8-10-2009

The Predictive utility of the model of multiple identity tracking in air traffic control performance

Ryan Hope

The Predictive Utility of the Model of Multiple Identity Tracking in Air Traffic Control Performance

by
Ryan M. Hope

A Thesis in

Applied Experimental and Engineering Psychology

Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Science

August 10, 2009

Rochester Institute of Technology
Department of Psychology

We Approve the Thesis of Ryan M. Hope

Date of Signature

. . . .

Esa M. Rantanen, Ph.D.
Associate Professor, Dept. of Psychology
Thesis Adviser and Chair of Thesis Committee

. . . .

Lauri Oksama, Ph.D.
The Academy of Finland and
The University of Turku, Finland
Reader

. . . .

Nicholas DiFonzo, Ph.D.
Professor, Dept. of Psychology
Reader

Acknowledgements

This work would not have been possible without the support and encouragement of my advisor, Dr. Esa Rantanen, under whose supervision I chose this topic and began the thesis. Thank you to all teachers who recommended my study to their students. With out your help I would not have had any participants and this thesis would not have been possible. I would also like to thank Dr. Nicholas DiFonzo and Dr. Lauri Oksama for agreeing to be members of my thesis committee. Finally, I cannot end without thanking my family, on whose constant encouragement and love I have relied throughout my time at RIT.

It is to these people that I dedicate this work.

Abstract

This paper reports an experiment designed to investigate whether the Model of Multiple Identity Tracking can be use to predict air traffic controller performance in terms of situation awareness. The experiment tests a number of predictions derived from the Model of Multiple Identity Tracking. This model posits that when tracking multiple moving objects, the location of 4-5 objects can be acquired in parallel but the identification of any one target needs serial attention. While one object is being identified, an error factor associated with the location of all other objects increases with respect to an objects speed and the amount of time since the object's location-identity binding was refreshed. On an individual basis, working memory limits the number location-identity bindings that be stored at any one time and long term memory makes familiar objects easier to track.

The experimental task consisted of tracking a set of moving objects for twenty seconds. The objects were 6-character strings; three letters followed by three numbers. After tracking the objects for twenty seconds, the participant was instructed to locate a target-objects. The time required to find the target object was recorded. The number of objects and the magnitude of direction changes (entropy) were manipulated. The main effect of number of objects was found to be significant. The main effect of entropy was found to be marginally significant. The pattern of results supports the idea that the Model of Multiple Identity Tracking (MOMIT) can be used to predict air traffic controller performance. A formula derived from MOMIT to predict completion time showed a good fit to the experimental data.

Table of Contents

Introduction **1**

Air Traffic Controller Performance 1

Aircraft Trajectory Predictability and Situation Awareness 2

Confounds of Location Predictability 3

Entropy: probability of direction change magnitude 4

Models of Visual Attention 5

Multiple Object Tracking 6

Visual Search 7

Model of Multiple Identity Tracking 8

MOMIT and Entropy 9

MOMIT and ATC Performance 9

Differences between classic MOT and ATC 10

Goals and Hypotheses 11

Method **13**

Participants: sampling procedure, size and power 13

Apparatus 13

Stimuli 13

Trial generation 14

Determining Direction of Travel for Objects 15

Object Movement 15

Independent variables 16

Number of moving objects: 4 levels (4,9,14,19) 16

Entropy of object: 3 levels (0.00, 0.69, 1.00) 16

Design . . . 17
Dependent variables . . . 18
Task . . . 18
Procedure . . . 19

Results **21**
Preliminary Analyses . . . 21
Outliers . . . 21
Sphericity . . . 21
Dispersion of initial object locations . . . 22
Analysis of Variance . . . 23
Planned Contrasts . . . 23
Omnibus ANOVA . . . 24
Analysis of Covariance . . . 24
Entropy and Displacement Per Second . . . 25
Model Fitting . . . 27

Discussion **32**
The generalizability of the observed effects . . . 32
Contributions to MOMIT . . . 34
Application to ATC . . . 34
Design Issues . . . 34
Conclusion . . . 35

References **36**

APPENDIX A: Figures **40**

APPENDIX B: RCode **45**

APPENDIX C: IRB Forms **49**

List of Figures

1 Displacement confounds 3
2 Entropy probability distributions 5
3 Classic MOT experiment 7
4 Experimental task 19
5 Completion time lineplots 29
6 Displacement lineplots 30
7 Model fit 31

List of Tables

1 Entropy probability distributions 17

2 Index of dispersion chi-square tests 23

3 Completion time means and standard deviations 23

4 Analysis of Variance 24

5 Analysis of Covariance 25

6 Linear regressions 26

7 Non-linear regressions 28

Introduction

The air traffic control (ATC) industry is constantly looking for new ways to improve performance of the controllers and reduce the number of accidents while at the same time increase efficiency and reduce costs. A predictive model of controller performance could help the ATC industry accomplish these goals. For example, automation systems relating to air traffic control could benefit from performance models. Models are useful for designers who want to know how users will interact with their systems. Oftentimes, there is not enough time or resources to test their designs with actual users. This is especially true for systems intended for trained users, like air traffic controllers. It can be expensive to train users with a prototype, but unskilled users do not interact with a product in the same way as a skilled user. Performance modeling aims to solve this problem by predicting how a skilled user will interact with a system. Simulation and analysis take the place of expensive training and testing. Performance models could also be used in an attempt to prevent controllers from making errors, for example, offloading work when a controller's performance is predicted to drop below a certain threshold (Charlton & O'Brien, 2002).

Air Traffic Controller Performance

A central factor in ensuring aircraft safety is the degree to which controllers have sufficient situation awareness (SA) to maintain safe separation of aircraft. SA is critical for controllers who must maintain up-to-date assessments of the rapidly changing location of each aircraft and their projected future locations relative to each other. Controllers typically call the mental model from which they base all their decisions the "picture". This picture is what researchers are referring to when they mention SA. Many definitions of SA have been developed; some are very closely tied to the avaition domain and some are more general. "A general definition of SA that has been found to be applicable across many domains breaks SA into 3 levels: level-1, the perception of the elements in the environment within a volume of space and time; level-2, the comprehension of their meaning; and level-3, the projection of their status in the near future" (Endsley & Garland, 2000, p. 5). Many of the technological changes being implemented to enable free flight involve the use of automation. Human operators acting as monitors of automated systems often exhibit problems in detecting system errors and performing tasks manually in the event of automation failures (Wickens & Hollands, 2000). With many automated systems, forming the higher levels of SA becomes significantly difficult (Carmody & Gluckman, 1993; Endsley & Kiris, 1995). A performance model for ATC needs to account for factors affecting all levels of situation awareness.

Aircraft Trajectory Predictability and Situation Awareness

In the current system, controllers gain information about how the aircraft is going to behave from knowledge of their assigned flight path and destination. There are a limited number of ways that aircraft will proceed through a given airspace according to a given flight plan and the aircraft intended activity in that sector (e.g., approach, departure, or en route). The controller can usually detect deviations from these norms quickly (Wickens, Mavor, & McGee, 1997; Wickens, Mavor, Parasuraman, & McGee, 1998). With the advent of technologies such as GPS and the Traffic Collision Avoidance System (TCAS), the concept of free flight is changing the way that air space is managed. With free flight, aircraft may come from almost any direction into a sector, change paths many times without controller action or approval, and depart the sector in almost any direction. With this loss of aircraft predictability comes lower situation awareness and subsequently, the ability of the controller to determine potential separation problems may be reduced (Endsley & Rodgers, 1996; Mogford, 1997; Endsley et al., 1997; Metzger & Parasuraman, 2001).

There are three aspects of an object's motion that can be manipulated which all affect the predictability or uncertainty associated with its location: (a) the velocity of an object (Oksama & Hyönä, 2008), (b) the rate at which an object changes direction, and (c) the magnitude of changes in direction. The common link between these three factors is that each affects the average displacement per second. Speed by definition is related to displace-

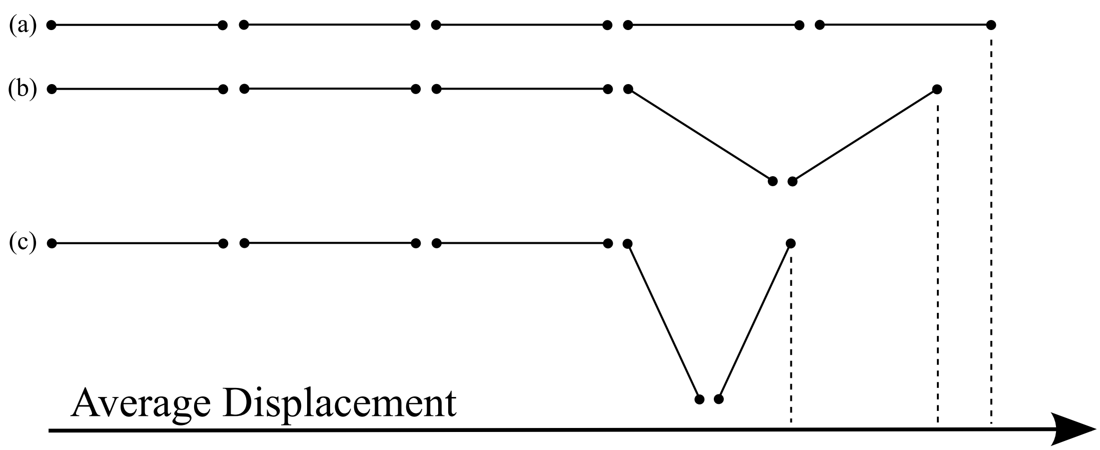


Figure 1: The individual line segments are supposed to represent an object traveling for 1 second at a constant speed, thus line a, b, and c represent the displacement of the object after 5 seconds. Path a and b have the same speed but different rates of direction change; the path with the higher rate of direction change has lower average displacement. Path b and c have the same speed and the same rate of direction change but different magnitudes of direction change; the path with the larger magnitude direction change has lower average displacement.

ment per second. When speed is held constant, an object that changes direction will have a smaller average displacement than an object that does not change direction, see lines (a) and (b) in *Figure 1*. When velocity and the rate of direction changes are held constant, an object with large magnitude direction changes with have a smaller average displacement than an object with small magnitude direction changes, see lines (b) and (c) in *Figure 1*. There is no way to separate the effects of velocity from from the effects of changing direction on displacement. Therefore, average displacement per second best quantifies the uncertainty of a moving objects location. Average displacement per second is similar to velocity in both magnitude and units. To calculate an average displacement per second, sample the displacement of an object over many seconds then divide the displacement by the sample duration. Consider the example in *Figure 1.* Lines a, b, and c represent the path of an 3 objects. Each line segment represents equal displacement of that object in one second. If the displacement of each object was sampled after each 1 second interval, the displacement of each object is equal. If the displacement of each object was sampled after 5 seconds however, the displacement of each object is different.

