EXAMINING METACOGNITIVE CONTROL: ARE THERE AGE-RELATED DIFFERENCES IN ITEM SELECTION DURING SELF-PACED STUDY?

A Dissertation
Presented to
The Academic Faculty

By

Jodi L. Price

In Partial Fulfillment
of the Requirements for the Degree
Doctorate of Philosophy in the
School of Psychology, College of Sciences

Georgia Institute of Technology
August, 2008

EXAMINING METACOGNITIVE CONTROL: ARE THERE AGE-RELATED DIFFERENCES IN ITEM SELECTION DURING SELF-PACED STUDY?

Approved by:

Dr. Christopher Hertzog, Advisor
School of Psychology
Georgia Institute of Technology

Dr. Anderson Smith
School of Psychology
Georgia Institute of Technology

Dr. Randall Engle
School of Psychology
Georgia Institute of Technology

Dr. John Dunlosky
Department of Psychology
Kent State University

Dr. Wendy Rogers
School of Psychology
Georgia Institute of Technology

Date Approved: April 23, 2008

ACKNOWLEDGEMENTS

With great appreciation I thank Dr. Christopher Hertzog, Dr. Richard Catrambone, and my committee members, Dr. Randall Engle, Dr. Wendy Rogers, Dr. Anderson Smith, and Dr. John Dunlosky, for their guidance and advice. Thank you for taking the time to share your expertise and mentor me throughout my graduate training.

I would also like to thank the members of the Hertzog Adult Cognition Lab, especially Teri Boutot, Chanteal Edwards, Ronit Greenberg, Dennis Hayes, Shannon Langston, Colin Malone, Alisha Monteiro, Rory Murray, and Starlette Sinclair, for their assistance in this project.

Thanks also to my graduate school colleagues and friends who made coming to school an engaging experience both in and out of the classroom. Huge thanks are due to those in my cohort as well as my lab and office mates for their support and friendship. Thank you for sharing the successful, and sometimes less than successful, but always interesting times with me.

I thank my family for their continuous love and support throughout the years. Thank you for cheering me on. Most of all I would like to thank my husband, David, for his unfaltering confidence in me and for making my goal his own. I cannot thank him enough for all the love, kindness, patience, and encouragement he has shown me throughout this process. It was he who made sure I found reasons to laugh each day and for that I am grateful.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Self-paced study involves choosing items for (re)study and determining how much time will be allocated to those items so as to maximize later recall, making it a viable venue for examining whether there are age-related differences in metacognitive control. Two prominent models have been proposed to account for item selection and study time allocation behaviors during self-paced study. The Discrepancy Reduction Model (DRM; Dunlosky & Hertzog, 1998; Nelson & Leonesio, 1988) suggests individuals will always select and allocate the most time to items that have not yet been learned, whereas the Region of Proximal Learning model (RPL; Metcalfe, 2002) predicts individuals will select the easiest unknown items and will only later select and allocate time to the more difficult items if time constraints permit, thus making distinctions among unlearned items graded by difficulty. Two experiments were conducted to examine whether younger and older adults' item selection and study time allocation behaviors would be more consistent with DRM or RPL model predictions. Across both experiments younger and older adults initially selected easier items for study, providing the first evidence to date that the RPL model would extend to older adults' self-paced study of heterogeneously difficult Spanish-English vocabulary pairs. However, both younger and older adults allocated more time to difficult than easier items. The assignment of point values to items in Experiment 2 affected how likely participants were to pursue each of four experimenter-determined task goals that either stressed the number of words recalled, points earned, or both. Whether point values initially favored recall of easy or difficult items interacted with time constraints to influence the basis (objective versus subjective difficulty) and

order of participants' item selections (Experiment 2). However, younger adults were better able to effectively allocate their study time to achieve self-determined (Experiment 1) and experimenter-determined goals (Experiment 2), indicating age-related differences in metacognitive control despite younger and older adults having similar memory self-efficacy ratings.

CHAPTER 1

INTRODUCTION

Self-paced study involves several variables that rely on metacognition and metacognitive control (e.g., monitoring, selection, and study time allocation decisions; Metcalfe & Kornell, 2005; Nelson & Narens, 1990). When presented with the to-be-learned material individuals must monitor and judge what they currently know and do not know, judgments that presumably affect which items are selected for study. Monitoring also plays a role after individuals have selected the items they will (re)study, as they monitor their progress and determine how much time will be allocated to each item and when to cease studying. Thus, the ability to accurately monitor and translate that monitoring into selection and time allocation behaviors that maximize later recall of the material makes self-paced study a viable arena for examining whether there are age-related differences in metacognitive control.

Metacognitive control comes into play because the decision to attend to some material often means other material will be given less attention, or ignored altogether. Of interest is how time constraints, the perceived difficulty of the material, perceptions of how well the material is already known, as well as the subjective and objective consequences of doing well influence decisions about what one will study and for how long. Also of interest is whether the ability to control item selection and the allocation of self-paced study time is maintained with age, or instead suffers age-related declines once one is removed from environments (e.g., educational or work) that frequently demand the skill.

How Items are Selected for Study

Much empirical research has been devoted to the question of how individuals select items for study (see Son & Metcalfe, 2000). Researchers have examined what types of items individuals will choose to study and allocate the most time using a variety of materials (e.g., concrete and abstract, related and unrelated word pairs, English-Spanish vocabulary pairs; Dunlosky & Connor, 1997; Dunlosky & Hertzog, 1997; Kornell & Metcalfe, 2006; Metcalfe, 2002; Metcalfe & Kornell, 2003, 2005; Nelson, Dunlosky, Graf, & Narens, 1994; Peligrina, Bajo, & Justicia, 2000; Thiede & Dunlosky, 1999). The results of these studies have yielded two prominent models about the basis of item selection during self-paced study: the Discrepancy Reduction Model (DRM; Dunlosky & Hertzog, 1998; Nelson & Leonesio,1988), and the Region of Proximal Learning (RPL) model (Kornell & Metcalfe, 2006; Metcalfe, 2002; Metcalfe & Kornell, 2003, 2005). The DRM and RPL model both predict individuals will select unknown rather than known items to study, but each model leads to very different predictions about *which* of the unknown items will be selected first.

The Discrepancy Reduction Model (DRM) of Self-Paced Study

According to the DRM (Nelson et al., 1994), when individuals are presented with items to study, they decide on a norm of study or desired level of learning for the items prior to studying them (Nelson & Narens, 1990). Once they begin studying, individuals are assumed to monitor how well they have learned each item and to compare their current level of learning to the desired level. If the current level of learning matches the norm of study then individuals are expected to stop studying that item and to select another item for study. However, if the current level of learning does not yet meet the

desired level, then the DRM predicts they will continue studying that item until the norm of study is achieved, either by allocating additional study time to the item or selecting that item for restudy. The same monitoring process is assumed to occur during the restudy phase, with the person monitoring the discrepancy between the current and desired level of learning, and studying the item until the perceived discrepancy between the two levels reaches zero (Nelson et al., 1994).

Because the DRM assumes the goal is to reduce the discrepancy between what individuals currently know and what they desire to know, the model predicts that individuals will always be most likely to select and devote the most time to the least well-known items, as reflected by their ease of learning judgments (EOLs; i.e., judgments about how easy or difficult it will be to learn the particular item; Underwood, 1966) and judgments of learning (JOLs; i.e., item-level judgments reflecting the likelihood of later being able to recall the items), in order to achieve zero discrepancy (Nelson & Leonesio, 1988). Thus negative correlations are expected between metamemory judgments (e.g., EOLs, JOLs) and item selection (Dunlosky & Hertzog, 1997). Moreover, if EOL monitoring affects control directly via a discrepancy-reduction mechanism, one would predict a negative correlation between metamemory judgments and study time allocation as well (Nelson & Leonesio, 1988).

Metcalfe (2002) proposed a slightly different interpretation of the DRM by suggesting that item difficulty would influence item selection. That is, she suggested that because individuals would be least likely to know the most difficult items, and because difficult items would likely elicit the lowest EOLs/JOLs, the DRM could be interpreted to suggest that individuals would be most likely to select the most difficult items because

these items would have the greatest discrepancy from the norm of study. Thus, Metcalfe equated item difficulty with the likelihood of an item being known or unknown, thereby suggesting the DRM would always predict the most difficult items would be selected for (re)study and allocated the most time. The distinction then is whether item selection behaviors, and the negative correlations between EOLs/JOLs and item selection, are assumed to depend on binary decisions about whether items are known or unknown or more fine-tuned decisions within the unknown (known) categories. Dunlosky and Hertzog (1998) implied that the negative correlations were driven by binary (i.e., recalled/ unrecalled) decisions, whereas Metcalfe emphasized grades of difficulty within unknown items as the basis for item selection and the negative correlations.

The DRM has remained a prominent account of self-paced study because a large portion of studies examining self-paced study behaviors have obtained the negative correlations predicted by the model, reflecting individuals' tendency to choose to study (Nelson et al., 1994) or restudy more difficult, presumably less well learned, items (Le Ny, Denhiere, & Le Taillanter, 1972). Son and Metcalfe (2000) provided a review of the self-paced study literature and found that within the 19 articles and 46 different conditions they reviewed, 35 of these supported the DRM predictions.

The Region of Proximal Learning (RPL) Model

The DRM is contrasted by Metcalfe's (2002) RPL model. Metcalfe suggested that after individuals assess the task, rather than selecting the least well known (i.e., most difficult) items for study, as proposed by the DRM, individuals select those items that are within their region of proximal learning, that is those items they feel they are most likely to be able to learn given their level of knowledge, goals, and task constraints (e.g., time

pressure). Specifically, the RPL model predicts that variables such as item difficulty, expertise of the participant, number of trials spent studying, and total study time available should affect the selection of items to which time is allocated and the susceptibility of the items to learning gains, with learning being reflected by a shift toward the study of items of progressively greater difficulty. The model further predicts that providing individuals with additional study time or additional study trials should produce a similar shift in the region toward more difficult items, but that the region of focus for a novice should be of lower difficulty than that of an expert.

Consistent with these predictions, Metcalfe (2002) found that children and adults unfamiliar with Spanish focused their time on easy and moderately difficult items, transitioning to more difficult ones when time constraints allowed. In contrast, experts focused on the difficult (and moderately difficult) items regardless of time constraints. Metcalfe and Kornell (2003) found that when novices were given more time and allocated some of it to the difficult items, the performance on those difficult items did not improve (i.e., what Nelson and Leonesio [1988] called the "labor in vain effect", p. 678), whereas additional study time did substantially help recall of the moderately difficult items.

Metamemory Judgments and Item Selection

Judgments of Learning

Nelson and Narens' (1990) theoretical model of metacognitive monitoring and control suggests that the subjective monitoring that occurs as individuals study and provide metamemory judgments (e.g., JOLs) will be used as a basis for controlling future study behavior. The DRM predicts that this subjective monitoring is used to identify and

select only the least well known items assigned the lowest JOLs, which should consistently yield a negative correlation between JOLs and item selection. In contrast, the RPL model suggests that the relationship between JOLs and item selection will vary depending on task and time constraints, and one's current knowledge level (e.g., whether one is an expert or a novice).

Judgments of the Rate of Learning (jROLs)

Dunlosky and Thiede (1998) and Metcalfe and Kornell (2005) hypothesized that an individual's assessment of how quickly one is learning items, which Metcalfe and Kornell referred to as judgments of the rate of learning (jROLs), will also influence the two processes that drive self-paced study: Choice and perseverance. Choice involves determining which items one will study and in which order, while perseverance involves a decision about how long to continue studying an item before switching to another. Metcalfe and Kornell hypothesized that individuals will continue studying if jROLs are high, but will cease studying when the perceived rate of learning decreases, unless individuals have mastery goals and are given unlimited time, in which case the DRM and RPL model predict the same thing (i.e., that individuals will study all items until they are learned with more difficult items taking longer to learn than easier items). However, time constraints, material difficulty, and an individual's self-expectations (e.g., memory self-efficacy) and motivation level are expected to influence the parameter value of the stopping jROL and decisions about whether to persist in studying or select another item for study. Thus, the RPL model predicts that with time constraints easier items and more difficult items will be studied less than moderately difficult items because easier items will be learned quickly and extremely difficult items will be learned very slowly, both of

which will cause the jROL to drop and inspire individuals to cease studying the item and select another. The RPL model thus assumes that study time allocation decisions are based on participants' jROLs, which can vary depending on task demands. In contrast, the DRM suggests study time allocation decisions are driven by a discrepancy reduction mechanism which is presumably not affected by perceptions of how quickly one is learning material (i.e., jROLs) and may require more study time than is warranted based on the rate of return in recall performance.

Metcalfe and Kornell (2005) investigated the choice component of novice college students' self-paced study decisions, who under time constraints could choose which of multiple items within each difficulty level they would study. They found that participants shown 3X3 grids containing 9 English words associated with 3 easy, 3 moderately difficult, and 3 difficult English-Spanish word pairs and given 45 seconds to select and study the intact word pairs, consistently chose easy items first, then the moderately difficult items, before the difficult items. However, after transitioning to the difficult items, participants did not always choose the difficult items exclusively but rather often went back to or stayed on easier items. The researchers concluded that their data indicated that the negative correlations relating choice of items to JOLs are due exclusively to people attempting to eliminate from further study the items they already know. However, once those items are eliminated then participants proceed in an order from the easiest to the most difficult rather than the reverse, consistent with the RPL model and inconsistent with the DRM.

Experimental Manipulations and Item Selection

Time Constraints

The DRM predicts that the least well known (most difficult) items will be selected for study, regardless of how much time individuals are allotted, whereas the RPL model predicts that only with increased time limits will individuals select and study the more difficult items. Yet only under conditions (e.g., time constraints) that force participants to prioritize which of the unlearned items they will study do the DRM and RPL model predict different selection behaviors, with the RPL model predicting that time constraints will cause individuals to select easier items first.

Metcalfe and colleagues (Metcalfe, 2002; Metcalfe & Kornell, 2003, 2005; Son & Metcalfe, 2000) observed differences in how individuals selected items for study depending on the time constraints and the task goals they were given. Participants given limited time to study material were more likely to focus on the easier items, whereas no preference was shown for easy or difficult material when participants were allotted sufficient time to study all the items. Metcalfe and colleagues therefore concluded that the perception of time pressure could influence the correlations found between JOLs and item selection/ study time behaviors.

Goals

Thiede and Dunlosky (1999) demonstrated that performance goals also influence the JOL and item selection/ study time relationship. Their hierarchical model of self-regulated study predicted that a high performance goal (e.g., try to recall 24 of the 30 words) would yield the standard (i.e., modal) negative correlations between JOLs and selection and study time, but that a low performance goal (e.g., try to recall 6 of the 30

words) would result in a positive relationship between JOLs and selection/time because participants should select the easiest items and study them for as little time as they felt necessary to achieve that low goal. Thiede and Dunlosky found a dissociation between how JOLs related to item selection (positively) and how JOLs related to study time allocation (negatively) under low performance goal conditions. However the relationship between JOLs and study time was negative for both high and low performance goals. This dissociation held regardless of whether the performance goals were provided by the experimenter or self-initiated by participants as a result of time constraints.

Presentation Format

Thiede and Dunlosky (1999) discovered that presentation format critically moderated how people regulated study via item selection by influencing the likelihood of individuals forming an item selection plan. Despite giving all participants a low performance goal (i.e., learn 6 items), participants given arrays only selected on average 8 to 9 items for restudy, whereas those required to select items one at a time selected on average 19 to 20 items. Thus JOLs were positively related to item selection in the array format but negatively related to item selection in the item-by-item format, with both correlations being reliably different from zero.

Dunlosky and Thiede (2004) further demonstrated that individuals were better able to select items in accord with a low performance goal when items were presented in a simultaneous rather than sequential format. Participants required to select items presented simultaneously were more likely to form an item selection plan than those in the sequential presentation condition, and those with higher working memory capacity (WMC) were better equipped to execute the plan, since it required attention and control,

than those with lower WMC. Thus the tendency to adhere to a low performance goal by selecting the easiest items for study, an effect that has been called the shift-to-easier-materials (STEM) effect, was present under simultaneous, but not sequential presentation methods.

Age-Related Differences in Item Selection during Self-Paced Study

Only a few studies have examined whether there are age-related differences in regulation of study. Dunlosky and Hertzog (1997) found that both younger and older adults chose to restudy items to which they had given the lowest JOLs, consistent with the DRM and the results Nelson et al. (1994) obtained with younger adults, but that the older adults had lower recall performance. Dunlosky and Connor (1997) found that although both younger and older adults selectively allocated more time to low JOL and/or not recalled items, younger adults had better recall. Moreover, allocation of study time mediated the age-related differences in recall such that controlling for the allocation of study time drastically reduced the age-recall relationship, suggesting possible age-related differences in self-regulation.

Manipulations that Influence the Motivation to Select and Recall Particular Items

Recent research conducted by Castel, Benjamin, Craik, and Watkins (2002) runs counter to the notion that older adults might have greater problems controlling their study and recall behaviors. They presented younger and older adults words to study that each had different point values associated with them, with certain words having greater value than others in terms of points earned for being able to recall them. They found that older adults were much more likely than younger adults to selectively control their study and

recall so as to focus on the items worth more points and optimize the number of points earned.

Dunlosky and Thiede (1998) examined the impact of points on younger adults' study time allocation behaviors and later recall of either Swahili-English vocabulary or concrete noun paired associates (PA). Participants were found to allocate increasing amounts of time across trials to the higher valued items, which resulted in small, but reliably higher recall levels for the higher valued items. Dunlosky and Thiede concluded that point values, despite having a small effect, did reliably influence how individuals chose to allocate their study time when presented with homogeneously difficult vocabulary or concrete noun PA. However, because the Dunlosky and Thiede point value manipulations were designed to examine study time allocation rather than item selection behaviors, it remains unknown how point values would influence which items participants would select for study and whether point values would interact with how difficult the items were if item difficulty were not held constant across items.

