

Running head: CHINESE SELF-PACED STUDY

Item selection and self-paced study when learning Chinese pictographs: Does the region
of proximal learning hypothesis apply to unfamiliar stimuli?

Rory Murray, Psychology

Senior Thesis

Spring 2010

Supervisor:

Dr. Christopher Hertzog

Abstract

This experiment examined the order of learning when faced with a time-constrained, difficult, and unfamiliar task by using 36 Chinese-English word pairs. Eighty younger adults (18 to 25 years old) and forty older adults (60 to 80 years old) participated in this experiment. Participants were asked to make an Ease of Learning Judgment (EOL) for each Chinese character. Then, there were two distinct learning trials in which the participants studied the same 36 Chinese-English word pairs in six 2X3 grids. The younger adults had a maximum of 30 seconds of study time allowed per grid. The older adults had 60 seconds per grid. The critical question was whether individuals would focus study on easier items that have not yet been learned, as predicted by Metcalfe's region of proximal learning model, or whether they would simply study all unlearned items with equal emphasis (as predicted by the discrepancy-reduction model (DRM)). Both age groups had relatively accurate EOL's corresponding to item complexity. Young adults showed no evidence for RPL predicated effects. Older adults exhibited behavior largely consistent with RPL predictions

Introduction

Self-regulated learning involves choosing which items to study, deciding how long to study the to-be-learned material, and monitoring item difficulty (Metcalf & Kornell, 2005). Individuals must decide if they have learned an item to a degree sufficient to allow for later remembering (Thiede, 1999). Self-regulated learning, in the form of selective, goal-directed study behavior, has been shown to increase recall performance especially for older adults (Canestrari, 1963; Hulicka & Wheeler 1976). The two most popular self-regulated-learning theories that pertain to the order of item selection when under constraints — in particular, time constraints — are Metcalfe's (2000) region of proximal learning (RPL) model and the discrepancy reduction model (DRM; Nelson, & Leonesio, 1988). Metcalfe's (2000) region of proximal learning (RPL) model suggests that individuals will allocate more effort toward studying and reviewing items they regard as potentially learnable, given their current level of learning. It predicts, then that people will selectively study the easier items they have not yet mastered. DRM predicts that individuals will allocate more effort toward studying and reviewing the items that are less well known. DRM can be interpreted as involving study of all unlearned items with equal emphasis (Dunlosky & Hertzog, 1998), or as placing greater emphasis on unlearned items that are more difficult.

Previous studies have utilized stimuli such as Spanish-English word pairs, with difficulty largely determined by the similarity of the Spanish word to its English equivalent. This procedure raises potential issues as to whether the results are stimulus-driven (Metcalf, 2002). Preliminary results from the National Foreign Language survey show that in 1997, 31% of all American elementary schools offered foreign languages

and 79% of these offered Spanish. Furthermore, Spanish and English share a Latin based root resulting in many similarities in base words and other linguistic characteristics. When exposed to an item with prior knowledge, participants are expected to study as predicted in the DRM model (Verkoeijen, Rikers, & Schmidt, 2005).

This study used Chinese-English word pairs. The Chinese characters are unfamiliar to most university students. This allows for a stronger test of the competing predictions of the RPL and DRM models by testing individuals with unfamiliar stimuli. Furthermore, the stimuli can be scaled with respect to their apparent difficulty. Chinese pictographs were used, with the number of strokes comprising a character used as an a priori measure of the difficulty of the Chinese character (Liu, Hao, Tan, & Weekes, 2008; Qian, Reinking, & Yang 1994). Intact (standard) Chinese-English items were used as stimuli for younger and older adults. In order to determine whether pictograph complexity was the major influence on perceived difficulty, a second young group was run, using random assignment of English words to pictographs. Stimuli in this condition would counteract English-Chinese interaction effects that have been previously been found. (Verhaeghen, Palfai, & Johnson 2006).

A working hypothesis of this study was that pictograph complexity would be directly related to perceived and actual item difficulty. Participants gave Ease of Learning Judgements (EOLs) for all Chinese items. We could then determine whether initial item selection and study time allocation were influenced by complexity and perceived complexity. Further analysis would be conducted to assess if character complexity was inversely related to the percentage of items recalled corresponding to the hypothesis that participants' EOL ratings would indicate more complex items as more

difficult to learn. Assuming that the character complexity hypothesis holds, in Trial 2 RPL predicts that people would favor the less complex items they did not recall at Trial 1, whereas DRM would predict participants would study all unknown items (or even favor the most difficult items they did not recall at Trial 1). Least known items would be defined as those items missed in Trial 1.

Age related differences in self-paced study behaviors are also of interest. It has been shown that older adults use an RPL learning style under similar conditions as younger adults (Price, Hertzog, & Dunlosky, in press). Older adults in these studies allocated more time to the difficult items, but showed a preference for selecting nominally easy items first.

However, there were several issues with the studies by Price et al. (in press), which based its methodology directly on earlier studies of RPL (e.g., Metcalfe, 2002; Kornell & Metcalfe, 2005). Given that younger adults are more likely to have had formal exposure to Spanish, age related performance differences may not be entirely accurate since 33% of American elementary schools offer Spanish courses (National Foreign Language Survey 2009). Furthermore, participants were informed in advance about items' normative difficulty (grouped into sets of easy, medium and hard items). Finally, the grids of items used for study were always ordered, from left to right, as easy, medium, and difficult. Thus, grid location was confounded with difficulty.

This study addressed all these issues. By using Chinese, the possibility for prior knowledge across all age groups can be better controlled as only 3% of American elementary schools teaching foreign language offered Chinese (National Foreign Language Survey 2009). The Chinese-English pairs and the two trial experimental set up

tested the hypothesis that age related differences in self paced study are stimulus driven. Further, it was hypothesized that there would be no age related differences in study behaviors and subjective difficulty interpretation. It was expected that the older adults will not perform as accurately as the younger adults during the recall task, as found by Verhaeghen et al., (2006) and by Price et al. (in press). To address earlier methodological limitations, items were not labeled with respect to a priori difficulty, and the location of items in grids was randomly assigned. This procedure with Chinese pictographs made it likely that item selection behavior and study time allocation based on perceived (or actual) difficulty could be detected by using pictograph complexity as an independent variable.

Method

Participants

One-hundred-six young adult participants between the ages of 18 and 25 and forty-four older adult participants between the ages of 60 and 80 participated. The younger adults were Georgia Tech students who received compensation in the form of 1.5 hours of course credit. The older adults consisted of people from the Atlanta area, who received an honorarium of \$20 for their participation. Participants were pre-screened to ensure they have no background or formal training in Chinese or any East Asian language with written Chinese influences such as Japanese.

To help evaluate the relationship between the complexity of Chinese pictographs and both perceived and actual difficulty in learning these characters, forty young adults were randomly assigned to a control condition in which the Chinese-English words were

randomly combined to form arbitrary pairings, rather than using the direct translation of the Chinese-English as vocabulary pairs. One hypothesis of this study is that complexity of the pictographs influences selection behavior and perceived item difficulty. The random re-mapping condition helped to determine whether the meaning of the English word also influences perceived difficulty, selection behaviors, and learning. It is possible that the difficulty of learning the pictograph-word association is determined by not only by the Chinese character complexity but also by the difficulty of learning the English equivalent (Verhaeghen et al., 2006).

Materials

Thirty-nine Chinese-English word pairs were used as stimuli. One third of these Chinese words consisted of one to four pen strokes and were classified as *easy* as shown in Figure 1 (top). Thirteen words that consist of five to nine strokes were classified as *medium* stimuli. (Figure 1 - middle) The remaining thirteen items consisted of ten to fifteen pen strokes and were classified as *hard* stimuli (Figure 1 - bottom). One word pair from each difficulty level was used in the instruction screens to familiarize participants with what is meant by *easy*, *medium*, and *hard* items. The other 36 Chinese-English word pairs were presented in six different 2 X 3 grids.