Entropy: probability of direction change magnitude. None of the past MOT research has attempted to quantify an object's motion in terms other than velocity, nor have they systematically manipulated how often an object changes direction or the magnitude of direction changes. Information theory could be used to predict the probability of a direction change of a certain magnitude. In information theory, entropy is a measure of the uncertainty (or predictability) associated with a random variable. When applied to moving objects, entropy could refer to predictability of an object making a direction change of certain magnitude. If the velocities and the rate of direction change of a set of moving objects are all equal, then the entropy of any one object could be quantified using a modified version of the entropy formula introduced by Shannon (1948):

$$H(x) = -1 \sum_{i=0}^{n} p(x_i) log_n p(x_i) \tag{1}$$

where n is the number of bins the range of possible direction changes is divided into, and $p(x_i)$ is the probability of turning in the direction associated each bin x. When there is no entropy, the probability of going straight (a change of direction of zero degrees) is 100% and the probability of all other directions is 0. When there is a medium amount of entropy, the probability of going straight is less than 100% and the probability of a direction change in a certain direction decreases as the magnitude of the direction change increases. When entropy is at its lowest, the probability of all magnitude direction changes are equal, see *Figure 2*. The deviation from Shannon's original equation is in the base of the loga-

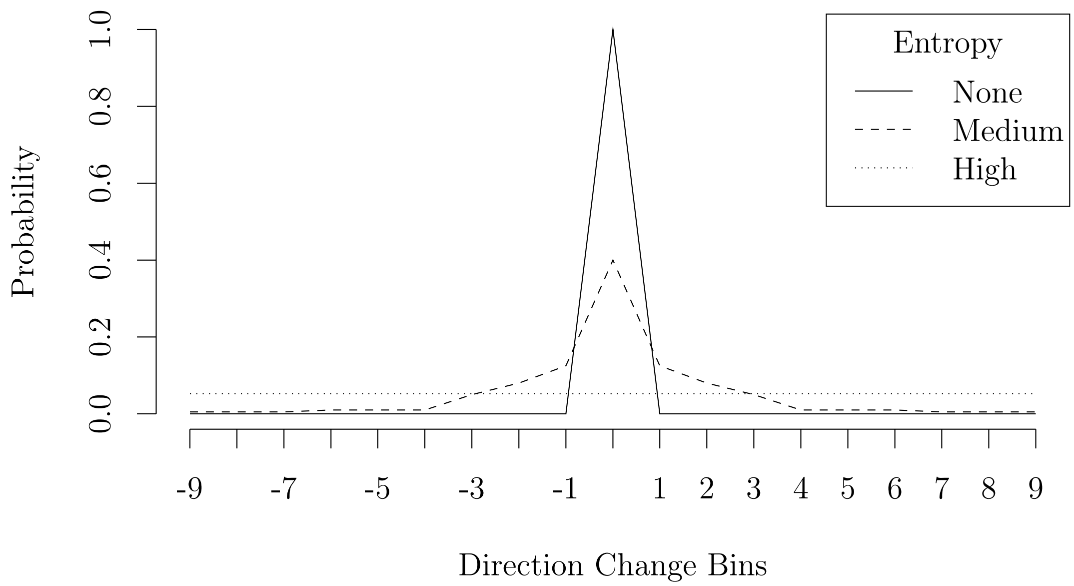


Figure 2: Entropy probability distributions for none, medium and high entropy conditions.

rithm. Shannon's equation used a base of 2 whereas *Equation* (1) uses a base of n. Using a base of n ensures that $H(x) = 1$ when the probability of every one of the n direction changes are equal (the highest entropy level) and $H(x) = 0$ when the probability of going straight is 100% (the lowest entropy level).

Models of Visual Attention

Most models of visual attention are based on fixed- or limited-capacity parallel processing. Fixed-capacity parallel models have been used to describe both general "static" visual attention (Bundesen, 1990; Logan, 1996, 2002; Bundesen, Habekost, & Kyllingsbaek, 2005) and "dynamic" visual attention (Cavanagh & Alvarez, 2005; Pylyshyn & Storm, 1988). An assumption common to the fixed-capacity parallel models is that multiple visual objects can be selected and spatially tracked in parallel. This tracking is done preattentively (Pylyshyn & Storm, 1988), attentively (Cavanagh & Alvarez, 2005) or a combination of both (Bundesen, 1990; Logan, 1996, 2002; Bundesen et al., 2005).

Pylyshyn and Storm (1988) most explicitly define fixed capacity at 4-5 “FINgers of INSTantiation” or FINST, visual indexes that move along with the moving objects, as if the fingers were glued to the tracked objects. The multi-focal model of Cavanagh and Alvarez (2005) posits four attentional foci, two of which are tracked from the left visual field and two from the right visual field (Alvarez & Cavanagh, 2005).

Multiple Object Tracking

Multiple object tracking (MOT) is an experimental paradigm, similar to ATC, designed to study how the human visual system tracks multiple moving objects. It was created by Pylyshyn and Storm (1988) in an attempt to test and illustrate their proposed theoretical mechanism called a Visual Index or FINST (for FINgers of INSTantiation). The FINST theory posits a small number of indexes or pointers that pick out and stay attached to individual objects in the visual field independent of any changes in their properties, allowing for the objects to be tracked. The theory was created to address the question of how conceptual descriptions can pick out individual visual objects despite the fact that descriptions themselves are insufficient in general to pick out tokens. The FINST theory claims that the tracking aspect of MOT is automatic and non-attentional, though others view it as illustrating split attention (Cavanagh & Alvarez, 2005).

A typical MOT task (shown in *Figure 3*) starts with a display of identical objects (t=1). Subsequently, a subset of “target” objects are cued with a brief flash to make them distinctive (t=2). After the cue, the targets stop blinking so that the “target” objects become once again indistinguishable from the other “distracter” objects. All objects then move in a random fashion for about 10 seconds (t=3). The motion then stops (t=4) and the observer’s task is to indicate all the tracked objects by clicking on each one using a computer mouse. In some studies instead of identifying all targets, the observers task is to

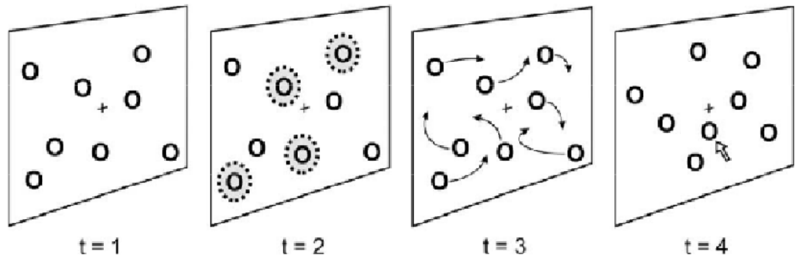


Figure 3: Sequence of events in a typical multiple object tracking experiment (Pylyshyn & Storm, 1988).

judge whether a particular object, flashed at the end of the trial, was a target (Pylyshyn, 2005). ATC could be considered a real life MOT task (without distracters). Controllers must maintain up-to-date assessments of the rapidly changing location of each aircraft and their projected future locations relative to each other (Endsley & Jones, 1996).

Visual Search

The MOT task has been used widely in the study of visual attention and particularly in the study of sustained multiple-locus of attention. However, the question of what is selected by visual attention is equally important as the question of how and under what conditions selection takes place. The previous research has shown that selection can be automatically induced by what some have called exogenous cues that are automatic and data-driven, or can be voluntarily allocated by symbolic or endogenous cues (Theeuwes, 1994). While Pylyshyn has provided abundant evidence that selected objects are available simultaneously, it is not clear whether they must be selected automatically (and preattentively) or whether some voluntary and perhaps serial process may be involved.

In a static visual search, depending upon the relationship between targets and distracters, search can vary from being either be extremely efficient and unaffected by the number of distracters in the field, to being inefficient and directly related to the number of distracters present. Both types of search can be observed when observers do not make eye movements, and so the differences reflect the variations in the efficiency of internal mechanisms of selection (Zelinsky & Sheinberg, 1997). Typically when targets and distracters differ on the basis of some salient, simple feature (e.g., orientation or color), search is efficient. This suggests that such simple features can be computed and compared in parallel across the visual field. Such computations can be carried out prior to any selection of one part of the field. In contrast, when search is directly related to the number of distracters present, it may be that the differences between targets and distracters are not computed prior to selection but only afterwards (Humphreys, 1996). "Early" selection theories stress that only simple visual properties, such as edges of particular spatial frequencies, colors and orientations, are coded prior to selection "pre-attentively" (Treisman, 1998). Such properties may be registered rapidly by special purpose detectors, which operate in parallel across the visual field. Higher-order representations of objects (conjunctions of features) can only be computed more slowly, following selection, and perhaps even at just one location at a time (Humphreys, Gilchrist, & Free, 1996). In ATC, objects only differ based on higher-order features. For example, all callsigns have 3 letters and 3 to 4 numbers; what differs is the combination of the letters and numbers. A recent study (Pylyshyn & Annan, 2006) supports the idea that selection of multiple targets, defined by features that do not capture attention in an automatic exogenous manner, requires that targets be visited serially.

Model of Multiple Identity Tracking

A new model titled the Model of Multiple Identity Tracking (MOMIT) accounts for both the parallel and serial aspects of dynamic visual attention. MOMIT is based on five premises: 1) Efficient maintenance of multiple dynamic objects requires serial refreshing (or reactivation) of identity-location bindings; if the bindings are not refreshed periodically, they will eventually be lost. The refreshing of existing bindings is assumed to be non-automatic (serial) and effortful requiring continual shifting of attention between targets. 2) The number of identity-location bindings that can simultaneously kept active in the episodic buffer is limited. Furthermore, the maximum number of bindings varies significantly between individuals. 3) Long-term memory (LTM) facilitates bindings; tracking performance is better for familiar than unfamiliar targets. 4) As targets continuously move, there is a location error in the spatial index, which is stored in visual short term memory, (VSTM). 5) The system responsible for switching attention during tracking also obtains location information of moving targets in parallel though peripheral vision. However, unlike the information provided by VSTM, this spatial information is not indexed (Oksama & Hyönä, 2008).

MOMIT and Entropy. In order to predict an object's future location (a sign of high level-3 situation awareness), it is possible that besides binding location with identity, the direction of travel and speed of an object might also be bound to identity. For example, in ATC, a controller might anticipate where each object is based on the object's last known velocity and direction of travel. Freyd and Finke (1984) and Finke and Shyi (1988) provide evidence supporting this form of mental extrapolation which they call *representational momentum.* When an object changes direction before it's identity-location binding is refreshed, then the anticipated object location might not be the actual object location. If controllers do use something like a velocity vector to anticipate where objects will be in the future, then increases in the entropy of an object should cause decreases in the controller's tracking performance and situation awareness.

MOMIT and ATC Performance. MOMIT has the potential to predict ATC performance. While MOMIT in its current form predicts object tracking accuracy, a formula could be derived to predict reaction time. A mathematical formula based on MOMIT which predicts search time consists of two components:

$$T = N * t \tag{2}$$

the average number of objects visited before finding the correct target N, and time needed to process one object t. If the number of objects to track is n, the probability of having access to a binding is m/n, when guessing probability is not considered. In addition to the binding capacity, MOMIT posits the probability of guessing affects the performance. P_{guess} is influenced by the number of response alternatives and by possible strategies adopted by the participant. In other words, the number of objects visited varies as a function of binding capacity m, the number of objects to track n, the probability of guessing (P_{guess}) non-remembered items 1 - m/n, and the toal number of objects n-1; see equation (3).

$$N = (\frac{m}{n} + (1 - \frac{m}{n}) * P_{guess}) * (n - 1) \tag{3}$$

Using equation (3) in place of N in equation (2), the final formula for search time is shown in equation (4).

$$T = (\frac{m}{n} + (1 - \frac{m}{n}) * P_{guess}) * (n - 1) * t \tag{4}$$

This formula resembles a non-linear function when n is small and becomes linear with a slope close to P_{guess} as n get larger. Thus, if P_{guess} was equal to 0.5 (chance levels), this formula would resemble the formula for a typical serial search at large values of n.

Differences between classic MOT and ATC

Although the classic MOT task is very similar to ATC, there is a fundamental difference that makes applying MOMIT to ATC difficult. There is no concept of targets and distracters in ATC. In ATC, all objects on controllers' displays are usually targets that need to be tracked. There are seldom distracter objects in ATC. Having distracters present allowed past research to use signal detection theory to quantify performance. Without distracters and eye tracking equipment, the only performance measure that does not modify the task too much, is a reaction time measure, the time required to find a target object in the set total objects*completion time*. Relating performance in terms of signal detection to performance based on completion time is not trivial.

Goals and Hypotheses

The present study sought to accomplish a number of goals. The first goal was to determine if research based on MOT tasks with distracters can be applied to tasks without distracters, like ATC. Accomplishing this goal required completion of two subgoals: (a) a novel MOT experimental paradigm more similar to real life ATC than past MOT tasks needed to be created, and (b) the predictions of past research needed to combined into a predictive formula performance. The new experimental task was created to more accurately resemble ATC by: (a) using objects that look like aircraft callsigns, (b) using speeds consistent with what is found on ATC displays, (c) using object set-sizes similar to ATC, and (d) requiring all objects to be tracked (no distracters). The second goal was to explore how entropy manipulations might affect tracking performance. Before the present study, entropy had never been systematically manipulated. Entropy could affect tracking performance in two ways. If an object's future location can be predicted based on it's current direction of travel, then tracking performance should decreases with increases in entropy. On the other hand, since average displacement decreases as a function of entropy, and displacement has been shown to be negatively correlated with tracking performance, increases in entropy could increase performance (Keane & Pylyshyn, 2006). In the experimental task used in the present study, the number of objects and the entropy of the objects were manipulated. Performance was measured by the time required to find a target object in the set of objects, or task completion time. The following predictions were made: (a) the time required find the target object should be based on equation (4), and (b) search time should increase as a function of entropy.

