

12-21-2016

Short-Term Retention of Location-Identity Bindings for Complex Objects

Alexander Nalbandian
atn7080@rit.edu

Follow this and additional works at: <http://scholarworks.rit.edu/theses>

Recommended Citation

Nalbandian, Alexander, "Short-Term Retention of Location-Identity Bindings for Complex Objects" (2016). Thesis. Rochester Institute of Technology. Accessed from

This Thesis is brought to you for free and open access by the Thesis/Dissertation Collections at RIT Scholar Works. It has been accepted for inclusion in Theses by an authorized administrator of RIT Scholar Works. For more information, please contact ritscholarworks@rit.edu.

DEPARTMENT OF PSYCHOLOGY, COLLEGE OF LIBERAL ARTS
ROCHESTER INSTITUTE OF TECHNOLOGY

Short-Term Retention of Location-Identity Bindings for
Complex Objects

by
Alexander Nalbandian

A Thesis in
Experimental Psychology

Submitted in Partial Fulfillment of the Requirements for the Degree of
Master of Science

December 21, 2016

We approve the thesis of Alexander Nalbandian:

Esa Rantanen, Ph.D.

December 21, 2016

Associate Professor, Department of Psychology

Faculty Adviser and Chair of the Thesis Committee

Andrew Herbert, Ph.D.

December 21, 2016

Professor and Chair, Department of Psychology

Reader

Lauri Oksama, Ph.D.

December 21, 2016

National Defence University Department of Behavioral Sciences and University of Turku, Finland

Reader

Acknowledgements

I would like to acknowledge and thank Alex Roughton for being an integral part in the programming of the experiment. I would like to thank Dr. Andrew Herbert and Dr. Lauri Oksama for their insight and time into this project as my committee. For his guidance and patience through my entire thesis process, I would like to thank Dr. Esa Rantanen. Finally, I would like to thank my family, Faith Andrews, and my friends for their constant support and love, because without them, this work would not be possible.

Abstract

Knowing what is where is an essential and complex component of human perception. This ability refers to the concept of Situation Awareness (SA). The underlying mechanisms for this ability show parallels to the tracking of a set of identical objects moving on a screen, a theory known as Multiple Object Tracking (MOT) (Pylyshyn & Storm, 1988). This theory was useful for investigating the fundamental factors of visual tracking, but lacked a connection to real-world scenarios. In an attempt to bridge that gap, Oksama and Hyönä (2008) created the Model of Multiple Identity Tracking (MOMIT), which includes unique identities for each object being tracked and posits a combination of peripheral and focal perception in tracking, as opposed to strictly peripheral in the MOT paradigm. This model was then applied to air traffic control (ATC) displays to create a predictive utility for analyzing controllers' performance (Hope, Rantanen, & Oksama, 2010). However, the call sign objects used in the MIT application study (Hope et al., 2010) only required the observer to remember a single letter from the object's identity, negating the need to memorize the entire identity. Using a similar structure of typical ATC call signs (6-7 character alphanumeric strings), the experiment investigated the study time duration necessary to acquire and retain the identity and location information of complex objects, and in effect, form a level 1 SA. Furthermore, the accuracy of the location-identity bindings formed in level 1 SA are also investigated for various object set sizes (4, 8, or 12 objects). The object array size of 4 had the highest accuracy. As the object set size increased from 4 objects to 12 objects, the error in location information and identity information increased. Finally, it was found that approximately 3 complex ATC-like objects were retained by participants in all conditions.

Contents

Acknowledgements	ii
Abstract	iii
Introduction	1
ATC Task and Performance	1
Cognitive Underpinnings of Controller Performance	2
Situation awareness (SA)	2
Serial and Parallel Information Processing	4
Visual Tracking	4
Multiple object tracking (MOT)	5
Serial and parallel tracking	5
Multiple identity tracking (MIT)	6
Model of Multiple identity tracking (MOMIT)	7
Visuo-Spatial Short-Term Memory (VSTM)	8
Capacity limitations	9
Subsystem internal collaboration	10
Purpose of the Research	10
Object Identity Information	11
Static vs. Dynamic Displays	12
Measuring Performance	12
Study Time	13
Performance over a Series of Trials	13
Hypotheses	13
Method	15
Participants	15
Apparatus	16

Task	16
Independent Variables	17
Dependent Variables	17
Design	17
Procedure	18
Results	19
Initial Analysis	19
Outliers	19
Accuracy vs. Reaction Time Trade Off	19
Analysis of Variance	20
Aggregation of Dependent Variables	20
ANOVA	20
Non-Linear Regression on Accuracy	21
General Non-Linear Regression	21
Fitting Non-Linear Regression	23
Location and Identity Error Analysis	23
Study Time and Strategies	24
Discussion	26
Situation Awareness	26
Serial and Parallel Tracking	28
VSTM	29
Design Limitation	29
Conclusion	30
Recommendations	30
References	32

Appendices	39
Appendix A: Figures	39
Appendix B: Tables	47
Appendix C: R Studio Code	52
Appendix D: IRB Form	58

List of Figures

1	Experimental task	40
2	Non-linear and Linear Regression for Mean Accuracy per Trial	41
3	Histogram of Accuracy per trial for 4 Objects	42
4	Histogram of Accuracy per trial for 8 Objects	43
5	Histogram of Accuracy per trial for 12 Objects	44
6	Scatterplots of Accuracy and Reaction Time in the Pilot Study	45
7	Study Time Graph	46

List of Tables

1	Object Speeds for MOT and MIT research	48
2	Object Identities for the Experiment	49
3	Location & Identity Error Table	50
4	Object Identities for Future Experiments	51

Introduction

Awareness of our surroundings is an essential part of our daily lives. Whether observing on-coming traffic while merging onto a highway or watching your favorite basketball players in a game on television, you are tracking specific features in the visual environment apart from the rest of the world. This is a fundamental task for certain jobs, such as air traffic control (ATC). In ATC, the controller must track all the aircraft in a particular airspace from a screen representation. Any confusion about the identities of aircraft on the screen could compromise the safety of the flights.

ATC Task and Performance

ATC involves complex tasks that require comprehensive awareness of a given situation. Air traffic controllers must be constantly vigilant of flights in their sector and provide instructions to pilots so they do not collide with each other. Research has shown that air traffic controllers form a mental “picture” of what is actually happening in the world from the digital information that they acquire from their displays to effectively direct air traffic (Mogford, 1997). Controllers’ displays provide relevant information, such as aircraft call sign (a unique identifier), type, altitude, speed, and other information, all of which controllers integrate into their picture. This picture, in turn, allows controllers to track flights and make control decisions about them. With the advances in automation technology (Next Generation, or NextGen technology), the role of air traffic controllers will likely shift from that of controlling flights to primarily monitoring flights in their sector (Durso & Manning, 2008; Metzger & Parasuraman, 2001). Additionally, flight deck automation, which is referred to as Free Flight, could allow pilots to have greater control over their flight plan (e.g., change course without air traffic controller approval). Therefore, the predictability of flight paths will decrease, which will weaken the strength of the air traffic controllers’ picture and their performance (Endsley, Mogford, Allendoerfer, Snyder, & Stein, 1997; Metzger & Parasuraman, 2001; Mogford, 1997). Ultimately, the critical role of air traffic controllers will remain a predominantly visual task of tracking aircrafts regardless of the future changes

in technology.

Although there are many measures of ATC performance in the human-technology system, psychological research on information processing (e.g., perception, memory, and attention) should provide a thorough understanding of ATC performance and solutions to domain-specific difficulties (Hopkin, 1980). The visual task in ATC monitoring and scanning of dynamic displays requires the understanding of “what is where” and constant updating of controllers’ knowledge about each flight (Shorrock, 2005, 2007). This process relates to the concept of situation awareness (SA) (Endsley, 1995a; Endsley, Sollenberger, & Stein, 2000), which will be discussed shortly.

Cognitive Underpinnings of Controller Performance

Gronlund, Ohrt, Dougherty, Perry, and Manning (1998) demonstrated that the importance of aircraft information to air traffic controllers rather than the frequency of interaction with the information better indicates what is retained in memory. Furthermore, this information is kept in working memory, which allows for easier access by the controller but is susceptible to error due to memory lapses or constraints (Shorrock, 2005). Kopardekar, Rhodes, Schwartz, Magyarits, and Willems (2008) found that the maximum workload capacity of air traffic in a sector to be based on the complexity of aircraft movement as opposed to the sheer number of aircraft. In low complexity conditions, the maximum number of aircraft controlled within a sector was found to be 24 for lab simulation. Niessen and Eyferth (2001) described a cognitive model of ATC (MoFi) that combined the concept of the mental picture and SA; this model includes a monitoring cycle and the learning and forgetting of all relevant information on the radar screen.

Situation awareness (SA). It is fundamental in ATC to know the positions of aircraft at all times to effectively understand air traffic flows. SA is a cognitive construct that is directly related to this point. Endsley (1995a) divided SA into three hierarchical levels: (1) perception of the elements in the environment, (2) the comprehension of the elements’ meanings, (3) the projection of the elements’ future states. The levels must be achieved

in this order. In ATC, level 1 SA is achieved through the visual information provided by the plan view displays (PVDs) and the verbal information from other air traffic controllers and pilots. This information is then integrated, and meaning is assigned to the elements in the environment (level 2 SA). Finally, the air traffic controller is able to predict future states of the elements, such as location or elevation, based on one's understanding of the environment (level 3 SA).

There are subjective (e.g., self ratings or observer reports) and objective (e.g., performance measures) ways to evaluate SA (Endsley, 1988, 1995b; Endsley et al., 2000). One key tool for measuring overall situation awareness (levels 1 to 3) is Situation Awareness Global Assessment Technique (SAGAT; Endsley et al., 2000). SAGAT works by freezing a simulation at random intervals and querying the participant on his or her understanding of a situation. While answering the queries, the participant is unable to view the display. In a similar way, Mogford (1997) measured performance of ATC trainees in an ATC simulator on their recall accuracy of aircraft information to compare the trainees' mental picture of the simulation to their SA. A portion of his analysis focused on the acquisition and recall of level 1 SA, specifically the identifiers (ATC call signs) and the location of each flight. Mogford (1997) found that ATC trainees were 55% accurate for identifiers and 86% accurate for locations of aircraft. Endsley et al. (2000) demonstrated that SAGAT measurements in an ATC simulation provided better indicators of SA than other measurement tools (i.e., subjective observer ratings and reaction time). The study compared the effectiveness of SAGAT in an ATC simulation to the tools measuring SA in an operational setting. SAGAT queries, such as recalling call sign identifiers and locations of aircraft, enabled researchers to draw similar parallels in studying level 1 SA.

SA in ATC allows for controllers to effectively comprehend and direct air traffic flows (Endsley et al., 1997; Mogford, 1997). If controllers are able to correctly perceive information on the PVDs and comprehend the meaning behind the visual elements of their displays, they can effectively plan and predict how those items interact and provide for safe air traffic. With the increased technological automation, the level 1 SA will likely be

negatively affected as air traffic controllers lose control in the system (Endsley & Kiris, 1995; Endsley et al., 1997, 2000). In a separate study, Kaber, Perry, Segall, McClernon, and Prinzl (2006) suggested that adaptive automation effectively helps controllers assess situations and improves controllers' SA for information acquisition (level 1 SA). Regardless of the future technology impact, there is great importance in understanding the acquisition and retention of perceptual components in the ATC task.

Serial and Parallel Information Processing

The serial processing and parallel processing of information are components of human cognition. Serial processing refers to conducting a single process to completion before starting the next process. This focused type of processing has a limited capacity that is based on the average processing time for a task as it relates to the overall processing time for a number of tasks (Townsend, 1990). Serial processing can provide more details on specific elements from within complex circumstances. Parallel processing refers to the ability to conduct multiple processes simultaneously. This kind of processing can have capacity limitations based on the nature of the task (Townsend, 1990). Additionally, parallel processing enables quicker analysis of information than serial processing but does not provide finite details. Both of these types of processing are integral in the visual perception of our surroundings (Treisman, 1986; Duncan & Humphreys, 1989; Bundesen, 1990; Wheeler & Treisman, 2002). More specifically, tracking of visual objects has been theorized to consist of both serial and parallel processing components (Hope, 2009; Oksama & Hyönä, 2008).

Visual Tracking

The core mechanisms of the visual tracking process have been extensively researched and theorized (Alvarez & Cavanagh, 2004; Oksama & Hyönä, 2004; Pylyshyn & Annan, 2006; Pylyshyn & Storm, 1988). Some research has applied these theories to simulations of real-world situations (Hope et al., 2010; Horowitz et al., 2007; Oksama & Hyönä, 2008), such as ATC. Before introducing the presented study, the foundation of object tracking

is discussed, starting with a review with the original multiple object tracking study by Pylyshyn (1989).

