Contextual memory and skill transfer in category search

JAMES A. KOLE, ALICE F. HEALY, DEANNA M. FIERMAN, AND LYLE E. BOURNE, JR.

University of Colorado, Boulder, Colorado

In three experiments, we examined transfer and contextual memory in a category search task. Each experiment included two phases (training and test), during which participants searched through category and exemplar menus for targets. In Experiment 1, the targets were from one of two domains during training (grocery store or department store); the domain was either the same or changed at test. Also, the categories were organized in one of two ways (alphabetically or semantically); the organization either remained the same or changed at test. In Experiments 2 and 3, domain and organization were held constant; however, categories or exemplars were the same, partially replaced, or entirely replaced across phases in order to simulate the dynamic nature of category search in everyday situations. Transfer occurred at test when the category organization or domain was maintained and when the categories or exemplars matched (partially or entirely) those at training. These results demonstrate that transfer is facilitated by overlap in training and testing contexts.

Searching for information needed to complete a desired action is a common everyday experience. Examples—such as finding the telephone number of a local department store, searching for a book title on Amazon, or locating the required troubleshooting steps to fix a computer glitch—are both familiar and numerous. The relevant information is generally available in a directory, in a manual, or on the Internet and needs to be searched for, often by category.

Consider an individual using a relatively unfamiliar word-processing program at his or her workplace. If a situation arises that requires the user to perform a novel action, the first step that the individual might take is to search through the online help manual for that action, which requires searching through menus of categories and subcategories for the required information, as well as verifying that a given item solves the problem at hand. Similar situations arise when searching through Web sites or even through an unfamiliar grocery store for certain products. A difficult search is likely to be inefficient and might result in failure with consequent stress and frustration on the part of the user.

Previous research has revealed that users' ability to locate information in a menu-based hierarchically organized system is variable and, in some instances, is highly inefficient. Whereas some studies have found that users can optimally locate (i.e., select the correct categories and subcategories without having to return to previous menus) target information as often as 95% of the time (e.g., Allen, 1983; Miller, 1981), others have found much lower rates. Frankhuizen and Vries (1980) reported optimal location of targets only 56% of the time, with participants accessing approximately twice the minimum

number of menus required, and in a study by Whalen and Latremouille (1981), users completely failed to locate the target information 28% of the time. In general, better performance is associated with smaller set sizes and more distinct categories (Dumais & Landauer, 1984), transfer to new category exemplars is better with a larger number of trained exemplars (e.g., Homa & Vosburgh, 1976), and performance improves with increased practice on the task or with expertise in the search domain (e.g., Salmeron, Canas, & Fajardo, 2005).

A complication that might impact performance, particularly on a menu-based category search task, is that category sets and category contents may be dynamic. That is, the set of categories used to partition the search space, as well as the items that constitute the search space, may change over time. For example, with regard to software programs, entirely new categories of functions may be introduced over time, or functions might be added to or removed from the program under an existing category. Even when the content of the categories and subcategories is static, users may be required to search for novel target information or through categories previously available but never before containing target information. Interesting questions arise from the changing nature of the search space and search targets with regard to the training and transfer of skill. In particular, does training the search task result in a generalizable search skill or strategy that is independent of the specific categories and category exemplars included in the training set? How do changes in the search space impact subsequent performance? Does training with one set of categories and/or category exemplars transfer to subsequent performance on the task with difference categories,

different category exemplars, or different targets? And if so, what principles underlie this transfer?

The ultimate goal of any training regimen is to develop skills that positively transfer beyond the conditions of training, such that high levels of performance may be maintained with task changes or in situations that require extrapolating the skills acquired during training. However, studies of transfer often reveal marked specificity, with changes in stimuli (e.g., Bourne, Healy, Pauli, Parker, & Birbaumer, 2005; Rickard, Healy, & Bourne, 1994), changes in the required responses (e.g., Healy, Wohldmann, Sutton, & Bourne, 2006), or changes in task procedures (e.g., Healy, Wohldmann, Parker, & Bourne, 2005) often resulting in a return to pretraining levels of performance. Previous researchers who have found specificity of training have examined primarily procedural or skill-based tasks. One notable exception to the finding of specificity occurs when tasks are rooted in declarative memory, as opposed to procedural memory, such as in a complex computer programming task (Pennington, Nicolich, & Rahm, 1995; cf. Singley & Anderson, 1989) or a fact-learning task (Kole & Healy, 2007; Postman & Underwood, 1973; Van Overschelde & Healy, 2001). For example, Kole and Healy (2007, Experiment 3) found that previously acquired knowledge can improve the learning of novel facts even from a different domain. In that experiment, participants learned a large number of facts about unfamiliar countries while associating these unfamiliar countries with individual people who were well known to them. It was found that the participants who used their prior knowledge of individuals learned and retained over twice the number of facts as did the participants who did not. The differential effect of task requirements (i.e., procedural or declarative) on transfer has been formalized in the empirically based procedural reinstatement principle (Healy, 2007), which generalizes previous observations that limited retention but robust transfer occur when tasks are primarily declarative in nature, whereas robust retention but limited transfer occur when tasks are primarily procedural (e.g., Ivancic & Hesketh, 2000; Lee & Vakkoch, 1996; Maxwell, Masters, Kerr, & Weedon, 2001). By this principle, training on a search task should yield transfer largely to the extent that declarative information is learned, as opposed to procedural information.

Transfer in a category search task has been examined in only a few studies. For example, Neisser and Beller (1965) trained participants to search for a set of category exemplars. Following training, the participants were given a new set of exemplars to search for from the practiced categories. These untrained exemplars were detected as accurately as were the trained category exemplars, thereby demonstrating complete transfer of training. Schneider and Fisk (1984, Experiment 1B) extended the findings of Neisser and Beller by demonstrating that the observed transfer effect was not due to general skill learning but was instead category specific. In Schneider and Fisk's experiment, participants were shown a category name and were subsequently shown three exemplars from different categories; their task was to identify which of the three was an exemplar of the presented category. After being

trained on a set of exemplars, the participants were given a transfer task in which the targets included trained exemplars, untrained exemplars from trained categories, and untrained exemplars from new categories. Performance on the transfer task was equivalent for trained and untrained exemplars from trained categories, which replicates the findings of Neisser and Beller, and performance was better for these two types of targets than for untrained exemplars from untrained categories. The results were consistent with category learning, as opposed to either individual item learning or general skill learning; however, Schneider and Fisk did not examine whether there was any transfer of training to untrained exemplars from untrained categories. That is, they did not compare performance on the transfer test for the untrained exemplars from untrained categories with performance at the start of training. Finding such transfer of training would indicate general skill learning in this task.

In the present study, we expand upon the findings of Neisser and Beller (1965) and of Schneider and Fisk (1984), using a task that is in some sense a reversal of the one that they employed. The search task that we used involved the search for a target within a hierarchically organized information structure. Specifically, participants were presented with a target, which was a category exemplar, and were required to search for that target first among a set of predefined categories, then among a list of exemplars within a category. Thus, rather than being presented with a category name and asked to select an appropriate exemplar of that category from a list of exemplars (as was the case in both Neisser & Beller, 1965, and Schneider & Fisk, 1984), in our task, the participants were presented with a category exemplar and were required, at least initially, to select the appropriate category from a list of category names, thereby more realistically simulating the nature of everyday search tasks involving menus of categories. This search task itself can be quite complicated and might involve, in addition to perceptual processes, some semantic processes, such as mapping one's own categorical labels to those provided in the task, category learning, and attending to different dimensions of the target in order to make a category membership judgment.

In the present study, we investigated some basic issues involved in the category search task. In Experiment 1, we examined whether training the search task on a circumscribed area of one domain would transfer to search both within and outside of that domain, or in other words, we examined whether training of the search task would result in a generalizable search skill that transcends memory for categories and exemplars. Finding such a result would demonstrate that transfer can be based on procedural information. We also investigated whether holding category organization constant, regardless of specific category content, would improve performance.

