Relational information between objects is available to guide search.

A Dissertation Presented

by

Joseph Charles Schmidt

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Experimental Psychology

Stony Brook University

May 2012

Stony Brook University
The Graduate School

Joseph Charles Schmidt

We, the dissertation committee for the above candidate for the
Doctor of Philosophy degree, hereby recommend
acceptance of this dissertation.

Gregory J. Zelinsky
Associate Professor Cognitive Science & Computer Science

Susan E. Brennan
Professor of Cognitive Science

Hoi-Chung Leung
Associate Professor of Biopsychology

Daniel Casasanto
Assistant Professor of Psychology
The New School for Social Research

This dissertation is accepted by the Graduate School

Charles Taber
Interim Dean of the Graduate School

Abstract of the Dissertation

Relational information between objects is available to guide search.

by

Joseph Charles Schmidt

Doctor of Philosophy

in

Experimental Psychology

Stony Brook University

2012

Objects in the real world exist relative to other objects, resulting in an intricate web of spatial relationships. Do we use this relational information when we search for objects? Current search theory suggests that object relationships can only be established using focal attention (Logan, 1994; 1995). If this is true, pre-attentive search guidance by relational information should be impossible. In a series of seven experiments, I demonstrate that search guidance by relational information *is possible*, even in the absence of real-world contextual constraints that may magnify relational guidance. Experiment 1 shows search guidance by relational information only, i.e. in the absence of target feature guidance. Experiment 2 indicates that relational guidance is evident in highly heterogeneous displays as well. Experiment 3 demonstrates that relational guidance does *not* affect search when targets are cued using text labels referring to four object classes, suggesting that the effective coding of relational information may require highly specific target features. Experiment 4 shows that relational guidance is selectively *not* expressed when functional relationships between objects are contrary to real-world expectations (e.g. a hammer *below* a nail), suggesting that relational guidance is affected by object spatial associations in long-term memory. Experiment 5 further demonstrates that with minimal practice there is a small automatic contribution to relational guidance, though with continued practice relational guidance increases or disappears depending upon task demands. Experiments 6 and 7 show that relational guidance is unaffected by various grouping cues, suggesting that object spatial relationships are not coded by low-level visual processes, but rather by higher order pointers that code the categorical spatial relationships between objects (above, below, left, right). Collectively, these experiments suggest that object spatial relationships are encoded into the guiding target template at preview, thereby making this relational information available to guide search and removing the need to assume a pre-attentive coding of relational information between peripherally viewed search objects.

Dedication Page

This work is dedicated to my loving and supportive wife Samantha and son Johnathan and in loving memory of my Mother, Elaine Schmidt.

Table of Contents

List of Figures

List of Tables

Tables

Acknowledgments

I would like to thank Gregory Zelinsky, Susan Brennan, Hoi-Chung Leung, and Daniel Casasanto for their thoughtful comments and suggestions in preparation of this dissertation. I would also like to thank the entire eyelab for discussing data and design issues with me as well as Fizza Hussain, Sara Mott, Nikki Silver, Amanda Rendall, Jane Joseph, Alice Bartoldus, Andrea Tountas, Xiaoge Tao, Shirley Ju, Samuel Levy, and Vincent Perrone for their help with data collection and stimulus generation. Lastly, I would like to thank Jiye Shen and the entire SR research support staff for their unparalleled customer service and advice in designing and analyzing this work.

Joseph Schmidt, Ph.D., ABD
Curriculum Vitae
April, 2012

Office Address:
Department of Psychology
Stony Brook University
Stony Brook, NY 11794-2500
Office: (631)632-4972
Cell: (631)433-3930
Email: jcschmidt@notes.cc.sunysb.edu

EDUCATION

Ph.D. in Experimental / Cognitive Psychology, 2012
Dissertation: **Relational information between objects is available to guide search.**
(Advisor: Gregory J. Zelinsky)
Stony Brook University

M.A. in Psychology, 2008
Thesis: **Visual search guidance is best after a short delay**
(Advisor: Gregory J. Zelinsky)
Stony Brook University

B.A. in Psychology (Magna cum Laude), 2005
Stony Brook University

RESEARCH FOCUS:

Broad: Visual Cognition, Attention, Visual Working Memory, Sensation and Perception and Human Factors

Specific: My main research focus has been to investigate the nuances of the visual working memory (VWM) representation used to guide search. Through the use of multiple oculomotor measures (and more recently neuro-imaging and electrophysiological measures) I have systematically manipulated the amount and type of information provided at cue to investigate what information is available to guide search. By better understanding the VWM representations used to guide search, I hope to gain insight into the processes involved in visual memory and how these processes are integrated into our broader cognitive function.

PUBLICATIONS:

Schmidt, J., & Zelinsky, G. J. (2011). Visual search guidance is best after a short delay. *Vision Research, 51, 535-545*

Schmidt, J., & Zelinsky, G. J. (2009). Search guidance is proportional to the categorical specificity of a target cue. *Quarterly Journal of Experimental Psychology, 62 (10),* 1904-1914.

Zelinsky, G. J., & Schmidt, J. (2009). An effect of referential scene constraint on search implies scene segmentation. *Visual Cognition, 17 (6),* 1004-1028.

TEACHING EXPERIENCE:

Mentoring / Student supervision:

2009-Present Oversee one student conducting an undergraduate research project, University of Stony Brook

2005-Present Oversee three to eight students enrolled in undergraduate research credit per semester, University of Stony Brook

Direct Instruction:

Spring 08 Recitation Instructor for Research Design and Methods
Lead Instructor: Anne Moyer, Ph.D., Stony Brook University

Fall 06 Recitation Instructor for Human Factors
Lead Instructor: Susan Brennan, Ph.D., Stony Brook University

Teaching Assistant:

Spring 09 Teaching Assistant for Sensation & Perception
Lead Instructor: Gregory Zelinsky, Ph.D., Stony Brook University

Fall 08 Teaching Assistant for Biopsychology: Drugs in the Brain
Lead Instructor: Patricia Whitaker, Ph.D., Stony Brook University

Fall 07 Teaching Assistant for Survey in Cognitive & Perception
Lead Instructor: Susan Brennan, Ph.D., Stony Brook University

Spring 07 Teaching Assistant for Survey in Biopsychology
Lead Instructor: Patricia Whitaker, Ph.D., Stony Brook University

PEER REVIEWED PRESENTATIONS:

MacNamara, A., Schmidt, J., Zelinsky, G. J. & Hajcak, G. (2011, Sept.). *Attention toward fearful and neutral faces presented under high and low working memory load: a combined ERP and eye-tracking study*. Presented at the 51st annual meeting of the Society for Psychophysiological Research, Boston, MA.

Schmidt, J., MacNamara, A., Hajcak, G., & Zelinsky, G. (2011). ERP correlates of the target representation used to guide search. Journal of Vision, 11(11), 1345. doi: 10.1167/11.11.1345

MacNamara, A., Schmidt, J., Zelinsky, G. J. & Hajcak, G. (2011). *The effect of working memory load on electrocortical and ocular measures of attention to fearful and neutral faces.* Presented at the annual Determinants of Executive Function and Dysfunction (DEFD) conference, "How do Executive Function and Emotion Interact?", Boulder, CO.

Schmidt, J., & Zelinsky, G. J. (2010). *Searching for two objects: Does knowing their relational alignment produce greater search guidance?* Journal of Vision, 10(7), 1310a

Schmidt, J., & Zelinsky, G. (2010) *When is visual Information sampled?* Annual meeting of the Eastern Psychological Association. Brooklyn, New York.

Zelinsky, G. J., & Schmidt, J. (2009). *Searching aerial images: Evidence for scene constraints in the absence of global context.* Journal of Vision, 9(8), 1195a

Schmidt, J., & Zelinsky, G. J. (2009). *Visual search guidance is best shortly after target preview offset.* Journal of Vision, 9(8), 1183a

Zelinsky, G. J., Yang, H., & Schmidt, J. (November 21, 2009). Categorical visual search. (Part of the symposium on Visual Knowledge) *Abstracts of the 50th Annual Meeting of the Psychonomic Society*, *14*, 25-26. Boston, MA.

Schmidt, J., & Zelinsky, G. J. (2008). *Visual search guidance increases with a delay between target cue and search.* Journal of Vision, 8(6), 317a.

Schmidt, J., & Zelinsky, G. J. (2007). *Manipulating the availability of visual information in search.* Journal of Vision, 7(9), 715a.

Schmidt, J., & Zelinsky, G. J. (2006). *How is gaze affected by cognitive load and visual complexity?* Journal of Vision, 6(6), 363a.

PROJECTS CURRENTLY UNDERWAY *(Draft in progress and/or data collected)*

Schmidt, J., MacNamara, A., Hajcak, G. & Zelinsky, G. J. (in preparation) *A neural marker of the representation used to guide visual search: CDA's interaction with the guiding representations.*

Schmidt, J., & Zelinsky, G. J. (data collected) *Functionally related and interacting objects are more easily encoded in VWM as indicated by CDA amplitude.*

Schmidt, J., Alexander, R., & Zelinsky, G. J. (data collected) *Updating and recency effects within a single spatiotemporal object.*

Schmidt, J., & Zelinsky, G. J. (data collected) R*elational information between objects is available to guide search.*

MacNamara, A., Schmidt, J., Hajcak, G. & Zelinsky, G. J. (in preparation). *The effects of working memory load and emotion on electrocortical and ocular indices of attention.*

Moore, R., Schmidt, J. & Zelinsky, G. J. (in preparation) *Working memory load leads to saccadic suppression of eye movements and avoidance of visual stimuli*

Schmidt, J., Moore, R. & Zelinsky, G. J. (data collected) *Working memory load leads to saccadic suppression of eye movements and avoidance of visual stimuli: Part II*

Ferri, J., Schmidt, J., Hajcak, G. & Canli, T. (in preparation) *Don't look now: Attentional deployment modulates the neural response to unpleasant stimuli*. Journal of Neuroscience

And several others still in the works!

GUEST LECTURES/COLLOQUIA

Fall 11 "CDA, VWM, Search Guidance, Processing Speed and Their Intertwined Relationship" Cognitive & Experimental Psychology Colloquium, Stony Brook University

Fall 10 "Searching for two objects: Does knowing the spatial relationship between the objects produce greater search guidance?" Cognitive & Experimental Psychology Colloquium, Stony Brook University

Spring 10 "Visual search guidance is best after a short delay" 2nd/4th Year Department Symposium, Stony Brook University

Fall 09 "When is information loaded into visual working memory to guide search?" Cognitive & Experimental Psychology Colloquium, Stony Brook University

Spring 09 "Search & Object Based Attention" Guest lecture for Sensation & Perception, Stony Brook University

Spring 09 "Psychophysical Techniques" Guest lecture for Sensation & Perception, Stony Brook University

Fall 08 "Visual Search Guidance Changes with a Delay between Target Cue and Search" Cognitive & Experimental Psychology Colloquium, Stony Brook University

Spring 08 "Manipulating the Availability of Visual Information in Search" 2nd/4th Year Department Symposium, Stony Brook University

Fall 07 "Manipulating the Availability of Visual Information in a Target Cue Mediates Guidance in Visual Search" Cognitive & Experimental Psychology Colloquium, Stony Brook University

Fall 07 "Attention part I" Guest lecture for Survey in Cognitive & Perception, Stony Brook University

Fall 06 "Gaze Aversion as it Relates to Visual Stimuli" Cognitive & Experimental Psychology Colloquium, Stony Brook University

SERVICE:

Ad hoc reviewer for Attention, Perception & Psychophysics, Journal of Vision, Psychonomic Bulletin & Review, and Vision Research

PROFESSIONAL AFFILIATIONS:

Vision Sciences Society

Eastern Psychological Association

Relational information between objects is available to guide search.

Objects in our environment exist in relation to each other; at work my coffee mug is to the right of my computer keyboard, and they both reside on top of my desk. If I ask someone to hand me my coffee mug, I may even reference it relative to the keyboard. Often this reference object is used as a landmark because its location is known or it is easy to find (Beun & Cremers, 1998; Mack & Eckstein, 2011; Talmy, 1983). Once attention has been directed to the keyboard, it must then be further directed in the indicated direction to locate the coffee mug (Logan, 1995). This simple example demonstrates that fluidly interacting with objects in the world requires the online analysis of spatial relations between objects.

A significant amount of work has already investigated the use of relational information in various contexts ranging in their cognitive demands. Broad higher level tasks like foraging for food (e.g. Tamara & Timberlake, 2011), locating and interacting with objects in the environment (e.g. Byrne & Crawford, 2010, Marsh, Spetch, & MacDonald, 2011), describing the location of an object (e.g. Schober, 1993, 1995) and navigation (e.g. Foo, Warren, Duchon, & Tarr, 2005; Marsh, Spetch, & MacDonald, 2011; Meade, Biro, & Guilford, 2005; Newman et al., 2007; Sturz & Bodily, 2010; Sturz, Cooke, & Bodily, 2011; Sturz, Kelly, & Brown, 2010; Sutton, 2002) are highly dependent on the use of landmarks and relational information. Even highly schematized tasks such as interpreting information from a scene seems to be influenced by the relational information of the objects contained within the scene (e.g. Mandler & Johnson, 1976). Many lower level processes also seem to implicitly use relational information. Memory for an object's identity or orientation in an explicit change detection task is more accurate when the absolute and/or relative spatial relations between the objects in the study and test array are the same, presumably because the relational information provides a broader context which aids memory

(Hollingworth, 2006, 2007; Jiang, Olson, & Chun, 2000; Vidal,Gauchou, Tallon-Baudry, & O'Regan, 2005). Similarly, if only the location of the test object changes or the non-test objects do not reappear, this too is believed to disrupt the context relative to the remaining objects in the scene, resulting in less accurate memory for the test object. This finding suggests that relational information between objects is not only maintained in working memory (WM) but it is closely intertwined with memory for an object's identity (Hollingworth, 2006, 2007; Hollingworth & Rasmussen, 2010; Jiang, Olson, & Chun, 2000; Olson & Marsheutz, 2005). Moreover, relational information is thought to provide a larger context that is also able to speed search performance among repeated array presentations (Chun, 2000; Chun & Jiang, 1998; Jiang & Chun, 2001). Whereas these array repetitions occur infrequently enough that they go unnoticed, search efficiency increases with repetition, possibly due to the implicit use of relational information during the search process (e.g. Chun & Jiang, 1998). Unlike the implicit effects just described, the use of explicit salient landmarks in search is thought to reduce the likelihood that objects are re-inspected; the explicit reference is believed to enable inhibition of return to operate more effectively (Becic, Kramer, & Boot, 2007; Peterson, Boot, Kramer, & McCarley, 2004). This inter-object effect is not only thought to affect memory and search for an individual object but it also affects the recognition of individual objects (Bar & Ullman 1996); when an individual object is presented along with a related interacting object, speeded recognition tends to be more accurate when they are interacting in a way congruent with real-world expectations and less accurate if they are interacting in an incongruent way (Green & Hummel, 2006), leading some to suggest that objects performing an action (e.g., a corkscrew opening a bottle) are perceptually bound and perceived as a unitized configuration or group (Green & Hummel, 2006; Humphreys et al., 2010; Riddoch et al., 2011; Roberts & Humphreys, 2011a, 2011b).

The current work extends research into the effects of relational information to a visual search task. People are highly adept at using relational information between objects in a search display to increase their search efficiency (e.g., Brockmole & Henderson, 2006; Neider & Zelinsky, 2006). When searching for my coffee mug at work, I will tend not to look on a shelf. However, if I was told that my coffee mug was placed on the shelf, would awareness of this spatial relationship improve search performance? It would clearly restrict the search space by confining attention to the shelf (Zelinsky & Schmidt, 2009); however, would the shelf be incorporated into the target representation? In this case one could search for the coffee mug alone, the coffee mug and the shelf, the coffee mug above "something", or the coffee mug above the shelf. These possible target representations will be the focus of this investigation. Can relational information between objects be included in the representation used to guide search?

Spatial terms are thought to impose a relevant space around an object, often referred to as a reference frame (Garnham, 1989; Levinson, 1996; Logan, 1995). This reference frame is essentially a flexible coordinate system with a number of parameters such as an origin, orientation, direction and scale/distance (Logan, 1995; Logan & Sadler, 1996). The majority of work investigating how reference frames are used to process relational information has occurred in the context of two primary tasks, a sentence verification task in which the observer's goal is to verify if "X is above Y" in a display containing some number of objects, and acceptability ratings in which observers rate how well the current display exemplifies the specified relation. These tasks have led to the conclusions that small distances between objects tend to not affect relational judgments (Logan & Compton, 1996), but long distances can (Carlson & Van Deman, 2004). However, when distractor objects are present, verification times increase even with small increases in distance (Logan & Compton, 1996), regardless of the location of the distractor

(Carlson & Logan 2001). These findings are consistent with: attention taking longer to traverse a larger distance between objects, distractors only affecting relational judgments by increasing the difficulty associated with locating the relevant objects, and only checking the relational information for the target objects. For quite some time, distance was thought to be unnecessary to verify directional terms (e.g. Logan & Sadler, 1996). However, recent evidence suggests that calculating the distance between objects is a necessary step in coding relational information, as successively repeating distances between objects (while not repeating objects, locations, or relations) results in faster verification times compared to not repeating distances, presumably because the distance parameter can be carried over from the previous relational judgment (Carlson & Van Deman, 2004). Similarly, direction seems to be encoded during the use of proximal terms such as near and far, despite directionality being thought to be irrelevant to distance judgments (Ashley & Carlson, 2007; Logan & Sadler, 1996). These last two points suggest that while encoding relational information, people encode additional features about the relevant pair of objects.

Researchers have tried to model the use of relational information for quite some time, both within (Hummel & Biederman, 1992) and between objects (Kosslyn, Chabris, Marsolek, & Koenig, 1992; Ullman, 1984). However, some of the most thorough investigations of relational information have come from Logan's many inquiries into the topic. In a series of eleven experiments, Logan (1995) argued that computing relational information requires many additional processes beyond those discussed in the attention literature. Logan points out that the attention literature often investigates the absolute position of a target object (e.g. Posner, 1980), that is "basic" relations. But relational terms such as above, below, left and right are relative or indexical relationships; object X may be above object Y and both may in fact be below you.

