



Published in final edited form as:

*Vis cogn.* 2012 December 1; 20(10): 1153–1163. doi:10.1080/13506285.2012.735718.

## Searching Through the Hierarchy: How Level of Target Categorization Affects Visual Search

Justin T. Maxfield<sup>1</sup> and Gregory J. Zelinsky<sup>1,2</sup>

<sup>1</sup>Department of Psychology, Stony Brook University

<sup>2</sup>Department of Computer Science, Stony Brook University

### Abstract

Does the same basic-level advantage commonly observed in the categorization literature also hold for targets in a search task? We answered this question by first conducting a category verification task to define a set of categories showing a standard basic-level advantage, which we then used as stimuli in a search experiment. Participants were cued with a picture preview of the target or its category name at either superordinate, basic, or subordinate levels, then shown a target-present/absent search display. Although search guidance and target verification was best using pictorial cues, the effectiveness of the categorical cues depended on the hierarchical level. Search guidance was best for the specific subordinate level cues, while target verification showed a standard basic-level advantage. These findings demonstrate different hierarchical advantages for guidance and verification in categorical search. We interpret these results as evidence for a common target representation underlying categorical search guidance and verification.

### Keywords

Categorical search; Search guidance; Basic-level advantage; Category verification; Eye movements

### Introduction

How does the categorization of an object impact its use as a target in a visual search task? Decades of research has shown that objects can be categorized at multiple levels in a hierarchy (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976; Murphy, 2002; Mack & Palmeri, 2011); the same object can be recognized as a beagle (subordinate), a dog (basic), or an animal (superordinate). Exploration of this categorical hierarchy led to the seminal finding of a *basic-level superiority effect*, one expression of which is that objects are categorized fastest at the basic-level (Rosch et al., 1976; Murphy & Brownell, 1985). Categorization theories typically explain the basic-level advantage in terms of a favorable balance existing between object specificity and object distinctiveness at the intermediate basic level (Murphy & Brownell, 1985). Whereas the features of objects at the subordinate level may be highly specific, these features tend to overlap with those from other object categories and therefore lack distinctiveness. Similarly, the features of superordinate objects are highly distinct, but the variability at this level means that the features of category members will generally lack specificity. Shape information at the basic level also tends to be representative of the category membership (Rosch et al., 1976), potentially adding to its advantage in visual tasks.

---

Address correspondence to: Gregory J. Zelinsky, Psychology B-240, Stony Brook University, Stony Brook, NY 11794-2500, Phone: 631-689-7087, Gregory.Zelinsky@stonybrook.edu.

One task that relies heavily on categorical representations is visual search. Embedded within a categorical search task is a category verification task. Each time gaze lands on a potential target, the fixated object must be verified as belonging to the target category, with the time needed for this to occur reflecting the match between that object and the target representation. These verification times have been shown to vary with the specificity of the target cue (Schmidt & Zelinsky, 2009; Malcolm & Henderson, 2009), as well as target typicality (Castelhano, Pottatsk, & Cave, 2008). We might therefore also expect categorical search targets defined at the basic level to be verified faster than those defined at higher or lower levels in the hierarchy.

Verification is only one component of search, there is also guidance—the attraction of attention by target features. Early work suggested that search is not guided to categorically-defined targets, targets specified by a word rather than a picture cue. Smith, Redford, Gent, & Washburn (2005) found at chance search performance in a combined search-categorization task using dot stimuli, and other studies using naturalistic scenes and objects also found little to no guidance when searching for categorically-defined targets (Foulsham & Underwood, 2007; Wolfe, Horowitz, Kenner, Hyle, & Vasan, 2004). More recent work using a categorical search task has reached an opposite conclusion, arguing for the existence of guidance to categorical targets. Yang & Zelinsky (2009) compared search guidance for pictorially-previewed and categorically-cued teddy bear targets and found significantly above chance levels of guidance using categorical cues, although this was less than what was found for pictorial previews. Other studies using random photorealistic objects (Schmidt & Zelinsky, 2009) and naturalistic scenes (Malcolm & Henderson, 2009) have shown that increasing the specificity of a categorical cue also improved search guidance. However, while these studies demonstrate that attention can be attracted to categorically-defined targets, it remains unclear how guidance might be affected by the hierarchical level of the categorical cue.