Method

Participants: sampling procedure, size and power

A priori power analyses were conducted prior to recruiting participants. Based on an $\alpha = 0.05$, $1 - \beta = 0.8$ and the desire to detect a medium effect size ($r = 0.3$), the target sample size was thirty (Cohen, 1988). Thirty one students (15 male, 16 female) volunteered to participate in the present study. Of the thirty one participants, 3 were hearing impaired. All participants had normal or corrected to normal vision. There were no restrictions with regards to who could participate, however, all participants were sampled from the undergraduate population of Rochester Institute of Technology in Rochester, NY. Participants were referred to the present study by psychology class professors and as compensation, given extra credit. The amount of extra credit offered was determined by each professor and was different from class to class.

Apparatus

The computers used to run the experiment were all Dell Optiplex GX260s with 1280 MB of ram and a 2.8 GHz Pentium 4 processor. The display for each computer was a 17" LCD running at a resolution of 1280x1024 (96 DPI) at 60Hz. The program used in the experiment was written in Java.

Stimuli

The stimuli in the present study, or the objects which the participants were instructed to track, were 6-character strings. The strings consisted of a 3-letter International Civil Aviation Organization (ICAO) airline designator (ex. SWA, AAL, UAL) picked randomly from a table of real airline designators, followed by a random 3 digit number between 99 and 1000; see *Figure 4* for an example. A mono-spaced font was used to ensure the width of all the objects remained constant. A font size of 12 was used to ensure focal vision was required to determine the identity of the object. Based on an average viewing distance of 16", each object was approximately 1.57 degrees of visual angle wide and 0.45 degrees of visual angle high requiring the participant to use focal vision to clearly see and recognize the object.

Trial generation. The initial object locations were picked randomly at the beginning of each trial. In an attempt to prevent this randomization from influencing the dependent variable, the function in the experimental program responsible for generating the initial object locations only used sets of objects that passed certain criteria (X,Y offset and X,Y variance). If a set of initial object locations did not meet the spatial randomness criteria, a new set of initial object locations was generated.

To ensure that the distribution of initial targets was centered on the screen, the average of the x and y coordinates of the initial target locations were calculated. To meet the required spatial randomness criteria for the present study, sets of targets needed to have an average x and y coordinate within a 20-pixel by 20-pixel box ($\pm$10 pixels in the x and y direction) relative to the center of the screen.

To ensure the initial target locations were not clustered in any one part of the screen, the variance of the x and y coordinates of the initial target locations was calculated. The selection criteria for the x and y variance was the resolution of the screen in one direction multiplied by that same resolution but first divided by the greatest common divisor (GCD) of the x and y resolution. The result was then multiplied by a constant; the larger this constant, the higher the minimum variance. During the coding stage of the experimental program, it was determined that setting the constant to 20 resulted in an even dispersion of objects on the screen in the x and y axis. For a 1280x1024 resolution, the GCD is 256.

The x and y resolution was then divided by the GCD to get the screen ratio; 1280 / 256 = 5, 1024 / 254 = 4. For the x direction the variance had to be greater than or equal to 1280 X 5 / 20 = 128000. The variance in the y direction had to be greater than or equal to 1024 X 4 X 20 = 81920.

Determining Direction of Travel for Objects. In addition to initial object location, the initial direction of travel (in degrees from 0 to 359) was also picked randomly at the beginning of each trial. The experimental program contained 3 static ten-thousand cell arrays (one for each level of entropy) containing the change in direction associated with, and in proportion to, the 19 bins listed in *Table 1.* For example, if the probability of a particular bin was 0.40, then 40% of the ten-thousand cells contained the direction change associated with that bin. When it was time for an object to change direction, a random number between -1 and 10000 was picked. This random number was then used to pick a direction change from a cell in the entropy array corresponding to the level of entropy of the current trial. The direction change was then added to the objects current direction of travel.

Object Movement. At the beginning of each trial when the direction of travel was first set, and after an object's direction of travel was updated following a change in direction, the direction of travel (in terms of degrees) was broken into its x and y velocity components by taking the sine and cosine of the direction of travel. The loop responsible for moving the objects performed the following steps on each target: (a) the x and y velocity components is added to the x and y location coordinates; (b) the new location is checked to determine if a screen border was crossed and if a border was crossed, the object's location was moved to being exactly on the border and the direction of travel was modified to mimic a mirror-like reflection off the border; (c) if a border was crossed the x and y velocity component was recalculated and a *steps* counter reset to 0, the *steps* counter was incremented if a border was not crossed; (d) the *steps* counter is compared k (the number of steps the object makes before changing direction) and if the steps counter is larger, a new direction of travel is picked; (e) if a new direction of travel was picked the x and y velocity components are recalculated, and the steps counter is reset to 0. After completing these steps for each object, the screen was repainted which updated the object locations. The object moving thread then slept for 20 ms before repeating the process.

Independent variables

Number of moving objects: 4 levels (4,9,14,19). Since the primary goal of the present study was to predict air traffic controller performance, the range of number of objects used was specifically chosen to mimic what a real controller might experience. ATC commonly requires tracking in excess of 40 objects. The number of moving objects used in the present study was also much larger than what was used in past research. In past MOT experiments, the most number of moving objects used at one time was typically around 8 to 10 objects in total with a subset of those objects (usually about 1 to 5) being target objects (objects which the participant is supposed to pay attention to or track). In the present experiment the highest number of moving objects was 19, and unlike past studies, any object could be a potential target. Since Pylyshyn and Storm (1988) and Oksama and Hyönä (2008) showed performance tracking is near perfect with around 4 or less objects, 4 was chosen as the lowest level of the number of objects factor.

Entropy of object: 3 levels (0.00, 0.69, 1.00). The probability distributions used with *Equation* (1) to calculate the three levels of entropy are shown in *Table 1.* The distributions were created first by dividing the range of possible direction changes (-35°to 35°relative to an objects current direction of travel) into 19 equal bins (*n=19*). The probabilities used in each level of entropy were intended to create a low, medium and high entropy condition. There was initially four levels of entropy. However, in order to reduce the length of the experiment, the low entropy condition ($H(x) = 0.34$) was dropped. As a result, the medium entropy condition did not turn out to be equal distant between the none and high entropy conditions. In the $H(x) = 0.00$ condition, each object traveled in one direction, the objects did not change direction as they move. For the $H(x) = 0.69$ condition, objects changed direction every 7 *steps*. For the $H(x) = 1.00$ condition, objects changed direction every $7{\pm}2$ *steps* determined randomly at the beginning of each trial for each target. One *step* refers to one complete cycle of the loop in the experimental program responsible for moving the objects. This loop will be explained in further detail momentarily. The number of *steps* an objects took before changing directions (k) remained constant for the duration of the trial.

Table 1: A table of probability distributions for three levels of entropy: none, medium and high. H(x) for each level of entropy was calculated with *equation* (1).

	Entropy Probabilities		
Directions	None	Medium	High
$0/9 * 35°$	1.0000	0.4000	0.0526
$1/9 * 35°$	0.0000	0.1250	0.0526
$-1/9 * 35°$	0.0000	0.1250	0.0526
$2/9 * 35°$	0.0000	0.0800	0.0526
$-2/9 * 35°$	0.0000	0.0800	0.0526
$3/9 * 35°$	0.0000	0.0500	0.0526
$-3/9 * 35°$	0.0000	0.0500	0.0526
$4/9 * 35°$	0.0000	0.0100	0.0526
$-4/9 * 35°$	0.0000	0.0100	0.0526
$5/9 * 35°$	0.0000	0.0100	0.0526
$-5/9 * 35°$	0.0000	0.0100	0.0526
$6/9 * 35°$	0.0000	0.0100	0.0526
$-6/9 * 35°$	0.0000	0.0100	0.0526
$7/9 * 35°$	0.0000	0.0050	0.0526
$-7/9 * 35°$	0.0000	0.0050	0.0526
$8/9 * 35°$	0.0000	0.0050	0.0526
$-8/9 * 35°$	0.0000	0.0050	0.0526
$9/9 * 35°$	0.0000	0.0050	0.0526
$-9/9 * 35°$	0.0000	0.0050	0.0526
H(x)	0.00	0.69	1.00

Design. The present study was a 4x3x5 fully-factorial within-subjects design. The 4 levels of the number of objects and the 3 levels of the entropy created 12 experimental blocks. Each participant performed 5 consecutive trials in each of the 12 experimental blocks creating 60 observations from each participant. The order of the 12 blocks was random for each participant instead of counterbalanced. With 31 participants, each block had 155 observations. There were 1860 observations in total.

Dependent variables

The main dependent variable was completion time, or the time required to find the target object starting from the moment the targets were masked and ending the moment the target object was clicked. Directly measured completion time was used to indirectly measure level 1 situation awareness. If a participant had maintained an awareness of the object identities and locations while they were moving, he or she should be able to click the target object immediately without checking identities of other objects, resulting in a very short completion time.

Task

The participant's task was to track a set of objects which moved on a computer screen for a fixed period of time after which the motion stops and a target object must be found, see *Figure 4*. At the beginning of each trial, the objects were drawn on the screen, but did not move (a). The objects remained motionless for n X 0.5 seconds where n was equal to the number of objects on the screen at that time. This ensured there was enough time for the participant to view the starting location of each object. After this "preview" period, the objects moved for 20 seconds (b). After 20 seconds of motion the objects stopped moving (c) and the identity of each object was masked (changed to "$$$$$$"). At the same time, a message at the bottom of the screen appeared instructing the participant to click on a target object (d). When the cursor was moved over a masked object, its identity was revealed; when the cursor was moved off an object, the object was re-masked. To complete a trial the participant was required to click on the target object.

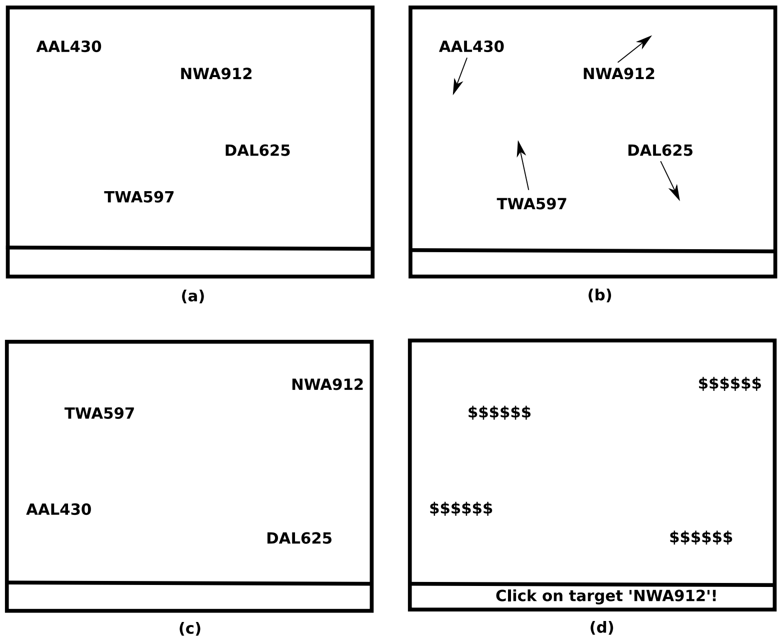


Figure 4: Sequence of events in the multiple object tracking task of the present experiment.

Procedure

Prior to starting the experiment, the task was explained to the participants and any questions from the participants were answered. The participants were then instructed to read and sign the consent form if they wanted to continue participating. Once the consent form was signed, the participant started the experimental trials. Participants were seated approximately 16" away from the screen. Before starting, participants were reminded to find the target object as quickly and accurately as possible. There were 65 trials in total; 1 block of 5 practice trials followed by 12 blocks (each block representing a combination of the independent variables) of actual trials. Before each trial, a splash screen was displayed for 2 seconds showing the block number and replicate. The order in which the blocks were presented were random for each participant to minimize the effects of practice and/or fatigue. The experiment took, on average, 45 minutes to complete. After the experiment was completed, any additional questions a participant might have had were answered.