A new computer-based testing program was developed using the Java programming language with young adult pilot testing conducted in July 2009. This program presents instructions and stimuli on a computer screen and participant responses are inputted using a mouse and keyboard. The program records participants' responses to all queries including the order in which items are (re)selected, how much time is allocated to each item, and the grid in which the items are associated. The program was

run on nearly identical Windows XP based machines running Java 6 update 14 Standard Edition.

Procedure

When participants first arrived they were asked, in chronological order, to sign the consent form; to fill out a personal data sheet; to answer the Memory Control Inventory (MCI); to complete the Advanced Vocabulary Test (AVT; Ekstrom, French, & Harman, 1976), an English vocabulary test; to complete the Pattern Comparison task which measures perceptual speed (Salthouse, 1996); and to complete the Listening Span task (Salthouse & Babcock, 1991) a measure for working memory capacity. Participants then started the computer-based task developed for this study. The program was ordered as follows: ease of learning (EOL) phase, Trial 1 learning grid phase, Trial 1 recall phase, Trial 2 learning grid phase, and Trial 2 recall phase.

In the EOL phase, the 36 to-be-studied Chinese characters were presented to participants in a unique individually randomized order and were not accompanied by their respective English counterparts. Participants were then asked to rate each item on a 1-9 Likert scale (1 being *easy* and 9 being *hard*), indicating how difficult it would be to learn the English meaning (Figure 2). After completing the EOL phase, participants began the Trial 1 learning phase. During the learning phase, all 36 Chinese characters were randomly assigned to be presented in six 2x3 Study Grids containing 6 Chinese characters (a selection of 2 *easy*, 2 *medium*, 2 *hard*) and a 30 second timer per grid (Figure 3). The English meaning of the Chinese character was only displayed after a participant had selected that item on the Study Grid. Upon selection, the timer would begin counting down to 0s until the participant had indicated they had finished studying

that item. Then, the English word would be removed from their view, and the timer would pause (Figure 4). The next Study Grid appeared when the timer reached 0s or the participant chose to move on. After having the opportunity to study all 36 word pairs, the recall task began wherein the 36 Chinese characters were again presented in a randomized order and participants were asked to type their English meanings. There was no recall accuracy feedback in either trial, and participants were not able to continue without entering an answer (Figure 5). This procedure was repeated in Trial 2 with the same 36 words. After completing the Chinese computer program, participants completed a questionnaire regarding strategies and self-rated performance during the computer-based task. Participants were debriefed upon completion of the experiment. There were no differences in the computer program for the randomized condition beyond the English-Chinese word pairings. Those in the randomized condition were verbally informed that the word pairs they studied were randomly paired only at the time of debriefing.

Results

Multilevel regression models (in SPSS 17.02, Field 2009) were used to analyze the data relevant to each dependent variable. This procedure allowed us to evaluate within-person regression functions relating stimulus complexity to other variables (see Hoffman & Rovine, 2007).

Ease of Learning (EOL) Ratings

I tested for differences between the three groups (Standard young, random young, and older adults) in relations of pictograph complexity to EOLs by including a centered

variable for character complexity (1 through 36 centered based on subtracting 18.5 from all values) in the model.

2 alternative models were tested to see if random effects for intercepts and random effects for complexity slopes were needed to account for variation in EOLs. Figure 6 shows the results of the model comparisons, using differences in -2LL to generate χ^2 tests of the null hypothesis of no random effects. The test for random effects in complexity slopes allowed for a variance in slopes and a covariance of intercepts and slopes.

Both tests were reliable, indicating individual differences in mean EOLs and in the effects of complexity on EOLs. Figure 7 presents the estimated fixed effects parameters on the EOLs from the model including random effects in intercepts and slopes.

There was a significant effect of condition $F(1, 5476) = 80.133, p < .001$. As seen in Figure 7 there was little difference between the intercepts of the Standard and Random conditions, indicating no effect of the random mappings on EOLs. This outcome supports the hypothesis that the number of line segments in the pictographs, not features of the English words the pictographs map to, are used to make EOL's. Given centering of complexity, the intercepts can be interpreted as group mean EOLs. The older adults on average had a higher mean EOL rating, indicating they perceived a greater overall difficulty in learning Chinese pictographs.

Complexity had a strong effect for all three conditions, $F(1, 5476) = 3891.83, p < .001$. There was also a reliable condition X complexity interaction, $F(1, 5476) = 33.976, p < .001$. The younger adults in both the random and standard conditions showed higher

sensitivity of rated ease of learning to complexity than older adults. However, the complexity effects were robust in all three groups. The two groups of younger adults yielded no difference in complexity slopes. Again, this outcome suggests that only the pictographs themselves, not their English meanings, influenced EOLs. (Figure 9)

Items Not Selected for Study

I first evaluated whether an item was studied or not studied, and whether choosing to study an item was related to the pictograph complexity. Figure 10 reports the overall frequency of items not selected for study. The younger adult conditions show little difference in the number of items not studied at all. Older adults appeared to skip more items on average than younger adults.

Items not selected for study were of interest to determine if there were any DRM or RPL like behaviors in not selecting an item. Another Multilevel regression model was run to predict whether an item would be skipped in Trial 1 as a function of complexity factored by condition type. The dependent variable was scaled as a binary outcome (0 = not studied, 1 = studied). Hence the fitted means in each group can be interpreted as the mean proportion of items studied (or not skipped) during Trial 1 for each group.

There was a significant effect of condition, $F(1, 5476) = 127.9, p < .001$. Older adults demonstrated a higher propensity to skip items, studying only about 77% of them. There was no reliable difference in the proportion of items studied by the Standard and Random young adult conditions. Complexity had a reliable effect, with slopes being reliably greater than 0, $F(1, 5476) = 55.88, p < .001$. There was also a reliable condition X complexity interaction effect, $F(1, 5476) = 6.91, p < .001$. The older adults were more

averse to selecting difficult items than the younger adults, as reflected in their shallower complexity slope.

Item Selection Order

A multilevel regression model was used to test differences between the three conditions (standard, random and older) for the order in which items were first studied. Excluded from this analysis were items that were not selected for study. By excluding items not studied, a more precise analysis of study selection behavior can be performed. The dependent variable reflects the order of item selection within each grid of 6 items.

It was important to determine whether initial item selection was based simply on grid location or on pictograph complexity, or both variables. Grid location was recorded as 1 being the top left item in a grid 3 being the top right item in a grid and so forth with 6 being the bottom right item in a grid. This coding represented the expected left to right selection pattern reflecting normative English reading patterns. Because items were randomly assigned to grid locations, it was possible to simultaneously estimate both item complexity effects and location effects.

The dependent variable was the order in which the variable was selected. It could differ from 1-6 if an item were restudied before another item was first studied. Separate models were created for Trial 1 and Trial 2. Both of these models included a centered variable for character complexity and a centered variable for grid location (1 through 6 centered based on subtracting 3.5 from all values). Condition was also an independent variable (standard, random, or older adult group). The Trial 2 regression also included Trial 1 accuracy as a factor to assist with hypotheses concerning RPL and DRM (which concern study of as-yet-unlearned items).

Condition effects were not significant for item selection with $F(1, 4838) = 1.955$, $p = .142$. Given the nature of the order dependent variable, this simply reflects the fact that the groups did not differ in the average amount of interpolated restudy before other items were first studied. The complexity and grid location effects were of greater interest.

Complexity had a large main effect with $F(1, 4838) = 649.571$, $p = .001$ and grid location also had a large main effect with $F(1, 4838) = 1249.942$, $p = .001$. However there is an interaction effect of condition with complexity with $F(1, 4838) = 13.368$, $p = .001$, and between condition and grid location with $F(1, 4838) = 13.553$, $p = .001$.

The older adults were more averse to selecting difficult items first. All conditions were influenced by grid location as a function of reading order with a strong regression coefficient. However, the complexity effects showed that standard reading order was not the only influence on item selection; pictograph complexity had an independent influence. The younger adults are more likely to select items based on reading order than the older adults.