In the present study we combine a standard category hierarchy manipulation with a search task to determine how search guidance and verification is affected by targets cued at different hierarchical levels. Specifically, will a basic-level superiority effect be found in a search task, and if so, will it be expressed in search guidance, target verification, or both? Answers to these questions will inform the relationship between search guidance and verification. Can guidance and verification processes use different representations of the target to best accomplish their respective tasks, or does the adoption of a target representation for guidance require that the same representation be used for verification (or vice versa)?

## Experiment 1

We conducted a standard category verification task in order to find a set of images that produce a basic-level categorization advantage. Finding stimuli that demonstrate this classic effect is a necessary first step towards our goal of studying the basic-level advantage in the context of a search task.

### Methods

**Participants**—Fifteen Stony Brook University undergraduates participated. All reported normal or corrected-to-normal visual acuity and that English was their native language.

**Stimuli and Apparatus**—Images of photorealistic objects were obtained from the Hemera collection and various web sources. Objects were selected to be typical members of their category at superordinate, basic, and subordinate levels. To quantify this, twelve different participants viewed a set of images and rated each twice, once for typicality and

again for image agreement (as in Snodgrass & Vanderwart, 1980), on a scale of 1 (high typicality/image agreement) to 7 (low typicality/image agreement) at each hierarchical level. From this set we selected 48 target objects having a mean typicality score of 1.77 and a mean image agreement score of 1.98. There were no significant differences between hierarchical levels for either type of rating ( $F = 2.18, p = .156$ ).

Stimuli were presented centrally on a flat-screen CRT monitor at a resolution of  $1024 \times 768$  pixels. Head position and viewing distance were fixed at 60 cm using a chinrest. Objects subtended  $\sim 2.5^\circ$  visual angle, the same as in the norming task, and category names were drawn in 18-point Tahoma font. Judgments were made by pressing the left and right index finger triggers of a game pad controller; trials were initiated with a button operated by the right thumb.

**Procedure**—Participants were shown a category name for 2,500 ms, followed by a fixation cross for 50 ms and then an object. They were instructed to respond whether the object belonged to the named category as quickly and accurately as possible. Category names were at subordinate, basic, and superordinate levels (Table 1). Over the course of the experiment, each picture was presented twice at each hierarchical level. To minimize contingency-based learning, one-third of the duplicate presentations of an object were both true, another third were both false, and the final third appeared once as true and once as false.

There were 160 trials per participant, half true and half false. Half of the false trials consisted of cueing the participant with a target category name then showing a random object from a different non-target superordinate category. The other half of the false trials contained lures in which the object was drawn from target images one hierarchical level above the cued target (e.g., taxi when cued with police car, or car when cued with truck). At the superordinate level, these lures were objects drawn from other superordinate categories, making them indistinguishable from the objects used on the other false trials. The purpose of the lures in the subordinate and basic conditions was to ensure encoding at the cued level (see Tanaka & Taylor, 1991).

## Results and Discussion

Error rates were less than 5% and showed no significant differences between any of the conditions ( $F = 1.85, p = .179$ ). Only correct trials were included in the subsequent analyses. Analysis of target-present categorization times revealed a significant basic-level advantage ( $F(1,12) = 12.47, MSE = 2,199, p < .001$ ). Objects cued at the basic level ( $M = 613$  ms) were verified significantly faster than both superordinate ( $M = 703$  ms,  $p < .01$ ) and subordinate levels ( $M = 643$  ms,  $p < .05$ ). Objects were also verified significantly faster following subordinate cues than superordinate cues ( $p < .001$ ). These findings replicate the standard basic-level advantage using our stimulus set.

## Experiment 2

Having obtained a basic-level superiority effect using our stimuli, in Experiment 2 we constructed search displays from these stimuli to investigate the expression of a basic-level advantage in the context of a search task.

## Methods

**Participants**—Twelve different Stony Brook undergraduates participated, all reporting normal or corrected-to-normal visual acuity and that English was their native language.