In Experiments 2 and 3, we examined the contribution of context memory and contextual reinstatement to transfer. We define *context* to be those elements of the category search task that are irrelevant to the search for a particular target, and context includes distractor categories (i.e., categories that do not contain the target) and distractor

exemplars (i.e., nontarget exemplars; for a similar definition of *context*, see Schneider & Fisk, 1984). We varied the set of targets on transfer trials either at the level of category (Experiment 2) or at the level of category exemplar (Experiment 3), such that targets either belonged to distractor categories or were distractors within a category during training. If context memory contributes to transfer, we would predict an advantage for transfer trials relative to an unpracticed baseline. In addition, during transfer trials, the distractor categories (Experiment 2) or category exemplars (Experiment 3) were either the same, partially replaced, or entirely replaced relative to those at training. Thus, this design also allowed us to examine whether transfer would occur in a graded fashion and whether reinstating context at test would facilitate transfer. If contextual reinstatement contributes to transfer, we would predict an advantage on transfer trials when distractor categories or category exemplars are the same as those during training. In addition, unlike the earlier studies by Neisser and Beller (1965) and Schneider and Fisk (1984), in the present study, we compared performance on new targets during transfer trials with performance at the start of training in order to evaluate the extent of general skill learning.

EXPERIMENT 1

In Experiment 1, participants completed a two-tiered category search task. The participants were presented with target items, which were category exemplars, along with a list of categories. The participants were to select a category to which they believed the target might belong; once this selection had been made, the contents of the selected category were displayed, which consisted of a set of exemplars. If the target was among the set and was selected by the participant, a new target was presented; if not, the search continued iteratively until the target had been located within a category. The categories and category exemplars were from one of two domains (grocery store, department store), and the categories were arranged according to one of two types of organization (alphabetical, semantic). There were two phases during the experiment (training, test); search domain and category organization were either held constant or changed from training to test.

The purpose of Experiment 1 was to examine two levels of skill transfer in a category search task. The first level of transfer was *domain-general transfer*, which may be defined as an improvement over an unpracticed baseline when the search domain is changed, such that both categories and category exemplars change. Thus, domain-general transfer represents the learning of a generalizable search strategy that is not dependent on memory for specific, previously encountered items and is therefore attributable to procedural knowledge (e.g., knowledge about how to search through a given category organization). The second level of transfer was *domain-specific transfer*, which may be defined as an improvement over an unpracticed baseline when the search is for different targets or through different categories within the same domain. Thus, domain-specific transfer represents the learning of a specific search strategy that may be dependent on mem-

ory for previously encountered information and that may be based on declarative knowledge. In the category search task used in the present experiment, this information may include memory for categories, category order, category exemplars, or category boundaries, such that new targets may be categorized more quickly.

Method

Participants

Thirty-two undergraduate students at the University of Colorado participated for either payment or course credit in an introductory psychology class. The participants were assigned to conditions according to a fixed rotation on the basis of their time of arrival for testing.

Materials and Apparatus

The experiment consisted of a training phase and a test phase. For both training and test, the participants were shown the name of an item (i.e., the target) and a list of categories on a computer (see the Appendix for categories and category exemplars for both the grocery store and department store domains). The participants were to search for the target by clicking on a category to which the target might belong, then searching through the category exemplars that appeared, and finally, clicking on the target name if it was listed. If the target was not listed in the selected category, the participants were to search within other categories from the main list until the target had been located.

The targets were objects found either in a grocery store or in a department store, and the categories were organized either alphabetically or semantically. The two particular organization types were selected because they are the most commonly used in everyday life. The targets themselves, their categorization, and their semantic organization were based on both intuition and actual observation in relevant stores; for example, in the grocery store domain, fruits and vegetables were listed together in the semantic organization condition, since they are typically found together in a grocery store. In the two phases of the experiment, the participants searched for targets in either the same domain or different domains (grocery store or department store), and the categories were organized in either the same manner or different manners (alphabetically or semantically).

Two lists were created for each domain, both of which contained the same set of 24 categories; however, the four exemplars associated with each category varied between lists. For example, in the grocery store domain, the beverage category in the first list contained *apple cider*, *milk*, *orange juice*, and *tomato juice*, whereas the same category in the second list contained *bottled water*, *cranberry juice*, *fruit punch*, and *tea*. The categories and exemplars were selected largely on the basis of perusal of the organization of actual grocery and department stores, rather than on any existing norms. Within each set of four exemplars, one exemplar was defined a priori to be a difficult search item, on the basis of both intuition and the likelihood of being classified in more than one category. Across participants, each list was used equally frequently during training and test in each of the between-subjects conditions, and for all of the participants, different lists were used for training and test.

For both training and test, there were four blocks of 24 trials; within each block, one exemplar from each of the 24 categories was presented as a target. Thus, a total of 96 targets were presented during each phase, with all four exemplars from each category in a given list serving once as a target. Within each block of 24 trials, 6 targets were classified as *difficult* and 18 as *easy*, and the targets were presented in a random order with the constraint that the final target in one block was not a member of the same category as the initial target in the following block. Between the training and test phases, the participants performed a 10-min irrelevant Stroop distractor task.

The computer program used for the search task was designed with RealBasic software. The primary experiment was run on an Apple iMac computer operating under OS 9, and the distractor task was run on an Apple iBook computer operating under OSX.

Design

The experiment consisted of a 2 (phase: training, test) $\times$ 4 (block within phase: 1–4) $\times$ 2 (item difficulty: easy, difficult) $\times$ 2 (domain: same, different) $\times$ 2 (category organization: same, different) mixed-factorial design. The first three factors varied within subjects, and the last two factors varied between subjects. Item difficulty was included as a factor here, but, for simplicity, that factor is omitted from the present summary because the results were generally consistent across the two types of items and the categorization into easy and difficult items was done informally. The dependent measures included the number of clicks (how many clicks were made before successfully locating a target; two is the minimum number of clicks corresponding to correct identification of the category and the exemplar within that category) and response time (RT) per click. We included the RT/click measure as an assessment of RT that is independent of the number of clicks. To remove the problem of outliers, statistical analyses involving the number of clicks variable were computed using medians as the summary measure for each participant.

Procedure

No practice trials were administered, because one primary point of interest was the speed with which the participants became familiar with the categories and category exemplars, and to examine transfer, an unpracticed baseline was required. The participants began each trial with the presentation of the target and by clicking the “ready” button, at which point a list of categories appeared on the screen. The participants clicked on a category entry to view the four exemplars associated with that category, which were always listed in alphabetical order. If the target was among the displayed exemplars, the participants clicked on the exemplar and began the next trial. However, if the target was not among the displayed exemplars, the participants clicked the “back” button to return to the list of categories, and the exemplars from the previously selected category disappeared. The target exemplar name was present on the left side of the screen at all times, so that the participants could easily remember the item for which they were searching.

At the end of training, the participants were told to contact the experimenter, at which point they immediately performed the distractor (Stroop) task. Following the distractor task, the participants completed the test phase, which proceeded in the same manner as training, with two exceptions: For half of the participants, the search domain changed, and for the other half, the domain remained the same; for half of the participants within each of these groups, the organization of categories (i.e., alphabetical or semantic) changed.

Results

In this and the subsequent Results sections, all significant results ($\alpha = .05$) are reported, and nonsignificant results are reported only when noted.

Training and Test

Number of clicks. Across the first three blocks, the number of clicks required to complete each trial decreased (Block 1, $M = 4.309$ clicks; Block 2, $M = 4.229$ clicks; Block 3, $M = 3.613$ clicks; Block 4, $M = 3.980$ clicks); the main effect of block was significant [$F(3,84) = 8.808$, $MS_e = 1.402$, $p < .001$, $\eta^2 = .239$]. There was also an interaction between category organization and phase [$F(1,28) = 4.205$, $MS_e = 2.509$, $p = .050$, $\eta^2 = .131$]. This interaction reflects the fact that the number of clicks decreased from training to test when category organization (alphabetical or semantic) was held constant (training, $M = 4.316$ clicks; test, $M = 3.922$ clicks); however, the opposite pattern was observed when category organi-

zation changed from training to test (training, $M = 3.852$ clicks; test, $M = 4.031$ clicks).