Multiple object tracking (MOT). To test how humans track identical objects, Pylyshyn and Storm (1988) pioneered a novel experimental paradigm, which is known as MOT. In the MOT experiment, the participant viewed a set of identical objects on a display and focused on the “+” in the center of the screen. The objects were momentarily highlighted to distinguish the targets (i.e., the ones to track) from the non-targets. Then, all objects moved, and the observer tracked for a duration of time while remaining focused on the “+” in the center of the screen. Finally, the objects stopped, and the observer identified the targets using a mouse. Pylyshyn and Storm determined that this type of tracking is accomplished through parallel processing of the peripheral changes while the participants focus on a stationary location. Pylyshyn and Storm developed a model that incorporated this concept of parallel processing called Fingers of Instantiation, or FINST. It represents the ability of a person to assign a visual index to a particular stimulus (the target), which isolates that stimulus while it is moving. The concept is similar to placing one’s finger on each object and moving the finger around as the target moves. Although there is a small number (4-5) of indexes or “fingers” that can stay on individual objects in the field of view, they are not influenced by any changes to the properties of the objects. Ultimately, this model posits that tracking in MOT is an automatic process and does not require focal attention on individual objects for a small number of targets.

Serial and parallel tracking. A major component of the visual tracking paradigm is the distinction between serial and parallel tracking. Pylyshyn and Storm (1988) supported the concept of parallel processing for object tracking with FINST; however, there has not been consensus on the topic. Following Pylyshyn and Storm (1988), Pylyshyn and Annan (2006) demonstrated that observers could track numerically labeled objects in parallel if given additional time to distinguish the object identities. Pylyshyn and Annan (2006) found that tracking with numeral designations was slightly poorer than with flashing objects (at 360 ms). They increased the duration of cue availability (i.e., the numeral designation)

from 360 ms to 1080 ms, and the difference in performance disappeared. d'Avossa, Shulman, Snyder, and Corbetta (2006) theorized that a limited capacity processor, associated with serial processing, is what limits selection of parallel processing for multiple moving objects. Using electrophysiological measures, Drew, Horowitz, and Vogel (2012) found that increasing the set size of tracked objects led to confusion between targets, while increasing the object speed led to losing the tracked objects in the MOT task, which supported the concept of serial processing during tracking.

Through using the simultaneous-sequential paradigm (Eriksen & Spencer, 1969) on a MOT task, Howe, Cohen, and Horowitz (2010) attempted to discern which type of model, parallel or serial, would best explain tracking performance. The simultaneous-sequential paradigm is composed of two types of tasks: the simultaneous task, in which all objects move at the same time, and the sequential task, in which only a subset of objects move at a time. A parallel model would predict that the independent movements of the stimuli would not influence performance, therefore the tracking performance would be consistent between conditions. A serial model would predict that observer accuracy would excel in the sequential condition rather than the simultaneous condition because of the isolated movements of targets. Howe et al. (2010) found that performance for the simultaneous condition was either higher than or equal to the sequential condition, which supported the parallel processing of multiple object tracking. Both processes are theorized to work in tandem when tracking objects that have unique identities.

Multiple identity tracking (MIT). The prior experiments that have been discussed all used stimuli (i.e., large indistinguishable circles) similar to the MOT task by Pylyshyn and Storm (1988). Horowitz et al. (2007) challenged this concept by adding identity to the MOT paradigm in the form of unique, nameable objects, such as cartoon animals. Participants would track a set of animal stimuli that moved and would then be hidden at prior designated locations. The participants would either respond by giving all the locations of the targets (i.e., standard condition) or a specific target (i.e., specific condition). Through a series of experiments, Horowitz et al. (2007) tested for interference in short-term memory

by increasing variability for tracking duration, visual and auditory cues, and by varying the number of unique stimuli. They found that the capacity (the number of correctly tracked identities) for the specific condition was lower than that of the standard condition. Horowitz et al. (2007) attributed their findings to the existence of two separate cognitive systems for tracking, one focused on the position information and the other on the identity information.

Model of Multiple identity tracking (MOMIT). Although visual tracking is predominantly considered a parallel process, Oksama and Hyönä (2008) theorized that serial components are involved when tracking unique identities. This theory was explained through the Model Of Multiple Identity Tracking (MOMIT), which is based on five major assumptions: (1) a serial mechanism is used to refresh the identity-location bindings for particular objects, (2) the episodic buffer has a capacity limit for simultaneously active identities and has high individual differences, (3) visuo-spatial short-term memory is used to store the location information, (4) long-term memory is used to create temporary bindings, and (5) peripheral vision is used as a parallel mechanism to continuously switch the attention to the object(s) that require refreshing.

To test these assumptions, Oksama and Hyönä (2008) used two experimental paradigms: partial report probe recognition (PRPR) and change detection (CD). The PRPR method was used to query information for a particular object's identity-location binding while minimizing the memory and verbal responses that could interfere with the tracking performance. The CD method tested the observers' ability to retain the identity-location bindings for all objects while reducing the semantic information of the identities stored in memory. The experimental tasks were similar in that the tracking was of a particular subset of moving objects and the object types were manipulated.

Unlike previous experiments (Pylyshyn & Annan, 2006; Pylyshyn & Storm, 1988), the participants were allowed to move their eyes around the display. Oksama and Hyönä (2008) found that object speed (Table 1), set size (ranging from 2 to 6 objects), and object type (familiar vs. pseudo) influenced tracking performance. As the object speed of the targets increased, tracking performance decreased (i.e., the faster the objects moved, the greater

the performance declined), and this effect was modulated by the set size of the targets. Additionally, as the set size increased, tracking performance decreased faster for pseudo-stimuli than for familiar stimuli. Based on the findings, MOMIT models participants' cognition and provides insights on the perceptual limitations of users

Visuo-Spatial Short-Term Memory (VSTM)

The “standard” model of working memory is composed of four sections: the central executive, the phonological loop, the visuo-spatial sketchpad, and the episodic buffer (Baddeley, 2001; Baddeley & Hitch, 1974). The central executive is an attentional control system that manipulates the other components (i.e., subsystems) and has a limited capacity for processing information. The phonological loop is connected to verbal short-term memory and processing, while the visuo-spatial sketchpad is connected to VSTM and processing (Baddeley & Hitch, 1974). Initially, the model did not consist of the episodic buffer but Baddeley (2001) incorporated it into the model to account for the interactions of subsystems along with long-term memory.

The purpose of VSTM is to store a representation of visual input for a short duration of time. The temporary storage allows for an attentional control unit to manipulate and comprehend the multimodal input (Baddeley & Hitch, 1974; Cowan, 2001; Luck & Vogel, 1997). Short-term representations are able to last through eye movements (Hollingworth, Richard, & Luck, 2008; Irwin, 1992) and are used for relevant, cognitive tasks. Features of visual representations, such as color, shape, or orientation, are stored in VSTM; more specifically, they are integrated and stored as objects rather than individual features (Luck & Vogel, 1997). Along with the feature representation and integration, spatial information is encoded into VSTM; spatial information tends to be connected to identifying features (Johnston & Pashler, 1990). The VSTM is used to bind the location and identity information of visual objects and store that information for recall (Oksama & Hyönä, 2008; Pinto, Howe, Cohen, & Horowitz, 2010; Wheeler & Treisman, 2002).

The VSTM is measured by one of four ways: (1) creation of a mental image (Brooks,

1967), (2) recall of visual stimuli after a delay (Oksama & Hyönä, 2008), (3) detection of change between sequential displays (Jiang, Olson, & Chun, 2000; Oksama & Hyönä, 2008), or (4) brief presentation of a stimulus in the periphery while fixating on a point. Using a change detection paradigm, Jiang et al. (2000) uncovered that VSTM encodes relational information among the visual objects and that spatial configuration is the basis for the relational information. Location information is associated with an object's relative location to other objects rather than its specific location. This finding connects to the work done by Allen, McGeorge, Pearson, and Milne (2006) that suggests the visuo-spatial sketchpad in Baddeley's model is used to store spatial and relational information of visual objects during tracking, while the central executive is used to track the objects. Oksama and Hyönä (2008) also assessed the location-identity bindings during tracking, using both change detection and recall procedures. They found, in agreement with Pylyshyn (1989), that the VSTM capacity for tracking is around 4-5 objects, regardless of object type (face or object).

Capacity limitations. Similar to the findings of Oksama and Hyönä (2008), approximately four objects have been largely considered to be the limit of VSTM, both for static (Cowan, 2001; Eng, Chen, & Jiang, 2005; Luck & Vogel, 1997) and dynamic displays (Alvarez & Cavanagh, 2004, 2005; Pylyshyn & Storm, 1988). There is, however, less agreement on the influence of the objects' complexity with regards to VSTM capacity. Luck and Vogel (1997) posited that although four objects tend to be the maximum that can be stored in VSTM, the features that can be integrated into four objects could be much greater than just four total features (up to 4 features each). The complexity of the visual object, or the information load associated with it, can influence the capacity of VSTM (Alvarez & Cavanagh, 2004; Eng et al., 2005). Alternatively, a shared reservoir of resources for location and identity information can be distributed flexibly based on the task demands (Alvarez & Franconeri, 2007; Zhang & Luck, 2008). This common resource pool can have a trade-off effect between the location and the identity information being retained (Cohen, Pinto, Howe, & Horowitz, 2011). However, repeated exposure to repeat identities allows for more accurate location information recalled (Pinto et al., 2010).

Subsystem internal collaboration. In the model of working memory from Baddeley (2000, 2001), the episodic buffer is the component that integrates information from the visuo-spatial sketchpad and the phonological loop as temporary representations. This information can be bound together with the effort from the executive component and the episodic buffer. Therefore, a verbal query of an object's identity can be translated, with help from the episodic buffer, from the phonological loop to the object features in the visuo-spatial sketchpad, or the visuo-spatial short-term memory. Using a dual-task experimental paradigm, Postle, Desposito, and Corkin (2005) found that object information is connected to verbal processing rather than the location information in working memory.

Purpose of the Research

It is critical to apply the correct research paradigms to understand the cognitive tasks involved in ATC. The experimental paradigms of MOT/MIT explore serial and parallel processing (Pylyshyn & Storm, 1988; Oksama & Hyönä, 2008) and have been theorized to most closely mimic ATC tasks (Hope, 2009). It should be noted, however, that the experimental tasks of MOT/MIT tend to have two subsets of objects (the targets and the non-targets), whereas in ATC all visual objects are potential targets. This distinction demonstrates the importance of the overall situation awareness of the participant in an ATC-like task.

Air traffic controllers must comprehend all elements in their displays before they are able to perceive greater meaning and ultimately, predict future states (Endsley, 1995a, 1995b; Mogford, 1997). The success and safety of those in flight is dependent on the controllers' SA. While a seemingly simple task, establishing Level 1 SA in ATC involves the encoding of complex identity and location information attached to visual objects. The encoding takes place in the VSTM (Johnston & Pashler, 1990; Luck & Vogel, 1997; Oksama & Hyönä, 2008; Wheeler & Treisman, 2002), combining the identity and location information together for future recall (Oksama & Hyönä, 2008; Pinto et al., 2010; Wheeler & Treisman, 2002). Although the VSTM has limits on encoding and recall of location-identity bindings (Cohen

et al., 2011; Luck & Vogel, 1997; Shorrock, 2005), research has shown that controllers are able to exceed those limitations (Kopardekar et al., 2008).

The goal of this study was to combine the concepts of SA and working memory with the MIT models to better understand the acquisition and retention of visual information of complex objects, like ATC call signs. To demonstrate the overlap of these theories, the present research used a similar testing methodology for SA on objects with complex identities that would require focal vision (i.e., a serial mechanism) to create the identity-location bindings theorized in MOMIT.

The MIT experimental paradigm that Hope (2009) augmented served as a foundation for this study. Hope (2009) and Hope et al. (2010) were the first to apply MOMIT to an ATC-like task to create a predictive utility model for ATC performance. Rather than drawings and faces, they used easily discernible aircraft call signs as the stimuli that were masked by dollar signs when queried. Additionally, there were no distractor objects, and the call signs were so small that the participants had to foveate on them for an accurate reading, which required serial processing. Hope et al. (2010) suggested that the level of SA of the participants was not at the level 3 that has been suggested by previous literature (Endsley et al., 2000), because the increase in perceivable predictability resulted in lower performance. Hope (2009) and Hope et al. (2010) posited that only a level 1 SA was accomplished in an ATC-like task for MIT, because the participants were unable to effectively predict the future state of the visual objects (i.e., level 3 SA). This study provides a novel methodology to probe the MOT/MIT experimental application to the ATC domain on the foundation of level 1 SA.