The Trial 2 results were similar to Trial 1. However, the use of Trial 1 accuracy as a variable meant that remaining effects were evaluating influences on changes in selection behavior at Trial 2. Accuracy in Trial has a reliable effect with $F(1, 4875) = 39.051$, $p < .000$. There was a strong complexity effect with $F(1, 4875) = 509.314$, $p < 0.001$. Likewise grid location continued to be important, $F(1, 4875) = 645.056$, $p < 0.001$. There was also a condition effect, $F(1, 4875) = 10.814$, $p < 0.001$ and interactions between condition and the other aforementioned variables condition X complexity $F(1, 4875) = 31.291$, $p < .001$, condition X trial 1 accuracy $F(1, 4875) = 5.682$, $p < .01$, and

condition X grid location $F(1, 4875) = 7.309, p < .001$. There was also an interaction between complexity and grid location, $F(1, 4875) = 5.275, p < .05$, and a three way interaction of condition X complexity X trial 1 accuracy $F(1, 4875) = 3.300, p < .05$. These results suggest that all three groups on average selected first items which they missed in trial one. However, older adults were less likely to do so for the complex items. There is a complexity interaction effect where participants are more likely to select for study items that are complex and missed in trial 1.

Total Study Time

A multilevel regression model was used to test differences between the three conditions (standard, random and older) in for total study time allocation. Separate models were created for Trial 1 and Trial 2. Both of these models included a centered variable for character complexity and factored by condition either Standard, random, or older adult. The Trial 2 regression also included Trial 1 accuracy as a factor.

For Trial 1, there was no significant effect of condition $F(1, 5476) = 0.488, p < 0.614$. However there was a condition X complexity interaction, $F(1, 5476) = 23.128, p < .001$. Complexity was also significant, $F(1, 5476) = 32.245, p < .001$. These results indicate that older adults on average allocated less time to more complex items than the other two groups. The younger adults show little differentiation in study time allocation in relation to difficulty.

For Trial 2, every factor except condition was significant in this model, $F(1, 5476) = 1.27, p = 0.293$. Accuracy in Trial 1 was a strong predictor, $F(1, 5476) = 100.324, p < .001$. The regression coefficients for Trial 1 accuracy point to a negative

relationship between accuracy in Trial 1 and study time allocation in Trial 2. This effect indicates that people spent less time studying items they recalled at Trial 1. This effect is seen across all conditions including the older adults. There was also a complexity effect, $F(1, 5476) = 25.181, p < .001$. This relationship was complicated by the condition complexity interaction, $F(1, 5476) = 40.27, p = 0.001$. The effect of complexity on study time was only seen in the older adult condition. It showed that the older adults continued to allocate less study time to the more complex items. Trial 2 study time allocation seems to be in a DRM like fashion in the younger adults, with lesser known items as determined by trial 1 accuracy allocated more study time than known items across all conditions. For older adults the effect of prior recall was also observed, but it was constrained by the tendency to spend less time studying the more difficult items, consistent with an RPL effect.

Recall Accuracy

A multilevel regression was performed for recall accuracy where accuracy was a 0 to 1 variable incorrect and correct respectively. This model included complexity as a predictor with group as a differentiating variable. Trial 2 included accuracy in Trial 1 as an additional factor. Items that were not selected for study were excluded from this analysis. Complexity was a centered variable so all results are at the mean value of 18.5 in character complexity.

There was a significant effect of condition, $F(1, 4838) = 108.24, p < 0.001$. Complexity was also significant, $F(1, 4838) = 353.65, p < .001$. There was also a condition X complexity interaction, $F(1, 4838) = 5.63, p < .01$. The centered intercepts

differed between the older and two younger adults, with the younger adults have a higher recall on average than the older adults. Complexity is a centered variable so these are the estimated percent recalled of the medium items. There was also an effect of complexity, with the more complex items have a lower recall rate than the less complex items. These results can be observed visually in Figure 19 at the item level. Older adults had a shallower complexity slope, indicating that they showed a lower probability of recalling easy items. Item by item analysis of characters was not conducted, but aggregate results can be seen in Figure 19.

For Trial 2 accuracy, the largest effect was seen for previous accuracy, $F(1, 5475) = 1354, p < .001$. There was an effect for condition, $F(1, 5475) = 245.16, p < .001$. Complexity also influenced recall, $F(1, 5475) = 209.84, p < .001$. There was also a condition X accuracy interaction, $F(1, 5475) = 28.38, p < .001$, and a complexity X accuracy interaction, $F(1, 5475) = 36, p < .001$. These results can be observed visually in Figure 21 at the item level. Item by item analysis of characters was not conducted.

Discussion

Standard Versus Random Conditions

No differences were observed between the young standard and young random conditions in EOL ratings, recall accuracy, selection order, or study time allocation. The possible interaction between the Chinese-English word pairs was not observed in any of these measures. The primary driver of perceived difficulty was character complexity as recorded from the EOL ratings (Figure 8). Likewise the character complexity effect was

present in recall with the less complex items having higher recall rates on average than the more complex items for both the young standard and the young random conditions. This experiment did not look at Chinese character item level interactions so it is possible that while there were no differences on average between the young standard and young random groups that there are item level differences. Further discussion will refer to both groups as young adults.

Young Adult DRM

Young adult study selection was primarily influenced by item location in Trial 1. There was a statistically significant effect for avoidance in selection of complex items. Likewise, for study order there was an effect for selecting less complex items earlier, but this effect was very small especially in light of the large location effects. Further, complexity has a statistically significant but very small effect on study time allocation. Young adults on average allocate nearly the same amount of time per item regardless of complexity (Figure 16).

In Trial 2, young adult item selection was still primarily influenced by item location and there was still a slight complexity aversion. However, the overall effect for younger adults was to select earlier in Trial 2 items that had not been recalled in Trial 1. Study time allocation was also higher for items missed in Trial 1, with no complexity effects. These results support a DRM hypothesis for younger adults at Trial 2. Further Trial 1 behavior more closely follows a DRM hypothesis than RPL. Selection and allocation in Trial 1 was nearly uniform regardless of complexity. This effect occurred despite EOL ratings indicating that young adults perceive that the more complex items

would be more difficult to learn. It may be possible that young adults are dropping easier items from study too soon (Pyc & Rawson 2009).

Prior research with young adults such as Metcalfe (2000) and subsequent studies that have reported RPL effects in young adults have used study grids with an easy to hard left to right item presentation order. In the present study, there were strong effects of location based on a left to right reading order. The previous research confounded the effects of difficulty with the effects of presentation order. The present study also utilized the Chinese character stimuli which allowed for character complexity to be used in analysis instead of the nominal easy, medium and hard of prior Spanish based studies. The complexity variable allowed for difficulty to be treated as a continuous variable with many degrees of freedom. It is possible that if item presentation order is controlled, such as in this experiment, that the effects of item difficulty in selection may not be significant for Spanish vocabulary studies.

Older Adults RPL

Are the older adults engaging in optimal study behavior for this task? First, older adults appeared to understand that item complexity is related to the difficulty of learning the items. The older adults showed less sensitivity of EOLs to complexity than younger adults. Nevertheless, these results indicate that the older adults are able to perceive that there are differences in item difficulties as a function of complexity.

These EOL effects are interesting in light of older adults' avoidance of difficult items. In Trial 1, older adults show a propensity to not select the most difficult items for study. That is, the items they did not select to study tended to be the more complex

items, supporting the RPL hypothesis. Older adults also preferred to select first less complex items, above and beyond the strong effect of a left to right reading order. Study time allocation for older adults showed more allocated time to less complex items in both Trial 1 and Trial 2 again supporting the RPL hypothesis. However, older adults did allocate more time to items missed in Trial 1. In sum, whereas younger adults appeared to select and study items in a DRM-fashion, older adults definitely showed an aversion to the more complex items (Figure 13). Although all of these effects are significant, older adults were not just studying items in their region of proximal learning. They still selected more complex items are still selected for study and the differences in study time allocation between the older adults and the young adults are small when controlling for items not selected for study.