The Motivating Effect of Goals

That older adults in the Castel et al. (2002) study were able to select and recall items to maximize the number of points earned raises questions about why age-related differences in metacognitive control were small in their study, given the larger age-related differences observed by Dunlosky and Connor (1997). One possibility is that because Castel et al. (2002) participants were provided extrinsic motivation (i.e., points) to recall items, whereas the Dunlosky and Connor (1997) and Dunlosky and Hertzog (1997) participants were asked to recall items simply for the sake of recalling them, that point values served to motivate the Castel et al. participants in a way the Dunlosky and

colleague participants were not. If task manipulations such as point values can affect participants' motivation and reduce observed age-related differences in metacognitive control then this would point to the age-related differences being more of a motivational or self-efficacy issue rather than a metacognitive issue.

West and colleagues (West, Bagwell, & Dark-Freudeman, 2005; West & Thorn, 2001; West, Thorn, & Bagwell, 2003; West, Welch, & Thorn, 2001) found that both younger and older adults were more likely to pursue and increase self- or experimenter-determined task goals when they received positive feedback (i.e., feedback indicating success), as opposed to no feedback. Yet goal setting only affected younger adults', but not older adults' self-efficacy (West & Thorn, 2001). Thus identifying goal and motivational manipulations that serve to improve or override older adults' lower levels of memory self-efficacy might in turn improve older adults' performance in memory tasks by affecting their willingness to learn and use effortful, but effective strategies to learn the material, as well as influence the norm of study they set for each item (Dunlosky & Hertzog, 1998). Unfortunately, no attempts have been made in the self-paced study literature to determine what participants' task goals are. Yet, because goals have been found to affect younger and older adults' task motivation (West & Thorn, 2001), it seems important to know whether there are age-related differences in the goals that individuals have in the context of self-paced study.

Such questions become important if item selection and study time allocation are to be taken or interpreted as evidence of metacognitive control. For example, if younger adults select items and allocate their study time in a different manner than older adults and the younger adults subsequently recall more items, how does one tease apart whether

the different recall levels are the result of differential item selection, the way time was allocated, age-related differences in episodic memory, or some other variable such as the implicit goals that participants in each age group brought to the task? The Dunlosky and Thiede (2004; Thiede & Dunlosky, 1999) results indicated that younger adults given lower performance goals approached self-paced study very differently than those given higher performance goals. It seems likely then, that if older adults enter self-paced study with lower goals than do the younger adults (e.g., as a result of lower memory self-efficacy; Dunlosky & Hertzog, 1998), that these goal differences could account for some, if not all, of the age-related differences in self-paced study that have been observed (Dunlosky & Connor, 1997).

Strategy Use and Vocabulary Learning

Also unknown are what methods or strategies, if any, younger and older adults in the cited studies have used to facilitate or hinder recall during self-paced study. Older adults have been found to be less likely than younger adults to spontaneously select and use normatively effective strategies (e.g., interactive imagery) to learn PA items (Dunlosky & Hertzog, 1998). Moreover, older adults tend to have greater difficulty decoding mediators (i.e., retrieving the mediator and then using the mediator to facilitate recall of the to-be-remembered words) they might have used (Dunlosky, Hertzog, & Powell-Moman, 2005). If older adults have greater problems using and decoding strategies when learning English PA, it remains possible they would have even greater difficulty applying strategies to learn foreign language vocabulary. This possibility stems from the vocabulary terms being unknown as well as the finding that even younger adults infrequently spontaneously use more effortful strategies such as imagery or the keyword

method (i.e., finding an English word that looks or sounds like the foreign word and then forming an image or sentence that links the keyword with the foreign word; e.g., remembering the Spanish word *mesa* means *table* by using the keyword *mess* and forming the sentence *a messy table*) during the early stages of learning foreign language vocabulary terms. This occurs despite research indicating that these more complex strategies yield greater retention than the less effortful rote repetition strategy (Sagarra & Alba, 2006). Thus it seems likely that if older adults are less likely than younger adults to use complex strategies when learning English-English PA, that they would be equally, if not more, unlikely to use these more effortful strategies when faced with learning Spanish-English vocabulary pairs. While this remains an empirical question, Gruneberg and Pascoe (1996) have demonstrated that older adults' ability to learn foreign language vocabulary is improved by using the keyword method.

Whether younger and older adults choose to use more effortful strategies in the context of studying foreign language vocabulary could therefore account for additional variance if age-related differences are observed in self-paced study and the resulting ability to recall the items. For example, even if participants in both age groups selected the same items for study and allocated the same amount of time to each item, if younger adults spontaneously generated and used mnemonic strategies to encode the items and the older adults did not, then this could account for younger adults having relatively higher recall levels and give the appearance of their also having greater metacognitive control. Thus, without knowing *what* participants are doing during allocated study time, it remains difficult to ascertain whether one age group is in fact being more efficient in their study time allocation and/or exhibiting greater metacognitive control.

Finally, whereas Dunlosky and colleagues demonstrated that both younger and older adults selected (Dunlosky & Hertzog, 1997) and allocated more time to low JOL items (Dunlosky & Connor, 1997), as predicted by DRM, it has been an open question whether Metcalfe's (2002) RPL model would extend to older adults. All work to date examining the RPL model's predictions has been limited to younger adult samples, leaving open the question whether novice older adults presented with items that vary in difficulty level would select easier items first, as predicted by the RPL model, or would instead select only the less well known (most difficult) items as predicted by the DRM. Also unknown is how experimental manipulations (e.g., assigning point values to the to-be-learned items), designed to affect participants' motivation, might impact the likelihood of participants selecting items of varying difficulty. If point values could "push" or cause individuals to be more or less likely to select particular items for study (e.g., Dunlosky & Thiede, 1998), and the manipulation affected younger and older adults in a similar fashion, then this would provide insight as to whether age-related differences in self-paced study are driven by motivation or metacognition.

Two experiments were conducted to investigate these issues. Of interest was whether the RPL findings that have been observed with younger adults would extend to older adults and whether older adults have different task goals than younger adults do, which might account for the presence or absence of age-related differences in self-paced study behavior and recall performance outcomes. In addition, the studies were designed to allow examination of which strategies, if any, younger and older adults used that might facilitate (or hinder) recall performance. Finally, the second experiment investigated how attaching point values to the to-be-learned items interacted with experimenter-provided

goals to influence item selection and strategy use behaviors, and later recall performance. Thus the second experiment examined whether self-paced study is driven more by motivational or metacognitive goals.

CHAPTER 2

EXPERIMENT 1

Dunlosky and colleagues' (Dunlosky & Connor, 1997; Dunlosky & Hertzog, 1997) prior research examining item selection and study time allocation behaviors with homogeneously difficult items (e.g., unrelated concrete noun PA) found that younger and older adults' choice and allocation behaviors were consistent with the DRM. Metcalfe and colleagues (Metcalfe, 2002; Kornell & Metcalfe, 2006; Metcalfe & Kornell, 2003, 2005) instead used items from obviously heterogeneous, discrete categories (easy, moderately difficult, and difficult) to investigate self-regulated study decisions and found that novice younger adults selected and allocated more time to easier rather than less well known English-Spanish vocabulary pairs. Experiment 1 was therefore designed to examine whether younger *and* older adults would select items in keeping with RPL model predictions or if instead age-related differences in item selection and study time allocation behaviors would emerge. Experiment 1 also examined whether there were age-related differences in task goals and strategy use.

Method

Design and Participants

Experiment 1 used a 3 (difficulty of the items: easy, moderately difficult, difficult) X 3 (age group: younger experimental, younger control, older) X 2 (presentation order: order 1 versus order 2) design, with item difficulty manipulated within subjects. Presentation order did not impact any of the variables of interest and so was ignored in all the analyses.

Sixty-four younger (30 males and 34 females; 64% Caucasian, 22% Asian, 8% African American, and 6% of Mixed racial background) and 31 older adults (11 males and 20 females; 74% Caucasian, 26% African American with *M* = 14.52, *SD* = 2.45 years of education) participated in the first experiment. Younger adults (*M* age = 20.05 years, *SD* = 1.60) were psychology students at the Georgia Institute of Technology and received extra credit for participating. Older adult participants (*M* age = 69.16 years, *SD* = 5.27) were normal, community-dwelling adults recruited from Atlanta, Georgia. They received a nominal fee of 35 dollars to compensate them for their time.

Participants in both age groups rated their overall health as very good (younger adult *M* rating = 1.92, *SD* = 0.78 and older adult *M* rating = 2.19, *SD* = 0.87, where 1 = excellent and 5 = poor) and all were pre-screened to ensure they were neither a bilingual Spanish speaker nor had taken Spanish courses in high school or college so that pre-existing knowledge of the Spanish language would not interfere with interpretation of the results. All participants were either native English speakers or, in the case of 5 younger adults, had been speaking English for at least 10 years.

Random assignment of the younger adults resulted in 32 participants in both the experimental (45 second) and control (90 second) conditions. All older adults received 90 seconds. Calculations produced in the G*Power 3 program indicated the number of participants tested was sufficient to detect a small effect (Faul, Erdfelder, Lang, & Buchner, in press).

Materials

Seventy-five of the 144 concrete and abstract noun, adjective, and verb English-Spanish vocabulary pairs from Metcalfe (2002) were used as stimuli (see Appendix A).

The subset of 75 items was selected based on recall norms collected by Kornell and Metcalfe (2006) such that only pairs with recall levels indicative of easy, moderately difficult, and difficult status were used. All items whose a priori category classification was not justified by Kornell and Metcalfe's data were excluded before 25 items in each category were selected. Across categories (i.e., easy, moderately difficult, and difficult), the selected words differed in terms of Spanish word length as well as probability of recall (as measured by Kornell & Metcalfe). Easy items had an average Spanish word length of 7.72 (1.86) letters and average recall of .95 (.02), whereas moderately difficult items had an average length of 8.08 (1.89) letters, with an average recall of .63 (.14), and difficult items were on average 11.12 (1.81) letters long, with an average recall rate of .06 (.04). To preview the recall results, the rates obtained in this experiment for younger adults are similar or higher than those obtained by Kornell and Metcalfe (2006), perhaps due to their using the full lists which had greater variability in each difficulty level.

Spanish-English, rather than English-Spanish, vocabulary pairs were used such that Spanish words served as the cues and their English counterparts as targets. Although Metcalfe (2002; Metcalfe & Kornell, 2003, 2005) used the opposite (i.e., English-Spanish pairs), the use of Spanish words as cues allowed EOL judgments to be collected on the Spanish words alone rather than the intact Spanish-English vocabulary pairs. This removed the possibility that younger adults might learn more than older adults through exposure to the intact pairs during the EOL collection phase, which would have made the age-group comparison of study behaviors difficult to interpret if differential learning in the EOL phase resulted in younger adults selecting more difficult items than the older

adults. Spanish-only EOLs alleviated these concerns and simultaneously increased the likelihood that participants' EOLs would be more predictive of item selection behaviors because the same item characteristics (e.g., word length, similarity to English words, etc.) that likely influenced participants' EOLs were likely to influence item selection.

Three (one of each difficulty level) of the 75 selected vocabulary pairs were used in the instruction screens. The other 72 pairs (24 of each difficulty level) were presented during the EOL collection, prestudy recall testing, encoding, delayed judgment of learning collection, and poststudy recall testing phases. Instructions and stimuli for all but the encoding phase of the criterion task were presented on desktop computers using Inquisit software, version 3.0.0.0 (2007). Participants completed all phases of the criterion task, with the exception of the encoding phase, using a mouse and standard *Qwerty* keyboard to record their metamemory judgments and recall responses. The Inquisit program automatically scored participants' prestudy and poststudy recall responses as correct if their typed responses were 100% correct (i.e., did not contain a spelling error). Scoring was later hand-checked using both a strict and lenient (gist-based) criterion so that participants were not penalized for spelling errors.

The encoding phase did not use a computer program due to the amount of information that had to be recorded during this phase. Instead, the 72 stimuli were presented on 5 x 8 inch note cards on which the Spanish cue was presented, centered on one side, in 50-point Arial Bold font and the intact Spanish-English word pair was presented, centered on the other side, in 35-point Arial Bold font. The font style and sizes were chosen to facilitate ease of reading and to provide a consistent maximum size that would allow words/pairs of all lengths to fit on the note card. Two random presentation

orders were formed by creating two sets of note cards and then randomly dividing, within each set of cards, the 24 note cards for each difficulty level into three sets of 8. Each set of 8 cards was then bound with a ring which maintained the two randomly determined set orders, but allowed the cards to easily be flipped once participants selected an item for study. Depending on which of the two orders the participant was supposed to receive, the proper set of note cards was prearranged on a 24 x 36 inch bulletin board which held the three sets of easy item note cards in three rows in the left column, the moderately difficult items in the middle column, and the difficult items in the right column. The order in which the sets of cards were arranged on the board corresponded with the arrangement of the Spanish words on paper-based grids on which the experimenter recorded whether an item was (re)selected, the order in which items were (re)selected, how much time was allocated to an item each time the item was selected, and which strategy participants reported using to study an item each time it was selected. Thus the two presentation orders allowed examination of order effects while reducing the set-up time and data recording burden that would have existed if each participant had received a completely random order.

A computerized version of the Personal Beliefs about Memory Inventory (PBMI; Lineweaver & Hertzog, 1998) was used to assess participants' memory beliefs and memory self-efficacy. The paper-based Memory Controllability Inventory (MCI; Lachman, Bandura, Weaver, & Elliott, 1995) was used to allow investigation of the relationship between participants' memory control beliefs and item selection behaviors. Because knowledge of the English language has been found to influence the ability to learn Spanish vocabulary (Sparks & Ganschow, 1993), the Advanced Vocabulary Test

(AVT; Ekstrom, French, & Harman, 1976) was used to measure participants' understanding of word meanings. The Listening Span (Salthouse & Babcock, 1991) task was used to assess participants' WMC based on Dunlosky and Thiede's (2004) finding that WMC influenced participants' ability to select and recall items. Finally, a measure of perceptual speed, the Letter Comparison task (Salthouse, 1996), was also used to allow assessment of whether age-related declines in processing speed might impact older adults' performance in the criterion task. The WMC, perceptual speed, and memory beliefs measures were included to provide insight about potential sources of age-related differences in study behaviors. However, the limited number of participants did not afford examination of individual differences in these tasks.

Procedure

Participants were tested individually. Participants first completed a demographic questionnaire and the memory beliefs and memory control questionnaires (i.e., the PBMI and MCI) before completing any other measures to ensure that any problems participants had on the other measures would not influence their responses on these questionnaires. Participants then completed the Listening Span, AVT, and Letter Comparison tasks, took an optional 5 minute break, and then began the criterion task.

Participants were told they would be given the opportunity to select Spanish-English vocabulary pairs to study during the study phase, but should first indicate how easy or difficult they thought each Spanish word would be for them to learn (i.e., complete the EOL collection phase). Each Spanish word was presented, in a random order, one at a time, for 3 seconds, in the middle of the screen under which a 9-point Likert scale appeared containing the word "easy" under the "1", "medium" under the "5",

and “hard” under the “9”. Participants provided their EOL judgments for each word by typing the number that corresponded with their judgment.

Participants began the prestudy recall testing phase after they provided their EOL judgments. Each Spanish word was presented again, one at a time, in a different random order, and participants were asked to type their best guess as to what the English equivalent for the Spanish word was. Participants received immediate feedback regarding whether the English word they typed was correct. If participants typed the English equivalent of the presented Spanish word the word “correct” appeared, whereas if their response was not the equivalent, the word “incorrect” appeared in the center of the screen for 2 seconds before the onset of the next Spanish word. The feedback was provided to reduce the likelihood that participants would select items during the study phase simply to check the accuracy of their responses given in the prestudy recall test. The test also provided an objective measure of which Spanish words participants already knew prior to beginning the study phase of the criterion task. Assessing prior knowledge of Spanish was important because both the DRM and RPL model predict that individuals will focus on unknown items, and because RPL predicts individuals will select items they can expect to learn within the given time constraints. The prestudy recall test allowed identification of which items were known, unknown, or seemed to be partially known (e.g., partially correct; i.e., in a transitional state; Atkinson, 1972) by each participant, thereby allowing the analyses of item selection behaviors to account for this.

Participants began the study phase after completing the prestudy recall test. Participants viewed introductory screens on the computer which briefly described strategies that have been found to be helpful in learning foreign language vocabulary

terms (e.g., rote repetition, keywords, imagery, and sentence generation; Lawson & Hogben, 1998). Participants were shown three new sample Spanish –English vocabulary pairs, one from each level of difficulty, to familiarize them with the types of items they could expect in each difficulty level. They were then asked to indicate what their global recall goal was, ignoring difficulty level (e.g., "In a minute you will be asked to study 72 Spanish-English vocabulary pairs. What is your overall goal for recalling items? In other words, how many of those 72 items are you going to try to learn and recall? Please enter a number between 0 and 72 below."), and then accounting for the three difficulty levels by providing three different differentiated goals for items from each difficulty level (e.g., "Now, because some items are easier to learn than others, I'd like to ask about your goals for learning items of different levels of difficulty in this session. You will be asked to study 24 easy/ 24 medium/ 24 hard vocabulary pairs. What is your goal for recalling easy/medium/ hard items? Please enter a number between 0 and 24 below."). Participants then provided global and differentiated recall predictions (e.g., "How many easy/ medium/ hard items do you expect to be able to recall?") for each type of item before they began the study phase.