Object Identity Information. This study used a similar approach to Hope (2009) by omitting distractor objects, because ATC requires controllers to be aware of all objects on their displays, not just a select few. Rather than easily discernible call signs, this study used highly confusable ATC-like call signs that would require the participants to retain more of the identity information and provide insight into VSTM. The participants studied an array of objects, where the objects mimicked aircraft call sign structure (e.g., ABC1234)

for a short duration (See Figure 1). This study further enhanced the experimental design of achieving level 1 SA from the visual displays by completely masking the objects on the screen when queried, a similar approach to that used by Mogford (1997) and SAGAT (Endsley et al., 2000). Additionally, the target of interest was designated using an audio cue, instead of a visual cue (Hope, 2009), which demonstrated the interactions of working memory. The audio query of visual information better simulates the ATC setting, while also minimizing visual interference (Baddeley & Hitch, 1974; Postle et al., 2005).

Static vs. Dynamic Displays. The object speeds used in MOT (Alvarez & Franconeri, 2007; Pylyshyn & Storm, 1988) and MIT experiments (Hope, 2009; Horowitz et al., 2007; Oksama & Hyönä, 2008) far exceeded the object speeds for ATC (Table 1). The speeds of realistic ATC objects on displays in operational settings are found to range between 0.02 to 0.26 degrees of visual angle (Rantanen, 2010) and are vastly different from traditional MOT/MIT tracking paradigms. Additionally, Hope (2009) suggested that participants recalled the last known location of the moving objects (i.e., level 1 SA) rather than predicting the trajectory of the objects (i.e., level 3 SA). For these reasons, this study used a static display as a closer approximation to realistic ATC to analyze the creation of location-identity bindings of the target objects (Nalbandian & Rantanen, 2015).

Measuring Performance. Unlike previous work that focused on item recognition (Hope, 2009; Oksama & Hyönä, 2008; Pylyshyn & Storm, 1988), this study primarily focused on the location recall. The location that the participant recalled (i.e., “click location”) was further analyzed to determine which object was the shortest Euclidean distance to the click location and if the identity matched the queried object’s identity. If the identities matched then the error was only in the distance from the click location; if the identities did not match, the click would demonstrate an identity error along with a location error. This criterion gave a more precise representation of the participant’s level 1 SA. For the purposes of demonstrating VSTM efficiency, reaction time was measured. The quicker the participants could retrieve the location information attached to the audio cue, the more easily the location-identity information was recalled.

Study Time. For each MOT/MIT task, there was a period of time when all of the objects were stationary (Hope, 2009; Oksama & Hyönä, 2008; Pylyshyn & Storm, 1988). This inspection time period was used to designate the targets for the participant and encode the location-identity information of the stimuli. In the Oksama and Hyönä (2008) study, the motionless inspection time allotted for initial identity-location bindings was a total of 4000 ms, which included highlighting of the tracked objects. Hope et al. (2010) provided a motionless study time of $N \times 500$ ms (N is the number of objects on the screen). The presented study hypothesized that there was insufficient time allotted to acquire the location-identity information for stationary objects with complex identities, specifically similar to ATC call signs, in a small set of trials. If the level 1 SA had been sufficiently established in prior research, the participants would have been accurate and efficient from the first trial. Therefore, this study selected the study time $N \times 700$ ms (N is the number of objects on the screen), which also served as a check on previous research.

Performance over a Series of Trials. The previous studies analyzed the data in a snapshot fashion, using a small set of trials (5-24) to test each condition (Hope, 2009; Oksama & Hyönä, 2008). In an attempt to better understand the creation of location-identity bindings, this study used a larger set of trials (50) for each condition (i.e., 4, 8, and 12 objects). By repeating each condition with the same method of inquiry over numerous trials, the formation of the location-identity bindings (level 1 SA) would be assessed. Using the power law of practice (Newell & Rosenbloom, 1981), the rate of learning of location-identity bindings for each condition was evaluated based on the accuracy of the participant's recollection to the actual display location. Based on the limitation of VSTM and results of previous research, the learning curve of trials for the 4-object condition will have a lower error measure curve compared to the 8- or 12-object condition, but all conditions will exhibit a power law curve.

Hypotheses. The acquisition of the level 1 SA for objects with complex identities, such as ATC call signs, is tedious and difficult. By utilizing similar static objects with complex, confusable identities, this study examined the limitation of human performance

on location-identity bindings for such objects in various set sizes.

1.) The location error and reaction time will be lower for the 4-object condition compared to the 8- or 12-object conditions.

2.) The performance over 50 trials should mimic a power law of practice progression for each condition. The progression will decrease in overall performance as object set size increases.

3.) As the object set size increases, the correct identification of the objects (SA) will decrease. Furthermore, this will demonstrate the capacity limitation of VSTM of four objects.

Method

In this study, the acquisition and retention of visual information in the ATC task, specifically the call signs, is analyzed in a controlled setting. Traditionally, call signs are six or seven character alphanumeric strings that represent the flight company (first 3 letters) and the flight number (last 3-4 digits) associated with the aircraft. However, the participants in this study were not trained as air traffic controllers and did not have the background knowledge that comes with that training. Therefore, the traditional call signs have no greater meaning than any random letter and number combination. A pilot study was conducted using 5 participants and consisted of testing 4-8 object arrays using easily distinguishable object identities (ABC123, DEF456, etc.). The participants consistently used the strategy of reducing the identity into a single letter. Air traffic controllers are, however, required to know the entire identity of each flight to effectively control air traffic flow. To better simulate this requirement, the call sign objects in this study were made to be highly confusable by consisting of the same letters and numbers sequenced in different orders, while preserving the call sign structure (3 letters followed by 4 digits). This required the participants to know more of the object's identity than a single letter.

Participants

A total of 45 participants were recruited from the student population at Rochester Institute of Technology. The participants were recruited through the SONA participation management system, or were referred by their professors, and signed up for designated time slots. Participants received extra credit from a professor (or SONA credits) for completing the study. Most participants had normal or corrected to normal vision. One participant did not have corrective lenses on while conducting the experiment and mentioned the need for them at the end of the experiment. Because an audio clip was used to identify the target object, deaf students were excluded from the experiment. Apart from this prerequisite, there were no restrictions on who could participate in this study. One hard of hearing student participated successfully. Two participants were highly dyslexic, which

was mentioned after completing the experiment. Of the 45 participants recruited, six were removed from analysis due to physical limitations (i.e., uncorrected vision, dyslexia, or hard of hearing) and complications during the experiment (e.g., repeating conditions). Therefore, the total number of participants used in this analysis was 39.

Apparatus

A MacBook Pro (2009) laptop computer was used to conduct the experiment. The computer had a 15-inch LCD screen, with dimensions of 276 mm \times 207 mm (10.87 in \times 8.15 in), to display the experimental stimuli. Screen resolution was 1024 pixels \times 768 pixels. Using an average distance from the screen of 406 mm (16 in), the whole screen is 37.5 degrees of visual angle (DVA) horizontal \times 28.6 DVA vertical. A Microsoft desktop standard mouse was used to move the cursor. PEBL programming language was used to create the experimental visualization and to collect the data.

Task

In the pilot study, the study time duration was self-paced; the participants would press the space bar when they felt they had sufficiently memorized the array. The average inspection time was found to be 244.74 ms, however, this inspection time is limited based on the single letter identity strategy that was used by the participants. Upon reviewing the previous work in MIT (Hope, 2009; Horowitz et al., 2007; Oksama & Hyönä, 2008) that used 166 ms to 750 ms for each object and incorporating the results from this pilot test, the inspection time was selected to be 700 ms for each object on the screen.

The participants' task was to study the entire set of objects presented on the display for a certain period of time ($N \times 700$ ms), at which point the screen went blank and the participant heard a pre-recorded spoken object identity; the participant then clicked on the remembered location of the queried object as quickly and accurately as possible. The stimuli were font size 10pt DejaVu Sans and each object was 7 characters long. The resolution of the stimuli is 40 px \times 8.5 px. The stimuli was measure to be 11 mm \times 2 mm. Again

assuming an average distance of 406 mm, the stimuli is 1.52 DVA horizontal and 0.32 DVA vertical. This sequence is demonstrated in Figure 1.

Independent Variables

The set size of the objects varied between trials (4, 8, and 12 objects with unique identities). The 4- and 8-object conditions were used to replicate previous research (Alvarez & Franconeri, 2007; Cowan, 2001; Pylyshyn & Storm, 1988; Oksama & Hyönä, 2008), while the 12-object condition provided a simulation of the call sign density of typical ATC displays. Each of the object sets was displayed stationary to examine the study time required to effectively create location-identity bindings.

Dependent Variables

The primary dependent variable for this experiment was the error measurement for the location of the object in question. The location error was measured as the distance between the click location and the actual location of the queried object. The click location was recorded for further analysis for identity error by comparing the click location to the array of object locations. Additionally, the response time was recorded, which was measured from the beginning of the audio query to the mouse click on the queried object.

Design

This study was conducted as a time-series, within-subjects design on the object set size condition (3 levels). The participants completed 50 consecutive trials for each condition. By counterbalancing the order of the three conditions before the experiment (e.g., 4-8-12, 12-8-4, 8-4-12, etc.), the experiment was designed to minimize any order or fatigue effects. In the pilot study, object identities were recycled, which had a confounding effect. For example, when shown an 8 object set size condition, the participant could potentially have a better chance because of the added exposure to lower set size objects. Therefore, this study utilized different identity groups for each condition (See Table 2).

Procedure

The experimenter explained the task to the participant and answered any questions about the experiment, after which the participant then read through and signed an informed consent form. Then, the participant was situated in front of the display. The participant was reminded to study all the objects on the display and to click on the location of the object whose identity was announced from the computer speakers. Participants were also reminded to be as accurate and quick as possible to properly assess the level 1 SA. Before the experiment began, 5 practice trials were run to help the participant learn the task. The participant then completed one set of 50 trials for each condition. A break was offered after each set of trials. Once the trials were completed, the participant was debriefed and was asked two questions: if any strategies were used, and if they were satisfied by the amount of study time. The participants were then allowed to ask any additional questions and thanked for their time.

Results

Analyses were conducted using Microsoft Access and R Studio. Microsoft Access was used to organize the data and to determine the location and identity error that will be discussed later. R Studio was used for inferential statistics and creation of graphs. While using R Studio, the following libraries were used: EZ, NLS, Lattice, and GGplot. The total number of participants used in this analysis was 39.

Initial Analysis

Outliers. There was a total of 150 trials for each participant, which resulted in a total of 5,850 data records. In the initial analysis of the data, it was determined that some of the response times occurred before it was possible for the participant to distinguish the object identities. The “distinguish time” was determined by the number of alphanumeric characters it would take to differentiate an object’s unique identity. Each object identity had three letters followed by four numbers, and the first and last characters were the same for all objects. Therefore, each object identity can be differentiated by the first five characters for all conditions. The distinguish time was the average time in which the fifth character was announced for each object identity in all conditions; the distinguish time was found to be 2777 ms. Using this time as a threshold for reaction time, 33 additional data points were removed, which left a total of 5,817 data points for analysis.

Accuracy vs. Reaction Time Trade Off. A Pearson Correlation was conducted to determine an interaction between the mean reaction time and mean accuracy (error measure) per participant for each object set size. For the 4-object condition, the correlation was $r(37) = 0.16$, $p=0.33$, for the 8-object condition $r(37) = -0.03$, $p=0.88$, and for the 12-object condition $r(37) = 0.15$, $p=0.37$. A very weak correlation is shown by each condition, but none of them were statistically significant.

Analysis of Variance

Aggregation of Dependent Variables. The reaction time and accuracy were compared to the number of trials to determine any overt trends. The accuracy increased as the trial number increased, but this was also influenced by object set size (i.e., decreasing accuracy as object set size increased). The reaction time stayed consistent around 5 seconds throughout the experiment but fluctuated in variance. To further analyze the noticeable trends in each dependent variable, the data were averaged by groups of trials per participant. Using the aggregate function in R-Code (See Appendix C), the accuracy was averaged for 5 trials per trial group; there was a total of 10 blocks.

ANOVA. Using the `ezAnova` function in R, a within-subjects design ANOVA (3 object set sizes \times 10 trial groups) was conducted to discern the effect of object set size and trial groups on accuracy and reaction time. The R code for this analysis is in Appendix C.

