. All of the items on screen are identical to one another, and the participant is required to track the targets for a discrete, rather short period of time (e.g. 8 to 10 seconds), though tracking durations as long as 10 minutes have been reported (Wolfe et al., 2007).

In comparison, in the real world one is faced with an extremely detailed visual environment. Considering the example of the driving environment, individual objects exist in 3 dimensions; their trajectories can vary independently of one another, or vary together in a complex manner. Both target vehicles (for example an encroaching vehicle that needs to be avoided, such as an emergency response vehicle) and distractor vehicles may enter and leave the scenario at random. Objects are usually
visually distinct (unique), and target acquisition may be ambiguous (i.e., there are no direct target cues, such as flashing off and on in the original object tracking paradigm; rather targets are chosen based on their situational relevance). There is also no obvious constraint on the visual angle of items to be attended to, and tracking duration is idiosyncratic and situation-dependent.

3.2.1 Composition of the visual display, presence of motion, and 3d environments

Realistic environments are invariably more complicated than typical laboratory environments. As we know from seminal works in the field of cognitive psychology, searching for targets within an array of distractors is more difficult when the targets and distractors share common elements (e.g., Treisman and Gelade's feature vs. conjunction search, 1980; Wolfe et al.'s guided search, 1989). In a realistic tracking environment such as the ultimate frisbee example discussed previously, targets and distractors will share many if not all of their characteristics. As such, visual search, an important skill in tracking, for example when a target is lost, becomes more difficult as the relatedness of targets and distractors increases. Additionally, increased heterogeneity of the distractor set has been shown to increase the difficulty of the visual search task (Humphreys et al., 1989), again indicating that the complexities inherent in realistic stimuli may lead to increased task difficulty.

Another factor that differentiates real-world scenes from simple laboratory stimuli is the presence of motion. In most real-world situations where simultaneous attention to multiple targets may be necessary, it is likely that in addition to the movement of the targets and distractors, there will be some degree of perceived motion that is associated with the background. Background motion has generally been shown to have an influence of how objects are perceived. For example Li et al. (2008) demonstrated that background motion (expanding and contracting dots) interfered with participants' ability to reliably indicate the direction of offset (left or right) for a superimposed ellipsoid array of
dots. Additionally, Beardsley and Vaina (2008) have demonstrated that when a field of dots (the background) had a left- or right-angled movement, it interfered with the ability of participants to accurately judge the trajectory of a smaller field of dots. Given these findings, it is evident that the presence of extraneous background motion could be expected to impact performance in a multiple object tracking task, in which trajectory information is very relevant.

Whereas multiple object tracking is typically tested using a 2-dimensional representation on a computer monitor, realistic stimuli can take on 3-dimensional form. Regarding visual depth and multiple object tracking, Viswanathan and Mingolla (2002, Experiment 2) used a classical multiple object tracking task performed across two planes of depth using stereo-vision goggles. The authors found that tracking performance was poorer when the targets and distractors were spread equally across multiple depth planes, indicating that the inclusion of depth in a tracking task can lead to poorer performance. Zelinsky and Neider (2008) tested participants in a modified multiple object tracking task, where the objects to be tracked were computer generated sharks, and the search environment was a virtual 3d fish-tank. Accuracy was high for 1-3 sharks (92%), but declined to 78% for 4 sharks (compared with ~90% for standard multiple object tracking with 4 targets), indicating that it may be harder to track objects in 3 dimensional space.

There is also evidence that some types of 3d motion take precedence over others. Imura et al. (2008) demonstrated that a when visual search target, defined by its motion relative to the distractor items (either receding amongst approaching items, or vice-versa) is seen to be approaching the viewer, response times are faster than when it is receding away from the viewer. Additionally with regards to tracking and motion, a recent study by Thomas and Seiffert (2010) demonstrates a direct effect of ego-motion (self-motion) on the ability to track multiple moving objects. In this study, participants tracked multiple coloured dots in four possible conditions. Using virtual reality goggles, participants either
walked in place beside the tracked objects, or sat passively in a wheelchair beside the tracked objects. It was thereby demonstrated that at the 3-target level, tracking was significantly better in the “passive stay” (sitting still in the wheelchair) and “active stay” (walking in place next to the tracked objects) conditions, as compared with either of the “moving” conditions. Interestingly, there were no differences at the 1-target level. The authors discuss a spatial-updating hypothesis, which rests upon the notion that tracking multiple objects, and updating information regarding one’s physical location in space may have overlapping cognitive components, thus introducing interference between the task elements. Applying these findings to the context of realistic object tracking, it is very likely that there will be important differences between how objects are tracked in the typical laboratory test, and how objects would be tracked in a realistic task such as driving an automobile.

Finally, it is important to note that although computer simulations may appear 3-dimensional, this is an illusion based on perspective in the 2-d view monitor. The extent to which findings can be generalized to real-world experience is dependent upon how much of the 3-d experience can be simulated. Modern simulators such as the DS-600c used in the current research employ multiple screens at different angles, and seat the driver within an actual vehicle, thus increasing the likelihood of accurately replicating a realistic 3d experience.

4. **Overall summary**

The current set of studies is an attempt to learn more about how we distribute our attentional resources among tasks in a realistic driving environment, and how this affects tracking, our knowledge of the tracked objects, and performance on standard driving indexes such as steering and headway maintenance, and hazard response time. In Experiment 1, I test how participants are able to perform
the multiple object tracking task in a simulated driving environment. This task is modeled as close to the original task (Pylyshyn and Storm, 1988) as the simulator will allow, while measuring driving performance in addition to tracking performance. In the second experiment I investigate participants’ ability to perform a change detection task for tracked vs. non-tracked items, (2a and 2b), and test their ability to respond to specific braking events for tracked vs. non tracked vehicles (2c). Experiment 3 investigates the effect of homogenous vs. heterogenous tracking environments on tracking ability in a realistic driving environment, and further, on standard indexes of driving ability, including lane keeping, standard deviation of headway distance, average headway distance, and hazard reaction times.

5. Experiment 1: Multiple object tracking in a simulated driving environment: multiple vehicle tracking

The purpose of this experiment was to model the multiple object tracking task in an applied, quasi-realistic environment. In particular, I was interested in demonstrating the accuracy with which objects (in this case vehicles) can be tracked in a typical highway driving scenario, reproduced in the DS 600c driving simulator. Based on the increased complexities of the realistic tracking environment, it is conceivable that accuracy would be somewhat lower than is typically found in multiple object tracking studies, although it has been previously demonstrated that multiple object tracking performance is robust to such additional complexities as tracking across occlusions (e.g., Horowitz et al., 2006). Additionally, I tested performance when the driver was required to operate the vehicle (average headway distance, standard deviation of headway distance, and standard deviation of steering), compared with when these operations are performed by the driving simulator. If the object indexes are maintained pre-attentively, as suggested by Pylyshyn (1989, 2009), then the addition of control requirements should have a minimal impact on performance, because pre-attentive processing
of information is assumed to be automatic (i.e., to require minimal or no attentional resources, e.g., Logan, 1992; Shiffrin and Schneider, 1977). However if attention is required to maintain the target indexes, as proposed by Scholl (2009), then the control manipulation should significantly reduce accuracy at the tracking task.

According to Wickens (2002) attention can be divided into either focal or ambient channels, in the sense that the division supports efficient time sharing, and that the channels use different neural structures, and are associated with qualitatively different types of information processing. Focal attention involves detailed tasks such as reading or pattern recognition and hazard detection, and is generally thought of as a foveal activity (i.e., one that takes place within the centre of the visual focus, or the fovea), though focused attention has been demonstrated outside the fovea, for example in Posner et al.’s (1980) cuing paradigm. Ambient attention, on the other hand, involves broad, unfocused background motion, as occurs for example when one is walking or driving in a particular direction. As such, ambient attention is considered to be important for sensing one’s physical orientation and motion. Given these definitions, I hypothesize that if tracking is a focal task, there should be minimal interference between tracking, and steering and headway control (which are thought to use the ambient attentional channel). Otherwise if tracking is an ambient task, it may interfere significantly with steering and headway control, and vice-versa.

5.1 Methods

Participants: 28 undergraduate students (6 males and 22 females) from the University of Guelph Dept. of Psychology participant pool participated in this study. The participants were students taking introductory psychology. Participants had normal or corrected to normal vision, and received 1

---

1 Due to a missing demographic question, participant ages are not available for Experiment 1, however were in the expected range for undergraduate students, 18-25.
course credit for their participation. All participants were screened for medical issues that could be dangerous in a driving simulator (epilepsy, heart conditions).

Apparatus and Stimuli: Tracking tasks took place within a Drivesafety DS 600C fixed-base driving simulator. The simulator body is a modified Saturn sedan, and the projection system provides a 300-degree wrap-around display consisting of 5 screens in the front and 1 in the back (250 and 50 degrees, respectively). The simulator display operates at 256 colours, with a refresh rate of 60 hz. A standard 3-lane freeway-style roadway simulation, with no turns, was modified such that 9 identical blue cars appear in front of the participant vehicle in a 3 x 3 grid, as shown in Figure 2. In the object tracking only condition, vehicles in the row closest to the driver (row 1) subtended an average visual angle of 2.97º, ranging from 2.01º at the most distant position, to 3.44º at the closest position. The vehicles in row 2 and 3 subtended an average of 2.1º degrees (range 1.9º to 2.3º) and 1.86º (range 1.72º to 2º), respectively. In the multiple task condition the precise visual angle of the object vehicles was similar to that in the object tracking only condition, but depended on participants’ performance on the headway maintenance task. During the lane change maneuver target and distractor vehicles moved across the screen laterally at a rate of approximately 3.4º per second, averaged across all 3 rows.

Figure 2: Experiment 1 vehicle tracking stimuli from ‘birds eye’ (left) and isometric (right) view angles
Design: Experiment 1 used a 3 x 2 repeated measures design. The first independent variable was the number of targets (1, 3 or 4). The maximum number of targets was 4 rather than 5, because there were a total of 9 rather than 10 objects in the display, and participants could otherwise have used a strategy whereby they track the non-targets to achieve higher accuracy scores. The second independent variable was task load, of which there were 2 levels: object tracking only (no steering or headway control required), and multiple-task (steering and headway control required, in addition to object tracking). A baseline condition (steering and headway control required, with no object tracking task) was also included in order to determine the characteristics of driving the simulator with no concurrent tracking task. The experiment consisted of 2 blocks of 30 trials each, where the order of blocks was counterbalanced; the baseline condition was run as a between-subjects factor, using a separate group of participants. Accuracy at the tracking task was measured using a full-report methodology, where the participant attempted to locate all of the cued targets. The main dependent variable was the percentage of target vehicles that were correctly identified (e.g., 3 / 4 targets = 75% accuracy). Additionally, the dependent variables average headway distance (average distance between the subject vehicle and the vehicles ahead of it), standard deviation of headway distance (i.e., standard deviation of the distance between the subject vehicle and the vehicles in front), and standard deviation of steering (i.e., weaving away from a straight-line path), were collected during each trial, and were measured in metres (m).

Procedure: Upon arrival, participants were given the information/consent forms, as well as a Simulator Sickness Questionnaire (SSQ) in order to determine their susceptibility to simulator sickness (i.e., motion sickness due to simulator usage). Participants with a high likelihood of becoming ill during the simulation did not take part in the study, but received full credit nonetheless. Once the participant was seated, the experiment began with the simulator in a stopped position, with the 9 object
vehicles directly ahead. The subject vehicle was placed between the left and middle highway lanes, directly on top of the dotted line, in order to maximize visibility for all 9 object vehicles. The participant was initially given a 5 minute training period (consisting of 5 one-minute trials) in which to become familiar with the simulator and task. In the object-tracking-only condition the subject vehicle, as well as the 9 object vehicles, accelerated automatically to 60 km/h. In the multi-task condition the 9 object vehicles accelerated automatically to 60 km/h, and the participant was required to maintain headway and lane position behind the object vehicles. During each trial the target vehicles (1, 3 or 4) flashed off and on five times, indicating which vehicles were to be tracked. At this point the vehicles began changing lane-position laterally on the roadway at a rate of 1 change every 2 seconds. Vehicles were allowed to change positions without maneuvering to avoid one-another. The lane-changing continued for 30 seconds, during which the vehicles progress through a random ordering of all possible lane positions. After the lane changes were complete, the simulation was paused, and the participant indicated which of the stimuli were the target vehicles using a ledger designed for this purpose, which is provided in Appendix A. When the trial was finished, the next trial commenced via a button press, and the experiment continued until completion. Each trial took approximately 55 seconds, including 5 seconds to cue the target vehicles, 30 seconds of tracking, ~15 seconds to report, and 5 seconds to reset the vehicle positions for the next trial.

5.2 Results

Tracking Accuracy: A 2 x 3 (task load x number of targets) ANOVA was performed on the object tracking accuracy data. Tracking accuracy score sheets for 5 participants were misplaced in the lab, and therefore the accuracy analysis is based on the data from 23, rather than 28 participants. The main effect of number of targets was significant at F(2, 44)=28.676, p<.000, partial $\eta^2=.566$, where
tracking accuracy was significantly lower when there were more targets to track. The main effect of task load was also significant at F(1, 22)=37.543, p<.000, partial η²=.631, where tracking accuracy was significantly lower when the subject was required to steer and maintain the simulator vehicle’s speed. Finally there was a task load x number of targets interaction, F(2, 44)=3.781, p<.05, partial η²=.147, indicating that when subjects were required to steer and maintain the vehicle’s speed, accuracy at tracking 3 or 4 vehicles was lower than when the subject was not required to steer and maintain headway.

Simple main effects for the accuracy data were analyzed using multiple pairwise comparisons with a Bonferroni correction. Accuracy when tracking and driving (dual task) vs. tracking without driving (single task) was compared for the 1, 3 and 4 target conditions. When there was 1 target to track, accuracy in the dual-task and single task conditions was equivalent, p>.05. However, in both the 3 and 4 target conditions, accuracy was significantly poorer in the dual task condition, as compared with the single task condition, p<.05 in both cases. The effect of Number of Targets in the dual task vs. single task condition was also compared. In the single task condition, the effect of Number of Targets was non-significant, p>.05, where accuracy scores were equivalent at 1, 3 and 4 targets. In comparison, for the dual task condition, accuracy was significantly lower, p<.05, when there were 3 or 4 targets to track.