Such deictic relationships prohibit simply directing attention in the specified direction; rather, you must first locate object Y and then object X's relative position must be calculated. Logan also acknowledged that this distinction between basic and deictic relations has existed in the field of linguistics for quite some time (e.g. Garnham, 1989), and his goal was to combine theories of attention that deal with basic terms, with theories of linguistics that deal with relative terms and the use of reference frames so as to explain how attention may be involved in computing relational information. According to Logan, to determine whether X is above Y, object Y must first be located, then a reference frame must be imposed on object Y. This reference frame must then be aligned with a spatial template to compute the region above object Y, object X must then be spatially indexed to create a marker for its location, and finally a goodness of fit must be computed to see if object X falls within a region that is indeed above object Y. Much of this was later extended and refined in the form of a computational model (Logan & Sadler, 1996), however the basic framework remained largely the same. This framework has been repeatedly tested (Ashley & Carlson, 2007; Carlson & Logan 2001; Carlson & Van Deman, 2004; Logan & Compton, 1996) and in fact ERP correlates of many of the steps originally outlined by Logan have been identified (Carlson, West, Taylor, & Herndon, 2002). More recently, however, an alternative simplified theory has been proposed to this multi-step model which suggests that relational information *could* be coded by simply recording a pointer between the objects as covert or overt attention shifts from one object to the other (Franconeri, Scimeca, Roth, Helseth, & Kahn, 2011). This active research topic has even spawned theories attempting to model relational judgments of functionally related interacting objects (Carlson, Regier, Lopez, & Corrigan, 2006; Regier & Carlson, 2001).

Few studies have investigated whether the relational information between objects can guide search. Much of the work using sentence verification tasks and acceptability ratings to investigate the use of relational information has allowed for multiple shifts of attention between the objects. Search guidance is somewhat different as it requires locating the target objects using peripheral vision, presumably before attention has shifted there. Search for a target defined by relational information within or between objects is known to be inefficient and difficult (Palmer, 1994; Wolfe, 1998) and possibly capacity limited (Huang & Pashler, 2005). But perhaps the most thorough investigation into this question was again by Logan in 1994, in which he had observers search for a + and – in a specific arrangement (e.g. + left of –) among distractor pairs flipped over the specified axis (e.g. + right of –). Logan demonstrated that search for relational information between objects using pictorial previews or linguistic text based cues was highly inefficient and it continued to be inefficient even after extended practice. He suggested that computing relational information between objects requires focal attention; relational information cannot be used or computed pre-attentively or in the absence of attention. Logan's conclusion would preclude the ability to guide search by relational information, as guidance signals are derived from parallel stages of processing across the field of view (Wolfe, 1994, 2007; Zelinsky, 2008). If focal attention must be directed to a pair of objects before relational information can be computed, it follows that attention cannot be guided to these objects on the basis of their relational information. It further follows from this argument that attention would shift randomly from object pair to object pair until the target is located, as the target spatial relationship would not be able to guide these movements of attention. As well, Logan's demonstration that search efficiency for a target defined by relational information does not improve with extended practice or with pictorial previews suggests that memory and practice for a pair of objects does not mitigate the need to check the relational information, or increase the efficiency of the checking process. The one exception to this rule is when the extended practice occurs with objects that are connected such that they create a single perceptual object (Logan, 1994). This suggests that each time a pair of objects is encountered the relational information must be rechecked to verify the relationship.

Aims:

Logan (1994) concluded that computing spatial relations requires focal attention, which would preclude using this information to guide search. The present series of experiments builds on this earlier effort by addressing two main aims. First, I replicated the parameters of Logan's now seminal work and used real-world objects and more sensitive eye movement dependent measures and demonstrated that guidance by relational information is indeed possible under these more naturalistic conditions. Second, I informed existing search theory by exploring the boundary conditions in which relational information affects search performance, and when it does not.

Overview:

The first experiment addresses Aim 1 of this study, the demonstration that relational information can guide overt attention (eye movements) in the context of a search task using visually complex real-world objects. The design of Experiment 1 replicates the parameters of Logan's (1994) fifth experiment in which the target pair was pictorially previewed and defined by the relational information between the target objects, i.e. the target pair and the distractor pairs use the same objects and only differ with respect to the relative locations of the individual objects (their relational information). It is possible that the perceptual self-similarity of the pluses and dashes and the exchanging of targets and distractors used in Logan's work may have made the relational judgment especially difficult, whereas the featurally rich, trial-unique, real-world objects used in the present work would make for an easier coding of object spatial relationships. As well, the measures of search efficiency used by Logan (1994), principally the slope of the set-size effect, can be affected by a number of factors, including the time spent looking at distractors and a shift in decision criteria. Given that relational judgments are often considered difficult and time consuming, it is not surprising that search for a relation between objects would be inefficient. However, inefficient search need not be unguided search. Eye-movement measures tend to be more powerful as they allow the direct measurement of overt search guidance relative to a clear chance baseline rather than inferring guidance from search slopes (e.g. Schmidt & Zelinsky, 2009). I also tested whether the relational information is coded differently over time; if encoding relational information is highly explicit and requires one or more shifts of attention (Franconeri et al., 2011), then guidance by relational information will be highly dependent upon preview duration. Homogenous displays having identical distractors were used in Experiment 1 in order to definitively test whether relational information can guide search, and to remain consistent with Logan (1994). However, the degree of homogeneity used in Logan's work and in the present Experiment 1 is rare; real world settings are often highly heterogeneous, and this would allow the features of the target objects to guide search. Demonstrating guidance by relational information in these more heterogeneous displays would suggest that this information may be useful more generally. Experiments 2-7 address Aim 2, parameterizing guidance from relational information and informing search theory as to how relational information is coded and used to guide search. In Experiment 2 all of the distractor pairs consisted of random non-target objects, half vertically oriented and half horizontally oriented. Target pairs either matched the preview's relational information or the objects swapped relative positions (rather than object X being on the right, it was moved to the left); guidance to the target pair when the relational information matched the preview was quantified and compared to when it did not match the preview. This difference between matched and swapped targets is a pure measure of the guidance resulting from the relational information; and will be referred to as relational guidance.

Experiment 3 further tested the generalizability of relational guidance by removing the pictorial previews and cuing the objects with categorical text cues. Replicating relational guidance with categorical cues would suggest that search can generally be guided by relational information in day to day searches, even when referenced on the fly. Failure to replicate relational guidance in this circumstance may suggest that relational information must be obtained from specific feature information, that is, visual features of a specific object rather than features of a category of objects.

Experiment 4 tested whether specific pairs of objects and their spatial relationships impact relational guidance. Recent evidence suggests that related interacting objects, in which the directionality of the relationship is implied by the objects themselves (i.e. a hammer above a nail), may be perceptually bound and preferentially processed (Green & Hummel, 2006; Gronau, Neta, & Bar, 2008). This suggests that our semantic knowledge about objects and their expected relationship biases our recognition of the individual objects involved in the relationship (see also Kroll, Schepeler, & Angin, 1986). This perceptual binding may affect when and if the relational information impacts search guidance. Finding that related non-interacting pairs (i.e. the hammer *below* the nail) selectively *fails* to show relational guidance would demonstrate that relational guidance is dependent upon the objects involved in the pair, suggesting that an effect of the relational information is subject to the ease in which the relational information between the objects can be coded.

Experiment 5 tested whether relational information between objects is automatically coded and able to guide search. To the extent that coding relational information is automatic, a task change should not affect relational guidance. However, if coding relational information is effortful and explicit, relational guidance should only be found when observers are instructed to the code this information. The extent to which relational information is automatically coded will be determined by comparing a condition in which observers are instructed to localize the target pair to a condition in which observers are instructed to localize the target pair and respond if the relational information matches the preview or not.

There are clearly many ways in which relational information *could* be represented. One possibility would be to represent the two objects as a single target representation, implicitly capturing the spatial relationship (the unified percept hypothesis). Another method might code the features from both objects separately, along with an explicit spatial reference between them (the explicit reference hypothesis). This second possibility would obtain features from both objects separately and use a categorical spatial term (above, below, left, or right) to establish a spatial relationship between them. The unified percept hypothesis asserts that grouping principals should impact the extent to which relational information guides search; when objects are grouped more strongly, they should be more easily unified and therefore available for guidance. The explicit reference hypothesis asserts that guidance need not be affected by perceptual grouping principles, as it is the categorical spatial term that is responsible for guidance. Experiment 6 tested these hypotheses by placing a picture frame around each object pair, thereby explicitly grouping the objects into a unitized percept that, according to the unitized percept hypothesis, may result in a magnified effect of relational guidance.

Experiment 7 tested a more standard grouping procedure by manipulating the distance between the objects in each pair. Spatial proximity is a strong grouping cue (e.g. Shihui, 2004); therefore grouping by proximity should be strongest when the objects are close and should drop off as the distance between the objects increases. When grouping is strong, relational information should have its maximal effect, assuming the objects are bound into a unified percept. Conversely, as distance between the objects increases relational guidance should decrease if a unified percept is used. Again, if the explicit reference hypothesis is correct, this grouping manipulation should also be ineffective.

Experiment 1

The goal of this experiment was to replicate the parameters of Logan's (1994) Experiment 5 in the context of random real-world objects that avoid the self-similarity inherent to the simple stimuli used in the earlier work. As well, this experiment will observe search guidance directly by analyzing eye movement measures rather than inferring guidance from search slopes. Like Logan, the current goal is to measure a pure effect of relational guidance, and to this end target and distractor pairs used the same objects on a trial by trial basis, prohibiting guidance from anything other than relational information (i.e., visual features of the target items). The target pair differed from the distractor pairs only in the relative positions of the objects, that is, the target was defined by the relational information between the objects. If the search for a relationally-defined target is indeed random, as suggested by Logan, then search in this task should be completely unguided and the proportion of trials in which the target is the first pair fixated should be at chance. However, if Logan was incorrect and search for relational information *is* guided, then the target pair should be fixated first significantly more often than chance would predict. As for how strong relational guidance might be, it is certainly the case that relational information is not a basic visual feature like color or orientation; early visual areas code color, orientation, and retinotopic location directly (e.g. Engel, Glover, & Wandell, 1997; Hubel & Livingstone, 1987; Livingstone & Hubel, 1984), but they do not code relative positions. Rather, relational information is very likely a constructed visual feature, and its contribution to search might therefore be expected to be small relative to chance. Nevertheless, any guidance by relational information would indicate that the relation has been encoded, that it is available to the guiding representation, and that it can be resolved using peripheral vision (see also Alvarez & Oliva, 2008; 2009; Rosenholtz, Huang, & Ehinger, 2012).

Experiment 1 also tested whether relational information is coded differently over time. If encoding this relationship requires spreading attention over a large region encompassing both objects, or requires shifting attention between the objects, then guidance by relational information should occur only when the preview duration is sufficiently long. Lastly, I also tested if relational guidance changes with the number of previewed objects. The one-object condition captures the situation in which someone references the “coffee mug to the left of the thing”. This provides some degree of additional information, yet not knowing what the coffee mug is to the left of may make the relational information less useful. Ordinarily, previewing part of the target would generally result in poorer search performance (Alexander & Zelinsky, in preparation), yet multiple target search is generally more difficult than single target search (e.g. Kaplan & Carvellas, 1965; Menneer, Barrett, Phillips, Donnelly, & Cave, 2007; Menneer, Cave, & Donnelly, 2009; Menneer, Donnelly, Godwin, & Cave, 2010; Moore & Osman, 1993; Yang & Zelinsky, in preparation), and in this case the added features from the target objects are also shared by the distractor objects. It is therefore unclear how previewing both target objects will affect search performance relative to previewing only one object. If search is improved after previewing both objects, this would likely reflect a more precise coding of the relational information, as the second object’s features are completely redundant with the distractors. Conversely, if search performance is worse after previewing both objects this would be consistent with interference from the distractor objects and from the increased working memory load associated with coding two objects.

Methods

Participants

Twenty-four undergraduate students from Stony Brook University participated for course credit. All had normal or corrected-to-normal vision and were native English speakers, by self-report.

Apparatus

Stimuli were presented at a screen resolution of 1280 x 960 pixels using a 19" flat-screen CRT ViewSonic SVGA monitor operating at a refresh rate of 100 Hz. The experiment was designed and controlled using the Experiment Builder software package (SR Research Ltd, version 1.10.165), running on an Intel core 2 (quad core) PC under Windows XP. Eye position was sampled at 1000Hz, using an EyeLink 1000 eye-tracker with default saccade detection settings. Observers' head position and viewing distance was fixed at 70 cm using a chin and forehead rest. All responses where made with a standard gamepad controller (Microsoft Sidewinder 1.0) by pressing the left and right index finger shoulder triggers; trials were initiated by pressing the x button operated by the right thumb.

Stimuli

Stimuli consisted of a target preview and a search display depicting 12 objects (Figure 1). Search displays were presented on a 21° × 21° white background, and consisted of six pairs of the previewed object/s (all Hemera Photo Objects - Gatineau, Quebec, Canada) arranged in a circle (8.73° radius from central fixation). Only the target pair matched the previewed relationship (e.g. X above Y), all distractor pairs differed from the target pair only by the relational information between the objects (e.g. X below Y). To construct the object arrays, six coordinates where chosen along the circle's circumference at 60° increments and a pair of objects was centered at each of these locations. Each object of a pair was offset 20 pixels (0.46°) from the selected locations, resulting in a distance of 40 pixels (0.92°) between the paired objects. Objects were resized such that the smallest bounding box enclosing each object was as close to 4225 pixels as possible (without exceeding this amount), and the resizing operation preserved the aspect ratio of the objects. This resulted in the object area subtending approximately 1.5°, and each pair of objects subtended approximately 1.5˚ x 3.92˚ (if vertically oriented). All objects were presented in color, and a different pair of objects was used on each trial. Within each condition target objects appeared equally often at each of the six possible locations.

Design and Procedure

There were two target preview conditions. The one-preview condition showed one of the objects from the target pair above, left, right, or below central fixation. The two-preview condition showed both objects of the target pair vertically or horizontally oriented, centered at central fixation. In both conditions, the edge of each object was offset 20 pixels (0.46°) from central fixation (equal to the offset of the objects in the search display), and a small cross was displayed at central fixation; observers were instructed to look at the cross rather than at the objects to assure that they began their search from the center of the screen, equidistant from each pair of objects. As well, in both preview conditions a bounding box was drawn around the pair of previewed objects accurately depicting the height and width of the target pair. In the one-preview condition the empty space served as a placeholder for the other object (see Figure 1). Previews were presented for either 400ms or 1600ms. This yields four conditions, a one-object and a two-object preview with a duration of 400ms and a one-object and a two-object preview with a duration of 1600ms. Observers were instructed to find the previewed object/s in the exact spatial relationship as the preview (e.g. X above Y, all distractors were X below Y), and while looking at the target pair hit either of the top shoulder trigger buttons to indicate that they localized the target pair. In case of inadvertent track loss or exceedingly long RTs, trials timed out after 10 seconds. These trials were counted as errors. Across observers all search displays were identical. Preview conditions were counterbalanced across observers and trials such that the four conditions appeared equally often on each trial. As well, the one-preview condition was counterbalanced such that half of the observers previewed object "X" and the other half previewed object "Y". Number of previewed objects and preview duration were blocked within subjects. Trial order was randomized within blocks. In total, there were 240 experimental trials, 60 per condition/block.

The experiment began with a calibration routine used to map eye position to screen coordinates. A calibration was not considered acceptable unless the average error was less than .40° and the maximum error was less than .90°. At the start of each block, observers were calibrated and asked to perform six practice trials in order to familiarize themselves with the task, stimuli, number of previews and preview duration. This was followed immediately by the block of experimental trials.

Trials began with observers fixating a central point and pressing the X button on the game pad. In addition to starting the trial, this served as a "drift check" for the eye-tracker to record any shift in the eye position since calibration. The fixation point was then replaced by a one or a two-object target preview for either 400ms or 1600ms, followed immediately by the search display until response or time-out. Feedback was given by sounding a tone after an incorrect/timed-out response. The entire experiment lasted approximately 50 minutes.

Measures & Predictions

By presenting one or both of the target objects for 400ms or 1600ms, Experiment 1 can independently examine the effects of the number of preview matching objects and preview duration on relational guidance. Although RTs are reported as a general measure of search performance, RT measures incorporate many perceptual and post perceptual processes unrelated to search guidance. For example, RT differences may result from shifting decision criteria, differences in the time needed to recognize a target, and differences in how quickly distractors can be rejected, all factors that lessen the usefulness of RTs as a measure of search guidance.

The crucial analysis to test for guidance by relational information is the proportion of trials in which the target pair is the first pair of objects fixated during search. If Logan (1994) was correct and search by relational information is indeed serial, then the target pair should be fixated first at chance levels. However, if relational information can guide search, the target should be fixated first more often than chance would predict, indicating that guidance can operate on relational information alone. In this case, because there are six pairs of objects in the search display, chance corresponds to 1/6 or 16.667% of trials. Search would be considered guided by relational information if the target pair is fixated first above this baseline rate. As the proportion of trials in which the target pair is fixated first increases, so too does guidance.

To investigate guidance effects that may not be evident at the outset of search but rather may evolve with further investigation of the search display, I analyzed the time until the target pair is fixated. The time to fixate the target is a measure of search guidance that excludes delays resulting from target recognition and target verification. It is possible that target verification may rely on different processes and therefore may have a different pattern of results compared to a measure of the time to fixate the target. Given that time to the target encompasses a greater portion of the search task while minimizing other processes unrelated to guidance, any guidance related effects should have the greatest likelihood of showing up in this measure. Faster times to fixate the target would suggest stronger search guidance, whereas slower times would suggest weaker search guidance.

Results & Discussion

RTs revealed no significant differences between conditions (See Table 1, all Fs $(1, 23) \leq$ 2.52, all ps $\geq$ .12); however, this says little about actual search guidance and the use of relational information in guiding search. Critically, as can be observed in Figure 2 all conditions resulted in a significantly greater than chance proportion of trials in which the target pair was fixated first (all ts(23) $\geq$ 2.18, all ps $\leq$ .04), demonstrating that relational information alone can be used to guide search. This also means that the search for a target defined by relational information is not completely serial, as suggested by Logan (1994). The effect of preview duration, number of previews and the interaction of the two all failed to reach significance for the proportion of trials in which the target pair was fixated first (all Fs $(1, 23) \leq$ 1.75, all ps $\geq$ .19). However, the time until the target pair was fixated (Table 1) revealed that the target pair was fixated significantly faster after a 400ms preview compared to a 1600ms preview $F(1, 23) = 7.30$, $p = .01$, and that the target pair was fixated significantly faster when previewed with both target objects compared to one $F(1, 23) = 5.53$, $p = .03$; the interaction of preview duration and number of previewed objects failed to reach significance, $F(1, 23) = 0.07$, $p = .80$.

This pattern of results suggests that relational guidance can be established very quickly, within a 400ms preview, and additional processing beyond 400ms results in more difficult target localizations as indicated by the time to target measure. Faster search after a short preview is consistent with earlier work using a single object pictorial preview (Wolfe, Horowitz, Kenner, Hyle, & Vasan, 2004)[1]. As well, because this experiment was designed to limit the use of target features and to restrict guidance to the use of relational information, the faster target localization after a two-object preview indicates that it may be easier to code relational information when both objects are previewed.