RT/click. The analysis of RT/click revealed main effects of block [$F(3,84) = 96.004$, $MS_e = 0.093$, $p < .001$, $\eta^2 = .774$] and phase [$F(1,28) = 20.382$, $MS_e = 0.112$, $p < .001$, $\eta^2 = .421$]. RT/click improved across blocks (Block 1, $M = 1.541$ sec/click; Block 2, $M = 1.126$ sec/click; Block 3, $M = 0.989$ sec/click; Block 4, $M = 0.979$ sec/click), and from training ($M = 1.226$ sec/click) to test ($M = 1.092$ sec/click).

Transfer

Following previous investigations (e.g., Healy et al., 2005; Healy et al., 2006), to examine transfer, we conducted analyses comparing the first block of training (which served as an unpracticed baseline) with the first block of test for each dependent measure; positive transfer occurred if there was improvement from the first block of training to the first block of test (see Table 1). These analyses, which include only the first block of each phase, provide a cleaner assessment of transfer than did the previous analyses, which included all blocks, because the additional blocks of test reflect training (relearning) as well as transfer. In the following analyses, the effects reported are limited to those involving the transfer variable (Training Block 1, Test Block 1).

Number of clicks. The analysis of number of clicks revealed an interaction between transfer and category organization [$F(1,28) = 4.416$, $MS_e = 1.755$, $p = .045$, $\eta^2 = .136$]. This interaction occurred because some positive transfer was evident when the category organization was the same between phases (training, $M = 4.547$ clicks; test, $M = 4.281$ clicks), but the opposite pattern (negative transfer) was evident when the category organization changed between phases (training, $M = 3.844$ clicks; test, $M = 4.562$ clicks). The positive transfer when the category organization was the same was not significant [$F(1,15) = 1.249$, $MS_e = 0.904$, $p = .281$, $\eta^2 = .077$], and the negative transfer when the category organization was different only approached significance [$F(1,15) = 3.446$, $MS_e = 2.399$, $p = .083$, $\eta^2 = .187$].

RT/click. The analysis of RT/click revealed a main effect of transfer [$F(1,28) = 8.690$, $MS_e = 0.161$, $p = .006$, $\eta^2 = .237$], with performance better during test ($M = 1.437$ sec/click) than during training ($M = 1.646$ sec/click). The amount of positive transfer depended on domain [$F(1,28) = 5.143$, $MS_e = 0.161$, $p = .031$, $\eta^2 = .155$]; greater improvement occurred when the search domain was the same during training and test (training, $M = 1.648$ sec/click; test, $M = 1.279$ sec/click) than when the search domain changed (training, $M = 1.643$ sec/click; test, $M = 1.595$ sec/click). Positive transfer was significant when the search domain was the same [$F(1,15) = 14.53$, $MS_e = 0.151$, $p = .002$, $\eta^2 = .492$], but not when it changed ($F < 1$).

Discussion

The purpose of the present experiment was to examine the type or level of transfer in a category search

Table 1
Transfer Performance (Mean Number of Clicks and Response Time [RT] per Click)
on the First Block of Training and First Block of Test in Experiments 1, 2, and 3

Condition	Number of Clicks				RT/Click			
	Train Block 1		Test Block 1		Train Block 1		Test Block 1	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
Experiment 1								
Same D, same O	4.500	0.365	4.250	0.299	1.531	0.092	1.154	0.083
Same D, diff O	3.625	0.287	4.500	0.447	1.765	0.181	1.403	0.101
Diff D, same O	4.594	0.416	4.312	0.338	1.511	0.097	1.561	0.110
Diff D, diff O	4.062	0.382	4.625	0.507	1.776	0.177	1.629	0.166
Experiment 2								
Match	4.213	0.198	4.546	0.234	2.294	0.132	1.502	0.076
Partial match	5.121	0.510	4.571	0.229	2.022	0.109	1.583	0.082
No match	4.442	0.210	4.429	0.242	2.127	0.082	1.845	0.100
Experiment 3								
Match	6.862	1.256	5.682	0.463	2.488	0.119	1.532	0.090
Partial match	7.454	0.672	5.803	0.620	2.360	0.131	1.505	0.076
No match	9.644	1.905	7.893	0.829	2.586	0.127	1.668	0.090

Note—D, domain; O, organization; diff, different; train, training.

task. Domain-general transfer occurs if, after training the category search task on one domain (e.g., grocery store items), performance is facilitated when performing the category search task on a different domain (e.g., department store items) and is attributable to procedural knowledge. Domain-specific transfer occurs if, after training the category search task on one domain, performance is facilitated when performing the category search task with different targets from the same domain (e.g., searching for new grocery store items) and may be attributable, at least in part, to declarative knowledge.

To begin, the results of this experiment clearly indicate that performance improves with practice in the category search task. For both training and test, performance on both dependent measures showed improvement across blocks and, in one instance (the measure of RT/click), across phases. Facilitation in category search may be attributable to multiple sources. First, declarative memory for categories and/or category exemplars might benefit performance—in the case of categories because candidate categories may be identified for a given target without having to perform a time-intensive search of the category list, and in the case of category exemplars because category membership for a given target may be identified, again without having to make a time-consuming judgment as to the likelihood that a target belongs to a category. A second possibility is that practice with a category search task results in category learning, such that participants obtain a general sense of how the search space is partitioned even without explicit memory for categories or associated category exemplars. In this case, the categorization of targets is more efficient, because targets are more readily classified into categories. A third possibility is that training the category search task results in a general search strategy that is not dependent on memory for categories, category exemplars, or category learning. For example, for semantically organized categories, participants might limit their search to the relevant category cluster, and for alphabeti-

cally organized categories, they might limit their search to categories that begin with the appropriate letter, so that the entire list of categories does not need to be searched. A final, less interesting possibility is that practice with the search task benefits from psychomotor fluency or proficiency in using the task interface.

Examining both domain-general and domain-specific transfer allows for the isolation of some of the factors underlying performance improvements. We found some evidence for domain-general transfer; there was a significant interaction between organization and transfer for the measure of number of clicks. Specifically, performance improved more from training to test when the category organization was the same, regardless of domain, than when it changed. We also found evidence for domain-specific transfer for the measure of RT/click; performance at the start of test was better than that at the start of training when the domain was held constant. With respect to the specific types of information playing some role in transfer, these findings are consistent with the role of memory for categories, category learning, the learning of a general search strategy, and psychomotor fluency but are inconsistent with the role of memory for category exemplars, because the set of exemplars associated with each category changed from training to test even when the search domain remained the same. Consequently, in a subsequent experiment (Experiment 3), we considered a search task that varied the continuity of category exemplars from training to test. However, first, in the next experiment (Experiment 2), we examined a search task that varied the continuity of categories from training to test.

EXPERIMENT 2

The results of Experiment 1 provide evidence for domain-specific transfer. The purpose of Experiment 2 was to examine possible bases for such transfer. More specifically, we considered whether, with respect to cat-

egories, memory for contextual information from training and the reinstatement of contextual information during test contribute to domain-specific transfer in the category search task. In this experiment, domain and category organization were both held constant across phases, so that examination of domain-specific transfer would not be diluted by other variables. However, the set of categories presented during training and test varied in the degree to which they matched. In one condition (the match condition), the set of categories was held constant from training to test, with half of the categories containing targets during training and the other half containing targets during test. In a second condition (the partial-match condition), half of the categories from training were replaced with new categories during test; the categories that were repeated across phases contained targets only during test. In a third condition (the no-match condition), the entire set of categories changed from training to test, again with half of the categories containing targets during test.

If memory for contextual information from training contributes to domain-specific transfer, there should be greater improvement from the first block of training to the first block of test when targets are located in categories that were presented as context (i.e., did not contain targets) during training, as in the match and partial-match conditions, relative to when targets are located in novel categories, as in the no-match condition. Furthermore, if the reinstatement of contextual information from training contributes to domain-specific transfer, there should be greater improvement from the first block of training to the first block of test for the match condition, in which categories are presented in the same context during both phases (i.e., all categories are held constant from training to test), than for the partial-match condition, in which context changes (i.e., categories are partially replaced from training to test). Thus, performance at test should be best for the match condition (in which all categories were shown previously), next best for the partial-match condition (in which only half of the categories were shown previously), and worst for the no-match condition (in which none of the categories were shown previously). Note, however, that the partial-match condition might suffer from interference similar to that obtained in part-set cuing studies (see, e.g., Del Missier & Terpini, 2009; Slamecka, 1968), thus leading to the worst performance in that condition rather than intermediate performance.