In order to get a better idea of approximately how many vehicles the participants were successfully tracking, an estimate of tracking accuracy if the participant was randomly guessing 1 or 2 of the locations was calculated, using sampling without replacement. For example, if there were 4 target objects out of 9 objects total, and the participant was tracking 3 of the 4 objects (i.e., guessing 1) the expected accuracy would be: 3 (the number tracked) + 1/6 (the probability of guessing one out of the remaining 6 items) divided by 4 (total items to be tracked) for an expected outcome of 79.2%.
Predicting tracking accuracy when guessing 2 out of 4 objects is slightly more complex, and involves taking the number tracked (2), plus the probability of guessing the remaining 2 objects correctly $2(2/7 * 1/6)$, added to the probability of guessing 1 of the items correctly $1(2/7 * 5/6 + 5/7 * 2/6)$, divided by 4 (total items to be tracked), for an estimated accuracy of 64.3% if guessing 2 of the 4 targets. The formulae are available in Figure 3 below. Obtained accuracy scores were compared with the predicted guessing scores using a one-sample t-test for each close comparison. Details are available in Table 1.

$$\frac{3 + (1/6)}{4} = 79.2\% \quad \quad \quad \frac{2 + 2(2/7 * 1/6) + 1(2/7 * 5/6 + 5/7 * 2/6)}{4} = 64.3\%$$

Figure 3: Calculation of predicted accuracy if guessing 1 (left) or 2 (right) out of 4 targets

Figure 4: Accuracy Data (Number of Targets X Task Load). Scores represent the percentage of correctly identified targets. Error bars represent standard error of the mean
Table 1: Obtained accuracy was compared with predicted guessing using a one-sample t-test for each close comparison (i.e., 85% vs. 11% would not be tested). Statistically equivalent scores, p<.05, are shaded the same on the horizontal line, and outlined.

**Driving Metrics:** Average headway distance, standard deviation of headway distance, and standard deviation of steering were analyzed using a separate ANOVA. An outlier analysis was performed for each metric, removing all trials that were >2.5 standard deviations away from the mean. This resulted in the removal of 1.2% of the data for average headway distance, 2.3% of the data for standard deviation of headway, and 4.5% of the data for standard deviation of steering. The difference in outlier removal across conditions was non-significant (p>.05). There were no significant differences in the average headway that drivers kept, p>.05, indicating that drivers generally drove at an equal distance regardless of how many targets there were; no difference in average headway was predicted. The standard deviation of headway distance increased significantly with increases in the number of tracking targets, $F(2, 54)=3.84$, p<.05, partial $\eta^2=.124$, indicating that tracking higher numbers of target vehicles can impact a driver’s ability to maintain a steady headway distance. Standard deviation of steering also differed significantly depending on how many objects were being tracked, $F(2, 54)=4.85$, p<.05, partial $\eta^2=.152$. This indicated that the drivers’ ability to consistently maintain a straight course was reduced as the number of items to be tracked increased. The finding that both SD of headway and SD of steering were affected by increases in the number of targets to track, indicates that there may be a substantial ambient component to tracking, as these are both considered to be
ambient-type tasks (Wickens, 2002).

Investigating the simple main effects for the driving metrics, using multiple pairwise comparisons with a Bonferroni correction, it was evident that average headway distance was equivalent for 1, 3 and 4 targets, p>.05. The standard deviation of headway distance was equivalent for 1 and 3 targets, which both were significantly lower than in the 4 target condition, p<.05. For the standard deviation of steering, again employing the Bonferroni correction for multiple comparisons, there were no significant differences. Considering the fact that the original ANOVA for this analysis was significant, p<.05, this non-significant outcome is likely a result of the more stringent criteria for multiple comparisons.

For each baseline condition an independent-samples t-test was performed, comparing the average baseline performance with the next closest score (e.g., baseline vs. track-3 for SD of steering). Average headway distance was significantly longer in the baseline condition, t(54) = -4.196, p<.05, indicating that participants naturally followed at a more distant position when there were no tracking requirements. The standard deviation of headway distance was lower in the baseline condition, t(54) = 2.384, p < .05, indicating that participants were able to maintain a steadier headway when they were not required to track any objects. Similarly, the standard deviation of steering was smaller in the baseline condition, t(54) = 6.670, p < .05, indicating that participants were able to maintain a steadier course (less weaving) when there were no tracking requirements.
Figure 5: Experiment 1 Driving Metrics - Graphs for Average Headway (left), SD of Headway (centre), and SD of Steering (right).

5.3 Discussion

Experiment 1 provides a number of interesting findings. First, it is apparent that, in the simplest condition where participants were not required to operate the vehicle, tracking performance was very good, with an overall accuracy of 80%. It is therefore obvious that participants were quite capable of keeping track of the physical location of a number of objects in the driving environment.

This number, however, was reduced significantly when the participants were required to operate the vehicle in addition to performing the tracking task, to approximately 65% when tracking 4 targets. This means that some qualifications need to be made before assuming that drivers can simultaneously track 4 or 5 objects in the driving environment. With regards to the driving metrics, it was also demonstrated that drivers performed more poorly when required to track higher numbers of vehicles. Specifically, drivers were less consistent at steering the vehicle and maintaining a specific headway when the number of items to be tracked was higher, though average headway distances did not differ.

Given that Wickens (2002) defines steering and headway maintenance as ambient attentional tasks, and that tracking was shown here to interfere with steering and headway maintenance, it can be reasoned that tracking may also be an ambient-attentional task, or at least contain a substantial ambient component. It is also interesting to note that the standard deviation of headway distance and the
standard deviation of steering were significantly poorer at all three levels of tracking (1, 3, 4), as compared to a baseline drive where no tracking took place. While this is consistent with the notion that increased workloads lead to decreased performance, it is also another indication that object tracking may be a more complex, composite task than previously thought.

It is also interesting to note that while the standard deviation of headway distance and the standard deviation of steering were affected by increases in the number of targets to be tracked, average headway distance was not. This difference can be interpreted using Michon’s (1985) three-part hierarchy of driving requirements. These include the Strategic level, which involves such issues as route choices, trip goals, and evaluations of costs and risks; the Tactical level, which involves dealing with situation-relevant action decisions (e.g., lane change, passing, emergency stop), and the Operational level, which involves the millisecond time-frame operations of maintaining the desired heading and acceleration values. When a driver decides upon his or her following distance, it may be considered a Tactical decision, in that it involves a situation-relevant factor (for example, nearness to the tracking objects). However when the standard deviation of steering and the standard deviation of headway distance increase, this is indication that the driver is having trouble at the Operational level – in this case, due to the added requirement of tracking the vehicles while driving. This is what appears to be happening in the present experiment, where both standard deviation of steering and standard deviation of headway distance were affected by the increase in tracking set size, but average headway distance was not.

6. Experiment 2a: Change detection for target vs. distractor vehicles in a multiple vehicle tracking task.

It has previously been demonstrated that form changes for target items in a multiple object tracking task are identified more rapidly than form changes for non-target items (Sears and Pylyshyn,
In Experiment 2a I was interested in following up this finding using a *change detection* task to determine the presence or absence of the attentional focus, and additionally, investigating how this effect transfers to a realistic tracking environment (driving), including the complex visual and control characteristics that are intrinsic to this environment. First, I introduce some more detail on the paradigm and my modifications to it.

**Attention, Onsets, and Change Detection**

Sears and Pylyshyn (2000) asked participants to determine whether one of a number of digital-type “8” characters, in a standard multiple object tracking task, changed into an “H” or an “E” when the appropriate lines were removed. The change was identified faster when it occurred at a tracking target, as compared with a distractor. The authors considered this to be an attention-demanding task, because a visual offset rather than a visual onset was used. Previous studies (e.g., Kramer and Hahn, 1995) have demonstrated that unlike visual onsets (which attract attention due to the appearance of a “local transient”), visual offsets, which involve removing part of a stimulus to reveal the character of interest, do not attract attention to themselves in a visual display. Given this control, and the fact that a performance advantage was still found for tracking targets as compared with distractors, the authors concluded that attention was facilitated for tracking targets, thus supporting the indexing hypothesis.

In the current experiment I wished to provide further support for the notion that attentional processing is facilitated for indexed objects, using a different paradigm, as well as to investigate how this effect is manifest in a realistic environment. I employed an adaptation of the *change detection* paradigm (e.g., Rensink et al., 2002, 1997), in which a pre- and a post-scene are compared and the participant is required to indicate whether or not a change occurred, and (in some versions) what that change was. Whereas a singular change in the visual scene will attract attention in the same manner as the onsets described above, it has been demonstrated that this effect can be negated by including a
global transient, in which the entire visual display goes blank for a brief interval (e.g. 100 msec) at the same time as the change occurs (Velichkovsky et al., 2002). This disrupts the attention-grabbing local transient that would otherwise occur. As such, attention must then be present at the object level before the change can be detected (e.g., Rensink et al., 1997). This paradigm therefore provides us another opportunity to investigate how attention is applied in a multiple object tracking scenario. Specifically, if changes are detected more accurately for tracked vs. non-tracked vehicles, this is further indication that the process which enables multiple object tracking involves the facilitation of attentional focus to the tracked objects.

Because the change detection task occurred at the end of the trial, there was no opportunity to determine whether it had an effect on driving metrics (average headway distance, SD of headway and SD of steering). As in Experiment 1, if tracking requires ambient rather than focal attention (Wickens 2002), both tracking accuracy and the operational-level driving metrics (SD of headway, SD of steering) should be poorer when there are more targets to track (i.e., when there are 3 targets, compared to when there is only 1). No effect of the number of targets on average headway (tactical level) was predicted. It was also expected that object tracking performance would be poorer when participants were required to operate the vehicle (i.e., steering the vehicle, and maintaining headway distance), compared to when no vehicle control was required. As for the change detection task, if attention is facilitated for indexed vehicles (that is, vehicles that were successfully tracked), then change detection performance for tracked (target) vehicles should be better than change detection performance for non-tracked (distractor) vehicles. This would support the notion that tracked (and therefore, indexed) vehicles benefit from facilitated access to the attentional system.
6.1 Methods

Participants: 20 undergraduate students (5 males and 15 females) between the ages of 18 – 24 from the University of Guelph Dept. of Psychology participant pool participated in this study. Average age was 18.9 years (SD = 1.7). Average driving experience was 2.84 years (SD = 1.46). Before arriving at the study, participants were asked to confirm that they had normal or corrected to normal vision, normal colour vision, and that they did not have any medical issues that could be dangerous in a driving simulator (epilepsy, heart conditions). Participants received 1 course credit for taking part in the study.

Apparatus and Stimuli: The apparatus was the same as for Experiment 1. The stimuli were six vehicles of the same general size and shape (mid-size sedan), arrayed on a standard 3 lane highway in a natural-seeming arrangement. The vehicles were composed of a heterogeneous selection of colours (see Table 2) from the simulator, representing a typical assortment of colours found on the roadway. During the vehicle change, one of the 6 vehicles would be replaced by another vehicle that did not duplicate any of the existing cars. There were 3 possible change vehicles in total, each of which differed in colour from the vehicles in the original tracking set. Vehicles in the row closest to the driver (row 1) subtended an average visual angle of 2.97º, ranging from 2.01º at the most distant position, to 3.44º at the closest position. The vehicles in row 2 and 3 subtended an average of 2.1º degrees (range 1.9º to 2.3º) and 1.86º (range 1.72º to 2º), respectively.

<table>
<thead>
<tr>
<th>Tracking Set</th>
<th>Possible Change Vehicles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grand Prix Green</td>
<td>Lexus Grey</td>
</tr>
<tr>
<td>Celica Purple</td>
<td>Grand Prix Tan</td>
</tr>
<tr>
<td>Grand Prix Red</td>
<td>Grand Prix (dark) Blue</td>
</tr>
<tr>
<td>Grand Prix White</td>
<td></td>
</tr>
<tr>
<td>Lexus (light) Blue</td>
<td></td>
</tr>
<tr>
<td>VW Yellow</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Tracking set for Exp. 2a and possible change-vehicles.
Design and Procedure: Experiment 2a used a 2 x 3 x 2 factorial design. Participants completed 60 trials in total, counterbalanced across all conditions. The first factor was Number of Targets (1, 3); the second factor was Change Type (target, distractor, none), and the third factor was Task Load (steering and headway control required, not required). The procedure was similar to that of Sears and Pylyshyn (2000), using vehicles in a simulated driving environment as in Experiment 1. Participants received 5 minutes of practice, after which they completed the 60 experimental trials. Each trial began by cueing the target vehicles, after which the vehicles commenced changing lanes in a random order. At the end of each trial a vehicle change occurred at one of the vehicle locations (see Figure 7). At this point participants were required to respond by first entering the locations of the tracked items (1 or 3) using the data entry form as in Experiment 1 (Appendix A). Immediately thereafter, they were asked to indicate the location of the change item by circling the relevant vehicle location. Participants were aware that a portion of the trials had no change. The ratio of target-, distractor-, and no-change trials was 40:40:20. In order to ensure that local transients did not attract attention, a global transient (blank screen) occurred for 60 milliseconds, during which the change took place. In order to ensure that participants were unable to predict the onset of the change, trials lasted 16, 18, or 20 seconds, after which the change detection task commenced; data collection however was
only performed for the first 15 seconds of tracking, to ensure equal difficulty across timings. The study was 60 minutes in total duration. The dependent variables were accuracy at change detection (i.e., the percentage of trials on which participants were able to correctly identify the exact location of the changed vehicle), as well as tracking accuracy (i.e., the percentage of tracking targets that were correctly located). Additionally, average headway distance, standard deviation of headway, and standard deviation of steering were measured.

