RICE UNIVERSITY

Foraging Memory: Retrieving Words from One and From Two Semantic Categories

by

Daniel Shields Glaser

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

Master of Arts

APPROVED, THESIS COMMITTEE:

Jim Dannemiller, Lynette S. Autrey
Professor
Psychology

Michael Byrne, Associate Professor
Psychology

David Lane, Associate Professor
Psychology

HOUSTON, TEXAS

APRIL, 2008

ABSTRACT

Foraging Memory: Retrieving Words from One and From Two Semantic Categories

by

Daniel Shields Glaser

Traditionally applied to an animal's search for food, the concept of foraging has been extended to include the search for information in such places as the Internet and libraries (Pirolli & Card, 1999). The premise behind the research reported here is that memory searching can also be construed as foraging. The goals of this investigation are to uncover mental factors that may affect memory production during memory search and to use this knowledge to guide a prediction of foraging production. Prior to testing, four such mental factors were identified: a time cost when producing an initial item from a different category (switch cost); a production benefit driven by a release of proactive interference (time-out benefit); a production cost caused by the additional mental load of executing an autonomous switching strategy (executive-decision cost); and sub-optimal allocation of time between categories. Experiment 1 tested whether switching between categories leads to a switch cost and/or time-out benefit by having subjects produce items from a category in a continuous three-minute block or multiple blocks that add to three minutes. Experiment 2 addressed the possibility of an "Executive-decision" cost by either allowing subjects to autonomously switch between categories or yoking them to another subject's switch schedule. Experiment 3 tested whether memory foragers divide their time optimally between categories.

Data from the first experiment demonstrated that like external foraging, moving from category to category (patch to patch) results in a production downtime. These data also demonstrated that switching production between domains may lead to a time-out benefit. Experiment 2 showed that the execution of an autonomous switching strategy leads to less production then when switching is forced. The third experiment demonstrated that, unlike animals, humans do not have an innate sense of how to divide their time between patches (categories) to maximize gain.

Our prediction was derived by having subjects produce category exemplars from a single category alone or from two categories at once. Data from single-category production trials as well as adjustments inspired by Experiment 1 through 3 were used to predict production from two categories. Though accurate, the flexibility of our prediction is limited. Research needed to allow for greater flexibility is discussed.

Acknowledgements

I would like to acknowledge all of those who assisted in the completion of my thesis. I give thanks to Drs. James Dannemiller, Michael Byrne, and David Lane for their time an advice offered as my committee members and special thanks to my former advisor Mike Watkins for his role in the development of my research.

TABLE OF CONTENTS

Foraging Memory: Producing Words from One and From Two Semantic Categories

The ability to bring information to mind selectively is surely a hallmark of being human. Such willful remembering is no small feat, given the potential number of memories available. Since the cognitive revolution swept the study of the human mind some half century ago, researchers have conceptualized memory as being organized into more or less specific domains (Tulving & Donaldson, 1972). Of interest here is memory searching involving multiple domains. If, to take a simple example, a given amount of time is allowed for the production of exemplars of two distinct categories, such as *professions* and *tools*, how many items are produced relative to the case when the items are retrieved solely from one category or the other? This is the question addressed in this thesis.

The question is framed in terms of efficiency, and more specifically with reference to Stephens's and Krebs's (1986) optimal foraging theory. This theory addressed how animals should divide their time between discrete food patches to maximize their rate of nourishment. Pirolli and Card (1999) have used the theory to ask whether humans optimally forage information from websites, books, articles, and so on. The research described in this thesis applies optimal foraging theory, and variants thereof, to the "searching" of human memory.

Basic Tenants of Optimal Foraging Theory

Stephens's and Krebs's (1986) optimal foraging theory (OFT) constitutes a model of a forager's ideal division of its search time into time searching for or locating patches (T_B)[1] and time exploiting patches (T_W)[2]. The average rate of gain (R) will equal the total

[1] Total time traveling to P patches

gain (G)[3] divided by the total time ($T_B + T_W$):

The gain function for an average patch, $g(t_w)$, is defined by the amount of gain (g) extracted after any given time spent exploiting a given patch (t_w). For example, a forager may extract 10 units of gain by 1 unit of time and 15 units of gain by 2 units of time. If the potential gain from an average patch is assumed to be finite, then, given patch depletion, the rate of gain will decline over time. In other words, a $g(t_w)$ function will monotonically increase but tend to do so at an ever slowing rate, eventually reaching an asymptote corresponding to exhaustion of the patch.

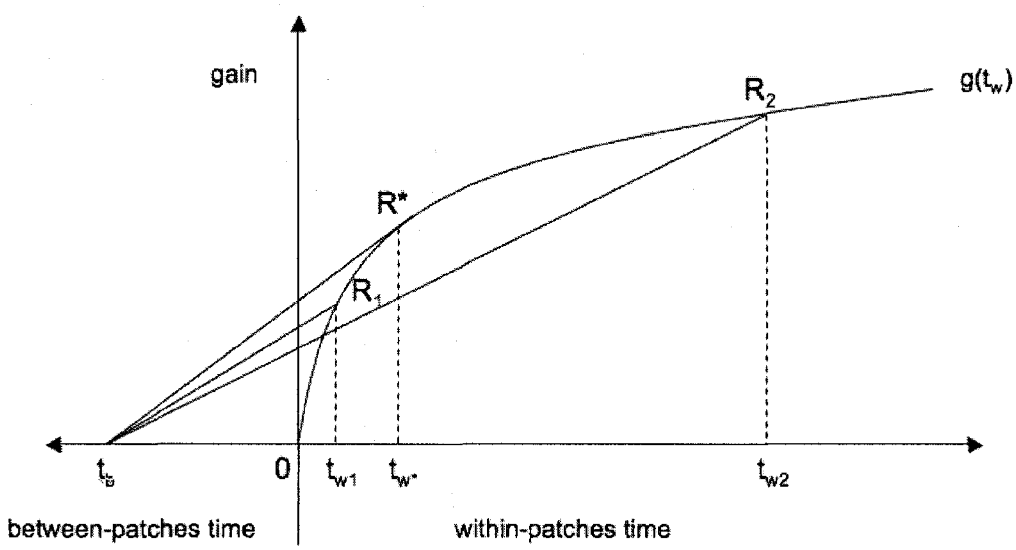


Figure 1. Maximizing rate of gain: the average slope (R) is maximized when the t_w corresponds to the point at which the line extending from t_b to $g(t_w)$ is tangential to $g(t_w)$.

---

[2] Total time spent within P patches
[3] Total gain from P patches
[4] $g(t_w)$ denotes the average gain function from a series of patches

Figure 1 shows an idealized gain function, with gain as the ordinate and time as the abscissa. The gain function begins at the t_w origin. Left of the origin is t_b, the time it took to access the patch. The fundamental problem for the forager is, for a given t_b and a given $g(t_w)$, what is the optimal time, t_w, to spend in the patch before searching for another patch? Shown are three hypothetical departure times: t_{w1}, t_{w2}, and t_w*. As is clear from Figure 1, the optimal switch time occurs when the slope of the line extending from t_b is tangential to $g(t_w)$, for it is at just this point that the average rate of gain over the total foraging time (t_b+t_w) is maximized. R_1 and R_2 represent sub-optimal slopes, which occur if too little or too much time (t_{w1} and t_{w2}, respectively) is spent exploiting the patch. For a given gain function, the optimal within-patch time, t_w*, varies with the time to find a patch, t_b, such that the shorter is t_b, the shorter will be t_w* (Figure 2). Also, generally speaking, the richer the average patch, the shorter will be t_w* (Figure 3).

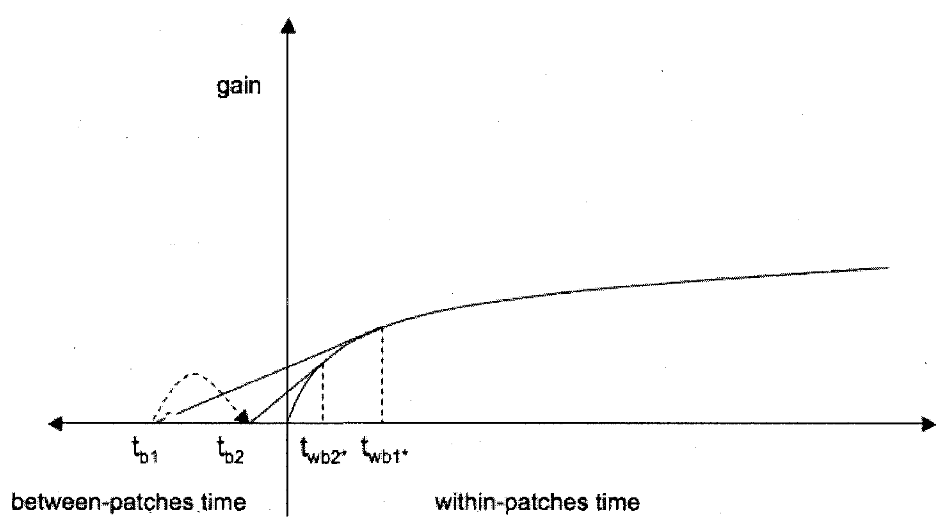


Figure 2. Between-patch enrichment: As t_b shortens, ($t_{b1} \rightarrow t_{b2}$) so does the optimal time within $g(t_w)$ ($t_{wb1*} \rightarrow t_{wb2*}$). Thus increasing the number of patches in a given area will

lead to a shorter t_b.

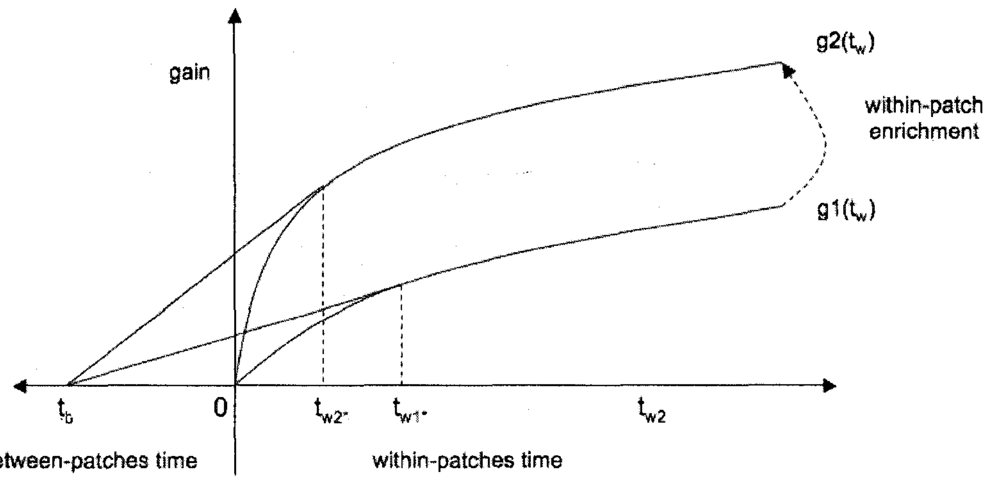


Figure 3. Within-patch enrichment: As a patch becomes more enriched, ($g1(t_w) \rightarrow g2(t_w)$) the optimal time within the patch ($g(t_w*)$) will generally shorten ($t_{wb1+}* \rightarrow t_{wb2}*$).

Unlike OFT, our prediction does not utilize parameterized gain functions nor does it calculate R or t_w. Because memory foraging is merely an analogy to ecological foraging is remains unclear whether it meets the underlying assumptions of OFT. In this investigation, we aim to develop a method of predicting the precise number of memory items foragers produce from two semantic categories (e.g., *professions* and *tools*) at each instant of a three-minute period using a unique non-parametric approach.

Human Memory Foraging

The prediction of memory production from two semantic patches is not unprecedented. Although they did not adopt a foraging perspective, Maylor, Charter, & Jones (2001) used a dual-category paradigm to observe whether searching for instances from two semantic domains is better characterized as parallel or serial. Unexpectedly, they found searching to be less efficient than predicted not only by the parallel model but also by the serial model, suggesting that only one category can be accessed at a time, with an additional cost, which the authors attribute to "time costs associated with switching between categories" (p. 1193).

The notion of switch time is, of course, equivalent to the between-patch time (t_b) of optimal foraging theory. It is entirely possible, however, that although optimal foraging theory adequately captures the activities of a bird searching for worms, it will not capture the complexities of a human searching memory. There are at least four potential factors that might need to be taken into account in characterizing memory searching: (i) a possible beneficial effect from category time-outs, (ii) a cost of switching between categories, (iii) "executive-decision" costs of dual-category searching, and (iv) suboptimal allocation of time.

Are the time-outs in dual-domain production beneficial?

The notion of a potential beneficial effect of dual-domain production has no counterpart in optimal foraging theory, but it could be incorporated by assuming two patches with re-generation powers. During memory search however, a facilitative effect of domain switching is suggested by a vast body of research that is usually referred to as demonstrating "release from proactive interference".

In a now classic meta-analysis of part of the proactive interference literature, Underwood (1957) showed that recall of syllable lists decreases sharply as a function of the number of prior lists learned. For example, Greenberg and Underwood (1950) found that when on each of four successive days a list of nonsense syllables was memorized and recalled on the next day, performance declined steadily from 69% to 25%. Such proactive interference is doubtless domain-specific, for it is unlikely in the extreme that a parallel decline occurred for memory for breakfast or the weather. Perhaps the best known findings for a release from proactive interference come from research with the Brown-Peterson paradigm in which recall of a sub-span set of items (typically word or alphabetic trigrams) decline drastically over the course of several seconds of distracter activity (such as counting backwards). Keppel and Underwood (1962) showed that this decline did not occur for the first trial, but built up only after the first two or three trials. Wickens (1970, 1972), in an extensive series of experiments with his collaborators, showed that this built-up proactive interference would partially dissipate or release if the type of item was changed, as from digits to letters or more germane to present purposes, from one semantic category to another.

Also relevant is evidence from the part-set cuing paradigm. Here subjects study a list of randomly selected words or, more commonly, of words selected from each of several distinct semantic categories, and then try to recall the words. The key finding is that providing a portion of the studied words at the time of test, in the guise of cues for recall of the remainder, actually depresses recall of the remainder (Slamecka, 1968). One way to conceptualize this finding is to assume that the nominal cues are a source of proactive interference for the remaining (to-be-recalled) items. Importantly, domain specificity has been demonstrated in this paradigm, in that the negative effect of part-set cuing is eliminated if the cues are from a different category from the items being cued (Mueller and Watkins, 1977).

Proactive interference is constrained not only by domain, but also by time. The proactive interference demonstrated in laboratory experiments will obviously not persist for a lifetime. Indeed, Loess and Waugh (1967) found proactive interference in the Brown-Peterson paradigm to diminish systematically as the inter-trial interval was increased to 2 minutes, with an appreciable reduction with an interval as short as 15 seconds (the shortest interval they investigated). Such evidence suggests the possibility that continually producing items from a single category gives rise to proactive interference. If so, such interference could be curtailed by switching between categories.

The evidence for proactive interference discussed to this point has all involved a study or learning phase as well as a recall phase. Perhaps, then, the locus of the effect is entirely at the study phase rather than at the time of recall. If proactive interference is not localized at the recall phase of the memory process, then switching between categories in the research to be reported here would not facilitate performance, for no study phase is involved.

The temporal locus of the release from proactive interference that comes from category switching in the Brown-Peterson paradigm has been investigated by Gardiner, Craik, and Birtwistle (1972). Their subjects were given four successive trials in which they tried to remember, for example, the names of three sports through a distracter task. In one trial block, items for the first three trials were indoor sports and those for the fourth trial were outdoor sports. Subjects who were never informed of this fact showed no release on the fourth trial, whereas those who were informed showed release regardless of whether they were informed of the subcategory switch at the study phase or only at the test phase. The implication is that a category switch prior to recall can reduce proactive interference.