Method

Participants

A total of 120 undergraduate students at the University of Colorado participated for either payment or course credit in an introductory psychology class. None of these students had participated in Experiment 1. The participants were assigned to conditions according to a fixed rotation on the basis of their time of arrival for testing.

Materials and Apparatus

As did Experiment 1, Experiment 2 consisted of two phases, training and test, separated by a 10-min irrelevant distractor task. Only one domain was used in the present experiment (grocery store items), and the categories were always organized alphabetically. The search task was slightly modified: To avoid creating new categories

or exemplars, the participants were presented with only 12 categories, rather than 24, and there were 24 trials during each phase, rather than 96. Also as in Experiment 1, four exemplars were associated with each category, three classified as *easy* and one classified as *difficult*. For both training and test, trials were divided into six blocks, with 4 trials (3 easy, 1 difficult) within each block, and search was only required within 6 of the 12 listed categories. That is, although 12 categories were presented to the participants, only 6 of those categories contained targets, with all four exemplars from each of these categories in a given list serving as targets.

The participants in the match condition were presented with the same set of 12 categories during the training and test phases, with 6 of the categories containing targets during training and the other 6 containing targets during test (see Figure 1A). Thus, in this condition, the categories that served as context (i.e., categories that did not contain targets) during training contained targets during test. Furthermore, because the same 12 categories were presented during the training and test phases, the context for the categories containing targets during test remained the same during the two phases.

The participants in the partial-match condition were presented with six categories from training, as well as six novel categories at test. The six repeated categories contained targets during test, but did not contain targets during training (see Figure 1B). Thus, as in the match condition, the categories that served as context during training contained targets during test. However, because half of the categories changed from training to test, the context for the categories containing targets during test changed across phases.

The participants in the no-match condition were presented with 12 novel categories at test, with only half of the categories containing targets (see Figure 1C). Thus, in this condition, the categories that contained targets during test did not serve as context during training, and the context for target-containing categories was new during test.

The 24 categories were divided into four sets of 6 categories. The assignment of category set to conditions (e.g., contained or did not contain targets during training or test) was counterbalanced across participants such that each category set occurred equally often in each condition. Although half of the categories in the partial-match condition and all of the categories in the match condition were repeated during test, the set of exemplars associated with these repeated categories was different between phases in all conditions, because different lists were used in the two phases. As in Experiment 1, across participants in each condition, each list was used equally frequently during training and test.

Design

The experiment consisted of a 2 (phase: training, test) $\times$ 6 (block within phase: 1–6) $\times$ 2 (item difficulty: easy, difficult) $\times$ 3 (condition: match, partial match, no match) mixed-factorial design, with the first three factors varied within subjects and the last factor varied between subjects. Again, the factor of item difficulty is not included in the present summary. The dependent measures included the number of clicks and RT/click.

Procedure

The procedure was the same as that of Experiment 1. The participants were presented with a target, clicked on the most appropriate category, and either clicked on the target name if they saw it listed among the category exemplars or clicked the “back” button to return to the category list and select a different category if they did not see the target listed as a category member. The first phase (training) was immediately followed by a 10-min distractor task (either a letter detection task or a Stroop task) and then by the second phase (test).

Results

Training and Test

Number of clicks. The analysis of the number of clicks revealed a main effect of block [$F(5,585) = 3.018$,

Figure 1. Category structure during training and test for the match (A), partial-match (B), and no-match (C) conditions of Experiment 2. Shaded categories indicate those that contain targets, and arrows indicate categories that were repeated across phases.

$MS_e = 8.974$, $p = .011$, $\eta^2 = .025$], with the number of clicks generally decreasing across blocks (Block 1, $M = 4.553$ clicks; Block 2, $M = 4.435$ clicks; Block 3, $M = 4.445$ clicks; Block 4, $M = 4.492$ clicks; Block 5, $M = 4.225$ clicks; Block 6, $M = 3.908$ clicks).

RT/click. The analysis of RT/click revealed main effects of phase [$F(1,117) = 26.823$, $MS_e = 0.603$, $p < .001$, $\eta^2 = .187$] and block [$F(5,585) = 152.359$, $MS_e = 0.395$, $p < .001$, $\eta^2 = .566$]. These main effects arise because RT/click decreased from training ($M =$

Figure 2. Mean response time (RT; in seconds) per click as a function of block and phase in Experiment 2. Error bars represent ± 1 standard error of the mean.

1.322 sec/click) to test ($M = 1.172$ sec/click), and generally improved across blocks (see Figure 2). The two-way interaction between phase and block was significant [$F(5,585) = 10.246$, $MS_e = 0.359$, $p < .001$, $\eta^2 = .081$], reflecting the fact that RT/click during training was longer than during test, but primarily for the first block (see Figure 2).

The improvement across blocks within each phase and from training to test depended on condition; the two-way interactions between condition and block [$F(10,585) = 2.513$, $MS_e = 0.395$, $p = .006$, $\eta^2 = .041$], and condition

and phase [$F(2,117) = 3.762$, $MS_e = 0.603$, $p = .026$, $\eta^2 = .060$] were both significant. The former interaction arose because there were larger differences between conditions during the first four blocks than during the last two blocks (see Figure 3). The latter interaction arose because performance gains across phases were greatest for the match condition (training, $M = 1.375$ sec/click; test, $M = 1.131$ sec/click), least for the no-match condition (training, $M = 1.295$ sec/click; test, $M = 1.245$ sec/click), and intermediate for the partial-match condition (training, $M = 1.295$ sec/click; test, $M = 1.140$ sec/click).

Figure 3. Mean response time (RT; in seconds) per click as a function of block and condition in Experiment 2. Error bars represent ± 1 standard error of the mean.

Transfer

As in Experiment 1, transfer was examined by comparing the first block of training to the first block of test for each of the dependent measures (see Table 1). In the following analyses, the effects reported are again limited to those involving the transfer variable (Training Block 1, Test Block 1).

Number of clicks. No significant effects involving the transfer variable were found in the analysis of number of clicks.

RT/click. The analysis of RT/click revealed a main effect of transfer [$F(1,117) = 57.279$, $MS_e = 0.532$, $p < .001$, $\eta^2 = .329$], as well as an interaction between transfer and condition [$F(2,117) = 5.127$, $MS_e = 0.532$, $p = .007$, $\eta^2 = .081$]. On average, RT/click showed positive transfer (Training Block 1, $M = 2.147$ sec/click; Test Block 1, $M = 1.644$ sec/click). However, transfer was significantly greater for the match and partial-match conditions than for the no-match condition [$F(1,118) = 5.403$, $MS_e = 0.548$, $p = .022$, $\eta^2 = .044$], which suggests that memory for categories from training contributes to transfer. Furthermore, transfer was greater for the match condition than for the partial-match condition [$F(1,78) = 4.104$, $MS_e = 0.606$, $p = .046$, $\eta^2 = .050$], reflecting the importance of reinstating category context during test. Nonetheless, transfer was significant for each group [match, $F(1,39) = 38.276$, $MS_e = 0.654$, $p < .001$, $\eta^2 = .495$; partial match, $F(1,39) = 13.762$, $MS_e = 0.559$, $p < .001$, $\eta^2 = .261$; no match, $F(1,39) = 8.303$, $MS_e = 0.382$, $p = .006$, $\eta^2 = .176$].

Discussion

In Experiment 2, we examined the contribution of memory for category information to transfer in the category search task. We first considered whether performance during test would be facilitated when targets were located in categories that served as context during training by varying the degree of match between training and test categories. Greater transfer was evident in terms of RT/click when the categories were repeated across phases (i.e., for the match and partial-match conditions) than when the categories were entirely new during test (i.e., for the no-match condition). However, planned comparisons revealed that the improvement across phases was significant even for the no-match condition, in which an entirely different set of categories was presented during test. This particular result suggests that procedural learning was transferred from training to test. Also, psychomotor learning of the task interface might have aided transfer. However, the advantage for the match and partial-match conditions over the no-match condition suggests that, in addition to procedural and psychomotor learning, declarative knowledge of category information contributes to domain-specific transfer from training to test. Furthermore, the numerical advantage for the partial-match condition over the no-match condition rules out any interference effects like those found with part-set cuing (Del Missier & Terpin, 2009; Slamecka, 1968).