**Apparatus**—Eye position was sampled at 1000 Hz using an Eye Link 1000 eye-tracker with default saccade detection settings. Calibrations were accepted when the average spatial error was less than  $0.49^\circ$ , and the maximum error was less than  $0.99^\circ$ . Head position and viewing distance were fixed at 65 cm using a chinrest. Equipment was otherwise the same as described for Experiment 1.

**Procedure**—The procedure for the search task paralleled the procedure for the category verification task. The identical subordinate, basic, and superordinate category names were shown for the same 2,500 ms duration, followed by the same 50 ms interval with a fixation cross. We added to these categorical cue conditions a picture preview condition in which participants were cued with the exact image of the search target. Also different from Experiment 1 was the presentation of a search array of six objects, each at  $7.8^\circ$  eccentricity (Figure 1), instead of just one centrally-positioned object. There were 256 trials, half target-present and half target-absent. In target-present trials a target was displayed with five distractors from random categories. Target-absent trials were again evenly divided into lure and non-lure search displays. Target-absent lure displays included five random objects and one object from the same basic category (on subordinate trials) or superordinate category (on basic trials) as the target. Target-absent non-lure displays consisted of six random objects from different non-target superordinate categories.

## Results and Discussion

Error rates are provided in Table 2. Significant differences between conditions were found in both the target-present and target-absent trials ( $F = 10.13, p = .001$ ). Post-hoc tests (LSD corrected) on the target-present data revealed that error rates were significantly lower for pictorial cues ( $p = .05$ ) and significantly higher for superordinate cues ( $p = .01$ ), compared to all other conditions. These errors can likely be explained by participants occasionally failing to recognize a target as a member of the cued category (see Murphy & Brownell, 1985). Error rates were more stable across the target-absent conditions, with only subordinate cues yielding higher errors ( $p = .01$ ). Only correct trials were included in the subsequent analyses.

Consistent with previous reports (Castelhano et al., 2008; Schmidt & Zelinsky, 2009), we divided our analysis of search performance into guidance and target verification epochs. Search guidance was defined as the time from the onset of the search display until the participant fixated the target (time-to-target). Supplementing this temporal measure of guidance was an analysis of the proportion of trials in which the target was the first fixated object during search. Verification was defined as the time between when a participant first looked at the target and when the target-present/absent button response was registered.

Turning first to search guidance, analyses of the target-present data revealed significant differences in time-to-target between conditions ( $F(3,33) = 31.42, MSE = 2,986, p < .001$ ). As shown in Figure 2a, pictorially-cued targets were fixated sooner than subordinate, which were fixated sooner than basic, which were fixated sooner than targets cued at the superordinate level ( $t = 2.53, p = .05$ ). Figure 2b shows a similar trend in the first fixated objects ( $F(3,33) = 12.41, MSE = 81, p < .001$ ). Post-hoc tests revealed that targets in the pictorial and subordinate conditions were fixated first most often (~40%), and did not differ significantly from each other ( $p = .981$ ). Targets cued at the basic level were fixated first significantly less often ( $p = .05$ ), and superordinate cues resulted in the lowest rate of first target fixations ( $p = .01$ ). In fact, while targets in the subordinate and basic conditions were fixated first more often than chance ( $t = 3.19, p = .01$ ), targets cued at the superordinate level were not ( $t(11) = 1.47, p = .169$ ). Initial saccade latency did not reliably differ between cueing conditions ( $F(3,33) = .678, MSE = 337, p = .572$ ), suggesting that the differences in first fixated objects were not due to a speed-accuracy tradeoff. Both the time-to-target and

first fixation guidance measures therefore generally support previous reports showing an increase in search efficiency with greater cue specificity (Schmidt & Zelinsky, 2009; Malcolm & Henderson, 2009). Critically, we found no evidence for a basic-level superiority effect in search guidance to targets.