Results

All analyses were done with the *R* statistical computing language. Analysis of variance computations were done with the *aov* function in the *stats* package for *R*. Linear and non-linear mixed-effects model computations were done with the *lmer* and *nlmer* functions in the *lme4* package for *R* (Bates, Maechler, & Dai, 2008; R Development Core Team, 2008). The independent variable completion time was non-normally distributed and thus log-transformed in all calculations and analyses (Rosenthal & Rosnow, 2008). The goodness of fit for mixed-effects models is compared using Akaike's information criterion (AIC). AIC is based on the concept of entropy, in effect offering a relative measure of the information lost when a given model is used to describe reality and can be said to describe the trade off between bias and variance in model construction, or loosely speaking that of precision and complexity of the model. AIC is computed using the maximized value of the likelihood function and differences in AIC of nested models can be tested using a chi-squared test since the likelihood function assumes that the underlying errors are normally distributed (Sakamoto, Ishiguro, & Kitagawa, 1986; Crawley, 2007).

Preliminary Analyses

Outliers. Potential completion time outliers were detected using the modified Z score introduced by Iglewicz and Hoaglin (1993). Z scores were calculated based on the log of "completion time divided by number of objects". The absolute value of only 4 of the 1860 Z scores were above the outlier criterion value of 3.5. For this reason, the complete data set was used in the remaining analyses.

Sphericity. Within-subjects ANOVA assumes, for all factors, that the different levels of each factor have equal varriance, or *sphericity.* Violating the sphericity assumption results in a loss of power; the F-ratios produced cannot be trusted. To test the data for unequal variance between groups, Levene's test for equality of group variability was used. Levene's test of the entropy factor on completion time was not statistically significant, F(2,1857)=0.73, p<0.49. Levene's test of the number of objects factor on completion time was statistically significant, F(3,1856)=213.14, p<0.001. To compensate for the violation of the sphericity assumption, either post-hoc corrections to the F-ratio (such as the Greenhouse-Geisser or Huynh-Feldt corrections) or an analysis technique that does not require sphericity, like mixed-effects models, can be used.

Dispersion of initial object locations. Since the initial object locations on each trial were random, it was possible that the dispersion of initial object locations formed patterns that might have affected completion time. In order to rate each trial's "quality of dispersion", the index of dispersion was used. The index of dispersion (Equation (5)) or variance-to-mean ratio is used as a measure to quantify how clustered or dispersed a set of observations is (Diggle, 1983).

$$I = \sum_{i=1}^{m} (n_i - \bar{n}^2)/[(m-1)\bar{n}] \tag{5}$$

The index of dispersion along the X and Y axis (I_x and I_y) was calculated for each trial. Twenty-four chi-square goodness-of-fit tests were per performed to determine if I_x and I_y were equally distributed among participants for each of the twelve factor combinations. The results of the chi-square tests can be found in *Table 2*. The distribution of I_x and I_y were unequal among participants for a few of the conditions with only 4 objects. The significance of the chi-square tests in some of the four-objects conditions was probably due to the index of dispersion being most reliable when $m > 6$. Any inequalities in the dispersion of the four objects between participants was unlikely to have an impact on their performance in the easiest of conditions.

Table 2: Results of the chi-square goodness-of-fit tests performed to determine whether the Index of Dispersion along the X and Y axis was equally distributed among participants (N=30) for each factor combination. A p-value < 0.05 indicates an unequal distribution among participants.

Factors		I_x			I_y		
Targets	Entropy	χ^2	p	sig.	χ^2	p	sig.
4	Low	48.25	0.019	**	22.79	0.824	
4	Medium	43.48	0.053	.	35.89	0.212	
4	High	74.73	0.000	***	65.81	0.000	***
9	Low	3.76	1.000		4.19	1.000	
9	Medium	7.55	1.000		4.59	1.000	
9	High	6.16	1.000		2.99	1.000	
14	Low	2.71	1.000		1.94	1.000	
14	Medium	1.96	1.000		1.37	1.000	
14	High	2.68	1.000		1.28	1.000	
19	Low	0.67	1.000		1.15	1.000	
19	Medium	0.91	1.000		1.08	1.000	
19	High	1.26	1.000		0.75	1.000	

Analysis of Variance

A table of means and standard deviations of completion time for each of the factor combinations can be found in *Table 3*.

Table 3: The means and standard deviations of completion time in miliseconds for each of the 12 experimental blocks, (N=155).

	No Entropy		Medium Entropy		High Entropy	
Objects	Mean (ms)	Std.Dev.	Mean (ms)	Std.Dev.	Mean (ms)	Std.Dev.
4	3121.67	1318.15	3137.95	1475.03	3086.85	1199.30
9	6347.84	3081.58	6352.80	3521.23	5471.20	2835.46
14	9321.14	5586.84	9057.77	5100.11	8798.21	5581.65
19	11413.81	6336.45	11375.95	6547.51	10758.10	6292.93

Planned Contrasts. A planned linear contrast of the targets factor using the lambda weight assignments of -3, -1, 1, and 3 for 4, 9, 14, and 19 targets respectively was performed for completion time. As predicted, the contrast for the targets was significant, t(30)=24.22, p<0.001, r=0.98, 1-beta=1. A planned linear contrast of the entropy factor using the lambda weight assignments of -1, 0 and 1 for the none, medium, and high entropy levels respectively was performed for completion time. As predicted, the contrast for entropy was significant, t(30)=-2.23, p=0.018, r=0.38. While the effect of entropy on completion time was significant, it was to the opposite direction of what was predicted; completion times were faster in the high entropy than in the no entropy conditions.

Omnibus ANOVA. A three-way (Number of Objects x Entropy x Replicate) within-subjects analysis of variance was performed, see *Table 4*. The main effect of number of objects was significant, F(3,90)=338.65, p<0.001. The main effect of entropy was marginally significant, F(2,60)=2.70, p=0.075. Neither the main effect of replicate or the interaction between number of objects and entropy were significant.

Table 4: Analysis of Variance: 4x3x5 Within-Subjects

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Error: Participants					
Residuals	30	26.9820	0.8994		
Error: Participants:Objects					
Objects	3	363.79	121.26	338.65	< 0.001
Residuals	90	32.23	0.36		
Error: Participants:Entropy					
Entropy	2	2.1633	1.0817	2.703	0.07516
Residuals	60	724.010	0.4002		
Error: Participants:Replicates					
Replicates	4	0.378	0.094	0.3173	0.866
Residuals	120	35.727	0.298		
Error: Participants:Objects:Entropy					
Objects:Entropy	6	1.283	0.214	0.6844	0.6624
Residuals	180	56.228	0.312		
Error: Within					
Residuals	1364	361.87	0.27		

Analysis of Covariance. It is possible that one of the random factors present in the experimental design affected completion time and thus reduced the effect size of entropy. In order to make sure that the randomization of initial object locations or the randomization of block order didn't have a significant impact on completion time, the goodness-of-fit of four mixed-effects models were compared; see *Table 5* for the degrees of freedom and AIC value for each model. Model 1 is the mixed-effects equivalent of a within-subjects ANOVA

Table 5: In order to determine if block order or the dispersion of initial object locations had an impact on completion time, mixed-effects models with varying covariates were created. *Model 1* is equivalent to a within-subjects ANOVA model. Model 2 included block order as a covariate; block order did not improve the quality of fit of the model in comparison to model 1. Models 3 and 4 included the index of dispersion measures as covariates; their presence did not improve the quality of fit of the model in comparison to model 1.

Model	Fixed Factors	Covariate	Random Factors	Df	AIC
1	Objects, Entropy		Participants	8	2995
2	Objects, Entropy	Block Order	Participants	9	3005
3	Objects, Entropy	I_x	Participants	9	3009
4	Objects, Entropy	I_y	Participants	9	3008

(*compare APPENDIX RCode 2 and 3*); number of objects and entropy were included as fixed factors and participants was included as a random factor. Model 2 is the same as model 1 with the addition of block order as a covariate. Models 3 and 4 are the same as model 1 with the addition of I_x and I_y as covariates respectively. The goodness-of-fit for models 2, 3 and 4 were worse than the model 1 which suggests that the randomization of block order and initial object locations did not influence completion time.

Entropy and Displacement Per Second

Since entropy had an inverse relationship with average displacement per second, the marginal significance of the entropy factor could be related to an unknown factor influencing displacement. Speed and the rate of direction change were held constant in the present study, so the only changes in average displacement per second should have come from the entropy factor. In order to compare the effect size of entropy to average displacement per second, two mixed-effects regressions were performed; both regressions included participants as a random factor.

The first regression (model 1) included number of objects and entropy as fixed continuous factors. The second regression (model 2) included number of objects and displacement per second as fixed continuous factors. In order to estimate the average displacement per second of each entropy level, the experimental program was modified to spawn one object at the center of the screen, then for 20 trials of each of the three levels of entropy, the object would move for 5 seconds in a random direction. After each trial, the displacement of the object from the center of the screen was calculated, divided by 5, then recorded. The average displacement per second (DPS) for the none, medium and high entropy conditions was 86.87, 83.83, 74.56 pixels per second respectively. Pixels per second can additionally be converted to degrees of visual angle. Based on an average viewing distance of 16 inches and 72 pixels per inch, 86.87, 83.83, and 74.56 pixels per second is equivalent to 4.32, 4.17, and 3.71 degrees of visual angle per second. This reaffirms that as entropy increases the displacement of the object decreases. Parameters estimates, standard errors, significance values and effect sizes are shown in *Table 6*.

Table 6: The probability and effect size values were calculated using a $df = 30$. The goodness-of-fit measured by AIC for model 1, 2 and 3 were 3046, 3048, and 3043.

Model	Parameter	Estimate	Std. Error	t value	p	r
5	(Intercept)	7.8105	0.03781	206.57	0.000	0.99
	Number of Objects	0.0770	0.00224	34.41	0.000	0.98
	Entropy	-0.0689	0.02993	-2.30	0.028	0.38
6	(Intercept)	7.2403	0.19810	36.55	0.000	0.98
	Number of Objects	0.0770	0.00224	34.43	0.000	0.98
	Displacement Per Second	0.0065	0.00239	2.72	0.011	0.44
7	(Intercept)	7.6324	0.19810	168.22	0.000	0.99
	Number of Objects	0.0770	0.00224	34.68	0.000	0.98
	Displacement Per Second	0.0032	0.00252	1.28	0.210	0.22
	Final Target Displacement	0.0002	0.00005	3.90	0.000	0.57

The results of these regressions confirm the inverse relationship between entropy and displacement per second. The relative similarity in significance and effect size suggest that entropy and displacement per second both measure the same aspect of an object's motion; see *Figure 5(a)* and *Figure 5(b)*. Both entropy and displacement per second quantify the amount of uncertainty associated with an object independent of the object's speed. In light of average displacement per second having a slightly larger effect size (r=0.44) than entropy(r=0.38), the final displacement of the target object was included as a covariate in addition to displacement per second and a third regression was performed. The effect size target displacement (r=0.57) was much larger than the average displacement per second (r=0.22).

A closer look at target-object displacement shows an odd interaction with the 14-object condition, see *Figure 6(a)* and *Figure 6(b)*. Target displacement should follow the same trend as average displacement, but there is an obvious deviation. This interaction is probably what caused the marginal significance of the entropy factor.