A related question is whether reinstating category context during test contributes to transfer. We found greater transfer (again in terms of RT/click) when categories com-

pletely overlapped from training to test, as in the match condition, than when categories only partially overlapped from training to test, as in the partial-match condition. This finding demonstrates that contextual information was learned and used during the search task and did facilitate transfer.

The results jointly suggest that with regard to categories, memory for contextual category information from training and the reinstatement of category context during test independently contribute to domain-specific transfer. A second possibility is that learning might be based on memory for exemplars or for which specific items are associated with a given category. This possibility could not be fully explored in Experiment 2, because the set of exemplars associated with each category changed from training to test. The contribution of memory for exemplar information is further explored in Experiment 3.

EXPERIMENT 3

In Experiment 3, we again examined whether contextual memory and contextual reinstatement contribute to transfer, but in this case with respect to category exemplars. In the present experiment, domain, category organization, and categories were all held constant during training and test. Conditions were compared in which (1) the set of four exemplars in a given category was held constant from training to test (match), (2) two exemplars were held constant and two were replaced with new exemplars during test (partial match), and (3) the set of four exemplars changed from training to test (no match).

If, with respect to category exemplars, memory for contextual information from training contributes to domain-specific transfer, there should be greater improvement from the first block of training to the first block of test when targets were presented as context (i.e., served as distractors) during training, as in the match and partial-match conditions, than when targets were entirely novel, as in the no-match condition. Furthermore, if the reinstatement of contextual information from training contributes to domain-specific transfer, there should be greater improvement from the first block of training to the first block of test for the match condition, in which exemplars were presented in the same context during the two phases (i.e., all exemplars were held constant from training to test) than for the partial-match condition, in which context changed (i.e., exemplars were partially replaced from training to test).

Method

Participants

A total of 48 undergraduate students at the University of Colorado participated for either payment or course credit in an introductory psychology class. None of these students had participated in Experiment 1 or 2. The participants were assigned to conditions according to a fixed rotation on the basis of their time of arrival for testing.

Materials and Apparatus

As did Experiments 1 and 2, Experiment 3 consisted of two phases, training and test, separated by a 10-min distractor task. Only one domain was used in the present experiment (grocery store items), and the same categories were listed during training and test, with cate-

ries organized alphabetically. The search task was slightly modified from that in Experiment 2: The participants were presented with the full set of 24 categories, rather than just 12 (because there was no comparison of trained and untrained categories), and there were 48 trials during each phase, rather than 24. Thus, there were 8 trials per block (6 easy, 2 difficult), rather than 4. Also as in Experiments 1 and 2, four exemplars were associated with each category, three classified as *easy* and one classified as *difficult*. For both training and test, only two exemplars from each category served as targets.

For the participants in the match condition, the same set of four category exemplars were associated with each category during the two phases; two of the exemplars served as targets during training, and the other two served as targets during test (see Figure 4A). Thus, in this condition, the two exemplars that served as context (i.e., were distractors) during training were targets during test. Furthermore, because the same four exemplars were presented during training and test, the context for test targets remained the same during the two phases.

For the participants in the partial-match condition, for each category during test, two exemplars were repeated from training and two were new. The two repeated exemplars were distractors during training and targets during test (see Figure 4B). Thus, as in the match condition, the exemplars that served as context during training served as targets during test. However, because the set of exemplars associated with each category partially changed from training to test, the context for test targets changed across phases.

For participants in the no-match condition, four novel exemplars were associated with each category at test, with two of the exemplars serving as targets in a given phase (see Figure 4C). Thus, in this condition, the exemplars that served as context during training did not serve as targets during test, and the context for targets was new during test.

Lists were constructed such that across participants, each exemplar was assigned to each condition (e.g., served as a target or a distractor during training or test) equally often.

Design

The experiment consisted of a 2 (phase: training, test) $\times$ 6 (block within phase: 1–6) $\times$ 2 (item difficulty: easy, difficult) $\times$ 3 (condition: match, partial match, no match) mixed-factorial design. The first three factors varied within subjects, and the last factor varied between subjects. Because the number-of-clicks variable did not yield significant effects in the transfer analysis of Experiment 2 (or in Experiment 1 for the factor of domain), we introduced another, possibly more sensitive, pure accuracy measure in Experiment 3. Specifically, in addition to number of clicks and RT/click, the dependent measures included the proportion of perfect trials (accuracy) in which the participants selected the correct category and exemplar without having to backtrack through the category and exemplar lists.

Procedure

The procedure was the same as those of Experiments 1 and 2, except that the distractor task always involved letter detection.

Results

Training and Test

Number of clicks. The analysis of number of clicks revealed main effects of phase [$F(1,45) = 16.549$, $MS_e = 26.225$, $p < .001$, $\eta^2 = .269$] and block [$F(5,225) = 2.499$, $MS_e = 27.920$, $p = .0316$, $\eta^2 = .053$]. The number of clicks decreased from training ($M = 7.286$ clicks) to test ($M = 6.058$ clicks) and generally improved across blocks (Block 1, $M = 7.223$ clicks; Block 2, $M = 6.276$ clicks; Block 3, $M = 6.201$ clicks; Block 4, $M = 7.597$ clicks; Block 5, $M = 6.583$ clicks; Block 6, $M = 6.152$ clicks).

Figure 4. Category exemplar structure shown for a sample category (Category 1) during training and test for the match (A), partial-match (B), and no-match (C) conditions of Experiment 3. Shaded exemplars are targets, and arrows indicate exemplars that were repeated across phases.

RT/click. The analysis of RT/click again revealed main effects of phase [$F(1,45) = 189.287$, $MS_e = 0.301$, $p < .001$, $\eta^2 = .808$] and block [$F(5,225) = 39.885$, $MS_e = 0.256$, $p < .001$, $\eta^2 = .470$]. RT/click decreased from training ($M = 1.803$ sec/click) to test ($M = 1.358$ sec/click) and improved monotonically across blocks (Block 1, $M = 2.023$ sec/click; Block 2, $M = 1.626$ sec/click; Block 3, $M = 1.524$ sec/click; Block 4, $M = 1.465$ sec/click; Block 5, $M = 1.448$ sec/click; Block 6, $M = 1.395$ sec/click).

The improvement across blocks depended on both phase and condition [$F(10,225) = 2.113$, $MS_e = 0.268$, $p = .025$, $\eta^2 = .086$]. This three-way interaction reflects the fact that there were smaller between-condition differences during the first two blocks of training and during the first, third, and last blocks of test than during the other blocks (see Figure 5). The component two-way interaction between phase and block was also significant [$F(5,225) = 11.569$, $MS_e = 0.268$, $p < .001$, $\eta^2 = .205$].

Accuracy. The main effect of condition was significant [$F(2,45) = 7.757$, $MS_e = 0.080$, $p = .001$, $\eta^2 = .256$]; accuracy was greater for the match ($M = .545$) and partial-match ($M = .570$) conditions than for the no-match condition ($M = .491$). The main effect of phase was significant [$F(1,45) = 40.600$, $MS_e = 0.053$, $p < .001$, $\eta^2 = .474$]; accuracy was lower for training ($M = .492$) than for test ($M = .579$). The three-way interaction between phase, block, and condition was significant [$F(10,225) = 2.618$, $MS_e = 0.078$, $p = .005$, $\eta^2 = .104$]; there was a more consistent advantage for the match and partial-match conditions across blocks during test than during training (see Figure 6). The component two-way interaction between phase and condition was also significant [$F(2,45) = 4.036$, $MS_e = 0.053$, $p = .024$, $\eta^2 = .152$].

The main effect of block [$F(5,225) = 4.677$, $MS_e = 0.082$, $p < .001$, $\eta^2 = .094$] was also significant; accuracy generally improved across blocks (Block 1, $M = .455$; Block 2, $M = .546$; Block 3, $M = .524$; Block 4, $M = .584$; Block 5, $M = .535$; Block 6, $M = .567$).