Turning to target verification, there were again significant differences between conditions ( $F(3,33) = 15.08$ ,  $MSE = 10,139$ ,  $p < .001$ ), but the pattern was very different from the one observed for search guidance (Figure 3). Targets cued at the basic level were verified significantly faster than those cued at either the subordinate or superordinate levels ( $t = 2.28$ ,  $p < .05$ ). Subordinate targets were also verified marginally faster than superordinate targets ( $t(11) = 2.1$ ,  $p = .06$ ), and pictorially previewed targets were verified significantly faster than any of the categorically cued targets ( $p < .001$ ). In contrast to the guidance data, the verification data are therefore consistent with the categorization literature; a basic-level advantage was found in the time taken to verify an object as a member of the target category.

## General Discussion

Our results demonstrate two distinct effects of categorization hierarchy on search, one appearing early during search guidance and the other later at target verification. Consistent with previous studies of visual search (Schmidt & Zelinsky, 2009; Malcolm & Henderson, 2009), as cue specificity increased, so too did search guidance. Consistent with the categorization literature (Rosch et al., 1976; Johnson & Mervis, 1997; Murphy & Brownell, 1985), a typical basic-level advantage was found in target verification times. Because a manual reaction time measure would have collapsed across these patterns, only through eye movement analysis and the division of search into separate guidance and verification epochs were we able to discern these distinct effects of categorical hierarchy on search.

We interpret these patterns of guidance and verification in terms of the principles of specificity and distinctiveness proposed in the categorization literature (Murphy & Brownell, 1985). Highly specific subordinate representations are best for guidance but not for verification, as their lack of distinctiveness requires relatively extensive feature checking in order to verify category membership—a police car and a taxi share many features, which increases verification difficulty. Basic representations are best for verification but are not optimal for guidance, as basic representations typically lack the specificity of subordinates—a police car is more specific than a car and might produce a guiding representation that better matches the search target. Arguably, pictorial cues maximize both distinctiveness and specificity relative to categorical cues. Just as experts have been shown to categorize subordinate-level objects as fast as basic (Tanaka & Taylor, 1991), a person given a pictorial cue becomes a sort of expert on the cued object, leading to a pictorial preview advantage in both guidance and verification.

The fact that guidance and verification processes are best served by different levels of hierarchical representation raises a theoretically important question—does search guidance and target verification use different target representations or a single common target representation? We found that subordinate-level cues produced strong search guidance, consistent with the use of a highly detailed representation, but slower verification relative to basic-level cues. This pattern suggests that participants could not flexibly adopt different target representations depending on the stage of the search task—the choice of representational level for one process (guidance) may lock in that representational level for the other (verification). This evidence for a common target representation also raises the intriguing possibility that the search guidance and recognition processes might be more similar than what has been believed, and perhaps one in the same.

Also interesting is the fact that pictorial and subordinate cues resulted in equally strong guidance, as measured by initial looks to the target. The percentage of first fixated objects is an accepted measure of early search guidance (Chen & Zelinsky, 2006; Schmidt & Zelinsky, 2011), and finding in this measure no evidence for a guidance difference between these two types of cues was unexpected. The search literature suggests that target guidance should always be better with a pictorial cue than a categorical cue (Yang & Zelinsky, 2009; Schmidt & Zelinsky, 2009). Our failure to replicate this finding means one of two things: Either the highly detailed target representation from a subordinate cue is surprisingly good, or the target representation from a picture preview is surprisingly bad, and in this case no better than a good categorical cue.

One direction of future work will explore more fully the conditions under which pictorial and categorical cues produce equivalent guidance, and the formation of categorical target representations in response to pictorial cues. Another direction of future work will train classifiers on superordinate, basic, and subordinate-level objects in order to quantify categorical distinctiveness throughout the hierarchy and to better understand basic-level superiority effects in categorical search.