Experiment 2

Experiment 1 addressed Aim 1, demonstrating that search can be guided by relational information alone. The remainder of the experiments will address Aim 2, parameterizing relational guidance so as to inform existing search theory as to how relational information may be coded to guide search. Search in a more natural setting includes many different objects at various orientations, and the increased heterogeneity in such displays may wash out any small contribution of relational information. Experiment 2 seeks to replicate the effects of relational guidance observed in Experiment 1 in the context of more heterogeneous displays. Importantly, in heterogeneous displays search can be guided by target features (see Figure 3). To observe a pure effect of relational information, half of the trials had the target preview match the search view, and in the other half of the trials the objects swapped relative positions (X above Y became X below Y). If relational information affects guidance in these heterogeneous displays then the relational match condition should show significantly greater search guidance than the relational swap condition. It is interesting to note, in this task, that the relational information is needed only for the manual response and therefore a reasonable strategy would be to guide search based on the target objects features only, and then to verify the relation. To the extent relational guidance is observed, that would suggest that the same features are used to guide search and verify the target. As well, since the target features are no longer redundant with the distractor features, previewing both objects would be expected to produce a larger improvement in search performance relative to previewing just one object compared to what was observed in Experiment 1.

Methods

Participants

Twenty-four undergraduate students from Stony Brook University participated for course credit. All had normal or corrected-to-normal vision and were native English speakers, by self-report.

Design and Procedure

Stimuli, apparatus, and procedures were identical to Experiment 1, with the following exceptions. Variability was added between the target and distractor objects; distractor pairs no longer used the same objects as the target pair, but rather were composed of other random objects (all objects were presented once per observer, all Hemera Photo Objects - Gatineau, Quebec, Canada) to allow for guidance based on the features of the target objects. On each trial, half of the objects in the search display were horizontally oriented and half were vertically oriented (Figure 3). Preview duration was held constant at 400ms. The previewed object/s were always present in the search display, but to encourage the use of the relational information the task was changed to target present/absent rather than 100% target present. On target present trials (50%) the target pair appeared in the same spatial relationship as the preview (e.g. X above Y); on target absent trials the spatial relationship was reversed (e.g. X below Y). Thus, on absent trials the previewed objects in the search display act as a lure allowing us to observe guidance relative to the features of the objects when the relational information is incorrect. As in Experiment 1, search displays were identical across observers, however conditions were counterbalanced across trials by changing the number of previewed objects and the relational information of the previewed objects such that each condition appeared equally often on each trial across observers.

Results & Discussion

If search can be guided by relational information in heterogeneous displays above and beyond guidance from target features, then the target pair should be fixated first significantly more often when the relational information matches the preview compared to when the objects swap relative positions. A similar pattern should also be found in the time to target measure. To the extent that observers use both target objects during search, the additional target features from previewing both objects should also result in a significantly greater proportion of trials in which the target is fixated first and shorter times to fixate the target, relative to when only one target object appeared at preview. Overall RTs are also expected to replicate the time to target data, showing faster search when the relational information matches the preview compared to when it is swapped, and faster search when both objects are previewed relative to when only one object is previewed.

The RT data (Table 2) revealed that search was significantly faster in the matched condition relative to the swapped condition, $F(1, 23) = 46.21$, $p < .001$, and the two-object preview also produced faster search relative to the one-object preview, $F(1, 23) = 8.73$, $p = .01$. However, the interaction of the two failed to reach significance, $F(1, 23) = 0.54$, $p = .47$. Because it is unclear whether the RT data is the result of true target guidance or not, I again analyzed the proportion of first object fixations. As can be observed in Figure 4, the target pair was fixated first significantly more often in the matched condition compared to the relational swapped condition, $F(1, 23) = 5.90$, $p = .02$; however, the effect of number of previewed objects and the interaction of the two failed to reach significance (both $Fs\ (1, 23) \leq 0.85$, both $ps \geq .36$). This pattern of data clearly demonstrates that coding relational information for a manual response results in relational guidance even in the context of highly heterogeneous displays, suggesting that in this task the same target features are used for search and verification processes. Consistent with the RT data and the first object fixated data, the time to target data (Table 2) also showed a clear benefit for the matched condition compared to the swapped condition, $F(1, 23) = 17.86$, $p < .001$. As well, the time to target data showed a clear advantage for having previewed both target objects $F(1, 23) = 8.24$, $p = .01$. Interestingly, the two-object advantage is of a similar magnitude as that observed in Experiment 1, suggesting that the benefit may come from a more complete target encoding rather than just the additional target features, although future work would clearly need to confirm this. This is supported by the results of a formal debriefing questionnaire in which observers explicitly reported the use of one target object even when both were presented. Lastly, the interaction of the relational information and of the number of previews failed to reach significance, $F(1, 23) = 1.94$, $p = .18$. These results clearly demonstrate guidance by relational information and a small search advantage for having previewed both target objects, even with highly heterogeneous search displays.

Experiment 3

Experiments 1 and 2 clearly demonstrated that search can be guided by relational information. However, both of these studies used pictorial previews in which the relational information could be obtained from a visual percept (the pictorial preview). Yet in the real world objects are often referenced using words. Experiment 3 further parameterized relational guidance by specifying the target objects using text cues referring to a target class instead of a pictorial preview or a specific object. In this case the target features must therefore apply to an entire class or category of objects rather than a specific object to achieve reasonable search guidance. If relational guidance does not require highly specific visual features, and if relational information can be obtained from self-generated categorical representations, then I should replicate the search guidance advantage when the relational information matches the text cue compared to when the objects swap relative positions. This would suggest that people can use relational information "on the fly" in the real world without the need to code relational information from a visual percept. However, if highly specific visual features are needed to accurately represent the relational information then guidance should be unaffected by the relational swap.

Methods

Participants

Forty undergraduate students from Stony Brook University participated for course credit. All had normal or corrected-to-normal vision and were native English speakers, by self-report.

Design and Procedure

Stimuli, apparatus, and procedure were identical to Experiment 2, with the following exceptions. All target objects were replaced with objects from four target categories (fish, butterflies, cars, and teddy bears) to allow for easy target referencing; search displays were otherwise identical. Pictorial previews were replaced with text cues (such as “butterfly above fish” for the two-object condition or “butterfly above ______” or “______ above fish” for the corresponding one-object conditions). Cue duration was increased from 400ms to 3000ms and was followed by a 1000ms ISI to allow enough time to read the cue, construct a guiding representation, and to allow for gaze to return to central fixation before the presentation of the search array. In addition to counterbalancing conditions over search displays and observers, all possible pairings of target categories and reference terms were used equally often.

Results & Discussion

If search can be guided by relational information in the absence of a pictorial preview then the relational guidance effects found in Experiments 1 & 2 should replicate in Experiment 3. This would suggest that relational information can be used with self-generated categorical representations and that the features used to code the relative positions need not be highly specific. However, if the use of relational information requires very specific visual features then our three primary measures of search performance should show equivalent performance regardless of the relative positions of the target objects.

The RT data (Table 3) revealed a clear advantage for the match condition compared to the swapped condition, $F(1, 39) = 72.52$, $p < .001$. However, both the effect of number of cued objects and the interaction of the two failed to reach significance (both $F(1, 39) \leq 0.75$, both $p \geq .39$). Both measures of search guidance, the proportion of trials in which the target pair was fixated first (Figure 5), and the time to fixate the target pair (Table 3), did not show any reliable differences (all Fs $(1, 39) \leq 2.22$, all ps $\geq .14$). Taken together these results suggest that categorical text cues are not sufficient to generate relational guidance and when multiple objects are cued, observers' search performance is unaffected. Despite the lack of guidance differences, the RT data did show a strong advantage for the matched relational information, suggesting that categorical text cues exert effects of relational information selectively during recognition/verification stages rather than on search guidance. Future work will have to examine whether this result is due to the use of text cues or whether it is specific to the use of categorical cues. More specifically, if observers are instructed to recall specific objects rather than a category of objects, the target features can be highly specific, which may allow for a more detailed coding of relational information and possibly the use of relational information in guiding search.

Experiment 4

All three experiments thus far have used arbitrary objects to test for relational guidance; this experiment investigated how the target objects themselves affect the use of relational information. Functional interactions between objects are known to bias relational judgments (Carlson-Radvansky, Covey, & Lattanzi, 1999); when observers are instructed to place a coin above a piggy bank they tend to place it close to the slot in the piggy bank rather than directly above it. More recent evidence suggests that functionally related interacting objects, or objects depicting an action, seem to be perceptually grouped, resulting in more accurate recognition of these objects relative to non-interacting related and unrelated objects (e.g. Green & Hummel, 2006; Humphreys et al., 2010). In fact, RT and neural imaging data suggest that related interacting objects may in fact be linked, such that they generate a unified representation (Gronau, Neta, & Bar, 2008). This linking of functionally-interacting/action-related objects (Green & Hummel, 2006; Humphreys et al., 2010; Riddoch et al., 2011; Roberts & Humphreys, 2011a, 2011b), e.g. a corkscrew opening a bottle, also results in benefits relative to semantically related objects (Riddoch, Humphreys, Edwards, Baker, & Willson, 2003). The preferential processing of action related interacting objects is likely due to the affordances associated with the implied action between the objects (Roberts & Humphreys, 2011a, 2011b); ‘active’ objects (the corkscrew) are more accurately recognized when oriented for use with the dominant hand, whereas ‘passive’ objects (the bottle), are more accurately recognized when presented in a functionally-interacting pair, relative to non-related objects, e.g. a match over a bottle (Roberts & Humphreys, 2011b). If related interacting objects are perceptually grouped or linked, this might result in the relational information being coded more effectively.

Interestingly, most of the work investigating related interacting objects has not required the coding of relational information. Often observers perform a single object verification task, with the other object presented as a distractor. Its presence and position seems to selectively affect performance despite the lack of any explicit use of relational information. Thus, it is unclear what may happen in our task, given that observers' must code the relational information. Assuming the earlier reported effects of functionally related and interacting objects translates into an effect in search guidance; relational guidance should be strongest with related interacting objects and should be weakest with related non-interacting objects compared to unrelated objects. This would suggest that relational information may be weighted by the relatedness and relative positions of the target objects (see also Carlson et al., 2006; Regier & Carlson, 2001). However, rather than a shift in the magnitude, it is possible that relational guidance may selectively exert its influence. The previous literature certainly suggests that related non-interacting objects should be the least well recognized and processed, which may result in a lack of relational guidance in this condition. Likewise, when the targets are related and interacting, and when the targets are unrelated and non-interacting, the implied function and the relational information are not conflicting, thus coding the relational information may be easiest in these circumstances.

Methods

Participants

Twenty-four undergraduate students from Stony Brook University participated for course credit. All had normal or corrected-to-normal vision and were native English speakers, by self-report.

Design and Procedure

Stimuli, apparatus, and procedure were identical to Experiment 2, with the following exceptions. Target pairs were replaced with objects in which the relational information is consistent with common functional associations; e.g. a hammer above a nail, or a spoon above soup (for a complete list of related pairs see Figure 6). The interacting nature of these objects provides a degree of directionality inherent to the relational information. Half of the trials maintained the related nature of the target pairs (related condition); the other half used unrelated objects (unrelated condition), e.g. a spoon above nail or a hammer above soup. On 50% of the trials, previews were presented with the directionality of the functional relationship preserved (interacting condition), i.e. a hammer above a nail; on the remaining 50% of the trials, the previews were presented with the directionality of the relationship contrary to the functional placement of the objects (non-interacting condition), i.e. a hammer below a nail. In both cases, unrelated trials used the object placements from the related condition. This resulted in 8 conditions, related & interacting, related & non-interacting, unrelated & interacting, and unrelated & non-interacting, presented with both matched targets and swapped targets in the search display. As in Experiment 2, half of the search displays were target absent, in which the relative locations of the previewed objects were swapped in the search display. Like Experiment 2, all distractors were trial unique; though unlike Experiment 2, the target objects were repeated once per condition to control for object specific effects such as ease of naming or imagability of the object. Due to the nature of the stimuli, across trials target pairs were no longer balanced across orientations (2/3 vertical, 1/3 horizontal), however each orientation occurred equally in each condition. Only the two object preview condition was tested. The number of search trials per condition was reduced from 60 to 30 to accommodate the fewer number of target pairs and the additional conditions. Conditions were completely interleaved within observer to discourage the use of long term memory, which could allow observers to ignore the previewed relation; if a particular block of trials only tested interacting objects (hammer above a nail), observers might ignore the previewed relational information and respond present/match if the search pair matched their real world knowledge about the target objects. Conditions were counterbalanced across observers and trials by changing the previews and targets in the search display such that each trial matched/mismatched equally often and the preview was presented with an interacting/non-interacting direction equally often. As well, eight sets of unrelated object pairings were generated such that each observer received a unique condition/target pair/search scene pairing.

Results & Discussion

Relational guidance might be affected by the semantic relatedness of the target objects, the functional interaction between these objects (whether their positions suggest an interaction or not), or both. If relational guidance is dependent upon the semantic relatedness of the target objects, then target relatedness should interact with relational guidance (the difference in guidance between the match and swap trials). If relational guidance is dependent upon the position of the target objects, then the functional interaction should likewise interact with relational guidance. If however, relational guidance is dependent upon both the semantic relatedness and position of the target objects, there should be a three way interaction between relational guidance, semantic relatedness, and functional interaction, suggesting that coding relational information is affected by the LTM memory associations of both the target objects and their expected arrangements.

RT data revealed a significant main effect of matching target relations (Table 4, $F(1, 21) = 7.24$, $p = .01$), demonstrating that matched targets are searched for and recognized faster than swapped targets[2]. RT data also revealed a significant main effect of functional interaction, $F(1, 21) = 6.50$, $p = .02$, suggesting that interacting targets are searched for and recognized faster than non-interacting targets. All other main effects and interactions were not significant, all $Fs(1, 21) \leq 1.26$, all $ps \geq .28$. The target fixated first analysis (Figure 7) revealed significant main effects of relational guidance and functional interaction, as well as a significant interaction between relational guidance, semantic relatedness, and functional interaction, all $Fs(1, 21) \geq 4.68$, all $ps \leq .04$. All other effects were not significant, all $Fs(1, 21) \leq 0.69$, all $ps \geq .41$. These analyses show that matched targets generate stronger search guidance than swapped targets, as do non-interacting targets relative to interacting targets, although this latter effect of functional interaction may be driven by the three-way interaction. The interaction demonstrates that relational guidance is selectively expressed in the related-interacting condition and in the unrelated-non-interacting condition. This selective expression of relational guidance suggests that LTM associations for the target objects impact the coding of relational information, as relational guidance was expressed only when the target semantic relationship and the target functional interaction were consistent with real world expectations.

Lastly, the time until the target was fixated was analyzed (Table 4) to obtain a more complete measure of search guidance. This measure revealed a significant main effect relational guidance, $F(1, 21) = 7.66$, $p = .01$, and a significant interaction of relational guidance and semantic relatedness, $F(1, 21) = 8.07$, $p = .01$, all other effects were n.s. all $Fs(1, 21) \leq 0.81$, all $ps \geq .36$. This interaction was driven by large effects of relational guidance in both unrelated conditions and the lack of relational guidance in both related conditions. Whereas the target fixated first analysis revealed a lack of relational guidance in the unrelated-interacting condition, this more holistic measure of search guidance revealed a rather large effect of the relational guidance. One explanation for this might be that the interacting but unrelated target objects resulted in the relational information being encoded *after* search commenced, and thus relational guidance exerted its influence only on later measures of guidance. Alternatively, the relational information may not have become useful in this condition until after the search displays were further evaluated (i.e. an object had already been fixated). Future work may seek to alter the preview duration to tease apart these possibilities; if encoding the relational information simply took longer in this condition, a longer preview should result in relational guidance being expressed in the first object fixated analysis as well. The predicted lack of relational guidance in the related non-interacting condition may have resulted from interference between the LTM associations of the related objects (their expected interacting arrangement) and the actual perceived arrangement of the related objects (non-interacting). However, it is puzzling that the related interacting condition expressed relational guidance in the first object fixated measure (Figure 7) but *not* in the time to target measure (Table 4), especially given that relational guidance was expected to be expressed rather strongly in this condition.

These results are surprising in the context of the existing literature. It may be the case that related interacting objects are processed such that they result in more accurate recognition, yet when the task requires the encoding of relational information these processes may interfere with the relational coding. Future work will need to test guidance processes for functionally related interacting objects when the task does not require relational coding to see if the explicit relational coding in this task may contribute to this unexpected data pattern. Nevertheless, the results clearly demonstrate that the semantic relatedness of the target objects and the positions of the target objects modulate the expression of relational guidance, as well as when this information is available to guide search. These findings serve to further parameterize relational guidance, suggesting that the LTM associations of the target objects impact the coding and the timing of when relational guidance is expressed.

Experiment 5

Experiment 5 tested whether relational information is automatically obtained from a preview and used to guide search. If relational information is *not* coded automatically, but rather must be explicitly and effortfully encoded, then a task change which removes the necessity to attend to the relational information should affect the expression of relational guidance. This should result in the match and swap conditions interacting with the task; there should be no difference in guidance between the match and swap conditions when the observers' task is to localize the target pair irrespective of the relational information. A relational guidance effect should be replicated when the task requires the coding of relational information. Conversely, if relational information is coded automatically then relational guidance should be observed even when the task does not require the coding of relational information; relational guidance should *not* interact with the task.

Methods

Participants

Twenty-four undergraduate students from Stony Brook University participated for course credit. All had normal or corrected-to-normal vision and were native English speakers, by self-report.

Design and Procedure

Stimuli, apparatus, and procedure were identical to Experiment 2, with the following exceptions. In separate blocks, observers performed two tasks, one required the use of the relational information and the other did not. In one block of trials observers performed the same present/absent task as in Experiment 2 (relation relevant task). In another block the observer's task was to localize the target pair irrespective of the relational information and, while looking at the target pair, to press either of the top shoulder buttons to indicate their selection (relation-irrelevant task). If they were looking at a non-target pair when they pressed the shoulder button this was counted as an incorrect response. However, in both tasks 50% of the trials had a matched spatial relationship and 50% had a swapped spatial relationship. Only the two-target condition was tested. Search displays were identical across observers, however conditions were counterbalanced across trials by changing the order of the tasks and the relational information of the previewed objects such that each condition appeared equally often on each trial across observers.

Results & Discussion

If relational information is coded automatically then significant relational guidance should be observed in the relation-irrelevant task. As well, the amount of relational guidance should not interact with the task. However, if relational guidance requires effortful and explicit encoding, there should not be significant relational guidance in the relation-irrelevant task and the task should significantly interact with relational guidance.