Indeed, the optimal older adult study behavior might be more RPL-like than their actual behavior. Recall after Trial 1 was substantially lower for the older adults than for the younger adults. These results were found despite the fact that older adults had been allocated 60 seconds (compared to the 30 seconds per study grid given young adults). There may have been a shock effect of being asked to learn the Chinese characters. The older adults may have difficulty attempting to learn an entirely unfamiliar stimulus when compared to the younger adults. Despite both all participants being screened for prior training in an East Asian language, there may also have been some cohort effects where the older adults have had less exposure to foreign stimuli such as Chinese than the younger adults, but these effects are difficult to quantify and were not measured in this study.

Future Directions and Applications

When presented with an unfamiliar stimuli, as in this study, young adults exhibit a DRM style of learning. Under the same conditions older adults exhibit an RPL style of learning. Future experiments should look examine if these learning strategies are due to individual differences within each age group. There may be young adults who exhibited RPL learning strategies and older adults who exhibited DRM learning strategies. It may also be the case that level of education affects self-regulated study behaviors. The older adults were selected based on having a high school education level or above; whereas the young adults were Georgia Tech undergraduates. Future studies should see if a DRM style of learning is correlated to level of education. It would also be interesting to see if there is a continuous shift from DRM to RPL over the adult life span. The present research indicates that there may be a shift in strategy with time starting with using a DRM strategy as a young adult and shifting to an RPL strategy with advanced age.

Another key future research area is if there learning strategies beyond selection order and study time allocation that enhance learning in self-paced situations? It is possible that participants, in particular young adult participants, spontaneously engage in self testing behaviors during self paced study (Murphy, Schmitt, Caruso, & Sanders, 1987). The fact that young adults allocate similar amounts of time to items regardless of complexity may indicate that young adults are engaging in self testing behaviors. Older adults may have lower recall accuracy and have complex item aversion on average due not exhibiting self testing behavior as frequently. The present research does not have a reliable way to definitively prove that a participant was self testing or another seemingly random selection pattern. Future research using eye-tracking or think aloud procedures

may be able to determine if participants are self testing and if there are any age related differences.

This research may be useful for any self-paced study situation such as online classes and language training. The age related differences may help in designing tools to aid in learning for older adults. It is clear that older adults favor an RPL learning style and that an RPL learning style may be optimal for studying difficult material for older adults. As was demonstrated by Nelson, Dunlosky, Graf, & Narens (1994) with Swahili-English word pairs, it is possible to shift people's learning behavior to a more optimal learning style, in their case, the more optimal learning style was DRM based. Techniques used to create an RPL learning style shift and ways to enhance these shifts -- such as only allowing participants to begin studying more complex items after mastering less complex items -- may increase overall recall accuracy for older adults.

References

- Canestrari, R. (1963). Paced and self-paced learning in young and elderly adults. *Journal of Gerontology, 18*, 165-168.
- Connor, L., Dunlosky, J., & Hertzog, C. (1997). Age-related differences in absolute but not relative metamemory accuracy. *Psychology and Aging, 12*(1), 50-71.
- Ekstrom, R. B., French, J. W. and Harman, H. H. (1976) *Kit of factor-referenced cognitive tests*, Educational Testing Service, Princeton.
- Field, A. (2009). *Discovering Statistics Using SPSS Third Edition*. Thousand Oaks California, SAGE Publications Ltd.
- Hines, J., Touron, D., & Hertzog, C. (2009). Metacognitive influences on study time allocation in an associative recognition task: An analysis of adult age differences. *Psychology and Aging, 24*(2), 462-475.
- Hulicka, I., & Wheeler, D. (1976). Recall scores of old and young people as a function of registration intervals. *Educational Gerontology, 1*(4), 361-372.
- Liu, Y., Hao, M., Shu, H., Tan, L., & Weekes, B. (2008). Age-of-acquisition effects on oral reading in Chinese. *Psychonomic Bulletin & Review, 15*(2), 344-350.
- Metcalf, J., & Kornell, N. (2005). A region of proximal learning model of study time allocation. *Journal of Memory and Language, 52*(4), 463-477. Retrieved November 25, 2008,
- Metcalf, J. (2002). Is Study Time Allocated Selectively to a Region of Proximal Learning?. *Journal of Experimental Psychology: General, 131*(3), 349-363.
- Murphy, M., Schmitt, F., Caruso, M., & Sanders, R. (1987). Metamemory in older adults: The role of monitoring in serial recall. *Psychology and Aging, 2*(4), 331-339.

- National Foreign Language Survey. (2009). National K - 12 Foreign Language Survey .
 In Center for applied Linguistics. Retrieved September, 20, 2009, from
<http://www.cal.org/flsurvey/prelimbrochure08.pdf>.
- Nelson, T., Dunlosky, J., Graf, A., & Narens, L. (1994). Utilization of metacognitive judgments in the allocation of study during multitrial learning. *Psychological Science*, 5(4), 207-213.
- Nelson, T., & Leonesio, R. (1988). Allocation of self-paced study time and the 'labor-in-vain effect.'. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(4), 676-686
- Pyc, M., & Rawson, K. (2009). Testing the retrieval effort hypothesis: Does greater difficulty correctly recalling information lead to higher levels of memory?. *Journal of Memory and Language*, 60(4), 437-447.
- Price, J. (2009). Examining metacognitive control: Are there age-related differences in item selection during self-paced study?
- Qian, G., Reinking, D., & Yang, R. (1994). The effects of character complexity on recognizing Chinese characters. *Contemporary Educational Psychology*, 19(2), 155-166.
- Salthouse, T. A., & Babcock, (1991). Decomposing adult age differences in working memory. *Developmental Psychology*, 27, 763-777.
- Son, L., & Metcalfe, J. (2000). Metacognitive and control strategies in study-time allocation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(1), 204-221.

- Thiede, K. (1999). The importance of monitoring and self-regulation during multitrial learning. *Psychonomic Bulletin & Review*, 6(4), 662-667.
- Thiede, K., & Dunlosky, J. (1999). Toward a general model of self-regulated study: An analysis of selection of items for study and self-paced study time. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(4), 1024-1037.
- Xiping, L., Weihai, T., & Ge, F. (2004). The Relationship Between Accuracy of Ease of Learning Judgment and Memory Performance. *Psychological Science (China)*, 27(1), 111-113.
- Verhaeghen, P., Palfai, T., & Johnson, M. (2006). Verbal labeling as an assimilation mnemonic for abstract visual stimuli: The sample case of recognition memory for Chinese characters. *Memory & Cognition*, 34(4), 795-803.
- Verkoeijen, P., Rikers, R., & Schmidt, H. (2005). The Effects of Prior Knowledge on Study-Time Allocation and Free Recall: Investigating the Discrepancy Reduction Model. *Journal of Psychology: Interdisciplinary and Applied*.

Figure 1

Easy

Character	Word Number	English	Strokes
人	0	PERSON	2
匕	1	SPOON	2
刀	2	KNIFE	2
山	3	MOUNTAIN	3
小	4	SMALL	3
中	5	MIDDLE	4
文	6	LANGUAGE	4
云	7	CLOUD	4
天	8	SKY	4
不	9	NO	4
日	10	SUN	4
王	11	KING	4

Medium

Character	Word number	English	Strokes
们	12	DOOR	5
北	13	NORTH	5
好	14	GOOD	6
在	15	NOW	6
同	16	SAME	6
合	17	CLOSE	6
光	18	LIGHT	6
各	19	SEPARATE	6

学	20	STUDY	8
性	21	NATURE	8
说	22	EXPLAIN	9
活	23	ALIVE	9

Hard

Character	Word number	English	Strokes
家	24	HOME	10
桃	25	PEACH	10
贼	26	THIEF	10
黄	27	YELLOW	11
猜	28	GUESS	11
剩	29	REMAIN	12
裂	30	CRACK	12
圆	31	CIRCLE	13
解	32	LOOSEN	13
搬	33	MOVE	13
满	34	FULL	13
幕	35	GRAVE	14

Chinese-English word pairs used in this study. This list was created explicitly for use in this and Dr. Price's study. In all cases, similarity was avoided if at all possible, the word pair selection criteria was as follows: stroke count, similarity to other Chinese characters on the list, similarity to other English words on the list, and similarity between Chinese characters of pinyin (the a roman letter form of Chinese primarily used to aid in

pronunciation which is not shown here). Note that words 0-11 are *easy*, 12-23 are *medium* and 24-36 are *hard*. The difficulties were determined by the number of strokes with 2-4 being easy, 5-9 being hard, and 11-14 being hard.