Sphericity was tested for both ANOVA's using the Mauchly Test of Sphericity. For the accuracy measure, the object set size and the interaction of conditions was found to violate sphericity ($p < 0.001$). The trial groups and interaction of conditions was found to violate sphericity for reaction time ($p < 0.001$). Therefore, the Greenhouse-Geisser correction was used on the ANOVA results.

An interaction was also found to be significant between the object set size and trial groups for accuracy, $F(14.76, 420.76) = 2.17$ ($p = 0.015$). The accuracy measure was significantly influenced by the compounding effect of object set size and trial groups. A main effect for accuracy was found to be significant for trial groups, $F(9, 342) = 8.62$ ($p < 0.001$), along with object set size, $F(1.64, 62.3) = 281.3$ ($p < 0.001$). A Bonferroni pairwise comparison was used to determine any significant differences within the conditions and trial groups for accuracy. All object set sizes were significantly different from each other, ($p < 0.01$). For the trial groups, the first trial block was significantly different from blocks 3-10 ($p < 0.02$) but not significantly different from block 2. Additionally, the second trial block was significantly different from blocks 7-10 ($p < 0.05$).

An interaction was also found between the object set size and trial groups for reaction

time, $F(18,684) = 2.96$ ($p < 0.001$). Similar to accuracy, reaction time was significantly influenced by the compounding effect of object set size and trial groups. A main effect for reaction time was found to be significant for trial groups $F(6.07, 230.78) = 12.71$ ($p < 0.001$), along with object set size, $F(2, 76) = 27.11$ ($p < 0.001$). Using a Bonferroni pairwise comparison, all object set sizes were significantly different from each other, ($p < 0.01$). A Bonferroni pairwise comparison was also used to determine any significant difference between the trial groups. For the trial groups, the first trial block was significantly different from blocks 2-10 ($p < 0.03$). The second trial block was significantly different from blocks 5-10 ($p < 0.01$), and block 4 was significantly different from block 10 ($p < 0.03$).

The accuracy measure was significantly different between the starting 2 blocks and the final 4 blocks, potentially showing the influence of a learning effect. The reaction time was significantly different between the starting 2 blocks and the final 5 blocks. A non-linear regression was used on the aggregated accuracy to discern any learning effects for each condition.

Non-Linear Regression on Accuracy

General Non-Linear Regression. The mean accuracy was plotted per trial for each condition to further analyze the significant main effect of trials on accuracy found in the ANOVA. The trend that is shown within the accuracy data resembles a learning curve (See Figure 2). An equation that is similar to this trend is the power law equation (Newell & Rosenbloom, 1981).

$$y = ax^b + c \tag{1}$$

In the power law equation, the y typically is the reaction time for a given trial, while the x represents the trial number. The a , b , and c parameters are constants that are fitted to specific curves. The a parameter represents the measure of the first trial, the b parameter represents the slope of the curve, and the c parameter represents the asymptote of the curve (Newell & Rosenbloom, 1981).

Using Equation 1 and the NLS function in R Studio (See Appendix C), the aggregated data is the basis for a general curve to be fitted to the data. For each object set size condition, the a, b, and c parameters were calculated for Equation 1. For the 4-object conditions, the parameters [a = 179.94, b = -0.58, c = 35.11] of the generalized curve were found, and all were determined to be significant ($p < 0.0013$). For the 8-object conditions, the parameters [a = -19.13, b = 0.47, c = 290.45] of the generalized curve were found and only the c parameter was determined to be significant ($p < 0.001$). For the 12-object conditions, the parameters [a = 31.29, b = -0.65, c = 287.77] of the generalized curve were found and only the c parameter was determined to be significant ($p < 0.001$).

The nested equation in Equation 1 (Equation 2) was also fitted to the aggregated data to compare the importance of the c parameter in the equation.

$$y = ax^b \tag{2}$$

For each object set size condition, the a and b parameters were calculated for Equation 2. For the 4-object condition, the parameters [a = 199.96 and b = -0.36] of the generalized curve were found, and both were determined to be significant ($p < 0.001$). For the 8-object condition, the parameters [a = 302.41 and b = -0.13] of the generalized curve were found and both parameters were determined to be significant ($p < 0.001$). For the 12-object condition, the parameters [a = 309.12 and b = -0.018] of the generalized curve were found and only the a parameter was determined to be significant ($p < 0.001$).

The ANOVA function in R was used to compare Equation 1 and Equation 2 fits to the aggregated data. For 4 objects, Equation 1 was found to be significantly better fit than Equation 2 to the data, $F(2,47)=5.08$ ($p=0.029$). For 8 objects, Equation 1 was found to be marginally better fit than Equation 2 to the data, $F(2,47)=3.36$ ($p=0.073$). For 12 objects, Equation 1 was found no to be a better fit than Equation 2 to the data, $F(2,47)=0.24$, ($p=0.62$). Therefore, Equation 1 was fitted on the Participant's data, and Figure 2 shows Equation 1 fitted to the data with the estimated parameters.

Fitting Non-Linear Regression. Using the `nls.List` function in R, Equation 1 was fitted to each individuals' data per object set size condition. This equation was poorly fitted to the data, because less than half of the data could produce parameter measurements. However, Equation 2 with the `nls.List` function was fitted to the participants' data and produced parameter estimates for each participant. The fit of the equation was tested using the Non-Linear Mixed Effect (`nlme`) function in R to determine the group parameter estimates, correlation between parameters, and goodness of fit for the model using Akaike's information criterion (AIC). AIC is used to compare models, in this case Equations 1 and 2, on their fit of the data. The lower the AIC, the better the fit of the model.

For 4 objects, the parameter estimates were found to be $a = 205.32$ ($p < 0.01$) and $b = -0.32$ ($p < 0.01$). There is a strong negative correlation between a and b ($r = -0.82$), and goodness of fit for the model was calculated (AIC = 23261.17). For 8 objects, the parameter estimates were found to be $a = 308.07$ ($p < 0.01$) and $b = -0.14$ ($p < 0.01$). There is a strong negative correlation between a and b ($r = -0.76$), and goodness of fit for the model was calculated (AIC = 25425.02). For 12 objects, the parameter estimates were found to be $a = 326.49$ ($p < 0.01$) and $b = -0.033$ ($p = 0.13$). There is a strong negative correlation between a and b ($r = -0.92$), and goodness of fit for the model was calculated (AIC = 26022.93).

Location and Identity Error Analysis

By creating a relationship database for the data (e.g., the dependent variables to independent variables and click location to object location) in Microsoft Access, the data were organized to complete the location and identity error analysis. To accomplish this analysis, the data were filtered to determine the intended object based on click location. The click location coordinates for each trial were compared to the coordinates of all the objects for that trial using Equation 3.

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (3)$$

The object with the shortest distance away from the click location was considered the intended target. By comparing the intended target with the queried object, the identity error could be determined (i.e., match or no match). If the intended target matched the queried object, the distance between those two points was considered the location error.

The purpose of this analysis is to measure the participants' ability to accurately recall the location and identity of ATC call signs. By measuring the click location to the queried object location, an objective measure of accuracy is created to measure the participants' SA of the objects. The mean location error in Table 3 shows the distance, in Degrees of Visual Angle (DVA), of the participants' recall location to the queried object when they correctly recalled the object. The DVA was calculated from the measured pixel distance using an average viewing distance of 406 mm (16 inches), however, participants were free to move their heads.

The correct identification also demonstrated the SA of the participant through the overall number of objects retained (the percentage correctly identified). Using the percentage of correctly identified objects and the number of objects, it was calculated that approximately 3 objects were retained in memory during the experiment for each condition.

Using the ANOVA function in R, a repeated measure one way ANOVA was calculated on the percent of correct identification of location identity bindings on object set size. Sphericity was found to be violated, therefore the degrees of freedom were corrected using the Greenhouse-Gressier correction. The percentage of correct identification was significantly affected by the object set sizes, $F(1.58,60.03)=306.74, p<0.001$.

Study Time and Strategies

At the end of the experiment, participants were debriefed and then provided feedback on their experience. On the topic of study time, participants largely felt that they had sufficient time to memorize the 4-object condition but insufficient time for the 8- and 12-object conditions. Additionally, several strategies to memorize the location and identity of objects prevailed among participants' responses. Chunking of the object identities made it

easier to memorize the objects. Some participants would memorize the spatial location and then the identities; using this method, the participants would create a spatial pattern of the objects or partitioned the screen into halves. Finally, some participants would use simple trial and error, a non-explicit feedback loop as the objects would reappear after their click, to learn the identity and location information.

Discussion

This study explored a connection between SA, VSTM, and the theory of location-identity bindings using MOMIT in an ATC-like task. The following findings were observed: 1) the Level 1 SA of participants significantly decreased as the object set size increased, 2) object set size was significant for the location and identity error, and 3) the learning effect was significant for the 4-object set size condition, minimally significant for 8 objects and absent for 12 objects set sizes.

This different experimental paradigm supported findings from prior research and the core principles of MOMIT. The influence of the object set size was determined by running an ANOVA on the accuracy and reaction time data for each trial group. The acquisition and retention of the location-identity bindings were found to be significantly influenced by the number of objects and the number of trials.

Situation Awareness

Similar to previous research (Mogford, 1997), the participants' SA was ascertained through information recall from a visual display. This approach simulated the effort of air traffic controllers to create and maintain a mental picture. While controllers must achieve all levels of SA to effectively perform their jobs, this study was focused exclusively on participants achieving level 1 SA. From the Hope et al. (2010) study, it was concluded that participants were unable to go beyond level 1 SA during tracking of call sign objects.

Given that the object identities had no greater meaning than the characters on the screen, the strength of the participants' SA was measured objectively in this study. Overall, the performance results were poor, which seems to indicate there was inadequate study time to achieve good level 1 SA. The 4-object condition showed to have the best performance (i.e., correct identification in Table 3) with the lowest response time. As object set size increased, performance decreased and response time increased. These results were expected, but it was striking that even for the 4-object condition, participants were misidentifying object identities approximately 10% of the time.

To understand the creation of location-identity bindings, and in effect SA, the performance was analyzed over a series of trials. This study used the accuracy measure to understand the practice effect on the location-identity bindings. The a, b, and c parameters in the equations were estimated by fitting the equations to the individuals' data. Due to the poor fit of the data to Equation 1, the analysis focused on the a and b parameters from Equation 2. Typically, learning curves are a function of reaction time; it is expected that reaction time should decrease over time for each continuous unit of activity performed. Reaction time alone is not the emphasis of the ATC task and therefore was not the focus of this study. Nevertheless, it was expected that participants would require less time for each query over the course of the trials, which was not seen in the results. Rather, the mean reaction time increased throughout the trials, particularly for the 8- and 12-object conditions. It is likely that participants were exhausted, frustrated or were attempting to be more deliberate with their selections. There is some degree of learning evident, however, as seen by the improved accuracy over the course of the trials.

Over the course of the trials, performance notably improved for the 4-object condition, as anticipated. The accuracy measure for the first trial, parameter a, was estimated to be 205.32 ($p < 0.01$) for the 4-object condition. The b parameter, which indicates the rate of learning, was the smallest for the 4-object condition ($b = -0.32$, $p < 0.01$), which was expected and demonstrated a learning effect. Compared to the 4-object condition, the 8-object condition had a higher error measure for the first trial ($a = 308.07$, $p < 0.01$) and lower learning rate ($b = 0.14$, $p < 0.01$). However, the learning effect was present, though minimally, for the 8-object condition. The 12-object condition had the highest starting error rate ($a = 326.49$, $p < 0.01$), but the learning effect was not present ($b = -0.033$, $p = 0.13$). It was expected that the 8- and 12-object conditions would have curves that were higher in error than the 4-object condition, however, all conditions were expected to show learning effects. A possible confounding variable that led to these results was the difficulty between the levels of object set sizes, which will be discussed in design limitations.

Serial and Parallel Tracking

By drastically increasing the difficulty of the object identities, the participants were required to commit more components of the identities to memory instead of a single letter or number (Hope, 2009; Hope et al., 2010). During the debriefing post-experiment, participants observed they had sufficient study time for the 4-object condition but inadequate time for the larger set sizes. This feedback aligns with the learning curve observations previously discussed. Contrary to what was assumed in previous work (Hope et al., 2010; Oksama & Hyönä, 2008), this experiment showed that study time and the number of objects do not have a linear relationship when acquiring location-identity bindings. If this assumption was correct, then the trend of performance for 8 and 12 objects should have mirrored that seen with the 4-object condition. Either more inspection time needs to be given when object set size exceeds 4 objects, or the process that ATC more closely resembles is that of a visual search task and less of a visual tracking task. In other words, the controllers are not committing all the information (object identity and location) to memory, but just searching for the object in question (Hope et al., 2010).