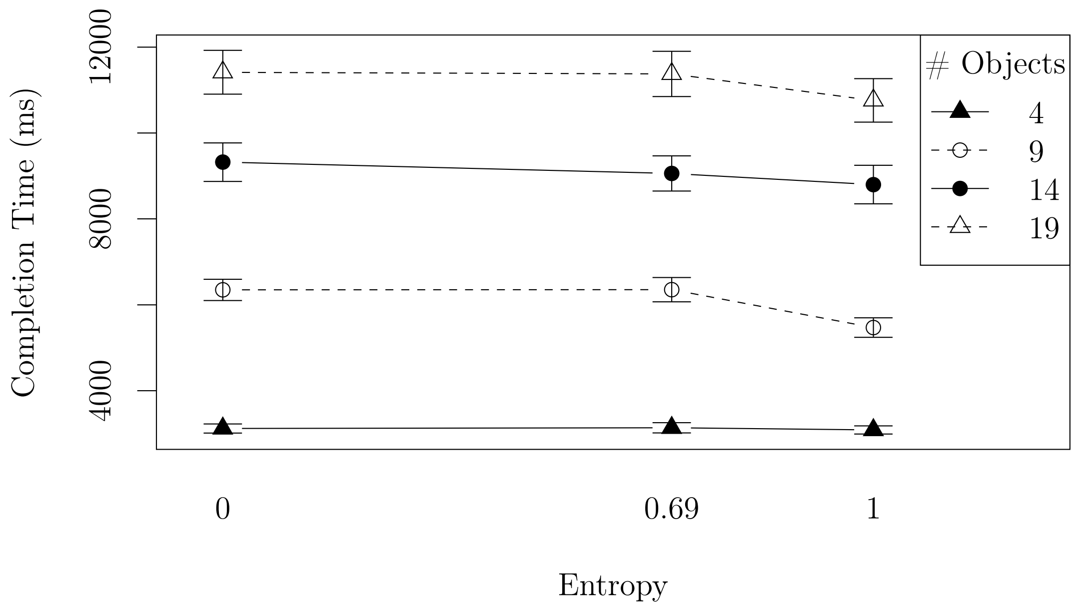
Model Fitting

The completion time data was fitted to *Equation* (4), which includes 5 parameters, two of which, t and m, are free parameters. The parameters that were not free were fixed (P_{guess}) or derived from the structure of the experiment (n). The parameter n was set to 4, 9, 14, or 19. The probability of guessing was set to reflect guessing strategy no better than chance, $P_{guess} = 0.5$. With respect to the free parameters, the static binding capacity m was assumed to be approximately 4 (consistent with Oksama and Hyönä (2008) and Pylyshyn and Storm (1988)) and a plausible range of 500-1000 ms for the time spend processing each object t, which is consistent with the time required to fixate on and recognize a static object, serially shift attention between objects (Oksama & Hyönä, 2008) and physically move the mouse cursor. The best fitting parameters and confidence intervals are shown in *Table 7*. *Model 8* fixed P_{guess} at 0.05, m and t were left as free param-

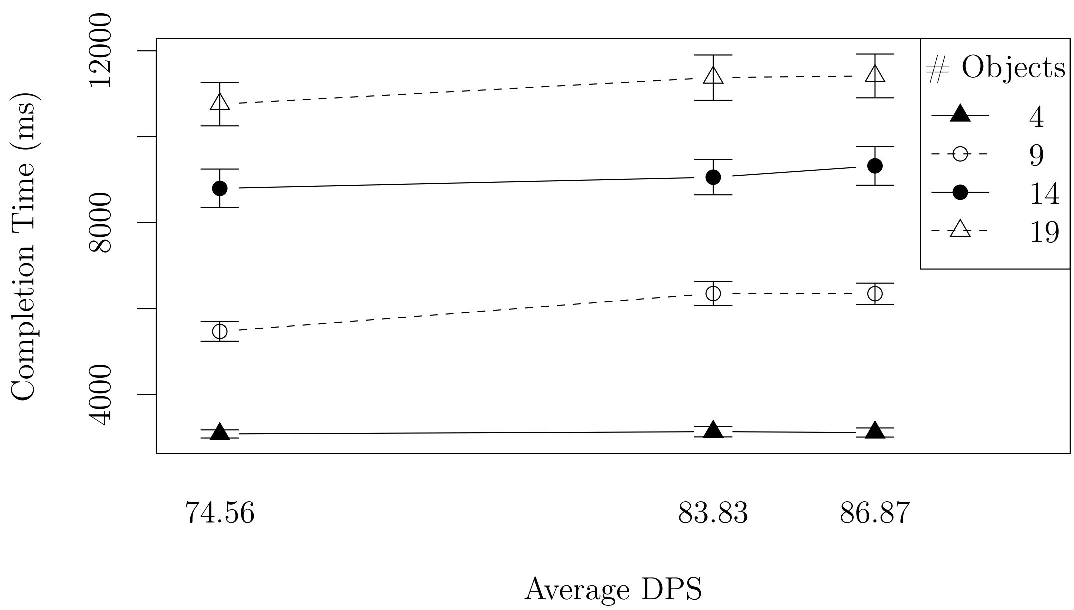
Table 7: Non-linear least squares regression models based on *Equation* (4).

Model	Parameter	Fixed Value	Estimate	Std. Error	t value	p	95% CI Lower	95% CI Upper	AIC
8	P_{guess}	0.5							2989
	m		5.50	0.51	10.76	0.000	4.57	6.59	
	t		0.82	0.03	27.58	0.000	0.76	0.88	
9	P_{guess}		0.42	0.02	18.59	0.000	0.38	0.47	2989
	m	4							
	t		0.97	0.02	42.74	0.000	0.93	1.02	
10	P_{guess}		0.41	0.01	27.58	0.000	0.38	0.44	2989
	m		3.83	0.16	24.33	0.000	3.51	4.13	
	t	1							

eters. *Model 9* fixed m at 4, P_{guess} and t were left as free parameters. *Model 10* fixed t at 1, P_{guess} and m were left as free parameters. The fit these models which was based on equation (4) provided a better fit to the completion time data (AIC 2989) than linear regression models (AIC>=2995) from *6*. Additionally, the best fitting values for P_{guess}, m and $\bar{s}$ are both psychologically plausible and consistent with past research. *Figure 7* shows a comparison of fit between the MOMIT based model and a standard linear model.

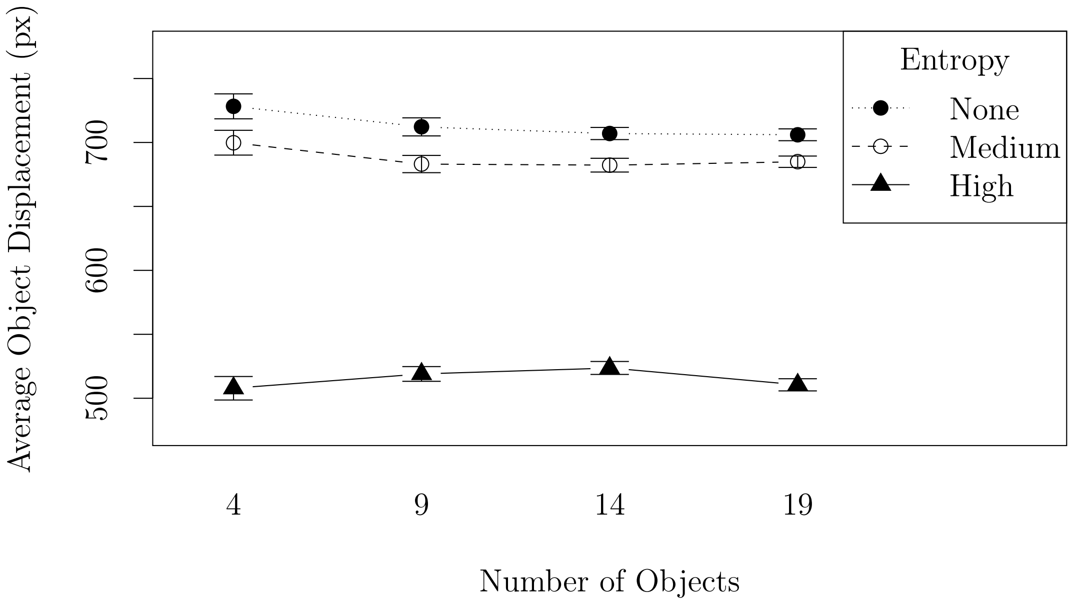


(a) Completion time with respect to entropy for each level of number of objects.

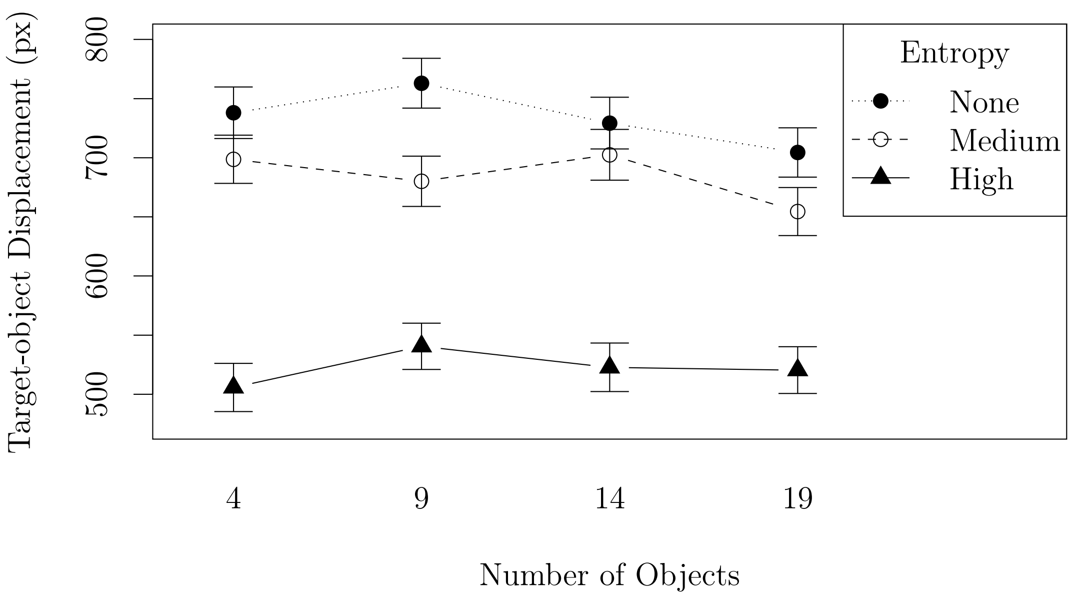


(b) Completion time with respect to average displacement per second for each level of number of objects.

Figure 5: Completion time with respect to entropy (a) and average displacement per second (b) for each level of number of objects.



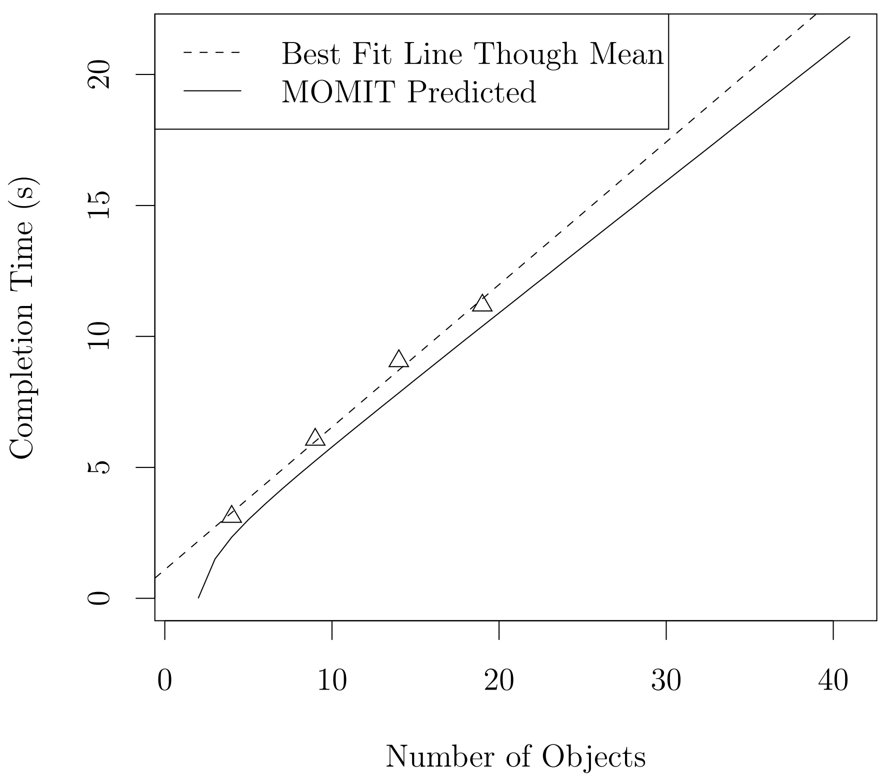
(a) Average displacement of all objects with respect to number of objects for each level of entropy.



(b) Target-object displacement with respect to number of objects for each level of entropy.

Figure 6: The average displacement of all objects (a) and the dispacement of the target-object only (b) for each level of number of objects.

Figure 7: The dashed line represents the best fit line though the mean completion time for each level of number of objects (represented by the triangle points). The solid line represents the completion time predicted by equation (4).