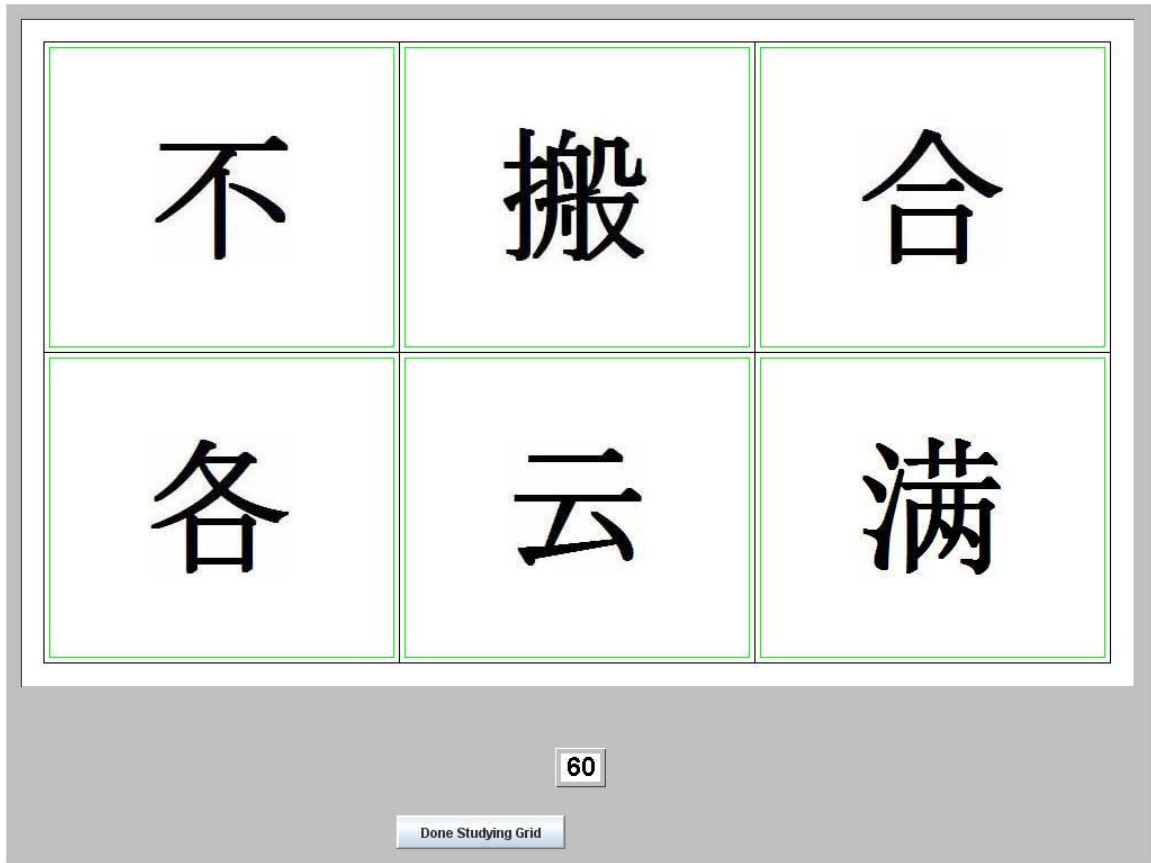
Figure 2



The screenshot displays a large white rectangular area in the center containing the Chinese character '不' (bù). Below this area is a gray control panel. At the top of the panel, the text reads: "Rate how easy or difficult this Chinese word will be to learn then click Continue." Below this text is a horizontal line with nine radio buttons numbered 1 through 9. Underneath the numbers, the word "Easy" is positioned under '1', "Medium" under '5', and "Hard" under '9'. A "Continue" button is located at the bottom center of the panel.

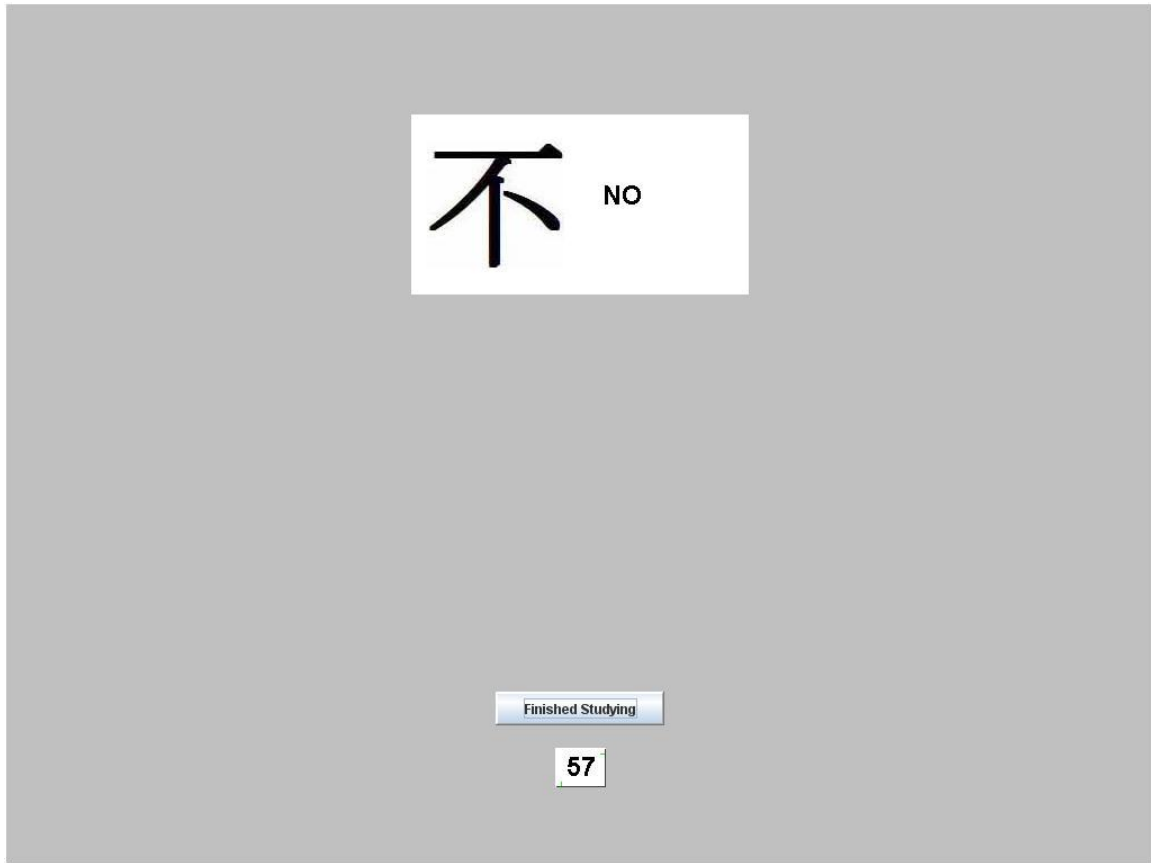
This is a screenshot from the computer program showing the EOL section. Ratings were on a 1-9 scale with 1 being *easy* and 9 being hard. Participants made this rating at the beginning of the experiment for all 36 Chinese characters. There was a 3 second delay in which the character was shown to prevent rapid screen skipping.