6.2 Results

Although there were catch trials in which there was no change, the main analysis in this section is a 2 x 2 x 2 ANOVA. The factors were: Number of Targets: (1,3); Trial Type (target change, distractor change), and Task Load (driving, not driving). To ensure that change detection tasks were accurately identified as either a target or a distractor change, only trials in which all tracked items are correctly identified were used in the analysis.

*Tracking Accuracy:* Accuracy at the tracking task was uniformly high (between 90 and 95 percent) in all conditions, indicating that participants were able to track the target vehicles with no...
difficulty. Tracking accuracy did not differ significantly across any of the factors (p>.05). Estimates of the accuracy with which participants would perform if they were to randomly guess one or two of the target items were calculated in the same way as for Experiment 1, and are included on Figure 8, as well as in Table 3 below.

![Tracking Accuracy Graph](image)

**Figure 8**: Experiment 2a tracking accuracy (Number of Targets X Task Load) for Experiment 2a. Scores represent the percentage of correctly identified targets. Error bars represent standard error of the mean.

<table>
<thead>
<tr>
<th></th>
<th>Targets</th>
<th>Distractors</th>
<th>Guess 1</th>
<th>Guess 2</th>
<th>Est. # Tracked</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tracking Only</td>
<td>1 targ</td>
<td>95.4%</td>
<td>95.0%</td>
<td>16.7%</td>
<td>n/a</td>
</tr>
<tr>
<td></td>
<td>3 targs</td>
<td>93.1%</td>
<td>92.1%</td>
<td>75%</td>
<td>60%</td>
</tr>
<tr>
<td>Tracking + Driving</td>
<td>1 targ</td>
<td>92.5%</td>
<td>92.5%</td>
<td>16.7%</td>
<td>n/a</td>
</tr>
<tr>
<td></td>
<td>3 targs</td>
<td>92.8%</td>
<td>93.6%</td>
<td>75%</td>
<td>60%</td>
</tr>
</tbody>
</table>

Table 3: Obtained accuracy was compared with predicted guessing using a one-sample t-test for each comparison. Statistically equivalent scores, p<.05, are shaded the same on the horizontal line, and outlined. In this case, all obtained accuracies were significantly higher than what would be expected if participants were guessing 1 or 2 of the items.

**Change Detection Accuracy**: Change detection accuracy was measured in two ways. First, because 20% of the trials had no change, it was possible to measure the percentage of false positives (i.e., indicating a change when no change occurred). The overall proportion of false positives was
10.4%, and did not differ significantly across Number of Targets or Task Load factors, p>.05. See Appendix G for a graph of the means for d-prime scores in each condition.

Accuracy at localizing the vehicle change was assessed using a 2 x 2 x 2 ANOVA. Change detection accuracy was measured by counting the number of trials on which participants correctly circled the exact location of the changed item on the data entry form (Appendix A). For target changes, participants correctly identified 96.1% of the changes, compared with 80.4% for distractor-change trials. This difference was statistically significant as predicted at $F(1,19)=7.843$, $p<.05$, Partial $\eta^2 = .292$. The effect of Number of Targets on change detection was also significant at $F(1,19)=4.538$, $p<.05$, Partial $\eta^2 = .193$. Surprisingly, in this case, change detection accuracy was significantly better when tracking 3 vehicles as compared with 1 vehicle.

Investigating the simple main effects for Change Detection Accuracy using multiple pairwise comparisons with a Bonferroni correction as in Experiment 1, the number of targets effect appears to be non-significant at each level of Change Type (target, distractor), though the difference in the no-steering / distractor condition is marginal at $p=.07$. Because the overall ANOVA for Number of Targets was significant, $p<.05$, this again is indication that the more stringent requirements of testing multiple comparisons makes it difficult to detect subtle differences. Finally, it is evident that the target-change vs. distractor-change difference is significant in both the steering-required and no-steering-required conditions, $p<.05$. 
Driving Metrics: Average headway distance, standard deviation of headway, and standard deviation of steering were analyzed using separate ANOVAs. An outlier analysis was performed for each driving metric, removing all trials that were >2.5 standard deviations from the mean for each subject, resulting in the removal of 1.3% of the data for average headway distance, 2.4% of the data for SD of headway, and 3.2% of the data for SD of steering. There were no significant differences in the percentage of outliers removed in each condition. Because the change detection task did not occur until the very end of the trial, when active driving had already ceased, there was no opportunity to measure the effect of the type of change on the driving metrics.

For average headway distance, the effect of Number of Targets was non-significant as expected, p>.05, with an average headway of 22.6 metres in the 1 target condition vs. 22.1 metres in the 3 target condition. For standard deviation of headway distance, the effect of Number of Targets was significant, $F(1,19)=5.836$, $p<.05$, Partial $\eta^2 = .235$, with a mean of 4.84 and 4.57 metres, for 1 and 3 targets, respectively. This effect is the reverse of what was expected, with deviations in headway being
significantly greater when there was one target to track. Finally, the standard deviation of steering did not differ significantly when there were 1 or 3 targets, p>.05; the averages were .105 metres and .108 metres, respectively, although greater deviations in the track-3 condition had been predicted.

![Figure 10](image)

Figure 10: Experiment 2a Driving Metrics: Average Headway Distance (left), Standard Deviation of Headway (middle) and Standard Deviation of Steering (right) in metres. Error bars represent standard error of the mean

### 6.3 Discussion

The high tracking accuracy found in all of the conditions is important, because it indicates that participants were very likely tracking (or not tracking, in the case of a distractor) an object when it underwent a change; this is vital for any interpretation of these findings. The fact that the number of targets had no impact on tracking accuracy is interpreted as a ceiling effect, in that the track-3 condition was easy enough to allow for optimal performance, on par with the track-1 condition. Note: the tracking task difficulty was intentionally low in this experiment, in order to provide an ideal environment for testing the change detection manipulation, i.e., one in which most of the targets will be successfully tracked. It is also interesting to note that the ease of this tracking task may be due to the colour composition of the tracking set used in this experiment. Specifically, there is a possibility that tracking performance is better when the tracking set is made up of uniquely coloured items, rather than items that are identical to one another (Makovski and Jiang, 2009). This question will be further addressed in Experiment 3.
Looking at the change detection accuracy, there are two important features to note. First, there is the number of targets effect, whereby performance at identifying the change was better when there were 3 targets as opposed to 1 target. This was unexpected, and at first glance seems counter-intuitive; after all, it might be expected that identifying a change would be easier when there is only one item to be tracked. One possible explanation involves the concept of object-based vs. space based attention. It may be that when tracking 3 objects, the viewer must adopt a broader attentional lens (e.g., Erikson and St. James, 1986), as compared with tracking 1 object. Although previous experiments such as Kramer and Hahn (1995) have demonstrated a lack of identity information for points intermediate to pre-cued locations, these cues were fixed to particular, and unchanging, space-based locations. In the present experiments, attention was necessarily allocated to the objects, in order that they could be tracked as they moved across the visual space. Considering this, and the fact that object-based attention has been shown to have significant processing advantages as compared with space-based attention (e.g., Atchley and Kramer, 2001), it is possible that change events at locations over which the tracked items might be expected to move were also facilitated. This would potentially explain the finding that changes are detected more accurately when tracking three targets, as compared with one.

The second item of interest here is the significant effect of change type. Specifically, the finding that subjects were more accurate at detecting changes at target locations, as compared with non-target locations, is supportive of the notion that tracking a visual object provides the visual system with facilitated access to some properties of the tracked object (Pylyshyn, 2004; 2009). That is, because tracked items were indexed, and because those indexes allowed for enhanced retrieval of information regarding the indexed item, processing was facilitated at a level over and above what is available for non-tracked (and therefore non-indexed) items.
Driving Metrics: Looking at the Average Headway Distance, it can be seen that drivers kept an equivalent distance regardless of whether there were one or three targets. Because it is assumed that deciding upon a headway distance is a tactical-level operation, involving situation-relevant action decisions (Michon, 1985), no difference was expected here. In contrast, because standard deviations of headway distance and steering were assumed to be operational level activities (i.e., the second-to-second actions required for maintaining vehicle speed and heading), it was expected that a more difficult task would result in poorer SD of Headway and SD of Steering. This did not occur. Instead, the effect Number of Targets on SD of Headway was in the opposite direction, with drivers demonstrating more variability in their headway when there was one target, as compared with three. For SD of Steering, there was no significant difference between tracking 1 and 3 targets.

A couple of things may be going on here. First of all, the track-3 task is not markedly harder than the track-1 task, as can be seen by the statistically equal tracking accuracies found in each condition. As such, any effects that were anticipated to occur as the result of increased task load, may not appear until a more difficult task is encountered. This could account for the lack of difference for the SD measures between 1 and 3 targets. The reversal seen for SD of Headway, with participants maintaining poorer control of their headway distance when there was one target to track, is also interesting. One possible explanation for this pattern of results involves not the difficulty of the task, but rather the simplicity. That is, the track-1 condition is very easy, and as a result, the participant may become somewhat bored or inattentive. In such a situation, the participant’s attention may wander, and as a result he or she may tend to maintain poorer control of the vehicle, resulting in the observed pattern.
7. Experiment 2b: Alternate implementation of change detection for target vs. distractor vehicles in a multiple vehicle tracking task: vehicle switching

Though Experiment 2a goes a long way towards determining the presence of facilitated attention for tracked vs. non-tracked vehicles, there are a few drawbacks to the design, which Experiment 2b was intended to ameliorate. First, the change detection task from Experiment 2a, which was based on the one-shot change detection paradigm (e.g., Velichkovsky et al., 2002), occurred at the end of the task, when the vehicles had stopped moving. As such, it is possible that the hypothesized tracking mechanism (i.e., the object indexes) may have relaxed somewhat due to the decreased requirement of “tracking” when the targets had come to a standstill. That is, because active tracking is no longer required, there was some concern that the expected advantage might not manifest despite its existence while actively tracking. Another issue was the change items themselves. In 2a, the change item was a new addition to the tracking set, and therefore could be located due to its originality – that is, the original tracking-set may have formed a perceptual group which was intruded upon by a new item, thus increasing the likelihood of an accurate detection. To avoid this problem, Experiment 2b involved the detection of a vehicle switch, where two vehicles in the tracking set instantaneously change locations, thus avoiding any new intrusions to the perceptual group. Finally, by making the vehicle switch occur in mid-trial rather than at the end of a trial, and by requiring a button-press response, in Experiment 2b I was able to collect response-time as well as accuracy data using the methodology described in the Design and Procedure section below (also see Figure 11). The predictions for Experiment 2b were similar to those of 2a, with the addition of response time as a dependent variable: facilitated performance (RT and Accuracy of Change Detection) when the change occurs for target as compared with a distractor vehicle. Participants were not required to drive the simulator (no steering or headway requirements) in Experiment 2b.
7.1 Methods

Participants: 30 undergraduate students (15 males and 15 females) between the ages of 18 – 28 from the University of Guelph Dept. of Psychology participant pool took part in this study. Average age was 20 years (SD = 2.8). Average driving experience was 4.03 years (SD = 2.6). Participants were asked to confirm that they had normal or corrected to normal vision, normal colour vision, and that they did not have any medical issues that could be dangerous in a driving simulator (epilepsy, heart conditions). Participants received 1 course credit for taking part in the study.

Apparatus and Stimuli: The apparatus was the same as for previous experiments. Six vehicles of similar shape (mid-size sedan), consisting of a heterogenous selection of colours (see Table 4), were arrayed on a standard 3 lane highway in a natural-seeming arrangement. Vehicles in the row closest to the driver (row 1) subtended an average visual angle of 2.97º, ranging from 2.01º at the most distant position, to 3.44º at the closest position. The vehicles in row 2 and 3 subtended an average of 2.1º degrees (range 1.9º to 2.3º) and 1.86º (range 1.72º to 2º), respectively.

<table>
<thead>
<tr>
<th>Tracking Set</th>
<th>Possible Vehicle Switches</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Celica Purple</td>
<td>1-2, 1-3, 1-4, 1-5, 1-6</td>
</tr>
<tr>
<td>2. Grand Prix Tan</td>
<td>2-3, 2-4, 2-5, 2-6</td>
</tr>
<tr>
<td>3. Grand Prix Red</td>
<td>3-4, 3-5, 3-6</td>
</tr>
<tr>
<td>4. Grand Prix White</td>
<td>4-5, 4-6</td>
</tr>
<tr>
<td>5. Grand Prix Green</td>
<td>5-6</td>
</tr>
<tr>
<td>6. Lexus (light) Blue</td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Tracking set and possible vehicle switches for Exp. 2b. For example, “1-2” means that Vehicle 1 (Celica Purple) and Vehicle 2 (Grand Prix Tan) switch places.
**Design and Procedure:** Experiment 2b was a one-independent-variable design, with *Change Type* (target, distractor) as the factor. The procedure was generally similar to Experiment 2a, with a few important differences. First of all, there were always 3 targets to track, in order to allow for the vehicle switching manipulation. There were no catch trials (trials on which no change occurred) in Experiment 2b. Instead of having a new vehicle appear during a change trial, in this experiment two of the vehicles (either two target vehicles, or two distractor vehicles) instantaneously switched locations. Participants were required to indicate that they noticed the vehicle-switch by pressing a button mounted on the steering wheel, and then indicate the location of each of the two vehicles that changed place, using a pencil-and-paper grid (Appendix A). Rapid response to perceived changes via the button press was encouraged. As before, there was five minutes of practice, after which there were 40 experimental trials, counterbalanced across all conditions. Each trial began by cueing the target vehicles, after which the vehicles began to change lanes in a random order. At a randomly timed interval during the tracking task, a global transient (screen flash) occurred, the simulation paused, and either two of the target vehicles, or two of the distractor vehicles switched places. The participant indicated his/her response with a button press, and the simulation resumed and continued until the end of the trial, at which point the participant indicated the location of the 3 target vehicles as in previous
experiments. A button press again unpaused the simulation, which continued with the next trial. The dependent variables were accuracy at tracking, accuracy at change detection, and response time for the change detection task.

‘Vehicle Switch’ Detection Task Flow

![Diagram of the task flow](image)

Figure 12: Task-flow diagram of procedure for Experiment 2b. For the ‘identify switched vehicles’ and ‘identify targets’ stages, the data collection form (Appendix A) was used.