This experimental paradigm (See Figure 1) provided deeper insight into the MOMIT theory about the creation of location-identity bindings in VSTM. The capacity limit of VSTM was first theorized to be 7 single feature items (Miller, 1956). However, when multiple features are introduced, that capacity decreases. The capacity limitation on object and identity acquisition and recall was affirmed to be 4-5 items in various studies (Baddeley, 2001; Cowan, 2001; Oksama & Hyönä, 2008; Pylyshyn & Storm, 1988). In this study, the average number of correctly recalled location-identity bindings was approximately 3 objects, which is likely attributed to the increase in complexity of the object identities. This study also provided an average area (location error) where participants correctly recalled the objects' identity (identity error). The participants were able to effectively recall a large portion of the objects' information without any reference and within an area of approximately 50-60 pixels (or 1.9 to 2.2 DVA based on a 16 inch viewing distance).

VSTM

Kopardekar et al. (2008) demonstrated that the maximum manageable air traffic that a controller can handle was 24 objects, albeit under the simplest conditions. The largest set size in this experiment was only 12 objects, and the performance on this condition was comparatively very poor. There is still a disconnect between what is being theorized as the mental model of controllers and the ability of the controllers. Even without object movement, the recollection of location and identities for objects in set sizes greater than 4 is no greater than approximately 3 objects. This finding seems to support prior research (Luck & Vogel, 1997; Vogel, Woodman, & Luck, 2001), which observed that visual working memory capacity depends on the number of entire objects rather than the number of different features in a display.

Additionally, the participants' working memory was tested by forcing memory recall of the objects' identities and locations from a blank screen as opposed to recognition of masked objects, as tested in previous works ((Hope et al., 2010; Horowitz et al., 2007; Oksama & Hyönä, 2008). During the debriefing post-experiment, participants described using spatial patterns of the objects, chunking the objects' identities (Cowan, 2001), and splitting the screen in half to determine where objects were positioned (Alvarez & Cavanagh, 2005).

Design Limitation

This experiment had several design limitations that restricted the generalizability of the study. First, the use of naïve college students as test subjects creates certain limitations given their limited exposure to the ATC task (i.e., encoding call sign structure). The identities were meaningless to the participants, which provided for an objective study of SA, but using participants more familiar with the ATC task would have likely resulted in higher performance.

Second, the object identities were increased to a difficulty that poorly resembled an air traffic controller's task. ATC call signs are made of a unique alphanumeric string, but it generally follows a conventional naming pattern, which ultimately communicates

key information to the controllers. As indicated by Gronlund et al. (1998), people more effectively recall information when there is meaning associated to it. The identities used in this study were especially difficult to remember as the object set size increased, given their highly confusable nature. In other words, the increase in difficulty between the 4- to 8-object set size was not equivalent to the increase from the 8- to 12-object set size. When considering such identities for future study, the confusability and meaning for ATC call signs should be equivalent between conditions to control this confounding effect.

Finally, there was no static reference point for the participants to gauge the location. In past research, the participant was given a point of reference, either at the center of the screen (Pylyshyn, 1989) or at the location of the masked objects (Hope et al., 2010; Horowitz et al., 2007; Oksama & Hyönä, 2008). Without a static reference point like a crosshair at the center of the screen, the participants created their own reference point using the mouse pointer. This step may have led to some error in click location, because the only reference point that was given to the participants was a moving point.

Conclusion

Ultimately, this study helped further the understanding of the location-identity binding retention limitations, using a multiple identity tracking experimental paradigm. It was found that level 1 SA can be precisely measured using a method that directly shows the error of recall location and in effect, identity information indirectly. The capacity limitation of visual objects, regardless of object set size, was consistent with prior research. When using complex object identities in further research, the study time should be further contemplated, as this study showed from the inconsistent learning between object set sizes that study time and identity complexity are not linearly correlated.

Recommendations

While this study attempted to more closely simulate the ATC task using confusable object identities, future research could examine various call sign identities. One example,

Table 4, would provide an equivalent increase in difficulty in each condition and still require most of the object identity to be acquired and retained. Additionally, this method better represents the ATC task while still requiring participants to remember entire identities. An alternative approach would be to compare highly confusable object identities with confusable realistic call signs (e.g., DAL1324 and DAL1234 for Delta Air Lines). This design could provide insight into the importance of the information on memory recall, which would also effectively compare level 1 and level 2 SA.

Another key component that warrants deeper analysis is the connection between study time and the number of objects. Contrary to what was posited in previous research (Cowan, 2001), study time was not a constant for the number of objects. Using multiple study times would provide insight into the relationship between set size and study time, which in this study was shown to be non-linear. If it were the case that location-identity bindings were created with a linear model, then the acquisition and retention for each condition should have had the similar learning curve trend.

References

- Allen, R., & McGeorge, P. (2004). Attention and expertise in multiple target tracking. *Applied Cognitive ...*, 18, 337–347. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1002/acp.975/abstract>
- Allen, R., McGeorge, P., Pearson, D. G., & Milne, A. (2006, June). Multiple-target tracking: A role for working memory? *Quarterly journal of experimental psychology (2006)*, 59(6), 1101–16. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/16885145> doi: 10.1080/02724980543000097
- Alvarez, G. A., & Cavanagh, P. (2004, February). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological science*, 15(2), 106–11. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/14738517>
- Alvarez, G. A., & Cavanagh, P. (2005, August). Independent resources for attentional tracking in the left and right visual hemifields. *Psychological science*, 16(8), 637–43. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/16102067> doi: 10.1111/j.1467-9280.2005.01587.x
- Alvarez, G. A., & Franconeri, S. L. (2007). How many objects can you track?: Evidence for a resource-limited attentive tracking mechanism. *Journal of Vision*, 7, 1–10. Retrieved from <http://www.journalofvision.org/content/7/13/14.short> doi: 10.1167/7.13.14.Introduction
- Baddeley, A. D. (2000, November). The episodic buffer: a new component of working memory? *Trends in cognitive sciences*, 4(11), 417–423. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11058819>
- Baddeley, A. D. (2001). Is working memory still working? *American psychologist*, 56(11), 851–864.
- Baddeley, A. D., & Hitch, G. J. (1974). Working memory. in g. a. bower (ed.). *The psychology of learning and motivation: Advances in research and theory*, 8, 47–89.
- Botterill, K., Allen, R., & McGeorge, P. (2011). Multiple-object tracking: the binding of spatial location and featural identity. *Experimental psychology*, 58(3), 196.

- Brooks, L. R. (1967). The suppression of visualization by reading. *Quarterly journal of experimental psychology*, *19*, 289-299.
- Bundesen, C. (1990, October). A theory of visual attention. *Psychological review*, *97*(4), 523-47. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/21146554>
- Cohen, M. a., Pinto, Y., Howe, P. D. L., & Horowitz, T. S. (2011, July). The what-where trade-off in multiple-identity tracking. *Attention, perception & psychophysics*, *73*(5), 1422-34. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/21380611> doi: 10.3758/s13414-011-0089-7
- Cowan, N. (2001, February). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *The Behavioral and brain sciences*, *24*(1), 87-114; discussion 114-85. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11515286>
- d'Avossa, G., Shulman, G. L., Snyder, A. Z., & Corbetta, M. (2006). Attentional selection of moving objects by a serial process. *Vision research*, *46*(20), 3403-3412.
- Drew, T., Horowitz, T. S., & Vogel, E. K. (2012). Swapping or dropping? electrophysiological measures of difficulty during multiple object tracking. *Cognition*, *126*(2), 213-22.
- Duncan, J., & Humphreys, G. W. (1989, July). Visual search and stimulus similarity. *Psychological review*, *96*(3), 433-58. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/2756067>
- Durso, F. T., & Manning, C. A. (2008). Air traffic control. *Reviews of human factors and ergonomics*, *4*(1), 195-244.
- Endsley, M. R. (1988). Design and evaluation for situation awareness enhancement. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 32, pp. 97-101).
- Endsley, M. R. (1995a). Toward a theory of situation awareness in dynamic systems. *Human factors: The journal of the human factors and ergonomics society*, *37*(1), 32-64.
- Endsley, M. R. (1995b). Measurement of situation awareness in dynamic systems. *Human factors: The journal of the human factors and ergonomics society*, *37*(1), 65-84.

- Endsley, M. R., & Kiris, E. O. (1995, June). The Out-of-the-Loop performance problem and level of control in automation. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *37*(2), 381–394. Retrieved from <http://openurl.ingenta.com/content/xref?genre=article&issn=0018-7208&volume=37&issue=2&spage=381> doi: 10.1518/001872095779064555
- Endsley, M. R., Mogford, R., Allendoerfer, K., Snyder, M., & Stein, E. (1997). Effect of free flight conditions on controller performance, workload and situation awareness: A preliminary investigation of changes in locus of control using existing technology. *Atlantic City, NJ: Federal Aviation Administration William J Hughes Technical Center*.
- Endsley, M. R., Sollenberger, R., & Stein, E. (2000). Situation awareness: A comparison of measures. *Proceedings of the Human Performance, Situation Awareness and Automation: User-Centered Design for the New Millennium*, 15–19.
- Eng, H. Y., Chen, D., & Jiang, Y. (2005). Visual working memory for simple and complex visual stimuli. *Psychonomic Bulletin & Review*, *12*(6), 1127-1133.
- Eriksen, C. W., & Spencer, T. (1969). Rate of information processing in visual perception: some results and methodological considerations. *Journal of Experimental Psychology*, *79*, 1-16.
- Gronlund, S. D., Ohrt, D. D., Dougherty, M., Perry, J. L., & Manning, C. A. (1998). Role of memory in air traffic control. *Journal of experimental psychology: Applied*, *4*, 263-280.
- Hollingworth, A., Richard, A. M., & Luck, S. J. (2008). Understanding the function of visual short-term memory: transsaccadic memory, object correspondence, and gaze correction. *Journal of experimental psychology: General*, *137*(1), 163-181.
- Hope, R. M. (2009). *The predictive utility of the model of multiple identity tracking in air traffic control performance* (Master's thesis). Retrieved from <https://ritdml.rit.edu/handle/1850/10642>
- Hope, R. M., Rantanen, E. M., & Oksama, L. (2010, September). Multiple identity tracking and entropy in an ATC-like task. *Proceedings of the Human Factors and Ergonomics*

- Society Annual Meeting*, 54(13), 1012–1016. Retrieved from <http://pro.sagepub.com/lookup/doi/10.1177/154193121005401303> doi: 10.1177/154193121005401303
- Hopkin, V. D. (1980). The measurement of the air traffic controller. *Human factors: The journal of the human factors and ergonomics society*, 22(5), 547-560.
- Horowitz, T., Klieger, S., Fencsik, D. E., Yang, K. K., Alvarez, G. A., & Wolfe, J. M. (2007). Tracking unique objects. *Attention, Perception, & ...*, 69(2), 172–184. Retrieved from <http://www.springerlink.com/index/T837N7601PU67813.pdf>
- Howe, P. D. L., Cohen, M. A., & Horowitz, T. S. (2010). The more often you see an object, the easier it becomes to track it. , 10, 1–15. doi: 10.1167/10.10.4.Introduction
- Irwin, D. E. (1992). Memory for position and identity across eye movements. *Journal of experimental psychology: Learning, memory, and cognition*, 18(2), 307-317.
- Jiang, Y., Olson, I. R., & Chun, M. M. (2000, May). Organization of visual short-term memory. *Journal of experimental psychology. Learning, memory, and cognition*, 26(3), 683–702. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10855426>
- Johnston, J. C., & Pashler, H. (1990). Close binding of identity and location in visual feature perception. *Journal of Experimental Psychology: Human Perception and Performance*, 16(4), 843-856.
- Kaber, D., Perry, C., Segall, N., McClernon, C., & Prinzl, L. (2006). Situation awareness implications of adaptive automation for information processing in an air traffic control-related task. *International Journal of ...*, 36, 447–462. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0169814106000229> doi: 10.1016/j.ergon.2006.01.008
- Kopardekar, P., Rhodes, J., Schwartz, A., Magyarits, S., & Willems, B. (2008). Relationship of maximum manageable air traffic control complexity and sector capacity. In *26th international congress of the aeronautical sciences (ICAS 2008)*, 15-19.
- Luck, S. J., & Vogel, E. K. (1997, November). The capacity of visual working memory for features and conjunctions. *Nature*, 390(6657), 279–81. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9384378> doi: 10.1038/36846