Figure 3



There were 2 words from each difficulty level with a random presentation order within the 2X3 grid. The countdown timer was not active on this screen. Participants were able to choose any of these six words for study. The green box indicates that a word has not been studied in this trial. After a word has been selected the green box turned to red and the timer reflected the time spent studying the word.

Figure 4



This is a screenshot after a word has been selected for study. The Chinese character was displayed with its English meaning. The countdown timer was active on the bottom of the screen until the participant pressed the ‘Finished Studying’ button.

Figure 5



The image shows a screenshot of a self-paced study interface. The main area is a large white rectangle with a gray border, containing the Chinese character '不' (bù) in the center. Below this area is a gray bar containing a text input field. The text 'Enter your answer here:' is positioned to the left of the input field. Below the input field, there is a line of smaller text: 'If what you are typing does not appear click in the text box then try again.' At the bottom center of the gray bar is a button labeled 'Continue'.

Participants were shown the Chinese character and asked to recall the English meaning on a screen that looks like this picture. Participants were not allowed to move to the next screen until an answer is entered.

Figure 6

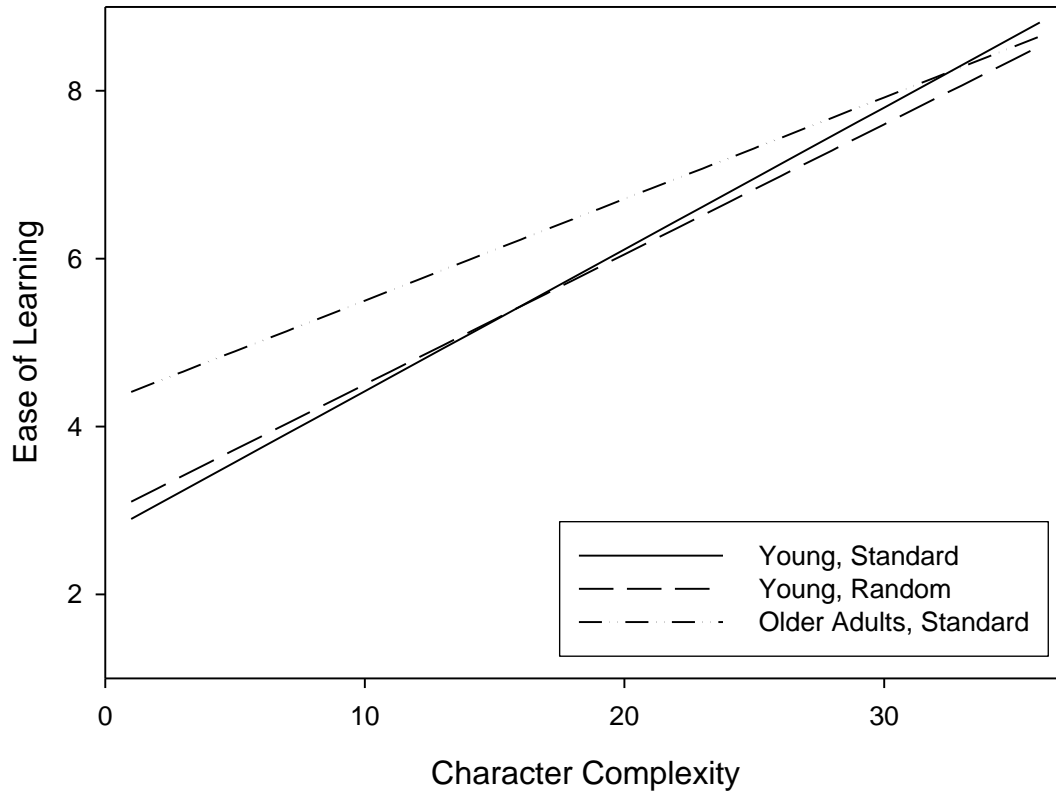
<i>Comparison of the hierarchical tested models based on LR Chi-SQ tests</i>			
Model	-2 Log Likelihood	χ^2 Differences	Degrees of Freedom
No Random Effects	22108.895		
Random Effects Intercept	20856.418	1252.44***	1
Random Effects	19999.082	857.34***	2
<i>Note *** p < .001</i>			

Figure 7

<i>Fitted Regression Equations for Effects of Complexity on EOLs</i>			
	Standard	Random	Older Adult
Intercept	5.866 (0.194)	5.822 (0.192)	6.532 (0.143)
Complexity	0.169 (0.014)	0.155 (0.014)	0.121 (0.010)
<i>Note</i> All significance calculated at the $p < .05$ level			

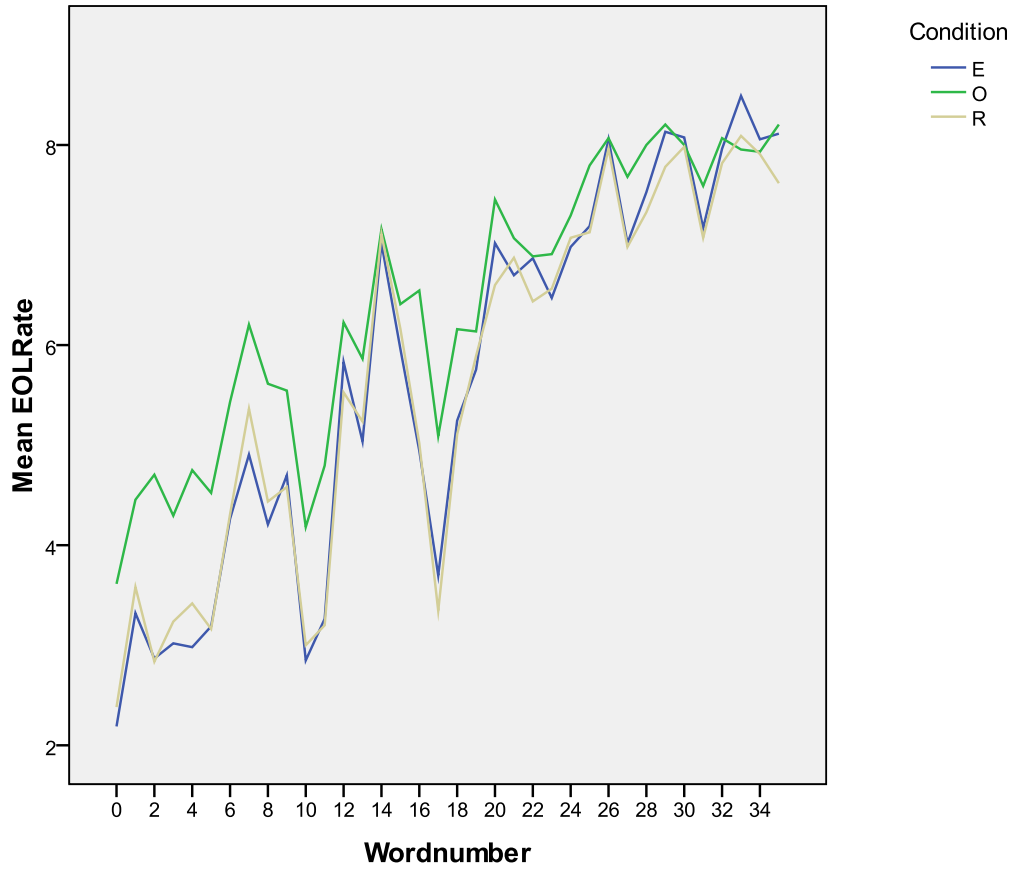
Figure 8

Regression Lines For EOL's as a Function of Character Complexity by Condition



Note There is little difference between the young adult groups.

Figure 9



Mean Ease of Learning Judgment by the word number for each condition (E standard young adult, O older adult, R random young adult). There appears to be little difference between age groups or condition.