Additional evidence that release from proactive interference can occur during the production, or recall, stage of the memory process comes from John Brown (1958). He introduced what can be thought of as a part-set cuing procedure without the study phase. Some subjects tried to recall as many of the 50 states of the USA as they could, whereas others were provided with 25 and tried to recall the remainder. These latter subjects were less successful in recalling the remaining 25 states than were those who were given no cues. Apparently, the states provided as cues proactively interfered with recall of the remainder under conditions in which there was no study, only production.

Even more relevant to the present concern is evidence from a procedure introduced by Alan Brown (1981). He showed that the speed of exemplar production was reduced by prior item productions from the same category even though the experimenter controlled the order of item recall by supplying the initial letter of the exemplar ("Fruit-A", "Fruit-P", "Fruit-M", etc). Note that the inverse relation between speed of response and (within-category) item order cannot be dismissed as an artifact of frequency of usage or of some other correlate of accessibility. A variation of Brown's finding has been replicated many times (Gunter, Clifford, & Berry, 1981; Loess & Waugh, 1967; Gardiner, Craik, & Birtwistle, 1972), giving ample evidence that continually producing items from the same domain slows, or makes less likely, retrieval of additional items of the same domain, but that switching domain will result in a time-out from that domain which could release, or ease, the accumulated proactive interference.

Is There A Cost To Switching Between Memory Domains?

Animals foraging in the "real world" take a finite time to switch between patches. Memory patches or domains, unlike worm or berry patches, are not separated in physical space. The notion of travel between memory patches is merely metaphorical, and there is no logical necessity of a switch cost. Whether there is a switch cost is an empirical question. The principal paradigm that cognitive psychologists have used for investigating the effects of task switching involves comparing performance on a trial of a particular task when that task directly follows a trial on either that same task or a different task. The dominant finding is that performance is slower when the task is changed from the preceding trial (Monsell, 2003); this slowing of performance has been assigned the label switch-cost. Some have argued that this cost is the result of a task-set inertia caused by memories of past tasks passively interfering with the interpretation and execution of the current task (Allport, Styles, & Hsieh, 1994). Others have asserted that a switch cost requires an effortful process where parameters of a new task set must be accessed from working memory and applied to the new task (Logan & Gordon, 2001; Evans, Meyer, & Rubenstein, 2001). The aim of the research to be reported here is not to distinguish between these possible origins but rather to determine whether such a cost is a factor during memory foraging. Specifically, we ask whether it take longer to produce a memory item when it is from a different domain than the previously produced item.

That there is a discernible cost to switching among categories is suggested by a study by Rohrer, Pashler, and Etchegaray (1998), who had subjects learn two categorical lists (such as *furniture items* and *countries*), and then recall the exemplars of either just one category (*table, chair, sofa, ...)* or of both categories in alternating category order (*chair, Mexico, table, England, ...*). The time required to produce items from the alternating runs took longer than items from the single category runs. It is unclear whether this conclusion will generalize to the research to be reported here, for the task demands differ from those of memory foraging in a couple ways. That a switch cost may be even greater than that reported by Rohrer et al. (1998) comes from Arlington and Logan (2004) who found that switching takes even longer when subjects are not explicitly cued when to switch. During memory foraging a switching strategy is executed autonomously (without cues) which may lead to a more exaggerated switch cost than was reported by Rohrer et al. (1998). Evidence that there may be no switch cost comes from the finding that preparation for a switch will attenuate or even eliminate switch cost (Dreisbach, Haider, and Kluwe, 2002; Altmann, 2004). A memory forager can prepare for a switch at any point because she controls her own switch schedule. As a result, switch cost may be greatly attenuated during memory foraging.

Does dual-domain production incur an additional executive cost?

A third factor that may need to be considered in an adequate account of memory search is a possible "executive-decision" cost, or cost of making decisions regarding the searching schedule. Put another way, the very execution of a switching strategy could impair within-domain item production. An executive-decision cost would violate the optimal foraging theory assumption that foragers will produce items from a specific patch at the same speed when there are multiple patches for the forager to choose from and when there is only the one specific patch. Arlington and Logan (2004) found that switch-cost in an autonomous switching paradigm is greater than when switches are explicitly cued. The possibility remains that the cost of deciding when to switch is ever-present, affecting both switch time and production after a switch. If so, the rate that items are produced from a category A may not be a true reflection of the rate that items are produced from category A when coupled with category B.

Is Time Allocation Sub-optimal?

Beyond switch-cost, release from proactive interference, and executive-decision costs, which all potentially affect within category production, there remains the possibility that human's do not optimally distribute their time between categories. Suboptimal time distribution would not affect the rate of item extraction from an individual patch but would affect the total number of words produced.

In a dual-domain foraging scenario a forager must choose how to allocate time between the more accessible (high) and less accessible (low) categories. If a forager does not differentiate the accessibility of the two categories, then he will likely spend an equal amount of time on the high and low categories for any time T. This particular foraging behavior leads to baseline word production. A strategic forager spends a greater proportion of time on the more accessible category than on the less accessible category, which would lead to greater performance at time T than if there were no differentiation between patches. For any time T there is an optimal amount of time to spend on the more accessible category (t_h*). The closer a forager is to meeting t_h* the closer performance will be to optimal. Experiment 3 will calculate the proportion of optimality foragers demonstrate when producing items from two categories. Baseline (0% optimal) performance occurs when an equal amount of time is spent on the more (high) and less (low) accessible categories. Optimal (100% optimal) performance occurs when a forager spends the correct amount of time on the more accessible category, which leads to the greatest production. For example, if the optimal amount of time (Optimal Division) to spend on the more accessible category after 180 seconds is 120 seconds, but a forager only spends 110 seconds (Actual Division) then foraging would be 67%[4] optimal. If the forager were to spend too much time on the greater category, for instance 130 seconds, then foraging optimality would also be 67%. This measure indicates how close optimal allocation is relative to baseline. Figure 4 illustrates an optimal and baseline prediction for a representative high and low accessible patch group at time T.

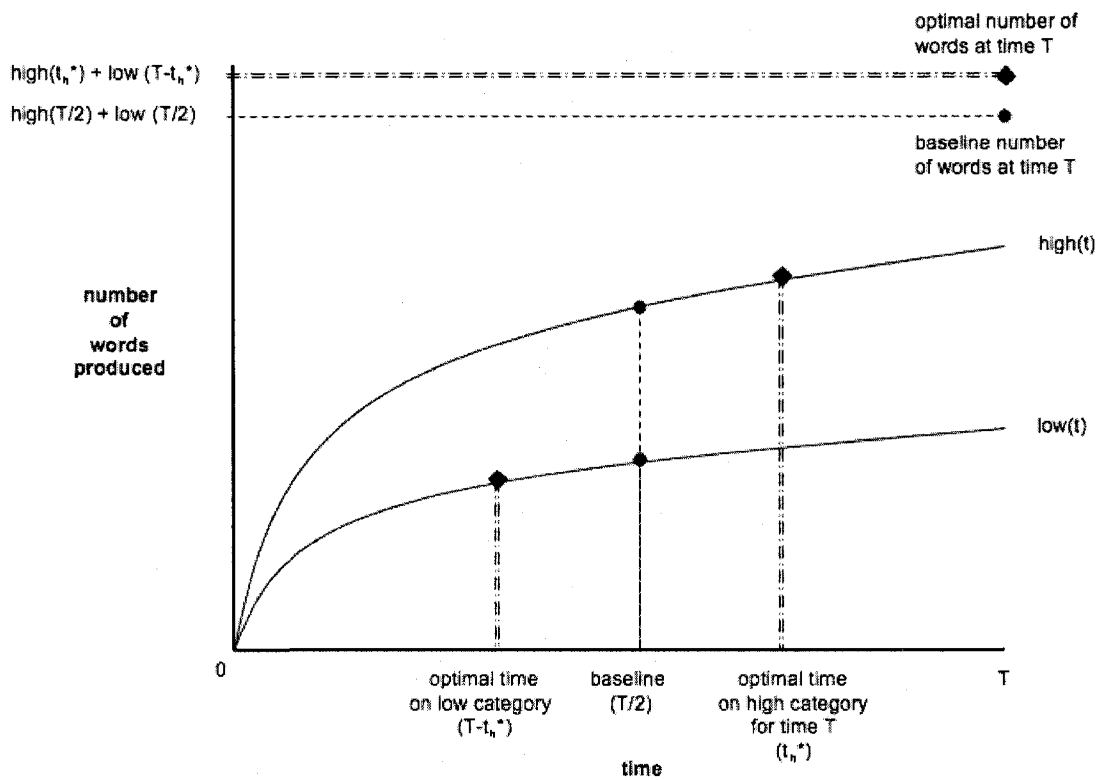


Figure 4. Illustrates baseline and optimal time distribution between a high and low single category gain functions and how the two time distributions may affect word production.

Outline of Experiments:

The following is a description of a series of four experiments designed to explore

---

[4] $$Optimality\_of\_Actual\_Division = 1 - \frac{|Optimal\_Division - Acual\_Division|}{Optimal\_Division - Baseline\_Division}$$

the memory foraging metaphor. Experiment 1 seeks evidence both for a switching cost and for release from proactive interference. Experiment 2 seeks evidence for an executive decision cost – that is, a cost of having to decide when to switch categories. Experiment 3 explores how effectively time is allocated between the two categories. At the end is the evaluation of our memory foraging prediction developed on the basis of the findings of the first three experiments. In the following three experiments, time spent typing exemplars was not considered in the analysis. This time was occluded because during actual memory, typing a memory item is not a necessary step when producing a memory item (bringing a memory item to mind).

Experiment 1

Experiment 1 was designed to test for a potential transitory cost (the switch cost) from a potential less transient benefit of changing categories. To this end, subjects produced exemplars from various categories for a total of 3 minutes with the category switching and number of switches per category controlled by the experimenter.

Method

Subjects. The subjects were 60 Rice University undergraduates of whom 39 were females.

Materials. The design utilized 12 semantic categories, and these were selected from the Toronto categorized word pool (Murdock, 1976). For each category, 48 exemplars were designated acceptable. These were the 48 most common responses of 100 students drawn from the same pool and working under similar conditions and time constraints as the subjects of this experiment. The categories and exemplars are listed in Appendix F.

Design. Counterbalancing aside, the only independent variable was number of response intervals into which the 3-minute (non-typing time) response time allowed for a given category was divided. There were three levels of this variable: a single interval of 180 seconds (the no-switch condition), 4 intervals of 45 seconds (the low-switch conditions), and 16 intervals of 11.25 seconds (the high-switch condition). This variable was manipulated within subjects. Thus, for any given subject, four categories (e.g., *tools, fruit, birds,* and *professions*) were assigned to the no-switch condition, four (e.g., *cars, colors, fish,* and *sports*) to the low-switch condition, and four (e.g., *vegetables, clothing items, 4-footed animals,* and *musical instruments*) to the high-switch condition. In all, there were 84 exemplar-production intervals [(4*1) + (4*4) + (4*16)]. The trial order was separately randomized for each subject within the constraint that adjacent intervals were assigned different categories. Finally, assignment of categories to the switch conditions was counterbalanced between subjects, with each category being assigned to the no-switch condition for one-third of the subjects, to the low-switch condition for another third, and to the high-switch condition for the remaining third.

Procedure. After instructions and four practice (two low-switch and two high-switch) trials, the subjects were told use the mouse to press the "Start" button displayed on the screen. This initiated an uninterrupted sequence of 84 exemplar production intervals of variable duration. For each interval, a category label was displayed in the center of the screen, directly above the response field. A running total of the number of retrieved exemplars (i.e., the score) for the category was displayed at the bottom of the screen, with the produced exemplars listed on the left side of the screen. Subjects typed each response into a text box and pressed the "return" key. If the exemplar matched one of the 48 pre-selected exemplars of the category, the category score was incremented by one, the exemplar was added to the response list, and the text box was cleared. Between intervals, the word "CHANGE" was displayed in the center of the screen for 333 ms. Except when a category was occurring for the first time, the already accumulated score and previous responses for the category were represented.

Results

The cumulative (gain) functions for the no switch, low switch, and high switch conditions are shown in Figure 5

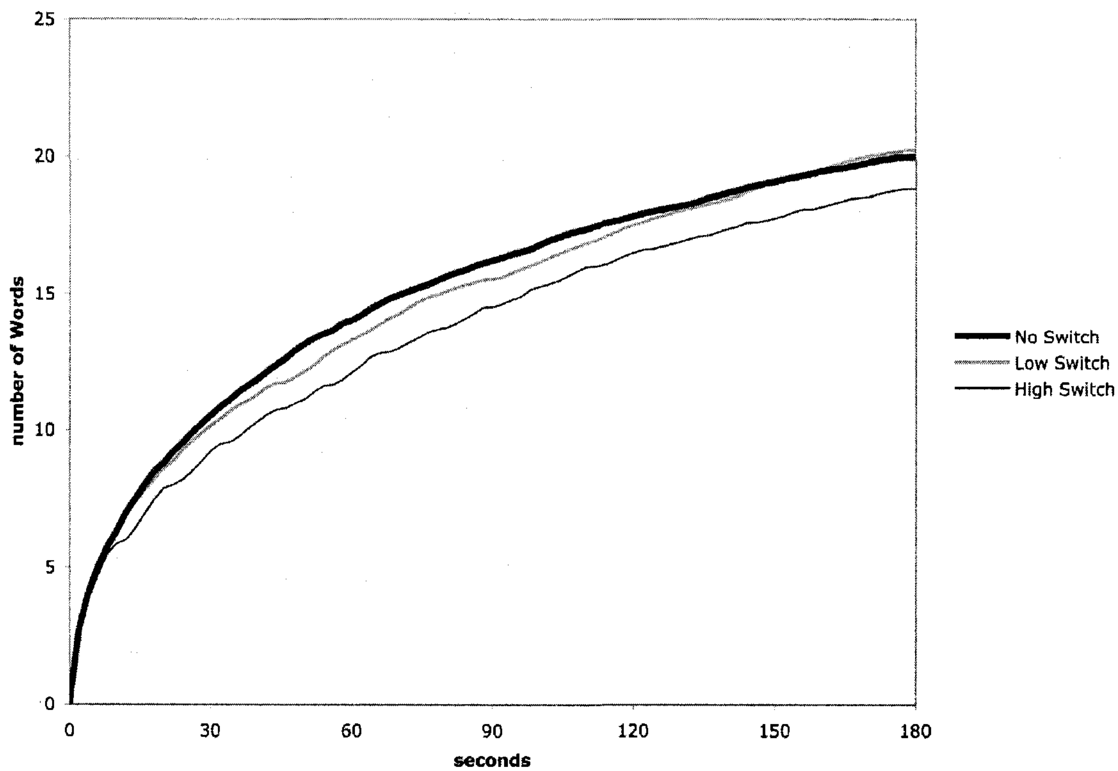


Figure 5. No Switch, Low Switch, and High Switch Functions: the summed effect of switch cost and time-out benefit for both the low and high switch conditions

The dependent variable is the sum of every ordinate at the conclusion of each the

720 intervals (quarter-seconds). This measure is an estimate of the area under a gain function. This nonparametric approach was used in lieu of curve fitting because of our interest in the different epics of each gain function not just the area. Of particular interest were the potential differential effects that switch cost and a release from proactive interference may have on the precise shape of the gain functions. Parameterizing these gain functions would obscure these shaping effects.

For purposes of statistical inference, categories (rather than subjects) served as the random variable. There was a main effect of switch condition that was unlikely to have arisen by chance, $F(2, 22) = 13.57$, $p < .001$. More particularly, the area under the high-switch gain function was smaller than that for both the no-switch condition, $t(11) = 6.29$, $p < .001$, and the low-switch condition, $t(11) = 4.10$, $p < .002$. No meaningful difference in area was observed between the no-switch and low-switch condition, $t(11) = .56$, $p = .59$.