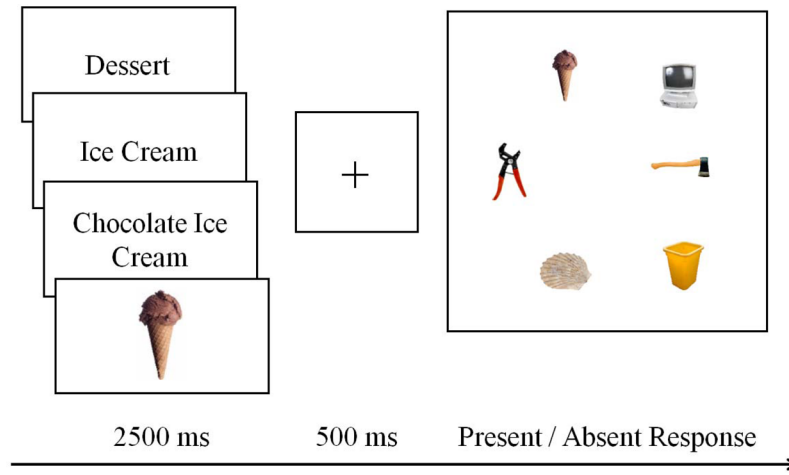
## Acknowledgments

We thank Christian Luhmann and Gregory Murphy for their helpful comments in the preparation of this manuscript, and all the members of the Eye Cog Lab for invaluable feedback. This work was supported by NIH Grant R01-MH063748 to GJZ.

## References

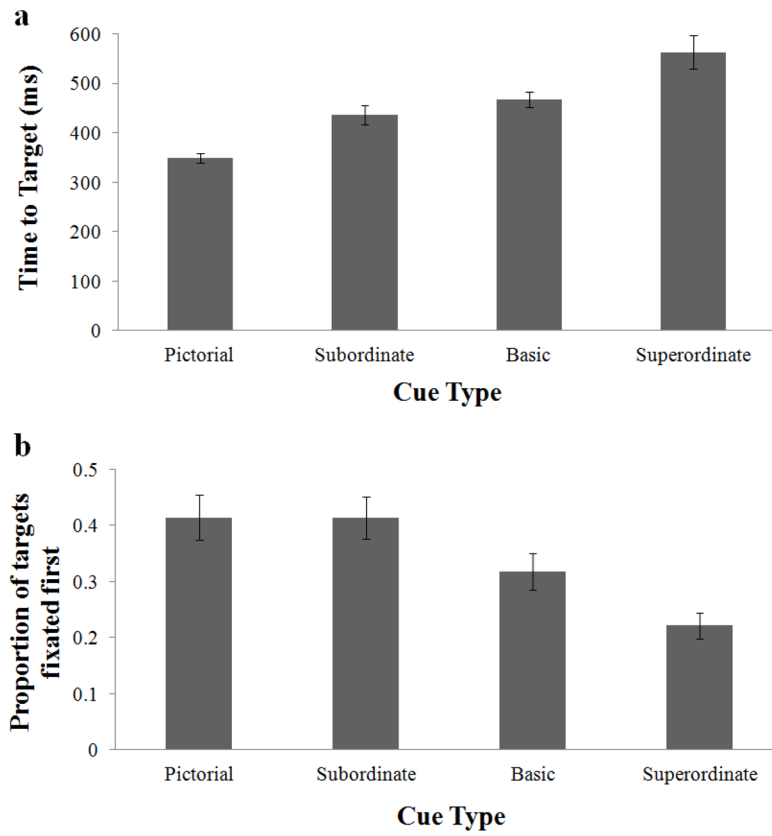
- Castelhano MS, Pollatsek A, Cave K. Typicality aids search for an unspecified target, but only in identification, and not in attentional guidance. *Psychonomic Bulletin & Review*. 2008; 15(4):795–801.10.3758/PBR.15.4.795 [PubMed: 18792506]
- Chen X, Zelinsky GJ. Real-world visual search is dominated by top-down guidance. *Vision Research*. 2006; 46:4118–4133.10.1016/j.visres.2006.08.008 [PubMed: 17005231]
- Foulsham T, Underwood G. How does the purpose of inspection influence the potency of visual salience in scene perception? *Perception*. 2007; 36:1123–1138.10.1068/p5659 [PubMed: 17972478]
- Johnson KE, Mervis CB. Effects of varying levels of expertise on the basic level of categorization. *Journal of Experimental Psychology: General*. 1997; 126:248–277.10.1037//0096–3445.126.3.248 [PubMed: 9281832]
- Mack ML, Palmeri TJ. The timing of visual object categorization. *Frontiers in Perception Science*. 2011; 2(165)10.3389/fpsyg.2011.00165
- Malcolm GL, Henderson JM. The effects of target template specificity on visual search in real-world scenes: Evidence from eye movements. *Journal of Vision*. 2009; 9(11):8, 1–13.10.1167/9.11.8 [PubMed: 20053071]
- Murphy GL, Brownell HH. Category differentiation in object recognition: Typicality constraints on the basic category advantage. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1985; 11:70–84.10.1037//0278-7393.11.1.70
- Murphy, GL. *The big book of concepts*. Cambridge, MA: MIT Press; 2002.
- Rosch EH, Mervis CB, Gray WD, Johnson DM, Boyes-Braem P. Basic objects in natural categories. *Cognitive Psychology*. 1976; 8:382–439.10.1016/0010-0285(76)90013-X
- Schmidt J, Zelinsky GJ. Search guidance is proportional to the categorical specificity of a target cue. *The Quarterly Journal of Experimental Psychology*. 2009; 62(10):1904–1914.10.1080/17470210902853530 [PubMed: 19459137]
- Schmidt J, Zelinsky GJ. Visual search guidance is best after a short delay. *Vision Research*. 2011; 51:535–545.10.1016/j.visres.2011.01.013 [PubMed: 21295053]