The computerized instructions told participants the study phase would include 8 different 3 X 3 study grids and that each grid would contain the Spanish portion of 9 Spanish-English vocabulary pairs --3 pairs normatively classified as easy, 3 moderately difficult pairs, and 3 difficult pairs. Participants were told that the easy Spanish items would be in the left column, moderately difficult in the center, and the difficult items in the right and that they would be able to select items for which they wanted to see the intact vocabulary pair. Older adults were told they would have 90 seconds to study items

within each grid. A younger adult control group also received 90 seconds per grid, whereas a second younger adult experimental group was allotted 45 seconds per grid, in keeping with the 45 seconds Metcalfe and Kornell (2005) gave their younger adult participants. The additional time given to older adults (and younger adult control participants) in each grid was designed to compensate for any age-related declines in processing speed (Salthouse, 1996) that might slow older adults' selection and allocation processes. However, because time constraints are known to influence item selection behaviors (e.g., Son & Metcalfe, 2000) the younger adult control group was included to allow direct comparison of their item selection behaviors to an older adult group given the same amount of time (i.e., 90 seconds per grid), as well as to a younger adult experimental group given half the time (i.e., 45 seconds per grid) to disentangle the impact of time on self-paced study decisions.

Once participants finished reading the computerized instruction screens they were moved to a seat placed approximately 24 inches in front of the bulletin board containing the note cards. Participants were again told the column locations of the easy, moderately difficult, and difficult items and informed they could use the allotted time to select and study any of the items within a grid. Participants were told to indicate which items they would like to study, one at a time, by pointing at or saying the Spanish word. Participants were told the experimenter would flip the selected note card over so the intact Spanish-English vocabulary pair was visible and would begin the countdown clock, which remained visible to participants so they could track how long they had to study or restudy items within the grid. They were instructed to study the intact pair as long as they liked (or until time ran out) and to indicate when they were finished studying that item, at

which point the experimenter stopped the clock. The experimenter then asked the participant which strategy, if any, they had used to study the item and recorded this information as well as the order in which items were selected and how much time was allocated to each item, before the participant selected the next item. This process continued until either time elapsed or participants indicated they had (re)studied all the items within a grid they wished to study, even if time still remained, at which point the next grid was presented. Once participants had viewed all 8 grids they returned to the computer and began providing their delayed judgments of learning (DJOL).

In the DJOL collection phase participants were shown the Spanish word from each of the 72 vocabulary pairs, one at a time in a random order, regardless of whether the item had been selected for study, and asked to rate how confident they were that they would be able to correctly recall the English equivalent on the recall test about 10 to 15 minutes later. Participants provided their ratings using a scale of 0 to 100 in which 0 indicated no confidence and 100 indicated 100% confidence the English equivalent would be correctly recalled.

After individuals provided their DJOLs, they provided global and differentiated poststudy prediction reports indicating how many items, total and within each difficulty level, they expected to be able to recall. Participants then completed the recall test by typing the English equivalent for each Spanish word before postdicting their performance, globally and for each item type. Participants then reported what their global and differentiated goal and recall predictions would be if they were given the opportunity to study the same items again to allow examination of how recall experience affected participants' task expectations. Participants were then debriefed and dismissed.

Results

Cognitive Ability Measures

Vocabulary

Younger and older adults' scores on the Advanced Vocabulary Test (AVT; Ekstrom, French, & Harman, 1976) were very similar. Younger adults chose the correct synonym an average of 16.84 (*SE* = .61) and older adults 18.74 (*SE* = 1.41) times out of the 36 items, a difference that was not reliable, F (1, 93) = 2.09, $p > .05$, partial $\eta^2 = .022$. Thus English vocabulary skills should not be an impediment to younger or older adults' learning of the Spanish-English vocabulary pairs (Sparks & Ganschow, 1993).

Processing Speed (PS)

Younger adults scored reliably higher than older adults on the Letter Comparison task (Salthouse, 1996; younger adult M = 25.69, SE = .53 versus older adult M = 17.03, SE = .82, out of 42 items), F (1, 94) = 63.14, $p < .001$, partial $\eta^2 = .472$. If processing speed affects participants' ability to select and allocate study time, then this justifies the additional study time afforded the older adults (and younger adult controls).

Working memory capacity (WMC)

Younger adults recalled significantly more items on the Listening Span task (Salthouse & Babcock, 1991) than older adults (younger adult M = 69.88, SE = .91 versus older adult M = 40.48, SE = 3.13, out of 81 items), F (1, 94) = 135.41, $p < .001$, partial $\eta^2 = .593$. These age-related differences in WMC could result in younger adults being better able to form and implement an item selection plan than older adults (Dunlosky & Thiede, 2004).

Memory Beliefs Measures

PBMI

Older adults were expected to have lower memory self-efficacy than younger adults. However, responses to PBMI items assessing participants' global memory self-efficacy and memory control beliefs revealed no age-related differences, F (1, 93) = 1.39, $p > .05$, partial $\eta^2 = .015$, and F (1, 93) = 0.05, $p > .05$, partial $\eta^2 = .001$, respectively.

MCI

None of the questions on the Memory Controllability Inventory produced age-related differences in ratings. Participants in both age groups rated their present ability (e.g., "I can remember the things I need to") and potential for improvement (e.g., "I can find ways to improve my memory") similarly, F (1, 93) = 2.36, $p > .05$, partial $\eta^2 = .025$, and F (1, 93) = 1.53, $p > .05$, partial $\eta^2 = .016$, respectively. Thus none of the memory belief variables measured on the two memory beliefs questionnaires that could be expected to influence goal setting and goal pursuit behaviors were found to differ across age groups.

Table 1. PBMI and MCI Memory Beliefs Ratings

	Younger- 45		Younger- 90		Older	
Judgment Type	*M*	*(SE)*	*M*	*(SE)*	*M*	*(SE)*
PBMI Global MSE	80.95	(2.66)	79.06	(2.60)	76.00	(3.03)
PBMI Control	84.74	(1.93)	85.55	(2.26)	84.53	(2.28)
MCI Present Ability	5.70	(.17)	5.50	(.15)	5.30	(.17)
MCI Potential Improve	5.21	(.15)	5.54	(.18)	5.11	(.20)

Note. Entries are means and standard errors for participants' memory beliefs ratings; Younger-45 = younger adults in the 45 second experimental condition; Younger-90 = younger adults in the 90 second control condition; PBMI Global MSE = overall memory self-efficacy; MCI Present Ability = ratings of how good their present memory is; MCI Potential Improve = participants' perceptions of how much they can potentially do things to improve their memory.

Ease of Learning Judgments

Participants' EOLs were analyzed to examine alignment with the Metcalfe (2002) and Kornell and Metcalfe (2006) normative difficulty classifications (i.e., easy, moderately difficult, difficult) and to determine whether there were age-related differences in participants' impressions of item difficulty. Participants provided their EOL judgments using a 1 (easy) to 9 (hard) scale, therefore alignment with the Metcalfe categories would be reflected by easy items having a mean rating between 1 and 3, moderately difficult items between 4 and 6, and difficult items between 7 and 9. Collapsing across age groups, participants' EOLs did align with Metcalfe's categories. The 24 items that Kornell and Metcalfe classified as easy were given a mean rating of 2.70 (*SE* = .13), moderately difficult a mean rating of 4.80 (*SE* = .13), and the difficult items were rated on average 6.95 (*SE* = .11). Although these average ratings are

consistent with Metcalfe's categories, the rated difficulty was higher, in large part because older adults rated items of each type as significantly more difficult than younger adults, F (1, 93) = 20.30, $p < .001$, partial $\eta^2 = .179$. Yet, because younger and older adults' ratings increased as difficulty increased, a main effect of item Difficulty was observed, F (2, 92) = 535.51, $p < .001$, partial $\eta^2 = .921$, but the Age X Difficulty interaction was not reliable, F (2, 92) = 1.74, $p > .05$, partial $\eta^2 = .036$ (see Table 2).

Goals

Prestudy Goals

Reliable age-related differences were found in participants' global goals, F (1, 93) = 12.10, $p < .01$, partial $\eta^2 = .115$. Both younger and older adults decreased their differentiated goals for more difficult items yielding a reliable main effect of Difficulty, F (2, 92) = 128.06, $p < .001$, partial $\eta^2 = .736$, but no reliable interaction of Age and Difficulty, F (2, 92) = 2.73, $p > .05$, partial $\eta^2 = .056$.

Post-Recall Goals

After participants studied and attempted recall of items they showed a significant shift in the goals they set for a hypothetical chance to restudy items a second time. The change in participants' global goals from prestudy to post-recall yielded main effects of Trial, F (1, 93) = 31.55, $p < .001$, partial $\eta^2 = .253$, and Age, F (1, 93) = 21.70, $p < .001$, partial $\eta^2 = .189$, but no Trial X Age interaction, $F < 1$. The shift in participants' differentiated goals also yielded main effects of Trial, F (1, 93) = 107.48, $p < .001$, partial $\eta^2 = .536$, Age, F (1, 93) = 17.72, $p < .001$, partial $\eta^2 = .160$, and item Difficulty, F (2, 92) = 205.91, $p < .001$, partial $\eta^2 = .817$, but these were qualified by a reliable Trial X Age X Difficulty interaction, F (2, 92) = 5.19, $p < .01$, partial $\eta^2 = .101$, as a result of

older adults no longer setting a lower goal for easy items, and instead setting lower goals for moderately difficult and difficult items (see Tables 2 and 3).

Table 2. Mean EOLs, Prestudy Goals, Prestudy and Poststudy Predictions

	Younger- 45		Younger- 90		Older	
Judgment	*M*	*(SE)*	*M*	*(SE)*	*M*	*(SE)*
Global EOL	4.62	(.16)	4.38	(.15)	5.48	(.18)
EOL Easy	2.46	(.17)	2.15	(.14)	3.53	(.21)
EOL Medium	4.60	(.20)	4.37	(.19)	5.45	(.22)
EOL Hard	6.79	(.18)	6.61	(.17)	7.45	(.18)
Global Goal 1	41.91	(2.59)	38.00	(2.72)	27.42	(2.93)
Goal 1 Easy	18.52	(.89)	18.87	(.89)	15.25	(1.08)
Goal 1 Medium	14.84	(.83)	12.02	(.79)	10.10	(.75)
Goal 1 Hard	8.89	(.70)	7.77	(.49)	6.82	(.69)
Global Prestudy Pred. 1	38.86	(1.94)	33.27	(1.54)	25.98	(2.05)
Prestudy Pred. 1 Easy	19.58	(.76)	18.58	(.58)	13.18	(.82)
Prestudy Pred. 1 Medium	12.06	(.78)	11.58	(.74)	8.96	(.71)
Prestudy Pred. 1 Hard	7.03	(.77)	7.56	(.61)	6.35	(.70)
Global Poststudy Pred.	39.38	(2.66)	42.03	(2.22)	29.71	(2.61)
Poststudy Pred. Easy	19.97	(.71)	21.63	(.78)	17.19	(.97)
Poststudy Pred. Medium	14.50	(.61)	13.66	(.73)	11.21	(.67)
Poststudy Pred. Hard	9.80	(.75)	8.87	(.75)	7.04	(.71)

Table 2 (continued).

Note. Entries are means (and standard errors) of individuals' metamemory judgments; Younger -45 = younger adults in the 45 second experimental condition; Younger-90= younger adults in the 90 second control condition; EOL = ease of learning judgment; easy = normatively easy items; medium = normatively moderately difficult items; hard = normatively difficult items; Pred= prediction.

Predictions

Participants were asked to provide global and differentiated predictions immediately after completing their goal reports to differentiate what they *hoped* (i.e., their goals) and actually *expected* to achieve in the criterion task. Participants provided predictions prior to study (prestudy predictions), after study but before the recall test (poststudy predictions), and for a hypothetical trial 2 after they had attempted recall of the vocabulary pairs (post-recall predictions), thus allowing examination of how predictions changed with task experience. Tables 2 and 3 contain these predictions.

Prestudy Predictions

Younger adults predicted significantly higher recall performance than older adults, globally, F (1, 93) = 15.68, $p < .001$, partial $\eta^2 = .144$, and on easy, F (1, 93) = 43.72, $p < .001$, partial $\eta^2 = .320$, and moderately difficult items, F (1, 93) = 9.87, $p < .01$, partial $\eta^2 = .096$, but not on the most difficult items, F (1, 93) = 1.22, $p > .05$, partial $\eta^2 = .013$.

Poststudy Predictions

After participants studied items, but before they attempted recall, younger adults predicted significantly higher recall performance than older adults globally and for all item types, F (4, 90) = 4.78, $p < .01$, partial $\eta^2 = .175$, (see Table 2).

Table 3. Mean DJOLs, Postdictions, and Post-Recall Goals and Predictions

	Younger- 45		Younger- 90		Older	
Judgment	*M*	*(SE)*	*M*	*(SE)*	*M*	*(SE)*
Global DJOL	63.12	(3.65)	72.12	(1.83)	50.50	(3.18)
DJOL Easy	88.80	(4.49)	97.79	(.51)	85.59	(4.03)
DJOL Medium	66.84	(4.26)	77.41	(2.43)	51.80	(4.60)
DJOL Hard	33.71	(4.42)	41.17	(3.25)	14.07	(2.42)
Global Postdiction	39.81	(2.26)	40.59	(2.60)	32.58	(2.70)
Postdiction Easy	20.70	(.55)	20.84	(.65)	17.29	(.90)
Postdiction Medium	12.97	(.73)	12.66	(.70)	10.95	(.73)
Postdiction Hard	6.33	(.68)	7.41	(.83)	6.90	(.77)
Global Goal 2	51.44	(2.73)	52.44	(2.67)	38.42	(3.28)
Goal 2 Easy	22.15	(.91)	22.82	(.76)	21.77	(.59)
Goal 2 Medium	17.76	(.77)	18.11	(.66)	15.12	(.81)
Goal 2 Hard	13.27	(.94)	13.79	(.69)	9.71	(.86)
Global Pred. 2	49.05	(2.17)	50.92	(1.68)	38.31	(2.62)
Pred. 2 Easy	22.50	(.56)	22.98	(.34)	19.83	(.72)
Pred. 2 Medium	18.13	(.77)	17.49	(.77)	14.60	(.82)
Pred. 2 Hard	11.96	(.92)	12.65	(1.03)	9.19	(.92)

Note. Entries are means (and standard errors) of individuals' metamemory judgments; Younger-45 = younger adults in the 45 second experimental condition; Younger-90= younger adults in the 90 second control condition; DJOL = delayed judgment of learning; easy = normatively easy items; medium = normatively moderately difficult items; hard = normatively difficult items; Goal 2 = post-recall goal for hypothetical trial 2; Pred 2 = predicted performance on hypothetical trial 2.

Post-Recall Predictions

Reliable age-related differences were found in participants' post-recall predictions because younger adults predicted higher recall performance than older adults both globally and for all item types, F (4, 90) = 6.06, $p < .001$, partial $\eta^2 = .212$ (see Table 3).

Item Selection

Of interest was whether individuals would initially select easy items as predicted by the RPL model or instead select difficult items, as the DRM would suggest. Also of interest was how participants' impressions of item difficulty would influence decisions to select or ignore different items, and whether subjective difficulty, as assessed by EOLs, or normative difficulty, as determined by the Kornell and Metcalfe (2006) norms, was more predictive of item selection behaviors. These issues were examined by comparing item selection behaviors across age groups while taking into account the time constraints (i.e., 45 versus 90 seconds) placed on participants as well as which items the prestudy recall results indicated participants already knew before beginning the study phase.

Selection Order

To evaluate whether participants selected items in a manner more consistent with the RPL model or the DRM, the item selection data were first analyzed as Metcalfe and Kornell (2005) had done. That is, all items normatively classified as easy were assigned a value of "1", moderately difficult items a value of "2", and difficult items a value of "3". The mean of all items selected first was then calculated (collapsing across grids), as was the mean for selections 2 through 9. If participants selected easier items first, as the RPL model predicts, then the mean of the first selection should be a number close to 1. If

instead participants selected difficult items first, as the DRM predicts, then the mean of the first selection should be a number close to 3.

As may be seen in Table 4, which contains the mean values of participants' first 9 selections, participants in both age groups selected easier items first, and then over selections slowly transitioned to selecting more difficult items, in keeping with the RPL model's predictions. However, older adults transitioned to choosing the more difficult items more slowly than younger adults. Thus, younger adults' mean level of selection was reliably higher than older adults' for the first six selections made. However, by the seventh selection, through the ninth selection, both younger and older adults had transitioned to selecting the more difficult items so there were no longer any age-related differences in the level of items selected.

The same pattern in selection behaviors was found when participants' prior knowledge of Spanish was taken into account by excluding all items that were gotten correct on the prestudy recall test and examining the mean level of difficulty of unknown items selected. Thus individuals still selected easier items first before transitioning to the more difficult items. However, the mean level of difficulty of each selection was slightly higher after accounting for prior knowledge, both because individuals were more likely to know the easier items, but also because cases where individuals selected an item for a brief period of time simply to check their prestudy recall test response were likely excluded.