Overall RTs were 951ms (56) and 1068ms (71) in the Relation Irrelevant Match and Swap conditions, respectively, and 1161ms (37) and 1303ms (53) in the Relation Relevant Match and Swap conditions, respectively (values in parentheses indicate standard error of the mean). Analysis of these data revealed a significant main effect of the relational information $F(1, 23) = 24.47$, $p < .001$, demonstrating that match trials were again faster than swap trials. Not surprisingly, the main effect of task was also significant, suggesting that the search judgment was faster when the relational information did not have to be verified, $F(1, 23) = 21.02$, $p < .001$. In both tasks the objects must be recognized and a motor response made, however, the relation relevant task also required recognizing the relational information and deciding which button to press and this generally took an additional ~200ms. Critically, the interaction between relational guidance and the task failed to reach significance $F(1, 23) = 1.04$, $p = .32$. To determine how much of these main effects were due to relational guidance, as opposed to verification, I again analyzed oculomotor measures of search. The proportion of trials in which the target was the first object fixated showed a marginally significant main effect of relational guidance, $F(1, 23) = 3.64$, $p = .07$, however, both the main effect of task and the interaction of task with relational guidance failed to reach significance (both $F(1, 23) \leq 0.95$, both $p \geq .33$, Relation Irrelevant Match 0.52 (0.02)and Swap 0.51 (0.02); Relation Relevant Match 0.53 (0.02)and Swap 0.49

(0.02)). As well, the time to target measure revealed a significant main effect of relational guidance, $F(1, 23) = 17.85$, $p < .001$, but again both the main effect of task and the interaction of task with relational guidance failed to reach significance (both $F(1, 23) \leq 1.53$, both $p \geq .22$, Relation Irrelevant Match 457ms (14)and Swap 476ms (16); Relation Relevant Match 451ms (10) and Swap 487ms (11)). Taken together these guidance measures suggest some degree of automatic encoding of relational information; relational guidance was found even when it was not required by the task.

However, further analysis revealed that this evidence for an automatic coding of spatial relations was subject to a practice or exposure effect. Analyzing the first block of trials and looking at the effect of task between subjects revealed a main effect of the relational information in time to target, $F(1, 22) = 5.42$, $p = .03$ (Table 5), and a trend in the same direction in first object fixated, $F(1, 22) = 1.90$, $p = .18$ (Figure 8); all effects of task and interactions were not significant, all $F(1, 22) \leq 0.34$, all $p \geq .60$. This supports the previous conclusion that there is some weak evidence for the automatic coding of relational information with minimal practice. However, when analyses were restricted to the second block of trials a significant interaction between task and relational guidance emerged in time to target, $F(1, 22) = 5.67$, $p = .03$ (Table 5), and a marginal interaction emerged in first object fixated, $F(1, 22) = 3.05$, $p = .09$ (Figure 8)). Relational guidance was evident in the relation relevant task (time to target, $t(22) = 3.48$, $p = .005$, and a marginal effect in first object fixated, $t(22) = 1.84$, $p = .09$), but there was no hint of relational guidance in the relation irrelevant task (time to target, $t(22) = 1.61$, $p = .13$, first object fixated, $t(22) = 0.37$, $p = .71$). Taken together these findings suggest that there may be some small automatic contribution to relational guidance with minimal practice (see footnote 3 for an additional discussion related to the effects of practice), but that this disappears with greater exposure to the task as observers learn to code the relational information more strongly in the relation relevant task and learn not to code the relational information in the relation irrelevant task. All of these findings suggest a degree of control in the coding of relational information for the purpose of guidance that is not normally associated with an automatic process.

Experiment 6

In this experiment, grouping effects on relational guidance were tested so as to inform existing search theory as to how relational information may be coded. Depending upon how the relational information is represented, grouping may or may not affect the weighting of relational information for search guidance. One method of representing relational information might be to combine the two objects into a single representation, essentially ignoring the object boundaries and implicitly capturing the relational information between the objects' features (referred to earlier as the unified percept hypothesis). Such grouping would likely be a low-level process, one that potentially even metrically represents the space between the two unitized objects. The partial evidence for automatic relational guidance in Experiment 5 is also consistent with low-level visual coding of relational information. Yet effortful encoding of relational information does seem to magnify relational guidance after practice (Experiment 5), suggesting that there may be more explicit processes operating in addition to the lower-level processes. Such an explicit code might keep separate the features of the two objects, connecting them only by the use of a higher-level and an explicit spatial reference (referred to earlier as the explicit reference hypothesis). According to this hypothesis, features might be extracted from the two objects, with these objects linked by the spatial reference. And because an explicit spatial reference is categorical, one would not expect it to be affected by grouping cues or metrical changes in distance; X to the left of Y will remain to the left of Y regardless of how strongly X and Y are grouped. If humans code relational information using an explicit categorical reference, it need not be affected by grouping cues.

If relational information is coded in the form of a low-level unitized percept, it might be affected by grouping cues, as such grouping effects are known to be mediated by low-level processes. However, if relational information is coded using an explicit categorical reference, then grouping cues would not be expected to have an effect. Experiment 6 increased both perceptual and semantic grouping by placing a picture frame around each pair of objects.

Methods

Participants

Fourty-eight undergraduate students from Stony Brook University participated for course credit. All had normal or corrected-to-normal vision and were native English speakers, by self-report.

Design and Procedure

Stimuli, apparatus, and procedure were identical to Experiment 2, with the following exceptions. Each pair of objects was placed on top of an image of a picture frame, both at preview and in the search display (see Figure 9), thereby creating the perception of a framed photograph of two random objects. The frame was sized to accommodate the widest and the tallest pair of objects on a trial by trial basis. All distractor pairs were presented in the same orientation as the target pair to allow the frame to be the same height and width across all pairs of objects on a given trial. The presence of the frame was a between subjects factor to prevent carry-over effects.

Results & Discussion

The no frame baseline condition should serve as a direct replication of Experiment 2, confirming that relational guidance occurs despite the orientation of the distractor pairs (see also Experiment 1). If the presence of the frame increases grouping, and relational information is coded as a unified percept, relational guidance should increase, interacting with the presence of the frame. In addition to this interaction, the presence of the frame may also produce poorer search performance and lower guidance overall relative to the no-frame baseline condition. This is because the placement of an identical frame behind each pair of objects would increase target-distractor similarity, a factor known to reduce search guidance and search efficiency (Alexander & Zelinsky, 2011; 2012; Duncan & Humphreys, 1989).

Analysis of the RT data in the no-frame baseline condition (Table 6) revealed that search was significantly faster in the matched condition relative to the swapped condition, $F(1, 23) = 46.58$, $p < .001$, however, both the main effect of number of previewed objects and the interaction of the relational information with the number of previewed objects failed to reach significance, both $Fs\ (1, 23) \leq 2.30$, both $ps \geq .14$. As can be observed in Figure 10, the target first object analysis revealed significant relational guidance, $F(1, 23) = 15.14$, $p = .001$, however, the effect of the number of previewed objects and the interaction of the two again failed to reach significance, both $Fs\ (1, 23) \leq 1.71$, both $ps \geq .20$. The time to target data (Table 6) showed clear relational guidance, $F(1, 23) = 52.70$, $p < .001$, and an advantage for having previewed both targets $F(1, 23) = 5.87$, $p = .02$, although without a hint of an interaction $F(1, 23) = .30$, $p = .59$. This pattern of data generally replicates the results of Experiment 2, confirming that the orientation of the distractors does not impact relational guidance.

The presence of the frame should result in a similar but magnified pattern of data if a unified percept is used to code relational information. Analysis of the RT data in the frame condition (Table 6) revealed that search was significantly faster in the matched condition relative to the swapped condition, $F(1, 23) = 27.27$, $p < .001$, and with a two-object preview relative to a one-object preview, $F(1, 23) = 6.29$, $p = .02$, but the interaction of the two failed to reach significance, $F(1, 23) = 0.28$, both $p = .60$. As well, the target first object analysis (Figure 10) revealed significant relational guidance, $F(1, 23) = 8.00$, $p = .01$, but the number of previewed objects and the interaction of the two failed to reach significance, both $Fs(1, 23) \leq 2.20$, both $ps \geq .15$. Lastly, the time to target data (Table 6) again showed significant relational guidance, $F(1, 23) = 15.48$, $p = .001$, however both the effect of number of previewed objects and the interaction of the two failed to reach significance, both $Fs(1, 23) \leq 0.83$, both $ps \geq .37$. These results generally confirm the effects of relational guidance on search when a frame is present; however the two object advantage was only evident in the RT measure, and not in the time to target measure.

The critical test of the unified percept hypothesis is the comparison between the frame and the no-frame conditions, which should reveal more relational guidance with a frame if relational information is coded using a unified percept. For RT, proportion target is the first object fixated, and in the time to the target measure, the presence of the frame failed to interact with relational guidance, all $Fs(1,46)\ ) \leq 0.98$, all $ps \geq .33$, suggesting that the increased grouping associated with the presence of the frame did not improve relational guidance (all other interactions were also n.s.). The presence of the frame did, however, produce marginally worse search guidance (both in proportion target is the first object fixated and in time to target, both $Fs(1, 46) \geq 2.68$, both $ps \leq .10$), likely because of the increased target-distractor similarity, although this decreased guidance failed to produce significantly slower search as measured by overall RTs, $F(1,46) = 0.57$, $p=.46$. Taken together these results provide no support for the unified percept hypothesis, and offer indirect support the explicit reference hypothesis.

Experiment 7

Experiment 6 tested a relatively novel grouping procedure and provided some initial evidence against the unified percept hypothesis. Experiment 7 again tests the effects of grouping on relational guidance, but this time using a more standard grouping principle. Proximity is generally considered a very strong grouping cue (e.g. Shihui, 2004), and therefore may be more effective at modulating how strongly relational information is represented. Grouping by proximity should be strongest when the objects are close, and should decrease as the distance between the objects increases. When grouping is strong, relational guidance should have its maximal effect, if a unified percept is used. Conversely, as the distance between the objects increases, relational guidance should decrease, again assuming a unified percept is used.

Methods

Participants

Twenty-four undergraduate students from Stony Brook University participated for course credit. All had normal or corrected-to-normal vision and were native English speakers, by self-report.

Design and Procedure

Stimuli, apparatus, and procedure were identical to the Experiment 6 no-frame condition, with the following exceptions. The distance between the objects in each pair (both at preview and in the search display) was manipulated in near and far conditions (see Figure 11). Distances between the object pairs were either 80 pixels (1.84°, double the distance from all prior experiments; far distance condition), relative to the smallest bounding box enclosing each object, or eliminated (close distance condition). Irregularly shaped targets or targets at a steep orientation were replaced so as to allow at least one point where the objects were very close to each other in the close distance condition (bounding boxes touching and the actual objects within the bounding boxes being less than 10 pixels from each other in both the match and swap conditions). The close condition allows for the relational information to be more efficiently coded relative to a single point, while the far distance condition should make this difficult. Only the two preview condition was tested as a distance manipulation should not affect a one object preview. Distance conditions were blocked within subject, and distance between the objects and relational match were counterbalanced across observers and trials.

Results & Discussion

If increased grouping results in a more effective coding of the relational information, then the close condition should have a larger difference in magnitude between the matched and swapped trials when compared to the far condition, resulting in an interaction of relational guidance and distance. Finding such an interaction between relational guidance and grouping would be consistent with the predictions of the unified percept hypothesis. However, if grouping is found not to affect relational guidance this would again provide indirect support for the explicit reference hypothesis.

Analysis of the RT data (Table 7) revealed a clear effect of the relational information, $F(1, 23) = 29.41$, $p < .001$, demonstrating that match trials were again faster than swap trials. However, both the main effect of distance and the interaction of distance and the relational information failed to reach significance, both $F(1, 23) \leq 1.01$, both $p \geq .32$. As can be observed in Figure 12 the target was fixated first significantly more often when the relational information matched the preview compared to when the objects swapped relative positions, $F(1, 23) = 30.21$, $p < .001$, replicating the effect of relational guidance. The target pair was also fixated first significantly more often in the far distance condition compared to the close distance condition, $F(1, 23) = 8.93$, $p = .01$, suggesting that guidance was in fact somewhat stronger when the target objects were widely separated. One possible explanation for this is a relationship between separation and the eccentricity of the closest object; as the distance between the objects increased, one object always moves closer to central fixation which may have increased guidance (see also Carrasco, Evert, Chang, & Katz, 1995; Wolfe, O'Neill, & Bennett, 1998). Alternatively, this may have occurred because clutter increased as the objects moved closer together (e.g. Rosenholtz, Li, & Nakano, 2007), making it more difficult to detect the target locations. Critically, distance and relational guidance did not interact, $F(1, 23) = 0.02$ $p = .90$. Analysis of the time until the target is fixated (Table 7) revealed a highly significant advantage for the matched compared to the swapped targets, $F(1, 23) = 30.07$ $p < .001$, but like the RT data both the effect of distance and the interaction of distance and the relational information failed to reach significance, both $F(1, 23) \leq 0.11$, both $p \geq .74$. This suggests that the guidance benefit for the far condition observed in the target first object analysis was a rather short-lived benefit[4].

Taken together these results and the findings from Experiment 6, suggest that relational guidance is unaffected by grouping cues. This provides converging evidence against the unified percept hypothesis, and indirect evidence in support of the explicit reference hypothesis[5]. Future work may also test the unified percept hypothesis by systematically changing the distance between the objects at preview and at search, as this might be expected to disrupt a low-level metrical coding of the objects' spatial relationship but should not affect a categorical coding of the relational information.

General Discussion

In a series of seven experiments, the relational information between objects was demonstrated to affect search guidance. Current theory suggests that focal attention is required to determine relational information between objects (Logan, 1994, 1995). If this is true, it should not be possible to use relational information to pre-attentively guide search. Experiment 1 used the design and parameters of Logan's now seminal work, in conjunction with real-world objects and more sensitive eye movement dependent measures, and demonstrated that relational guidance is indeed possible under these conditions (Aim 1). Experiment 1 used extremely homogenous displays to rule out guidance by anything other than relational information. Experiments 2-7 parameterized the use, generalizability and boundary conditions in which relational information affects search guidance (Aim 2). Experiment 2 increased the variability in the search displays to show that relational guidance persists even when specific target features are available to guide search. Experiment 3 removed the pictorial previews and linguistically cued the targets, and found that this manipulation resulted in no relational guidance. Experiment 4 examined how functional relations and the relative positions of the target objects affect relational guidance. Experiment 5 addressed the automaticity of relational guidance, and in Experiments 6 and 7 relational guidance was found to *not* be affected by the manipulation of object grouping.

Despite a close replication of the design and parameters from Logan (1994 experiment 5), our interpretation of the data led us to fundamentally different conclusions. Logan suggested that attention must first be directed to a pair of objects in order to determine their spatial relationship, precluding pre-attentive search guidance by relational information. There are two likely reasons why relational guidance was found in the present study but not by Logan (1994).

First, Logan used highly self-similar pluses and dashes whereas the present work used real-world color objects. The perceptual self-similarity of the pluses and dashes likely weakened any relational guidance that may have existed. As well, the pluses and dashes of one arrangement were targets on some trials and distractors on other trials, a phenomenon known to cause less efficient search (e.g. Becker, 2008; Lamy, Yashar, & Ruderman, 2010; Maljkovic & Nakayama, 1994). The use of trial-unique, real-world, color objects likely maximized the chances of finding relational guidance in this study. One could certainly envision a scenario in which, relational guidance scales with the saliency of the arranged objects. Objects that create a high contrast, very salient (Itti & Koch, 2000) boundary may allow for the relational information to be more easily discerned in peripheral vision when compared to less salient, low contrast pairs. Future work may seek to explore this potentially interesting relationship. More importantly, however, Logan arrived at his conclusion because search for a specific relation is highly inefficient, as measured by the change in RT as a function of the change in the number of search objects. Specifically, RTs increased as items were added to the search display, resulting in a *large* search slope. Efficient search is easy to determine; search slopes that are near zero indicate that adding additional distractors does not affect search speed, a phenomenon known as pop-out (Treisman & Gelade, 1980). Of course, in real world search targets rarely pop-out. The feature complexity of objects instead causes search guidance to exist on a continuum; search is guided strongly to some objects (perhaps not even requiring an eye movement to detect a target) and weakly to others (perhaps resulting in many eye movements and inefficient search). Given this range of slopes commonly reported in the search literature, quantifying inefficient search becomes a problem. Slopes greater than ~30ms per item are often taken as evidence for a lack of guidance (e.g. Wolfe & Horowitz, 2004), yet Logan observed slopes in excess of ~85ms per item in target present displays and ~125ms per item in target absent displays. What does this mean for actual search guidance? As Experiment 5 showed, verifying relational information is a very time consuming process. By chance alone, as the set-size increases distractors are more likely to be visited; given how long relational judgments take, search will likely be inefficient due to recognition factors rather than guidance factors. Inefficient search therefore doesn't indicate a lack of guidance. Measures of eye-movements make it possible to determine a clear chance baseline; a lower-bound on search performance indicating the complete absence of guidance. The presence of relational guidance in this task, suggests that tasks considered highly inefficient (e.g. Greene & Wolfe, 2011; Wolfe & Myers, 2010) need not be unguided in terms of actual eye-movements to the target. In this study I demonstrated that search can be guided by a non-basic (i.e., constructed) visual feature, a spatial pointer indicating the spatial relationship between two objects. This challenges the commonly held belief that "Efficient search is, therefore, a necessary but not sufficient property for showing the presence of a guiding feature" (Wolfe & Horowitz, 2004, pp. 3) and calls into question the very definition of search guidance. If one defines guidance as the above-chance direction of eye movements to a target rather than some arbitrary efficiency cut-off, then the range of guiding features will broaden far beyond the basic features (e.g., color, orientation, etc.) typically considered by this literature and will include a variety of constructed features, many of which may produce quite inefficient search.

Experiment 1 clearly demonstrated relational guidance; however, it did so in the context of highly homogenous displays so as to rule out guidance from features of the target objects. Experiment 2 showed that relational guidance generalizes to a more naturalistic context. Due to the heterogeneity of objects in the real-world, features of the target objects are commonly available and used to guide search. By replicating relational guidance in the context of multiple orientations and trial unique distractors, this experiment demonstrated that relational guidance persists despite the presence of target feature guidance, and that it is not limited to the somewhat unnatural conditions described in Experiment 1 and Logan (1994). More generally, this demonstration suggests that relational guidance may contribute to more natural search settings and situations outside of the laboratory. On this note, in the real-world relational guidance may indeed influence search to a far greater extent than what was observed in these tasks. Because knowing that mugs generally sit on flat surfaces impacts what regions are inspected (Torralba, Oliva, Castelhano, & Henderson, 2006), it is likely that relational guidance would interact with these contextual cues, perhaps increasing its influence.

Experiment 3 sought to further test the generalizability of relational guidance by removing the pictorial preview and cuing the target objects linguistically with a text based cue. Targets were referenced with a categorical text label referring to one or two out of four target classes. This experiment failed to show reliable relational guidance, suggesting that target features may need to be highly specific in order to effectively code relational information between them. An interesting direction for future work will be to determine if relational guidance is specific to pictorial previews, or if recalling specific objects from LTM results in target features precise enough to effectively code relational information.