7.2 Results

A simple ANOVA was performed to determine the effect of the vehicle switch manipulation on Tracking Accuracy and Change Detection Accuracy.

*Tracking Accuracy*: Overall tracking accuracy for Experiment 2b was high at 88.23%. The ANOVA revealed that tracking accuracy did not differ significantly between the target-vehicle and distractor-vehicle switches (average tracking accuracy of 87.9% and 88.57% respectively), p>.05. Estimates of performance when guessing 1 or 2 items was calculated as in previous experiments,
indicating that participants were performing better than would be expected if they were guessing one of the items and tracking the other two (see Table 5).

![Tracking Accuracy](image)

Figure 13: Percentage of correctly identified targets for Experiment 2b, for target and distractor changes. The lines represent the estimated accuracy if participants were randomly guessing 1 or 2 of the vehicle locations.

<table>
<thead>
<tr>
<th>Targets</th>
<th>Distractors</th>
<th>Guess 1</th>
<th>Guess 2</th>
<th>Est. # Tracked</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 targets</td>
<td>87.9%</td>
<td>88.6%</td>
<td>75.0%</td>
<td>60.0%</td>
</tr>
</tbody>
</table>

Table 5: Obtained accuracy was compared with predicted guessing using a one-sample t-test for each comparison. Statistically equivalent scores, p<.05, are shaded the same on the horizontal line, and outlined. In this case, all obtained accuracies were significantly higher than what would be expected if participants were guessing 1 of the items.

*Change Detection Accuracy:* For the change detection (vehicle switch) analysis, only trials on which 100% of the vehicles were accurately tracked were used. Accuracy of change detection was calculated such that indicating both change locations for a trial resulted in a score of 100% for that trial, while correctly indicating only one of the change locations would result in a score of 50% for that trial. The ANOVA revealed that, contrary to the prediction, accuracy of change detection was statistically equivalent for target and distractor switches, (62.57% and 61.81%, respectively), p>.05. Chance performance at detecting the 2 switch vehicles out of the 6 vehicle tracking set would be the
probability of guessing both right: \( \left( \frac{2}{6} \times \frac{1}{5} \right) \times 2 = 0.132 \), plus probability of guessing 1 right: \( \left( \frac{2}{6} \times \frac{4}{5} \right) + \left( \frac{4}{6} \times \frac{2}{5} \right) = 0.528 \) = \( \frac{66}{100} \), divided by 2, equals 33%. A one-sample t-test indicated that obtained accuracy for target change trials (65.79%, which is equivalent with accuracy distractor change trials at 66.21%) was significantly higher than chance, \( t = 9.837, p<.05 \).

**Response Time (RT) Data:** A separate ANOVA was performed to assess whether there was a significant difference between button-press response times for target, as compared with distractor changes. An outlier analysis was performed on the RT data, comparing each trial to the mean for that subject, and removing all trials that were >2.5 standard deviations from the mean, resulting in the removal of 1.7% of the data. With regards to the speed of the change detection tasks, there was a significant effect, \( F(1, 29)=4.550, p<.05 \), Partial \( \eta^2 = 0.136 \), indicating that participants were significantly faster at detecting target changes, as compared with distractor changes. This is as predicted, and the overall pattern (no accuracy effect, but an effect on response times) is similar to that found by Sears and Pylyshyn (2000).

![Figure 14: Accuracy (left) and Response Times (right) for the change detection task. Accuracy at the change detection task was scored such that for each trial, locating both changes gave a score of 100%, and locating 1 change gave a score of 50%. Change detection RT indicates the speed at which participants located the target or distractor change via a button press. Error bars represent the standard error of the mean.](image-url)
7.3 Discussion

Overall tracking accuracy was high, indicating that participants were able to track the vehicles most of the time. This is important, as it is vital that tracked and non-tracked vehicles are identified correctly in order for the analysis to be valid, and only correct trials were used in the analysis. A change detection trial was only marked correct if the participant correctly identified the exact location of the change in the post-trial paper test, rather than simply noting that a change occurred. Interestingly, in Experiment 2b there was no significant effect of change type on detection accuracy – participants were equally accurate at identifying the locations of target and distractor vehicle-switches. While this result is somewhat puzzling, one possible explanation exists in the fact that the change detection accuracy levels were rather low at 60%. Although this is significantly higher than chance performance (33%), change detection performance in Experiment 2b is low when compared with the change detection scores from Experiment 2a (~95% for targets and ~80% for distractors). One likely contributing factor here is that the “vehicle switch” version of the change detection task requires two responses (locating each of the changed vehicles), which is twice as many as in standard change detection, thereby increasing the difficulty of the task. In the response time data we do see the expected pattern, with target changes being identified more rapidly than non-target changes. Once again, only trials on which all targets were correctly identified, and on which the vehicle switch was also correctly identified, were used in the RT analysis. It is noted that Sears and Pylyshyn’s (2000) original paper also found no effects of change type on tracking accuracy, and significant effects of change type on RT. As such, Experiment 2b can be taken as additional support for the notion that indexed items enjoy facilitated access to focused visual attention.
8. **Experiment 2c: Braking event detection for target vs. distractor vehicles in a multiple vehicle tracking task.**

Experiment 2c also investigated responses to events at target vs. distractor vehicles. Instead of using abstract vehicle changes as in 2a and 2b, in 2c I looked at the time it took for a participant to depress the brake pedal, after seeing a braking event at either a target or distractor location. The methodology allowed for the collection of reaction time and accuracy data, with detection events occurring at a random time in the middle of each trial. The previous experiments have provided further evidence that indexed items enjoy facilitated access to attention, in that tracked items have better RT and accuracy scores for change detection tasks, as compared with non-tracked items. This conclusion hinges on the notion that the tasks are attention demanding, in that their successful performance requires the application of visual attention. In Experiment 2c I investigated detections for events that are arguably much less attention demanding: brake onsets. Because of the visual onset that occurs when a brake light is switched on, attention is attracted by the local transient, (visual onset, e.g., Kramer and Hahn, 1995), rather than having to be directed by top-down cognitive processes. If the advantage previously shown for target vs. non-target items was due to some third factor (e.g., participant strategies), then the target advantage should manifest in this case as well (Hypothesis A). Alternately, if the previously demonstrated advantage occurred because attention could be applied more rapidly to a task that required attention, then that advantage should disappear when the event detection task has a visual onset (local transient) which precludes the necessity of applying visual attention (Hypothesis B).
8.1 Methods

Participants: 19 undergraduate students (6 males and 13 females) between the ages of 19 – 22 from the University of Guelph Dept. of Psychology participant pool participated in this study. Average age was 20.1 years (SD = 1.08). Average driving experience was 4.05 years (SD = 1.06). Participants were asked to confirm that they had normal or corrected to normal vision, normal colour vision, and that they did not have any medical issues that could be dangerous in a driving simulator (epilepsy, heart conditions). Participants received 1 course credit for taking part in the study.

Apparatus and Stimuli: The apparatus was the same as for previous experiments. Six vehicles of a heterogenous colour selection (See Table 6) and similar shape were arrayed on a standard 3 lane highway in a natural-seeming arrangement. Vehicles in the row closest to the driver (row 1) subtended an average visual angle of 2.97º, ranging from 2.01º at the most distant position, to 3.44º at the closest position. The vehicles in row 2 and 3 subtended an average of 2.1º degrees (range 1.9º to 2.3º) and 1.86º (range 1.72º to 2º), respectively. Braking events occur using each vehicle’s built-in brake-light panels, though the vehicle did not alter speed during the event.

<table>
<thead>
<tr>
<th>Tracking Set</th>
</tr>
</thead>
<tbody>
<tr>
<td>Celica Purple</td>
</tr>
<tr>
<td>Grand Prix Tan</td>
</tr>
<tr>
<td>Grand Prix Red</td>
</tr>
<tr>
<td>Grand Prix White</td>
</tr>
<tr>
<td>Grand Prix Green</td>
</tr>
<tr>
<td>Lexus (light) Blue</td>
</tr>
</tbody>
</table>

Table 6: Tracking Set for Exp 2c: Brake Onsets
Design and Procedure: Experiment 2c used a 1-factor design. The factor was Braking Event Location (target, distractor). As in Experiment 2b, there were always 3 targets and 3 distractors for a total of 6 vehicles. The procedure was generally similar to that of Experiment 2a and 2b. Each trial began with 5 minutes of practice, after which the participant completed 40 Experimental trials, equally (and randomly) distributed between target and distractor brake onsets. The target vehicles were visually cued by flashing off and on for 3 seconds, after which all of the vehicles began to change lanes in a random order. During each trial 1 braking event occurred at either 16, 20, or 24 seconds into the trial, distributed randomly and evenly across all trials. Each trial had 1 braking event. For each event, the brake lights of either a target or a distractor vehicle were briefly activated (500ms), and the participant was required to respond using the simulator's brake pedal. After the braking event, tracking continued until the end of the trial, at which point participants were required to indicate which were the target items. Each trial lasted approximately one minute; the dependent variables were speed and accuracy at event detection, as well as accuracy at the tracking task. Tracking accuracy was collected as in previous experiments. Event detection was measured by taking the difference between the event onset, and the time at which the participant depressed the brake pedal (i.e., Braking RT). Event detection accuracy was collected at the end of each trial, by requiring the participant to indicate the
precise object location (either at a target object, or a distractor object) of the braking event.

8.2 Results

A 1-factor ANOVA was performed to determine whether there were any significant differences in accuracy of tracking and brake event detection, between the target onset and distractor onset conditions.

Tracking Accuracy: Once again, overall tracking accuracy was high, whereby the percentage of correctly identified targets was 93% averaged across all conditions. Participants’ accuracy at the tracking task did not differ significantly regardless of whether the braking vehicle was a target or a distractor, p>.05. Actual tracking performance was compared with the expected outcome if participants were guessing one or two of the targets. This analysis indicated that the obtained accuracy was significantly higher than would be expected if participants were guessing one target and tracking the other 2 (see Table 7).

![Tracking Accuracy graph](Figure 16)

Figure 16: Percentage of correctly identified targets for Experiment 2c, for target and distractor changes. The lines represent the estimated accuracy if participants were randomly guessing 1 or 2 of the vehicle locations. The error bars represent the standard error of the mean.
### Estimated Guessing

<table>
<thead>
<tr>
<th></th>
<th>Targets</th>
<th>Distractors</th>
<th>Guess 1</th>
<th>Guess 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 targets</td>
<td>92.0%</td>
<td>93.8%</td>
<td>75.0%</td>
<td>60.0%</td>
</tr>
</tbody>
</table>

Table 7: Obtained accuracy was compared with predicted guessing using a one-sample t-test for each comparison. Statistically equivalent scores, p<.05, are shaded the same on the horizontal line, and outlined. In this case, all obtained accuracies were significantly higher than what would be expected if participants were guessing 1 of the items.

**Event Detection Accuracy:** Only trials for which all tracked items were correctly identified were used in the analysis. In order for a change detection trial to be marked accurate, the participant had to identify the exact location of the brake-onset event using the data collection form (Appendix A). Detection of braking events differed significantly for target vs. distractor events, \(F(1, 18) = 16.318, p<.005, \text{Partial } \eta^2 = .475,\) with an accuracy of 87.7% for target events and 73.7% for distractor events. This advantage for event detection accuracy was not predicted.

**Response Time to Brake Onset:** Response times to brake-onsets for target and distractor vehicles were analyzed using a separate ANOVA. An outlier analysis was performed, removing all trials that were >2.5 standard deviations from the mean for that subject, resulting in the removal of 3.9% of the data. Target onsets were detected at an average of 1018 ms (standard error = 52 ms), while distractor onsets were detected at an average of 1024 ms, (standard error = 54 ms) a difference that is non-significant, \(p > .05,\) indicating that, as predicted, participants were equally fast at responding to target and distractor braking events.
8.3 Discussion

The results of Experiment 2c are mixed, but informative. First, there was a fairly strong (~14%) accuracy advantage for target vs. distractor brake onsets. This is interesting because it runs counter to the prediction, that when a visual onset is used, the detection advantage for tracked vs. not-tracked items should disappear. Instead, we see the change detection accuracy advantage persisting. The fact that participants were better able to report the location of the target onsets, as compared with the distractor onsets, may suggest that the “object indexing” mechanism is robust enough to provide localization advantages which persist over a short-to-medium term duration. There is however one consideration that must be taken into account, as we will see below. Interestingly, the prediction of no facilitation for tracked, as compared with not-tracked items held true for the response time measure. As predicted, there were no differences between response times for braking events at target and distractor locations. This indicates that the attentional advantage (at least in terms of speed) for target changes was negated when an attention-grabbing cue (visual onset) was present at the change object.
One point that may shed some light on the reason that the target advantage disappeared for response times, but not for event detection accuracy, is that the accuracy measure was taken at the end of each trial, rather than immediately coinciding with the event itself. That is, after noticing the change in the middle of the tracking task, participants had to continue to track that particular car until the trial ended, and then enter the correct location on the data collection form (Appendix A). There are a number of things that could have happened during this interval. One possibility is that participants could have been more likely to lose track of a distractor-event vehicle, as compared with a target-event vehicle. If the item itself was lost, then it would be impossible for the participant to achieve a correct localization of the braking event for that trial, thus producing the described pattern. Despite this possible explanation, it is important also to note that different processes may be at work when detecting sudden onsets in the middle of a tracking task, versus successfully tracking the locations of the change-items until the very end of the trial. As discussed in Experiment 3, it is also possible that there is a visual working memory component involved when localizing the changes, but not when responding rapidly to visual onsets.

Given the above uncertainties with the event detection accuracy measure, the response time measure may be more indicative of what is actually occurring at the precise moment of an event onset. That is, at the immediate time of the braking event, tracking an item provides no reaction time advantage; rather, RT to both target and distractor onsets are equivalent. This was arguably a result of the local transients / onsets that occur with a braking event, which override any attention-allocation advantage that was due to the presence of object indexes.