Table 4. Mean Difficulty Level of First Nine Selections

	Younger- 45		Younger- 90		Older	
Selection	*M*	*(SE)*	*M*	*(SE)*	*M*	*(SE)*
			Ignoring Prior Knowledge			
1	1.67	(.06)	1.65	(.07)	1.27	(.06)
2	1.91	(.10)	1.89	(.08)	1.45	(.08)
3	2.01	(.09)	2.05	(.08)	1.52	(.09)
4	2.36	(.06)	2.47	(.06)	2.17	(.06)
5	2.55	(.08)	2.65	(.06)	2.29	(.07)
6	2.58	(.09)	2.70	(.09)	2.33	(.08)
7	2.93	(.05)	2.80	(.08)	2.66	(.10)
8	2.98	(.03)	2.90	(.05)	2.81	(.09)
9	3.00	(.00)	2.90	(.07)	2.75	(.10)
			Accounting for Prior Knowledge			
1	2.30	(.21)	1.89	(.11)	1.63	(.36)
2	2.32	(.14)	2.14	(.09)	1.95	(.15)
3	2.28	(.08)	2.24	(.06)	1.98	(.16)
4	2.38	(.06)	2.58	(.06)	2.23	(.06)
5	2.60	(.08)	2.72	(.05)	2.33	(.08)
6	2.65	(.10)	2.78	(.09)	2.38	(.09)
7	2.98	(.02)	2.80	(.08)	2.71	(.10)
8	2.98	(.03)	2.91	(.06)	2.90	(.09)
9	3.00	(.00)	2.90	(.07)	2.85	(.12)

Note. Mean selection level where 1 = easy, 2 = moderately difficult and 3 = difficult.

Item Selection and EOLs

The items participants selected for study were examined to see if they had been given different EOLs than items participants chose to ignore. Collapsing across age group and difficulty level, it was found that participants selected items they perceived as more difficult (*M* of items selected = 5.34, *SE* = .03), and chose to ignore easier items (*M* of items ignored = 3.17, *SE* = .07). When the selection/EOL relationship was examined separately for each age group, younger adults were found to ignore items of an easier level than older adults, but to select items of a similar level. Items in the normatively easy and moderately difficult categories (Kornell & Metcalfe, 2006) were more likely to be selected by both younger and older adults if they had been given higher rather than lower EOLs. In contrast, normatively difficult items were more likely to be selected for study if they had been given lower, rather than higher, EOLs (see Table 5).

Differences were found in the EOL/selection behaviors of younger adults allotted 45 seconds versus those given additional time. As shown in Table 5, younger adults in both time conditions selected the more difficult of easy items. However, whereas younger adults in the 90 second condition continued to do the same (i.e., select the harder items) for moderately difficult items, those in the 45 second condition ignored those moderately difficult items to which they had given higher EOLs and focused their time on moderately difficult items with lower EOLs. Moreover, although the selections of younger adults in both time conditions reflected a tendency to ignore the most difficult of the difficult items, items ignored by younger adults in the 45 second condition had higher EOLs than those ignored by younger adults in the 90 second condition.

These patterns held both when prior knowledge of the Spanish language was ignored and when it was accounted for by excluding all items that individuals had gotten correct on the prestudy recall test. Only the EOLs associated with items in the normatively easy and moderately difficult categories changed when selections accounted for prior knowledge. Mean EOLs for selected and ignored items changed more in the easy than the moderately difficult category, and not at all in the difficult category, which was expected because participants were more likely to know or correctly guess what the easy items were than items in the other categories.

Participants' selection behaviors were consistent with both the DRM and RPL model's prediction that individuals seek to exclude known items to instead focus on the subset of unknown. However, the fact that individuals were more likely to ignore the easiest of the easy and moderately difficult items, but ignore the most difficult of the difficult items is more consistent with RPL model predictions than the DRM. Specifically, individuals in both age groups seemed to be trying to select items they thought they were capable of learning (i.e., within their range of proximal learning), while excluding those they deemed easy or too difficult to learn given the time constraints placed on them.

Table 5. Mean EOLs for Non-selected and Selected Items

EOL Type	Non-Selected *M*	Non-Selected *(SE)*	Selected *M*	Selected *(SE)*
		Younger Adults – 45 Seconds		
Overall	4.76	(.19)	4.82	(.09)
Easy Items	2.11	(.07)	3.16	(.12)
Moderate Items	5.00	(.36)	4.58	(.07)
Difficult Items	7.18	(.14)	6.73	(.07)
		Younger Adults – 90 Seconds		
Overall	4.10	(.27)	4.67	(.09)
Easy Items	1.95	(.06)	2.90	(.13)
Moderate Items	3.56	(.46)	4.47	(.08)
Difficult Items	6.79	(.30)	6.63	(.07)
		Older Adults		
Overall	4.70	(.17)	5.60	(.06)
Easy Items	2.14	(.11)	3.98	(.10)
Moderate Items	4.68	(.68)	5.44	(.09)
Difficult Items	7.77	(.14)	7.34	(.07)

Note. Entries are means (and standard errors) of individuals' metamemory judgments as a function of whether items were selected for study or not. EOL = ease of learning judgment; easy = normatively easy items; moderate = normatively moderately difficult items; difficult = normatively difficult items.

Item Selection and Subjective Versus Objective Difficulty

The prior analyses examined whether subjective difficulty ratings varied as a function of normative difficulty level for items selected and ignored by participants. Also of interest was whether participants' subjective perception of item difficulty (i.e., EOLs) or objective (i.e., normative) difficulty was more predictive of the order in which items were selected. To examine this issue, Goodman-Kruskal gamma correlations were calculated 1) between participants' EOLs and the order in which items were selected, 2) between selection order and EOLs placed into bins that corresponded with normative difficulty levels (e.g., EOLs of 1-3 were placed in bin "1", 4-6 in bin "2", and 7-9 in bin "3"), and 3) between normative difficulty level (i.e., 1, 2, or 3) and selection order.

If participants made global distinctions between items they deemed easy, moderately difficult and difficult, then the EOL bins should produce the highest gamma correlations. If participants were instead making more fine tuned judgments, based on their own item-level EOLs, when deciding which items they would select and in what order, then the EOLs should produce the highest correlations. However, if normative difficulty, as determined by Kornell and Metcalfe (2006), was more predictive of selection order then the norm gamma should be the highest.

As may be seen in Table 6, the gammas for individuals in both age groups were much higher when calculated based on normative difficulty and selection order than on subjective difficulty. However, both younger and older adults had higher "bin" gammas than item-level EOL gammas, and younger adults in the 90 second condition had higher gammas across all three types than participants in the other age group/ time condition. The gamma correlations did not change when prior knowledge of Spanish was accounted

for (i.e., when known items on the prestudy recall test were excluded), perhaps because individuals were still likely to select easier items before the moderately difficult and difficult items.

Table 6. Gamma Correlations

	Younger- 45		Younger- 90		Older	
Gamma Type	*M*	*(SE)*	*M*	*(SE)*	*M*	*(SE)*
Item-level EOLs/Order	.50	(.04)	.57	(.02)	.33	(.06)
Bin EOLs/ Order	.55	(.04)	.64	(.02)	.39	(.06)
Norm EOLs/Order	.72	(.06)	.83	(.03)	.54	(.08)
Order 1/ Order 2	.40	(.07)	.41	(.05)	.47	(.04)
Order 2/ Order 3	n/a		n/a		.74	(.04)
EOLs/Prestudy Recall	.86	(.01)	.86	(.01)	.76	(.03)
EOLs/Poststudy Recall	.64	(.03)	.63	(.02)	.66	(.03)
DJOLs/Prestudy Recall	.82	(.07)	.95	(.01)	.89	(.03)
DJOLs/Poststudy Recall	.85	(.06)	.90	(.02)	.83	(.07)

Note. Entries are means (and standard errors) of Goodman-Kruskal gamma correlations; Younger-45 = younger adults in the 45 second experimental condition; Younger-90 = younger adults in the 90 second control condition; Item-level EOLs/Order = gammas calculated between selection order and participants' item level EOL judgments; Bin EOLs/ Order = gammas calculated between selection order and EOLS placed into easy, moderate and difficult bins; Norm EOLs/ Order = gammas calculated between selection order and normative item difficulty; Order1/Order 2= gammas calculated between the selection order for the first time items were selected for study and the second time. Order 2/ Order 3 = a gamma calculated between the selection orders for items selected more than twice.

Together, these results suggest that individuals attended to having been told which items were normatively easy, moderately difficult, and difficult, more than they attended to their own subjective impressions of item difficulty. However, when individuals attended to their subjective perceptions of item difficulty in order to select items, they seemed to do so at a global, rather than a fine-tuned, level as indicated by the bin gammas being higher than the item-level EOL gammas were. In addition, the fact that younger adults given additional study time had higher gammas suggests they might have been able to take the time to strategically choose which items they would select, something younger adults in the 45 second condition did not have time to do and the older adults either did not attempt to or could not do. Finally, the fact that the gammas were positive indicates that individuals were more likely to select easier items first and the more difficult items later, providing additional evidence for RPL-consistent selection behaviors.

Reselected Items

Additional gammas were calculated between the first time individuals selected an item and additional selections of the same item to examine whether gamma values would increase or decrease (see Table 6). If gamma values increase then this would suggest that individuals become increasingly likely to reselect items in a similar order the second time around (i.e., easy, moderately difficult, then difficult in both selection orders as opposed to easy items selected first in order 1, but last in additional selection orders), which could reflect drop-out of items already deemed learned. If gamma values were instead found to decrease, it would suggest that individuals are reselecting items in a random fashion, ignoring difficulty level.

All order gammas were found to be positive and of a moderate size, suggesting individuals selected easy, then moderately difficult, then difficult items, in keeping with RPL model predictions. The gammas calculated between selection order 1 and selection order 2 were found to be slightly but not significantly larger for older adults than younger adults, F (1, 68) = 1.27, $p > .05$, partial $\eta^2 = .018$. Gamma values increased across multiple selections of items (i.e., when gammas were calculated between orders 2 and 3, and 3 and 4), suggesting that items were being dropped on each successive selection order, which served to increase the gamma correlation for items that continued to be restudied.

It is important to note that the interpretation of these order gammas was made difficult because of age- and time-related differences that emerged in item reselection behaviors. Specifically, older adults were more likely than younger adults to reselect items a second time (older adults' M number of reselections = 11.71, SE = 1.84 versus younger adults' M = 4.39, SE = .66), F (1, 93) = 21.05, $p < .001$, partial $\eta^2 = .185$, and younger adults with more time reselected items a second time more often than those with less time (90 second younger adults' M number of reselections = 5.69, SE = 1.09 versus 45 second younger adults' M = 3.09, SE = .70), F (1, 62) = 4.00, $p = .05$, partial $\eta^2 =$ .060. No other time-related differences were found in younger adults' reselection behaviors, however additional age-related differences were observed in the tendency to reselect items a third, F (1, 93) = 16.54, $p < .001$, partial $\eta^2 = .151$, and a fourth time, F (1, 93) = 6.15, $p < .05$, partial $\eta^2 = .062$, because only one younger adult, in the longer time condition, reselected a single item a third time (M = .02, SE = .02) compared to older adults who reselected items for a third time on average 2.13 (.75) times, and only

older adults selected items a fourth time (M = .42, SE = .24). Thus, because fewer and fewer individuals' item selection behaviors contributed to the gamma correlations, the values should have increased (e.g., because a single person's behavior is more likely to be consistent than multiple individuals' selection behaviors). It therefore remains an open question whether the increased gamma values were an artifact of item dropout, participant dropout, or both as fewer and fewer items were reselected by fewer and fewer individuals.

Study Time Allocation

The amount of time participants allocated to easy, moderately difficult, and difficult items was examined within and across grids to determine if there were age- or time-related differences in how participants chose to use their allotted study time. Analyses of the total amount of time allocated to studied items within each difficulty level revealed that participants in both age groups (and time conditions for younger adults) allocated the most time to difficult items and the least amount of time to easy items. However, as may be seen in Table 7, older adults allocated significantly more time to easy items than younger adults did, F (1, 93) = 37.22, $p < .001$, partial $\eta^2 = .286$, whereas younger adults allocated significantly more time to the moderately difficult items than the older adults, F (1, 93) = 17.52, $p < .001$, partial $\eta^2 = .159$. No age-related differences existed in the amount of time devoted to the difficult items, F (1, 93) = 1.58, $p > .05$, partial $\eta^2 = .017$.

Reliable differences emerged in time allocation behaviors as a function of how much total time participants were allotted. Specifically, younger adults in the longer time condition allocated significantly more time to studying the difficult items than did

younger adults with less time, F (1, 62) = 19.73, $p < .001$, partial $\eta^2 = .241$. This yielded a reliable difference in the overall amount of study time allocated to all items, F (1, 62) = 13.29, $p < .01$, partial $\eta^2 = .176$, with those in the longer time conditions spending more time studying items overall than younger adults who only had 45 seconds to study items. Thus whereas younger adults allocated similar amounts of time to easy and moderately difficult items, regardless of time allotted to them, younger adults in the longer time condition took advantage of the additional time to focus on the more difficult items, making their time allocation for difficult items more similar in magnitude to the older adults' who also received 90 seconds.

When study time allocation was examined at the grid level, younger adults initially allocated similar amounts of time to easy and moderately difficult items. However, by the third grid, younger adults in the 90 second condition began to spend increasing amounts of time on the moderately difficult and difficult items, thus making their times for these items more like the older adults than their younger 45 second counterparts. Moreover, by grids 7 and 8, the older adults actually allocated more time to the moderately difficult items than the younger adults in either condition, whereas the younger adults with 90 seconds utilized that time to focus on the most difficult items, spending more time on them than participants in either age group. These differences in study time allocation across age groups/time conditions, grids, and levels of difficulty yielded a reliable interaction between Grid, Difficulty level, and Condition, F (28, 160) = 1.56, $p < .05$, partial $\eta^2 = .214$, when analyzed using repeated measures GLM analyses. Significant two-way interactions were also observed between Grid number and Difficulty level, F (14, 79) = 3.89, $p < .001$, partial $\eta^2 = .408$, between Age/Time Condition and

Difficulty level, F (4, 184) = 5.22, $p < .01$, partial $\eta^2 = .102$ and between Grid number and Age/Time Condition, F (14, 174) = 1.86, $p < .05$, partial $\eta^2 = .130$.

Table 7. Mean Study Time Allocation

	Younger- 45		Younger- 90		Older	
Grid/ Item Type	*M*	*(SE)*	*M*	*(SE)*	*M*	*(SE)*
Grid 1						
Easy	4.19	(1.60)	2.56	(.70)	9.10	(1.90)
Medium	15.22	(1.55)	15.41	(1.52)	23.68	(2.44)
Hard	15.44	(1.31)	28.53	(3.22)	31.74	(3.80)
Grid 2						
Easy	2.28	(.45)	2.25	(.66)	4.71	(.61)
Medium	15.78	(1.53)	17.97	(1.90)	29.42	(2.81)
Hard	18.69	(1.37)	29.25	(2.76)	31.42	(3.75)
Grid 3						
Easy	1.09	(.39)	1.19	(.31)	4.23	(.66)
Medium	16.03	(1.16)	19.53	(1.93)	24.55	(1.80)
Hard	19.75	(1.19)	31.34	(2.89)	31.65	(3.43)
Grid 4						
Easy	1.50	(.26)	1.44	(.32)	3.61	(.45)
Medium	14.94	(1.41)	21.31	(2.36)	26.68	(2.57)
Hard	20.38	(1.59)	34.19	(3.29)	29.06	(3.73)

Table 7 (continued).

Grid 5						
Easy	1.91	(.67)	1.41	(.39)	5.68	(.77)
Medium	13.78	(1.16)	20.88	(2.02)	25.45	(2.91)
Hard	20.16	(1.49)	33.78	(3.21)	32.42	(4.17)
Grid 6						
Easy	1.66	(.43)	1.31	(.28)	4.68	(.65)
Medium	16.06	(1.38)	19.37	(1.86)	25.68	(2.49)
Hard	19.19	(1.26)	33.66	(2.97)	30.19	(4.00)
Grid 7						
Easy	1.78	(.41)	1.63	(.38)	6.90	(.73)
Medium	12.84	(.98)	16.41	(1.92)	23.06	(2.83)
Hard	21.78	(1.48)	38.06	(3.70)	31.00	(4.12)
Grid 8						
Easy	1.25	(.31)	1.38	(.45)	4.74	(.94)
Medium	15.69	(1.37)	17.91	(1.88)	26.55	(2.81)
Hard	21.03	(1.43)	36.75	(3.58)	28.23	(3.76)

Note. Values are means and standard errors of how much time individuals in each age group/ time condition allocated to items of each difficulty level; Younger-45= younger adults in the 45 second experimental condition; Younger-90= younger adults in the 90 second control condition; Easy = normatively easy items; Medium = moderately difficult items; Hard = difficult items.

Percentage of Time Allocated to Each Level

Because younger adults in the 45 second time condition had half as much time to study items as younger and older adults that received 90 seconds, of interest was whether

individuals would allocate the same percentage of time to items in each difficulty level. To examine this issue, the amount of time participants allocated to items of each type was divided by the total amount of time they spent studying all items.[1] These analyses revealed that older adults allocated proportionally more time to easy items, F (1, 93) = 24.74, $p < .001$, partial $\eta^2 = .210$, and moderately difficult items, F (1, 93) = 5.70, $p < .05$, partial $\eta^2 = .058$, than younger adults in either time condition. However, younger adults allocated significantly more time to the difficult items than did older adults, F (1, 93) = 15.31, $p < .001$, partial $\eta^2 = .141$.

Separate analyses of only younger adults' data revealed that the percentage of time allocated to easy items was similar regardless of time constraints, F (1, 62) = 2.92, $p > .05$, partial $\eta^2 = .045$, whereas those in the 45 second condition allocated, on average, a reliably larger percentage of their time to moderately difficult items than younger adults given 90 seconds did, F (1, 62) = 5.77, $p < .05$, partial $\eta^2 = .085$. In contrast, those given 90 seconds allocated significantly more time to difficult items than those given less time, F (1, 62) = 7.81, $p > .01$, partial $\eta^2 = .112$. Thus, whereas younger adults with 45 seconds allocated 5% (1.00) of their time to easy items, 41% (2.00) to moderately difficult, and 54% (2.00) to the most difficult, younger adults with 90 seconds to allot spent 3% (1.00) on easy items, 36% (1.00) on moderately difficult, and 61% (1.00) on difficult items. In contrast, older adults devoted 11% (2.00), 44% (3.00), and 45% (4.00) to items in each difficulty level, respectively. Thus, younger and older adults selected items in a similar manner (i.e., easiest first), but allocated time differently with older adults allocating proportionally more time to easy and moderately difficult items than younger adults who instead allocated significantly more of their time to the difficult items. Yet, on average

all participants allocated more (absolute and relative) time to the more difficult items than easier items.