- Metzger, U., & Parasuraman, R. (2001). The role of the air traffic controller in future air traffic management: An empirical study of active control versus passive monitoring. *Human factors: The journal of the human factors and ergonomics society*, *43*(4), 519-528.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological review*, *63*(2), 81-97.
- Mogford, R. H. (1997). Mental models and situation awareness in air traffic control. *The international journal of aviation psychology*, *7*(4), 331-341.
- Nalbandian, A., & Rantanen, E. M. (2015). Learning of location-identity bindings: Development of level 1 situation awareness in an air traffic control-like task. In *Proceedings of the 18th international symposium on aviation psychology* (pp. 159–164). Dayton, OH: ISAP.
- Newell, A., & Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the law of practice. *Cognitive skills and their acquisition*, *1*, 1-55.
- Niessen, C., & Eyferth, K. (2001). A model of the air traffic controller's picture. *Safety Science*, *37*, 187–202. Retrieved from <http://www.sciencedirect.com/science/article/pii/S092575350000485>
- Oksama, L., & Hyönä, J. (2004). Is multiple object tracking carried out automatically by an early vision mechanism independent of higher order cognition? An individual difference approach. *Visual Cognition*, *11*(5), 631–671. Retrieved from <http://www.tandfonline.com/doi/abs/10.1080/13506280344000473>
- Oksama, L., & Hyönä, J. (2008, June). Dynamic binding of identity and location information: a serial model of multiple identity tracking. *Cognitive psychology*, *56*(4), 237–83. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/17451667> doi: 10.1016/j.cogpsych.2007.03.001
- Pinto, Y., Howe, P. D., Cohen, M. A., & Horowitz, T. S. (2010). The more often you see an object, the easier it becomes to track it. *Journal of vision*, *10*(10), 1-15.
- Postle, B. R., Desposito, M., & Corkin, S. (2005, March). Effects of verbal and nonver-

- bal interference on spatial and object visual working memory. *Memory & cognition*, 33(2), 203–12. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1262675&tool=pmcentrez&rendertype=abstract>
- Pylyshyn, Z. W. (1989, June). The role of location indexes in spatial perception: a sketch of the FINST spatial-index model. *Cognition*, 32(1), 65–97. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/2752706>
- Pylyshyn, Z. W. (2004). Some puzzling findings in multiple object tracking: I. tracking without keeping track of object identities. *Visual cognition*, 11(7), 801–822.
- Pylyshyn, Z. W., & Annan, V. (2006, January). Dynamics of target selection in multiple object tracking (MOT). *Spatial vision*, 19(6), 485–504. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/17278524>
- Pylyshyn, Z. W., & Storm, R. W. (1988, January). Tracking multiple independent targets: evidence for a parallel tracking mechanism. *Spatial vision*, 3(3), 179–97. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/3153671>
- Rantanen, E. (2010). *On aircraft symbol velocities on different plan view displays*.
- Shorrock, S. T. (2005, October). Errors of memory in air traffic control. *Safety Science*, 43(8), 571–588. Retrieved from <http://linkinghub.elsevier.com/retrieve/pii/S0925753505000664> doi: 10.1016/j.ssci.2005.04.001
- Shorrock, S. T. (2007). Errors of perception in air traffic control. *Safety science*, 45(8), 890-904. (Errors of perception in air traffic control)
- Townsend, J. T. (1990). Serial vs. parallel processing: Sometimes they look like tweedledum and tweedledee but they can (and should) be distinguished. *Psychological Science*, 1(1), 46-54.
- Treisman, A. (1986). Features and objects in visual processing. *Scientific American*, 255(5), 114–125.
- Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 27(1), 92-114.

- Wheeler, M. E., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General*, *131*(1), 48–64. Retrieved from <http://doi.apa.org/getdoi.cfm?doi=10.1037/0096-3445.131.1.48> doi: 10.1037//0096-3445.131.1.48
- Zhang, W., & Luck, S. J. (2008). Discrete, fixed-resolution representations in visual working memory. *Nature*, *453*, 233-235.

Appendices

Appendix A: Figures

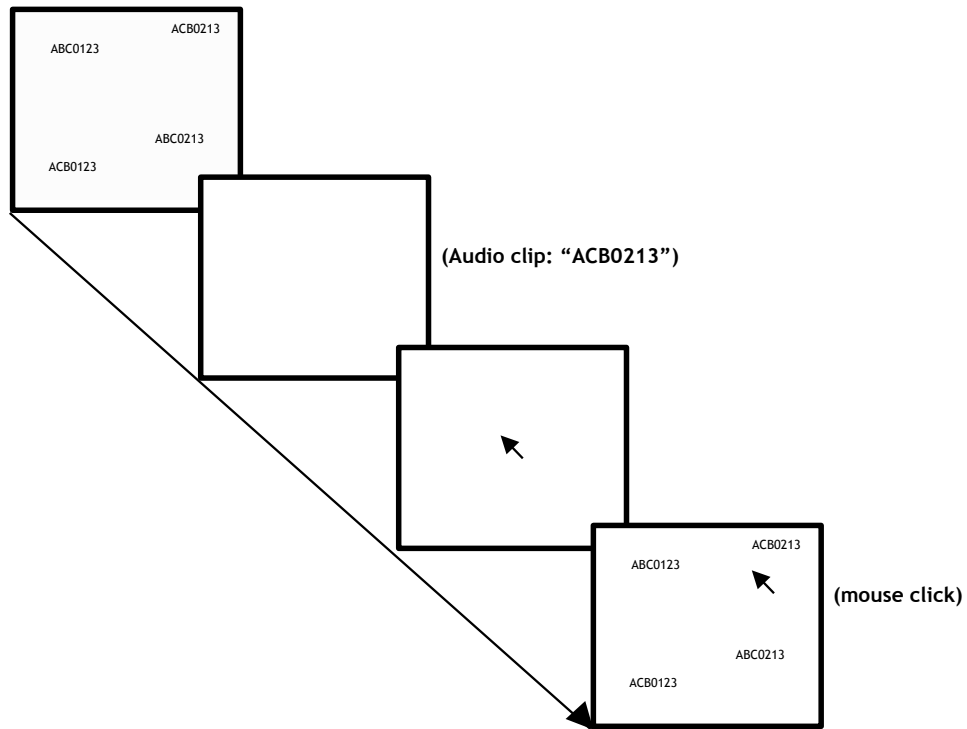


Figure 1. First, the participant tracked the set of identities. After a period of studying, the whole display went blank, and an audio clip of a particular identity played. The participant then moved the cursor to the location of the identity in question. When the participant clicked on the location, all the identities were shown in their respective locations.

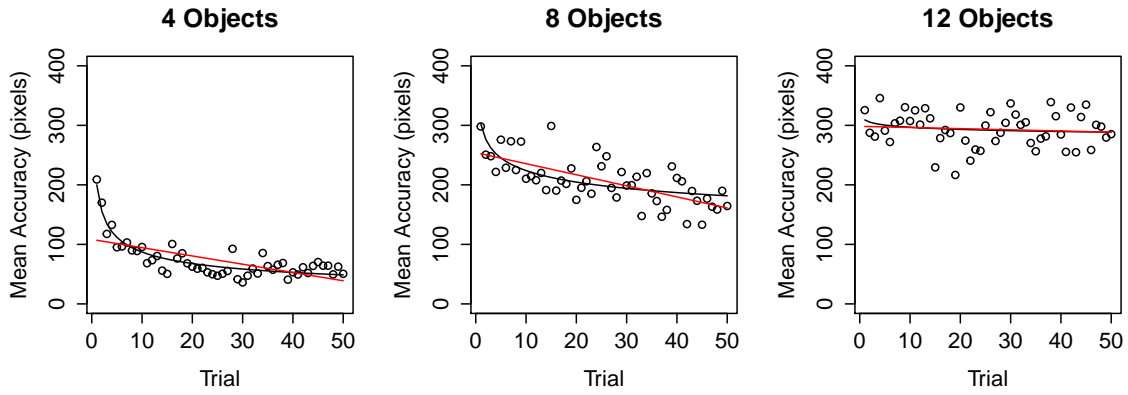


Figure 2. Non-linear and Linear Regression for Mean Accuracy per Trial. The red line represents the linear regression fitted to the data. The black line represents the non-linear regression fitted to the data.

4 Objects

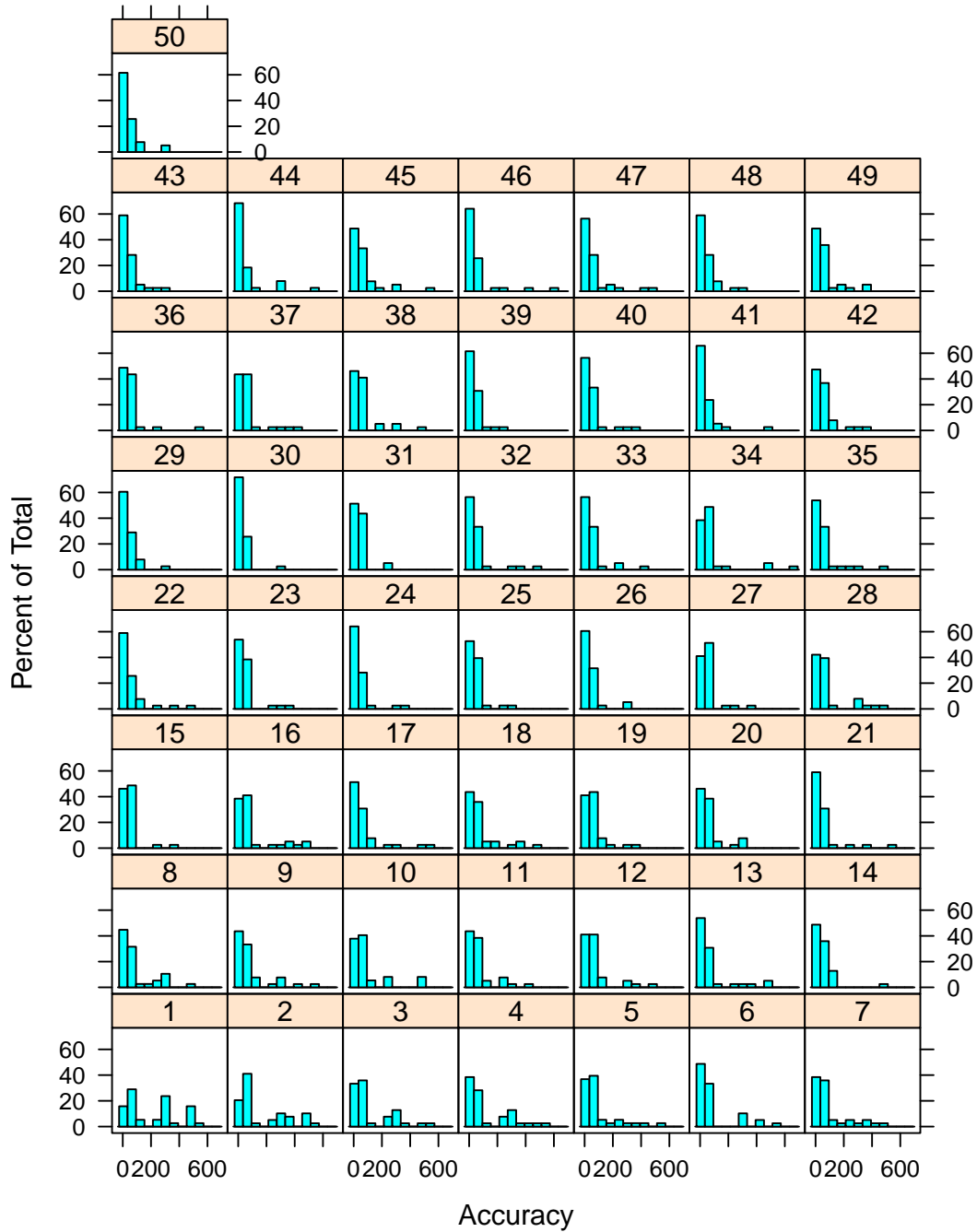


Figure 3. Histogram of Accuracy per trial for 4 Objects for each Participant.