Experiment 4 further parameterized relational guidance by showing that it is affected by prior expectations of the functional and semantic relations between the target objects, and that it is selectively expressed in certain conditions and at certain times during search. Early measures of search, such as the proportion of trials in which the target was the first pair fixated, showed that related-interacting object pairs and unrelated-non-interacting object pairs selectively showed relational guidance. However, later guidance measures, such as the time to fixate the target, showed relational guidance only in the two unrelated conditions. The only condition that completely failed to exhibit relational guidance across measures was the related-non-interacting condition, perhaps due to the relational information conflicting with the expected object arrangement from LTM. Interestingly, this is the precise condition under which extremely poor recognition has been previously reported in the literature (e.g. Green & Hummel, 2006). This clearly demonstrates that LTM associations for the target objects and their expected arrangement impact the use of relational information. However, and unlike the object verification paradigm used in previous studies, when the task requires encoding relational information and search guidance is the measure of performance, there is not a large difference between the related-interacting and the related-non-interacting condition in later measures of search. Perhaps the encoding of relational information in this search task interfered with the later verification processes that are needed to find differences between interacting and non-interacting objects. Future work might explore the possibility that related objects may be processed differently when relational information must be encoded versus when it is superfluous to the task. Lastly, the unrelated-non-interacting condition seemed to exhibit delayed relational guidance, as it was found only after the first object had been fixated. Perhaps simply presenting a longer preview would produce search guidance in early measures as well in this condition.

Experiment 5 moved beyond parameterizing relational guidance to investigating how relational information is coded by asking if it is automatically available or if it must be explicitly and effortfully encoded to produce relational guidance. Results initially supported the automatic coding of relational information, but further analysis showed that after sufficient practice relational guidance was significantly modulated by the task. When the task required an explicit coding of the relational information, relational guidance was evident; when the task did not require coding of the relational information, the target objects were coded in a spatially invariant way such that guidance was unaffected by the relational information. It is unclear from these results what initially caused the coding of relational information even when it was irrelevant to the task. Perhaps observers needed some initial practice with the stimuli to learn to code the objects in a spatially invariant way. This would suggest that the relational information may initially be automatically primed. Alternatively, the initial instruction period which described both tasks may have biased observers to code the relational information even when it was irrelevant to the task. Future work will have to test this using a between subjects design to determine if this was the result of the initial instruction period or if relational information is initially automatically processed to some minimal extent. What is clear is that after some practice relational guidance is modulated by task requirements. This suggests that the process of encoding the relational information for the manual response automatically results in a change to the guiding template, suggesting that both guidance and recognition procedures may in fact use the same representation.

Experiments 6 and 7 showed that relational guidance is completely unaffected by various grouping cues. When each pair of objects was placed on top of a picture frame to create the percept of a framed picture of two random objects, this explicit unitization of the objects should have increased both semantic and perceptual grouping processes. Similarly, the manipulation of the separation between objects using near and far conditions should have differentially engaged low-level grouping by proximity. However, in both experiments grouping did not impact the amount of relational guidance that was produced. These results suggest that object spatial relationships are not coded by low-level visual processes, at least for those processes that would be expected to be modulated by grouping cues. Rather, relational information between objects for the purposes of guiding search is likely to be coded by higher order spatial pointers. This observation, combined with the results from Experiment 5, provides theoretical insight into the mechanism of relation guidance, suggesting that the coding of spatial relationships between objects may in some sense be categorical (left of, above, etc.) and not the type of metrical coding associated with lower-level visual processes.

In addition to demonstrating relational guidance this work also has clear implications for multiple target search. Multiple target search is generally more difficult than single target search (e.g. Kaplan & Carvellas, 1965; Menneer et al., 2007; Menneer et al., 2009; Menneer et al., 2010; Moore & Osman, 1993; Yang & Zelinsky, in preparation) however, the reason for this is debated. Search could be more difficult because the features from both targets are combined or unified, allowing for guidance from both targets simultaneously. Alternatively, search could be more difficult because it is guided only by one of the targets, then after fixating an object, both target representations would be compared to the visual pattern for the purposes of making a target or distractor judgment. The former option would create poorer search performance by including too many target features (the other target); the latter option would create poorer search performance due to the wrong target being searched for on some percentage of trials. This work suggests that neither view is entirely correct. The grouping by frame (Experiment 6) and the grouping by proximity (Experiment 7) studies argue against a unitized percept of the two targets. However, this work also argues against search guidance being entirely based on a single target, as evidenced by the repeated 2 target advantage and guidance by relation information between objects. Taken together, this work suggests that it is possible to simultaneously represent the features from both objects, but that the features of each object remain separate and in this case, linked by a spatial relationship. However, it is also worth noting that the current studies depicted two targets in the search display simultaneously and in close spatial proximity, a procedure not typically used in the multiple target search literature. The multiple target aspect of the current task also places a sort of upper bound on the level of guidance that might be expected during search, which could be another reason why some of the reported effects of relational guidance were small.

In conclusion, these experiments clearly demonstrate that object spatial relationships can be encoded into the target representation formed at preview, possibly through the use of a categorical explicit reference, and that this relational information is then available to guide search. This conceptualization of relational guidance removes the need to assume a pre-attentive coding of relational information between peripherally viewed search objects, which was the dominant view embraced by previous theory on this topic (Logan, 1994). Moreover, previous theory had assumed that the inefficient search that is typically observed to targets defined by relational information meant that that this search was unguided; the small but extremely robust effect of relational guidance that is reported in this series of experiments requires theories to now include a relational feature among the other appearance-based and contextual features that are known to guide search.

Endnotes

[1] Observers took significantly longer to fixate the first pair of objects after a long preview compared a short preview $F(1, 23) = 7.12$, $p = .01$, yet achieved similar levels of guidance as measured by the proportion of trials in which the target was the first pair fixated. This speed-accuracy trade-off is consistent with a general slowing of search speed, suggesting an attentional effect.

[2] Two subjects were removed from all analyses because they failed to look at the target on the majority of trials, making it impossible to reliably calculate the proportion of trials in which the target is the first pair fixated and the time to fixate the target for these subjects.

[3] Re-examining experiments 1 & 2 revealed similar trends towards practice effects. However, the blocked design resulted in between subjects' comparisons that generally failed to reach significance. Interestingly, the two-preview advantage appears to primarily emerge from early trials. Thus, with practice observers may learn to improve one-preview search and to code the relational information more effectively. Future work will have to more closely examine these potentially very interesting practice effects in the context of an interleaved design.

[4] Observes took marginally longer to fixate the first pair of objects in the far distance condition compared to the close distance condition $F(1, 23) = 3.62$, $p = .07$. This speed-accuracy trade-off may have contributed to the more accurate target first fixations in the far distance condition.

[5] A version of this study in which the objects overlapped by 1-5 pixels was also conducted (N=40). Connecting the pair of objects so as to create the percept of a single object (Logan, 1994) should have enabled the relational information to be calculated relative to a single point. However, an identical pattern of results was obtained.

References

Alexander, R. G., & Zelinsky, G. J. (2011). Visual similarity effects in categorical search. *Journal of Vision, 11*(8). doi: 10.1167/11.8.9

Alexander, R. G., & Zelinsky, G. J. (2012). Effects of part-based similarity on visual search: The Frankenbear experiment. *Vision Research, 54*(0), 20-30. doi: 10.1016/j.visres.2011.12.004

Alexander, R. G., & Zelinsky, G. J. (in preparation) Part-based target representations in visual search.

Alvarez, G. A., & Oliva, A. (2008). The representation of simple ensemble visual features outside the focus of attention. *Psychological Science, 19*(4), 392-398.

Alvarez, G. A., & Oliva, A. (2009). Spatial ensemble statistics are efficient codes that can be represented with reduced attention. *Proceedings of the National Academy of Sciences, 106*(18), 7345-7350. doi: 10.1073/pnas.0808981106

Ashley, A., & Carlson, L. A. (2007). Encoding direction when interpreting proximal terms. *Language and Cognitive Processes, 22*(7), 1021-1044. doi: 10.1080/01690960701190298

Bar, M., & Ullman, S. (1996). Spatial context in recognition. *Perception, 25*(3), 343-352. doi: 10.1068/p250343

Becic, E., Kramer, A. F., & Boot, W. R. (2007). Age-related differences in the use of background layout in visual search. *Aging, Neuropsychology, and Cognition, 14*(2), 109-125. doi: 10.1080/13825580701202167

Becker, S. I. (2008). The mechanism of priming: Episodic retrieval or priming of pop-out? Acta Psychologica, 127(2), 324-339. doi: 10.1016/j.actpsy.2007.07.005

Beun, R.-J., & Cremers, A. H. M. (1998). Object reference in a shared domain of conversation. *Pragmatics & Cognition, 6*(1-2), 121-152.

Brockmole, J. R., & Henderson, J. M. (2006). Recognition and attention guidance during contextual cueing in real-world scenes: evidence from eye movements. Quarterly Journal Of Experimental Psychology (2006), 59(7), 1177-1187.

Byrne, P. A., & Crawford, J. D. (2010). Cue reliability and a landmark stability heuristic determine relative weighting between egocentric and allocentric visual information in memory-guided reach. *Journal of Neurophysiology, 103*(6), 3054-3069. doi: 10.1152/jn.01008.2009

Carlson-Radvansky, L. A., Covey, E. S., & Lattanzi, K. M. (1999). 'What' effects on 'where': Functional influences on spatial relations. *Psychological Science, 10*(6), 516-521. doi: 10.1111/1467-9280.00198

Carlson, L. A., & Logan, G. D. (2001). Using spatial terms to select an object. *Memory & Cognition, 29*(6), 883-892.

Carlson, L. A., Regier, T., Lopez, W., & Corrigan, B. (2006). Attention unites form and function in spatial language. *Spatial Cognition and Computation, 6*(4), 295-308. doi: 10.1207/s15427633scc0604_1

Carlson, L. A., & Van Deman, S. R. (2004). The space in spatial language. *Journal of Memory and Language, 51*(3), 418-436. doi: 10.1016/j.jml.2004.06.004

Carlson, L. A., West, R., Taylor, H. A., & Herndon, R. W. (2002). Neural correlates of spatial term use. *Journal of Experimental Psychology: Human Perception and Performance, 28*(6), 1391-1407. doi: 10.1037/0096-1523.28.6.1391

Carrasco, M., Evert, D., Chang, I., & Katz, S. (1995). The eccentricity effect: Target eccentricity affects performance on conjunction searches. Attention, Perception, & Psychophysics, 57(8), 1241-1261. doi: 10.3758/bf03208380

Chun, M. M. (2000). Contextual cueing of visual attention. *Trends in Cognitive Sciences, 4*(5), 170-178. doi: 10.1016/s1364-6613(00)01476-5

Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology, 36*(1), 28-71. doi: 10.1006/cogp.1998.0681

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review, 96*(3), 433-458.

Engel, S. A., Glover, G. H., & Wandell, B. A. (1997). Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cerebral Cortex, 7*(2), 181-192. doi: 10.1093/cercor/7.2.181

Foo, P., Warren, W. H., Duchon, A., & Tarr, M. J. (2005). Do Humans Integrate Routes Into a Cognitive Map? Map- Versus Landmark-Based Navigation of Novel Shortcuts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(2), 195-215. doi: 10.1037/0278-7393.31.2.195

Franconeri, S. L., Scimeca, J. M., Roth, J. C., Helseth, S. A., & Kahn, L. E. (In Press). Flexible visual processing of spatial relationships. *Cognition*(0). doi: 10.1016/j.cognition.2011.11.002

Garnham, A. (1989). A unified theory of the meaning of some spatial relational terms. *Cognition, 31*(1), 45-60. doi: 10.1016/0010-0277(89)90017-6

Green, C., & Hummel, J. E. (2006). Familiar interacting object pairs are perceptually grouped. *Journal of Experimental Psychology: Human Perception and Performance, 32*(5), 1107-1119. doi: 10.1037/0096-1523.32.5.1107

Greene, M. R., & Wolfe, J. M. (2011). Global image properties do not guide visual search. Journal of Vision, 11(6). doi: 10.1167/11.6.18

Gronau, N., Neta, M., & Bar, M. (2008). Integrated contextual representation for objects identities and their locations. *Journal of Cognitive Neuroscience, 20*(3), 371-388. doi: 10.1162/jocn.2008.20027

Hollingworth, A. (2006). Scene and position specificity in visual memory for objects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*(1), 58-69. doi: 10.1037/0278-7393.32.1.58

Hollingworth, A. (2007). Object-position binding in visual memory for natural scenes and object arrays. *Journal of Experimental Psychology: Human Perception and Performance, 33*(1), 31-47. doi: 10.1037/0096-1523.33.1.31

Hollingworth, A., & Rasmussen, I. P. (2010). Binding objects to locations: The relationship between object files and visual working memory. *Journal of Experimental Psychology: Human Perception and Performance, 36*(3), 543-564. doi: 10.1037/a0017836

Huang, L., & Pashler, H. (2005). Attention capacity and task difficulty in visual search. *Cognition, 94*(3), B101-B111. doi: 10.1016/j.cognition.2004.06.006

Hubel, D., & Livingstone, M. (1987). Segregation of form, color, and stereopsis in primate area 18. *The Journal of Neuroscience, 7*(11), 3378-3415.

Hummel, J. E., & Biederman, I. (1992). Dynamic binding in a neural network for shape recognition. *Psychological Review, 99*(3), 480-517. doi: 10.1037/0033-295x.99.3.480

Humphreys, G. W., Yoon, E. Y., Kumar, S., Lestou, V., Kitadono, K., Roberts, K. L., & Riddoch, M. J. (2010). The interaction of attention and action: From seeing action to acting on perception. *British Journal of Psychology, 101*(2), 185-206. doi: 10.1348/000712609x458927

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research, 40*(10-12), 1489-1506.

Jiang, Y., & Chun, M. M. (2001). Selective attention modulates implicit learning. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 54A*(4), 1105-1124. doi: 10.1080/02724980042000516

Jiang, Y., Olson, I. R., & Chun, M. M. (2000). Organization of visual short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(3), 683-702. doi: 10.1037/0278-7393.26.3.683

Kaplan, I. T., & Carvellas, T. (1965). SCANNING FOR MULTIPLE TARGETS. *Perceptual and Motor Skills, 21*(1), 239-243. doi: 10.2466/pms.1965.21.1.239

Kosslyn, S. M., Chabris, C. F., Marsolek, C. J., & Koenig, O. (1992). Categorical versus coordinate spatial relations: Computational analyses and computer simulations. *Journal of Experimental Psychology: Human Perception and Performance, 18*(2), 562-577. doi: 10.1037/0096-1523.18.2.562

Kroll, N. E., Schepeler, E. M., & Angin, K. T. (1986). Bizarre imagery: The misremembered mnemonic. Journal of Experimental Psychology: Learning, Memory, and Cognition, 12(1), 42-53. doi: 10.1037/0278-7393.12.1.42

Lamy, D., Yashar, A., & Ruderman, L. (2010). A dual-stage account of inter-trial priming effects. Vision Research, 50(14), 1396-1401. doi: 10.1016/j.visres.2010.01.008

Levinson, S. C. (1996). Frames of reference and Molyneux's question: Crosslinguistic evidence. In P. Bloom, M. A. Peterson, L. Nadel & M. F. Garrett (Eds.), *Language and space.* (pp. 109-169). Cambridge, MA US: The MIT Press.

Livingstone, M., & Hubel, D. (1984). Anatomy and physiology of a color system in the primate visual cortex. *The Journal of Neuroscience, 4*(1), 309-356.

Logan, G. D. (1994). Spatial attention and the apprehension of spatial relations. *Journal of Experimental Psychology: Human Perception and Performance, 20*(5), 1015-1036. doi: 10.1037/0096-1523.20.5.1015

Logan, G. D. (1995). Linguistic and conceptual control of visual spatial attention. *Cognitive Psychology, 28*(2), 103-174.

Logan, G. D., & Compton, B. J. (1996). Distance and distraction effects in the apprehension of spatial relations. *Journal of Experimental Psychology: Human Perception and Performance, 22*(1), 159-172. doi: 10.1037/0096-1523.22.1.159

Logan, G. D., & Sadler, D. D. (1996). A computational analysis of the apprehension of spatial relations. In P. Bloom, M. A. Peterson, L. Nadel & M. F. Garrett (Eds.), *Language and space.* (pp. 493-529). Cambridge, MA US: The MIT Press.

Mack, S. C., & Eckstein, M. P. (2011). Object co-occurrence serves as a contextual cue to guide and facilitate visual search in a natural viewing environment. *Journal of Vision, 11*(9). doi: 10.1167/11.9.9

Maljkovic, V., & Nakayama, K. (1994). Priming of pop-out: I. Role of features. Memory & Cognition, 22(6), 657-672.

Mandler, J. M., & Johnson, N. S. (1976). Some of the thousand words a picture is worth. *Journal of Experimental Psychology: Human Learning and Memory, 2*(5), 529-540. doi: 10.1037/0278-7393.2.5.529

Marsh, H. L., Spetch, M. L., & MacDonald, S. E. (2011). Strategies in landmark use by orangutans and human children. *Animal Cognition, 14*(4), 487-502. doi: 10.1007/s10071-011-0382-9

Meade, J., Biro, D., & Guilford, T. (2005). Homing pigeons develop local route stereotypy. *Proceedings of the Royal Society B: Biological Sciences, 272*(1558), 17-23. doi: 10.1098/rspb.2004.2873

Menneer, T., Barrett, D., Phillips, L., Donnelly, N., & Cave, K. (2007). Costs in searching for two targets: dividing search across target types could improve airport security screening. *Applied Cognitive Psychology, 21*(7), 915-932.

Menneer, T., Cave, K. R., & Donnelly, N. (2009). The cost of search for multiple targets: Effects of practice and target similarity. *Journal of Experimental Psychology: Applied, 15*(2), 125-139. doi: 10.1037/a0015331

Menneer, T., Donnelly, N., Godwin, H. J., & Cave, K. R. (2010). High or low target prevalence increases the dual-target cost in visual search. *Journal of Experimental Psychology: Applied, 16*(2), 133-144. doi: 10.1037/a0019569

Moore, C. M., & Osman, A. M. (1993). Looking for two targets at the same time: One search or two? *Perception & Psychophysics, 53*(4), 381-390.

Neider, M. B., & Zelinsky, G. J. (2006). Searching for camouflaged targets: Effects of target-background similarity on visual search. *Vision Research, 46*(14), 2217-2235.

Newman, E. L., Caplan, J. B., Kirschen, M. P., Korolev, I. O., Sekuler, R., & Kahana, M. J. (2007). Learning your way around town: How virtual taxicab drivers learn to use both layout and landmark information. *Cognition, 104*(2), 231-253. doi: 10.1016/j.cognition.2006.05.013

Olson, I. R., & Marshuetz, C. (2005). Remembering 'what' brings along 'where' in visual working memory. *Perception & Psychophysics, 67*(2), 185-194.