Discussion

The present study investigated an observer's ability to track and maintain multiple uniquely identified objects in a dynamic environment similar to air traffic control. The following findings were found: (1) a highly significant object set-size effect was observed; completion time increased as a function of number of objects, (2) a marginally significant effect of entropy was observed; completion time decreased as a function of entropy, and (3) a significant effect of displacement per second was observed; completion time increased as a function of displacement per second.

The mathematical reformulation of MOMIT to predict completion time captured both the effect of object set-size and binding capacity and provided a decent quantitative fit to the data, see *Figure 7.* The main aspects of the model are that continuous attention switching along with the storage of the spatial indexes in VSTM, bindings stored in the episodic buffer, and LTM representations are intimately involved the creation and maintenance of identity-location bindings. Based on these principles, MOMIT correctly predicted the effects of object set-size and displacement per second. A look at the raw data seems to indicate that participants spend more time per object when there are a few objects and have a higher accuracy in finding the target object on the first try. This observation further supports participants using a mental model to track multiple moving objects when the number of total objects is close to the binding capacity.

The generalizability of the observed effects

The observation that tracking performance deteriorates as a function of object set-size is a finding consistent with numerous other studies with similar tracking tasks (Yantis, 1992; Pylyshyn & Storm, 1988; Oksama & Hyönä, 2008). The influence of entropy, however, is a phenomenon that has not been systematically manipulated or studied. The finding that performance increased with increases in entropy is opposite of hypothesis (2) and does not support Finke and Shyi (1988) concept of *representational momentum.* This observation could be explained a number of ways. The strongest explanation is that this type of task is simply too difficult to develop higher-order representations such as velocity vectors. This observation could also be a result of the relatively short duration of each trial; maybe there was not enough time to develop higher-order representations in the trials with large object set-sizes.

The lack of evidence supporting the prediction of object locations could also be caused by the nature of the task. It is possible that either the experimental task did not require participants to anticipate where objects would be in the future, or the task was too difficult for novices to actually anticipate future locations of each object.

This opposite relationship of entropy and performance however can be explained by the tight relationship of entropy and displacement per second. Past studies have shown that velocity affects tracking performance (ex. Saiki (2002); Oksama and Hyönä (2008)) and if the objects do not change direction while moving, then velocity is a measure of displacement per second. However, the objects did change direction, so velocity was really measuring distance per second, not displacement per second. The systematic manipulation of displacement through changes in entropy, all while keeping velocity constant, is a unique contribution of the present study to the field of object tracking research. The small effect size of the entropy manipulations is most likely due to the relatively small differences in the displacement per second once converted to degrees of visual angle. The difference in displacement per second between the low and high entropy conditions was only 0.61 degrees of visual angle per second. In comparison, Oksama and Hyönä (2008) used velocity manipulations that resulted in differences in velocity as large as 9.7 degrees of visual angle per second; the smallest difference in velocity was 3.9 degrees of visual angle per second.

Contributions to MOMIT

The relationship between displacement per second and the location error component of MOMIT is quite evident. The lower the displacement per second, the closer an object stays to its last known location. Thus, the location error will be lower with high entropy than it will be with low entropy. This fits well with location error component from the original MOMIT. Since the concept of displacement captures not only changes in velocity but changes in direction of travel, average displacement per second is a more robust measure of the uncertainty of an object from motion than speed. This distinction should be made more clear and MOMIT should be adjusted to use average displacement per second instead of distance per second in the location error component. The present paper also provides a formula based on the components of MOMIT which predicts reaction time instead of accuracy.

Application to ATC

The predictive utility of MOMIT with respect to air traffic controller performance seems to the strong. The past MOT research appears to hold up on tasks without distracters (something that was unclear at the start of this experiment). Even in the presence of lots of noise, a result of both individual differences in tracking performance and the random nature of the experimental design, equation (4) still had a decent fit to the data. In addition, the new experimental paradigm created in this study can be used as a base for future research.

Design Issues

Despite the main effect of entropy being significant, there was an issue with the design of the experiment that could influence the true effect size of the entropy factor. When an objects motion resulted with a collision with the edge of the screen, the objects motion was reflected like a mirror. This caused the total displacement of the object, from the moment of reflection (and/or until any other direction change), to decrease instead of increase. This factor is most likely confounded with the entropy factor since both end up affecting the total displacement of an object and is most likely the cause of the odd interaction in the 14-object conditions. Another potential confound of this study comes from the familiarity of the callsigns used for objects. Oksama and Hyönä (2008) showed that familiarity facilitates tracking performance. It is possible that the callsign prefix, the 3-character airline designator, could be more familiar to one participant than another; this was not controlled for in the present study.

Conclusion

All in all, the new experimental paradigm used in the present study provides a MOT task that is similar to a very basic ATC task which should be suitable for use in future studies involving multiple object tracking performance without distracters. Performance on the experimental task of the present study was also consistent with the results of Oksama and Hyönä (2008). The decent fit of MOMIT to the experimental data supports the use of MOMIT to predict ATC performance.

References

Alvarez, G., & Cavanagh, P. (2005). Independent resources for attentional tracking in the left and right visual hemifields. *Psychological Science*, *16(8)*, 637-643.

Bates, D., Maechler, M., & Dai, B. (2008). lme4: Linear mixed-effects models using s4 classes [Computer software manual]. Available from `http://lme4.r-forge.r-project.org/` (R package version 0.999375-28)

Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, *97*, 523-547.

Bundesen, C., Habekost, T., & Kyllingsbaek, S. (2005). A neural theory of visual attention: Bridging cognition and neurophysiology. *Psychological Review*, *112*, 291-328.

Carmody, M., & Gluckman, J. (1993). Task specific effects of automation and automation failure on performance, workload, and situation awareness. In R. Jensen & D. Neumeister (Eds.), (p. 167-171). Columbus, OH: Department of Aviation, The Ohio State University.

Cavanagh, P., & Alvarez, G. (2005). Tracking multiple targets with multifocal attention. *Trends in Cognitive Sciences*, *4*, 417-423.

Charlton, S. G., & O'Brien, T. G. (2002). *Handbook of human factors testing and evaluation.* Lawrence Erlbaum Associates.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Crawley, M. J. (2007). *The r book.* West Sussex, England: John Wiley and Sons Ltd.

Diggle, P. J. (1983). *Statistical analysis of spatial point patterns.* New York, New York: Academic Press.

Endsley, M., & Garland, D. (2000). *Situation awareness analysis and measurement.* Mahwah, New Jersey.

Endsley, M., & Jones, D. (1996). Situation awareness in air traffic control. In A. O. . G. Salvendy (Ed.), *Advances in applied ergonomics. proceedings of the first international conference of applied ergonomics* (p. 394-397). West Lafayette, IN: USA Publishing.

Endsley, M., & Kiris, E. (1995). The out-of-the-loop performance problem of level and control in automation. *Human Factors*, *37(2)*, 381-394.

Endsley, M., Mogford, R., Allendoerfer, K., Snyder, M., , & Stein, E. (1997). *Effect of free flight conditions on controller performance, workload, and situation awareness (dot/faa/ct-tn97/12)* (Tech. Rep.). Atlantic City International Airport: Fedral Aviation Administration William J. Hughes Technical Center.

Endsley, M., & Rodgers, M. (1996). Attention distribution and situational awareness in air traffic control. *Proceedings of the 40th Annual Meeting of the Human Factors and Ergonomics Society*, 82-85.

Finke, R. A., & Shyi, G. C.-W. (1988). Mental extrapolation and representational momentum for complex implied motions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 112-120.

Freyd, J. J., & Finke, R. A. (1984). Representational momentum. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*, 126-132.

Humphreys, G. (1996). Neurophysiological aspects of visual attention and eye movements . a synopsis. In *Visual attention and cognition* (p. 73-78). Elsevier Science.

Humphreys, G., Gilchrist, I., & Free, L. (1996). Search and selection in human vision: Physiological evidence and computational implications. In *Visual attention and cognition* (p. 79-93). Elsevier Science.

Iglewicz, S., & Hoaglin, D. (1993). *How to detect and handle outliers.* Milwaukee: ASQC Quality Press.

Keane, B. P., & Pylyshyn, Z. W. (2006). Is motion extrapolation employed in multiple object tracking? tracking as a low-level, non-predictive function. *Cognitive Psychology*, *52*, 346-368.

Logan, D. (1996). The code theory of visual attention: An integration of space-based and object-based attention. *Psychological Review*, *103(4)*, 603-649.

Logan, D. (2002). An instance theory of attention and memory. *Psychological Review*, *109*, 376-400.

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*, *43(4)*.

Mogford, R. (1997). Mental models and situation awareness in air traffic control. the international journal of aviation psychology. , *7(4)*, 331-341.

Oksama, L., & Hyönä, J. (2008). Dynamic binding of identity and location information: A serial model of multiple identity tracking. *Cognitive Psychology*, *56*, 237-283.

Pylyshyn, Z. (2005). *Seeing and visualizing: It's not what you think.* Cambridge, MA.

Pylyshyn, Z., & Annan, V. (2006). Dynamics of target selection in multiple object tracking (mot). *Spatial Vision*, *19(6)*, 485-504.

Pylyshyn, Z., & Storm, R. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, *3*, 179-197.

R Development Core Team. (2008). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Available from `http://www.R-project.org` (ISBN 3-900051-07-0)

Rosenthal, R., & Rosnow, R. L. (2008). *Essentials of behavorial research: Methods and data analysis* (3rd ed.). New York, NY: McGraw-Hill.

Saiki, J. (2002). Multiple-object performance tracking: Limitation in maintenance and transformation of perceptual objects. In J. Hyönä, W. H. D.P. Munoz, & R. Radach (Eds.), *The brain's eye: Neurobiological and clinical aspects of oculomotor research* (p. 133-148). Amsterdam: Elsevier Science.

Sakamoto, Y., Ishiguro, M., & Kitagawa, G. (1986). *Akaike information criterion statistics.* D. Reidel Publishing Company.

Shannon, C. (1948). A mathematical theory of communication. *Bell System Technical*

Journal, *27*, 379-423,623-656.

Treisman, A. (1998). Features and objects: The fourteenth barlett memorial lecture. *Quarterly Journal of Experimental Psychology*, *40(A)*, 201-237.

Wickens, & Hollands, J. (2000). *Engineering psychology and human performance* (3r ed.). New Jersey: Prentice Hall.

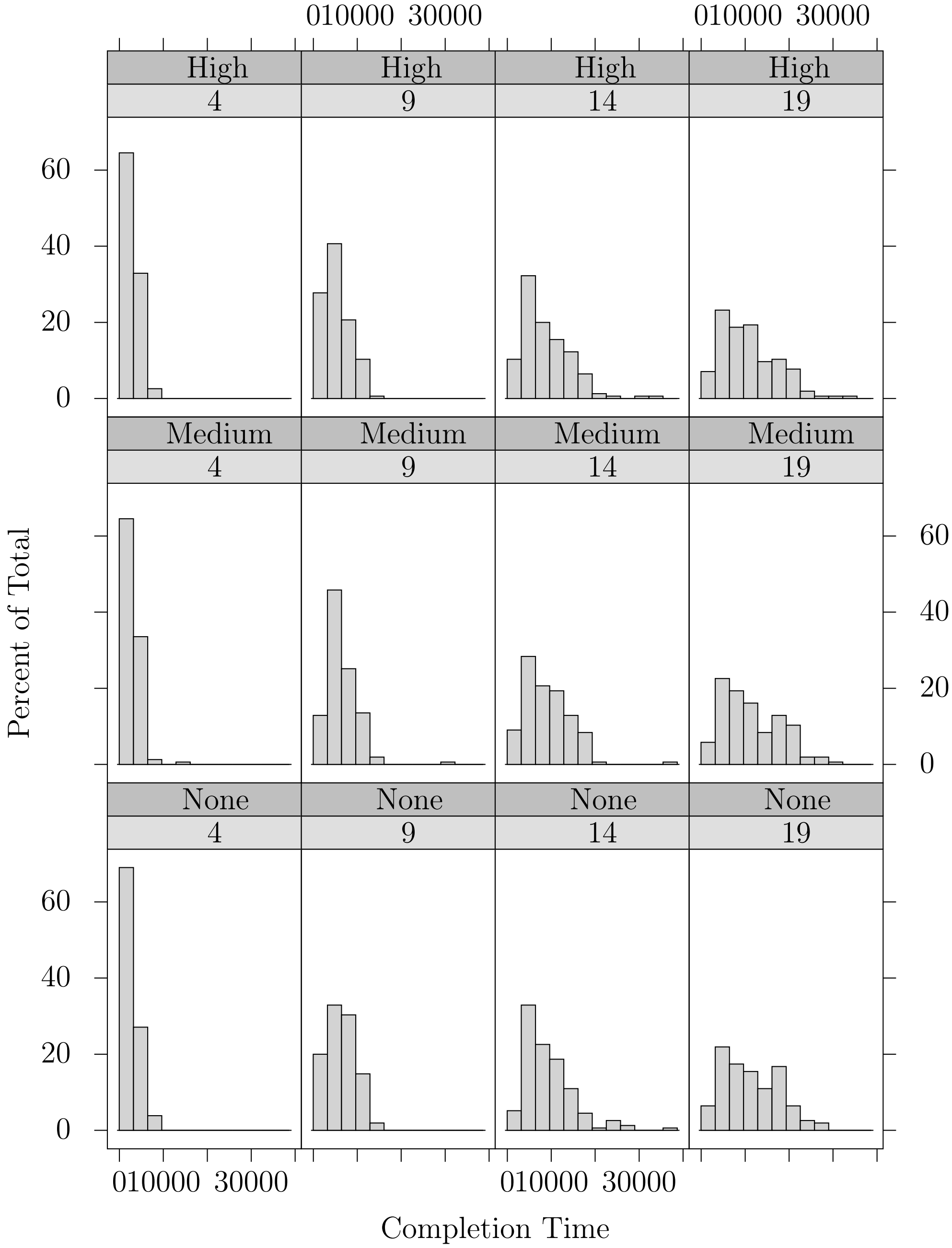
Wickens, Mavor, A. S., & McGee, J. (1997). *Flight to the future: human factors in air traffic control*. Washington, D.C.: National Academy Press.