8.4 Experiment 2 summary

In both this experiment, and in the previous experiments (2a and 2b), participants were required to make responses to changes at target and distractor locations. However, when the task required the
participant to allocate attention to the item in question (i.e., in the absence of any local visual transients), responses were significantly more accurate (2a) and rapid (2b) for target as compared with distractor vehicles. This supports the notion that indexes for tracked objects may provide the attentional system with rapid access to the spatial co-ordinates of the items being tracked. However, when participants were provided with robust visual cues (onsets) that grabbed attention at the changed item, the speed advantage provided by the object indexes was negated, as the local visual onsets were able to direct attention regardless of whether the object was being tracked or not. The results of Experiment 2c also indicate that even in the presence of visual onsets, which directed attention to the change item, participants were able to localize target changes more accurately than distractor changes. This finding is interesting in light of the fact that, when there are visual onsets available to direct attention, the reaction time difference for detecting changes at target and distractor locations disappeared – that is, both targets and distractors were detected with equal rapidity. This again may be taken as an indication that two different processes are at work.

9. Experiment 3: Visual working memory and tracking-set composition

There is evidence that although object identity is not necessarily available to the participant in a tracking task (e.g., Pylyshyn, 2004), there is nevertheless an advantage to be had when the tracking items are visually distinct (e.g., Makovski and Jiang 2009). Specifically, when the targets and distractors had unique identities (for example when each item had a unique colour, as in the “heterogenous” condition of Makovski and Jiang) tracking accuracy was improved. In that study, the authors presented viewers with a modified multiple object tracking task, using coloured discs instead of the identical circles usually employed. The result of this manipulation was an increase in tracking accuracy when individual object detail was available, as can be seen in Figure 18, from the 2009 paper.
Horowitz et al. (2007) also investigated the effect of unique item identity in a multiple object tracking task, and found facilitated tracking performance. In this case the stimuli were unique, namable objects (cartoon figures), and the unique stimuli benefitted tracking performance significantly. Interestingly, both papers bring up the notion of attention as a dual system – one for object locations in visual space, and one for object identity. Traditionally, cognitive processes are thought to be carried out using a general, resource-limited processing mechanism, which has been referred to as executive function, or working memory (e.g., Allen et al., 2006). Working memory has been further subdivided by sensory modality, including a distinct construct referred to as visual working memory. When one studies a picture for a brief interval, and then is asked to recall details, it is this visual working memory that is employed. The interesting conjecture that arises out of Horowitz et al., (2007) and Makovski and Jiang (2009), is whether the “location system” is what Pylyshyn is referring to with the FINST / indexing theory, and further, that the object identity information is maintained in parallel with object location, by visual working memory.

In Experiment 3 I tested the effects of item heterogeneity on tracking performance, determining the extent to which tracking was enhanced in a realistic tracking environment in which heterogeneity is the norm (driving), and further, how the composition of the tracking set affected
standard measures of driving ability, including steering and headway maintenance. It was predicted that tracking performance would increase when unique item detail was available, as per Makovski and Jiang (2009). With regards to steering and headway maintenance, because the added object detail should not change the ambient nature of the tracking task, it was expected that the homogeneity / heterogeneity of the tracking set would not interfere with driving performance. My prediction here was that there would be no effect of tracking-set composition on any of the driving metrics, SD of steering, average headway, or SD of headway. However, because the addition of a secondary task (steering / headway maintenance) would demand a greater level of attention (ambient or otherwise) it was considered possible that this additional requirement would interfere with visual working memory – indexing associations and thus remove or reduce the benefit of tracking heterogeneous items. I therefore predicted a reduced benefit of heterogeneity in the steering / headway maintenance-required condition. Finally, pilot testing revealed that participants tend to try and memorize the colour set (e.g., by mentally rehearsing the target colors, such as saying ‘red-green-blue, red-green-blue’ to themselves). In order to determine the effect of this potentiality, a between-subjects manipulation, Articulatory Suppression, in which the participant was required to repeat a word to him/herself, (thus blocking out any articulatory maintenance of the target colours) was included. If the participants use mental rehearsal to perform the task, then tracking accuracy should decrease in the articulatory suppression condition.

9.1 Methods

Participants: 48 undergraduate students (31 males, 17 females) between the ages of 19 – 27 from the University of Guelph Dept. of Psychology participant pool participated in this study. Average age was 20.29 years (SD = 1.6). Average driving experience was 4.3 years (SD = 1.7). Participants were asked to confirm that they had normal or corrected to normal vision, normal colour vision, and
that they did not have any medical issues that could be dangerous in a driving simulator (epilepsy, heart conditions). Participants received 1 course credit for taking part in the study.

*Apparatus and Stimuli:* The apparatus was the same as in Experiments 1 and 2. The tracking set was made up of 8 vehicles, which consisted of either a homogenous colour set (i.e., all the same), a heterogeneous assortment of colours (all different from one another), or a replication of the “Paired-4” condition from Makovski and Jiang (2009), in which the 4 targets and the 4 distractors each made up an identical heterogenous set (for example: red, blue, green, yellow; red, blue, green, yellow). Vehicle colours were distributed randomly across all vehicles at the start of each trial, and there were a total of 6 unique colour sets – 2 for each condition. The cars were situated on a 3-lane roadway in a realistic arrangement as in previous experiments. Two possible versions of each condition (so, 2 versions of the homogenous, 2 versions of the heterogenous, and 2 versions of the “paired-4”) were available to be presented during the experiment.

<table>
<thead>
<tr>
<th>Homogenous</th>
<th>Heterogenous</th>
<th>“Paired Four”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Targets</td>
<td>Distractors</td>
<td>Targets</td>
</tr>
<tr>
<td>Red</td>
<td>Red</td>
<td>Red</td>
</tr>
<tr>
<td>Red</td>
<td>Red</td>
<td>Green</td>
</tr>
<tr>
<td>Red</td>
<td>Red</td>
<td>Blue</td>
</tr>
<tr>
<td>Red</td>
<td>Red</td>
<td>Cyan</td>
</tr>
</tbody>
</table>

Table 8: Example of colour conditions used in Experiment 3
Design and Procedure: Experiment 3 used a 3 x 2 x 2 mixed factorial design. The first factor was Tracking Set Composition (homogeneous, heterogenous, and “paired-four” shown above in Table 8) and the second factor was Task Load (vehicle control required, no vehicle control required), both of which were within-subjects manipulations. The third manipulation, varied between-subjects, was Articulatory Suppression, in which half of the participants were required to repeat the word “the” while performing the tracking task, in order to prevent the mental rehearsal of the target-item colours.

The procedure was generally the same as in Experiments 1 and 2. Participants were required to track 4 targets amidst a total of 8 vehicles on a simulated 3-lane roadway. After completing 5 practice trials (approximately 5 minutes), the participants completed a total of 48 trials, across which the colour-type and task-load factors were equally counterbalanced. For each trial the tracking task lasted for 30 seconds, after which participants were asked to indicate, using the pen-and-paper method, which vehicles were the targets (full report technique). The dependent variables were accuracy at the tracking task (% of correctly identified targets), the standard deviations of steering and headway distance, as well as average headway distance.
9.2 Results

Tracking Accuracy: Overall tracking accuracy was 80%. A 3 x 2 x 2 mixed ANOVA was performed. The between-subjects factor Articulatory Suppression was non-significant, F(1, 47)=2.239, p=.141, \( \eta^2 = .046 \), and as such the conditions were combined. The effect of Tracking Set Composition was significant at F(2,94)=9.019, p<.001, \( \eta^2 = .161 \), indicating that, as expected, tracking accuracy was higher in the Heterogenous condition, as compared with the Homogenous and Paired-4 conditions. Likewise, the predicted effect of Task Load was significant at F(1,47)=16.061, p<.001, \( \eta^2 = .255 \), indicating that tracking accuracy was poorer when the participant was required to operate the vehicle. Finally there was a Tracking Set Composition X Task Load interaction, consistent with the prediction, significant at F(2,94) = 4.142, p < .05, \( \eta^2 = .081 \), indicating that the effects of the Tracking Set Composition differed across levels of the Task Load manipulation (see Figure 20).

Investigating the simple main effects for Experiment 3 accuracy scores, using a Bonferroni correction for multiple comparisons, it is evident that while the requirement of driving the vehicle significantly reduces tracking accuracy in the Heterogenous and Homogenous conditions, p<.05, there is no significant difference in the Paired 4 condition. Likewise, it is evident that accuracy in the Heterogenous condition is significantly higher than in either the Paired-4 or Homogenous conditions, p<.05, while accuracy in the Paired-4 and Homogenous conditions is statistically equivalent.

Estimated accuracy if the participants were guessing 1 or 2 items is given in Table 9. One-sample t-tests were used for each close comparison between estimated guessing, and actual performance. Tracking accuracy was significantly better in all cases than would be expected if participants were guessing 2 items, (p<.05), and equivalent with what would be expected if they were guessing 1 item – the only exception being in the no-driving / homogenous condition, where tracking accuracy was significantly better than would be expected if they were guessing one item, (p<.05).
Figure 20: Percentage of correctly identified targets for Experiment 3. Error bars represent standard error of the mean.

**Estimated Guessing**

<table>
<thead>
<tr>
<th>Est. # Tracked</th>
<th>Tracking Only</th>
<th>Heterogenous</th>
<th>Paired-4</th>
<th>Homogenous</th>
<th>Guess 1</th>
<th>Guess 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tracking Only</td>
<td>Heterogenous</td>
<td>Paired-4</td>
<td>Homogenous</td>
<td>Guess 1</td>
<td>Guess 2</td>
</tr>
<tr>
<td>Tracking Only</td>
<td>87.5%</td>
<td>80.4%</td>
<td>81.7%</td>
<td>80.0%</td>
<td>62.5%</td>
<td></td>
</tr>
<tr>
<td>Est. # Tracked</td>
<td>&gt;3</td>
<td>3</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tracking + Driving</td>
<td>78.3%</td>
<td>76.4%</td>
<td>75.7%</td>
<td>80.0%</td>
<td>62.5%</td>
<td></td>
</tr>
<tr>
<td>Est. # Tracked</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 9: Obtained accuracy was compared with predicted guessing using a one-sample t-test for each close comparison. Statistically equivalent scores, p<.05, are shaded the same on the horizontal line, and outlined.

**Driving Metrics:** The effects of Tracking Set Composition on standard deviation of steering, average headway distance, and standard deviation of headway distance were investigated using 3 separate 3 x 2 mixed ANOVAs, with Articulatory Suppression as a between-subjects factor. An outlier analysis was performed for each of the driving metrics, removing all trials that were > 2.5 standard deviations from the mean score of each metric, which resulted in the removal of 1.39% of the data for average headway distance, 3.47% of the data for SD of headway, and 4.86% of the data for SD of steering. The effect of Articulatory Suppression on each of the driving metrics was non-significant, p>.05, and the conditions were therefore combined. Individual means, for suppression and non-
suppression conditions (respectively) are: for SD of Steering, .172 and .183 metres; for Average Headway Distance, 20.27 and 22.7 metres; and for SD of Headway Distance, 4.29 and 4.39 metres.

For the average headway distance that drivers kept, the effect of Tracking Set Composition was non-significant as predicted, $p>.05$, indicating that headway distances were statistically equivalent in all conditions. For Standard Deviation of Headway Distance, the effect of Tracking Set Composition was significant at $F(2,94)=3.33$, $p<.05$, partial $\eta^2=.066$, contrary to the prediction, indicating that the ability of the driver to maintain a steady headway was affected by the colour composition of the tracking set. Specifically, headway deviations were lower in the Paired-4, as compared with the Heterogenous or Homogenous conditions. For SD of Steering, the effect of Tracking Set Composition was significant at $F(2,94)=5.607$, $p<.05$, partial $\eta^2=.107$, indicating that steering deviations differed depending upon the composition of the tracking set. This difference was not expected. Specifically, deviations were lower in the Heterogenous condition, as compared with the Paired-4 and Homogenous conditions.

![Figure 21: Experiment 3 driving metrics: Graphs of effects for Average Headway Distance (left), SD of Headway (middle), and SD of Steering (right). Error bars represent standard error of the mean.](image)

### 9.3 Discussion

In Experiment 3 it has been shown that, as predicted, the composition of the tracking set does influence the ability of a participant to track multiple vehicles in a simulated roadway environment.
Specifically, when the tracking set is made up of a heterogenous selection of colours, tracking accuracy was higher than when the tracking set was either all one colour (Homogenous), or when there was one target and one distractor in each colour (Paired-4). This is interesting because it indicates that, at least in terms of the composition of the stimuli, the classical multiple object tracking task may fail to take into account an important aspect of tracking objects in a realistic environment – namely that such heterogeneity often exists in naturalistic environments, and can be used to benefit performance.

Even more interesting for the present analysis is the interaction between Tracking Set Composition and Task Load. Here, it is shown that while the above effect exists when the participant is not required to operate the vehicle, this effect disappears entirely when the participant is required to steer the vehicle and maintain a specific headway. That is, when driving the car, any advantage that the heterogenous condition provided is absent, and performance is equal across all levels of Tracking Set Composition. This finding ties nicely in with Makovski and Jiang (2009), who concluded that visual working memory (VWM) was the binding force that allowed for the benefits of a Heterogenous tracking-set. Specifically, because the secondary task uses visual working memory resources, these resources are not available to bind the object identity to the object index – thus no benefit is present in the Multiple Task condition.

Finally, there were significant effects of Tracking Set Composition on both SD of Headway Distance and SD of Steering. This was not predicted; rather, based on the logic that added object detail would not alter the ambient characteristics of the task, and that steering and headway maintenance are thought to be ambient tasks (Wickens, 2002), no effects were expected. One possible interpretation of the current effects is that steering and keeping headway are not purely ambient tasks. Another, more interesting interpretation is that adding object detail to the items does not only change the focal nature of the items, but also the relational (therefore ambient) nature of the objects. That is, because there is
extra information available for each item, the ambient attentional system may have to do less work to track these items, thus freeing up resources that can be used to improve performance at other ambient tasks, such as steering and headway maintenance.