8 Objects

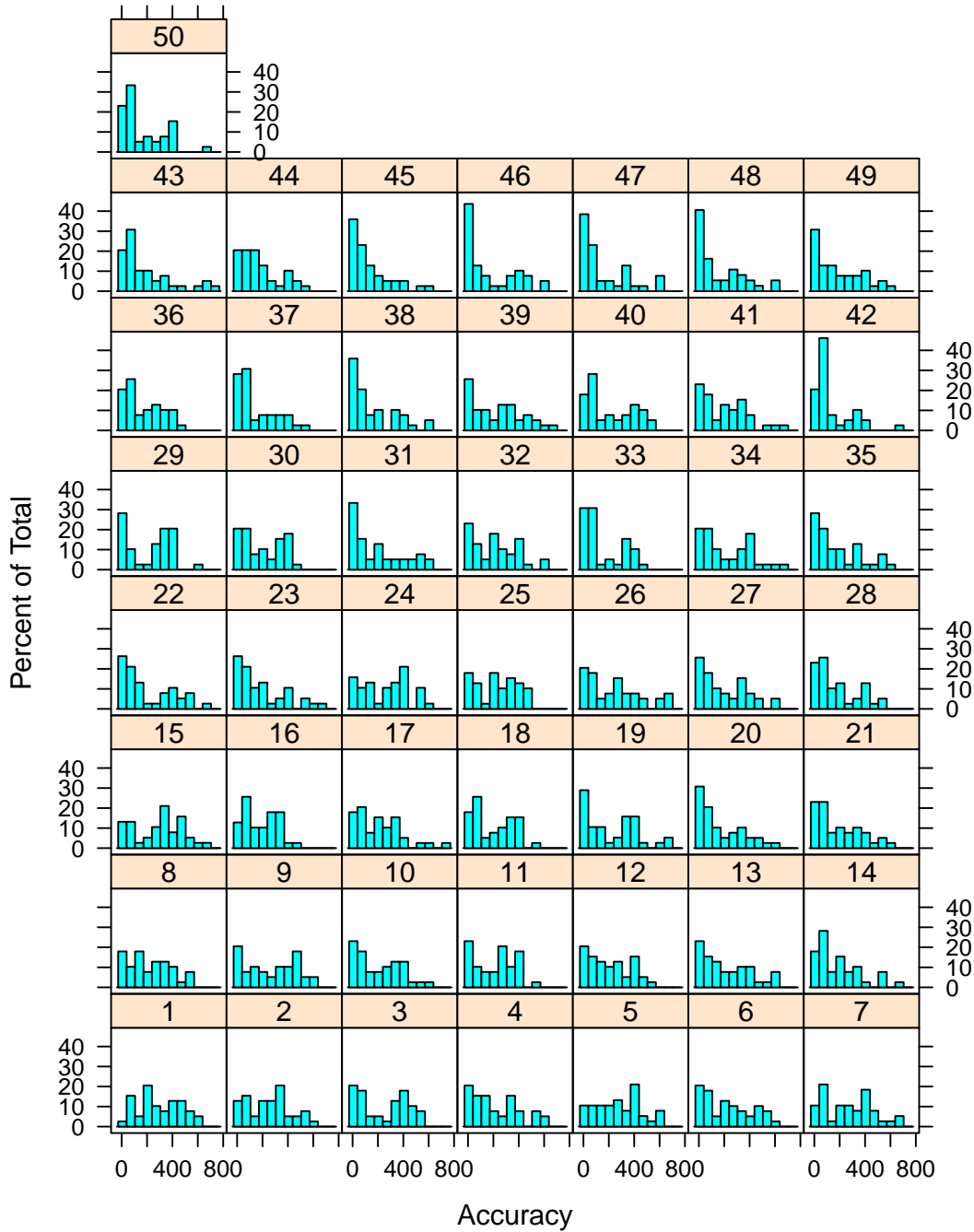


Figure 4. Histogram of Accuracy per trial for 8 Objects for each Participant.

12 Objects

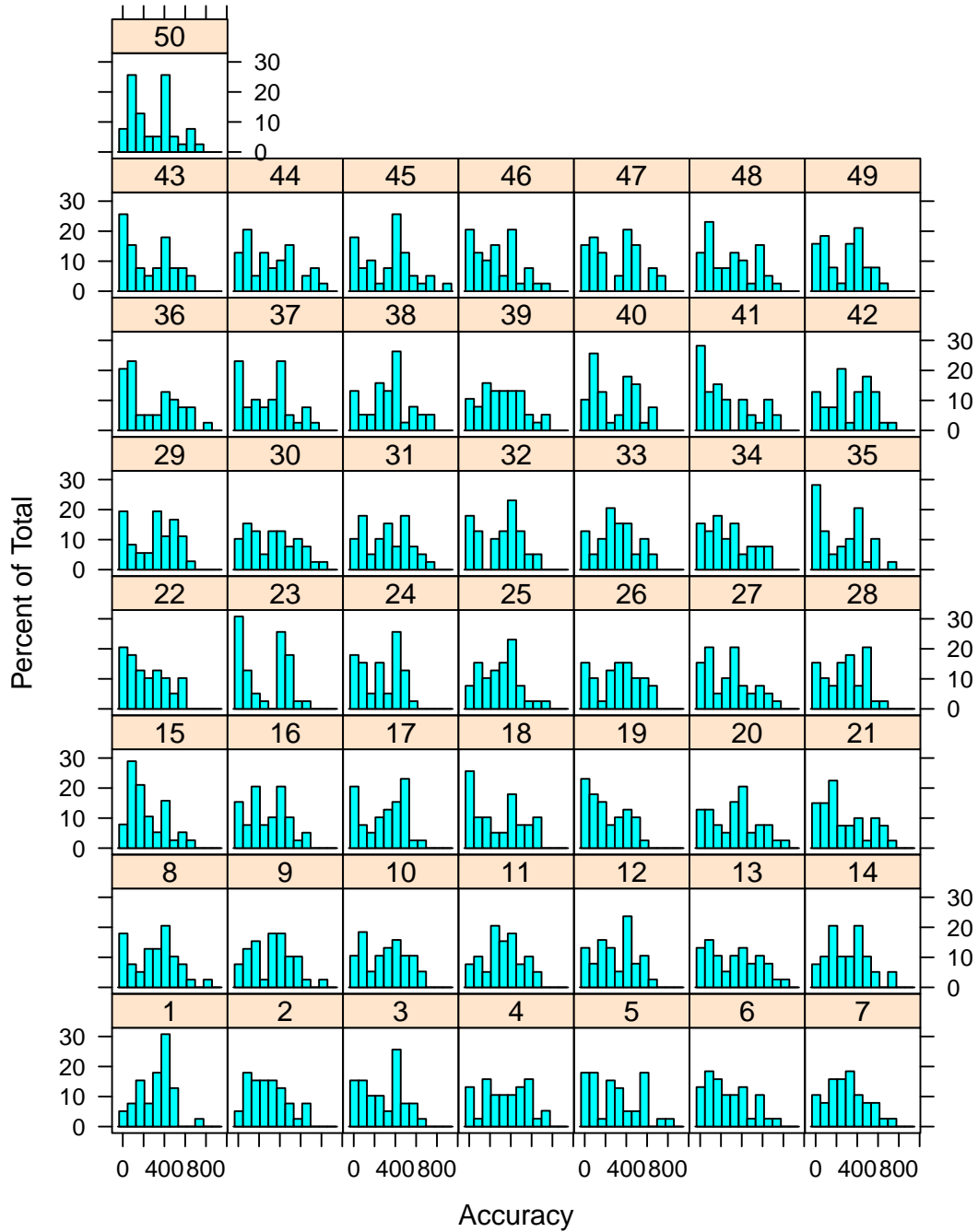
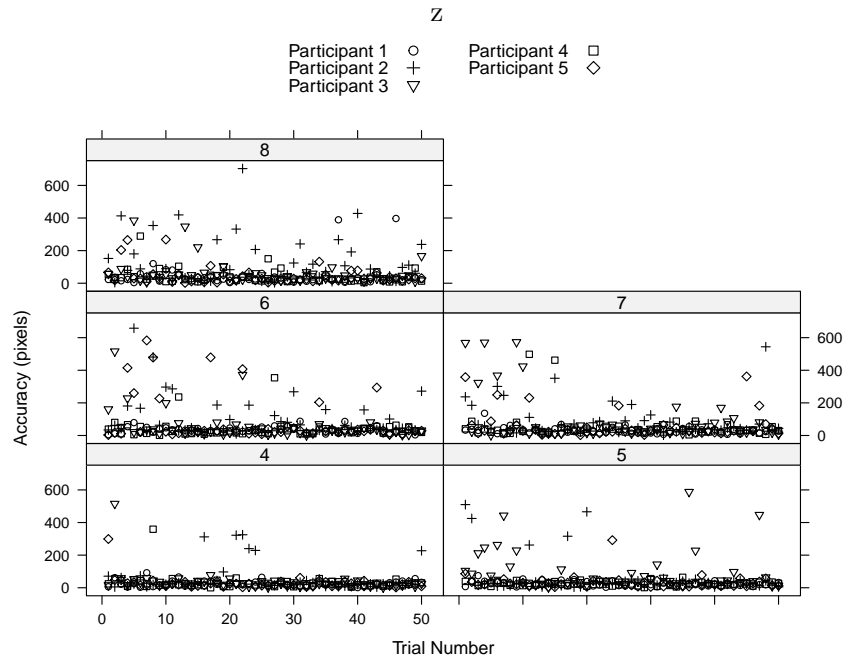
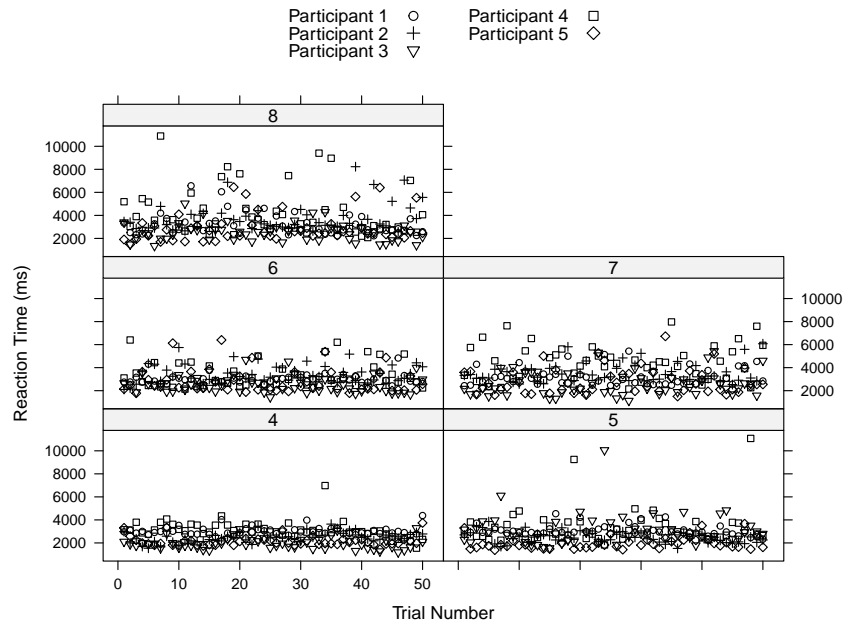


Figure 5. Histogram of Accuracy per trial for 12 Objects for each Participant.



(a) Pilot Scatterplot of Accuracy per Trial Number for each Condition.



(b) Pilot Scatterplot of Reaction Time per Trial Number for each Condition.

Figure 6. Scatterplots of Accuracy and Reaction Time measures per trial number for each object condition from the pilot study.

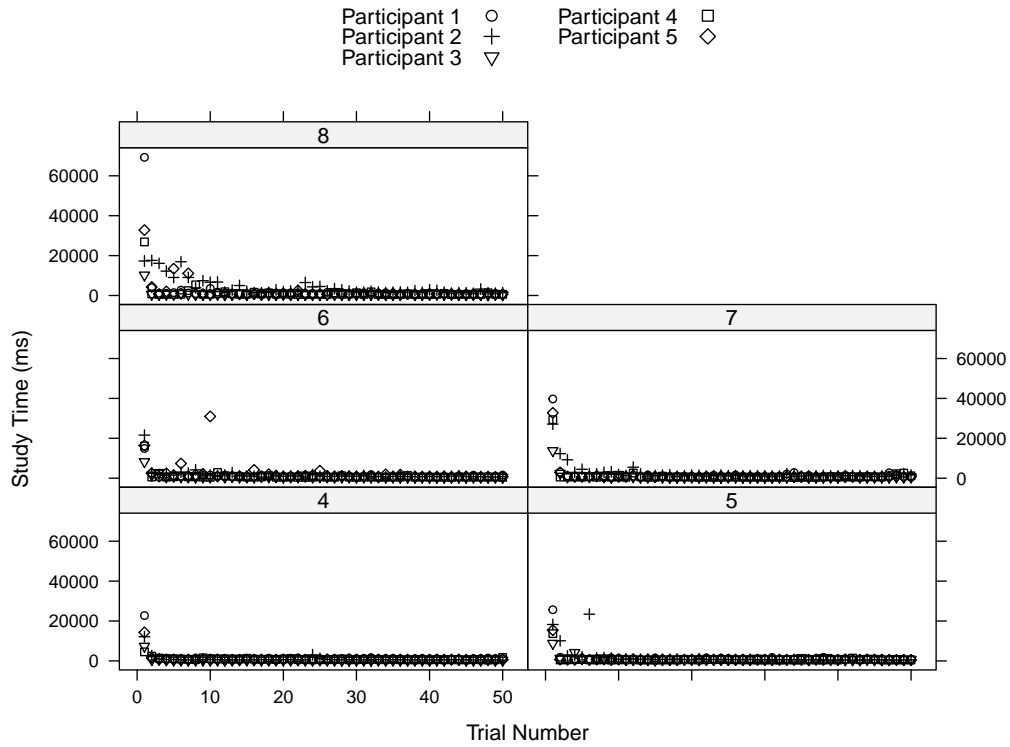


Figure 7. Study Time for each object set size for each participant. With a self-paced study interval, participants typically studied for a long period of time for the first trial followed by drastically shorter periods..