Calculating Switch Cost. To determine the switch cost, the average time in seconds needed to make an acceptable response after a switch is compared to the time needed to make an acceptable response after the same point in time but in the absence of a switch. The relevant data are summarized in Figure 6. Overall, the mean response time was longer following a switch by 1.81 seconds than that measured from the same instant but with no switch correct response with no switch $t(11) = 20.691$, $p=.000$. No reliable switch cost was found in the low switch condition ($M = 1.02$), $t(11) = 1.825$, $p = .56$; however there was a reliable switch cost in the high switch condition ($M = 1.98$), $t(11) = 15.043$, $p < .001$. There was no discernable difference in switch cost between the high and low switch conditions $t(11) = .532$ $p = .606$. The weighted average of all switch costs combined from both the low and high switch condition was 1.81 seconds with the high switch condition having 4 times the number of switches than the low switch condition. No significant difference in switch times was observed between the high and low switch conditions. Figure 6 illustrates these 3 switch-costs.

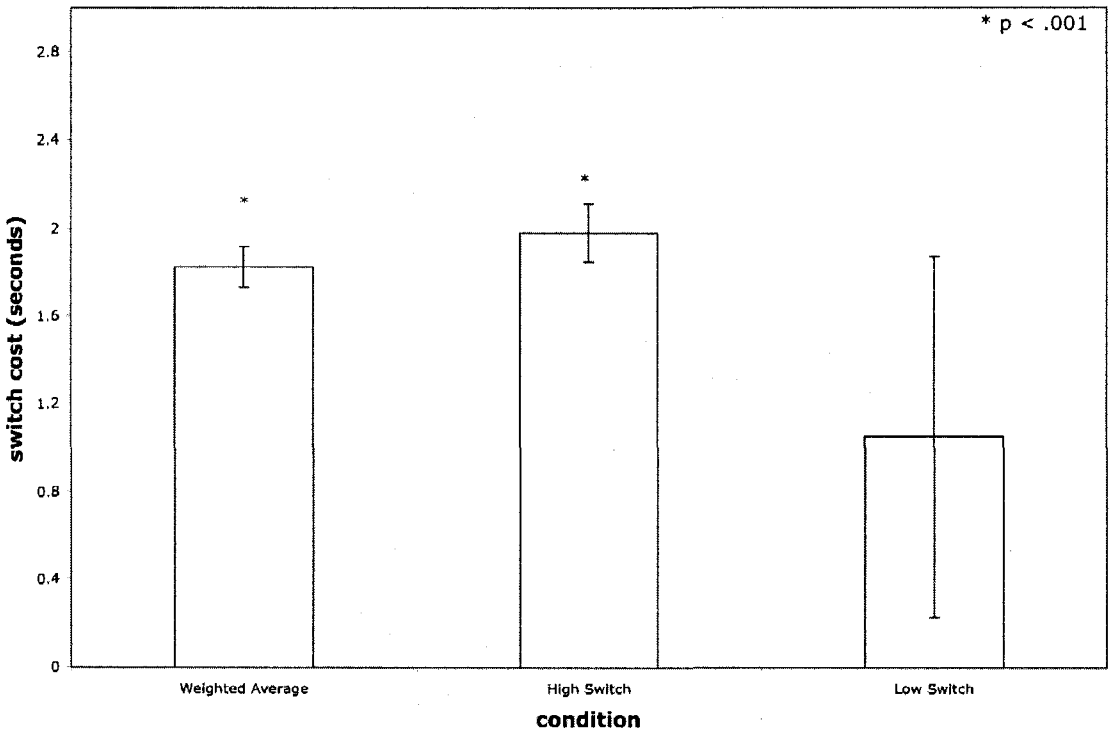


Figure 6. Average time cost for all switches, and for switches in the high-switch and low-switch conditions. The weighted average is closer to the high-switch mean than to the low switch mean because those were five times as many switches in the high-switch condition than in the low-switch condition. An asterisk indicates that a mean is greater than zero. None of the means are significantly different than the other.

Calculating the Time-out Benefit. Strictly speaking, the switch cost as computed in the preceding section is the net effect of an actual switch cost and any possible concomitant benefit to exemplar production starting or returning to a category afresh – as would occur, for example, with any reduction in "proactive interference." To assess any benefit of having just spent time on another category compared to the same category, the switch cost needs to be taken into account. To this end, the duration following each critical instant (i.e., the time at which a switch occurred in the case of the switch condition and the same point in elapsed time in the case of the no-switch condition) to the first acceptable response is discounted. For example, suppose that in one of the switch conditions a single subject's first acceptable response following the critical elapsed time of 45 seconds is made at 47 seconds. The 2 seconds that it took to make the response are removed from the data in constructing an abridged gain function. If no acceptable response within the 11.25 or 45-second was within the response interval, then the entire interval was removed. This correcting procedure was performed for each response interval of each of the 12 categories for each of the 40 subjects. Because each subject's gain functions were unique, so was the extent of its truncation. Also, because all conditions are identical up to the first switch, the data prior to the first switch were omitted in constructing the abridged functions. To preclude potential selection artifacts, area comparisons were restricted to durations that included at least half the subjects' data in both the no-switch and switch condition.

Overall, the switch conditions abridged functions ($M = 658.8$ word*quarter seconds) yielded more area than the no-switch condition abridged function ($M = 568.2$ word*quarter seconds), indicating that production after a switch is greater than after a critical instant $t(15) = 2.84$, $p = .016$. See Figure 7 for an illustration of the production advantage from switching.

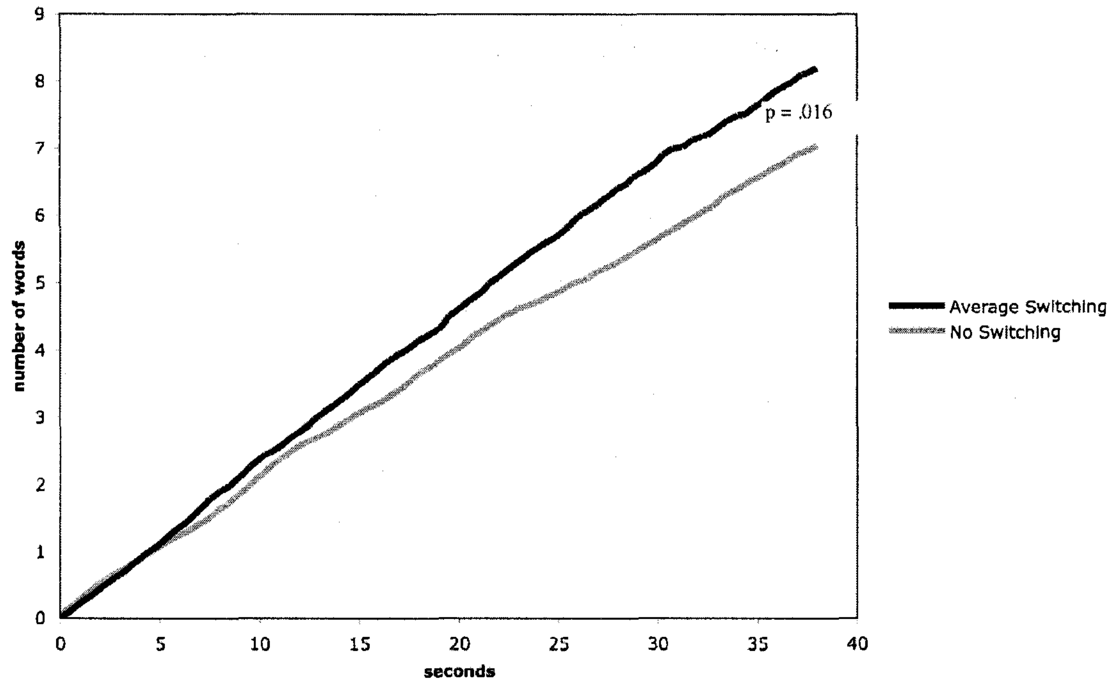


Figure 7. Average and No-Switch Corrected Functions: the area difference between the two lines illustrates the positive production effect of switching.

Specifically, each switch across both conditions led to a production advantage of 1.39%. A more detailed analysis of these data demonstrated that the total time-out benefit in the high-switch condition is more exaggerated than in the low-switch condition. This result is perhaps not surprising considering that there were also five times as many time-out periods in the high-switch condition. The production benefit per switch between the high and low–switch conditions were comparable, with each switch in the high-switch condition leading to a 1.38% production advantage and a 1.43% advantage in the low switch condition $t(11) = .21$ $p < .89$. See Figure 8 and 9 for an illustration of high-switch and low-switch conditions production advantage.

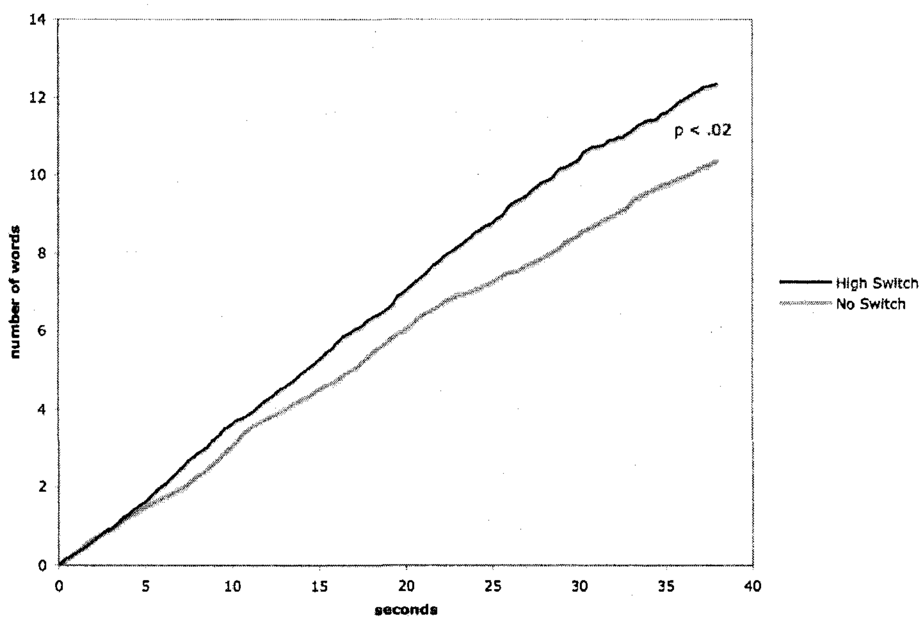


Figure 8. High and No-Switch Abridged Functions: the area difference between the two lines illustrates the cumulative positive production effect of 15 timeouts in the High-Switch condition

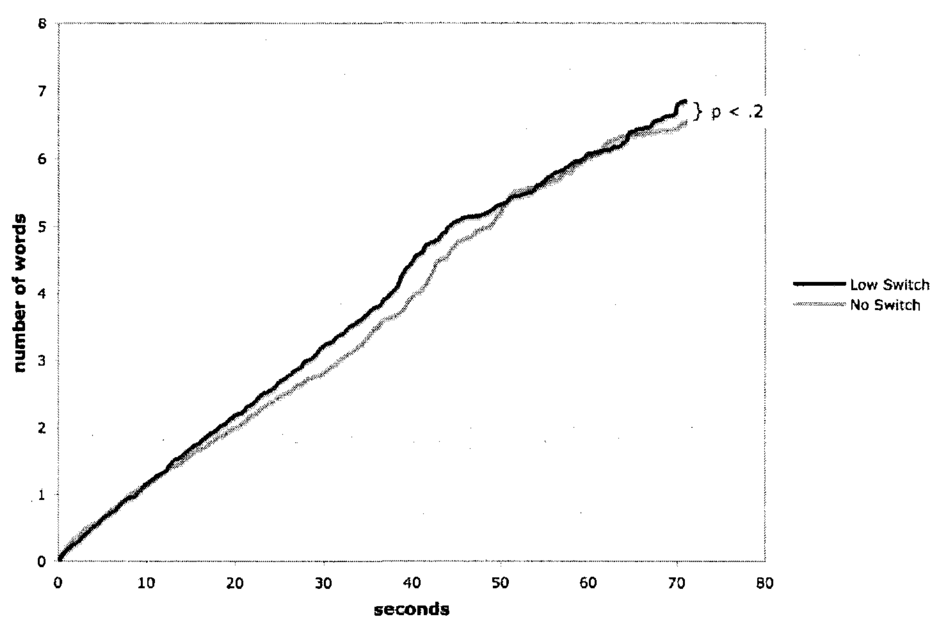


Figure 9. Low and No-Switch Abridged Functions: the area difference between the two

lines illustrates the positive production effect of a time out in the Low Switch condition.

Discussion

This first experiment has demonstrated both a cost and a benefit of switching between memory domains. This cost was, on average, an additional 1.8 seconds to produce an acceptable response following a category switch. This cost can be seen as analogous to that of an animal's travel time from one food patch to the next. The benefit is manifest only after the first response following a switch. Specifically, productivity between this first response and the next switch was 1.39% higher than if no switch had occurred. This time-out benefit, as well as the switch cost, will be utilized when predicting word production from dual domains. (See Appendix B)

Experiment 2

Experiment 2 addresses the potential of an executive-decision cost by using a yoked design. In one condition subjects were given two categories and were told produce as many exemplars as possible using the two categories (implies switching at will). The switch times for these subjects were recorded. These switch times (e.g., 12 seconds...40 seconds...64 seconds) for the executive subject were then forced upon the yoked subject. If there is an executive-decision cost when foraging memory, the single patch gain functions for the executive group will be lower than the yoked subjects who are explicitly cued when to switch and, in turn, do not have to bear the potential cost of executing a switching strategy.

Method

Subjects. Sixty Rice University undergraduates participated in the study.

Materials. Materials are as in Experiment 1.

Design. An experiment session consists of six double-category trials, with each trial extending for 3 minutes of non-typing time. For three of the trials, the subject decided when to switch between the two categories. For the other three, the switching was in accordance with the previous subject's switching schedule. To fully implement this scheme, the first subject was tested in two sessions. In the first session he was given three free-switching trials. In the second session, which was at the very end of the experiment, he was given the remaining three forced switching trials. The ordering of the free and forced switching conditions was counterbalanced such that for half of the subjects the executive trials precede the forced switching trials and for the other half of the subjects the yoked trials precede the executive trials. Figure 10 illustrates the design.

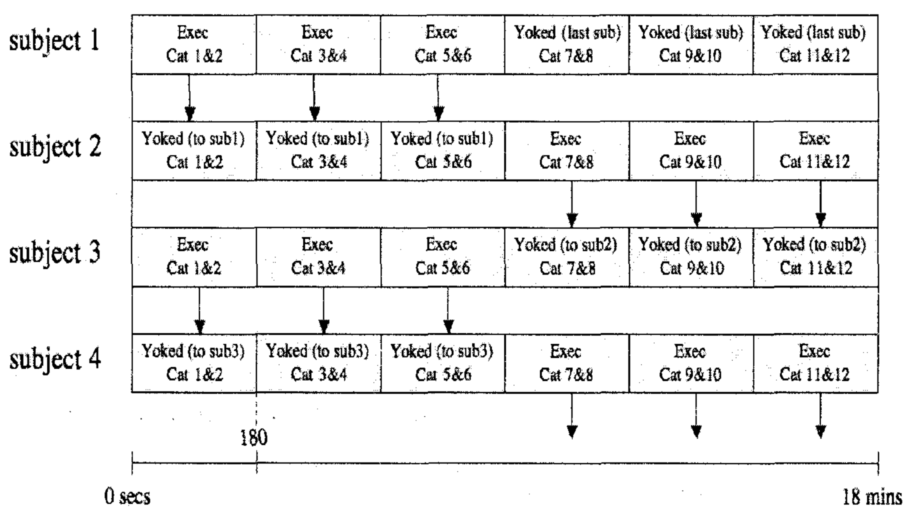


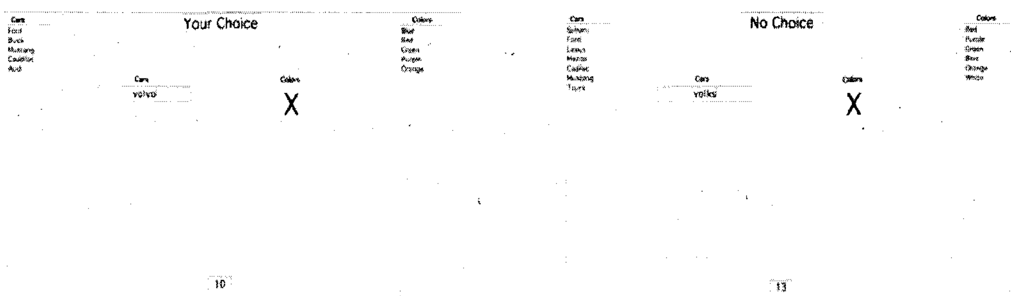
Figure 10. Illustrates the design of Experiment 2. Specifically, it demonstrates how subject n's Yoked trials are controlled by subject n-1's Executive switch schedules and how subject n's Executive trials control (→) subject's n+1's Yoked trials. There was one

exception to this Yoked patter. Subject 1's Yoked trials were decided by the last subject's Executive trial.