Palmer, J. (1994). Set-size effects in visual search: The effect of attention is independent of the stimulus for simple tasks. *Vision Research, 34*(13), 1703-1721. doi: 10.1016/0042-6989(94)90128-7

Peterson, M. S., Boot, W. R., Kramer, A. F., & McCarley, J. S. (2004). Landmarks help guide attention during visual search. *Spatial Vision, 17*(4-5), 497-510. doi: 10.1163/1568568041920230

Posner, M. I. (1980). Orienting of attention. *The Quarterly Journal of Experimental Psychology, 32*(1), 3-25. doi: 10.1080/00335558008248231

Regier, T., & Carlson, L. A. (2001). Grounding spatial language in perception: An empirical and computational investigation. *Journal of Experimental Psychology: General, 130*(2), 273-298. doi: 10.1037/0096-3445.130.2.273

Riddoch, M. J., Humphreys, G. W., Edwards, S., Baker, T., & Willson, K. (2003). Seeing the action: neuropsychological evidence for action-based effects on object selection. Nature Neuroscience, 6(1), 82-89.

Riddoch, M. J., Pippard, B., Booth, L., Rickell, J., Summers, J., Brownson, A., & Humphreys, G. W. (2011). Effects of action relations on the configural coding between objects. Journal of Experimental Psychology: Human Perception and Performance, 37(2), 580-587. doi: 10.1037/a0020745 10.1037/a0020745.supp (Supplemental)

Roberts, K., & Humphreys, G. (2011a). Action-related objects influence the distribution of visuospatial attention. The Quarterly Journal of Experimental Psychology, 64(4), 669-688. doi: 10.1080/17470218.2010.520086

Roberts, K., & Humphreys, G. (2011b). Action relations facilitate the identification of briefly-presented objects. Attention, Perception, & Psychophysics, 73(2), 597-612. doi: 10.3758/s13414-010-0043-

Rosenholtz, R., Huang, J., & Ehinger, K. A. (2012). Rethinking the role of top-down attention in vision: effects attributable to a lossy representation in peripheral vision. [Hypothesis & Theory]. Frontiers in Psychology, 3. doi: 10.3389/fpsyg.2012.00013

Rosenholtz, R., Li, Y., & Nakano, L. (2007). Measuring visual clutter. Journal of Vision, 7(2), 1-22.

Schmidt, J., & Zelinsky, G. J. (2009). Search guidance is proportional to the categorical specificity of a target cue. *The Quarterly Journal of Experimental Psychology, 62*(10), 1904-1914.

Schober, M. F. (1993). Spatial perspective-taking in conversation. Cognition, 47(1), 1-24. doi: 10.1016/0010-0277(93)90060-9

Schober, M. F. (1995). Speakers, addressees, and frames of reference: Whose effort is minimized in conversatins about locations? Discourse Processes, 20(2), 219-247. doi: 10.1080/01638539509544939

Shihui, H. (2004). Interactions between proximity and similarity grouping: an event-related brain potential study in humans. *Neuroscience Letters, 367*(1), 40-43. doi: 10.1016/j.neulet.2004.05.098

Sturz, B. R., & Bodily, K. D. (2010). Encoding of variability of landmark-based spatial information. *Psychological Research/Psychologische Forschung, 74*(6), 560-567. doi: 10.1007/s00426-010-0277-4

Sturz, B. R., Cooke, S. P., & Bodily, K. D. (2011). Solving for two unknowns: An extension of vector-based models of landmark-based navigation. *Journal of Experimental Psychology: Animal Behavior Processes, 37*(3), 368-374. doi: 10.1037/a0022938

Sturz, B. R., Kelly, D. M., & Brown, M. F. (2010). Facilitation of learning spatial relations among locations by visual cues: Generality across spatial configurations. *Animal Cognition, 13*(2), 341-349. doi: 10.1007/s10071-009-0283-3

Sutton, J. E. (2002). Multiple-landmark piloting in pigeons (Columba livia): Landmark configuration as a discriminative cue. *Journal of Comparative Psychology, 116*(4), 391-403. doi: 10.1037/0735-7036.116.4.391

Talmy, L. (1983). How language structures space. In J. L. P. A. e. Herbert L. Pick (Ed.), *Spatial orientation: Theory, research, and application* (pp. 225-282). New York: Plenum Press.

Tamara, C., & Timberlake, W. (2011). Route and landmark learning by rats searching for food. *Behavioural Processes, 86*(1), 125-132. doi: 10.1016/j.beproc.2010.10.007

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual Guidance of Eye Movements and Attention in Real-World Scenes: The Role of Global Features in Object Search. Psychological Review, 113(4), 766-786.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. Cognitive Psychology, 12(1), 97-136.

Ullman, S. (1984). Visual routines. *Cognition, 18*(1-3), 97-159. doi: 10.1016/0010-0277(84)90023-4

Vidal, J. R., Gauchou, H. L., Tallon-Baudry, C., & O'Regan, J. K. (2005). Relational information in visual short-term memory: The structural gist. *Journal of Vision, 5*(3). doi: 10.1167/5.3.8

Wolfe, J. M. (1994). Guided search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review, 1*(2), 202-238.

Wolfe, J. M. (1998). What can 1 million trials tell us about visual search? *Psychological Science, 9*(1), 33-39. doi: 10.1111/1467-9280.00006

Wolfe, J. M. (2007). Guided Search 4.0: Current Progress with a model of visual search. *In W. Gray (Ed.), Integrated Models of Cognitive Systems* (pp. 99-119): New York: Oxford.

Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience, 5*(6), 495-501.

Wolfe, J. M., Horowitz, T. S., Kenner, N., Hyle, M., & Vasan, N. (2004). How fast can you change your mind? The speed of top-down guidance in visual search. *Vision Research, 44*(12), 1411-1426.

Wolfe, J. M., & Myers, L. (2010). Fur in the midst of the waters: Visual search for material type is inefficient. Journal of Vision, 10(9). doi: 10.1167/10.9.8

Wolfe, J. M., O'Neill, P., & Bennett, S. C. (1998). Why are there eccentricity effects in visual search? Visual and attentional hypotheses. Perception & Psychophysics, 60(1), 140-156.

Yang, H. & Zelinsky, G. J. (in preperation) Too many features spoil the guidance: Multiple target search is inefficient due to feature mismatch

Zelinsky, G. J. (2008). A theory of eye movements during target acquisition. *Psychological Review, 115*(4), 787-835.

Zelinsky, G. J., & Schmidt, J. (2009). An effect of referential scene constraint on search implies scene segmentation. *Visual Cognition*, *17*(6-7), 1004-1028. doi: 10.1080/13506280902764315

Table 1. *RTs and time to fixate the target for correct trials in Experiment 1*

Number of Previews	One		Two	
Preview Duration (ms)	**400**	**1600**	**400**	**1600**
RT (ms)	1635 (76)	1712 (81)	1553 (60)	1670 (85)
Time to Target (ms)	874 (29)	931 (30)	849 (24)	899 (29)

Note: Values in parentheses indicate standard error of the mean (SEM).

Table 2. *RTs and time to fixate the target for correct trials in Experiment 2*

Number of Previews	One		Two	
Spatial Relationship	**Match**	**Swap**	**Match**	**Swap**
RT (ms)	1194 (51)	1304 (48)	1063 (44)	1196 (45)
Time to Target (ms)	493 (16)	552 (22)	470 (13)	502 (13)

Note: Values in parentheses indicate standard error of the mean (SEM).

Table 3. *RTs and time to fixate the target for correct trials in Experiment 3*

Number of Previews	**One**		**Two**	
Spatial Relationship	**Match**	**Swap**	**Match**	**Swap**
RT (ms)	1823 (83)	2038 (97)	1792 (75)	1979 (84)
Time to Target (ms)	629 (17)	635 (16)	621 (14)	634 (16)

Note: Values in parentheses indicate standard error of the mean (SEM).

Table 4. *RTs and time to fixate the target for correct trials in Experiment 4*

Semantic Association	Related				Unrelated			
Functional Interaction	**Interacting**		**Noninteracting**		**Interacting**		**Noninteracting**	
Spatial Relationship	**Match**	**Swap**	**Match**	**Swap**	**Match**	**Swap**	**Match**	**Swap**
RT (ms)	1025	1134	1066	1138	1071	1113	1094	1150
SEM	(67)	(99)	(79)	(75)	(116)	(66)	(108)	(75)
Time to Target (ms)	414	414	407	416	403	434	398	438
SEM	(13)	(9)	(10)	(13)	(9)	(11)	(11)	(13)

Table 5. *RTs and time to fixate the target for correct trials in Experiment 5*

Block	**One**				**Two**			
Relational Information	**Irrelevant**		**Relevant**		**Irrelevant**		**Relevant**	
Spatial Relationship	**Match**	**Swap**	**Match**	**Swap**	**Match**	**Swap**	**Match**	**Swap**
RT (ms)	977	1090	1222	1383	926	1046	1100	1224
SEM	(78)	(96)	(51)	(91)	(84)	(108)	(51)	(49)
Time to Target (ms)	471	497	466	482	443	456	435	492
SEM	(23)	(31)	(17)	(22)	(23)	(21)	(16)	(14)

Table 6. *RTs and time to fixate the target for correct trials in Experiment 6*

Grouping Condition	No Frame				Frame			
Number of Previews	One		Two		One		Two	
Spatial Relationship	Match	Swap	Match	Swap	Match	Swap	Match	Swap
RT (ms)	1233	1354	1179	1341	1290	1482	1205	1380
SEM	(46)	(56)	(58)	(71)	(55)	(83)	(58)	(73)
Time to Target (ms)	496	543	473	530	549	608	513	597
SEM	(14)	(14)	(15)	(17)	(30)	(39)	(23)	(44)

Table 7. *RTs and time to fixate the target for correct trials in Experiment 7*

Grouping Distance	Close		Far	
Spatial Relationship	**Match**	**Swap**	**Match**	**Swap**
RT (ms)	1101 (50)	1241 (57)	1120 (51)	1233 (53)
Time to Target (ms)	445 (11)	491 (18)	442 (12)	489 (15)

Note: Values in parentheses indicate standard error of the mean (SEM).

Figure Captions

Figure 1. An example preview and search view used in Experiment 1. Previews were presented for either 400 ms or 1600 ms and contained either one or both target objects.

Figure 2. The proportion of trial in which the target pair was the first pair fixated during search in Experiment 1 (correct trials only), as a function of number of previewed targets and preview duration. Chance is indicated by the dashed line and is equal to .167. Error bars indicate one standard error of the mean (SEM).

Figure 3. An example preview and search view used in Experiment 2. Previews contained either one or both target objects and matched the search view perfectly (spatial match conditions) or the objects swapped relative postions (spatial mismatch conditions).

Figure 4. The proportion of trials in which the target pair was the first pair fixated during search in Experiment 2 (correct trials only), as a function of previewed relation (match or swap) and the number of previewed targets. Error bars indicate one standard error of the mean (SEM).

Figure 5. The proportion of trials in which the target pair was the first pair fixated during search in Experiment 3 (correct trials only), as a function of cued relation (match or swap) and the number of cued targets. Error bars indicate one standard error of the mean (SEM).

Figure 6. A complete list of all related pairs and functional directions used in Experiment 4.

Figure 7. The proportion of trials in which the target pair was the first pair fixated during search in Experiment 4 (correct trials only), as a function of previewed relation (match or swap), semantic relatedness and functional interaction of the targets. Error bars indicate one standard error of the mean (SEM).

Figure 8. The proportion of trials in which the target pair was the first pair fixated during search in Experiment 5 (correct trials only), as a function of previewed relation (match or swap), task and block number. Error bars indicate one standard error of the mean (SEM).

Figure 9. An example preview and search view used in Experiment 6 (frame condition). Previews contained either one or both target objects and matched the search view perfectly or the objects swapped relative postions.

Figure 10. The proportion of trials in which the target pair was the first pair fixated during search in Experiment 6 (correct trials only), as a function of previewed relation (match or swap), the number of previewed targets and frame presence. Error bars indicate one standard error of the mean (SEM).

Figure 11. An example preview and search view used in Experiment 7 depicting both near and far distance conditions. Previews either matched the search view perfectly or the objects swapped relative postions.

Figure 12. The proportion of trials in which the target pair was the first pair fixated during search in Experiment 7 (correct trials only), as a function of previewed relation (match or swap), and the distance between the objects. Error bars indicate one standard error of the mean (SEM).

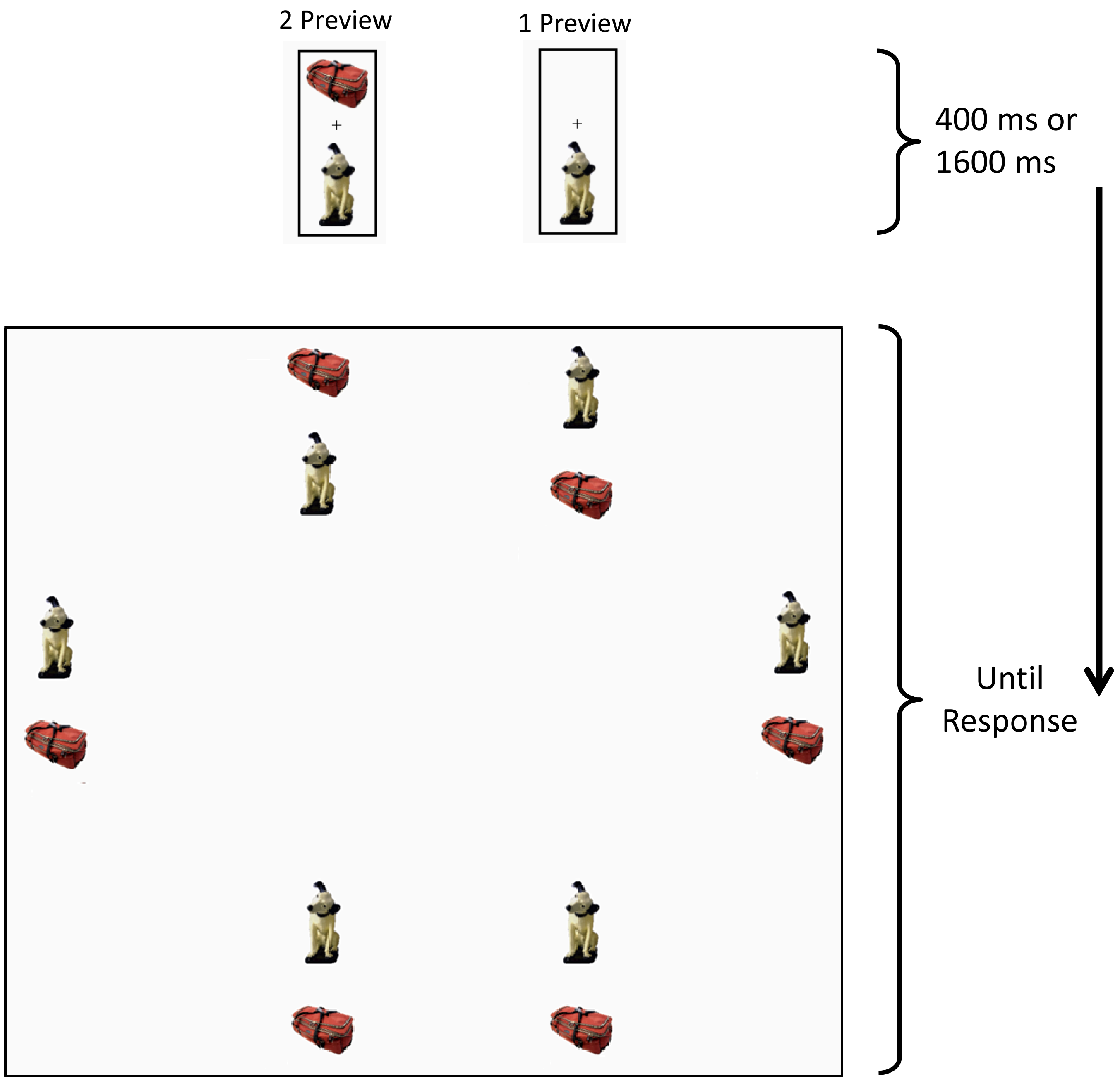
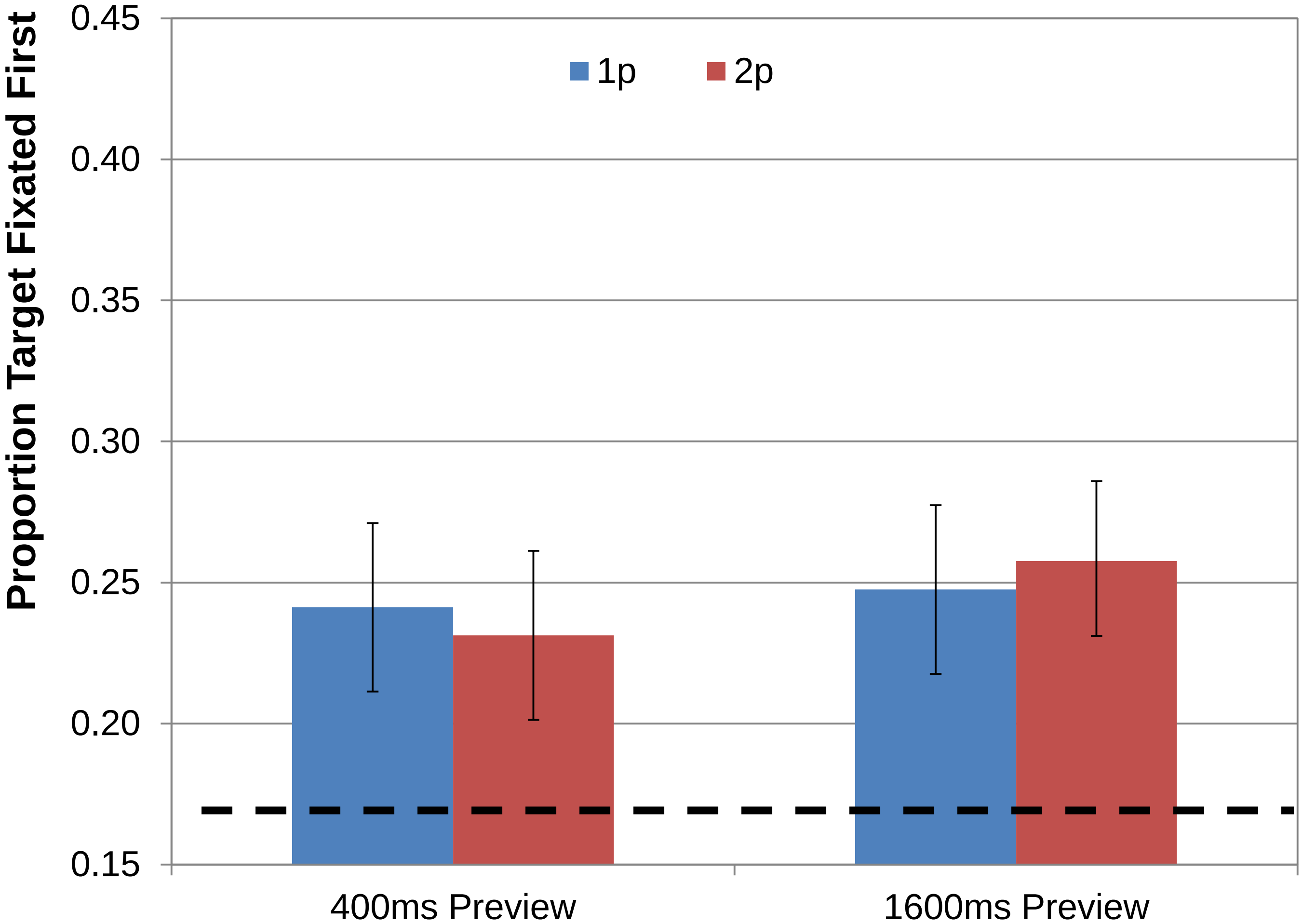