Transfer

As was the case in Experiments 1 and 2, transfer was assessed by comparing Block 1 of training with Block 1 of test for each measure (see Table 1). The effects reported are again limited to those involving the transfer variable (Training Block 1, Test Block 1).

Number of clicks. There were no significant effects involving number of clicks.

RT/click. The main effect of transfer was significant [$F(1,45) = 139.558$, $MS_e = 0.285$, $p < .001$, $\eta^2 = .756$], with RT/click decreasing from Training Block 1 ($M = 2.478$ sec/click) to Test Block 1 ($M = 1.568$ sec/click).

Accuracy. The main effect of transfer was significant [$F(1,45) = 11.777$, $MS_e = 0.050$, $p = .001$, $\eta^2 = .207$], with accuracy increasing from Training Block 1 ($M = 0.400$) to Test Block 1 ($M = 0.511$). The amount of transfer depended on condition [$F(2,45) = 3.348$, $MS_e = 0.050$, $p = .044$, $\eta^2 = .130$], with transfer larger for the match and partial-match conditions (training, $M = .395$; test, $M = .555$) than for the no-match condition (training, $M = .408$; test, $M = .422$) [$F(1,46) = 4.444$, $MS_e = 0.052$, $p = .041$, $\eta^2 = .088$], but statistically equivalent for the match condition (training, $M = .408$; test, $M = .510$) and the partial-match condition (training, $M = .383$; test, $M = .601$) [$F(1,30) = 2.123$, $MS_e = 0.051$, $p = .156$, $\eta^2 = .066$]. Transfer was significant for both the match and partial-match conditions [match, $F(1,15) = 6.059$, $MS_e = 0.027$, $p = .026$, $\eta^2 = .288$; partial match, $F(1,15) = 10.261$, $MS_e = 0.074$, $p = .006$, $\eta^2 = .406$] but was not significant for the no-match condition ($F < 1$).

Discussion

Training led to performance improvement from training to test and across blocks within each phase. As was that of Experiment 2, the primary purpose of Experiment 3 was to examine the contribution of contextual information—in this case, memory for exemplar information—to transfer of training in the category search task. Specifically, we

Figure 5. Mean response time (RT; in seconds) per click as a function of block, phase, and condition in Experiment 3. Error bars represent ± 1 standard error of the mean.

Figure 6. Mean accuracy as a function of block, phase, and condition in Experiment 3. Error bars represent ± 1 standard error of the mean.

examined whether performance during test would be facilitated when distractors from training served as targets during test. We found evidence for transfer of training, with improvements from the first block of training to the first block of test on multiple measures of performance. However, in terms of accuracy, greater transfer was evident when exemplars were repeated across phases, thus suggesting that memory for exemplar information, not just category information, contributes to domain-specific transfer and that exemplar context, like category context, affects learning and transfer.

A second issue examined in Experiment 3 is whether reinstating exemplar context during test contributes to transfer. In this case, we found transfer at least as great when exemplars only partially overlapped from training to test (i.e., for the partial-match condition) as when the categories completely overlapped from training to test (i.e., for the match condition). Thus, the present experiment suggests that, whereas memory for exemplar information contributes to domain-specific transfer, the reinstatement of full exemplar context is not necessary, perhaps because of the small search set (four exemplars in contrast to the 12 categories used in Experiment 2). The difference in the size of the search set is likely to be responsible for the differing results for the contrast between the match and partial-match conditions in Experiments 2 and 3. In any event, the results of both Experiments 2 and 3 suggest that reinstating partial contextual information about categories and exemplars does facilitate transfer of training.

GENERAL DISCUSSION

This study was motivated by the widespread everyday necessity of search. Search may be involved as either a primary (e.g., fact finding) or a secondary (e.g., searching for information required to complete an action or achieve

a goal) component of a task. Previous research has indicated that search is often highly inefficient or error prone (e.g., Dumais & Landauer, 1984; Frankhuizen & Vries, 1980; Whalen & Latremouille, 1981). Using a limited number of categories (12 or 24) through which a search must be made, we also found that performance was far from optimal, with the participants accessing the main category list approximately twice more than necessary, on average, across the three experiments. Thus, in addition to identifying factors that might contribute to search inefficiency, we examined whether appropriate training can increase search efficiency. Across the three experiments, we found gains on various measures of performance across blocks within a phase, as well as across phases.

A second motivation for this study was the fact that category-based search tasks are often dynamic: Entirely new categories are sometimes added to category menus, and new items might be added to existing categories. Training is only truly successful to the extent that its effects withstand time and modifications to the task; training individuals on a task seems hardly worth the effort if minor changes to the task negate any positive effects yielded from training. The dynamic nature of the search task raises questions as to the effect of such changes on subsequent performance. For example, do changes in search domain, categories, and category exemplars negatively impact transfer performance? Or does previous experience with the task benefit performance, or transfer, regardless of such qualitative changes in the task? Specifically, in this study, we investigated whether training on a category search task transfers to subsequent task performance when changes are made to either the search domain or category organization (Experiment 1), to the set of categories (Experiment 2), or to the exemplars associated with each category (Experiment 3), even when different search targets are used during training and test.

Previous researchers investigating transfer have found that skills are often specific to the conditions of training, with changes in stimuli, responses, or procedures sometimes completely eliminating training effects (e.g., Bourne et al., 2005; Healy et al., 2005; Healy et al., 2006; Rickard et al., 1994). In the present study, we did find evidence for transfer. Specifically, in Experiment 1, we found that performance for one measure (number of clicks) tended to improve from training to test when category organization was held constant but tended to decrease when category organization changed, and we also found that transfer of training occurred for another measure (RT/click) when search domain was held constant. The first finding suggests a type of domain-general transfer; that is, even with unfamiliar categories and category exemplars, performance improved more when organization was consistent than when it was inconsistent across phases, presumably because of the development of a strategy that facilitates search through that particular type of organization after the participants discover the basis for category organization. For example, when categories are organized semantically, participants might use a strategy in which they search through clusters of categories rather than through each category individually, and when categories are organized alphabetically, the participants might use a strategy in which they first generate a candidate category, then focus their search on the area of the category list that contains categories that are alphabetically close to the candidate category. The finding that performance at test depends on organizational consistency has implications for the design of categories in a search environment (see Dumais & Landauer, 1984), in that search may be made easier simply by using a single abstract organization (e.g., semantic) across different search categories. In contrast, the second finding represents domain-specific transfer, whereby familiarity with domain-specific information, such as category names and exemplars, can improve performance. This result implies that experience with categories and/or category exemplars can improve performance, even when the search is for novel targets.

The purpose of Experiments 2 and 3 was to isolate the mechanisms of domain-specific transfer; in particular, we examined the roles of memory for context and contextual reinstatement (see, e.g., Wright, Shea, Li, & Whitacre, 1996, for a discussion of contextual reinstatement) during test for both categories and category exemplars in domain-specific transfer effects. Context was manipulated by holding constant, partially replacing, or entirely replacing categories (Experiment 2) or category exemplars (Experiment 3) from training to test. Not surprisingly, we found greatest improvement across phases when there was a match between the categories and exemplars presented during training and those presented at test. For categories, transfer was greatest for the measure of RT/click when there was a full match between training and test category sets, and there was less transfer when there was only partial overlap. However, there was still significant transfer when an entirely new set of categories was presented during test. For category exemplars, transfer occurred for the measure of accuracy when there was either

a full or a partial match between category exemplar sets, and there was no significant transfer when an entirely new set of category exemplars was presented during test. Thus, both experiments demonstrate the important role in transfer from training to test of memory for context and contextual reinstatement for both categories and exemplars, although the evidence for transfer depends on the specific measure of performance.