Figure 10

<i>Comparison of the frequency of items not selected for study</i>			
Model	Trial 1 Skipped	Trial 2 Skipped	Total Selections
Standard	138	134	1908
Random	151	150	1984
Older Adult	349	317	1584
<i>Note</i>			

Figure 11

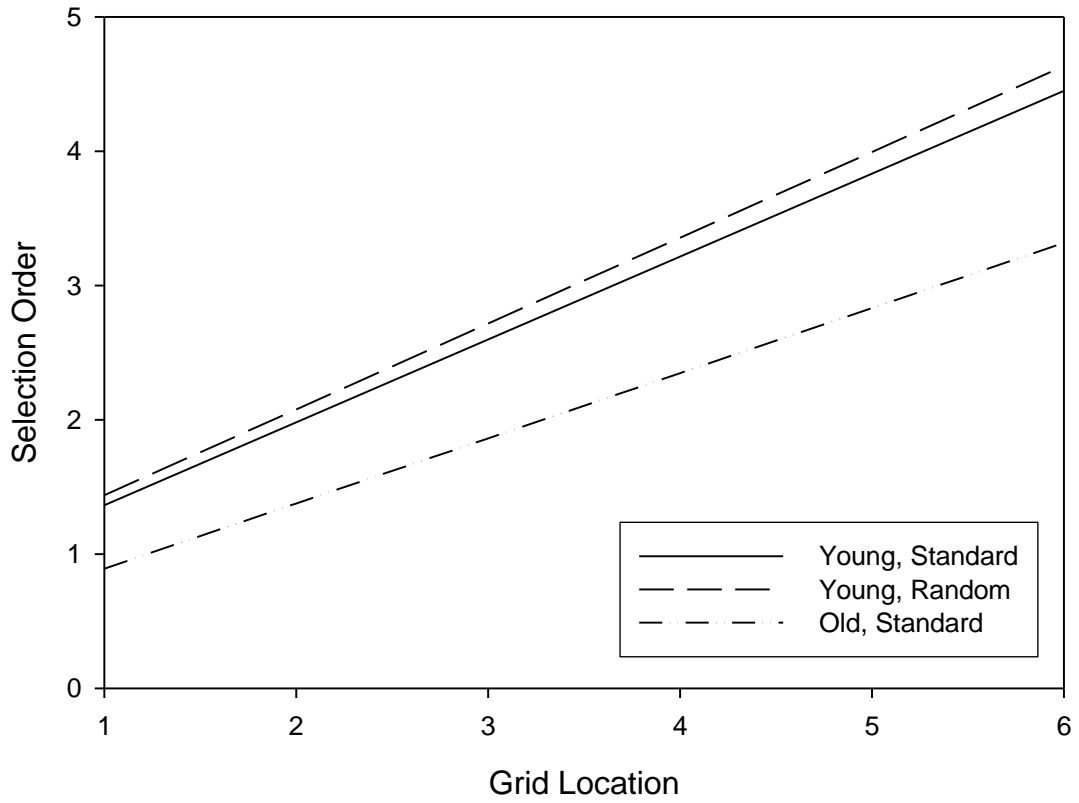
<i>Fitted Proportion of Items Selected for Study and the Slopes of Complexity</i>			
	Standard	Random	Older Adult
Mean items studied	0.925 (0.01)	0.891 (0.01)	0.774 (0.007)
Complexity	-0.002 (0.001)	-0.002 (0.001)	-0.005 (0.007)
<i>Note</i> All significance calculated at the $p = 0.05$ level			

Figure 12

<i>Significant Fixed Effects Per each Condition in Selecting Items With no Random Effects and Excluding Items Skipped in Trial 1</i>			
	Standard	Random	Older Adult
Intercept	3.717 (0.064)	3.676 (0.063)	3.801 (0.049)
Complexity	0.054 (0.006)	0.049 (0.006)	0.079 (0.004)
Grid Location	0.563 (0.037)	0.59 (0.037)	0.406 (0.028)
<i>Note</i> All significance calculated at the $p = 0.05$ level. The positive value for complexity indicates a preference for selecting easy items first.			

Figure 13

Regression Lines For Selection Order as a function of Grid Location and Character Complexity per age Group



Note The effects of grid location overwhelm the effects of character complexity.

Figure 14

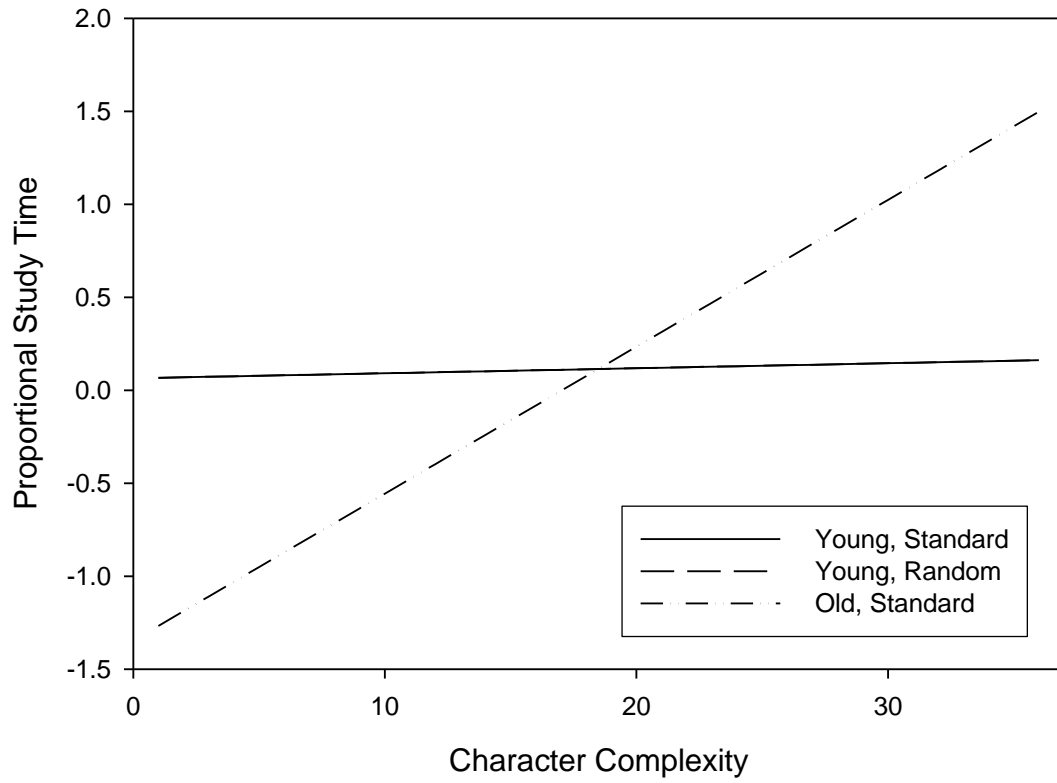
<i>Significant Fixed Effects Per each Condition in Selecting Items With no Random Effects and Excluding Items Skipped in Trial 2</i>			
	Standard	Random	Older Adult
Intercept	3.67 (0.071)	3.56 (0.070)	3.88 (0.050)
Complexity	0.0502 (0.0071)	0.0478 (0.0070)	.0977 (0.0050)
Grid Location	0.507 (0.042)	0.433 (0.041)	0.347 (0.029)
Trial 1 Accuracy	-0.334 (0.162)	-0.153 (0.161)	-0.668 (0.139)
Complexity*Accuracy	-.008 (0.014)	.003 (0.014)	-0.033 (0.011)
<i>Note</i> All significance calculated at the $p = 0.05$ level. The positive value for complexity indicates a preference for selecting easy items first.			

Figure 15

<i>Significant Fixed Effects Per each Condition in Total Time Allocation With no Random Effects in Trial 1</i>			
	Standard	Random	Older Adult
Intercept	0.114 (0.210)	0.114 (0.210)	0.114 (0.210)
Complexity	0 (0.0027)	0 (0.0027)	-0.0016 (0.002)
<i>Note All significance calculated at the p = 0.05 level</i>			

Figure 16

Regression Lines For Proportional Study Time as a function of Character Complexity per age Group



Note The two young adult groups are indistinguishable.