Figure 1



Proportion Target Fixated First
0.45
0.40
0.35
0.30
0.25
0.20
0.15
1p
2p
400ms Preview
1600ms Preview

Figure 2

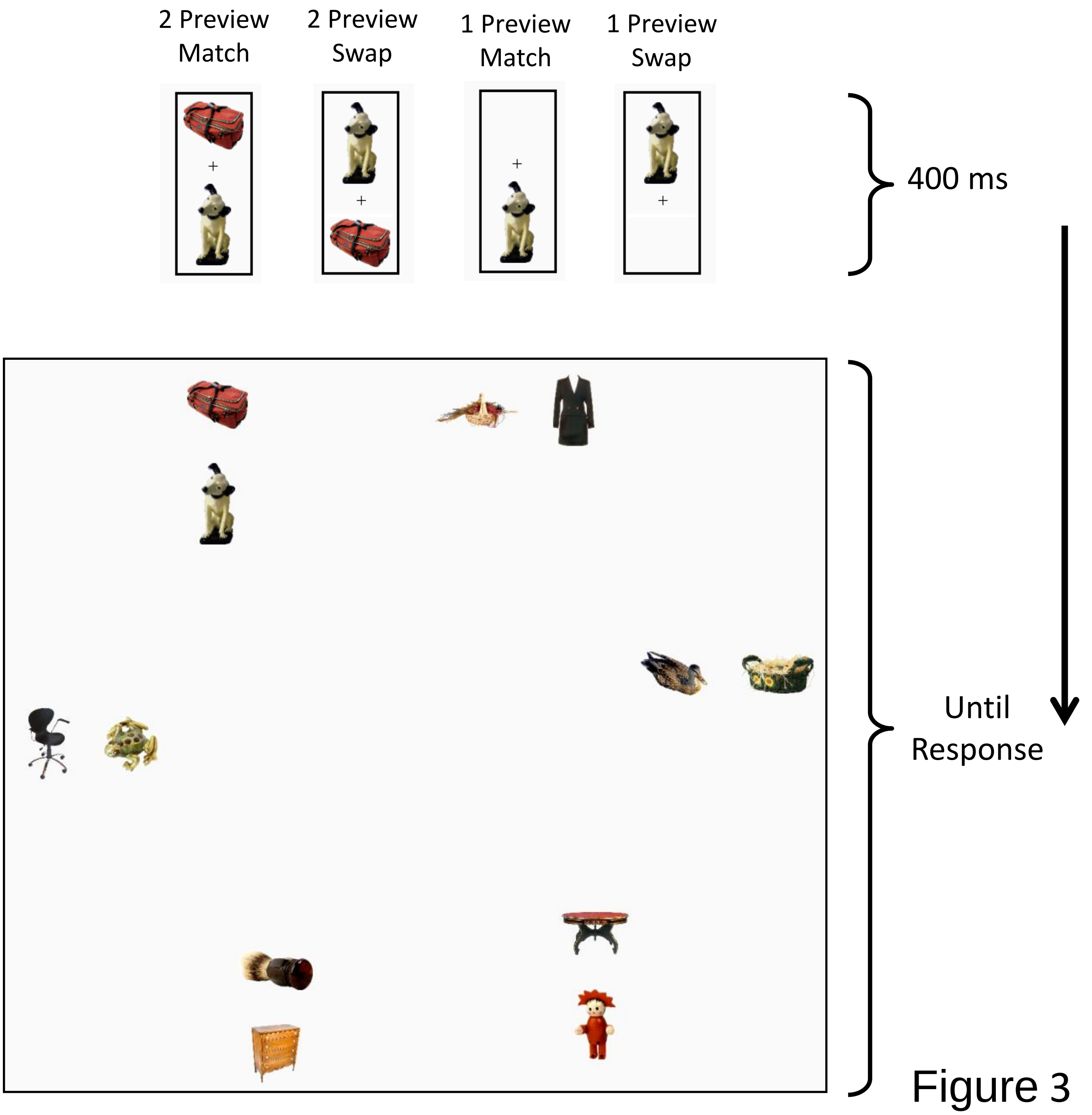

2 Preview Match
2 Preview Swap
1 Preview Match
1 Preview Swap
400 ms
Until Response

Figure 3

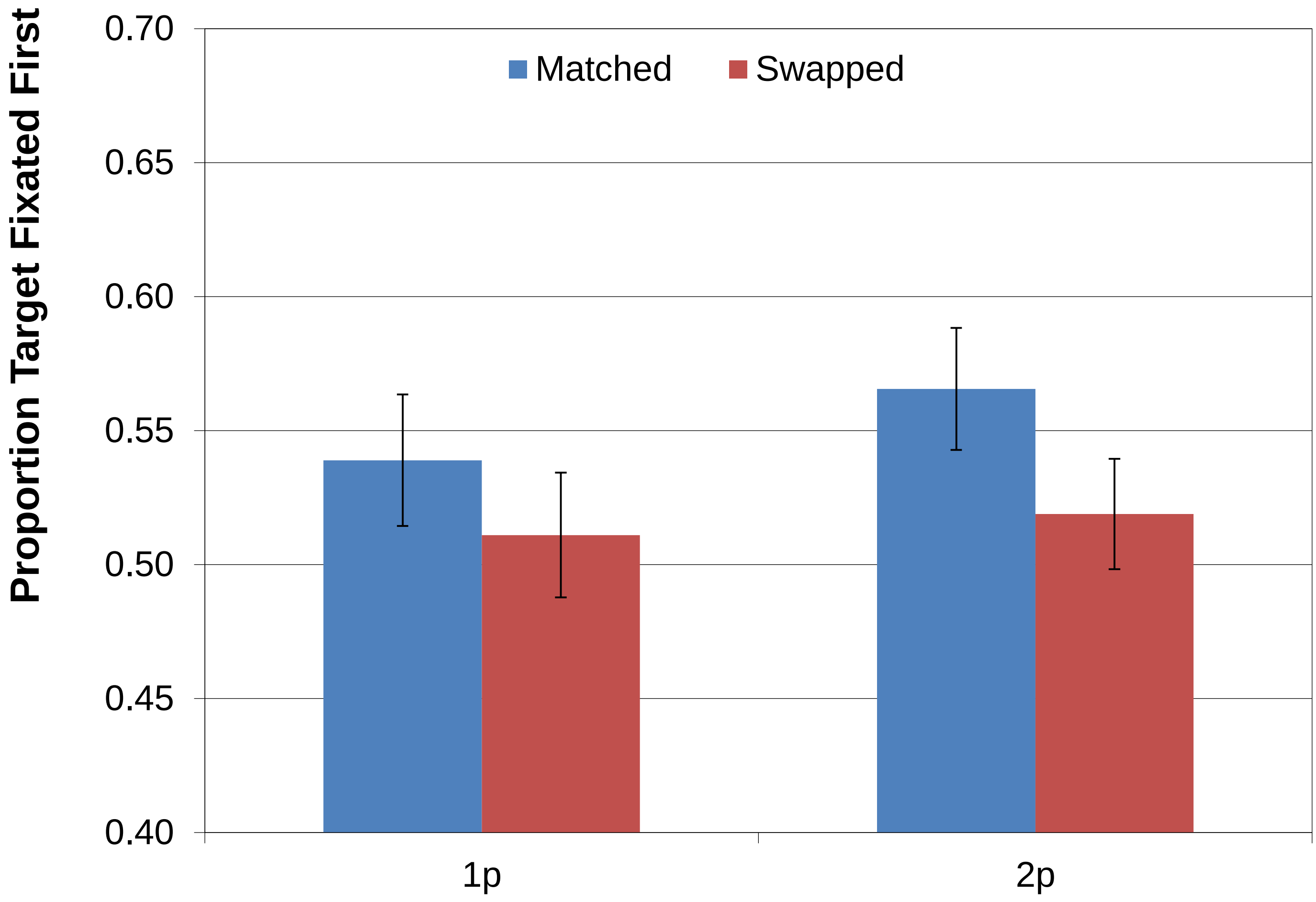

Proportion Target Fixated First
0.70
0.65
0.60
0.55
0.50
0.45
0.40
Matched
Swapped
1p
2p

Figure 4

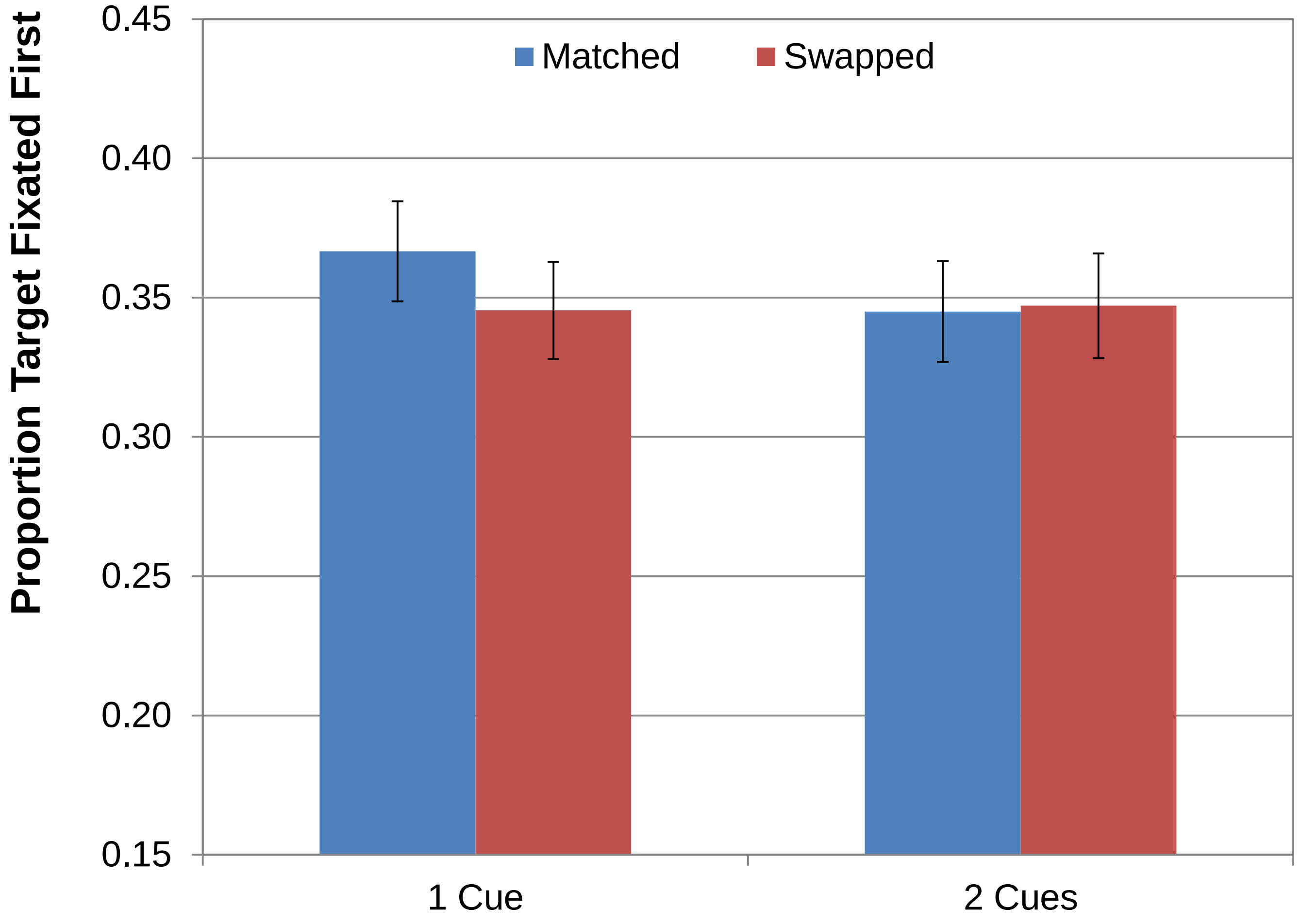

Proportion Target Fixated First
0.45
0.40
0.35
0.30
0.25
0.20
0.15
Matched
Swapped
1 Cue
2 Cues

Figure 5

Experimental

First Object	Direction	Second Object
hammer	above	nail
sparrow	above	nest
Christmas tree	above	present
garbage bag	above	pail
pitcher	above	glass
spoon	above	soup
soap	above	soap dish
mixer	above	bowl
coin	above	piggy bank
kettle	above	tea cup
cake slice	above	plate
cork screw	above	wine bottle
bottle	above	recycling bin
laptop	above	desk
axe	above	wood
light bulb	above	lamp
alarm clock	above	night table
pickaxe	above	rock
match	above	pipe
lighter	above	candle
screwdriver	left	screw
key	left	lock
mouse	left	cheese
dart	left	dartboard
flower	left	humming bird
letter	left	mailbox
paintbrush	left	easel
baseball	left	mitt
bolt	left	wrench
carrot	left	rabbit

Practice

First Object	Direction	Second Object
crown	above	pillow
hot dog	above	bun
pot	above	stove
camera	above	tripod
watering pot	above	plant
ice cream	above	cone
orange	above	juicer
toothpaste	above	toothbrush
bullet	left	gun
treat	left	dog
trailer	left	truck
horse	left	carriage

Figure 6

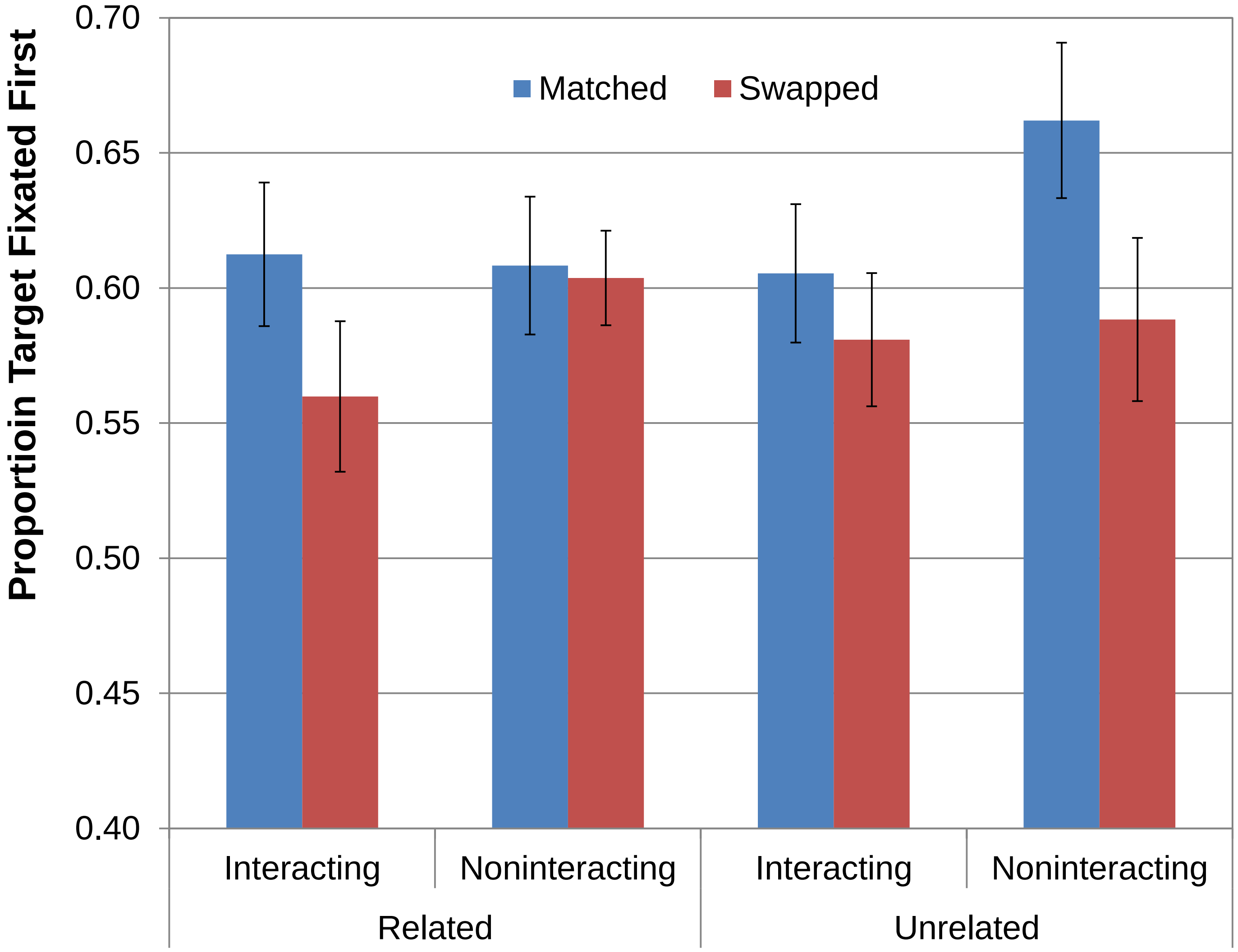

Proportioin Target Fixated First
0.70
0.65
0.60
0.55
0.50
0.45
0.40
Matched
Swapped
Interacting
Noninteracting
Interacting
Noninteracting
Related
Unrelated