Strategy Use

Because time allocation varied across age groups and time conditions, of interest was how individuals were utilizing this time to encode the vocabulary pairs. As previously noted, even if individuals selected the same items and allocated the same amount of time to those items, if one engaged in more effective processing (e.g., by using a normatively more effective strategy) then that individual could produce higher recall rates than the individual that used a normatively ineffective strategy.

Participants' strategy reports were examined as a function of age and allotted time to see if these factors influenced the types of strategies individuals reported using for items in each difficulty level. Strategy reports for restudied items were compared and normatively less effective strategies ignored if a more effective strategy was reportedly used either before or after the less effective strategy. Table 8 contains the proportion of times each strategy was reportedly used to study easy, moderately difficult and difficult items, for both age groups, and for younger adults in both time conditions.

Across both age groups and time conditions, collapsing across difficulty level, individuals were most likely to report using the keyword method strategy to encode items. This was a surprising finding, given the additional effort required to use this strategy and prior research in which younger adults were found unlikely to use more effortful strategies (Sagarra & Alba, 2006) and to instead rely on the normatively less effective rote repetition strategy.

Table 8. Reported Strategy Use

	Younger- 45			Younger- 90			Older		
Strategy/ Level	E	M	D	E	M	D	E	M	D
Rote Repetition	.17	.22	.32	.07	.14	.28	.09	.23	.31
Sentence Gen.	.01	.13	.09	.02	.18	.16	.06	.15	.15
Imagery	.04	.14	.16	.01	.18	.15	.08	.17	.19
Keyword Method	.50	.33	.21	.68	.39	.20	.45	.30	.14
Combination	.01	.06	.05	.01	.05	.09	.01	.01	.01
Not implemented	0	.02	.05	0	.01	.05	.01	.03	.05
None	.26	.10	.13	.23	.06	.06	.31	.11	.16

Note. Entries are the proportion of times each strategy was used to study items of each difficulty level; Younger-45= younger adults in the 45 second experimental condition; Younger-90= younger adults in the 90 second control condition; Sentence gen. = sentence generation; Imagery = interactive imagery; Combination = combination of rote and either imagery, sentence generation or the keyword method; Not implemented = means participant tried to implement a strategy but was unable to effectively do so; None= means no strategy was attempted or used.

When overall strategy use was examined within each level of difficulty, the keyword method was again the most prevalent strategy used for easy and moderately difficult items, but for difficult items individuals were more likely to use rote repetition. This held true for older and younger adults, both ignoring and accounting for time allotted, however, younger adults given 45 seconds were significantly more likely to use rote for moderately difficult items than younger adults given more time, F (1, 62) = 4.40, $p < .05$, partial $\eta^2 = .066$. Younger adults given the additional study time utilized it to select and use all the normatively effective strategies (i.e., keyword method, interactive

imagery and sentence generation) at a higher rate than younger adults given less time, but only the difference in sentence generation use for difficult items was reliable, F (1, 62) = 9.95, $p < .01$, partial $\eta^2 = .138$. So although younger adults in both time conditions utilized effective strategies, more so for easy and moderately difficult items than difficult, the increased time pressure placed on younger adults in the 45 second condition resulted in them being more likely to report using an ineffective strategy (rote repetition) or no strategy at all.

All participants frequently reported not using a strategy on easy items, perhaps reflecting the perception that a strategy was not needed due to many of the easy Spanish words being cognates for English words with which they were familiar. Yet, for easy items older adults were significantly more likely to report using no strategy than younger adults in either time condition, F (1, 93) = 9.29, $p < .01$, partial $\eta^2 = .091$, with similar trends for moderately difficult, F (1, 93) = 2.24, $p > .05$, partial $\eta^2 = .024$, and difficult items, F (1, 93) = 2.46, $p > .05$, partial $\eta^2 = .026$. Older adults were also more likely than younger adults to report that they unsuccessfully tried to implement a strategy, F (1, 93) = 4.32, $p < .05$, partial $\eta^2 = .040$, used sentence generation, F (1, 93) = 6.10, $p < .05$, partial $\eta^2 = .062$, or imagery for easy items, F (1, 93) = 6.49, $p < .05$, partial $\eta^2 = .065$. In contrast, younger adults reported significantly higher use of the keyword method for the difficult items, F (1, 93) = 5.78, $p < .05$, partial $\eta^2 = .058$. Thus older adults used a greater variety of strategies at a higher rate than younger adults for easy items, whereas younger adults were more likely to use the effective keyword method for difficult items. It is important to note however, that many younger and older adults had difficulty understanding the keyword method strategy and whether they should report using the

keyword method as opposed to no strategy for the easy items that were cognates for which they relied on the similarity of the Spanish and English words. Thus it remains possible that participants' confusion resulted in overestimation and/or underestimation of actual keyword method strategy use.

Delayed Judgments of Learning

Younger adults were more confident than older adults in their ability to recall moderately difficult, F (1, 92) = 17.63, $p < .001$, partial $\eta^2 = .161$ and difficult items, F (1, 92) = 28.75, $p < .001$, partial $\eta^2 = .238$ (see Table 3).

Recall Performance

Recall performance on both the prestudy and poststudy recall tests was scored initially by the computer. All items were then hand checked to ensure participants were not penalized for spelling errors. In addition, participants' responses were scored a second time using a lenient (i.e., gist-based) scoring system. The lenient score was included to allow examination of whether older adults were more likely to recall gist based information about the Spanish words, even if their responses were not considered correct when strictly scored. Thus each response was evaluated for its similarity to the correct response and items that were deemed to be semantically related in gist were considered correct in the lenient coding. For example, if participants typed *pregnant* instead of *pregnancy* their response would have been scored as incorrect for the strict score, but correct for the lenient score. Responses that were semantically related to the correct response, but seemed to reflect guessing as opposed to knowledge of what the word meant were scored as incorrect (e.g., responses such an *antique* for the word *antiguo,* or *brilliant* for the word *brillantez,* were considered incorrect and assumed to

reflect random guessing based on surface characteristics of the Spanish word rather than knowledge the words meant *ancient* and *brightness*, respectively). A list of acceptable and unacceptable gist words was formed and the author and one other coder hand scored items, with the author addressing any novel items that did not appear on the list of items deemed acceptable. In this way, all items were strictly and leniently scored.

The accuracy of participants' strict and gist-scored items was then analyzed using multivariate analyses of variance. Analyzing the data as a function of actual versus gist based responses did not change the pattern of recall results. Therefore, only the results based on the strict scoring are discussed, but means associated with both types of recall may be found in Table 9.

Prestudy Recall Test Performance

Younger adults' prestudy recall performance indicated they either knew more Spanish words upon entering the experiment or were more adept than older adults at deciphering/ guessing what the Spanish words meant, F (1, 93) = 16.73, $p < .001$, partial $\eta^2 = .152$. This was especially true for easy items, F (1, 93) = 21.00, $p < .001$, partial $\eta^2 = .184$, many of which were cognates. However, no age-related differences existed for moderately difficult items because participants in both age groups were unlikely to know or correctly guess their meaning, $F < 1$. Prestudy recall of difficult items was zero for both age groups. These differences yielded a reliable Age X Difficulty interaction, F (2, 92) = 11.24, $p < .001$, partial $\eta^2 = .196$.

Poststudy Recall Test Performance

Participants' poststudy recall performance indicated that younger adults learned more both overall and within each difficulty level than older adults, F (3, 91) = 15.93, $p <$

.001, partial η^2 = .344. Importantly, although younger adults in the 90 second condition had slightly higher recall than 45 second younger adults for moderately difficult and difficult items, these differences were not reliable, F (1, 62) = 2.53, $p > .05$, partial η^2 = .039 and F (1, 62) = 0.98, $p > .05$, partial η^2 = .015, respectively.

Table 9. Prestudy and Poststudy Recall Test Performance

	Younger- 45	Younger- 90	Older
Test/Level	*M* *(SE)*	*M* *(SE)*	*M* *(SE)*
Prestudy Test			
Easy	.85 (.02)	.88 (.01)	.71 (.04)
Medium	.02 (.01)	.04 (.01)	.03 (.01)
Hard	0	0	0
Overall Actual	.29 (.01)	.31 (.01)	.25 (.02)
Overall Gist	.30 (.01)	.31 (.01)	.25 (.02)
Poststudy Test			
Easy	.97 (.01)	.98 (.01)	.88 (.02)
Medium	.64 (.04)	.72 (.03)	.39 (.05)
Hard	.27 (.04)	.33 (.04)	.09 (.02)
Overall Actual	.63 (.03)	.68 (.02)	.45 (.03)
Overall Gist	.65 (.03)	.70 (.02)	.45 (.03)

Note. Entries are means (and standard errors) of individuals' prestudy and poststudy recall test performance; Younger-45 = younger adults in the 45 second experimental condition; Younger-90= younger adults in the 90 second control condition; Easy = normatively easy items; Medium = normatively moderately difficult items; Hard = normatively difficult items; Actual = strict scoring; Gist = lenient scoring.

Poststudy Recall Test Performance and Strategy Use

The recall performance of participants was also evaluated as a function of which strategies participants had reported using during encoding to determine if the conditional probability of recalling items was higher when normatively effective strategies were used than when normatively ineffective or no strategies were used.

The values reflecting the conditional probability of recalling items at each level of difficulty after using each type of strategy appear in Table 10. Collapsing across age groups and timing conditions the keyword method resulted in the greatest probability of recall, in keeping with prior research (Gruneberg & Pascoe, 1996; Sagarra & Alba, 2006). However, which strategies produced the best recall varied as a function of difficulty level such that younger adults were likely to recall easy items regardless of which strategy was used, whereas older adults had higher recall of easy items after using sentence generation, the keyword method, or rote repetition. Younger and older adults had higher recall of moderately difficult and difficult items if they used a normatively effective strategy (e.g., imagery, sentence generation, or the keyword method). However, the trend was for younger adults to have higher recall than older adults across all difficulty levels, regardless of strategy use, with the exception being recall of easy items encoded with rote repetition and difficult items encoded using a combination of strategies, $F < 1$.

Because younger adults recalled more than older adults across all difficulty levels, age-related differences in the conditional recall values were only examined after collapsing across difficulty level. Age-related trends were apparent, but not reliable, for sentence generation, F (1, 71) = 2.55, $p > .05$, partial $\eta^2 = .035$, a combination of

strategies, F (1, 32) = 1.79, $p > .05$, partial $\eta^2 = .053$, and no strategy, F (1, 62) = 2.30, $p > .05$, partial $\eta^2 = .036$, but not imagery, $F < 1$. Only the rote, F (1, 79) = 8.65, $p < .01$, partial $\eta^2 = .099$, keyword method, F (1, 85) = 10.86, $p < .01$, partial $\eta^2 = .113$, and no implementation strategies yielded reliable age-related differences in recall, F (1, 34) = 4.31, $p < .05$, partial $\eta^2 = .112$.

Ironically, after collapsing across difficulty levels, younger adults given less time had higher conditional probability of recall with all strategies, but especially normatively effective ones (i.e., keyword method, sentence generation, interactive imagery) than younger adults given 90 seconds. Although these time-related differences were not reliable, $F < 1$, the fact that younger adults given half the time were able to achieve similar recall levels as those given more time is impressive and suggests that younger adults in the 45 second condition were able to utilize the effective strategies to compensate for having less time to study items. Thus despite the fact that younger adults in the 90 second condition both allocated more time to the difficult items and reported using the keyword method at a higher rate than younger adults given 45 seconds, the additional time spent on difficult items would seem to have been a labor in vain (Nelson & Leonesio, 1988) since the additional time did not yield significantly higher recall of these items. The labor in vain effects are consistent with Metcalfe and Kornell's (2003) findings with younger adult novices who also failed to benefit from allocating more time to difficult items. Thus in both the present study and the Metcalfe and Kornell study, those given less time actually had a higher rate of return in recall performance for each second spent studying items than those given more time to study items.

Table 10. Conditional Probability of Recall as a Function of Strategy Used

	Younger- 45	Younger- 90	Older
Strategy	*M* (*SE*)	*M* (*SE*)	*M* (*SE*)
Rote Repetition			
Easy	.87 (.09)	.96 (.04)	.87 (.07)
Medium	.52 (.07)	.50 (.08)	.29 (.07)
Hard	.29 (.06)	.20 (.04)	.05 (.02)
Sentence Gen.			
Easy	1.00 (.00)	1.00 (.00)	.88 (.07)
Medium	.61 (.07)	.73 (.05)	.51 (.09)
Hard	.51 (.10)	.40 (.05)	.18 (.08)
Imagery			
Easy	1.00 (.00)	1.00 (.00)	.71 (.12)
Medium	.76 (.06)	.69 (.06)	.55 (.09)
Hard	.46 (.09)	.39 (.07)	.20 (.08)
Keyword Method			
Easy	.95 (.03)	.99 (.01)	.88 (.03)
Medium	.76 (.03)	.80 (.04)	.45 (.06)
Hard	.34 (.06)	.36 (.06)	.16 (.06)
Combination			
Easy	1.00 (.00)	1.00 (.00)	.50 (.00)
Medium	.60 (.12)	.68 (.13)	.33 (.33)
Hard	.18 (.10)	.34 (.10)	.17 (.11)

Table 10 (continued).

Not implemented			
Easy	n/a	n/a	.00 (.00)
Medium	.54 (.18)	.20 (.20)	.09 (.05)
Hard	.11 (.06)	.17 (.08)	.01 (.01)
None			
Easy	.99 (.01)	1.00 (.00)	.80 (.07)
Medium	.45 (.11)	.55 (.11)	.33 (.10)
Hard	.23 (.07)	.17 (.08)	.04 (.02)

Note. Entries are means (and standard errors) of individuals' poststudy recall test performance as a function of which strategy was used to encode the item (i.e., the conditional probability of recall); Younger-45 = younger adults in the 45 second experimental condition; Younger-90 = younger adults in the 90 second control condition; Sentence gen. = sentence generation; Imagery = interactive imagery; Combination = combination of rote and either imagery, sentence generation or the keyword method; Not implemented = means participant tried to implement a strategy but was unable to effectively do so; None= means no strategy was attempted or used; Easy = normatively easy items; Medium = normatively moderately difficult items; Hard = normatively difficult items.

Postdictions

Participants' estimates of how many items they got correct, both globally and for each level of difficulty, were analyzed for age- and time-related differences. Younger adults believed they recalled a higher number of easier, F (1, 93) = 16.02, $p < .001$, partial $\eta^2 = .147$, and moderately difficult items than older adults, F (1, 93) = 4.45, $p < .05$, partial $\eta^2 = .046$. However, no age-related differences were found for difficult items, $F < 1$, reflecting both age groups' lack of confidence in having recalled the more difficult

items. No differences were found in younger adults' postdicted performance as a function of how much time they had been allotted to study items, $F < 1$ (see Table 3).

Relative Accuracy of Metamemory Judgments

The accuracy of participants' metamemory judgments may be measured both in terms of absolute accuracy (i.e., the alignment between participants' mean recall performance and mean judgments) and relative accuracy (i.e., resolution), which assesses how well individuals differentiated, at the item level, those items they would be more or less likely to recall. Because participants' item-level impressions were considered to be more likely to influence self-paced study decisions, analyses focused on relative rather than absolute accuracy (but see Appendix B for mean absolute accuracy values).

Goodman-Kruskal gamma correlations were calculated between participants' recall performance and their metamemory judgments (i.e., EOLs and DJOLs). If participants recalled items they rated as easier to learn (EOLs) or for which they expressed greater confidence (DJOLs) then they would have high, positive gamma correlations, with perfect resolution reflected by a +1 gamma correlation. However, if individuals were unable to identify which items they would be more or less likely to learn and recall and gave low confidence ratings to items that were later recalled and high ratings to forgotten items, then individuals would have lower gamma correlations, with no resolution reflected by a gamma of 0. Of interest was whether there would be age-related differences in resolution and whether younger adults afforded extra study time would be better able to distinguish between items than those given less time.

EOLs

Younger adults had significantly higher resolution than older adults for the prestudy, F (1, 92) = 13.24, $p < .001$, partial $\eta^2 = .126$, but not the poststudy recall test, for which resolution did not differ reliably, $F < 1$ (see Table 6). As expected, EOLs correlated more highly with pretest knowledge of easy items than with poststudy recall performance, F (1, 92) = 89.49, $p < .001$, partial $\eta^2 = .493$. These differences across tests combined with the age-related differences to yield a reliable Age X Test interaction, F (1, 92) = 13.86, $p < .001$, partial $\eta^2 = .131$. Younger adults' gammas did not differ as a function of study time constraints. The same pattern of results was found when the data were analyzed as a function of gist recall.

DJOLs

No age- or time-related differences existed in the resolution of participants' DJOLs with poststudy recall, $F < 1$ (see Table 6). The pattern of results did not change when analyzed as a function of gist rather than strict recall.

Discussion

The present experiment provided the first evidence to date that older adults, like younger adults, selected items in a manner more consistent with RPL model than DRM predictions when presented with items that were heterogeneous with regard to difficulty. It also demonstrated that RPL selection order effects can be found regardless of whether the Spanish words are used as cues, as in the present experiment, or targets (e.g., as in Kornell & Metcalfe, 2006; Metcalfe, 2002; Metcalfe & Kornell 2003, 2005). That cue-target order did not matter could be because of the emphasis participants placed on normative difficulty information and because individuals were told about the normative

difficulty of items in both the present experiment and the Metcalfe experiments. The magnitude of the gammas calculated between selection order and either item-level or binned EOLs, or normative difficulty level indicated that normative difficulty critically influenced item selection decisions more than subjective difficulty. This is important because older adults perceived items of all types to be more difficult to learn than younger adults did, despite there being no age-related differences in participants' responses on the memory beliefs questionnaires. Perhaps the focus on normative rather than subjective difficulty is why younger and older adults selected items in a similar, RPL-consistent order, despite the age-related differences in EOLs.