Procedure. The experiment began with instructions and a practice trial for both the free-switching and forced-switching conditions. After any questions were answered, the subject pressed the 'Start' button to begin the experiment proper. Directly after, a subject was shown a screen with two category names at the top of the screen, one about a third of the way and the other two-thirds of the way from the left- hand side. The response fields, one for each category, were located a little below their respective category names. Trials began with the cursor arbitrarily placed in one of the fields. In the Executive trials, which were identified by the message "Your Choice" on the top of the screen, subjects were given freedom to switch categories (i.e., shift the cursor to the other response field) by pressing the return key whenever the active response field was empty. For example, if the left category was *fruit*, and *apple* was typed and then entered by way of the return key, the response field would be cleared and *apple* added to the response list for that category, the trial score would be incremented by 1 point, and the *fruit* response field would be cleared. The subject could then either type another *fruit* exemplar or press the return key again to shift the cursor to the other response field. For Yoked trials, which were identified by the message "No Choice" on the top of the screen, subjects entered exemplars in the same fashion as in Experiment 1. During these trials the subject was not allowed to switch between the two categories. Before a yoked subject was forced switched, the screen would read "Change" for 333ms and the cursor was moved to the other category. As in Experiment 1, trial length did not include typing time. Each subject underwent three executive and three yoked trials. Figure 11 illustrates the experiment screen and the two types of switch schedules.

a Executive Screen

b Yoked Screen



c Executive Schedule

time (secs)	action
2	"Ford"
4	"Buick"
10	"Mustang"
19	switch to colors
25	"Blue"
...	...
...	...
180	end

d Yoked Schedule

time (secs)	action
2	"Subaru"
4	"Ford"
10	"Lexus"
12	"Mazda"
19	forced switch
25	"Red"
...	...
180	end

Figure 11. (a) An example screen from the Executive condition. (b) An example screen from the Yoked Condition. (c) An example action schedule from an Executive subject. (d) An example action from a Yoked subject.

Results

As in Experiment 1, the dependant variable was area under the gain functions, which were calculated using the numerical integration method. We compared not the yoked and executive-decider double-category gain function but the individual category gain functions, which we extracted from both group's double-category gain functions. A raw comparison of the yoked and executive double-category functions would reflect the benefit the executive subject gains by being allowed to spend a disproportionate amount of time on her stronger category. For instance imagine an executive decider is a musician and must choose how to divide her time between musical instruments and cars. This executive decider will likely decide to spend more time on musical instruments. The yoked subject, in contrast, may have equal knowledge of cars and musical instruments but is forced to spend most of his time on musical instruments. Comparing only individual patch extraction rates eliminates the executive group's time-distribution advantage, which is unrelated to whether the cost of deciding affects within- category production.

A second executive-decider advantage may occur when subjects mentally switch and produce an exemplar without pressing the return key (the exogenous sign of a switch). If this mental "pre-switch" occurs frequently, executive-deciders would have faster responses after the overt switch. To control for this potential advantage, the amount of time needed to make the first correct response after a switch was controlled for by omitting the portion of the gain function between a switch and the first correct response (as in Experiment 1).

Results indicated that the extracted single patch functions of the yoked subjects yielded a greater area than those of the executive-decision subjects $t(11) = 2.31$, $p = .02$, bolstering the claim that autonomously switching between categories results in an attentional cost that interferes with the task of producing items. Figure 12 illustrates this effect. Executive-decision cost (EC) is calculated by subtracting the proportion of area from the executive-decider extracted functions to the yoked extracted gain function from one[5]. The average EC was calculated as 5.8%. See Appendix C for analysis details.

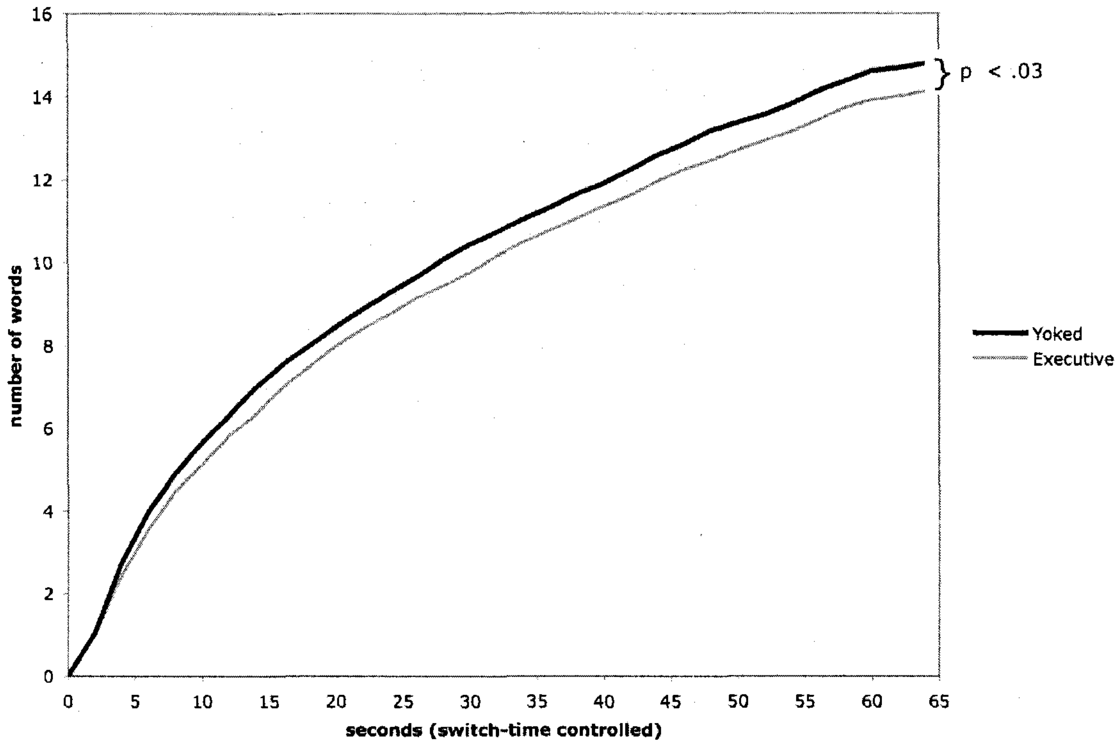


Figure 12. Yoked and Executive Single Category Function illustrates the area difference between the average Yoked and Executive single category function.

Experiment 3

The goal of this final experiment was to test whether memory foragers divide their time optimally between a higher and lower accessible category. This was done by mathematically, determining the optimal distribution of time between a higher and lower productive category pairings and comparing optimal time distribution to how subjects actually distribute their time during double category trials.

---

[5] $EC = 1 - \frac{area\_of\_executive\_decider}{area\_of\_yoked}$

Method

Subjects. Thirty Rice University undergraduates, 17 of whom were male, participated.

Materials. The stimuli consisted of 30 exemplars of each of 16 semantic categories, for a total of 480 words. Exemplars were selected from the Toronto Category Words List (Murdock, 1976) at the discretion of the experimenter. In addition, three 30-item "word-start" categories (words beginning with mor …, tri …, and del …, respectively) were formed for use in a single-category and a double category practice trial. These exemplars were entered into desktop computers, which controlled the experiment. A subject's response was accepted as correct only if it matched one of the 30 exemplars of the category. To reduce frustration from typos, matching was on the basis of the first four letters of the response without the subject's knowledge. For this reason, care had been taken to ensure that no two of the computer's exemplars of a category had matching first four letters.

Design. Counterbalancing measures aside, there were two independent variables: elapsed time and number of categories. Elapsed time was the time spent retrieving for a given trial, disregarding typing time. This "thinking" time extended to 3 minutes. The number of categories variable refers to whether, within a given trial, examples were being produced from one or two categories.

All subjects were tested in 4 blocks of 3 trials, for a total of 12 trials. The 16 categories were assigned to the 12 trials in essentially the same order for all subjects. More specifically, half of the subjects were assigned to the single and double categories as follows: professions, clothing, sports/bugs; tools, fruit, body parts/musical instruments;

fish, birds, colors/four-footed animals; U.S states, vegetables, cars/flowers. And the other half were assigned thus: professions/clothing, sports, bugs; tools/fruit, body parts, musical instruments; fish/birds, colors, four-footed animals; U.S states/vegetables, cars, flowers. The 12 trials were therefore arranged such that performance on a pair of categories by half of the subjects under the single-category condition could be used to predict performance by the other half of the subjects on the same two categories under the double-category condition.

Procedure. After receiving detailed instructions and both a single-category and a double-category practice trial, subjects underwent a mixed sequence of 3-minute single-category and double-category trials. For single-category trials, a single category name was displayed at the top center of the screen, and the subjects typed exemplars of that category into a text-field located just below. After typing each exemplar, they pressed the return key, which cleared the text-field and, if the entry matched one in the computer's word bank for that category, the trial score shown at the bottom of the screen was incremented by 1 point. In addition the exemplar was added to a list of accepted responses shown at the side of the screen. The purpose of this list was to reduce response repetition.

For double-category trials, two category names were shown at the top of the screen, one about a third of the way and the other two-thirds of the way from the left-hand side. The response fields, one for each category, were located a little below their respective category names. Trials began with the cursor arbitrarily placed in one of the fields. Subjects could switch categories (i.e., shift the cursor to the other response field) by pressing the return key whenever the active response field was empty. For example, if the left category was *fruit*, and *apple* was typed and then entered by way of the return key, the response field would be cleared and *apple* added to the response list for that category, the trial score would be incremented by 1 point, and the *fruit* response field would be cleared. The subject could then either type another *fruit* exemplar or press the return key again to shift the cursor to the other response field.

Results

To determine how closely subjects' time allocation strategy aligns with optimal time allocation, we first determined the optimal proportion of time subjects should spend on the high category at every interval I. This proportion was calculated using data from single category trials (e.g. the two functions observed from Group B's first and second trial). In each pair there was a more (high) and less (low) accessible category. At each interval I (720 corresponding to the number of quarter seconds in 3 minutes) of each block for each subject we calculated the optimal number of intervals each subject should have spent on the high category if they were to forage the same two categories in a double-category trial. Results indicate that at any interval I, subjects should have spent 68% of the intervals on the high category.

Next we calculated how subjects actually divided their time between categories. This was calculated by observing every entry (e.g "artist" for the category professions) during the double-category trials and calculating the average proportion of intervals spent on the high category to the number of total intervals spent between both the high and low categories. Was observed that subjects spent only 52% of their time on the greater category. A pair-wise t-test using category pairings as the random variable (8 comparisons) indicated a clear difference between the optimal percentage of time to spend on the high category and the actual amount of time subjects spent on the high category, t = 4.4(7), p = .003, demonstrating that humans do not naturally divide their time optimally when foraging between to differentially accessible categories.

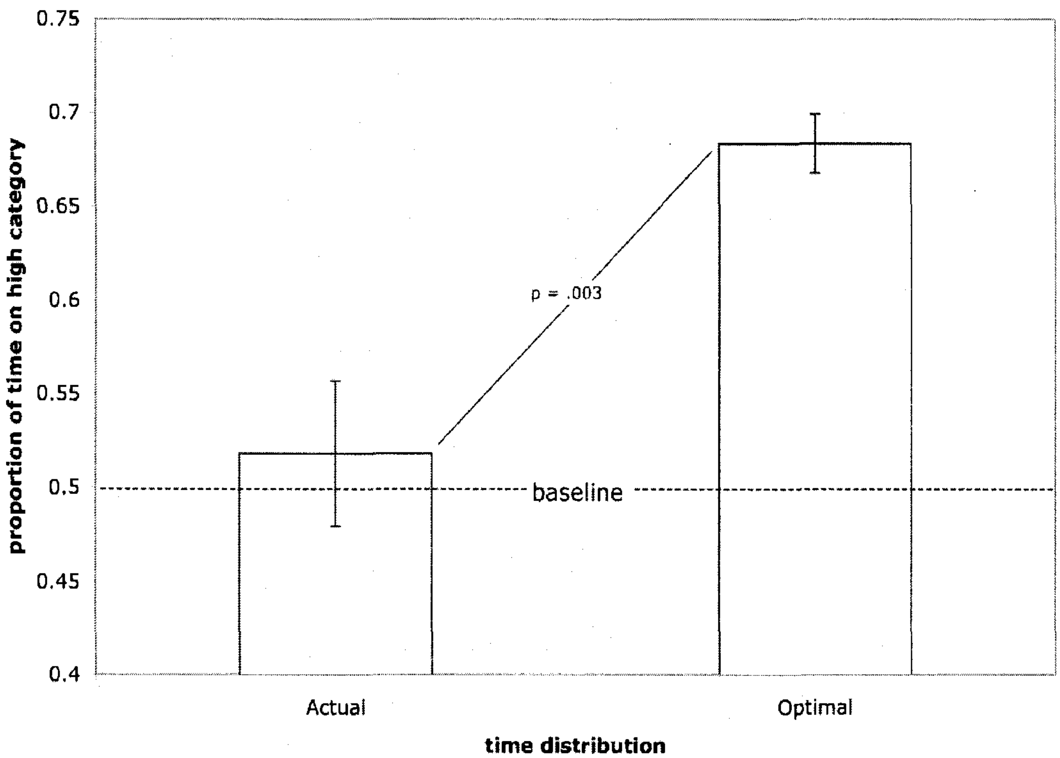


Figure 10 illustrates the difference between the proportion of time subjects should have spent on the greater category (Optimal) and how much time they actually spent on the greater category (Actual). The standard error bar on the Optimal bar indicates that there are differences in the optimal amount of time to spend on the 8 high categories.

We found that subjects differentiated between the high and low function only 10% of what is optimal, indicating that subjects are extremely poor at knowing how to divide their time between two differentially accessible categories to maximize the number of produced words. Moreover, the proportion of time subjects spent on the greater category did not significantly differ from baseline $t(7) = .47$, $p = .625$, indicating that subjects do not only undervalue the higher category but they do not act to distinguish between the two categories; spending approximately 50% of their time on both. The question of whether subjects are unaware of the value of the high category and/or they do not know how to utilize the higher category to maximize gain is beyond the scope of this research. The tendency for subjects to spend too little time on the higher category will be accounted for in our foraging prediction. See Appendix D for Experiment 3 details.

Predicting Dual Domain Word Production

In the following section our method of predicting word production from two semantic categories is described. It is expected that our prediction will be a closer resemblance to actual double-category performance than an expanded version of Maylor et al's (2001) serial prediction because our prediction accounts for switch cost, executive-decision cost, time-out benefit, and the tendency for subjects to not spend enough time on the more accessible category.