Wickens, Mavor, A. S., Parasuraman, R., & McGee, J. (1998). *The future of air traffic control: human operators and automation.* Washington, D.C.: National Academy Press.

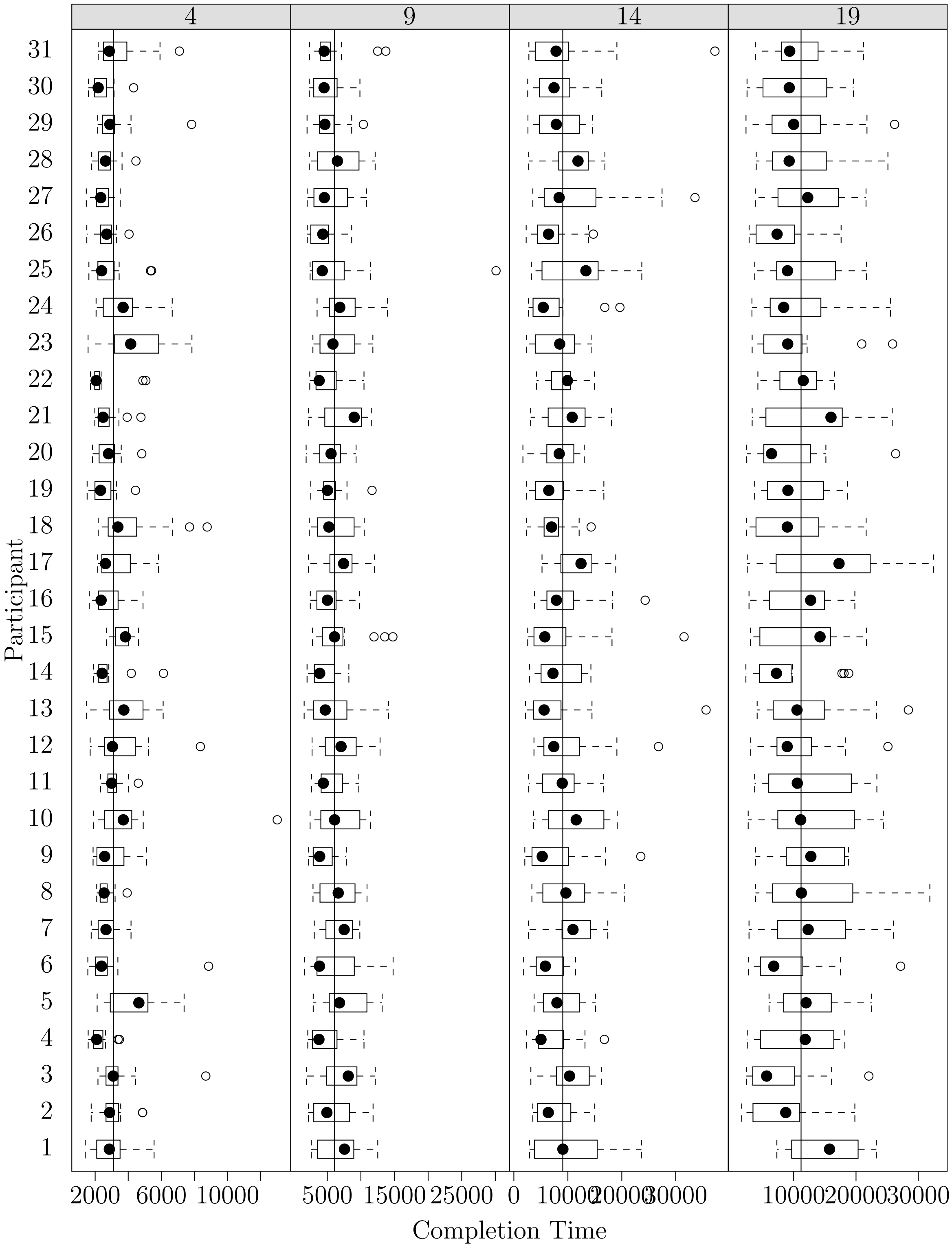
Yantis, S. (1992). Multielement visual tracking: Attention and perceptual organization. *Cognitive Psychology*, *24*, 295-340.

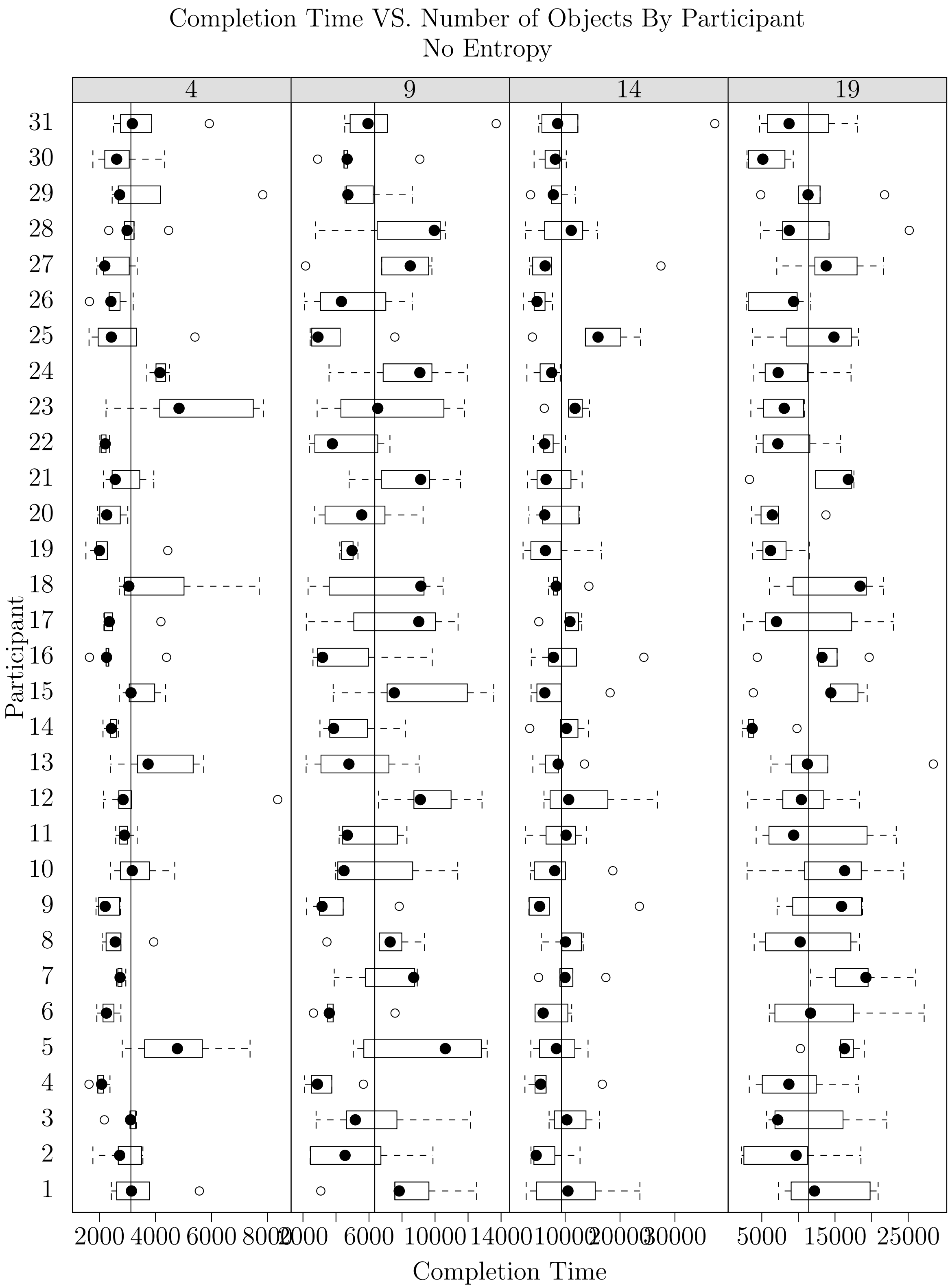
Zelinsky, G., & Sheinberg, D. (1997). Eye movements during parallel-serial visual search. *Journal of Experimental Psychology*, *23(1)*, 244-262.