Figure 7

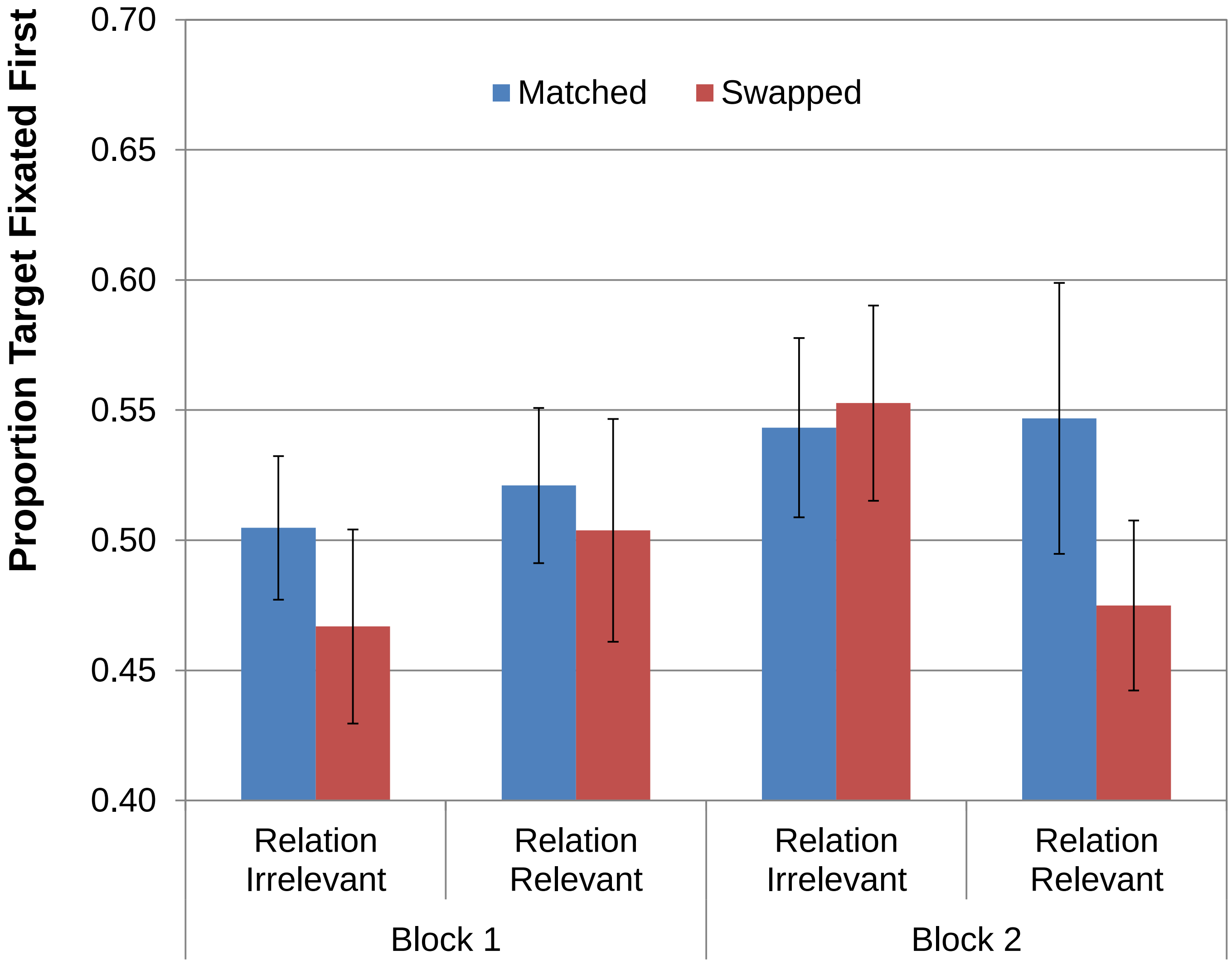

Proportion Target Fixated First
0.70
0.65
0.60
0.55
0.50
0.45
0.40
Matched
Swapped
Relation Irrelevant
Relation Relevant
Relation Irrelevant
Relation Relevant
Block 1
Block 2

Figure 8

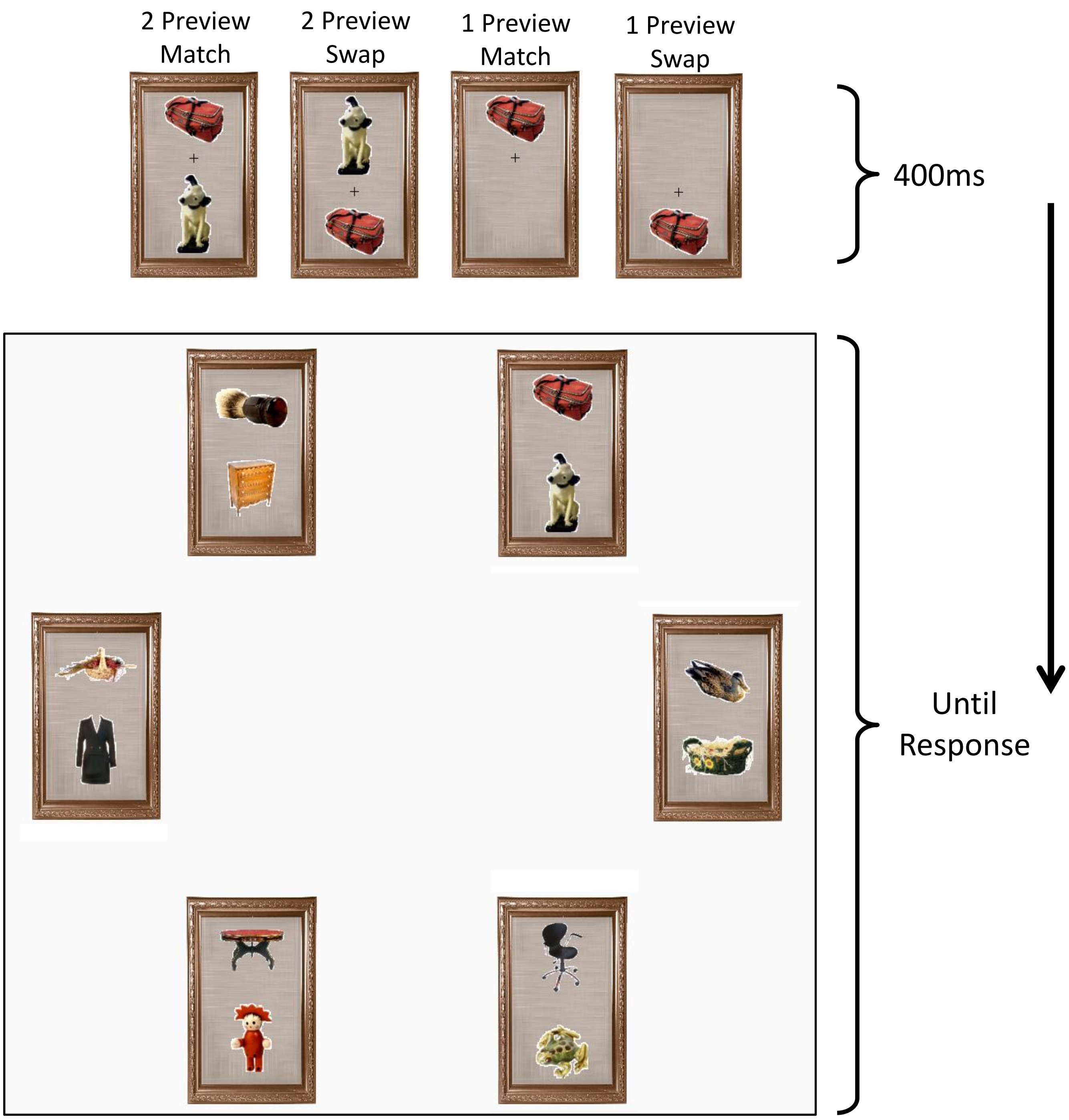
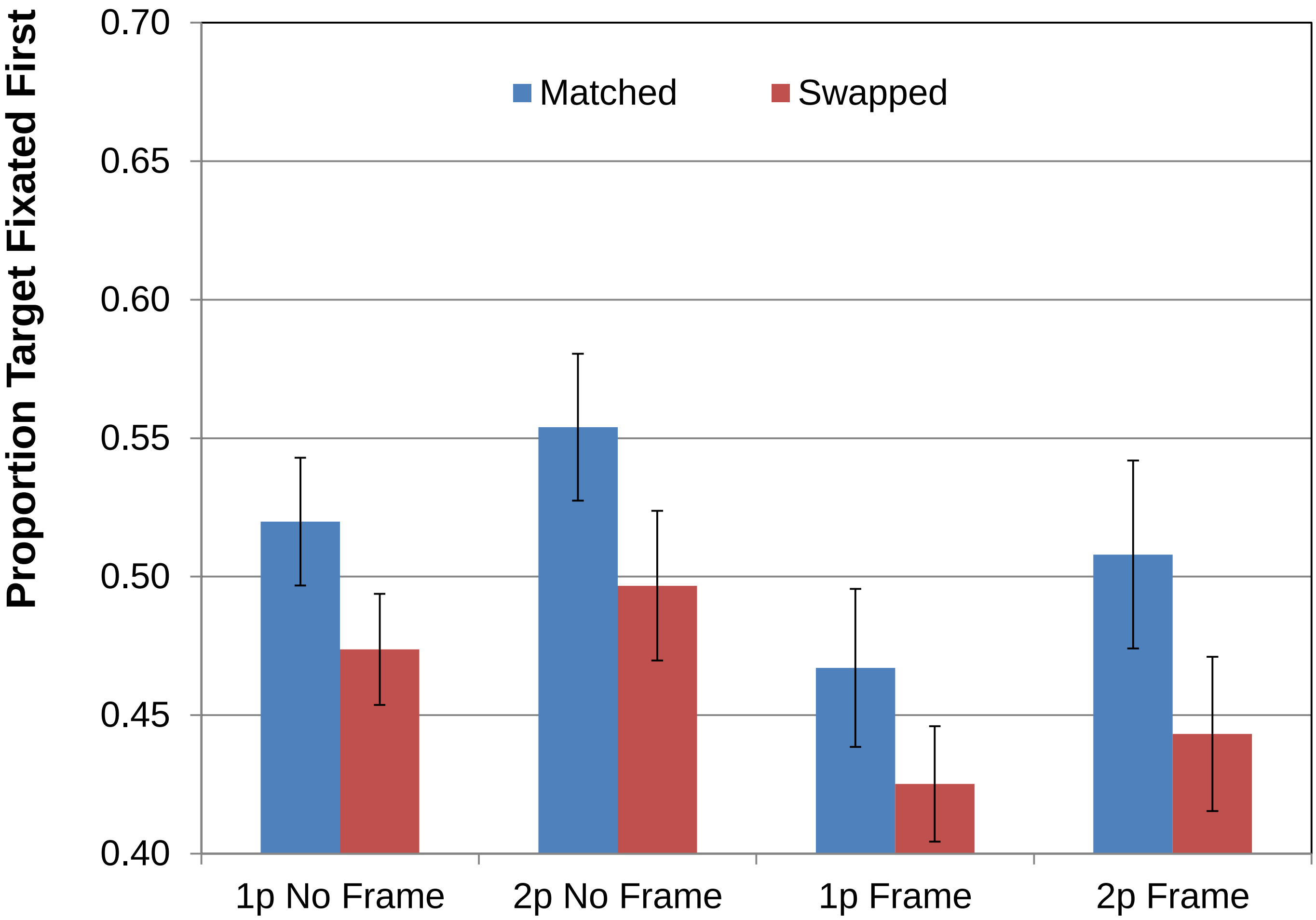

2 Preview
Match
2 Preview
Swap
1 Preview
Match
1 Preview
Swap
400ms
Until
Response

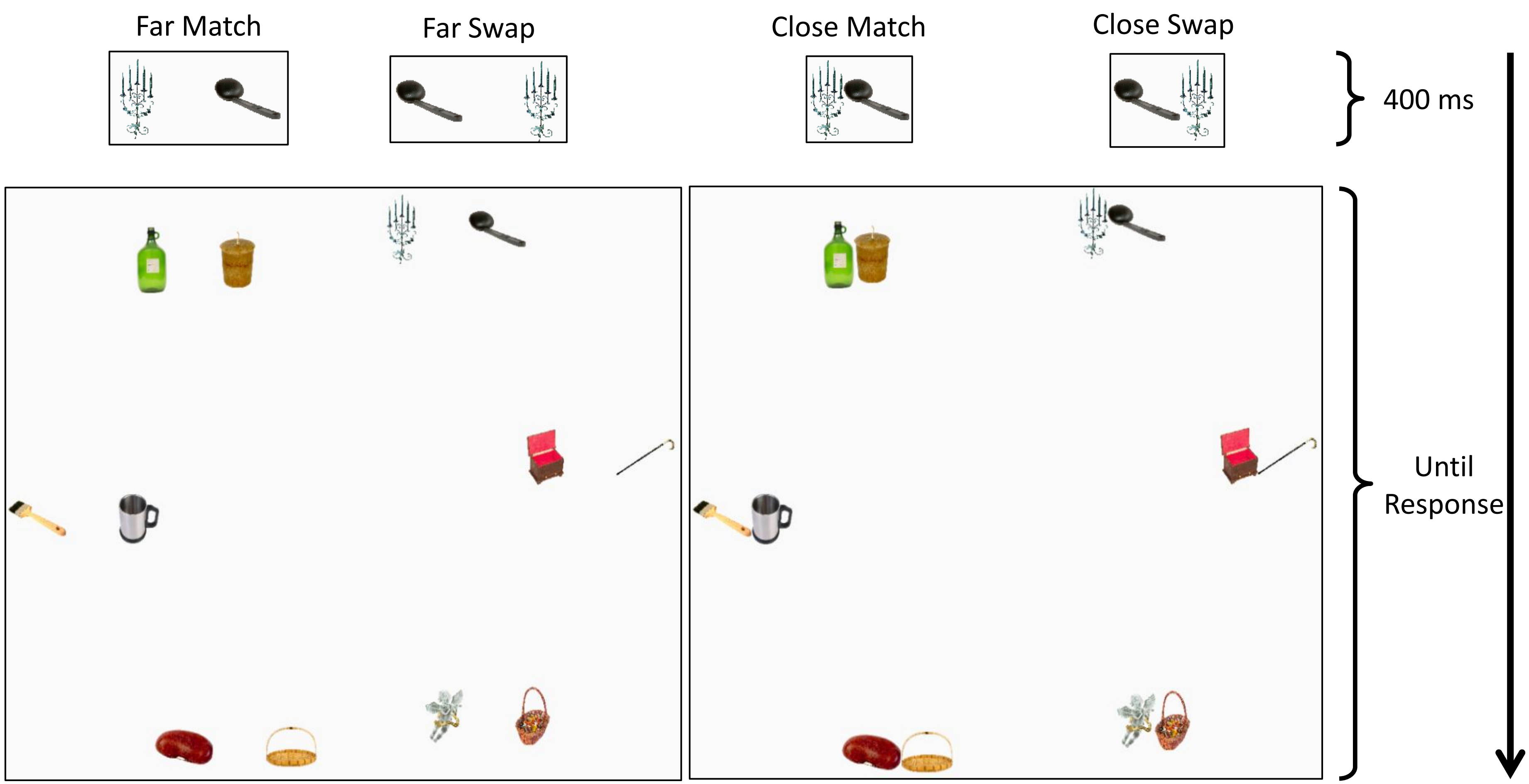
Figure 9



Proportion Target Fixated First
0.70
0.65
0.60
0.55
0.50
0.45
0.40
Matched
Swapped
1p No Frame
2p No Frame
1p Frame
2p Frame

Figure 10

Figure 11


Far Match
Far Swap
Close Match
Close Swap
400 ms
Until
Response

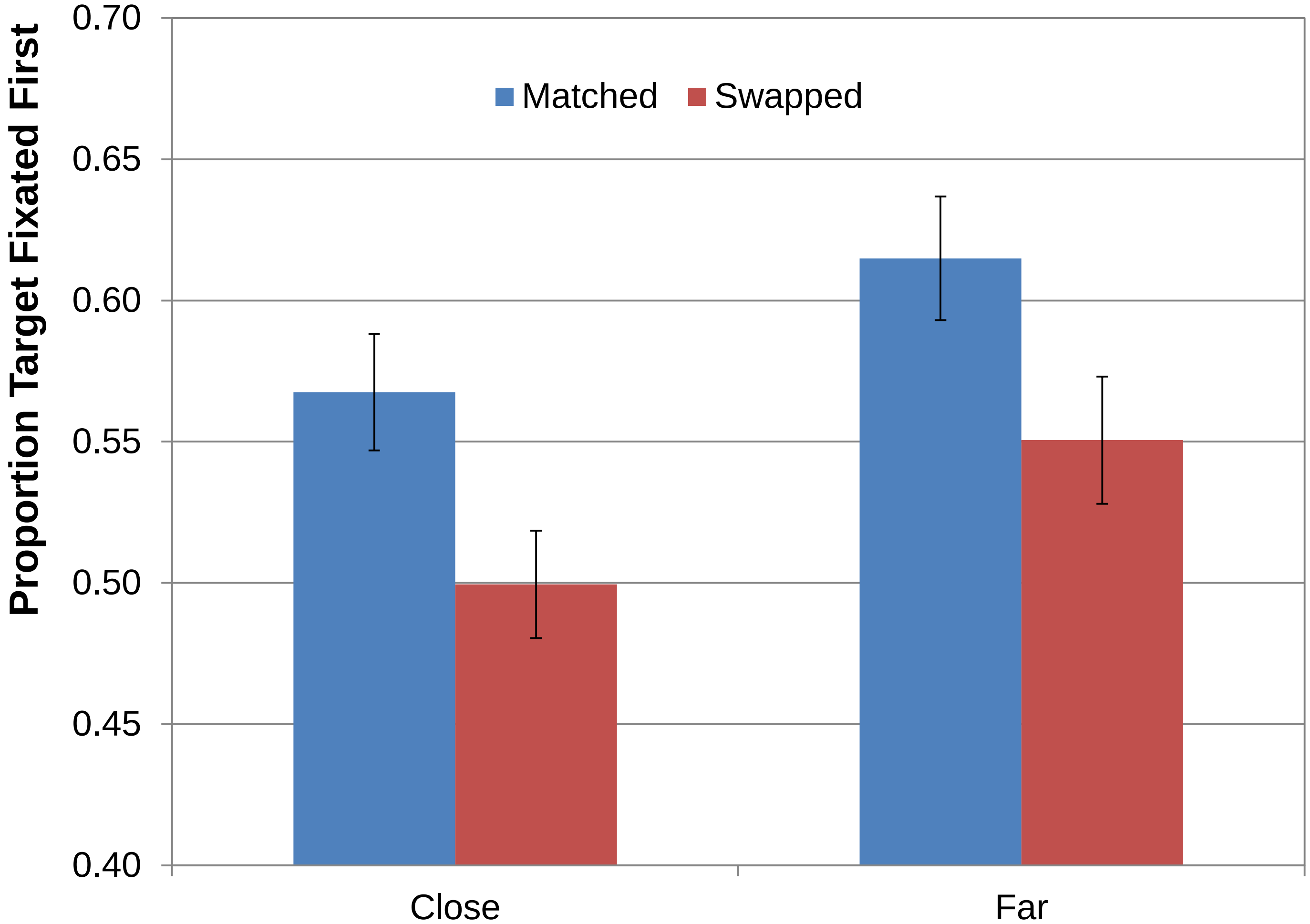

Proportion Target Fixated First
0.70
0.65
0.60
0.55
0.50
0.45
0.40
Matched
Swapped
Close
Far

Figure 12