Derivation of our prediction occurred in two stages. First, observed single-category functions (OS) were adjusted for "would-be" [6] switch cost, time-out benefit, and executive-decision cost, which we predicted would alter OSs so that they mimic their corresponding, extracted single-category functions (ESs). Second, we accounted for the tendency for subjects to not spend enough time on the more accessible category by equalizing the time distribution of high and low composite[7] pairs to the degree

[6] "would-be" referring to factors that would have taken place if switching had occurred
[7] OSs that have been corrected for switch cost, time-out benefit, and executive-decision cost

demonstrated by subjects in Experiment 3.

Accounting for Time-out Benefit and Executive-decision Cost

Experiment 1 demonstrated that each timeout leads to a production benefit of 1.44%, after switch-cost is controlled. To incorporate this production benefit into our prediction, we determined the average number of switches to each category during double category trials and multiplied this number times 1.39 (the expected production gain from each switch). Experiment 2 demonstrated that autonomous switching between two categories leads to a 5.8% reduction in area across a production interval. The joint effect of time-out benefit and executive-decision cost for each interval of each category was calculated by subtracting executive-decision cost from the expected time-out benefit and then adding 1. This category *corrector* value (16 corrector values were derived. One for each category) was multiplied times each ordinate value of each subject's OS functions, which yielded a new corrected function which reflects the expected time-out benefit and executive decision cost. If the corrector value was greater than one then there was more expected time-out benefit than expected executive decision cost and the function will be corrected upwards. The opposite result would occur if the expected executive-decision cost were greater than the expected time-out benefit.

Accounting for Switch Cost.

Experiment 1 also demonstrated that each switch leads to 1.81 seconds of lost production time. In this section we will describe how switch cost is applied to the *corrected* category gain functions, which results in our final estimate of ESs.

Our initial step was to determine the total expected switch cost. This was done by observing double-category trials from Experiment 3 and determining the average number of switches to each category at every interval for each category. For example, if out of 15 people, one subject switched to category *A* during the 180th interval, then this interval of category *A* is assigned an average switch number of 1/15. After monitoring all switches to category *A* from all 15 subjects, we calculated the average cumulative number of expected switches at each of the 720 intervals of each category. Using these data, we calculated a cumulative switch cost function, which defines the expected cumulative number of switches to category *A* at each interval multiplied times the cost per switch. This cumulative function was used to adjust every corrected category gain function so that they reflected the lost production that may have occurred if switching had occurred. For example, if an entry was made at second 45 during a single category *A* trial, then the ordinate value of the cumulative switch-cost function at second 45 for category *A* was added to the recorded time of that given response. Thus, if the expected cost due to switching were 3 seconds after 45 seconds of production, then the adjusted time of response would read 48 seconds. This same correcting procedure was repeated for every interval of every corrected function. The end result was 240 (16 functions for each of the 8 high and 8 low categories) category gain functions that have all been corrected for executive-decision cost, time-out benefit, and switch cost. These 240 corrected functions instead of the original 240 OSs were used in the final prediction of double-category performance.

Accounting for Sub-optimal Time Allocation

Experiment 3 indicated that subjects at any one interval during a double-category trial are demonstrating only 10% optimality in their time distribution between the high and low categories from baseline (equal distribution of time between a high and low category). We posit that this sub-optimality must be accounted for to accurately predict double-category performance. To incorporate this sub-optimality, we first determined the number of intervals that should be spent on the high category for each of the 720 intervals for all 120 high/low category pairings (pre-derived in Experiment 3). We then determined the number of intervals that was closest to equal distribution (baseline). To determine the number of intervals a subject would have likely spent on the high category (for each of the 720 intervals of a 180 second production session), we added 10% of the difference in the number of intervals between baseline and the optimal number of intervals to the baseline number of intervals. This calculation was repeated for all 720 intervals (the domain of the final prediction). Any point I along a predicted double-category function was the ordinate sum of i intervals on the high category and $I-i$ intervals on the low category. An individual prediction was made for each of the 120 high/low function pairs. Our final prediction was the average of these 120 functions.

Testing Our Prediction

At heart, evaluation of our model involved only 2 critical comparisons: the composite single function / ES comparison, which tests the joint validity of our switch cost, time-out benefit, and executive-decision cost adjustments and the predicted/observed double category function comparison which tests the validity of the sub-optimal time allocation adjustment, assuming the corrected and extracted single functions demonstrate no meaningful difference in area.

As a result of the potential switch cost, time-out benefit, and executive-decision cost that may occur during switching, we predicted that ESs would have less area than an OSs. We observed that the area disparity between OS (M = 2197.9 word*quarter-seconds) than the ES (M = 2121.9 word*quarter-seconds) did not reach conventional level of significance $t(15) = 1.66$, $p = .059$. However a difference in area is suggested. (See Figure 13)

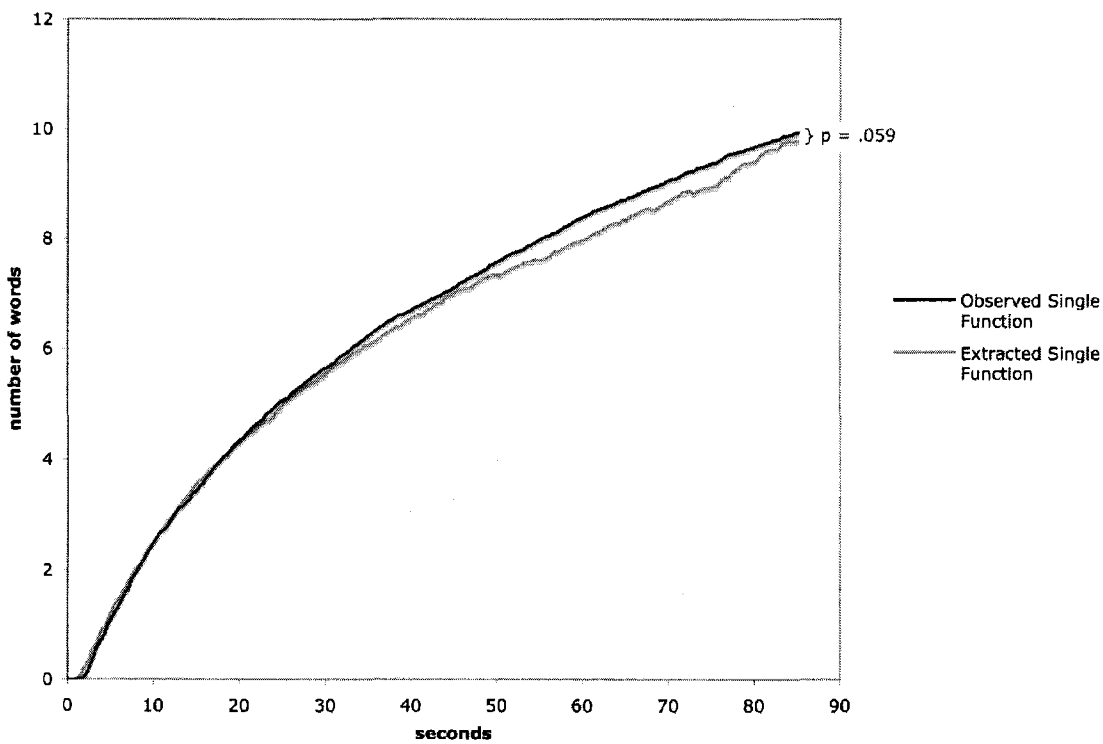


Figure 13 illustrates both the average gain function from a single category trial and the average extracted function from a double category trial.

As predicted, no meaningful difference in area was observed between the average composite function (M = 2140 word*quarter-seconds), which does control for the potential benefits and costs when switching between categories, and the average OS (M = 2197.9 word*quarter-seconds), t(15) = .377, p=.71 (Figure 12).

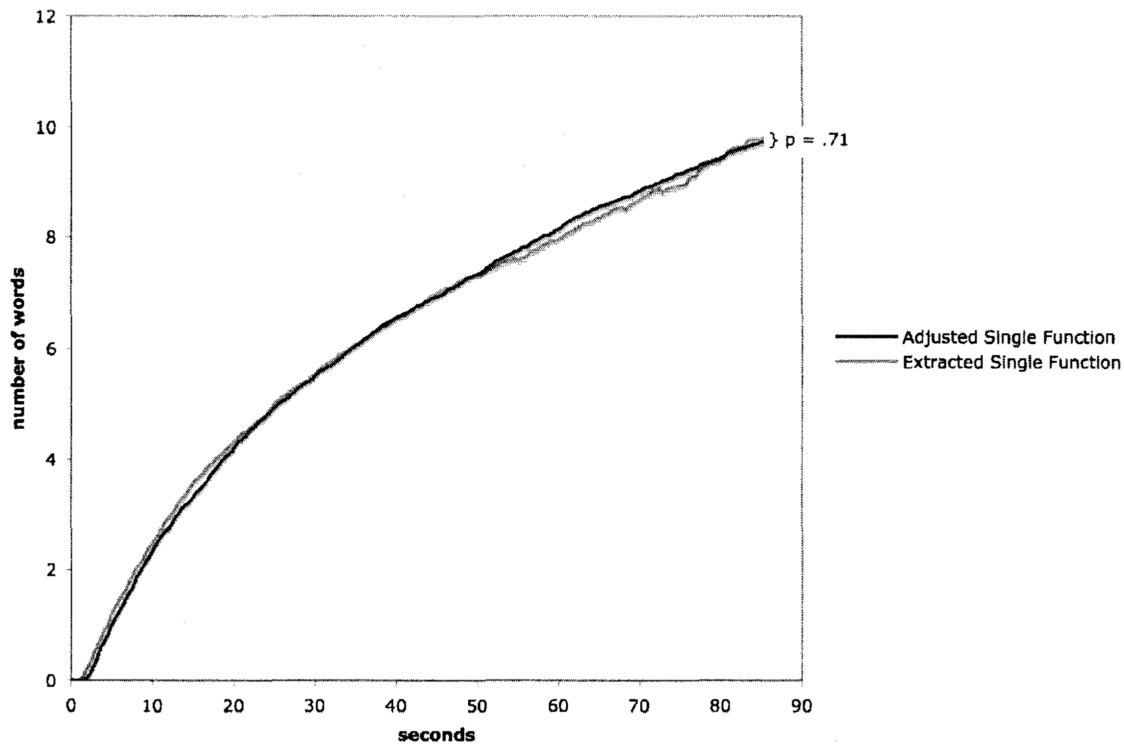
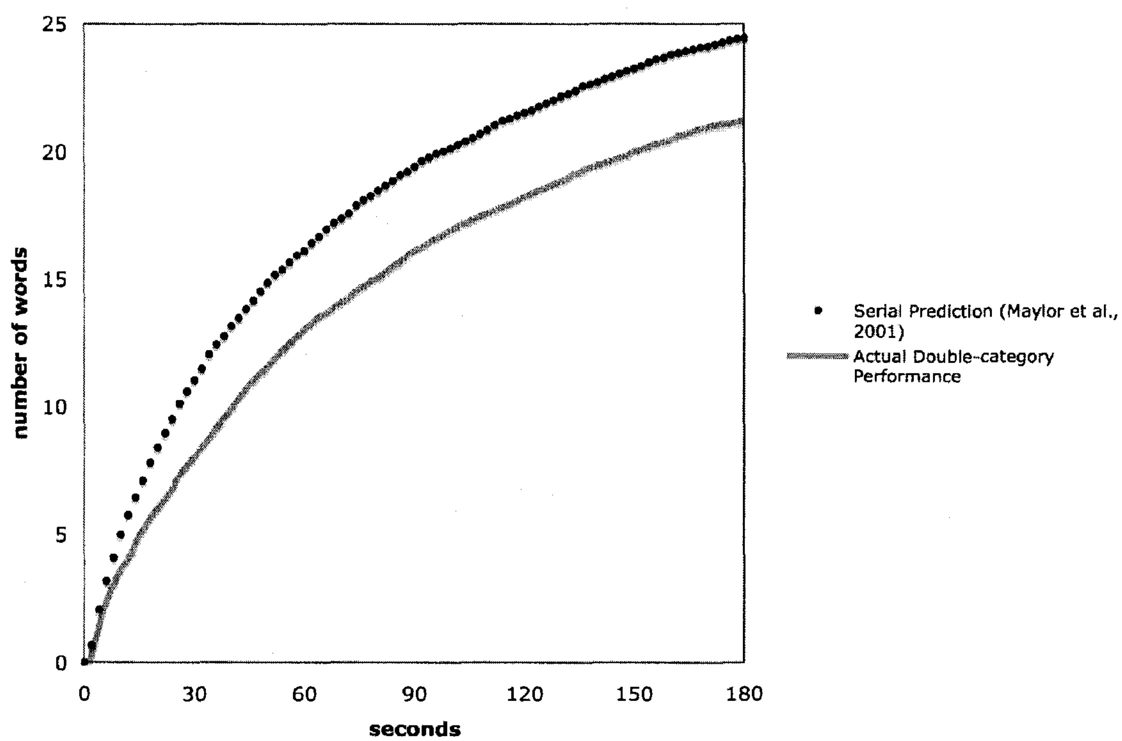


Figure 14 illustrates both the average adjusted single gain function and average extracted single gain function. It was predicted that these two functions are the same.

As is shown in both Figure 13 and 14, both the functions did not extend the full 180 seconds. Though ES have a potential range of 0 to 180 seconds, on average, these functions extend to about the half way mark of 180 seconds. Beyond 90 seconds, the amount of data represented drops dramatically because there were relatively few subjects that had large disparities in time distribution compared to those with near equal distribution of time distribution. For the purpose of comparison we decided to extend the analysis domain to the interval that the average ES included the data of at least 50% of the subjects, which was interval 342 or second 85.5 (of 180).

The second comparison concerns the effectiveness of our sub-optimal time allocation parameter to assist in the accurate prediction of double-category gain. We predict that actual double category gain will not align with a serial prediction that does not account for switch cost, executive-decision cost, time-out benefit, and the time-equalizing principle; much like the Maylor et. al. (2001). Indeed, there was a difference between these two gain functions with every serial prediction function yielding greater area than actual performance[8].



[8] The serial prediction overestimated double-category production for all eight of the category pairs tested for in Experiment 3

Figure 13 illustrates both the predicted double-category gain using Maylor et al.'s (2001) serial prediction and actual double-category gain.

However, using our prediction method we accurately predicted performance. There was no area difference between our (final) predicted function (10622.5 word*quarter-seconds) and actual performance (10457.8 word *quarter-seconds) $t(7) = .596$, $p = .57$

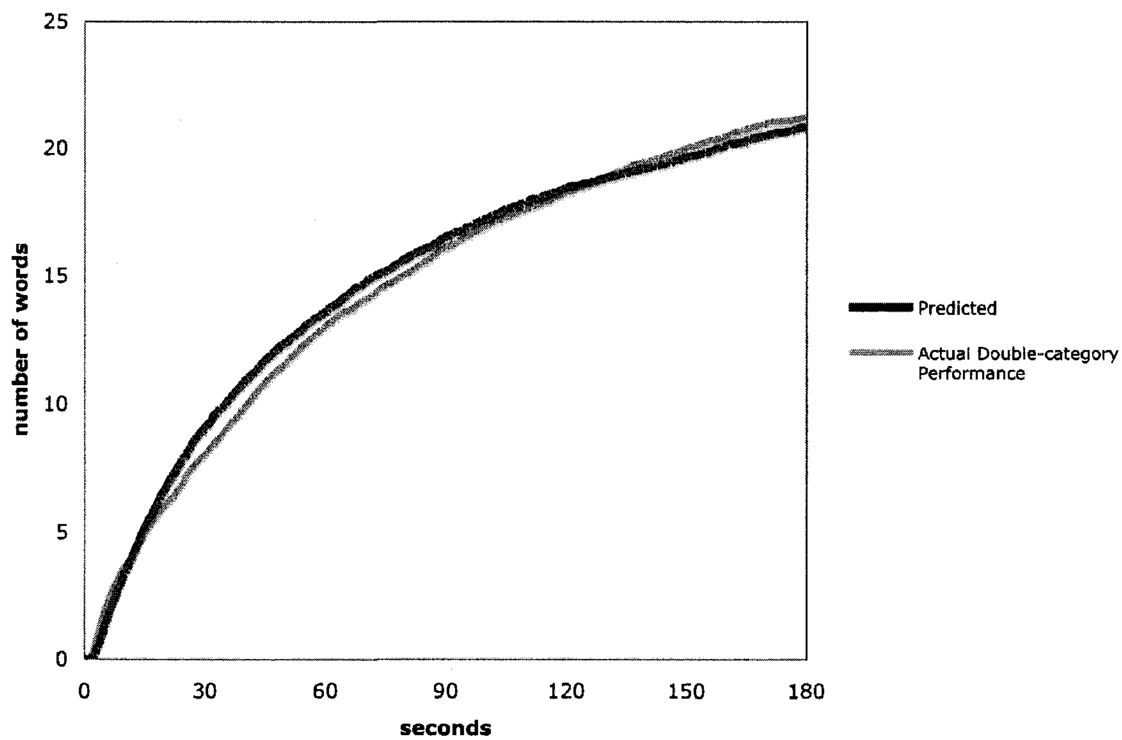


Figure 14 illustrates both our double-category gain prediction and actual double category gain.