- Smith JD, Redford JS, Gent LC, Washburn DA. Visual search and the collapse of categorization. *Journal of Experimental Psychology: General*. 2005; 134:443–460.10.1037/0096-3445.134.4.443 [PubMed: 16316285]
- Snodgrass JG, Vanderwart M. A standardized set of 260 pictures: Normed for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*. 1980; 6(2):174–215. [PubMed: 7373248]
- Tanaka JW, Taylor M. Object categories and expertise: Is the basic level in the eye of the beholder? *Cognitive Psychology*. 1991; 23:457–482.10.1016/0010-0285(91)90016-H
- Wolfe JM, Horowitz TS, Kenner N, Hyle M, Vasan N. How fast can you change your mind? The speed of top-down guidance in visual search. *Vision Research*. 2004; 44:1411–1426.10.1016/j.visres.2003.11.024 [PubMed: 15066400]
- Yang H, Zelinsky GJ. Visual search is guided to categorically-defined targets. *Vision Research*. 2009; 49:2095–2103.10.1016/j.visres.2009.05.017 [PubMed: 19500615]

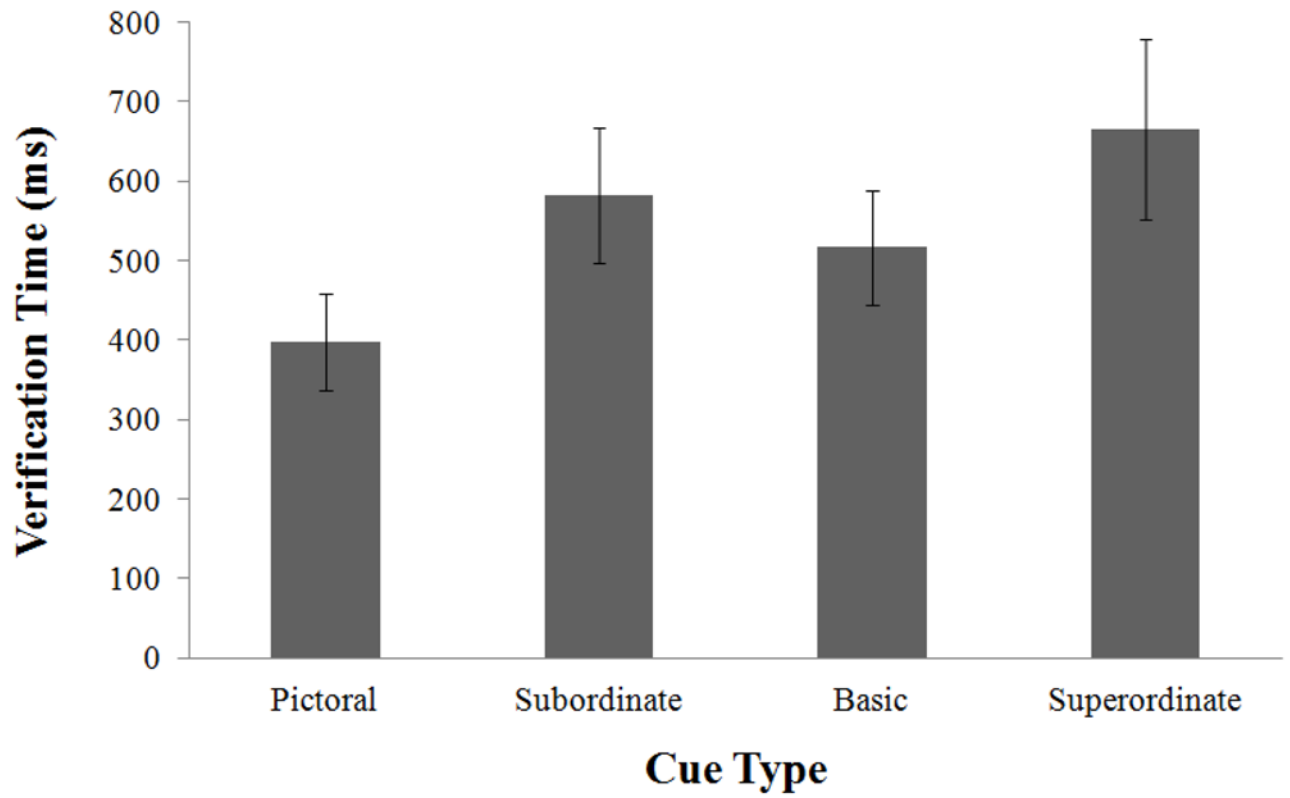


**Figure 1.** Procedure for the hierarchical search task used in Experiment 2.





**Figure 2.** Guidance measures for categorical and pictorial cues in Experiment 2. (a) Time-to-target, (b) proportion of trials in which the target was the first object fixated. Error bars show one standard error.



**Figure 3.** Verification times for pictorial and categorical cues in Experiment 2. Error bars show one standard error.

**Table 1**

Categorical targets used in category verification (Experiment 1) and search tasks (Experiment 2)