The last motivating factor of this study was to account theoretically for the transfer effects observed across experiments. We considered that transfer in our search task might be based on several different mechanisms, including (1) the learning of a general search strategy (which is a type of procedural learning), (2) category learning, (3) memory for categories, or (4) memory for category exemplars (the last three of which are types of declarative learning). First, we did find evidence for procedural learning of a general search strategy in Experiment 1, in that performance gains across phases were greater when category organization was the same across phases, regardless of whether the search domain was also held constant. This particular transfer effect represents the learning of a general search strategy because transfer, under these circumstances, cannot be due to category learning, memory for categories, or memory for category exemplars. Specifically, it seems plausible that the participants utilized organization information to guide their search. For instance, as was previously mentioned, when categories were organized semantically, the participants may have used this metacategory information to constrain their search by ruling out clusters of categories that appear to be semantically unrelated to the target and continuing on to the next cluster of categories. Second, evidence for declarative category learning was also provided in Experiment 1 by the finding of domain-specific transfer. We found that performance improved from training to test when the search domain and categories were held constant across phases, even though the participants were searching for entirely new exemplars. This finding suggests that the participants learned, to some extent, category boundaries within the search domain during training so that they were better able to categorize novel targets during test. Third, evidence for the role of declarative memory for both categories and category exemplars was provided by Experiments 2 and 3, where we found that greater transfer occurred when there was an overlap between categories and category exemplars from training to test. The results from this study are inconsistent with those from studies suggesting that training is highly specific to the learning situation (e.g., Healy et al., 2005; Healy et al., 2006), so that participants would learn only about how to classify targets that were encountered during training. In contrast, the results are consistent with theories that predict the greatest transfer when there is an overlap in the content between training and test (e.g., Singley & Anderson, 1989; Thorndike, 1906) and that greater transfer is observed when it is based on declarative information rather than on procedural information (e.g., Healy, 2007). Our study informs these theories, however, by suggesting that incidental memory for the training context also contributes to transfer effects.

In summary, in contrast to studies in which a severe specificity of training was found (e.g., Healy et al., 2006), the present study suggests that gains achieved during training can survive, to some extent, changes in the content of the task, perhaps because of the declarative rather than procedural nature of the task (see, e.g., Healy, 2007). This finding is promising, especially for those designing training programs, in that it suggests that training need not be inclusive of all possible test conditions in order for a benefit to be gained and that benefits occur even when the task changes dynamically across time. Pragmatic recommendations that follow from this study are, first, to hold category organization schemes constant when possible, since performance is best when category organization is held constant, regardless of domain or specific content, and, second, to hold category and exemplar information constant when possible. Although these conclusions might not apply to the training of all tasks, they should certainly apply to a wide variety of everyday search tasks that employ a hierarchical category structure like that in the present experiment, including, for example, searching through a telephone book, a grocery store, or a software menu system.

AUTHOR NOTE

This work was supported by Army Research Office Grant W911NF-05-1-0153 to the University of Colorado. We thank Alexandra Hindash for help with data collection, Carolyn Buck-Gengler for help with computer programming, and David Wright for helpful comments about this research. Correspondence concerning this article should be addressed to J. A. Kole, Department of Psychology and Neuroscience, Muenzinger Building, University of Colorado, Boulder, CO 80309-0345 (e-mail: james.kole@colorado.edu).

REFERENCES

- ALLEN, R. B. (1983). Cognitive factors in the use of menus and trees: An experiment. *IEEE Journal on Selected Areas in Communications*, *1*, 333-336.
- BOURNE, L. E., JR., HEALY, A. F., PAULI, P., PARKER, J. T., & BIRBAUMER, N. (2005). The influence of stimulus array on training of a speeded response. *American Journal of Psychology*, *118*, 385-411.
- DEL MISSIER, F., & TERPINI, C. (2009). Part-set cuing in option generation. *Memory & Cognition*, *37*, 265-276. doi:10.3758/MC.37.3.265
- DUMAIS, S. T., & LANDAUER, T. K. (1984). Describing categories of objects for menu retrieval systems. *Behavior Research Methods, Instruments, & Computers*, *16*, 242-248.
- FRANKHUIZEN, J. L., & VRINS, T. G. M. (1980). Human factors studies with viewdata. In *Proceedings of the Ninth International Symposium on Human Factors in Telecommunications* (pp. 1-7). Red Bank, NJ: Bell Labs.
- HEALY, A. F. (2007). Transfer: Specificity and generality. In H. L. Roediger III, Y. Dudai, & S. M. Fitzpatrick (Eds.), *Science of memory: Concepts* (pp. 271-275). New York: Oxford University Press.
- HEALY, A. F., WOHLDMANN, E. L., PARKER, J. T., & BOURNE, L. E., JR. (2005). Skill training, retention, and transfer: The effects of a concurrent secondary task. *Memory & Cognition*, *33*, 1457-1471.
- HEALY, A. F., WOHLDMANN, E. L., SUTTON, E. M., & BOURNE, L. E., JR. (2006). Specificity effects in training and transfer of speeded responses. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *32*, 534-546. doi:10.1037/0278-7393.32.3.534
- HOMA, D., & VOSBURGH, R. (1976). Category breadth and the abstraction of prototypical information. *Journal of Experimental Psychology: Human Learning & Memory*, *2*, 322-330. doi:10.1037/0278-7393.2.3.322
- IVANCIC, K., & HESKETH, B. (2000). Learning from errors in a driving simulation: Effects on driving skill and self-confidence. *Ergonomics*, *43*, 1966-1984. doi:10.1080/00140130050201427
- KOLE, J. A., & HEALY, A. F. (2007). Using prior knowledge to minimize interference when learning large amounts of information. *Memory & Cognition*, *35*, 124-137.
- LEE, Y., & VAKOCH, D. A. (1996). Transfer and retention of implicit and explicit learning. *British Journal of Psychology*, *87*, 637-651.
- MAXWELL, J. P., MASTERS, R. S. W., KERR, E., & WEEDON, E. (2001). The implicit benefit of learning without errors. *Quarterly Journal of Experimental Psychology*, *54A*, 1049-1068. doi:10.1080/02724980143000073
- MILLER, D. P. (1981). The depth/breadth trade-off in hierarchical computer menus. In *Proceedings of the Human Factors and Ergonomics Society 25th Annual Meeting* (pp. 296-300). Santa Monica, CA: Human Factors & Ergonomics Society.
- NEISSER, U., & BELLER, H. K. (1965). Searching through word lists. *British Journal of Psychology*, *56*, 349-358.
- PENNINGTON, N., NICOLICH, R., & RAHM, J. (1995). Transfer of training between cognitive subskills: Is knowledge use specific? *Cognitive Psychology*, *28*, 175-224. doi:10.1006/cogp.1995.1005
- POSTMAN, L., & UNDERWOOD, B. J. (1973). Critical issues in interference theory. *Memory & Cognition*, *1*, 19-40.
- RICKARD, T. C., HEALY, A. F., & BOURNE, L. E., JR. (1994). On the cognitive structure of basic arithmetic skills: Operation, order, and symbol transfer effects. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *20*, 1139-1153. doi:10.1037/0278-7393.20.5.1139
- SALMERON, L., CANAS, J. J., & FAJARDO, I. (2005). Are expert users always better searchers? Interaction of expertise and semantic grouping in hypertext search tasks. *Behaviour & Information Technology*, *24*, 471-475. doi:10.1080/0144329042000320018
- SCHNEIDER, W., & FISK, A. D. (1984). Automatic category search and its transfer. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *10*, 1-15. doi:10.1037/0278-7393.10.1.1
- SINGLEY, M. K., & ANDERSON, J. R. (1989). *The transfer of cognitive skill*. Cambridge, MA: Harvard University Press.
- SLAMECKA, N. J. (1968). An examination of trace storage in free recall. *Journal of Experimental Psychology*, *76*, 504-513. doi:10.1037/h0025695
- THORNDIKE, E. L. (1906). *The principles of teaching: Based on psychology*. New York: Seiler. doi:10.1037/11487-000
- VAN OVERSCHELDE, J. P., & HEALY, A. F. (2001). Learning of non-domain facts in high- and low-knowledge domains. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *27*, 1160-1171. doi:10.1037/0278-7393.27.5.1160
- WHALEN, T., & LATREMOUILLE, S. (1981). The effectiveness of a tree-structured index when the existence of information is uncertain. In D. Phillips (Ed.), *Telidon behavioral research II: The design of videotex tree indices* (pp. 3-12). Ottawa: Department of Communications.
- WRIGHT, D. L., SHEA, C. H., LI, Y., & WHITACRE, C. (1996). Contextual dependencies during perceptual-motor skill acquisition: Gone but not forgotten! *Memory*, *4*, 91-108. doi:10.1080/741940665