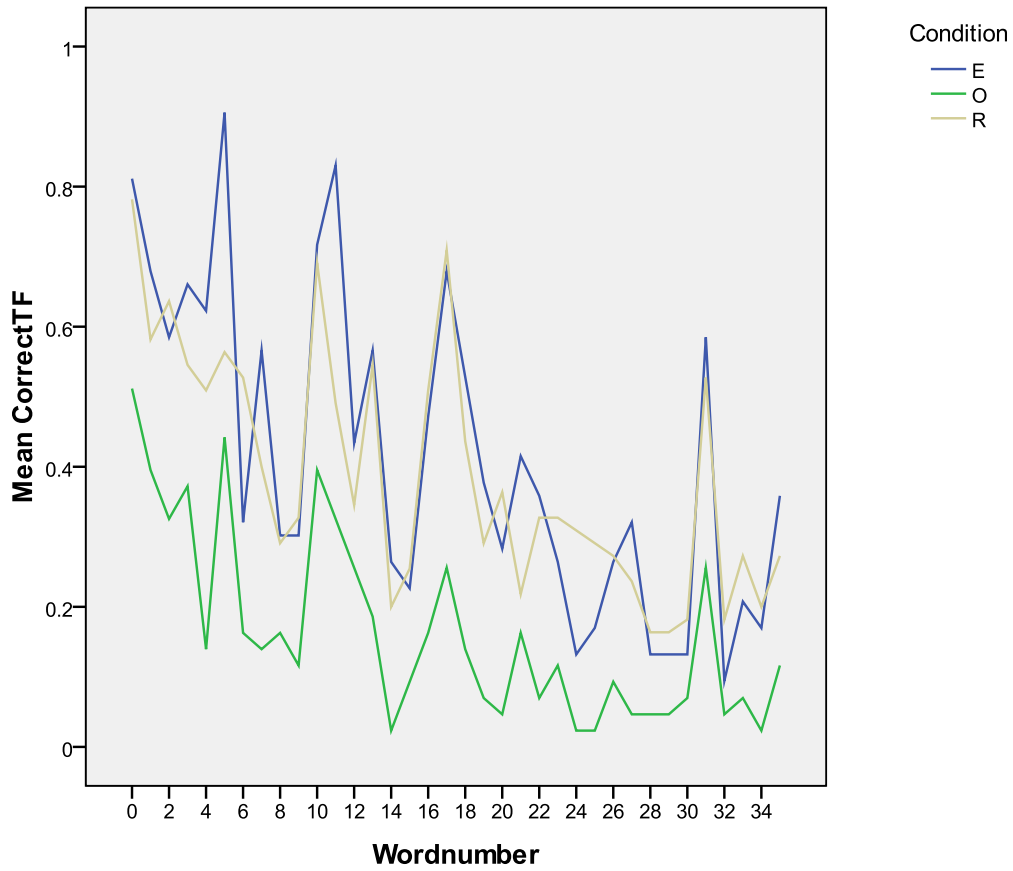
Figure 17

<i>Significant Fixed Effects Per each Condition in Total Time Allocation With no Random Effects in Trial 2</i>			
	Experimental	Random	Older Adult
Intercept	0.953 (0.003)	0.95 (0.003)	0.954 (0.002)
Complexity	0 (0.003)	0 (0.003)	-0021 (0.002)
Trial 1 Accuracy	-0.0032 (0.007)	-0.0041 (0.007)	-0.0061 (0.006)
Complexity*Accuracy	0 (0.0006)	0 (0.0006)	-0.0035 (0.0005)
<i>Note All significance calculated at the p = 0.05 level.</i>			

Figure 18

<i>Fitted Regression Line of Accuracy for Trial 1</i>			
	Standard	Random	Older Adult
Mean Accuracy	0.425 (0.017)	0.405 (0.016)	0.192 (0.013)
Complexity	-0.015 (0.0061)	-0.011 (0.0016)	-0.009 (0.0012)
<i>Note</i> All significance calculated at the $p = 0.05$ level			

Figure 19

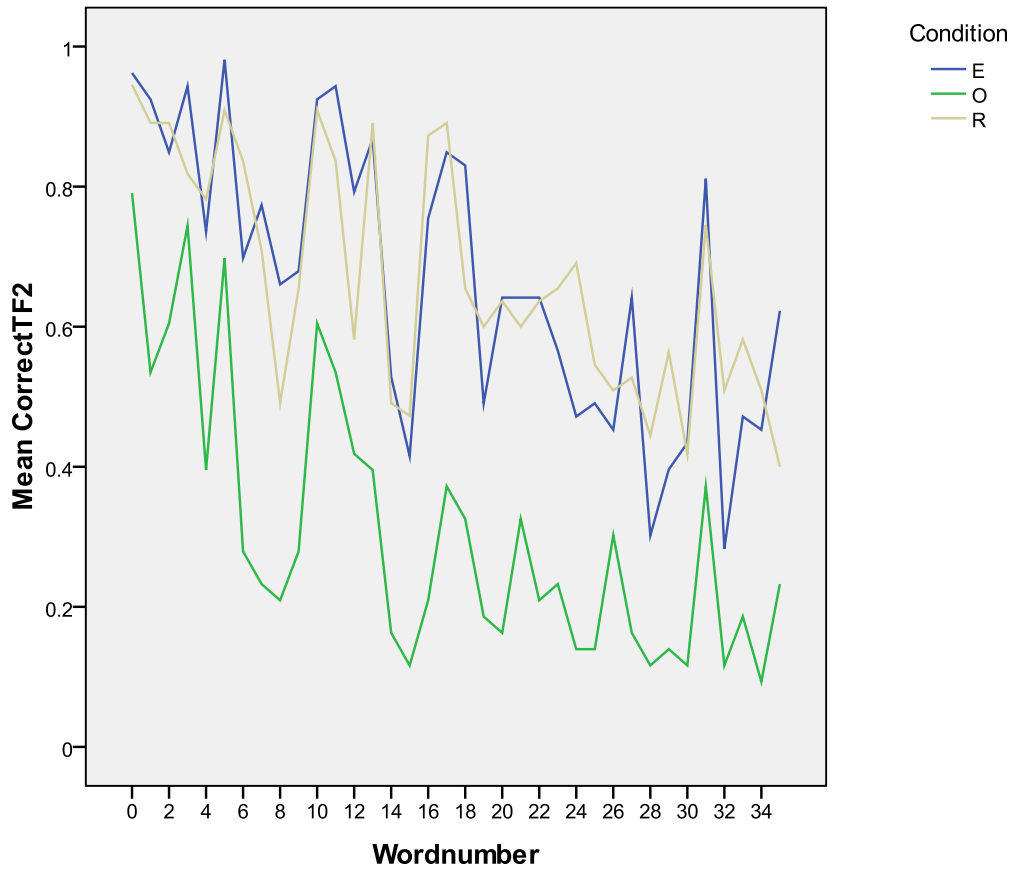


Mean number correct in trial 1 by the word number for each condition (E standard young adult, O older adult, R random young adult). There appears to be little difference between conditions, but the older adults show a lower performance.

Figure 20

<i>Significant Fixed Effects for Trial 2 Accuracy</i>			
	Experimental	Random	Older Adult
Mean Accuracy	0.523 (0.017)	0.542 (0.017)	0.238 (0.012)
Complexity	-0.008 (0.003)	-0.008 (0.003)	-0.008 (0.001)
Trial 1 Accuracy	-0.397 (0.04)	-0.400 (0.04)	.656 (0.034)
Complexity*Accuracy	0.007 (0.003)	0.007 (0.003)	0.008 (0.002)
<i>Note</i> All significance calculated at the $p = 0.05$ level.			

Figure 21



Mean number correct in trial 2 by the word number for each condition (E standard young adult, O older adult, R random young adult). There appears to be little difference between conditions, but the older adults show a lower performance.