<b>Superordinate</b>	<b>Basic</b>	<b>Subordinate</b>
Vehicle	Car	Police Car
		Taxi
		Race Car
	Boat	Sail Boat
		Cruise Ship
		Speed Boat
	Plane	Passenger Airliner
		Biplane
		Fighter Jet
	Truck	18 Wheeler
		Fire Truck
		Pickup Truck
	Furniture	Cabinet
Filing Cabinet		
China Cabinet		
Chair		Folding Chair
		Office Chair
		Dining Room Chair
Bed		Twin Bed
		Canopy Bed
		Bunk Bed
Table		Coffee Table
		Dining Room Table
		End Table
Clothing	Pants	Jeans
		Dress Pants
		Pajama Pants
	Shirt	Dress Shirt
		T-shirt
		Long Sleeve Shirt
	Hat	Baseball Hat
		Knit Cap
		Cowboy Hat
	Jacket	Winter Jacket
Windbreaker		
Trench Coat		
Dessert	Ice Cream	Chocolate Ice Cream
		Mint Choc. Chip Ice Cream
		Strawberry Ice Cream

---

<b>Superordinate</b>	<b>Basic</b>	<b>Subordinate</b>
	Pie	Pecan Pie Blueberry Pie Lemon Meringue Pie
	Cookie	Oreo Chocolate Chip Cookie Sugar Cookie
	Cake	Chocolate Cake Wedding Cake Bundt Cake

---

**Table 2**

Mean error rates by condition in Experiment 2

Condition	Pictorial	Subordinate	Basic	Superordinate
Target Present	4.69 (1.7)*	9.89 (2.3)	11.18 (2.4)	22.83 (3.6)**
Target Absent	3.13 (.94)	9.38 (2.1)**	1.8 (.61)	4.69 (.71)

Note:

\*  $p < .05$ ,\*\*  $p < .01$ . Values in parentheses indicate one standard error of the mean.