Appendix B: Tables

Table 1.

Object speeds for each relevant MOT and MIT experiments.

Research Literature	Object Speed Range (Degrees of Visual Angle/s)
Pylyshyn & Storm, 1988	Experiment 1 = 1.25 to 15 Experiment 2 = 1.25 to 15
Allen & McGeorge, 2004	1.25 to 9.4
Pylyshyn, 2004	0 to 7.02
Alvarez & Franconeri, 2007	Experiment 1 = 0 to 42 Experiment 2 = 7 to 14
Horowitz et al., 2007	Experiment 1-7 = 4.6 to 29.7
Oksama & Hyönä, 2008	Experiment 1 = 2.6 to 10.7 Experiment 2-3 = 6.3
Hope, 2009; Hope et al., 2010	4.32
Operational ATC (Rantanen, 2010)	0.02 to 0.26
Botterill, Allen, & McGeorge, 2011	7

Table 2.
Object Identities for the Experiment.

4 Objects	8 Objects		12 Objects		
ABC0123	DEF4567	DFE4567	GHI7890	GHI9780	GIH8790
ACB0123	DEF5467	DFE5647	GHI7980	GHI9870	GIH8970
ABC0213	DEF4657	DFE4657	GHI8790	GIH7890	GIH9780
ACB0213	DEF6457	DFE6457	GHI8970	GIH7980	GIH9870

Table 3.

Mean Location Error, Location Error Range, and Identity Error Percentage for each condition.

Condition (Objects)	Mean Location Error (DVA)	Location Error Range (DVA)	Correct	
			Identifi- cation (%)	Approximate Number of Recalled Objects
4	1.45	0.04 - 7.14	88.24	3.53
8	1.55	0.08 - 7.37	41.11	3.29
12	1.91	0.11 - 5.74	23.37	2.80

Note. The mean location error is only for the correctly identified targets, not for all data points. DVA stands for Degrees of Visual Angle. The DVA was calculated using an average viewing distance of 406 mm although participants were free to move their heads. The Location Error Range is lower for the 12-object condition than the other two conditions due to the limited area for the number of objects on the screen.

Table 4.
Object Identities for Future Experiments.

4 Objects	8 Objects		12 Objects		
ABC0123	DEF4567	GHI7890	JKL1234	MNO5678	PQR9012
ACB0123	DFE4567	GIH7890	JLK1234	MON5678	PRQ9012
ABC0213	DEF5467	GHI7980	JKL1324	MNO5768	PQR9102
ACB0213	DFE5647	GIH7980	JLK1324	MON5768	PRQ9102

Appendix C: R Studio Code

The R code that produced Figure 3, 4, and 5:

```
##Histograms of Accuracy each trial for each condition.----

library(lattice, pos=4)
histogram(Participant ~ Accuracy | factor(Trial.Num),
+ pch=16, auto.key=list(border=TRUE),
+ par.settings=simpleTheme(pch=16), scales=list(x=list(relation='same'),
+ y=list(relation='same')), main="4 Objects", data=S39.Stim.Num.4)
#A distinctive positive skew is demonstrated early on in the first couple of trials.

histogram(Participant ~ Accuracy | factor(Trial.Num), type="p", pch=16,
+ auto.key=list(border=TRUE), par.settings=simpleTheme(pch=16),
+ scales=list(x=list(relation='same'), y=list(relation='same')),
+ main="8 Objects", data=S39.Stim.Num.8)
#A distinctive and smooth positive skew is demonstrated after Trial 12 + 13,
#however, there is a regress demonstrating that uncertainty in locations

histogram(Participant ~ Accuracy | factor(Trial.Num), type="p", pch=16,
+ auto.key=list(border=TRUE), par.settings=simpleTheme(pch=16),
+ scales=list(x=list(relation='same'), y=list(relation='same')),
+ main="12 Objects",data=S39.Stim.Num.12)
#A distinctive and smooth positive skew is never completely demonstrated.

##Histograms of RT each trial for each condition.----
library(lattice, pos=4)
histogram(Participant ~ RT | factor(Trial.Num),
+ pch=16, auto.key=list(border=TRUE), par.settings=simpleTheme(pch=16),
+ scales=list(x=list(relation='same'), y=list(relation='same')), main="4 Objects",
+ data=S39.Stim.Num.4)
#A distinctive positive skew is demonstrated early on in the first couple of trials.

histogram(Participant ~ RT | factor(Trial.Num), type="p",
+ pch=16, auto.key=list(border=TRUE), par.settings=simpleTheme(pch=16),
+ scales=list(x=list(relation='same'), y=list(relation='same')), main="8 Objects",
+ data=S39.Stim.Num.8)
#A distinctive and smooth positive skew is demonstrated after Trial 12 + 13,
#however, there is a regress demonstrating that uncertainty in locations

histogram(Participant ~ RT | factor(Trial.Num), type="p",
+ pch=16, auto.key=list(border=TRUE),
+ par.settings=simpleTheme(pch=16), scales=list(x=list(relation='same'),
+ y=list(relation='same')), main="12 Objects",data=S39.Stim.Num.12)
#A distinctive and smooth positive skew is never completely demonstrated.
```

The R code that produced the ANOVAs:

```
###!!!Aggregating Accuracy based on IV's----

#load S39.Rdata
# Aggregate Accuracy measures into new table based on Participant,
# Stim.num, Trial groups (10).

Agg.accuracy = aggregate(S39$Accuracy,
+ list(Participant = S39$Participant, Stim.Num = S39$Stim.Num,
+ Trial.Groups.10 = S39$Trial.Groups.10), mean)

#Designate variables as factors
Agg.accuracy$Trial.Groups.10 =factor(Agg.accuracy$Trial.Groups.10)
Agg.accuracy$Stim.Num =factor(Agg.accuracy$Stim.Num)
Agg.accuracy$Participant =factor(Agg.accuracy$Participant)

#Sort by Participant, Stimulus Number, And Trial Bins.
Agg.accuracy = Agg.accuracy[order(Agg.accuracy$Participant,
+ Agg.accuracy$Stim.Num, Agg.accuracy$Trial.Groups.10),]

#Rename x to Mean Accuracy.
names(Agg.accuracy)[4]<- "Mean.Accuracy"

#Aggregate Accuracy Standard deviation measures in new table.
Agg.accuracy.sd = aggregate(S39$Accuracy,
+ list(Participant = S39$Participant, Stim.Num = S39$Stim.Num,
+ Trial.Groups.10 = S39$Trial.Groups.10), sd)

#Add new variable to table..
Agg.accuracy["Standard.Deviation"] = NA
Agg.accuracy$Standard.Deviation = Agg.accuracy.sd$x

#Add Median to table.
Agg.accuracy.median = aggregate(S39$Accuracy,
+ list(Participant = S39$Participant, Stim.Num = S39$Stim.Num,
+ Trial.Groups.10 = S39$Trial.Groups.10), median)
Agg.accuracy["Median"] = NA
Agg.accuracy$Median = Agg.accuracy.median$x

###!!!!ANOVA on Agg. Accuracy with package ezANOVA [PLAN B]!!!
library( ez )
ezANOVA(data=Agg.accuracy, dv=Mean.Accuracy, wid=Participant,
+ within=(Trial.Groups.10,Stim.Num), type= 3 ,detailed=TRUE)
```



```

with(Agg.accuracy, pairwise.t.test(Mean.Accuracy, Trial.Groups.10,
+ p.adjust.method="bonferroni", paired=T))
with(Agg.accuracy, pairwise.t.test(Mean.Accuracy, Stim.Num,
+ p.adjust.method="bonferroni", paired=T))

library(schoRsch)
anova_out(acc.ezanova, print = TRUE, sph.cor = "GG", mau.p = 0.05,
+ etasq = "partial", dfsep = ", ")

###!!Aggregating Reaction Time based on IV's----
#load S39.Rdata
# Aggregate Reaction Time measures into new table based on Participant,
+ Stim.num, Trial groups (10).
Agg.RT = aggregate(S39$RT, list(Participant = S39$Participant,
+ Stim.Num = S39$Stim.Num, Trial.Groups.10 = S39$Trial.Groups.10), mean)

#Designate variables as factors
Agg.RT$Trial.Groups.10 =factor(Agg.RT$Trial.Groups.10)
Agg.RT$Stim.Num =factor(Agg.RT$Stim.Num)
Agg.RT$Participant =factor(Agg.RT$Participant)

#SoRT by Participant, Stimulus Number, And Trial Bins.
Agg.RT = Agg.RT[order(Agg.RT$Participant,
+ Agg.RT$Stim.Num, Agg.RT$Trial.Groups.10),]

#Rename x to Mean Reaction Time.
names(Agg.RT)[4]<- "Mean.RT"

#Aggregate Reaction Time Standard deviation measures in new table.
Agg.RT.sd = aggregate(S39$RT, list(Participant = S39$Participant,
+ Stim.Num = S39$Stim.Num, Trial.Groups.10 = S39$Trial.Groups.10), sd)

#Add new variable to table..
Agg.RT["Standard.Deviation"] = NA
Agg.RT$Standard.Deviation = Agg.RT.sd$x

#Add Median to table.
Agg.RT.median = aggregate(S39$RT, list(Participant = S39$Participant,
+ Stim.Num = S39$Stim.Num, Trial.Groups.10 = S39$Trial.Groups.10), median)
Agg.RT["Median"] = NA
Agg.RT$Median = Agg.RT.median$x

library(ez)
ezANOVA(data=Agg.RT, dv=Mean.RT, wid=Participant, within=(Trial.Groups.10,Stim.Num),
+ type= 3 ,detailed=TRUE)

```

```
with(Agg.RT, pairwise.t.test(Mean.RT, Trial.Groups.10, p.adjust.method="bonferroni",
+ paired=T))
with(Agg.RT, pairwise.t.test(Mean.RT, Stim.Num, p.adjust.method="bonferroni",
+ paired=T))

library(schoRsch)
anova_out(rt.ezanova, print = TRUE, sph.cor = "GG", mau.p = 0.05,
+ etasq = "partial", dfsep = ", ")
```

The R code that produced Table 3:

```

#####Location and Identity Errors----

L.I.Error.data = S39[!(S39$Correct.Identity==0),]
L.I.Error.data.4 = subset(L.I.Error.data,Stim.Num==4)
L.I.Error.data.8 = subset(L.I.Error.data,Stim.Num==8)
L.I.Error.data.12 = subset(L.I.Error.data,Stim.Num==12)

summary (L.I.Error.data.4)
summary (L.I.Error.data.8)
summary (L.I.Error.data.12)

#Location Error Averages for all data points
loc.error.4.mean =mean(S39$Location.Error
+ [S39$Correct.Identity==1&S39$Stim.Num==4])
loc.error.8.mean =mean(S39$Location.Error
+ [S39$Correct.Identity==1&S39$Stim.Num==8])
loc.error.12.mean =mean(S39$Location.Error
+ [S39$Correct.Identity==1&S39$Stim.Num==12])

#Location Error Ranges for all data points
loc.error.4.range =range(S39$Location.Error
+ [S39$Correct.Identity==1&S39$Stim.Num==4])
loc.error.8.range =range(S39$Location.Error
+ [S39$Correct.Identity==1&S39$Stim.Num==8])
loc.error.12.range =range(S39$Location.Error
+ [S39$Correct.Identity==1&S39$Stim.Num==12])

#Identity Error Matched Percentages
#Table of counts for factors
table(S39$Stim.Num,S39$Identity.Error)
?table
#Percentages
ID.Error.Table = as.data.frame(table(S39$Stim.Num, S39$Identity.Error))

IError.4<-1710/1938*100
IError.8<-798/1941*100
IError.12<-453/1938*100

IError.4/100 * 4
IError.8/100 *8
IError.12/100*12

```