APPENDIX

Table A1
Categories and Category Exemplars for the Grocery Store Domain in Experiment 1

Category	List	Category Exemplars			
baby care	1	baby food	talcum powder	teething ring	<i>wet wipes</i>
	2	bibs	infant formula	<i>mild shampoo</i>	pacifiers
bakery	1	croissants	English muffins	<i>hamburger buns</i>	rolls
	2	cinnamon rolls	<i>doughnuts</i>	focaccia bread	raisin bread
baking	1	coconut	<i>chocolate</i>	powdered sugar	wheat flour
	2	<i>almonds</i>	corn starch	molasses	salt
beverages	1	apple cider	<i>milk</i>	orange juice	tomato juice
	2	bottled water	cranberry juice	fruit punch	<i>tea</i>
breakfast goods	1	<i>bagels</i>	eggs	jam	fruit pastries
	2	bacon	<i>coffee</i>	marmalade	sausage patties
canned goods	1	condensed milk	chicken soup	<i>peas</i>	pork and beans
	2	clams	fruit cocktail	<i>ravioli</i>	spam
cereal	1	bran flakes	cream of wheat	<i>granola</i>	oatmeal
	2	corn flakes	<i>puffed corn</i>	raisin bran	shredded wheat
cleaning	1	bleach	dishwasher detergent	<i>paper towels</i>	sponges
	2	brooms	mops	<i>soap</i>	window cleaner
condiments	1	<i>horseradish</i>	mustard	peanut butter	tabasco sauce
	2	honey	maple syrup	<i>olives</i>	relish
dairy	1	butter	half-and-half	<i>ice cream</i>	yogurt
	2	cottage cheese	cream cheese	sour cream	<i>whipped cream</i>
deli	1	cheddar cheese	jello salad	<i>pickles</i>	sandwiches
	2	bologna	egg salad	<i>ham</i>	meatballs
desserts	1	cakes	candy sprinkles	ice cream cones	<i>pies</i>
	2	cookies	<i>frozen custard</i>	fudge sauce	marshmallows
dry goods	1	<i>pasta sauce</i>	pinto beans	rice	spaghetti
	2	<i>crackers</i>	couscous	jello	macaroni
frozen foods	1	dumplings	<i>lasagna</i>	pie crust	waffles
	2	<i>blueberries</i>	juice concentrate	mixed vegetables	TV dinners
fruits	1	bananas	<i>caramel apples</i>	peaches	pineapple
	2	<i>mandarin oranges</i>	mango	papaya	plums
health foods	1	<i>natural toothpaste</i>	rice cakes	tofu	trail mix
	2	hummus	<i>organic lemonade</i>	soy cheese	vitamins
meats	1	<i>beef jerky</i>	hamburger	salmon	steak
	2	lamb	pork chops	<i>roasted chicken</i>	scallops
paper products	1	<i>cupcake liners</i>	facial tissue	napkins	wax paper
	2	<i>diapers</i>	notebooks	plastic wrap	stationery
personal hygiene	1	deodorant	razors	<i>toilet paper</i>	toothbrush
	2	aspirin	bandages	<i>breath mints</i>	shaving cream
picnic	1	charcoal	lighter fluid	ketchup	<i>watermelon</i>
	2	cole slaw	<i>paper plates</i>	potato chips	potato salad
snacks	1	candy bar	corn chips	<i>pudding</i>	pretzels
	2	mixed nuts	peanuts	raisins	<i>salsa</i>
soft drinks	1	coke	lemon-lime soda	<i>seltzer water</i>	root beer
	2	cherry cola	cream soda	ginger ale	<i>sports drink</i>
spices	1	basil	cayenne	cinnamon	<i>vanilla</i>
	2	dill	ginger	<i>minced garlic</i>	nutmeg
vegetables	1	carrots	cucumbers	mushrooms	<i>tomatoes</i>
	2	beets	onions	parsley	<i>salad dressing</i>
arts and crafts	1	glitter	stickers	<i>tongue depressor</i>	yarn
	2	beads	<i>buttons</i>	multicolored feathers	pipe cleaners
bath	1	hair dryer	<i>small rug</i>	toothbrush holder	towel
	2	scale	<i>scissors</i>	tissue dispenser	washcloth

APPENDIX (Continued)

Table A2
Categories and Category Exemplars for the Department Store Domain in Experiment 1

Category	List	Category Exemplars			
bedroom	1	nightlight	<i>nightstand</i>	pillow	quilt
	2	blanket	jewelry box	<i>mirror</i>	sheet
books	1	<i>greeting cards</i>	children's stories	magazine	romance novel
	2	books on tape	classic novels	<i>diary</i>	mysteries
electronics	1	cell phone	computer	<i>DVDs</i>	printer
	2	<i>alarm clock</i>	battery	palm pilot	speaker
furniture	1	bookshelf	cabinet	dresser	<i>kitchen table</i>
	2	card table	coffee table	desk	<i>lampshade</i>
gardening	1	mulch	seeds	<i>sunhat</i>	trowel
	2	<i>bug spray</i>	plants	potting soil	watering hose
gifts	1	chocolates	Easter basket	<i>perfume</i>	picture frame
	2	flower bouquet	candle	<i>lava lamp</i>	stuffed animal
kitchen/dining	1	<i>birthday candles</i>	candle holder	silverware	tablecloth
	2	<i>bleach</i>	blender	salt and pepper shakers	toaster oven
lighting	1	halogen lamp	lamp oil	<i>outdoor lantern</i>	table lamp
	2	candelabra	<i>flashlight</i>	light bulb	light stand
men's clothing	1	boxers	sweater	tie	<i>winter hat</i>
	2	flannel pants	suspenders	<i>t-shirt</i>	undershirt
office	1	pencils	pens	markers	<i>trash can</i>
	2	folders	<i>glue</i>	hole puncher	stapler
outdoor	1	hiking boots	sleeping bag	<i>sunblock</i>	tent
	2	compass	<i>matches</i>	water bottle	whistle
party supplies	1	balloons	<i>green lipstick</i>	invitations	streamers
	2	colored napkins	<i>crayons</i>	decorations	paper hat
patio	1	<i>chair</i>	doormat	garden ornaments	wicker furniture
	2	flamingo statue	hammock	outdoor swing	<i>sprinkler</i>
personal hygiene	1	hair ties	lotion	maskara	<i>soap dish</i>
	2	<i>aloe vera</i>	hair conditioner	soap	thermometer
sewing	1	fabric	needles	<i>string</i>	thread
	2	pins	<i>sequins</i>	sewing machine	thimble
shoes	1	dress shoes	high heels	<i>pantyhose</i>	sneakers
	2	cleats	clogs	<i>running socks</i>	sandals
sports	1	dart board	frisbee	football	<i>sports watch</i>
	2	bike helmet	dumbbells	<i>heavy-duty gloves</i>	tennis racket
storage	1	cardboard box	<i>clothes hamper</i>	plastic bin	storage box
	2	cookie tin	file cabinet	<i>food containers</i>	jars
swimming	1	<i>flip flops</i>	floats	goggles	swim cap
	2	beach ball	<i>beach towel</i>	ear plugs	water wings
toys	1	board games	<i>childs' baseball glove</i>	dolls	koosh balls
	2	model train	<i>puzzles</i>	remote control car	modeling clay
travel	1	<i>baby stroller</i>	backpack	money belt	suitcase
	2	<i>atlas</i>	luggage tag	map	toiletory kit
women's clothing	1	jeans	shorts	tank top	<i>wallet</i>
	2	<i>bathing suits</i>	khakis	scarf	shawl

Note—Separate lines for each category indicate different lists (List 1 is the top line and List 2 is the bottom line), and difficult items are italicized. For example, List 1 included the baby care items *baby food*, *talcum powder*, *teething ring*, and *wet wipes*, whereas List 2 included the baby care items *bibs*, *infant formula*, *mild shampoo*, and *pacifiers*.

(Manuscript received March 12, 2009;
revision accepted for publication July 9, 2009.)