APPENDIX A: Figures

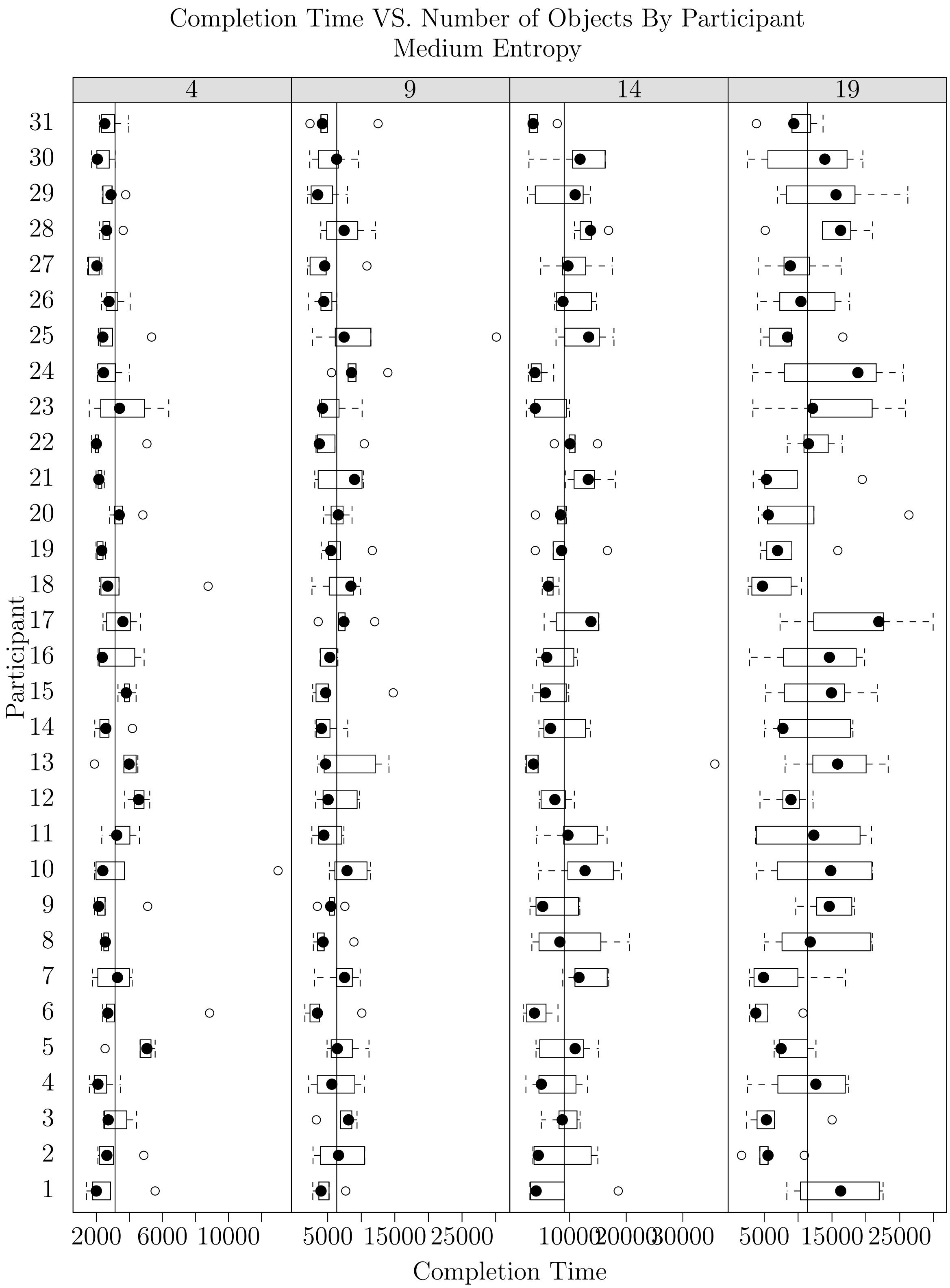


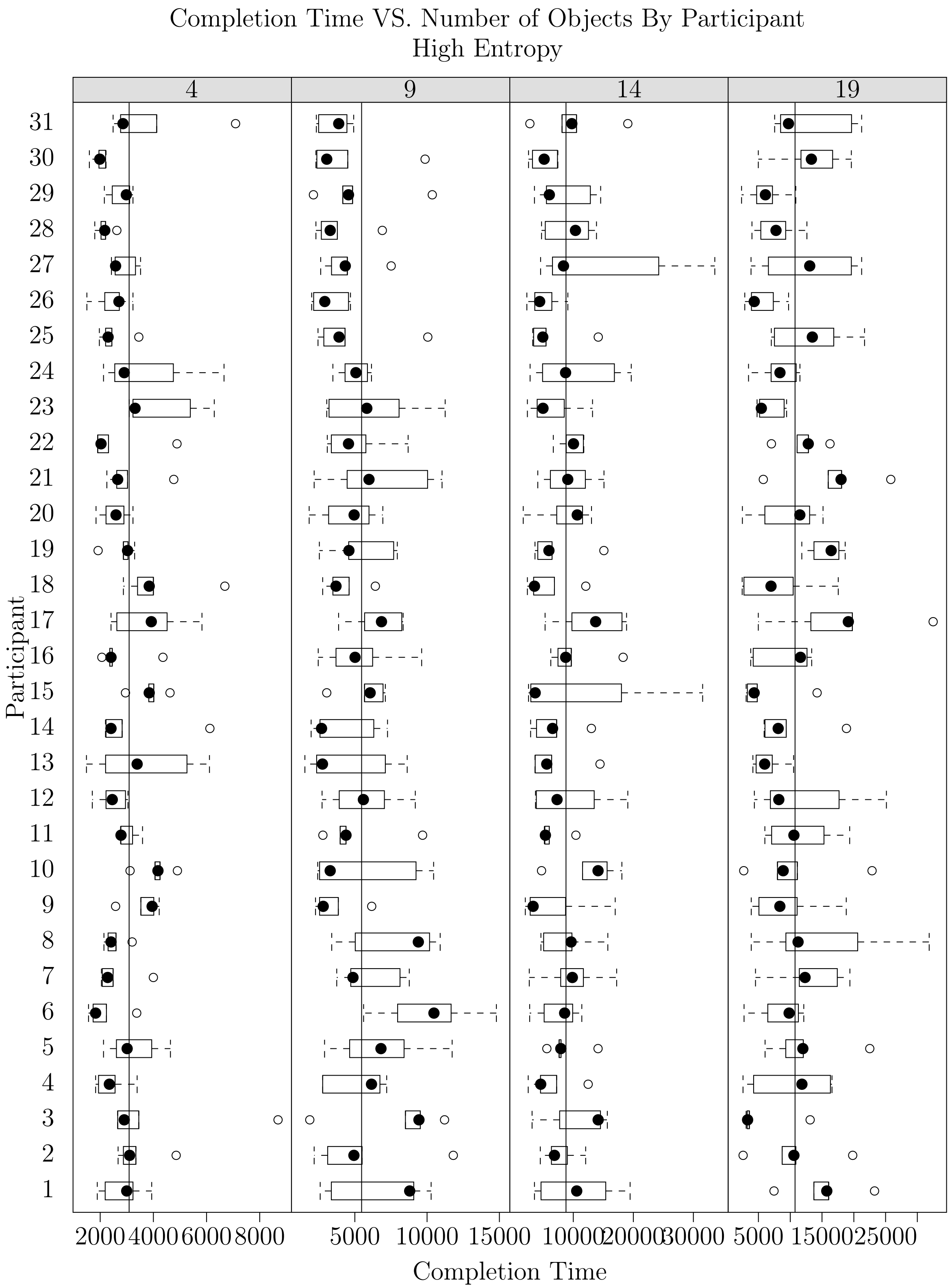
Completion Time VS. Number of Objects By Participant





Completion Time VS. Number of Objects By Participant
No Entropy
4
9
14
19
Participant
31
30
29
28
27
26
25
24
23
22
21
20
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
2000
4000
6000
8000
2000
6000
10000
14000
10000
20000
30000
5000
15000
25000
Completion Time


Completion Time VS. Number of Objects By Participant
Medium Entropy
4
9
14
19
Participant
31
30
29
28
27
26
25
24
23
22
21
20
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
2000
6000
10000
5000
15000
25000
10000
20000
30000
5000
15000
25000
Completion Time



Completion Time VS. Number of Objects By Participant
High Entropy
4
9
14
19
Participant
31
30
29
28
27
26
25
24
23
22
21
20
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
2000 4000 6000 8000
5000 10000 15000
10000 20000 30000
5000 15000 25000
Completion Time

APPENDIX B: RCode

1. Datafile

```
> datafile = "http://people.rit.edu/rmh3093/master.csv"
> master = read.table(datafile, header = T)
> master$Entropy = factor(master$Entropy, ordered = T, levels = c("Low",
+     "Medium", "High"))
> master$Objects = as.ordered(master$Targets)
> master$Replicates = as.factor(master$Replicate)
> master$BlockOrder = as.factor(master$Block_Order)
> master$Participants = as.factor(master$Participant_ID)
> names(master)
```

```
 [1] "Participant_ID"       "Trial_Order"          "Block_Order"
 [4] "Replicate"            "BlockN"               "BlockC"
 [7] "Targets"              "VA"                   "VA2"
[10] "DPS"                  "E"                    "Entropy"
[13] "H"                    "X_Offset"             "Y_Offset"
[16] "X_Variance"           "Y_Variance"           "D_x"
[19] "D_y"                  "SNR_x"                "SNR_y"
[22] "I_x"                  "I_y"                  "Variance_Ratio"
[25] "Offset_Ratio"         "Average_Displacement" "Target_Displacement"
[28] "AD.TD"                "TD.AD"                "Completion_Time"
[31] "Visited_Targets"      "CT.VT"                "CT.T"
[34] "SP"                   "Resp1"                "SA1"
[37] "SA2"                  "SA3"                  "SA4"
[40] "T1"                   "T2"                   "Objects"
[43] "Replicates"           "BlockOrder"           "Participants"
```

```
> summary(aov(log(Completion_Time) ~ Objects * Entropy + Replicates +
+     Error(Participants/(Objects * Entropy + Replicates)), master))

Error: Participants
          Df  Sum Sq Mean Sq F value Pr(>F)
Residuals 30 26.9820  0.8994

Error: Participants:Objects
          Df Sum Sq Mean Sq F value    Pr(>F)
Objects    3 363.79  121.26  338.65 < 2.2e-16 ***
Residuals 90  32.23    0.36
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Error: Participants:Entropy
          Df  Sum Sq Mean Sq F value  Pr(>F)
Entropy    2  2.1633  1.0817   2.703 0.07516 .
Residuals 60 24.0104  0.4002
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Error: Participants:Replicates
            Df Sum Sq Mean Sq F value Pr(>F)
Replicates   4  0.378   0.094  0.3173  0.866
Residuals  120 35.727   0.298

Error: Participants:Objects:Entropy
                 Df Sum Sq Mean Sq F value Pr(>F)
Objects:Entropy   6  1.283   0.214  0.6844 0.6624
Residuals       180 56.228   0.312

Error: Within
            Df Sum Sq Mean Sq F value Pr(>F)
Residuals 1364 361.87    0.27
```

3. Mixed Effects

```
> m1 = lmer(log(Completion_Time) ~ Objects + Entropy + (1 | Participants),
+     master)
> anova(m1)
```

```
Analysis of Variance Table
        Df Sum Sq Mean Sq  F value
Objects  3 363.79  121.26 432.2429
Entropy  2   2.16    1.08   3.8556
```

APPENDIX C: IRB Forms

R·I·T

Rochester Institute of Technology

RIT Institutional Review Board for the
Protection of Human Subjects in Research
141 Lomb Memorial Drive
Rochester, New York 14623-5604
Phone: 585-475-7673
Fax: 585-475-7990
Email: hmfsrs@rit.edu

Form C
IRB Decision Form

TO: Ryan Hope; Esa Rantanen

FROM: RIT Institutional Review Board

DATE: **October 20, 2008**

RE: **Decision of the RIT Institutional Review Board**

Project Title – The Predictive Utility of the Model of Multiple Identity Tracking (MOMIT) in Air Traffic Control Performances

The Institutional Review Board (IRB) has taken the following action on your project named above.

☒ Approved, no greater than minimal risk

Now that your project is approved, you may proceed as you described in the Form A. **Note that this approval is only for a maximum of 12 months; you may conduct research on human subjects only between the date of this letter and 10/20/09**

You are required to submit to the IRB any:

- **Proposed** modifications and wait for approval before implementing them,
- Unanticipated risks, and
- Actual injury to human subjects.

Return the Form F, at the end of your human research project or 12 months from the above date. If your project will extend more than 12 months, your project must receive continuing review by the IRB.

Heather Foti
Associate Director, Office of Human Subjects Research

Revised 10-18-06

INFORMED CONSENT FORM

Rochester Institute of Technology

Title of Project: The predictive utility of the Model of Multiple Identity Tracking in Air Traffic Control Performance

Investigators in Charge:

Mr. Ryan Hope
MS Candidate
Dept. of Psychology.
Rochester Inst. of Technology
Tel. (716) 308-1835
Email: rmh3093@rit.edu

A. Explanation of the Project.

1. You are being asked to participate in an experiment that seeks to quantify the level of situation awareness you aquire while performing a multiple object tracking task.

2. Situation awareness is the perception of environmental elements within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future.

3. This research will be used to test and refine a mathematical model which can assist air traffic controllers by predicting when their situation awareness falls below a critical level.

4. During this study, you will monitor a computer screen and track airplane callsigns (groups of 3 letters and 3 numbers) as they move around on the screen. At the end of each trial, the identity of the targets will be masked and you will be asked to click on a particular target. When you move the mouse pointer of a particular callsign it will become unmasked and its original identity will be revealed.

5. This experiment should take about 45 minutes to complete.

6. This research poses no risk to you beyond that of which you would normally experience operating a computer.

7. As a benefit of participating in this experiment, your future air travel experiences could be safer.

B. Your rights as a research participant

1. I will be happy to answer any questions you have about the study at any time. Mr. Hope may be contacted at the telephone number and e-mail addresses shown above. If you have questions about your rights as a research subject, you can call collect the Rochester Institute of Technology Institutional Review Board at (585) 475-7673, or e-mail hmfsrs@rit.edu.

2. No subsequently published results will contain any information that could be associated with individual participants. All data will be stored and secured only on the investigator's computer.

3. Your participation is wholly voluntary. Your decision to participate, or to not participate, or to withdraw from the study during the experiment will in no way influence your relationship with the researcher or your co-workers.

4. You may refuse to participate or may discontinue participation at any time during the project without penalty or loss of benefits to which you are otherwise entitled.