Looking at the standard deviation of headway distance, the advantage appears in the Paired-4 condition (where each tracking target has an identical distractor paired with it). This was unexpected, and at first glance defies easy explanation. In terms of accuracy at tracking the objects, ‘Paired-4’ is on par with the homogenous condition, and better than the heterogenous condition. As for why headway deviations would be lower in this condition, it could be that the combination of having visually distinct items, yet having some overlap in the overall tracking set composition (i.e., the ‘paired’ items), has an impact on how difficult people perceive the task to be, and therefore impacts how close they follow (note: in 2a drivers followed closer when there were 3 targets, as compared with 1 target, which would support this interpretation). This is also consistent with previous research (e.g., de Waard et al, 2008; Lansdown et al., 2004) demonstrating that headway and steering can be affected by the mental workload on the driver, for example in a situation where multi-tasking is required. In Lansdown et al., in particular, it was shown that headway distances decrease markedly when multiple tasks are involved while driving.

To gain further understanding of these effects, Michon’s (1989) 3-part hierarchy of driving tasks may be employed. These are: Strategic, for planning routes etc.; Tactical, for situation-relevant action decisions; and Operational, for the second-to-second operations of maintaining the desired heading and acceleration values. In relation to the current experiment, it can be seen that participants are better at keeping the vehicle in a straight line (SD of Steering) when the tracking set is heterogenous. This task falls into the Operational, or second-to-second activity required to operate the vehicle. Thus, because the individual object detail may decrease the overall difficulty of the task, more
resources are then available to maintain a steady steering angle. For the standard deviation of headway distance, however, this effect does not manifest in the same fashion. Instead, SD of headway distance is equivalent for Heterogenous and Homogenous tracking sets, but better in the ‘Paired-4’ condition. As the Paired-4 and Homogenous conditions are roughly equivalent in difficulty (as determined by equivalent tracking accuracy scores), this may yet be an indication that some element of the ‘Paired-4’ display was conducive to maintaining a steadier headway – that it was in some way easier due to the consistent colour associations. Notice that in this case, headway distances were statistically equal for all three colour conditions. In general, these findings may be taken as further evidence that our ability to track objects is aided by the presence of unique object information, and that this benefit is reduced or removed when a difficult enough secondary task is introduced.

10. General Discussion

In this series of experiments I have constructed a vehicle tracking methodology that attempts to transfer the commonly used multiple object tracking task (Pylyshyn and Storm, 1988) to a simulated driving environment. This is important for a number of reasons: first, because the driving environment is a good example of a real-world situation in which people need to pay attention to many different things at the same time. Even in the simplest driving scenario, one can expect situations in which simultaneous, or near simultaneous attention to multiple moving vehicles or people is necessary in order to avoid a collision. A classic example of such a situation is the left-hand turn at a signal, where there are potentially multiple points at which an intrusion could occur. The notion that the multiple object tracking task is applicable to the driving environment is further attested to by the fact that several authors (e.g., Fencsik et al, 2007; Feria, 2008; Horowitz et al., 2007; Oksama and Hyona, 2004) have made this connection in varying ways. Indeed, driving is one of the most commonly given
examples in the literature, when trying to explain the usefulness of the classical object tracking task. As discussed in this paper, there are numerous reasons that the original computer-desktop multiple object tracking task, with its 2-dimensional display and objects, 2-colour display, and lack of secondary task requirements, may not accurately predict how objects will be tracked in a real-world environment. Therefore, in order to maintain maximal compatibility with the rich literature base that already exists on multiple object tracking, while at the same time taking into account some of the complexities that exist in a realistic tracking environment such as driving, I have set out to elaborate on some classical object tracking findings, and to test some of the predictions that stem from the extensive research that exists on multiple object tracking.

8.1 Summary

In Experiment 1 I tested the simplest implementation of what I will refer to as multiple vehicle tracking. Participants were required to keep track of the locations of 1, 3, or 4 vehicles out of 9 total as they moved about a simulated roadway, while either operating the vehicle (steering and maintaining headway), or having the vehicle drive itself. Two important findings came out of this experiment. First, it was demonstrated that object tracking ability for identical vehicles on an apparent 3-d roadway is comparable to object tracking ability for identical circles on a computer desktop, in that tracking accuracy is generally high, and decreases with larger tracking sets. While the overall percentage accuracy demonstrated in Experiment 1, when tracking 4/9 targets, is lower than what is found in classical multiple object tracking studies which require the participant to track 5/10 objects, this can to some extent be attributed to the presence of self-motion, which has been shown to reduce object tracking performance (Thomas and Seiffert, 2010), as well as the other characteristics of object tracking in a realistic environment, such as apparent 3d and shared object trajectories.
The second important finding to come out of Experiment 1 is the fact that requiring the participant to steer the vehicle, and maintain a desired headway distance, has a negative impact on tracking accuracy. Because self-motion was present in all conditions, the reduced performance was not due simply to the presence of self-motion. Rather, Experiment 1 presents evidence that it is specifically the requirement of \textit{self-directed} motion (i.e., the requirement of operating the vehicle, in this context) that leads to the decrement in performance. It is therefore obvious that any estimate of tracking ability in a real-world environment would require one to take into account all of the various task-load factors that would come into play.

The Task Load x Number of Targets interaction found in Experiment 1 is also interesting. In this case, it is shown that the rate at which increasing numbers of targets impairs tracking accuracy is greater for situations in which a secondary task is required. Thus, the amount of free attentional capacity (i.e., not used by a secondary task) that the subject is able to devote to the task, has an impact on how well he or she can deal with increases in the number of tracking targets. To clarify, while there was only an approximate 5\% decrease in tracking accuracy between 1 and 4 targets in the no-task condition, in the driving condition that decrease was approximately 20\%, indicating that when focusing on a single task only, the effect of an increased number of targets is largely mitigated, at least in the 1-4 target range. This effect also operated in the opposite direction, whereby larger tracking sets resulted in poorer steering and headway maintenance, though average headway distance was not affected. The fact that average headway distance was not affected is interpreted as being a result of the fact that the standard deviation of steering and headway distance are Operational-level activities, while headway distance could be considered a Strategic choice (Michon, 1985).

Experiment 2 tested 3 different ways of measuring how tracking an object provides some benefit, when a task requires that attentional processes must be applied to that object. This is directly
related to Pylyshyn’s (1989) notion of an object indexing system for keeping track of moving points in physical space. Experiment 2a employed a classical “one-shot” change detection task as described by Velichkovsky et al. (2002), and succeeded in demonstrating that changes for tracked objects were detected more rapidly than changes for non-tracked objects. This supports the notion that attentional processing for tracked objects is facilitated, as compared with objects that are not being tracked.

This interpretation is further supported by Experiment 2b, which employed a modified version of the change detection task in which two of the vehicles switched places instantaneously (vehicle-switching manipulation). In this case, a button-press response allowed for the collection of reaction-time data, which indicated that responses for target vehicles were faster than responses for distractor vehicles. Surprisingly the accuracy of detection for targets and distractors was equal, and generally low (60%) when compared to standard change detection tasks (e.g., 95%, Velichkovsky et al., 2002), and to Experiment 2a (90%). This low change detection accuracy is likely due to the increased difficulty of the “vehicle switch” manipulation, which requires the detection of 2 change locations rather than 1; this also has possible implications as for why the target and distractor changes were detected at an equal rate (i.e., a possible ‘floor’ effect). Experiment 2b can be taken as further evidence that attentional processing is facilitated for tracked objects.

Experiment 2c applied the reverse logic – if the change could be detected without the directed application of focused visual attention, then there should be no advantages for tracked objects. For the speed of brake onset detection, this is what we found – brake onsets for both target and distractor vehicles were detected with equal rapidity. For accuracy at detecting the exact location of the brake onset, there was an advantage for target-change trials. This difference is interesting because it suggests that while the sudden onset of a brake light is enough to negate any RT difference between tracked and non-tracked vehicles, localization of the changed item is still better for tracked vs. non-
tracked items. This finding however must be considered in light of the fact that Experiment 2c contained a potential confound, where in the distractor-change condition, participants had to track the location of the distractor vehicle (change item) as well as the target vehicles – thus, the change detection accuracy difference in Experiment 2c is potentially interesting, but not conclusive.

Experiment 3 set out to test another prediction derived from the object tracking literature: that the composition of the tracking set will have an effect on the accuracy with which items can be tracked. The composition of the tracking set was varied, such that the cars were either all the same colour (homogenous), all different colours (heterogenous), or one of each colour (paired-four). As expected, it was found that tracking performance was improved when unique object colours were available to the observer. In addition, it was shown that increased task load (i.e., operating the vehicle as opposed to passive viewing) eradicates the advantage that was found for a heterogenous tracking set.

8.2 Implications

In Experiment 1 it was confirmed that object tracking performance, as tested in the classical MOT paradigm, can be adapted for use in a modern driving simulator. Expected effects of set size and task load were demonstrated; likewise these variables were shown to interact. It is important to note that, because even such well-practiced tasks as keeping the vehicle in a straight line and at a specific headway interfere with tracking ability, such task requirements must be included in any prediction of object tracking performance.

The findings of Experiment 1 also relate to the conclusions of a recent study conducted by Thomas and Seiffert (2010). In the Thomas and Seiffert study, the authors demonstrate the negative effects of self-motion on tracking performance. Specifically, they show that when the observer is in
motion (either walking around the display, or being pushed in a wheelchair), tracking performance is poorer than when the observer stays still (either walking in place, or standing still).

![Figure 22: Conditions from Experiment 1 of Thomas and Seiffert (2010)](image)

Interestingly, the present research reflects the difference in Thomas and Seiffert’s study between the “active move” and the “passive move” conditions. That is, driving the vehicle would be “active move”, and having the vehicle drive itself would be “passive move”. While Thomas and Seiffert’s analysis demonstrated no main effect of “Active vs. Passive” moving, it could be postulated that the obvious null effect in their 1-Target condition negated any chance of detecting an effect in the 3-Target condition. Inspecting Figure 22, it does appear as though there is a trend whereby when tracking 3 targets, performance in the “active move” condition is somewhat poorer than tracking performance in the “passive move” condition. Experiment 1 supports this notion, in that we have demonstrated concretely that tracking performance is poorer when the subject is required to drive the vehicle, compared with when the vehicle speed and steering is controlled by the computer. It is also considered that walking may be a substantially more automatic type of “active move” than driving.

Experiment 2a succeeded in demonstrating that changes for tracked objects were detected more accurately than changes for non-tracked objects. This supports the notion that attentional processing for tracked objects is facilitated, as compared with objects that are not being tracked. Experiment 2b further supported this interpretation. As a result of Experiment 2 in its entirety, there is significant
evidence that tracking a set of objects (such that they would be indexed by Pylyshyn’s theoretical FINST construct), can provide both speed and accuracy advantages when making decisions that require the application of focused visual attention (i.e., when there are non-attention-grabbing changes, such as visual offsets or changes that are obscured by a global transient). The results of Experiment 2b and 2c in particular support the notion that it is only when focused attention is required (such as when there are no local onsets to guide attention), that tracked items are detected more rapidly, once again indicating the value of the indexing hypothesis.

Experiment 3 investigated the extent to which the colour composition of the tracking set impacted the ability to track multiple objects. This question is interesting because in Pylyshyn’s original formulation of his FINST (Fingers of INSTantiation) object indexing system, the visual indexes are theorized to operate independently of the visual characteristics of the tracked object. That is, it is theorized that the indexes operate at an early-visual level, and simply keep a “tag” or “index” on the physical location of the object. In this conceptualization, the indexes do not encode the visual properties of the object, such as shape, colour, or texture. Studies such as Scholl et al., (1999) and Pylyshyn (2004) tend to support this interpretation, in that they found little or no evidence of memory for the identity of tracked objects.

There is, however, some recent evidence that unique object properties do have an impact on tracking ability. Specifically, Makovski and Jiang (2009) have shown that when the circles in a classical multiple object tracking task contain unique colours (though there is no requirement to memorize or report those colours), tracking performance improves markedly. At face value, this would seem to contradict Pylyshyn’s interpretation of how object indexing works. Indeed, the present Experiment 3 demonstrated that in a simulated vehicle tracking environment, tracking performance was likewise improved when unique object colours were available to the observer. The current
research, consistent with past studies (Bahrami et al., 2003; Makovski & Jiang, 2009), has demonstrated that there are definite effects of object properties, on the ability of individuals to track those objects. Pylyshyn’s original 1989 theory as to how multiple objects are tracked envisions an indexing system that is “preattentive”, and uninfluenced by object properties. Research in the current vein tends to upset this classical interpretation, in the sense that it has repeatedly been demonstrated that the individual properties of the object have an impact on object tracking performance.

8.2.1 Preserving the notion of an object oriented, pre-attentive indexing system (saving the wild FINSTS)

Does this mean that our concept of an object indexing system needs to be revised, such that an object index keeps track not only of the physical location, but of the physical properties of the object? There is an alternate explanation, one that is hinted at by the conclusions of Makovski and Jiang’s (2009) Experiment 4a. Makovski and Jiang attribute the tracking advantage for uniquely coloured items to the fact that visual working memory (VWM) can be used to differentiate the objects from one another, and that this system works in a separate but parallel manner with the object indexing system. Thus, it is possible that while Pylyshyn’s indexing system keeps track solely of the physical location of the target objects, our innate ability to remember the identity of a visual object is also at work, enabling us, for example, to rapidly regain a target that was lost during tracking, which would be missed if it weren’t for the fact that its unique identity retained its visual working memory trace.

Makovski and Jiang supported their conjecture on this subject by demonstrating that, when a visual working memory task (they used a centrally-presented “1-back” colour memory task) was run concurrently with a multiple object tracking task using either identical or unique colours, performance at both the tracking task and 1-back memory task was reduced; moreover, tracking performance for the
uniquely coloured items suffered more from the inclusion of a visual working memory task, as compared with tracking performance for the homogenous items. This was taken as evidence that both the tracking of uniquely-coloured items, and the performance of a 1-back colour memory task employ visual working memory, and therefore interfere with one another.