Also in keeping with RPL model predictions was the impact that time constraints had on participants' study time allocation behaviors. Thus, although younger and older adults allocated significantly more time to the moderately difficult and difficult items than easier items, consistent with DRM rather than RPL predictions, younger adults given 45 seconds spent proportionally less time on difficult items than younger adults given 90 seconds. Instead, those with less time prioritized moderately difficult items to a greater extent than did those given more time, a finding that is more consistent with RPL than DRM predictions, the latter of which would predict more time for difficult (i.e., less well known) items regardless of time constraints.

Although not predicted by either the DRM or the RPL model, time constraints also impacted strategy use. Younger adults given more time reported using normatively effective strategies at a higher rate than younger adults given less time, but strategy use was similar across age groups making it unlikely that strategy use was the reason for older adults' significantly lower recall for all item types. Instead, differences in the

efficient allocation of study time seemed to account for both the age-related (Connor & Dunlosky, 1997) and time-related differences found in recall. Younger adults given 45 seconds to study each grid were more efficient in allocating time, as reflected by their faster transition to more difficult items and ability to utilize effective strategies, despite less time, to maximize the likelihood of recall. Thus the fact that younger adults given less time had similar recall rates as those given twice as much time would seem to suggest that giving younger adults additional time to study items only led them to labor further in vain (Nelson & Leonesio, 1988).

The data from the present experiment do not provide a definitive answer as to why the additional study time afforded difficult items by those in the 90 second condition did not yield significantly higher recall performance for those items. However, Metcalfe and Kornell's (2005) discussion of how time constraints interact with item difficulty and jROLs to influence study time allocation decisions may account for why those given 90 seconds allocated more time to difficult items. Moreover, their description of how the information uptake functions likely differ for easy, moderately difficult, and difficult items may account for why the additional time spent on difficult items was a labor in vain. Metcalfe and Kornell suggested that jROLs will have the greatest impact for difficult items when individuals are under time pressure, such that individuals will likely require higher rates of learning to continue studying difficult items as time pressure increases. Yet because difficult items have very shallow information uptake functions and require extended study time to be learned, the likelihood that the perceived rates of learning will be sufficient for those under greater time pressure to persist in studying these items is decreased. In contrast, because moderately difficult items have functions

that increase slowly and steadily for a long time, it is more likely that participants' jROLs will be sufficient for participants to continue studying these items. These predictions are consistent with the results obtained in the first experiment. That is, younger adults in the 45 second condition, who had half the time and used proportionally more of their allotted time to study items than did younger and older adults given 90 seconds (see Footnote 1), ceased studying the most difficult items earlier than those given 90 seconds, suggesting the greater time pressure caused them to set a different parameter stopping value for their jROLs than those given more time. If those given 90 seconds did in fact set a different stopping value, this would account for the longer study times and the shallow information uptake functions for difficult items would explain why the additional time did not benefit recall. Thus difficult items were either learned in a relatively short period of time by participants, regardless of time constraints, or were so difficult they did not benefit from additional study time because jROLs caused individuals to stop studying before additional learning gains were achieved.

The shift in the overall pattern of goal and prediction means across trials (i.e., from prestudy to post-recall for the hypothetical trial 2) suggested that older adults were aware that they were less likely to learn the more difficult items. Prestudy goals indicated they had lower expectations than younger adults, both overall and for easy items. However, post-recall goals and predictions indicated older adults expected to do as well as younger adults on easy items, while still expecting poor performance for more difficult items. That the smallest age differences existed for easy items suggests older adults recognized this and expected these items to be easier to (re)learn if studied again. Thus it seems possible that older adults' slower transition to studying difficult items was based

on metacognitive monitoring. Yet, younger adults given additional time also had slower transition rates to the more difficult items. Of interest then was whether participants, young and older, could be inspired to transition to the more difficult items sooner if given extrinsic motivation and experimenter-determined (as opposed to participant-determined) task goals. These questions were addressed in the second experiment.

CHAPTER 3

EXPERIMENT 2

Younger and older adults in Experiment 1 exhibited similar item selection (i.e., easier first) and similar study time allocation behaviors when given equal study time (i.e., younger and older adults given 90 seconds). These behaviors were presumably driven by participants' attempts to achieve their self-determined goals. It was unknown, however, whether younger and older adults would be able to implement, adapt to, and achieve experimenter-provided goals. Of interest was whether experimenter-provided goals and point values would serve to motivate participants' task performance. In particular, if individuals initially showed a tendency to select easy items for study, could assigning higher point values to the more difficult items cause individuals to change their item selection strategy such that difficult items are selected more than the easy items? The Castel et al. (2002) findings suggested it would be worthwhile to empirically evaluate whether point values would motivate or "push" younger and older adult participants to select items of a particular difficulty level.

Of additional interest was how point values would interact with experimenter-provided goals that stressed both the number of items recalled and number of points earned, and whether strategy use would differ for higher valued items than lower valued items. Younger and older adults in Experiment 1 were unlikely to use more effortful strategies and instead relied on rote repetition for difficult items. Would individuals show increased tendency to invest the time and energy necessary to utilize more effortful, but also more effective, encoding strategies (e.g., keyword method or sentence generation)

for high valued (difficult) items, but continue to study lower valued (difficult) items with ineffective strategies (e.g., rote repetition)? These questions were examined in the second experiment by manipulating point values and the experimenter-provided task goals participants were given. The manipulation of points and goals allowed direct examination of how well participants in each age group were able to exert metacognitive control. That is, because the task involved changing goal and point structures, participants had to strategically form and adapt their plans for item selection and study time allocation to achieve the goals and obtain the specified number of points.

Method

Design

The experiment was a 4 (Goal type: High word/High point goal-HH, High word/Low point goal- HL, Low word/High point goal- LH, and Low word/Low point goal- LL) X 3 (Item difficulty: easy, moderately difficult, difficult) X 3 (Point values: favor recall of easy items, difficult items, or are neutral) X 2 (age group: younger versus older) X 2 (Point order: easy items favored first versus difficult items favored first) design. All factors except age group and point order were manipulated within subjects. Goal type was fully crossed with the point value and item difficulty factors (see Figure 1) such that for each goal type, participants were asked to study 3 different 2 X 3 grids, each containing 2 easy, 2 moderately difficult, and 2 difficult Spanish-English vocabulary pairs, which had point values associated with them that either favored recall of the easy items, difficult items, or were neutral (i.e., did not favor recall of items from any particular difficulty level). Thus each of the four goal types had 3 grids, one with each type of point structure, associated with it to allow examination of how well individuals

were able to adapt their item selection and study time allocation behaviors in order to achieve the respective experimenter-provided goals.

Goal Type	Point Structure		
Words/Points	Favor Easy	Neutral	Favor Difficult

High/High

E	MD	D
12	8	4
10	6	2

E	MD	D
6	6	6
8	8	8

E	MD	D
2	6	10
4	8	12

High/Low

E	MD	D
12	8	4
10	6	2

E	MD	D
6	6	6
8	8	8

E	MD	D
2	6	10
4	8	12

Low/High

E	MD	D
12	8	4
10	6	2

E	MD	D
6	6	6
8	8	8

E	MD	D
2	6	10
4	8	12

Low/Low

E	MD	D
12	8	4
10	6	2

E	MD	D
6	6	6
8	8	8

E	MD	D
2	6	10
4	8	12

Figure 1. Layout of the Full Factorial Design for Experiment 2

Note. Within each grid E= easy items, MD= moderately difficult items, and D= difficult items. For goal types: High/High required recall of 9 words worth 80 points; High/Low = 9 words and 42 points; Low/High goal required 4 words and 44 points; Low/Low goal required 4 words worth 12 points.

Participants

Sixty younger (33 males and 27 females; 58% Caucasian, 30% Asian, 10% African American, and 2% of Mixed racial background) and 31 older adults (10 males and 21 females; 68% Caucasian, 32% African American with *M* = 14.65, *SD* = 3.49 years of education) participated in the second experiment. Younger (*M* age = 19.57 years, *SD* = 1.50) and older adult participants (*M* age = 68.74 years, *SD* = 4.78) were recruited, compensated, and prescreened for Spanish exposure as in Experiment 1. All were native English speakers with the exception of one younger adult who had been speaking English since age 2. Both younger and older adults rated their health as very good (*M* rating = 2.00, *SD* = 0.76 and *M* rating = 2.39, *SD* = 0.96, respectively).

Participants were tested in groups of up to 7 people, but younger and older adults were tested separately. Random assignment of the younger adults resulted in 31 participants in the experimental (30 second) and 29 in the control (60 second) condition. All older adults were given 60 seconds. These numbers were again deemed sufficient to detect a small effect (Faul et al., in press). Participants within these age/time conditions were further subdivided randomly into either a Point Order 1 or Point Order 2 condition. Participants randomly assigned to Point Order 1 first viewed a grid whose point values favored recall of easy items before seeing the neutral and final grid which favored recall of the most difficult items. Those randomly assigned to Point Order 2 instead first viewed a grid that contained points which favored recall of the most difficult items before seeing the neutral and final grid which favored recall of the easy items (see Figure 2). Thirty-three younger and 13 older adults received point order 1, whereas 27 younger and 18 older adults received point order 2. Although the order of points was not expected to

matter, the two orders were included to allow examination of possible order effects, albeit with very small numbers in each order.

Order 1

Favor Easy

E	MD	D
12	8	4
10	6	2

Neutral

E	MD	D
6	6	6
8	8	8

Favor Difficult

E	MD	D
2	6	10
4	8	12

Order 2

Favor Difficult

E	MD	D
2	6	10
4	8	12

Neutral

E	MD	D
6	6	6
8	8	8

Favor Easy

E	MD	D
12	8	4
10	6	2

Figure 2. Point Order 1 versus Point Order 2

Note. Within each grid E= easy items, MD= moderately difficult items, and D= difficult items; Order 1= the first of the three grids that participants saw for each of the four goals had points that favored recall of easy items, the second grid was neutral, and the final grid had points that favored recall of difficult items; Order 2 = the first of the three grids that participants saw for each of the four goals had points that favored recall of difficult items, the second grid was neutral, and the final grid had points that favored recall of the easy items.

Materials

The same 75 Spanish-English word pairs used in Experiment 1 were used in Experiment 2. Three pairs were used in the instruction screens, 72 pairs in the task itself (24 easy, 24 moderately difficult, 24 difficult). One third of the items within each

difficulty level were assigned point values that favored recall of the difficult items (see Figure 3), one third of the items received point values that favored easy items (see Figure 4), and the remaining third of the items were given point values that were neutral (see Figure 5). This allowed the creation of 12 different 2 X 3 grids, 4 for each type of point distribution, to be used during the encoding phase. Two items of each difficulty type were randomly selected to appear in each of 12 different 2 X 3 grids. Three grids were associated with each of the four experimenter-determined goals.

Easy Items	**Moderately Difficult Items**	**Difficult Items**
2	6	10
4	8	12

Figure 3. Grid with Point Values Favoring Selection and Recall of Difficult Items

Easy Items	**Moderately Difficult Items**	**Difficult Items**
12	8	4
10	6	2

Figure 4. Grid with Point Values Favoring Selection and Recall of Easy Items

Easy Items	Moderately Difficult Items	Difficult Items
6	6	6
8	8	8

Figure 5. Grid with Neutral Point Values Favoring Selection and Recall of All Items

The experimenter-determined goals each specified how many words the participant should strive to recall as well as how many points they should try to earn by recalling those words. The word goal (i.e., how many words they should try to recall) was designed to either require recall of a high number of words (i.e., try to recall 9 of the 18 vocabulary pairs) or a low number of words (i.e., try to recall 4 of the 18 vocabulary pairs). The point goal (i.e., how many points they should try to earn) was also either a high or a low point goal. Word recall goals (high and low) were crossed with the point goals (high and low) to obtain 4 distinct task goals (i.e., High word/ High point, High word/Low point, Low word/ High point, and Low word/Low point).

The *High word/High point* goal instructed participants to try to recall 9 words and obtain 80 points. The high (word and point) goals were designed to try and force participants to study items of varying difficulty levels from multiple grids rather than just one grid in order to achieve the goal. That is, the high word goal made it more likely that participants would have to study items from all three grids associated with the goal, whereas the lower word goal could be achieved by studying items from only one or two grids. The point goal was determined by summing the highest point values one could obtain if one were to select the three items within each grid that contained the highest

point values. Thus across the three 2 X 3 grids one could earn 80 points by selecting and later recalling 3 easy, 3 moderately difficult, and 3 difficult items (i.e., the 22 points from the easy column and 8 points from the 1 moderately difficult item in the "favor easy" grid, the three items worth 8 points each in the "neutral" grid, and the 8 point moderately difficult item and 22 points from the difficult items in the "favor difficult" grid).

A similar method was used to determine the low point goal for the *High word/Low point* goal condition. The high word goal still required 9 words to be recalled, but the lower point goal meant participants only had to achieve 42 points (i.e., the sum of the 9 words with the lowest point values across the three 2 X 3 grids). The *Low word/High point* goal required 4 words and 44 points (i.e., the sum of the 4 highest valued items across the three grids), whereas the *Low word/ Low point* goal required 4 words and 12 points (i.e., the sum of the 4 lowest valued items across the three grids). As with the word goals, the point goals were designed to inspire participants to select items of different difficulty levels to achieve each goal. Therefore, although each point goal could be pursued in a variety of ways (e.g., recalling a single easy 12-point item in the low point goal condition), participants were somewhat constrained by the corresponding word goal such that to achieve both parts of the goal, they had to take into account both how many words they should study/recall, but also the point values associated with those items, thus requiring even greater metacognitive control.

Unlike Experiment 1, all phases of the criterion task were computerized. Instructions and stimuli were presented on the computer screen and responses were input using the mouse and keyboard. The computer program, developed in C# (Microsoft Visual Studio, Version, Microsoft Corporation, 2005), recorded participants' responses to

all queries, the order in which items were (re)selected, how much time was allocated to each item, and with which of the four experimenter-provided goals the items were associated. Participants prestudy and poststudy recall test responses were initially scored by the computer and then hand scored as described in Experiment 1.

The same background measures used in Experiment 1 were used in the second experiment to form hypotheses regarding potential explanations for any age-related differences observed in item selection behaviors. Limited sample size again prevented examination of individual differences.

Procedure

As in the first experiment, participants completed the memory questionnaires (i.e., PBMI and MCI), and cognitive ability measures (Listening span, AVT, and Letter Comparison tasks) before beginning the criterion task. The EOL collection and prestudy recall test phases occurred as in Experiment 1. Participants then read instruction screens that told them they would have the opportunity to select items for study. They were shown brief strategy descriptions (i.e., rote repetition, imagery, keyword method, and sentence generation), as in Experiment 1, and were then told they would be shown 12 different grids containing items with different assigned point values and that they should try to earn enough points, by correctly recalling items on the recall test, to achieve each of four experimenter-determined goals. Participants were instructed how to select items for study and shown a sample study grid so they would know what to expect.

Participants were presented with one of the four goals, in a random order, one at a time. The order of goals was randomly determined for each participant, but counterbalanced across all participants.[2] Once participants indicated they had read and

understood the goal, they were asked to reiterate what the point portion of the goal was by typing the correct number before being allowed to proceed. They had to also correctly indicate what the word portion of the goal was to progress. These goal check questions were designed to ensure participants attended to both aspects of each goal. If participants were unable to correctly answer either portion of the goal check questions, the goal description reappeared before participants were asked to again answer the goal check questions. After three incorrect responses, the program required the participant to get the experimenter, who then worked with the participant to ensure s/he understood the goal before moving on. Participants then provided separate ratings to indicate how confident they were they would be able to achieve the word portion and the point portion of the goal. These ratings were collected using a 0 to 100 scale in which 0 indicated no confidence and 100 reflected 100% confidence the particular goal would be achieved.

Once participants passed the goal check questions and provided their confidence ratings for a particular goal, they began the planning phase in which they were shown 3 grids containing points in the order they would appear in during the study phase, but without any vocabulary words. The planning screens were designed to both familiarize participants with how the points would be distributed as well as to allow them time to strategically plan which items they would need to select in order to achieve the goal. They were informed the point values would change across the grids associated with each goal, but that the location of easy (left column), moderately difficult (middle), and difficult (right column) items within each grid would remain constant. The fact that some points favored recall of easy items, some favored recall of difficult items, and that some were neutral with respect to item difficulty level, or that some individuals received the

points in opposite order, was never mentioned to participants. Younger adults were given 60 seconds in the planning phase and older adults 120 seconds. Participants were not required to use all the allotted planning time.

After the allotted planning time elapsed, or participants indicated they were finished planning, the first of the three grids associated with the goal appeared. The experimenter-determined goal appeared at the top of the study screen so as to alleviate any concern that individuals might fail to achieve the goal simply because they forgot it. Only the Spanish portion of the Spanish-English vocabulary pairs appeared in the grids, as well as the point value they could earn by correctly recalling each item (see Figure 6).

aeropuerto - 12 ?	sangre - 8 ?	alcantarilla - 4 ?
bufalo - 10 ?	alimento - 6 ?	requebrajadura - 2 ?