In short, the proximity of our prediction to actual performance indicates that consideration of switch cost, time-out benefit, executive-decision cost, and the tendency for foragers to spend too little time on the more accessible category are sufficient to predict dual-domain memory production.

General Discussion

The goal of the research reported here has been to explore how humans "forage" for exemplars of semantic categories. More specifically, it has been to predict the gain, or cumulative production function for exemplars of two semantic categories from the gain functions for the two categories derived one at a time. Four factors were identified as relevant to the prediction: a switch cost, a time-out benefit, an "executive" cost of deciding when to switch, and an insufficient sensitivity to the difference in yield between categories. Let us consider these factors in turn.

Switch cost

That switching from one task to another can incur a time cost has been extensively documented (Wylie & Allport, 2000; Jersild, 1927), but it could reasonably be argued that that the switch in these demonstrations is typically between more disparate tasks than the two tasks (i.e., producing words of one semantic category and producing words of another semantic category) in the research reported here. Rohrer et al. (1998) demonstrated that it takes greater time to produce studied category exemplars in an alternating run (e.g., shirt, doctor, sock, accountant) than in a blocked run (e.g., shirt, sock, pants, hat). These authors explained this result, not in terms of switch-cost related language, but by distinguishing between serial and parallel retrieval processes. Specifically it is posited that producing examples across categories involves a serial retrieval process, where item production is limited to one at a time; while producing items from the same category involves a parallel retrieval process, allowing multiple items to be retrieved simultaneously (faster). Findings from Experiment 1 were compatible with Rohrer et al.'s finding. Specifically, we found that the time to produce an item after a switch was greater than when there was no switch.

Though Experiment 1 corroborated Rohrer et al.'s finding that producing an item from a category takes longer when the preceding item was produced from a different category than from the same category, our estimation of switch cost was greater than Rohrer et al.'s estimation by 1.25 seconds. We posit that this difference can be attributed primarily to differences in the subject's task preparation and the type of memory being addressed. In Rohrer et al.'s paradigm, subjects were exposed both to the categories and potential items from each category prior to test. Thus, before test, subjects were aware that they would be switching from either category *A* to *B* or to *B* to *A* and which category items would be acceptable. Switch-cost literature has shown that advanced preparation for switch can reduce or eliminate the cost of switching (Dreisbach, Haider, & Kluwe, 2002; Altmann, 2004). In Experiment 1, subjects were unaware of the categories and the order in which they would be switched. This lack of preparation may partly explain why our switch cost estimate was greater. Another potential reason for the switch-cost estimate difference is that Rohrer et al.'s subjects may have produced items from short-term memory, where our subjects, by necessity produced items from semantic memory. It is well documented that semantic information is organized (Tulving & Donaldson, 1972). Data from Experiment 1 demonstrate that moving between organizations (semantic categories) leads to a significant time cost. Information in short-term memory is less structured, and it may therefore require less time to traverse between than semantic memory.

Though Experiment 1 clearly showed a switch cost, our precise estimation of switch cost may not be the same as when subjects are switching between categories in a dual-domain memory foraging scenario. Unlike Experiment 1, when foraging, subjects can prepare for a task switch because they can control when to switch and are aware of both patches. To this extent Experiment 1 could have overestimated switch cost during memory foraging because subjects were not aware of the next category. On the other hand, we could also have underestimated switch costs because switch cost is greater when there is no explicit cue (Arlington & Logan, 2004). During memory foraging subjects must provide their own cues[9].

Time-out benefit

To test for a time-out benefit, we observed performance after completion of a switch. We used a correct response as an indicator of a completed switch because it made certain that a switch had indeed occurred. As was predicted, there was increased productivity after a switch than after the same critical instant when no switch occurred. Dual-task literature has focused on the cost of switching, while and proactive interference literature has focused on the benefit of switching; Experiment 1 elucidated how both a cost and benefit can occur together when memory foraging.

The time-out benefit analysis concerned production only after the first response, which reflects a successful switch. The possibility remains, however, that a release from proactive interference occurs during a switch (as well as after) and is partially easing the costly mental shift that occurs when switching categories. If this were the case, then the

[9] The cost of switching between categories, in the context of memory foraging, may actually be the sum of a cost of a mental shift to a different task set and the beneficial effect of reduced proactive interference.

mental shift cost and time-out benefit may have both been underestimated.

Executive-decision Cost

Memory foragers must constantly decide whether to switch to a different category or to continue producing 42 items from the same category during double-category trials. Those supplying the predictor (single-category) functions, however, have only the task of producing items. Divided-attention literature has consistently demonstrated that dividing attention between tasks has a negative effect on task performance. Researchers explain this result with the claim that humans have a limited amount of attention to allocate and that increasing the number of tasks increases attentional demand, leading to fewer available attentional resources and thus greater difficulty performing tasks. Though concurrent tasks compete for attentional resources, task performance is affected by task similarity (Allport, 1972; Triesman and Davies, 1973). We expected that the task of recalling items from categories and the task of monitoring switching are either so different that there would be minimal interference and no reliable affect on production or that the two tasks are closely related enough such that they would interfere with one another and diminish production. We did not expect that within-category production would be greater in the Executive-decider condition. Experiment 2 addressed this hypothesis by giving subjects both the choice to switch (Executive) or forcing them to switch (Yoked) and comparing extracted gain function from each condition. After controlling for switch time, we found that production was greater within categories in the Yoked condition, when subjects did not have to choose when to switch categories. This result suggests that there is a substantial attentional cost when making and executing a switch strategy and that the predictor gain functions (gain functions from single category trials) need to be adjusted so that they represent this cost.

Sub-optimal time allocation

Perhaps most unclear prior to this research was how effectively subjects would divide their time between categories. Animals have demonstrated an innate ability to divide their time between patches optimally, however research on this behavior in humans is relatively scarce. Related research on humans is exclusively concerned with the search for external sources of information (Pirolli and Card, 1999). Experiment 3 was an initial attempt at measuring human effectiveness at searching for internal information. We found that subjects do not naturally optimize their time between two categories. Specifically, there is a tendency to dramatically undervalue the more accessible category.

Future Directions

Though we were able to predict performance, our prediction is limited to just two domains and requires the observation of actual performance to determine a *would-be* switch schedule. This limitation does not make an ideal model. Potential improvements of our foraging model are the accommodation of factors such as the number, semantic variance, and the perceived relative accessibility of the given domains. Below is a discussion of how these three foraging parameters may affect switch cost, time-out benefit, executive-decision cost, optimality of time allocation, and the frequency of switches.

Increasing the number of domains

When foraging from only two domains, the only possible switch is to the single other category. Preparation in this context is not difficult because there is only one option. As the number of domains increase the number of switching options will also increase. As a result, preparation may become more difficult and less complete, leading to a more time-consuming switch. Another possible ramification is a corresponding increase in switch rate due to the general strategy of extracting the most accessible items from each category. Such an increase may lead to a higher cost per switch as suggested in Experiment 1 and demonstrated in Arlington & Logan (2004).

Increasing the number of domains may also exaggerate the Executive-decision cost by increasing the complexity of the switch-monitoring task. With three or more domains, not only would a forager need to constantly monitor when to switch but also where to switch. This added task complexity might further tax the forager's attentional resources. The extent that increasing the number of domains affects Executive-decision cost can be addressed in a paradigm similar to Experiment 2 where executive control and the number of domains are both manipulated.

Experiment 3 demonstrated that subjects do not divide their time optimally between just two categories. Increasing the number of domains may exasperate this deficit leading to even less optimal performance. Manipulating the number of single category trials and categories in the multiple category trial within an Experiment 3 block would address this question.

Changing the semantic distinctiveness of domains

The semantic distinctiveness of the domains within a group may affect time-out-benefit. An example of two domains with a low semantic distinctiveness is *Girls Names* and *Boys Names;* an example of two domains with a high semantic distinctiveness is *Tools* and *Countries*. Time-out benefit, which we posit, is driven by a release of proactive interference, may increase as the average semantic distinctiveness of the categories increases because the amount of proactive interference released has repeatedly shown to be affected by semantic distinctiveness (Gardiner et al. 1972). A paradigm similar to Experiment 1, which manipulates the semantic variance of the domains, would address this question. Switch cost may also be affected by semantic distinctiveness. Rohrer et al. (1998) demonstrated that producing an additional item from the same category reflects a parallel retrieval process while producing an item from a different category reflects a serial retrieval processes. Rohrer et al. suggested that parallel retrieval may occur because of semantic priming, which occurs when the activation of a category name (e.g., *tools*) automatically activates examples from that category (e.g, *hammer*, *nails, ...*) at the same time (Loftus & Loftus, 1974). As categories become less semantically distinct, the probability that one category will prime items from the other categories increases, which may lead to faster production when producing an item from a different category.

Changing the relative accessibility of competing domains

In Experiment 3 we demonstrated that memory foragers are only 10% optimal when dividing their time between a high and low accessible category. Unknown, however, is whether performance would become more or less optimal by manipulating the relative accessibility of the competing domains. To answer this question, it may be beneficial to elucidate the precise cognitive deficit that is driving memory foragers to not spend enough time on the more accessible category. One potential source is simply that foragers do not know how to solve the problem of how to divide their time between categories to maximize gain. Another more subtle possibility is that foragers do not accurately judge the relative accessibility of each category, in which case, foragers may be dividing their time optimally with respect to their faulty representations of each category's accessibility. In contrast, if high/low category pairs were rated as having the same level of accessibility then, sub-optimal performance could be at least partly attributed to a faulty awareness of the two categories relative accessibility. If sub-optimal performance is purely a problem-solving deficit (not knowing how to optimally divide time between a high and low category even with the understanding that the two categories have different levels of accessibility), then changing the relative accessibility of category groupings, or the field of domains, should not have a significant bearing on the optimality of performance. In contrast, directly changing the relative accessibility of grouped categories may have a direct effect on how clearly a forager can judge their relative accessibilities (judging the relative accessibility of Swahili words vs. English words may be an easier judgment than the relative accessibility of *Clothing items* vs. *Professions*), which would potentially alter the optimality of a forager's time distribution.

Predicting switch rate

An additional omission of our model is the inability to predict a natural switch rate and how this is affected by the above manipulations. Being able to predict the average switch rate is important because switches are the driving force behind time-out benefit and switch cost. It seems likely that as the number of categories increase the switch rate would increase as well and as the accessibility gap between categories increases that switching would decrease. However, it is virtually unknown whether factors such as semantic variance of the categories would also affect the amount of switching.

Similarities and Differences to Ecological Foraging

An additional goal of this research, beyond predicting word production from dual domains, was to determine whether memory foraging is isomorphic to ecological foraging to the extent that memory foraging could be characterized using Stephens' and Krebs' (1986) Optimal Foraging Theory. Three fundamental assumptions of this theory are a between patch time, serial patch extraction, and non-dynamic gain functions (the average of all patches is not affected by switching). Rohrer et al. (1998) as well as results from Experiment 1 demonstrated that there is a between patch (category) time, which can be likened to the time an animal spends traversing the distance between two patches. In addition, previous literature (Rohrer et al., 1998, Maylor et al. 2001) has provided strong evidence that only one category can be accessed at a time; just as an animal in the physical world can occupy only one space at a time. Unlike ecological foraging however, and in direct conflict to a fundamental assumption of Optimal Foraging Theory, is time-out benefit. Evidence from Experiment 1 as well as a large body of literature has demonstrated that switching to a different category can lead to a higher probability of retrieval. As a result, the accessibility of a memory patch may change as memory foragers switch from and return to it. The effect of this memory phenomenon precludes memory foraging from being predicted using an optimal foraging model and necessitates the formation of a unique prediction method. Evidence from this research indicates that memory and ecological foraging share features but are not isomorphic.

Conclusion

This thesis introduced the notion of memory foraging and provided a method for predicting memory foraging performance in the simplest scenario possible. Our findings suggest that memory foraging can be predicted with a high level of accuracy if switch cost, release from proactive interference, executive-decision cost, and sub-optimal time allocation are accounted for. It can be argued that memory foraging is one of the most prevalent mental processes, considering the constant need for access to internal information. We discovered that humans do not forage memory with maximum efficiency, which is likely manifested in slower performance in daily memory tasks. Memory foraging, however, is a complicated issue, and it clearly requires substantially more investigation before the factors that affect human's ability to produce items from multiple memory sources, in its many contexts, are fully understood.

References

Allport, A., Styles, E. A. & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umilt 'a & M. Moscovitch (eds.), *Attention and performance XV: Conscious and unconscious information processing,* pp. 421–452. Hillsdale, NJ: Erlbaum.

Altmann, E. M. (2004) Advanced preparation in task switching. Psychological Science, 15(9), 616-622.

Arrington, C. M., & Logan, G. D. (2005). Voluntary task switching: Chasing the elusive homunculus. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 31, 683-702.

Allport D.A., Wylie G. (2000) Task switching, stimulus–response bindings, and negative priming. In: Attention and performance XVIII (Monsell S, Driver J, eds), pp. 35–70. Cambridge, MA: MIT Press.

Brown A.S. (1981) Inhibition in Cued Retrieval. *Journal of Experimental Psychology: Learning Human and Memory,* 7, 204-215.

Brown, J. (1958). Some Tests of the Decay Theory of Immediate Memory. *Quarterly Journal of Experimental Psychology 10*, 12-21.

Dreisbach, G. Haider, H. & Kluwe, R. H. (2002). Preparatory processes in the task switching paradigm: Evidence from the use of probability cues. Journal of *Experimental Psychology: Learning, Memory, & Cognition*, 28, 468-483.

Elizabeth A. M., Chater N., & Jones G. (2001). Searching for two things at once: Evidence of exclusivity in semantic and autobiographical memory retrieval. *Journal of Memory and Cognition*, 29, 1185-95

Evans, J.E., Meyer, D.E., Rubenstein, J.S., (2001). Executive control of cognitive processes in task switching. Journal of Experimental Psychology: Human Perception and Performance, 27(4), 763-797.

Gardiner, J. M., Craik, F. I. M., & Birtwistle, J. (1972). Retrieval cues and release from proactive inhibition. *Journal of Verbal Learning and Verbal Behavior, 11*, 778-783

Greenberg, R., & Underwood, B.J. (1950). Retention as a function of stage of practice. *Journal of Experimental Psychology, 40*, 452-457.

Gunter, B., Clifford, B., & Berry, C. (1980). Release from proactive interference with television news items: Evidence for encoding dimensions within televised news. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 216-233.

Keppel, G., & Underwood, B.J. (1962). Proactive inhibition in short - term retention of single items. *Journal of Verbal Learning and Verbal Behavior, 1*, 152 - 161.

Lavie, N. (1995) Perceptual load as a necessary condition for selective attention. *of Experimental Psychology: Human Perception and Performance*, 21, 451-468.