Interestingly, I have found further support for this interpretation in Experiment 3 of the present studies. Specifically, when a secondary task was included (in this case, the requirement of steering and maintaining a desired headway distance), the advantage at tracking uniquely-coloured cars disappeared (see Figure 20 from Experiment 3, above). Although it is arguable to what extent steering and maintaining headway require visual working memory, per se, it is obvious that they do require some level of extra mental effort, which we may assume has a working memory component. In any case, just as Makovski and Jiang demonstrated, when a secondary task is required, the advantage that was previously seen for uniquely-coloured objects disappeared. This provides us with the theoretically interesting conjecture that it is visual working memory, or rather perhaps just working memory in general, that is responsible for the advantage that is seen when tracking uniquely coloured objects, and that, in some sense, working memory operates in a separate but parallel manner, to aid in our ability to keep track of multiple unique moving objects.

8.3 Limitations

Perhaps the most obvious limitation of the current body of work is that it is a hybrid project, straddling two concurrent but separate fields of knowledge. This is a limitation in the sense that any such approach may lead to compromises in both of the areas of interest. Driving simulator research is often used to investigate societal or medical issues (teen texting, the use of cell phones, ADHD medication in the driving environment), and it is presupposed that the researcher is trying to replicate a
particular real-world environment in which to answer a particular question. In the current work, the “driving environment” is much less realistic, in terms of task and scenario goals.

On the other hand, classical cognitive psychology, though it has come a long way from the tachistoscope and mechanical counters of the past, nevertheless often continues to use abstract computer displays to tease apart the elements that make up human information processing systems. This abstraction allows for tight control of the various factors, allowing for greater certainty as to the operating factor when making conclusions. By attempting to take a classical object tracking task, and reproduce it in the driving environment, I have in these experiments given up a measure of control over the stimuli, to gain a realistic-seeming tracking environment. That is, for example, the cars move laterally at a speed controlled by the simulator, rather than a parameter that I was able to set. Likewise the range of stimuli (cars) was limited to that provided by the simulator software – the general notion being that more choices are determined by the driving simulator environment than would be desirable in a simple 2-d classical multiple object tracking task.

Attempting to generate a list of the ways in which this hybrid approach has impacted the research from the perspective of either a driving researcher, or a cognitive psychologist, a few issues are evident. First, from the perspective of a driving researcher, it is obvious that the scenarios generated to test our questions are to a large extent abstract, and not representative of a realistic driving situation. It could be argued that the tasks used do represent what it is like to drive on a busy 3-lane expressway with traffic directly in front; but even given this interpretation, the similarity ends there, in that cars flash on and off, are sometimes grouped with 8 other identical cars, and tend to generally defy the rules of the road (no signaling, etc), while at other times pausing and restarting. Given, however, the cognitive-oriented nature of the goals of this research, it may be that these compromises were unavoidable.
In terms of the cognitive nature of these studies, limitations also exist. As mentioned above, somewhat less control was possible over the various factors. That is, in a standard multiple object tracking program, it is possible to easily set the speed of the objects, the trajectory of the objects, whether or not they obscure one another, and a number of other possible factors. In this case, it was necessary to rely on the possible scripting activities of the simulator software (e.g., lane changes) to mimic the standard multiple object tracking task. The end result is that many of the parameters used to implement the “multiple vehicle tracking” task were inherent to the simulator hardware and software, in contrast with the standard MOT task in which all elements are strictly under experimenter control.

8.4 Future Directions

Perhaps the most useful direction this research could take is to test “vehicle tracking” ability in more and more realistic environments. The purpose of the current work was to mimic as closely as possible the classical multiple object tracking task, and transport it into a driving simulator environment in which we could make conclusions regarding the use of realistic stimuli (i.e., cars on a roadway). Future research may be geared less towards demonstrating the level of tracking ability that is obtainable in a driving environment, and more towards learning the specific characteristics of situations which commonly arise when operating a motor vehicle. For example, I return to the classic situation of a car turning left across traffic at a green signal (i.e., an “unprotected left turn”). It would be possible to closely model this scenario in a driving simulator, in order to understand what happens to a driver’s attentional focus as there are more and more objects to attend to. The “number of targets” manipulation could be applied to the number of entities that are relevant to the driver when trying to make the turn, such as pedestrians encroaching from the sidewalk to the roadway; vehicles from the opposite direction trying to continue straight across the path of the intended turn; and cyclists using the bicycle lane at a speed intermediate to both cars and pedestrians. To further sketch this idea out, it
could be arranged that the situational relevance of the object itself could be used to denote target-hood (i.e., items that the driver needs to pay attention to), and a cueing system could be set up whereby the driver is required to press a button if he/she notices a cue that appears at one of the object locations. In such a manner, the ability of a driver to distribute attention across complex, increasingly realistic driving situations could be tested. Other examples of common driving scenarios which may be modeled include passing on an expressway; passing on a country road; navigating a parking-lot; negotiating a round-a-bout… basically any situation in which the driver needs to attend to multiple points in order to avoid collisions and maintain a safe driving environment.

The other side of the coin, of course, is the secondary (or tertiary!) activities that the driver is engaged in. Once we have a standard measure of how many objects a driver can track in a given situation (i.e., the current work, or any of the examples given in the above paragraph), it is then possible, just as I have done in the current work, to include additional tasks which may be of interest to the researcher. It is one thing to demonstrate the effects of simply steering and maintaining the speed of the subject vehicle; in our modern driving environment the number of in-vehicle devices that are available to, or required for use by the driver, is increasing almost daily. In such a fashion we could then determine the actual detriment to performance accrued by the use of a particular device, and moreover, measure it in a manner that would potentially be more telling than simply measuring how much the driver swerves in each condition, or how well he/she maintains speed. Using such a framework, vehicle tracking ability could become a useful tool in understanding the various attentional impairments that arise from the use of in-vehicles devices of all types. That is, there is the potential that very specific deficits in performance may be located, like “cognitive blind spots” that occur due to particular device- or situation-related complexities. Finally, there are also numerous technologies emerging which are intended to reduce the workload of the driver (Adaptive Cruise Control, In-Vehicle
Warning Systems), and once again the framework described above could be employed to test the efficacy of these systems.

11. Conclusion

The current set of experiments has demonstrated not only that the classical multiple object tracking can be replicated in a realistic driving environment, but that factors intrinsic to the driving environment – namely the requirement of steering the vehicle and maintaining a desired headway, have a significant negative impact on the ability of participants to track the objects in question. This has implications for multiple object tracking research in general, as it indicates a necessary caveat when claiming that the classical MOT task can be used to understand the task of driving. Specifically, while some decrement in performance may be attributed to such issues as self-motion and (apparent) 3-dimensionality of the display, the most severe deficit seems to occur when one is required to actually drive the vehicle (that is, steering and maintaining a specific headway), rather than just sitting in it passively. While object tracking in a driving environment is possible, and even quite robust, it is obvious that the simplest requirements of everyday driving experience have a significant impact on one’s ability to interpret the physical environment.

The present work has also upheld the notion that items which are selected for “object tracking”, generally enjoy benefits that non-tracked objects do not. This is a fairly subtle advantage, as visual onsets, which in nature usually accompany an interesting or dangerous event, tend to override the response-time advantage for tracked items (this was seen in Experiment 2c). However, the phenomenon remains that when an item or group of items is selected and tracked, processing functions for which full attentional focus is required (for example identifying a word, or a letter, or localizing a
change), are facilitated, as compared processing for non-tracked items. The notion of “object indexes”, as evidenced by the facilitated processing of tracked items, be they FINSTS or otherwise, has been generally upheld by the present Experiment 2.

The tracking advantage for uniquely coloured objects (Makovski and Jiang, 2009) was also replicated (Experiment 3), indicating that object features play some role in tracking. The fact that the Tracking Set Composition and Task Load factors interacted is particularly interesting. It may be considered a small irony that while using uniquely coloured items does result in a significant tracking advantage, this advantage is negated by the requirement of steering and maintaining the speed of a vehicle! That is, although the visual complexity that is inherent in any realistic environment is an aid to object tracking, this advantage must be guarded against high task-load, lest the added requirements overload the system and negate any processing benefits; even simply steering the vehicle and maintaining headway seems to reduce the effect to statistical non-significance. This has particular implications for drivers who are under more-than-usual cognitive load, and likewise elderly drivers, who are known to be more sensitive to increases in task load.

Because the inclusion of a secondary task (steering and maintaining headway) negates the tracking advantage for uniquely-coloured objects, it is theorized that this advantage is due to the action of working memory. Although Makovski and Jiang (2009) theorize that it is visual working memory per se, the category is subsumed by the broader category of working memory in general, which their task could also be said to tap. As such, and as a result of the current work, I theorize that the tracking advantage for uniquely-coloured objects is contingent upon the availability of working memory. The specific mechanism is still somewhat uncertain – a good candidate is the ability to re-locate a specific item when it is lost during tracking, though subtler operations may also be at work. As could be expected, more research will be required to answer this question.
Overall, this research has demonstrated concretely that classical multiple object tracking can be employed in a driving environment, and how performance is impacted by active driving requirements. It has also applied the methodology in order to answer some questions about how vehicles are tracked in a simulated driving environment. Event detection advantages for tracked items, and accuracy advantages for uniquely-coloured tracking sets, were among the phenomena observed. Given further development of the vehicle tracking paradigm, there is potential to answer more specific questions about how objects are tracked in realistic driving situations, and how this is affected by the specific task requirements of that environment.
References


Eriksen, C. W.; St. James, J. D. (1986). Visual attention within and around the field of focal attention:


APPENDIX A: LANE CHANGE DATA COLLECTION FORM

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
</tr>
<tr>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
</tr>
<tr>
<td>25</td>
<td>26</td>
<td>27</td>
<td>28</td>
<td>29</td>
<td>30</td>
</tr>
<tr>
<td>31</td>
<td>32</td>
<td>33</td>
<td>34</td>
<td>35</td>
<td>36</td>
</tr>
<tr>
<td>37</td>
<td>38</td>
<td>39</td>
<td>40</td>
<td>41</td>
<td>42</td>
</tr>
<tr>
<td>43</td>
<td>44</td>
<td>45</td>
<td>46</td>
<td>47</td>
<td>48</td>
</tr>
<tr>
<td>49</td>
<td>50</td>
<td>51</td>
<td>52</td>
<td>53</td>
<td>54</td>
</tr>
<tr>
<td>55</td>
<td>56</td>
<td>57</td>
<td>58</td>
<td>59</td>
<td>60</td>
</tr>
</tbody>
</table>

Mark Target Locations with an "X"
APPENDIX B: Consent Form

Study Title: TRACKING MULTIPLE VEHICLES IN A SIMULATED DRIVING ENVIRONMENT

Investigators:                       Project Director:               Student Investigator (PhD thesis)
Dr. Lana M. Trick, Ph.D             Martin Lochner, MA
Department of Psychology             Department of Psychology
University of Guelph               University of Guelph
519-824-4120 (ex. 53518)            519-824-4120 (ex. 58932)
ltrick@uoguelph.ca                  mlochner@uoguelph.ca

If you have any questions or concerns, please feel free to contact us

PURPOSE OF THE STUDY
This is a study of visual attention in a simulated driving environment. We are interested in learning about your ability to track (visually follow) multiple vehicles in a realistic 3-dimensional space. This study is intended to benefit our understanding of how a driver keeps track of multiple vehicles in the driving environment.

PROCEDURES
For this study you will be driving our Drivesafety DS-600c simulated research vehicle, which provides a realistic driving environment from the seat of a real car chassis. During the study you will see a set of 6 identical cars in front of you. Some of these cars (1, 2, 3, or 4) will flash on and off for a few seconds, indicating that these are the target cars which you are required to keep track of. The cars will then begin to switch places, and your job is to maintain attention on the target vehicle(s). After each trial the scenario will stop and you will be required to indicate on the provided forms, which cars were the target cars. The total duration of this experiment will be 60 minutes, and you will be compensated for your time with 1 credit towards your psychology course grade.

POTENTIAL RISKS AND DISCOMFORTS
In driving simulations such as this, some people experience a temporary feeling of light-headedness, dizziness, or feeling of mild nausea. However, for people who pass the initial simulator adaptation pre-screening test (administered before the beginning of the driving) this will not be very likely to occur. Nonetheless, if you begin to feel uncomfortable at any time, or if you wish to stop for any other reason, the study will be halted, and you will still receive credit for your participation.

POTENTIAL BENEFITS TO PARTICIPANT AND / OR SOCIETY
By participating in this experiment you will benefit by learning more about the strengths and limitations of your ability to track (attend to) vehicles in a realistic 3d environment. You will also gain insight regarding the process by which vehicles in the ‘critical zone’ surrounding the driver are monitored and incorporated into the driver’s awareness. You will be benefitting the scientific community by helping us better understand how drivers pay attention to multiple vehicles in the driving environment.

PAYMENT FOR PARTICIPATION
For participation in this experiment you will receive 1 credit towards your final grade in your psychology course.

CONFIDENTIALITY
Every effort will be made to ensure confidentiality of any identifying information that is obtained in connection with this study. Your data will be referenced by subject number only, and will be kept in a

88
locked office for a period of not more than 10 years, after which it will be destroyed by deletion (electronic files) and shredding of documents (paper files). Individual information will remain confidential, and data will be referred to in group-averages only.

PARTICIPATION AND WITHDRAWAL
You can choose whether to be in this study or not. If you volunteer to be in this study, you may withdraw at any time without consequences of any kind. You may exercise the option of removing your data from the study. You may also refuse to answer any questions you don’t want to answer and still remain in the study. The investigator may withdraw you from this research if circumstances arise that warrant doing so.

RIGHTS OF RESEARCH PARTICIPANTS
You may withdraw your consent at any time and discontinue participation without penalty. You are not waiving any legal claims, rights or remedies because of your participation in this research study. This study has been reviewed and received ethics clearance through the University of Guelph Research Ethics Board. If you have questions regarding your rights as a research participant, contact:

Research Ethics Officer
Telephone: (519) 824-4120, extension 56606
University of Guelph
E-mail: sauld@uoguelph.ca
437 University Centre
Fax: (519) 821-5236
Guelph, ON N1G 2W1

SIGNATURE OF RESEARCH PARTICIPANT / WITNESS
I have read the information provided for the study “Tracking Multiple Objects” as described herein. My questions have been answered to my satisfaction, and I agree to participate in this study. I have been given a copy of this form.