Figure 6. Sample Layout of Grids in Experiment 2

Participants selected items by clicking on the item, which caused a new screen to appear that contained the intact vocabulary pair and a countdown clock. The clock immediately began to count down from either 30 or 60 seconds (depending on age/ time condition) until time either elapsed or the participant indicated they were finished studying the item

(e.g., by clicking the “finished studying” button). A strategy report screen appeared after each item was studied that queried participants as to which strategy, if any, they used to study the item. Once they provided their strategy report and clicked “enter” the 2X3 grid reappeared. A two-color (green = unstudied; red = studied) scheme was used within the 2 X 3 grid so that individuals could easily tell which items had already been studied and which items remained. Participants were aware they could reselect previously studied (i.e., red) items as long as time remained. The clock resumed counting down once the next item was selected. After the age- or condition- specific time elapsed on the first grid, the second grid appeared, with the task goal again stated at the top. The process repeated until time elapsed on the 3 grids associated with the first goal. A “finished studying this grid” button was included at the bottom of each grid so that participants were not forced to spend more time than they wished on any grid.[3]

Once participants finished studying all three grids associated with the first goal, instruction screens appeared that detailed what the second goal was. Participants were given age-appropriate planning time and then answered the goal check questions for that goal before selecting items for study from each of the three 2 X 3 grids until time ran out or they indicated they were done. The process was repeated for the third and fourth goals.

After time elapsed or was terminated on all twelve grids, participants provided delayed JOLs and then completed the poststudy recall test. They then postdicted their recall performance, both globally and for each item type. They were then asked to indicate what their global and differentiated goals would be if they were given an opportunity to study the same 72 items again (i.e., as was done in Experiment 1). After

participants indicated what they hoped to achieve in a second trial, they were debriefed and dismissed.

Results

Cognitive Ability Measures

No time-related or point order-related differences were found in any of the ability measures thus removing concern that differences in cognitive ability would cloud interpretation of any observed time- or point-related effects.

Vocabulary

Older adults scored slightly (M = 18.71, SE = 1.46), but not reliably higher than younger adults (M = 17.47, SE = .62) on the Advanced Vocabulary Test, $F < 1$. Thus English vocabulary ability should not have differentially impacted how well younger and older adults were able to learn the Spanish vocabulary.

Perceptual Speed

Younger adults (M = 25.12, SE = .58), scored significantly higher on the Letter Comparison task than older adults (M = 18.26, SE = .81), F (1, 89) = 47.86, $p < .001$, partial $\eta^2 = .350$. If perceptual speed impacted participants' ability to make selection and time allocation decisions, then older adults were at a disadvantage.

Working Memory Capacity

Younger adults (M = 65.87, SE = 1.06) had reliably higher span scores than older adults (M = 56.89, SE = 1.78) on the Listening Span Task, F (1, 89) = 107.69, $p < .001$, partial $\eta^2 = .548$. However, relative to experiment 1, the younger adults scored lower and the older adults substantially higher, producing less discrepancy in younger and older adults' scores than existed in Experiment 1. This could prove beneficial given the

arguably more difficult task in Experiment 2, and Dunlosky and Thiede's (2004) finding that increased WMC facilitated forming and implementing selection plans.

Memory Beliefs Measures

No time- or point order-related differences were found in participants' memory beliefs alleviating concern that differences would cloud interpretation of the results.

PBMI

As in Experiment 1, no age-related differences were found in participants' responses to any of the memory belief items that could be expected to influence their performance in the criterion task (e.g., Global MSE or control beliefs).[4]

MCI

Younger and older adults' MCI ratings were also very similar, $F < 1$.

Table 11. PBMI and MCI Memory Beliefs Ratings

	Younger- 30		Younger- 60		Older	
Judgment Type	*M*	*(SE)*	*M*	*(SE)*	*M*	*(SE)*
PBMI Global MSE	73.74	(3.42)	74.66	(3.95)	78.61	(2.93)
PBMI Control	81.37	(2.09)	84.10	(1.94)	84.79	(2.75)
MCI Present Ability	5.48	(.25)	5.53	(.14)	5.35	(.22)
MCI Potential Improve	5.10	(.21)	5.28	(.14)	5.24	(.23)

Note. Entries are means and standard errors for participants' memory beliefs ratings; Younger-30 = younger adults in the 30 second experimental condition; Younger-60 = younger adults in the 60 second control condition; PBMI Global MSE = overall memory self-efficacy; MCI Present Ability = ratings of how good their present memory is; MCI Potential Improve = participants' perceptions of how much they can potentially do things to improve their memory.

Ease of Learning Judgments

As in Experiment 1, participants' EOLs aligned with the Kornell and Metcalfe (2006) normative difficulty categories (see Table 12) and older adults viewed items of all types as more difficult than younger adults, F (1, 89) = 35.06, $p < .001$, partial $\eta^2 = .283$. Because participants' ratings increased with item difficulty, a reliable main effect of difficulty was found, F (2, 88) = 207.01, $p < .001$, partial $\eta^2 = .825$, which combined with the age-related differences in ratings to produce a reliable Age X Difficulty interaction, F (2, 88) = 3.64, $p < .05$, partial $\eta^2 = .076$.

Table 12. Mean EOLs, DJOLs, Postdictions, and Post-Recall Goals

	Younger- 30		Younger- 60		Older	
Judgment	*M*	*(SE)*	*M*	*(SE)*	*M*	*(SE)*
Global EOL	4.52	(.14)	4.58	(.18)	5.76	(.18)
EOL Easy	2.23	(.16)	2.61	(.30)	3.86	(.30)
EOL Medium	4.44	(.18)	4.49	(.22)	5.83	(.21)
EOL Hard	6.91	(.17)	6.65	(.21)	7.58	(.20)
Global DJOL	60.67	(2.26)	58.20	(2.32)	43.09	(3.92)
DJOL Easy	95.20	(1.43)	95.75	(1.13)	74.13	(5.43)
DJOL Medium	60.00	(3.71)	53.78	(3.94)	41.27	(5.06)
DJOL Hard	26.82	(3.23)	25.06	(3.46)	13.88	(3.38)
Global Postdiction	31.42	(2.13)	28.31	(2.31)	26.65	(3.10)
Postdiction Easy	18.84	(.91)	18.28	(.96)	13.77	(1.32)
Postdiction Medium	10.97	(.99)	9.14	(1.24)	8.52	(1.06)

Table 12 (continued).

Postdiction Hard	6.39	(1.04)	5.83	(.99)	3.97	(.63)
Global Goal 2	51.45	(3.07)	52.66	(3.30)	37.61	(4.08)
Goal 2 Easy	23.32	(.35)	21.62	(1.19)	18.00	(1.46)
Goal 2 Medium	17.42	(1.12)	18.00	(1.02)	13.52	(1.45)
Goal 2 Hard	11.90	(1.30)	13.76	(1.36)	8.65	(1.27)

Note. Entries are means (and standard errors) of individuals' metamemory judgments; Younger-30 = younger adults in the 30 second experimental condition; Younger-60= younger adults in the 60 second control condition; EOL= ease of learning judgment provided using scale of 1-easy to 9-hard; DJOL = delayed judgment of learning in which 0 was no confidence and 100 was complete confidence in ability to correctly recall each Spanish word's English counterpart; Postdiction = judgments reflecting how many items participants think they got correct on the recall test; Goal 2 = post-recall goal for hypothetical trial 2; Easy = normatively easy items; Medium = normatively moderately difficult items; Hard = normatively difficult items.

Goals

Because participants in Experiment 2 were provided with experimenter-determined goals rather than participants setting their own (as participants did in Experiment 1), the focus of analysis was different and examined participants' goal pursuit and goal attainment behaviors. That is, whether individuals selected enough items with high enough point values to allow them to achieve the goal if those items were later correctly recalled. However, because individuals could select enough items but then later have problems recalling enough items to actually achieve the goals, these issues (i.e., pursuit and attainment) were considered separately.

Goal Pursuit

To examine whether individuals pursued each of the goals and if there were differences in pursuit as a function of low and high word and point demands, the number

of items the participant selected for study was summed for each grid, as were the points associated with those items. The experimenter-determined word and point goal values were then subtracted from the sum of those that each participant selected. For example, if a participant selected 7 words to study for the Low/Low goal, which only required 4, then the participant would have attempted (7 [selected] – 4 [required]) +3 items. Thus negative discrepancies indicated the participant either did not select enough words to meet the word goal or did not select items with high enough point values to achieve the point portion of the goal. Values of zero or higher indicated the participant either selected exactly the right amount or more, respectively. Those with zero or positive discrepancy were considered to have pursued the goal. Those with negative values were considered to have not attempted the goal.

Goal Achievement

Goal achievement was examined in the same way as goal pursuit by subtracting the number of words/points participants had been told to try and achieve from how many they actually recalled/earned. Thus individuals who successfully recalled as many or more items as required by the goal had positive discrepancies and were considered to have achieved the word portion of the goal. Similarly, those who recalled enough items with high enough values to have positive discrepancies, relative to the number of points the goal dictated, were considered to have achieved the point portion of the goal.

As expected, participants in both age groups and time conditions were more likely to pursue and achieve the easier portions of the word and point goals than the harder, but younger adults were more likely than older adults to pursue and achieve the hard portions of the goals. As may be seen in Table 13, older adults' tendency to pursue and achieve

the low point goal was high, but dropped when paired with the high word goal. In contrast, younger adults' pursuit and attainment of the low point goal was largely unaffected by whether it was paired with the low or high word goal. Both younger and older adults were less likely to pursue and achieve the high point goal than the low point goal, but especially when it was paired with a high word goal. In fact, the frequency of high point goal achievement dropped 54% in older adults and 15% in younger adults when paired with a high word as opposed to low word goal. That high word goals served to reduce pursuit of low and high point goals for older adults and high point goals in younger adults is interesting because, if anything, the higher word goal should have increased the number of items selected which should have in turn increased point pursuit and achievement.

Similar patterns were found in younger and older adults' word pursuit and achievement behaviors. Pursuit of low word goals was similar regardless of points, but achievement dropped slightly when paired with a high point goal. Both younger and older adults were more likely to pursue the high word goal when it was paired with a high point goal, however, the high/high goal pairing hurt older adults', but helped younger adults' high word goal achievement.

Interesting differences emerged in the frequency of younger adults' goal pursuit and achievement behaviors as a function of how much study time they had. Although no time-related differences were found for low word/ low point pursuit and achievement, high word/low point goal pursuit was reduced for those given less time, and yet younger adults in both time conditions achieved the goal. Younger adults in both time conditions were more likely to pursue the high point goal when it was paired with the high word as

opposed to the low word goal. However, 20% more in the 30 second condition pursued the high/high goal than those in the 60 second condition, which resulted in 20% more of those given less time achieving the goal than those given more time.

When word goals were examined, younger adults were found to have similar goal pursuit regardless of time, but those with less time achieved the high word/ low point goal more often than those in the 60 second condition. In contrast, those given less time were less likely than 60 second younger adults to pursue either of the low word goals. This did not affect the likelihood of achieving the low word goal when it was paired with the low point goal, but did result in 30 second younger adults being less likely to achieve the low word goal when it was paired with the high point goal.

Together, these data indicated that if the goals stressed points (i.e., high point conditions), then increasing the word goal from low to high decreased the likelihood that younger and older adults would achieve the point portion of the goal. In contrast, if the goals stressed words (i.e., high word conditions), then increasing the point goal served to decrease the likelihood that the word goal would be achieved, more so for older than younger adults. This occurred despite the fact that individuals in both age groups (and time conditions) were more likely to pursue the higher word goal when it was paired with the higher point goal. Thus individuals seemed to focus on whichever aspect of the goal was most demanding. This held true when younger and older adults were compared as well as when younger adults in the two time conditions were compared.

Table 13. Frequency of Pursuit and Achievement of Word Goals and Point Goals

	Younger- 30		Younger- 60		Older	
Goal	*% No* (N)	*% Yes* (N)	*% No* (N)	*% Yes* (N)	*% No* (N)	*% Yes* (N)
Low Word/ Low Point Goal						
Point Goal Pursued	0	100 (31)	0	100 (29)	3 (1)	97 (30)
Point Goal Achieved	0	100 (31)	0	100 (29)	0	100 (31)
Word Goal Pursued	10 (3)	90 (28)	3 (1)	97 (28)	3 (1)	97 (30)
Word Goal Achieved	0	100 (31)	0	100 (29)	10 (3)	90 (28)
High Word/ Low Point Goal						
Point Goal Pursued	7 (2)	93 (29)	3 (1)	97 (28)	7 (2)	93 (29)
Point Goal Achieved	0	100 (31)	0	100 (29)	29 (9)	71 (22)
Word Goal Pursued	13 (4)	87 (27)	14 (4)	86 (25)	16 (5)	84 (26)
Word Goal Achieved	26 (8)	74 (23)	41 (12)	59 (17)	77 (24)	23 (7)
Low Word/ High Point Goal						
Point Goal Pursued	7 (2)	93 (29)	0	100 (29)	10 (3)	90 (28)
Point Goal Achieved	16 (5)	84 (26)	7 (2)	93 (27)	39 (12)	61 (19)

Table 13 (continued).

Word Goal Pursued	7 (2)	93 (29)	0	100 (29)	7 (2)	93 (29)
Word Goal Achieved	7 (2)	93 (29)	0	100 (29)	10 (3)	90 (28)
High Word/ High Point Goal						
Point Goal Pursued	13 (4)	87 (27)	24 (7)	76 (22)	19 (6)	81 (25)
Point Goal Achieved	52 (16)	48 (15)	72 (21)	28 (8)	93 (29)	7 (2)
Word Goal Pursued	0	100 (31)	17 (5)	83 (24)	10 (3)	90 (28)
Word Goal Achieved	23 (7)	77 (24)	38 (11)	62 (18)	71 (22)	29 (9)

Note. Entries are the percentage of participants in each age group/time condition that did or did not pursue and/or achieve the word and point goals associated with each of the four experimenter determined goals. The low/low goal required 4 words and 12 points; the high word/low point goal required 9 words and 42 points; the Low word/ High point goal required 4 words and 44 points; the High word/High Point goal required 9 words and 80 points in order to achieve it.

Participants' goal pursuit and goal achievement behaviors were further examined by calculating the mean discrepancy between what they were told to achieve and what they actually recalled and earned in points. As may be seen in Tables 13 and 14, the trend across all four goals was for older adults to attempt more but achieve less than younger adults. Thus older adults on average studied more items worth more points, but recalled fewer items and therefore earned fewer points than younger adults. The age-related differences found in attempted words, F (1, 89) = 4.64, $p < .05$, partial $\eta^2 = .050$, and attempted points, F (1, 89) = 4.35, $p < .05$, partial $\eta^2 = .047$, were only reliable for the low word/ low point goal. However, younger adults recalled more words, F (1, 89) = 15.27, $p < .001$, partial $\eta^2 = .146$, and achieved more points than older adults, F (1, 89) = 15.52, $p < .001$, partial $\eta^2 = .149$, despite both achieving more than was necessary to achieve the low/low goal. Younger adults also recalled more words, F (1, 89) = 29.97, $p < .001$, partial $\eta^2 = .252$, and earned more points than older adults for the high word /low point goal, F (1, 89) = 30.84, $p < .001$, partial $\eta^2 = .257$. A similar pattern was found for the low word/ high point goal, with reliable differences in words F (1, 89) = 13.42, $p < .001$, partial $\eta^2 = .131$, and points, F (1, 89) = 14.34, $p < .001$, partial $\eta^2 = .139$, as well as the high word/high point goal, F (1, 89) = 31.64, $p < .001$, partial $\eta^2 = .015$, and F (1, 89) = 31.20, $p < .001$, partial $\eta^2 = .005$, respectively.

Effects of point order emerged because those given Order 1 recalled more words, F (1, 89) = 9.01, $p < .01$, partial $\eta^2 = .092$, and earned more points, F (1, 89) = 9.94, $p < .01$, partial $\eta^2 = .100$, on both the low word/ low point goal, as well as the high word/ low point goal, F (1, 89) = 6.98, $p < .01$, partial $\eta^2 = .073$, and F (1, 89) = 7.32, $p < .01$, partial $\eta^2 = .076$, respectively. Because Order 1 initially rewarded recall of easy items,

whereas Order 2 initially rewarded recall of difficult items, participants fared better for low point goals when given Order 1 because they could earn enough points without being forced to study difficult items. Order effects did not exist for the high point goals or the (low or high) word goals because they required participants to study items across multiple levels of difficulty, often across multiple grids, thus removing the focus on studying only a few easy items.

No time-related differences were found when younger adults' mean word and point totals were compared with the experimenter-determined word and point goals.

Table 14. Mean Discrepancy in Specified Goals and Goal Achievement

	Younger- 30	Younger- 60	Older
Goal	*M* (*SE*)	*M* (*SE*)	*M* (*SE*)
Low Word/ Low Point Goal			
Attempted Points	57.03 (7.46)	69.31 (7.04)	81.16 (6.89)
Achieved Points	48.26 (3.57)	52.14 (3.10)	33.74 (3.53)
Studied Words	6.00 (0.98)	7.56 (0.94)	9.39 (0.96)
Words Correct	4.65 (0.48)	5.10 (0.41)	2.65 (0.50)
High Word/ Low Point Goal			
Attempted Points	52.06 (5.03)	45.79 (5.33)	57.81 (6.37)
Achieved Points	29.61 (2.88)	24.62 (3.25)	5.16 (3.59)
Studied Words	4.23 (0.67)	3.55 (0.73)	5.26 (0.89)
Words Correct	1.06 (0.37)	0.48 (0.46)	-2.16 (0.48)

Table 14 (continued).