Logan, G. D., & Schulkind, M. D. (2000). Parallel memory retrieval in dual-task situations: I. Semantic memory. Journal of Experimental Psychology: Human Perception & Performance, *26*, 1072-1090.

Logan, G. D., & Gordon, R. D. (2001). Executive control of visual attention in dual-task situations. *Psychological Review, 108,* 393-434.

Loess, H., & Waugh, N.C. (1967). Short-term memory and intertial interval. *Journal of Verbal Learning and Verbal Behavior*, *6*, 455-460

Loftus, G. R., & Loftus, E. F. (1974). The influence of one memory retrieval on a subsequent memory retrieval. Memory & Cognition, 2, 467-471.

Jersild, A. T. (1927). Mental set and shift. *Archives of Psychology*, Whole No. 89.

Johnston, W.A., Greenberg, S. N., Fisher, R.P., & Martin, D. W.(1970). Divided attention: A vehicle for monitoring memory processes. *Journal of Experimental Psychology, 83,* 164-171.

Marios, R., Ivanoff, J. (2005). Capacity limits of information processing in the brain. *Trends in Cognitive Science,* 9, 296-305.

Martin, D. W. (1970). Residual processing capacity during verbal organization in memory. *Journal of Verbal Learning & Verbal Behavior, 9,* 391-397.

Maylor E. A., Chater N., and Jones G. V. (2001) Searching for two things at once: Evidence of exclusivity in semantic and autobiographical memory retrieval. *Memory & Cognition 2001, 29 (8),* 1185-1195.

McCann, R., & Johnston, J. (1992). Locus of the single-channel bottleneck in Dual-task interference. *Journal of Experimental Psychology: Human Perception and Performance*, 18, 471-484.

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences, 7*, 134-140.

Mueller, C.A., & Watkins, M.J. (1977) Inhibition from Part-Set Cuing: A Cue-Overload Interpretation. *Journal of Verbal Learning and Verbal Behavior,* 16, 699-709.

Murdock, B.B., Jr. (1976). Item and order information in short-term serial memory. *Journal of Experimental Psychology: General, 105*, 191-216

Pirolli P. and Card S. (1999) Information Foraging. *Psychological Review, 106,* 643-675

Rohrer, D., Pashler, H., & Etchegaray, J. (1998). When two memories can and cannot be retrieved concurrently. *Memory & Cognition, 26*, 731-739.

Slamecka, N. J. (1969). Testing for associative storage in multitrial free recall. *Journal of Experimental Psychology, 81,* 557–560.

Stephens, D. W., & Krebs, J. R. (1986). Foraging theory. Princeton, NJ:Princeton University Press.

Treisman, A., & Davies, A.. (1973). Divided attention to ear and eye. In S. Kornblum (Ed.) Attention and Performance IV, Academic Press, 101-117.

Trumbo, D., & Milone, F. (1971) Primary task performance as a function of encoding, retention, and recall in a secondary task. *Journal of Experimental Psychology*, 40, 245-255.

Tulving, E & Donaldson, W. (1972). *Organizations of Memory,* Academic Press, New York.

Unnderwood, B.(1957) Interference and forgetting. *Psychological Review* 64:49-60.

Waugh, N.C., & Loess, H. (1967) Short-Term Memory and Intertrial Interval. *Journal of Verbal Learning and Verbal Behavior,* 6, 455-460

Wickens, D.D. (1970) Encoding categories of words: An empirical approach to meaning, *Psychological Review, 77*, 1–15.

Wickens, D.D. (1972) Characteristics of word encoding. In: A. Melton and E. Martin, Editors, Coding processes in human memory, Winston, Washington, DC (1972).

Williams, M.D., & Hollan, J.D.(1981). The process of retrieval from very long-term memory. *Cognitive Science, 5,* 87-119.

Wylie, G., & Allport, A. (2000). Task switching and the measurement of "switch costs.

Psychological Research, 63, 212-233.

Appendix A

Analysis functions used in Experiment 1, 2, 3, and in the testing of the model

When comparing functions the dependent variable was area. We defined the (1) area as the sum of all the ordinates at every time interval of a function.

(1) Area …(A) the sum of all ordinates of every interval of a gain function (func).

$$A(func) = \sum_{I=1}^{NI} func(I)$$

Each experiment produced many (2) gain functions. These functions are defined as the number of words produced by interval *I*. These functions may also be specified by (S)ubject, (C)ategory, and (Cond)ition.

(2) Gain_Single_Specific … (GSS) the number of words produced by interval I specified by *S, C,* and *Cond*

$$GSS(I,S,C,Cond) = number\_of\_words\_produced(I,S,C,Cond)$$

Appendix B

Analysis details for Experiment 1

Variables and constants used in Experiment 1:

Below are variables and constants used in Experiment 1 analysis functions. This sections should be used as a reference when reading the next section.

I ... interval identifier, ranges from 1 to 180

Cond ... identifies the switch condition or the switch condition that the control no-switch condition is compared to, either high- or low-switch

P... represents one of p production intervals. A *P* is an interval in which production was allowed before switching or the passing of a critical instant. In the high switch condition *P* ranged from 1 to 16 in the low-switch condition *P* ranged from 1 to 4. *P* is also used a switch reference. *P*=2 refers to the first switch.

Production_Interval_Length ... (*PIL*) represents the number of intervals in a *P*, which was either 45 or 180.

First_Response_Type ...(*FRT*) indicates the type of first response latency, which is either after a switch (First_Response_After_Switch_Specific) or after a critical instant (First_Response_After_Critical_Specific). These two response latency types are discussed in greater detail in the next section

C ... category identifier within a *Cond*, ranges from 1 to 4

S ... subject identifier, ranges from 1 to 60

NI ... the number of intervals within a GSS (= 180)

NPH ... the number of *P*s in the high-switch condition (= 16)

NPL ... the number of *P*s in the low-switch condition (= 4)

NC ... Number of categories within a *Cond* (= 4)

NS ... Number of subjects total (= 60)

NSS ... Used when averaging composite functions. Represents the number of subjects at interval *I* for a composite function specified by *C*, *Cond*, and *S*, ranges from 20 to 1

Analysis functions used to calculated switch cost:

To calculate switch cost we identified the amount of time it required to make a correct response after a switch in a switch condition and subtracted the amount of time it required to make a correct response in the no-switch control condition. Therefore we were interested in two types of first-response latencies: (3) The amount of time needed to make a response after a switch and (4) the amount of time to make a response after a non-switch (critical instant) in the no switch condition.

(3) First_Response_After_Switch_Specific … (FRASS) the amount of time to make a response after a switch specified by *P, C, Cond,* and *S*

(4) First_Response_After_Critical_Specific … (FRACS) the amount of time to make a response in the high switch condition after a critical instant specified by *P, C, Cond,* and *S*

We calculated the average first response time for the (5) high- and (6) low-switch conditions as well as the (7) high and (8) low-switch no-switch comparison condition. Calculating one of these four average first response times required the identification of first the response time at each switch (P) in each category (C) function within a condition for each subject (S).

(5) First_Response_After_High_Switch_Average …(FRAHSA) the average amount of time to make a correct response after a switch in the high-switch condition.

$$FRAHSA = \frac{\sum_{S=1}^{NS}\sum_{C=1}^{NC}\sum_{P=2}^{NPH}\mathrm{FRASS}(P,C,high-switch,S)}{(NPH-1)*NC*NS}$$

(6) First_Response_After_Low_Switch_Average …(FRALSA) the average amount of time to make a correct response after a switch in the low-switch condition.

$$FRALSA = \frac{\sum_{S=1}^{NS}\sum_{C=1}^{NC}\sum_{P=2}^{NPL}\mathrm{FRASS}(P,C,low-switch,S)}{(NPL-1)*NC*NS}$$

(7) First_Response_After_High_Critical_Average …(FRAHCA) the average amount of time to make a correct response after a high frequency critical instant

$$FRAHCA = \frac{\sum_{S=1}^{NS}\sum_{C=1}^{NC}\sum_{P=2}^{NPH}\mathrm{FRACS}(P,C,high-switch,S)}{(NPH-1)*NC*NS}$$

(8) First_Response_After_Low_Critical_Average ...(FRALCA) the average amount of time to make a correct response after a high frequency critical instant

$$FRALCA = \frac{\sum_{S=1}^{NS}\sum_{C=1}^{NC}\sum_{P=2}^{NPL} \mathrm{FRACS}(P, C, low - switch, S)}{(NPL - 1) * NC * NS}$$

To determine the (9) switch cost for the high-switch condition we subtracted FRAHCA from FRAHSA; to determine the (10) switch cost in the low-switch condition we subtracted FRALSA from FRALCA.

(9) Switch_Cost_High_Average ... (SCHA) the average switch cost in the high-switch condition

$$SCHA = FRAHSA - FRAHCA$$

(10) Switch_Cost_Low_Average ... (SCLA) the average switch cost in the low-switch condition

$$SCLA = FRALSA - FRALCA$$

The (11) average switch cost across the high- and low-switch condition were calculated weighting *SCHA* five times as high as *SCLA* because the high-switch condition required the subjects to switch five times as often than low-switch condition and thus *SCHA* represented five times the amount of data.

(11) Weighted_ASC ... (WASC) the weighted average of ASCH and ASCL. ASCH is weighted 5 times as high because the high switch condition yielded 5 times the number of switch data

$$WASC = \frac{5 * SCHA + SCLA}{6}$$

Analysis functions used to calculate time-out benefit:

Time-out benefit was operationalized as the percentage difference in total item productivity caused by a switch. To isolate time-out benefit, every first response time (which includes switching time) and the first production interval (no switching occurred until after the first production interval) was removed from each GSS in both the switch and no-switch conditions, which resulted in (12) switch-cost controlled composite functions reflecting only the time after subjects have demonstrated active production of category items. Removing the entire first-response time, however, was not ideal because it reflects not only the amount of time to switch but the amount of time to find the first item after a switch for the switch conditions in the switch conditions and pure production time for the no-switch condition (ideally there should be no time removed from the no-switch condition because there is no switch cost). With our data, it is not possible to identify when a switch is over and searching within a new patch begins. We decided to remove the entire first-response time to make certain that all switch cost was controlled for.

The domain for each composite function is equal to 180 intervals (*NI*) minus the number of intervals in the first production interval (*PIL*) minus the total amount of first response time. Because there were five times the amount of first-response times in the high-switch condition, these functions tended to have more restricted domains then those from the low-swtich condition.

(12) GS_Cost_Controlled_Specific … (GSCCS) a composite GSS function after removing the first production interval and the time between every switch and the time of first response for production intervals 2 to NP(H/L) or the time between every critical instant and the corresponding first response for production intervals 2 to NP(H/L).

$$c = \mathrm{FRT}(2,C,Cond,S)$$
$$\alpha = \mathrm{FRT}(P,C,Cond,S) + I - PIL$$
$$b = \mathrm{GSS}(PIL + c,C,S)$$
$$\lambda(I,P) = \begin{cases} \mathrm{FRT}(P,C,Cond,S) + I & if \quad I \leq PIL - \mathrm{FRT}(P,C,Cond,S) \\ \mathrm{FRT}(P,C,Cond,S) + I + \lambda(a,P+1) & if \quad I > PIL - \mathrm{FRT}(P,C,Cond,S) \end{cases}$$
$$\mathrm{GSCCS}(I,PIL,FRT,Cond,C,S) = \mathrm{GSS}(PIL + c + I + \lambda(I + c - PIL,3),C,S) - b$$
$$I \in [1, NI - (\sum_{P=1}^{NP} \mathrm{FRT}(P,C,Cond,S)) - PIL]$$

To determine an average composite function for each condition, the (13) number of subjects whose GSCCS function extended to interval 0 to 720, were identified. This was necessary because as *I* increased, average GSCCS functions were represented by fewer and fewer people because of first-response time deletions (individual differences in first response times led to GSCCSs with various domains). So that the (14) average composite function remained representative we omitted the average GSCCA function at and above the *I* (15) where fewer than half of the subjects were represented.

(13) Number_Subjects_at_Interval … (NSI) the number of subjects represented at interval I on category C of a GSCCS

(14) GSCC_Average … (GSCCA) The average of all GSCCs from a given *Cond* and *FRT*. As *I* increases, the number of subject's data that is represented decreases and stops at ICO.

$$\mathrm{GSCCA}(I,PIL,FRT,Cond) = \frac{\sum_{C=1}^{NC} \frac{\sum_{S=1}^{NS} \mathrm{GSCCS}(I,PIL,FRT,Cond,C,S)}{\mathrm{NSI}(I,C)}}{NC} \qquad I \le \mathrm{ICO}(FRT)$$

(15) Interval_Cut_Off … (ICO) represents the interval at which half the GSCCs within a specified *FRT* did not continue past due to deletion.

The average composite gain function for the (16) high-switch condition, (17) high-switch control, (18) the low–switch condition, and the (19) low-switch control were derived using GSCCA.

(16) GS_Switch_Controlled_High_Average … (GSSCHA) The average of all GSCCs in the high switch condition.

$$\mathrm{GSSCHA} = \mathrm{GSCCA}(I,45,FRASHS)$$

(17) GS_Critical_Controlled_High_Average … (GSCCHA) The average of all GSCCs in the no switch condition when compared to the high switch condition.

$$\mathrm{GSCCHA}(I) = \mathrm{GSCCA}(I,45,FRACHS)$$

(18) GS_Switch_Controlled_Low_Average … (GSSCLA) The average of all GSCCs in the low switch condition.

$$\mathrm{GSSCLA}(I) = \mathrm{GSCCA}(I,180,FRASLS)$$

(19) GS_Critical_Controlled_Low_Average … (GSCCLA) The average of all GSCCs in the no switch condition when compared to the low switch condition.

$$\mathrm{GSCCLA}(I) = \mathrm{GSCCA}(I,180,FRACLS)$$

Time-out benefit was operationalized as the average percentage difference in productivity caused by each switch. To find how switching affected productivity, we used switch controlled composite functions that exclusively reflected productivity after a switch (GSSCHA and GSSCLA) as a comparison to functions in which there was no switching (GSCCHA and GSCCLA). (20) Time-out benefit was calculated by determining the difference in area of two functions and dividing the difference by the number of switches. Using *TOB,* the (21) time-out benefit for high-frequency switching and (22) low-frequency switching was calculated.

(20) Time-Out-Benefit … (TOB) the percentage difference in area due to the benefit of leaving and returning to a category. A positive TOB reflects a production benefit when switching while a negative TOB reflects a production cost. For the analysis the domain for A1 and A2 was the smallest ICO of the two.

$$\mathrm{TOB}(A1, A2, P) = \frac{\dfrac{A1 - A2}{A2}}{P - 1}$$

(21) *TOB_High… (TOBH)* is the TOB from high frequency switching

$$TOBH = \mathrm{TOB}(\mathrm{A}(GSSCHA), \mathrm{A}(GSCCHA), 16)$$

(22) *TOB_Low* ... (TOBL) is the TOB from low frequency switching

$$TOBL = \mathrm{TOB}(\mathrm{A}(GSSCLA), \mathrm{A}(GSCCLA), 4)$$

The (23) average *TOB* for the high- and low-switch condition were calculated weighting *TOBH* five times as high as *TOBL*. We used this weighting because the high-switch condition required the subjects to switch five times as often as the low-switch condition and thus *TOBL* represented five times the amount of data.

(23) Weighted_Average_TOB … (WATOB) the weighted average of *TOB* across the high and low switch condition

$$WATOB = \frac{5 * TOBH + TOBL}{6}$$

Appendix C

Analysis details for Experiment 2: calculating executive-decision cost

Variables and constants used in Experiment 2:

Below are variables and constants used in Experiment 2 analysis functions.

Interval...(*I*) interval identifier, ranges from 1 to 180

Condition...(*Cond)* identifies either the (E)xecutive-decision or (Y)oked condition

Category ... (*C*) category identifier within a *Cond*, ranges from 1 to 6

Subject_Pair ... (SubP) a subject pair indicator. A SP at a given category will have a unique switch schedule, which is decided by the Executive in the SubP, ranges from 1-20.