__________________________________
Name of Participant (please print)

__________________________________
Signature of Participant

__________________________________
Date

__________________________________
Name of Witness (please print)

__________________________________
Signature of Witness

__________________________________
Date
APPENDIX C: Subject Information Form

Screening and participant characteristics questionnaire

If you are willing to participate, we would like to find out a little about you before we start….

1. Name (please print): ___________________ ___________________
   Last               First

2. Sex: Male    Female

3. When is your birthday? ________________ (Month / Day / Year)

4. Have you ever been diagnosed with any problems with your eyes or vision? Yes No
   If yes, what type of eye or vision problems? ______________________

5. Have you ever been diagnosed with any problems with your ears or hearing? Yes No
   If yes, what type of ear or hearing problems? ______________________

6. Have you ever been diagnosed with Attention Deficit Disorder? Yes No
   If so, are you currently taking any medications? Yes No
   Which medications? ______________________

7. Is there anything else that you think may be useful for us to know about you? If so, please tell us in the space below.

__________________________________________________________________________________

__________________________________________________________________________________

__________________________________________________________________________________

__________________________________________________________________________________

__________________________________________________________________________________

__________________________________________________________________________________

__________________________________________________________________________________

__________________________________________________________________________________

–
APPENDIX D: Simulator Adaptation Syndrome Pre-Screening form

Simulator Adaptation Pre-Screening Questionnaire:

Part A: Driving Questions

At what age did you first start driving? _____

When was the last time you drove? _____ days

How many times a month do you drive on average? _____

How long do you drive on an average day? _____

How far do you drive on an average day? _____

What kind of driving do you do most often? Urban_____ Rural_____ Highway_____

Is there anything else you would like to add about your driving experience?

Part B: General Medical History Questionnaire

Do you have heart problems or have you had a heart attack? Yes No

Have you ever had a stroke, brain tumour or head trauma? Yes No

Do you experience lingering effects from stroke, brain tumour or head trauma? Yes No

Do you suffer from epileptic seizures? Yes No (If yes, please describe)

Do you have inner ear problems (vertigo)? Yes No

Do you have diabetes for which insulin is required? Yes No

Do you have problems with low blood sugar (hypoglycaemia)? Yes No

Are you currently taking any medications that make you feel extremely dizzy or nauseated? Yes No

*If the participant answered yes to any of the above questions indicate that they may be at a higher risk of problems resulting from simulator exposure and ask if they want to continue. If the participant answered yes to two or more of the above questions do not permit them to continue in the study.*
Part C: Specific Predictors

Some participants feel uneasy after participating in studies using a simulator. To help identify people who might be prone to this feeling we would like you to answer the following questions.

Do you experience migraine headaches?  
Yes  No

Do you experience claustrophobia?  
Yes  No

Do you have any history of motion sickness?  
Yes  No  
(If yes, please describe where and when)

Have you ever experienced dizziness or nausea while looking at a wide screen (e.g. Silver City or Omnimax Theatre)?  
Yes  No

Do you experience dizziness or nausea while reading in a moving car?  
Yes  No

Do you prefer to be the driver rather than the passenger, because otherwise you experience dizziness or nausea?  
Yes  No

If the participant answered yes to any of the above questions indicate that they may be at a higher risk of problems resulting from simulator exposure and ask if they want to continue. In particular, viewing a computer screen may cause eye-strain and eye-strain triggers migraines for some migraine sufferers; the confined space may be a challenge for claustrophobics; people who have experienced dizziness or nausea as a result of motion (especially recently) or while viewing wide screen movies may experience similar symptoms in a simulator. However motion sickness experienced on a boat is much more typical. We are especially worried about people who get car sickness or motion sickness on a train or those who cannot read in a moving car.
APPENDIX E: Simulator Adaptation Syndrome Post-Test form

Subject Name ___________________
Subject Code ________________
Date _______________________

Post-Experiment Questionnaire (SSQ)

There is a small risk associated with driving in the driving simulator. Some individuals experience feelings of dizziness or nausea, and an increase in body temperature, which are symptoms of a temporary condition called Simulator Adaptation Syndrome. We are tracking the severity of any discomfort felt by those who drive in the driving simulator.

1. How many times have you been in the driving simulator? (Check one)
   - First time______
   - Second Time______
   - More than two times______

2. Please rate the following symptoms of discomfort on a scale of 0 to 3, where 0 = none, 1 = slight, 2 = moderate, and 3 = severe.

<table>
<thead>
<tr>
<th>Symptom</th>
<th>Rating</th>
<th>Nausea</th>
<th>Oculomotor</th>
<th>Disorientation</th>
</tr>
</thead>
<tbody>
<tr>
<td>General Discomfort</td>
<td>______</td>
<td>___</td>
<td>___</td>
<td>___</td>
</tr>
<tr>
<td>Fatigue</td>
<td>______</td>
<td>___</td>
<td>___</td>
<td>___</td>
</tr>
<tr>
<td>Headache</td>
<td>______</td>
<td>___</td>
<td>___</td>
<td>___</td>
</tr>
<tr>
<td>Eyestrain</td>
<td>______</td>
<td>___</td>
<td>___</td>
<td>___</td>
</tr>
<tr>
<td>Difficulty Focusing</td>
<td>______</td>
<td>___</td>
<td>___</td>
<td>___</td>
</tr>
<tr>
<td>Increased Salivation</td>
<td>______</td>
<td>___</td>
<td>___</td>
<td>___</td>
</tr>
<tr>
<td>Sweating</td>
<td>______</td>
<td>___</td>
<td>___</td>
<td>___</td>
</tr>
<tr>
<td>Nausea</td>
<td>______</td>
<td>___</td>
<td>___</td>
<td>___</td>
</tr>
<tr>
<td>Difficulty Concentrating</td>
<td>______</td>
<td>___</td>
<td>___</td>
<td>___</td>
</tr>
<tr>
<td>Fullness of Head</td>
<td>______</td>
<td>___</td>
<td>___</td>
<td>___</td>
</tr>
<tr>
<td>Blurred Vision</td>
<td>______</td>
<td>___</td>
<td>___</td>
<td>___</td>
</tr>
<tr>
<td>Dizzy (Eyes Open)</td>
<td>______</td>
<td>___</td>
<td>___</td>
<td>___</td>
</tr>
<tr>
<td>Dizzy (Eyes Closed)</td>
<td>______</td>
<td>___</td>
<td>___</td>
<td>___</td>
</tr>
<tr>
<td>Vertigo</td>
<td>______</td>
<td>___</td>
<td>___</td>
<td>___</td>
</tr>
<tr>
<td>Stomach Awareness</td>
<td>______</td>
<td>___</td>
<td>___</td>
<td>___</td>
</tr>
<tr>
<td>Burping</td>
<td>______</td>
<td>___</td>
<td>___</td>
<td>___</td>
</tr>
</tbody>
</table>

For Experimenter Use Only

<table>
<thead>
<tr>
<th>Weight</th>
<th>Nausea</th>
<th>Oculomotor</th>
<th>Disorientation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Totals</td>
<td>______</td>
<td>______</td>
<td>______</td>
</tr>
</tbody>
</table>

93
APPENDIX F: Debriefing Form

Object Tracking in a Realistic 3d Environment

Thank you for participating in this study. We were interested in finding out about the ability to attend to multiple moving objects in a realistic 3d virtual environment; specifically, a high-fidelity driving simulator. What you have just completed is a version of the ‘Multiple Object Tracking’ task first devised by Z. Pylyshyn and R. Storm (1988). Unlike standard replications of this task, we have programmed the driving simulator to provide multiple moving target vehicles, in order to test your ability to keep track of the ‘target’ vehicles as they move around the screen. Multiple Object Tracking is generally understood to be important to real-world activities, where you have to keep track of a number of independent moving objects at once (e.g. pedestrians, cyclists, other motorists). Traditional object tracking studies show that when items are moving in this way among distracters, most young adults can keep track of up to 4 at a time, though there may be differences between individuals. These differences may have practical implications for day-to-day tasks. If you have problems tracking multiple items at once this may put you at risk when driving, particularly at busy intersections or when turning left across traffic. In these cases it is essential that you keep track of all the important road users (pedestrians, cyclists, other vehicles). If one is “lost” the consequences could be fatal. Similarly, the research suggesting that certain athletes have enhanced abilities to track multiple things at once, and this gives them a special advantage in sports such as hockey (the positions of all the other players, as well as the puck).

The purpose of our investigation is to determine whether the Multiple Object Tracking Paradigm, as originally devised by Pylyshyn & Storm, can be applied to object tracking in realistic environments. Studies on Multiple Object Tracking consistently report that object tracking is important for every-day activities such as driving, keeping track of children on a playground, and playing group sports, however there is no direct evidence reported for these claims – instead, object tracking is tested in a 2d desktop-computer environment. There are reasons to suspect that object tracking in a complicated 3d environment may actually be more difficult than the 2d laboratory version; if this is the case, then we should be made aware of the differences in how many objects can be tracked in realistic 3d environments versus standard 2d object tracking studies. This experiment is a follow up to previous work using longer tracking durations and animated tracking backgrounds, attempting to simulate tracking in a realistic environment. If you would like more information on this study, or if you are interested in participating in a future study for this project, please contact Martin at mlochner@uoguelph.ca. Thank you again for your participation. If you think of any questions, or if you want to find out more about the results of the study, please contact us, either by phone, or e-mail.

Dr. Lana Trick
Dept. Of Psychology
University of Guelph
Phone: 519-824-4120 (ext 53518)
E-mail: ltrick@uoguelph.ca

Martin Lochner (mlochner@uoguelph.ca)
Ryan Toxopeus (rtoxopeu@uoguelph.ca)
Blackwood Hall 119J, 519-824-4120 (ext 56848)
Appendix G: D-prime means plot

Plotting $d'$ for Target & Distractor Changes

<table>
<thead>
<tr>
<th></th>
<th>1 Target</th>
<th>3 Targets</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Targ_Driving</strong></td>
<td>3.086355114</td>
<td>3.971201501</td>
</tr>
<tr>
<td><strong>Targ_Not Driving</strong></td>
<td>2.448779964</td>
<td>2.882013776</td>
</tr>
<tr>
<td><strong>Dist_Driving</strong></td>
<td>1.957428743</td>
<td>2.699326079</td>
</tr>
<tr>
<td><strong>Dist_Not Driving</strong></td>
<td>1.823634233</td>
<td>2.170078088</td>
</tr>
</tbody>
</table>
## Appendix H: Correlations – Accuracy vs. Driving Metrics

### Experiment 1 Correlations

<table>
<thead>
<tr>
<th></th>
<th>Tracking Accuracy</th>
<th>SD Steering</th>
<th>SD Headway</th>
<th>Avg Headway</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tracking Accuracy</strong></td>
<td>Pearson Correlation</td>
<td>1.000</td>
<td>-0.239</td>
<td>-0.008</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td></td>
<td>0.063</td>
<td>0.952</td>
</tr>
<tr>
<td><strong>SD Steering</strong></td>
<td>Pearson Correlation</td>
<td>-0.239</td>
<td>1.000</td>
<td>-0.059</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>0.063</td>
<td>0.656</td>
<td>0.023</td>
</tr>
<tr>
<td><strong>SD Headway</strong></td>
<td>Pearson Correlation</td>
<td>-0.008</td>
<td>-0.059</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>0.952</td>
<td>0.656</td>
<td>0.080</td>
</tr>
<tr>
<td><strong>Avg. Headway</strong></td>
<td>Pearson Correlation</td>
<td>0.041</td>
<td>-0.297</td>
<td>0.229</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>0.755</td>
<td>0.023</td>
<td>0.080</td>
</tr>
</tbody>
</table>

### Experiment 2a Correlations

<table>
<thead>
<tr>
<th></th>
<th>Tracking Accuracy</th>
<th>SD Steering</th>
<th>SD Headway</th>
<th>Avg Headway</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tracking Accuracy</strong></td>
<td>Pearson Correlation</td>
<td>1.000</td>
<td>-0.172</td>
<td>-0.094</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td></td>
<td>0.137</td>
<td>0.416</td>
</tr>
<tr>
<td><strong>SD Steering</strong></td>
<td>Pearson Correlation</td>
<td>-0.172</td>
<td>1.000</td>
<td>0.152</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>0.137</td>
<td>0.198</td>
<td>0.123</td>
</tr>
<tr>
<td><strong>SD Headway</strong></td>
<td>Pearson Correlation</td>
<td>-0.094</td>
<td>0.152</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>0.416</td>
<td>0.198</td>
<td>0.686</td>
</tr>
<tr>
<td><strong>Avg Headway</strong></td>
<td>Pearson Correlation</td>
<td>0.199</td>
<td>-0.179</td>
<td>0.047</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>0.077</td>
<td>0.123</td>
<td>0.686</td>
</tr>
</tbody>
</table>

### Experiment 3 Correlations

<table>
<thead>
<tr>
<th></th>
<th>Tracking Accuracy</th>
<th>SD Steering</th>
<th>SD Headway</th>
<th>Avg Headway</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tracking Accuracy</strong></td>
<td>Pearson Correlation</td>
<td>1.000</td>
<td>-0.393</td>
<td>-0.312</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td></td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td><strong>SD Steering</strong></td>
<td>Pearson Correlation</td>
<td>-0.393</td>
<td>1.000</td>
<td>0.479</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>0.000</td>
<td>0.000</td>
<td>0.363</td>
</tr>
<tr>
<td><strong>SD Headway</strong></td>
<td>Pearson Correlation</td>
<td>-0.312</td>
<td>0.479</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td><strong>Avg Headway</strong></td>
<td>Pearson Correlation</td>
<td>-0.072</td>
<td>-0.077</td>
<td>0.436</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>0.393</td>
<td>0.363</td>
<td>0.000</td>
</tr>
</tbody>
</table>