Production_Interval_Length...(*PIL*) the number of intervals in a P specified by P, C, *Cond, SubP.*

NP ... number of production intervals specified by *SubP* and *C*

NSubP ... number of subject pairs (= 20)

P...production interval identifier, ranges from 1 to *NP*(*SubP*, *C*)

NC ... number of categories within a *Cond* (= 6)

NSS ... number of subjects at *I*, specified by *C*, *Cond*, *SubP*, ranges from 20 to 1

Analysis functions used for calculating Executive-decision Cost:

To calculate Executive-decision Cost we compared the area, not of the double-category functions of the Executive and Yoked condition but of the area of the (24) extracted single functions of the two conditions. If extracted single functions from the executive-decision condition have less area then extracted single functions from the yoked condition then there is a production cost associated with executing an autonomous switching strategy.

(24) GS_Extracted_Specific ... (GSES) a GS function that is extracted from a double-category trial which is specified by *C, S,* and *Cond*

$$\mathrm{GSES}(I,C,SubP,FRT) = \sum_{i=1}^{I} \mathrm{words\_produced}(i,C,SubP,FRT)$$

Before comparing the area of the executive and yoked executive functions, the time between a switch and the first response after a switch was removed from all of the GSES functions resulting in (25) switch-cost controlled composite functions. The rationale for this correction is that executive subjects may have endogenously switched and produced an item to the other category prior to an exogenous switch (pressing the return key). Yoked subjects only have reason to endogenously switch after being forced. At issue is production while at a patch. Equalizing switch-cost differences by removing all first response times isolates this production period with greater precision. After deriving each function we calculated the (26-28) average composite function for each condition

(25) GSES_Switch_Cost_Controlled_Specific … (GSESSCCS) a residual GS function after removing the time between every switch and the corresponding *FRAS(E)xecutive or FRAS(Y)oked.*

$$\alpha = \mathrm{FRT}(P,C,Cond,SubP) + I - PIL(P,C,Cond,SubP)$$

$$\lambda(I,P) = \begin{cases} 0 & if \quad a \le 0 \\ \mathrm{FRT}(P,C,Cond,SubP) + I + \lambda(a,P+1) & if \quad a > 0 \end{cases}$$

$$\mathrm{GSESSCCS}(I,FRT,Cond,C,SubP) = \mathrm{GSES}(\mathrm{FRT}(2,C,Cond,SubP) + I + \lambda(I,3),C,SubP,FRT)$$

$$I \in [1, NI - (\sum_{P=1}^{NP} \mathrm{FRT}(P,C,Cond,SubP))]$$

(26) GSESSCC_ Average … (GSESSCCA) the average GSESSCCS by condition specified by *FRT* and *Cond.*

$$\mathrm{GSESSCCA}(I,FRT,Cond) = \frac{\sum_{C=1}^{NC} \frac{\sum_{SubP=1}^{NSubP} \mathrm{GSESSCCS}(I,FRT,Cond,C,SubP)}{NSS(I,C,Cond,SubP)}}{NC}$$

(27) GSESSCCE_Average …(GSESSCCEA) the average GSESSCC in the Executive condition

GSESSCCY_Average …(GSESSCCYA) the average GSESSCC in the Yoked condition

As in Experiment 1, an average composite function does not represent the same number of subjects at each *I* because of the individual composite function's various domains. To keep the area comparison representative we located the interval where no fewer than half of the subjects were represented for each category pairing in both the Executive and Yoked average composite functions and then selected the smallest of these two values as the Interval Cut Off (ICO), the last *I* considered in the final area comparison. By comparing the (28, 29) area of the average composite functions in each condition we were able to calculate (30) Executive-decision Cost which represents the production of executing an autonomous switching strategy.

(28) Area_GSESSCCEA ... (AGSESSCCEA) the area of AGSESSCCE up to interval ICO

(29) Area_GSESSCCYA ... (AAGSESSCCY) the area of AGSESSCCE up to interval ICO

(30) Executive_Cost ... (EC) the area proportion cost for those in the executive-decision condition

Appendix D

Analysis details for Experiment 3: Determining the optimality of actual time distribution during dual-domain word production

Variables and constants used in Experiment 3:

Block … (*B*) is comprised of two single category trials and a double category trial, ranges from 1 to 4.

Entry … (*E*) a produced word or a switch

Group … (*G*) refers to either group A or group B. Group A's blocks are organized as double category trial, single category trial, single category trial; Group B's blocks are organized as single category trial, single category trial, double category trial. With both groups all categories are in both single category and double category trials.

NI ... number of intervals (= 720)

NS ... number of subjects in each group (= 15)

NB ...number of blocks (= 4)

NG ... number of groups (= 2)

Below are functions used in the Experiment 3 analysis

The goal of Experiment 3 is to determine how optimally subjects allocate their time between a (31)more and (32)less lucrative category. We calculated the degree of this optimality by first finding the average (33,34) proportion of time subjects should (optimal) spend on a the more lucrative category at each interval of a double category trial and then compared the optimal value to subjects (35) actual division of time between categories. The degree of optimality is the proximity of actual division to optimal division from baseline (50% on both the high and low category). If the (36) optimality index is 1 then actual and optimal division are the same. An optimality index of 0 reflects that subjects are spending an equal amount of time on both categories and that subjects show no awareness of how to properly divide their time between a more and less lucrative category to maximize production from dual domains.

(31)*high*… the single category gain functions in each category pairing that yielded the greater area specified by B, S and G

(32)*low* … the single category gain functions in each block that yielded the lesser area specified by B, S and G

(33)Optimal_I … (OI) is the amount of time to spend on the high category in a double category trial to optimize the number of words produced at each interval

$$OI(I,B,S,G) = \operatorname{argmax}_i(high(i,B,S,G) + low(I-i,B,S,G)) \quad i \in [1,I]$$

(34)Optimal_Division … (OD) is the average optimal proportion of intervals to spend on the high category across all intervals.

$$OD = \frac{\left(\sum_{I=1}^{NI}\sum_{B=1}^{NB}\sum_{S=1}^{NS}\sum_{G=1}^{NG}\frac{\mathrm{OI}(I,B,S,G)}{I}\right)}{NI*NB*NS*NG}$$

(35)Actual_Division … (AD) is the measure of how subjects actually divided their time at every entry

$$AD = \frac{\left(\sum_{S=1}^{NS}\sum_{B=1}^{NB}\sum_{G=1}^{NG}\sum_{E=1}^{NE(S,B,G)}\frac{time\_on\_high(S,B,G,E)}{time\_total(S,B,G,E)}\right)}{NS*NB*NG*\left(\sum_{S=1}^{NS}\sum_{B=1}^{NB}\sum_{G=1}^{NG}NE(S,B,G)\right)}$$

(36)*Optimality_Index … (OptI)* is the measure of how optimal subjects are at dividing their time between a high and low category. An *OptI* of 0% would occur if subjects *AD* is 50%. An OptI of 100% would occur if subjects *AD* is equal to *OD*.

$$OptI = 1 - \frac{|OD - AD|}{OD - .5}$$

Appendix E

Predicting dual-domain production

In the following section our method for predicting word production from two semantic categories is described. Our model is expected to capture the meaningful variance between Maylor et al's (2001) serial prediction and actual double-category performance, which did not account for switch cost, executive-decision cost, time-out benefit, or the time equalizing tendency.

Derivation of our prediction occurred in two stages. First, individual subjects observed single-category functions (OSs) from Experiment 3 were adjusted for "would-be" (would-be referring to factors that would have taken place if switching had occurred) switch-cost, time-out benefit, and executive-decision cost, resulting in predictor single category functions. Then these pairs of these functions were combined using the time-equalization tendency to form our final dual-domain prediction

Variables and constants used to predict dual-domain production:

Time … (*T*) identifies time, ranges from 0 to 180 seconds

Interval …(*I*) identifies the interval number, ranges from 0 to 720

Group …(*G*) identifies group, ranges from 1 to 2

Subject …(*S*) identifies subject within a group, ranges from 1 to 15

Block … (*B*) identifies trial block within a group, ranges from 1 to 4. Each trial block contains a *high* and *low* category.

Category …(*C*) identifies category from one of the single-category trials or the double-category trial within a block, ranges from 1 to 2

High_Low_PS… (*HLPS*) is a category specified by *L* (*high* or *low*) and *B*.

NI ... number of intervals (= 720)

NS ... number of subjects in each group (= 15)

NB ...number of blocks (= 4)

NG ... number of groups (= 2)

Switch_Cost ... *(SC)* refer to 11 (= 1.81)

Time_out_benefit ... (*TOB*) refer to 23 (= 1.013)

Executive-decision_Cost ... (*EC*) refer to 30 (= .058)

Optimality_Index ... (*OptI*) refer to 36 (=.10)

Analysis functions used to predict dual-domain production.

Calculating and Applying the Expected Time-out Benefit and Executive-decision Cost:

Experiment 1 demonstrated that each timeout leads to a production benefit of 1.13%, after switch-cost is controlled. To incorporate this production benefit into our prediction, we determined the amount of (37) "expected" timeout benefit by calculating the average number of switches to each category by the end of every interval (38) using double-category trial data from Experiment 3 and multiplied the number times *TOB* (1.013) Experiment 2 demonstrated that autonomous switching between two categories leads to a 5% reduction in area across a production interval, which we termed Executive-decision Cost (EC).

(37) Expected_Total_Time_Out_Benefit ... (ETTOB) expected time-out benefit by the end of *I*, specified by *C*.

$$\text{ETTOB}(T,C) = \text{average\_number\_of\_switches}(I,\text{D}(C)) * TOB$$

(38) Double ... (D) represents a gain function in a double category trial, specified by C.

Both ETTOB and *EC* affect word production, therefore the combined effect of these two corrections will move each point of an OS upwards, downwards, or not at all, depending on whether the sum of the executive cost (-) and ETTOB (+) is positive, negative, or 0. (39) ETTOB and EC were applied to each (40) OSSs synchronously by combining the two effects into one corrector value. The order that these parameters are applied would affect the final prediction and there is no theoretical reason mandating precedence of one factor over the other (while foraging, *EC* and ETTOB occur at the same time).

If $\frac{EC}{\text{ETTOB}} > 1$, then the time-out benefit will outweigh the executive-decision cost, and each ordinate value of an OSS was multiplied by a (39) JECETTOBC greater than one; if this ratio was below one than each *T* was multiplied times a JECETTOBC less than one because the *EC* cost will have outweighed the *ETTOB.*

(39) Joint_EC_ETTOB_Coefficient ... (JECETTOBC) the joint cost of EC and ETTOB. Is a coefficient that is multiplied times the ordinal value time *T* on an OSS, specified by

C.

JECETTOBC$(I,C) = 1 - EC +$ ETTOB(I,C)

(40) Observed_Single_Specific … (OSS) an observed gain function from a single category trial from Experiment 3 specified by *C* and *S*.

(41) Gain_Adjusted_JECETTOB_Single_Specific … (GAJECETTOBSS) an observed single category function adjusted for ETTOB and *EC* specified by *S* and *C*.

GAJECETTOBSS$(I,C,S) =$ OSS$(I,C,S) *$ JECETTOBC(I,C)

Calculating and Applying Switch Cost.

Experiment 1 demonstrated that each switch incurs a cost of 1.89 seconds. In this section, the method for accounting for this "would-be" switch cost to the GAJECETTOBSS functions is discussed. The end result is our final estimate of extracted gain functions from double category trials. The initial step was to determine the expected would-be switch cost by the end of each time interval by monitoring double-category trials from Experiment 3 and determining the number of switches to each category at every interval. For example, if out of 15 people, one subject switched to category *A* during the 180th interval, then this interval of category *A* is assigned an average switch number of $\frac{1}{15}$. (42) After monitoring all switches to category *A* from all 15 subjects, (43) the average number of expected switches at each of the 720 intervals was calculated. Using these data, we calculated a (44) cumulative switch cost function which defines the expected cumulative number of switches to category *A* at each *I* multiplied times the cost per switch, 1.89 seconds.

(42) Number_of_Switches_Specific … (NOSS) the number of switches to a D(C) during time *I* by subject S.

(43) Number_of_Switches_Average … (NOSA) the average number of switches to a category from a double-category trial during time *I*.

$$\text{NOSA}(C,I) = \sum_{S=1}^{NS} \text{NOSS}(D(C),S,I)$$

(44) Cumulitive_Switch_Cost_Average … (CSCA) the average cumulative amount of switch cost at the end of each *I*.

$$\text{CSCA}(C,I) = \sum_{I=1}^{NI} \text{NOSA}(C,I) * SC$$

CSCA was then used to adjust every GAJECETTOBSS so that they reflected the displacement that may have occurred if switching was present. For example, if an entry was made at second 45 (interval 181) during a single category *A* trial and the expected cost due to switching were 3 seconds after 45 seconds of production, then the adjusted timestamp would read 48 seconds. (45) This procedure was repeated for every timestamp of every GAJECETTOBSS which resulted in a new predictor functions.

(45) Predictor_Specific … (PS) an observed single category function adjusted for time-out benefit and executive-decision cost, and switch cost specified by *S* and *C*.

The average PS for each category acted as the random variable when testing the accuracy of our switch cost, time-out benefit and executive-decision cost corrections. Extracted functions, unlike the observed single category functions did not necessarily extend the full 720 intervals because they were divided between category *A* and *B* (intervals spent on A and intervals spent on B summed to 720). As a result, switches to categories were far more frequent at intervals 1 to 360 than intervals 361 to 720.

Accounting for Sub-optimal Time Allocation

Experiment 3 indicated that subjects at any one interval are demonstrating only 10% optimality in their time distribution. To determine the number of intervals subjects may have spent on the high category for every interval of double-category trials, the number of intervals that is 10% of the distance between baseline and the optimal number of intervals was calculated for each PS pair. (46) The optimal number of (i)ntervals for each *I* of each PS pair was calculated by finding the interval that leads to the highest sum for the argument high_PS(i) + low_PS(*I*-i).

(46) Optimal_Predictor_Interval_Specific … (OPIS) the optimal number of intervals to spend on the *high* PS category, specified by *B, G,* and *S*.

$$\text{OIPS}(I,B,G,S) = \quad \text{argmax}_i(\text{PS}(i,HLPS(high,B),G,S) + \text{PS}(I-i,HLPS(low,B),G,S)) \quad i \in [1,I]$$

With the knowledge of (47) how many intervals subjects would likely spend on the high category for each *I* in a PS pairing we made (48) a dual-domain gain prediction for each PS pair and used (49) the average of all these PS-pair predictions as the final prediction of dual-category performance for the 8 category pairings used in Experiment 3 .

(47) Predicted_Distribution_Intervals_Specific … (PDIS) the predicted number of intervals subjects would have spent on the *high* PS category, specified by *B*, *G*, and *S*.

$$\mathrm{PDIS}(I,B,G,S) = \arg\min_i\left(\left|(i-\frac{I}{2}) - (\mathrm{OPIS}(I,B,G,S) - \frac{I}{2}) * OptI\right|\right) \quad i \in [1,I]$$

(48) Final_Prediction_Specific … (FPS) the double category gain prediction, specified by B, G, and S.

$$\mathrm{FPS}(I,B,G,S) = \mathrm{PS}(PDIS(I,B,G,S),high,B,G,S) + \mathrm{PS}(I - PDIS(I,B,G,S),low,B,G,S)$$

(49) Final_Prediction … (FP) the average of all FPSs. When comparing FP to observed double-category gain in Experiment 3, the number of blocks across groups acted as the random variable.

$$\mathrm{FP}(I) = \frac{\sum_{G=1}^{NG}\sum_{S=1}^{NS}\sum_{B=1}^{NB}\mathrm{FP}(I,B,G,S)}{NG * NS * NB}$$