Low Word/ High Point Goal			
Attempted Points	39.74 (5.72)	46.34 (5.70)	52.58 (6.61)
Achieved Points	25.10 (4.37)	18.69 (2.33)	5.35 (3.65)
Studied Words	7.00 (0.94)	8.31 (0.94)	9.39 (0.99)
Words Correct	5.55 (0.64)	4.76 (0.36)	2.84 (0.52)
High Word/ High Point Goal			
Attempted Points	21.29 (3.37)	16.55 (4.64)	22.45 (4.86)
Achieved Points	-4.13 (3.50)	-11.24 (3.06)	-32.06 (4.05)
Studied Words	4.87 (0.58)	3.90 (0.75)	5.39 (0.73)
Words Correct	1.65 (0.50)	0.66 (0.42)	-2.26 (0.56)

Note. Entries are the mean difference between what the goal asked participants to achieve and what they actually achieved thus positive numbers indicate the goal was achieved, negative that the goal was not; Younger-30 = younger adults in the 30 second experimental condition; Younger-60 = younger adults in the 60 second control condition; The low/low goal required 4 words and 12 points; the high word/low point goal required 9 words and 42 points; the Low word/ High point goal required 4 words and 44 points; the High word/High Point goal required 9 words and 80 points in order to achieve it.

Goal achievement confidence ratings

Because participants were provided with goals rather than asked to specify their own, participants were asked to indicate how confident they were that they would be able to achieve the word and point portions of each goal. Age-related differences were found in participants' confidence ratings for all four goals because older adults' ratings were reliably lower than younger adults', F (2, 88) = 8.46, $p < .001$, partial $\eta^2 = .161$ (see

Figure 7). Thus, older participants expected to be less likely to achieve the four goals and in fact were.

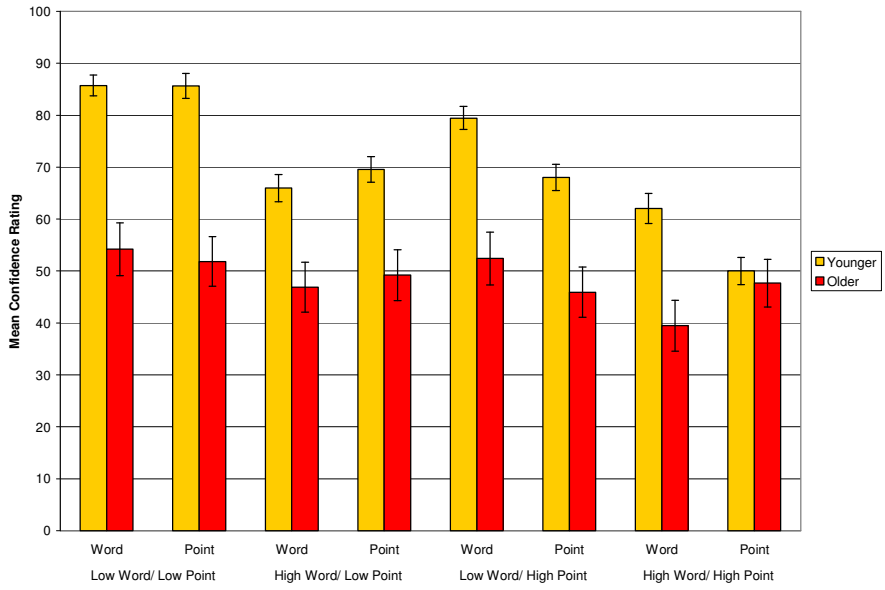

Figure 7. Word and Point Confidence Ratings as a Function of Goal

Item Selection

Selection Order

As in Experiment 1, each item was assigned a value of 1 (easy), 2 (moderately difficult), or 3 (difficult) based on normative difficulty level (Kornell & Metcalfe, 2006) and these were averaged for participants' first six selections, collapsing across the 12 grids.[5] The analyses examined whether the mean level of difficulty for the first six selections was more consistent with RPL model or DRM predictions. Consistent with the results in Experiment 1, both younger and older adults selected easier items first and only later transitioned to more difficult items (see Table 15). However, the mean level of difficulty associated with the first five selections was lower for older than younger adults.

In contrast, by the sixth selection, older adults selected items of greater difficulty than younger adults did. These age-related differences in the difficulty of selections were only reliably different for the second selection, F (1, 89) = 5.48, $p < .05$, partial $\eta^2 = .058$.

When selection order was reexamined after excluding items that participants had gotten correct on the prestudy recall test, the pattern of results was the same, however the mean difficulty level of each selection was slightly higher. A reliable age-related difference remained for the second selection, F (1, 85) = 7.39, $p < .01$, partial $\eta^2 = .080$, and the third selection was also found to differ after controlling for prior knowledge, F (1, 87) = 3.95, p =.05, partial $\eta^2 = .043$.

Mean selection level was further examined as a function of which point order participants had been given (i.e., Order 1: easy rewarded first versus Order 2: difficult rewarded first). No order-related differences were found in the difficulty level of participants' selections when prior Spanish knowledge (i.e., prestudy recall performance) was ignored. However, when items were discarded that were answered correctly on the prestudy recall test and the data were reexamined, the first selection was found to differ as a function of point order. Specifically, those that had received Order 2 (M = 2.08, SE = .11) selected items significantly more difficult than those given Order 1 (M = 1.79, SE = .09), F (1, 79) = 4.79, $p < .05$, partial $\eta^2 = .057$. This suggests that Order 2 participants were attending to the point values when deciding which items to select, and therefore selected the more difficult items earlier than those in Order 1 who could select easy items to achieve the same points.

Younger adults' selection data were examined for time-related differences in selection order, but no differences were found. This held true when prior knowledge was

accounted for as well. Yet order effects were found in younger adults' data that mirrored those in the overall data. Specifically, those given Order 1 initially selected items of lower difficulty (i.e., those with higher point values), whereas those given Order 2 initially selected more difficult items (i.e., those with greater point value for them). Across the next four selections (i.e., selections 2-5) those in Order 1 selected increasingly more difficult items, whereas those in Order 2 did not transition to more difficult items at the same rate, despite having initially selected items that were more difficult. On the fifth selection the difference in mean difficulty across the two orders was reliable, F (1, 48) = 5.71, $p < .05$, partial $\eta^2 = .106$. As in the overall data, when prior Spanish knowledge was accounted for, the difficulty selected by those given Order 2 for their first selection was reliably higher than those given Order 1, F (1, 53) = 6.65, $p < .05$, partial $\eta^2 = .111$.

Table 15. Mean Difficulty Level of First Six Selections by Point Order

	Younger- 30		Younger- 60		Older	
Selection	*M*	*(SE)*	*M*	*(SE)*	*M*	*(SE)*
			Point Order 1			
1	1.57	(.12)	1.63	(.16)	1.49	(.17)
2	1.71	(.12)	2.01	(.18)	1.57	(.13)
3	2.26	(.10)	2.54	(.11)	2.06	(.10)
4	2.39	(.10)	2.49	(.15)	2.30	(.15)
5	2.74	(.09)	2.63	(.14)	2.61	(.15)
6	2.49	(.17)	2.51	(.24)	2.85	(.15)

Table 15 (continued).

			Point Order 2			
1	2.06	(.24)	1.53	(.15)	1.48	(.13)
2	2.02	(.17)	1.80	(.16)	1.55	(.13)
3	2.38	(.13)	2.00	(.11)	2.12	(.13)
4	2.35	(.18)	2.23	(.12)	2.24	(.09)
5	2.16	(.28)	2.43	(.19)	2.55	(.13)
6	2.30	(.25)	2.75	(.11)	2.58	(.17)

Note. Entries are the average level of difficulty of items selected in each of the first selections, collapsing across grids; Younger-30= younger adults in the 30 second experimental condition; Younger-90= younger adults in the 90 second control condition; Normatively easy items were assigned values of 1, moderately difficult values of 2, and difficult values of 3 prior to averaging; Point order 1 first favored recall of easy items; Point order 2 first favored recall of more difficult items.

The prior analyses examined selection order means for participants' first six selections, collapsing across the 12 grids. However, collapsing across all grids could potentially mask the impact of points and point order on participants' selection decisions. Thus additional selection order means were calculated for participants' first six selections collapsing across 1) grids with points that favored easy items, 2) neutral grids, and 3) grids with points that favored moderately difficult items. Consistent with the prior results and RPL model predictions, older and younger adults selected easier items first regardless of point order or time constraints. Yet as may be seen in Figures 8 through 10, when grids had points that favored recall of easy items, those given Point Order 1 selected easier items initially but transitioned to selecting items that were on average more difficult than those given Point Order 2. In contrast, for grids with points favoring

recall of difficult items, Point Order 1 participants initially selected easier items than those given Point Order 2. Neutral grids were more likely to show the RPL-consistent easy to difficult study pattern for those given Order 1 than Order 2. Thus grids with points favoring easy items and grids favoring difficult items were likely to cross for the two point orders as a result of individuals in all age groups adjusting their selections based on what the points were rewarding. The fact that younger and older adults' selections varied as a function of point order suggests individuals in both age groups and time conditions were attending to point values when deciding in which order they would select items.

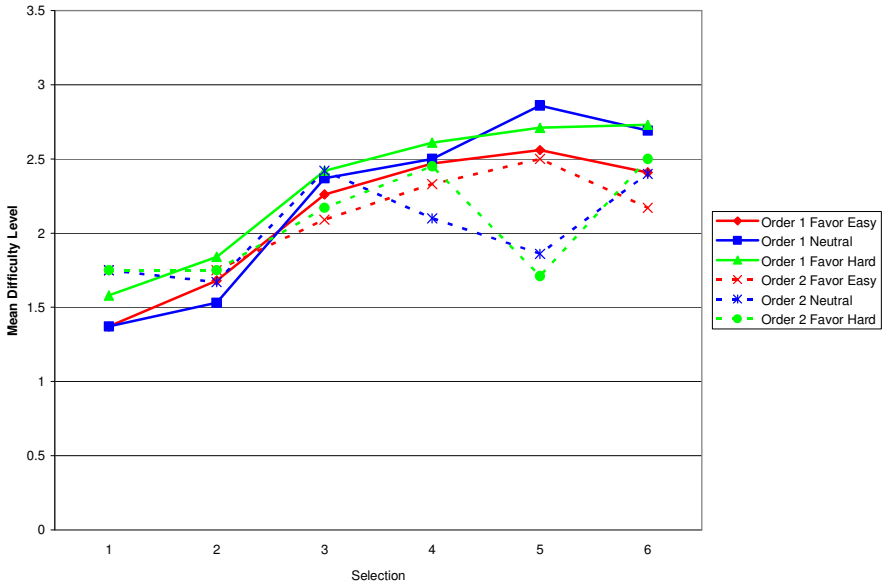

Figure 8. Experimental Condition Younger Adults' Selections for Point Orders 1 and 2 as a Function of Grid Type

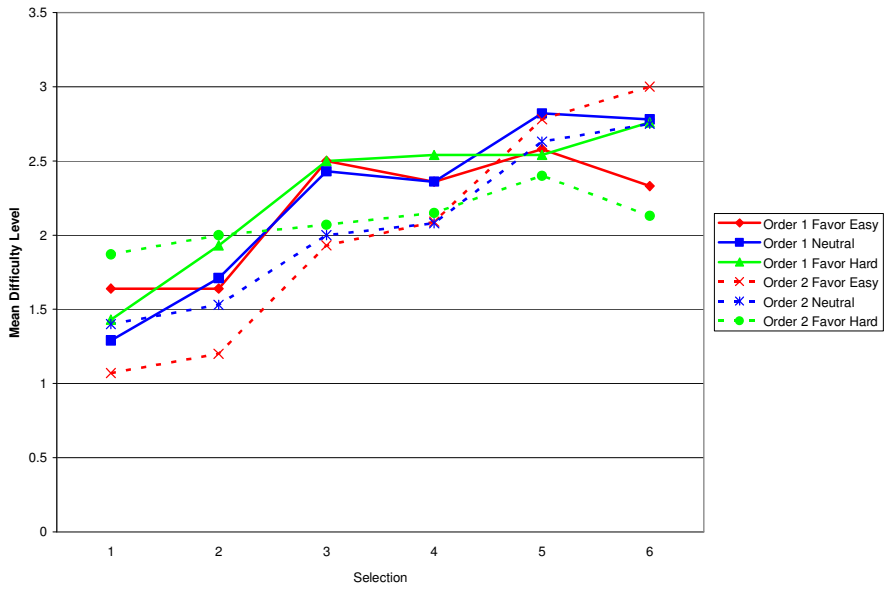

Figure 9. Control Condition Younger Adults' Selections for Point Orders 1 and 2 as a Function of Grid Type

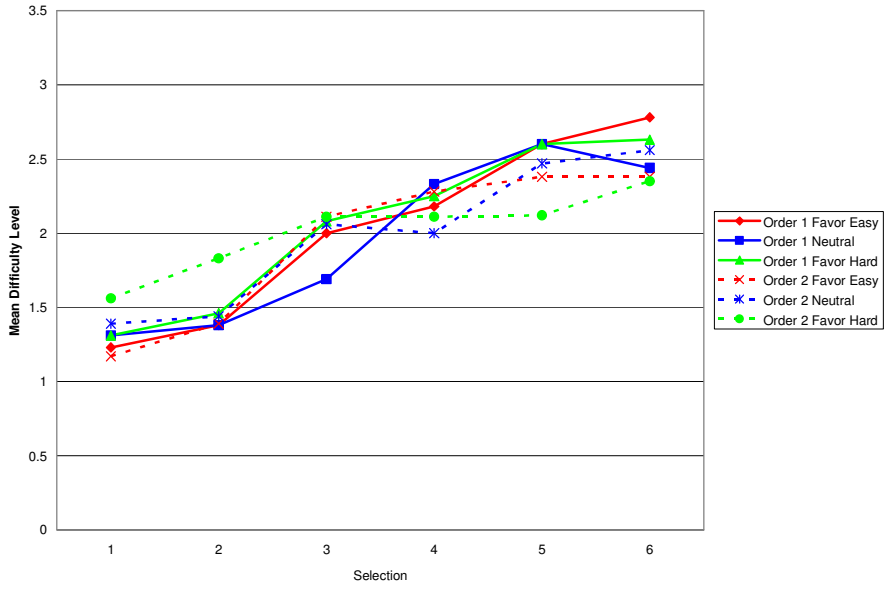

Figure 10. Older Adults' Selections for Point Orders 1 and 2 as a Function of Grid Type

Item Selection and EOLs

Items selected for study and ignored were examined for potential differences in perception of difficulty, as was done in Experiment 1. Collapsing across age group and difficulty level, it was found that participants selected items they perceived as easier (*M* of items selected = 4.91, *SE* = .04), and chose to ignore more difficult items (*M* of items ignored = 5.08, *SE* = .07). This also held true when the selection/EOL relationship was examined separately for each age group (see Table 16). Items in the normatively easy category (Kornell & Metcalfe, 2006) were more likely to be selected by both younger and older adults if they had been given higher rather than lower EOLs, but there was little differentiation between the mean EOL for selected and ignored items. In contrast, the items normatively classified as moderately difficult and difficult were more likely to be selected for study if they had been given lower, rather than higher, EOLs. These selection patterns contrasted those in Experiment 1 in which individuals in both age groups were more likely to choose items they had rated as more difficult for normatively easy and moderately difficult items and only ignored harder items in the difficult category. Thus overall, those in Experiment 1 chose to ignore easy items and selected items viewed as more difficult, whereas those in Experiment 2 did the opposite.

When EOL/selection behaviors were examined as a function of time constraints, younger adults in both time conditions were found to select items of a similar difficulty level, but those given more time ignored items they deemed more difficult whereas those given less time ignored items they viewed as easier.

Table 16. Mean EOLs for Non-selected and Selected Items

EOL Type	Non-Selected *M*	Non-Selected *(SE)*	Selected *M*	Selected *(SE)*
		Younger Adults – 30 Seconds		
Overall	4.48	(.10)	4.55	(.07)
Easy Items	2.13	(.09)	2.30	(.08)
Moderate Items	4.71	(.14)	4.34	(.09)
Difficult Items	7.25	(.10)	6.74	(.07)
		Younger Adults – 60 Seconds		
Overall	4.70	(.11)	4.53	(.07)
Easy Items	2.31	(.14)	2.77	(.10)
Moderate Items	4.75	(.16)	4.40	(.09)
Difficult Items	7.16	(.10)	6.39	(.08)
		Older Adults		
Overall	6.42	(.12)	5.56	(.06)
Easy Items	3.86	(.23)	3.87	(.10)
Moderate Items	6.55	(.19)	5.67	(.10)
Difficult Items	7.87	(.11)	7.44	(.08)

Note. Values represent mean EOL ratings for items that were or were not selected as a function of difficulty level. EOLs were collected on a 1 to 9 scale for which 1 was easy and 9 was difficult. Easy = normatively easy items; Moderate = moderately difficult; Difficult = normatively difficult items.

Perhaps the most interesting finding regarding the EOL/selection behavior relationship was that goals interacted with EOLs to influence younger, but not older adults' item selection behaviors. That is, an interesting interaction existed in younger adults' EOLs for selected versus non-selected items as a function of which goal was being pursued. For goals with lower point demands (i.e., the low word/ low point and high word/ low point goals) younger adults tended to ignore the harder items (i.e., items to which they had given higher EOLs) and to instead select the easier. In contrast, for goals with higher point demands (i.e., the low word/ high point and high word/ high point goals), younger adults tended to ignore the easier items and instead selected harder items for study (see Table 17). These data provided direct evidence that younger adults were basing their item selection decisions on the point values and goals, as much if not more than the subjective difficulty of items. Yet older adults' selections seemed to be based more on a desire to select easier and ignore more difficult items, as reflected by the lower EOL means for selected and higher EOL means for ignored items across all four goals.

Table 17. Mean EOLs for Non-selected and Selected Items as Function of Goal Demands

	Non-Selected		Selected	
Goal (Word/Point)	*M*	*(SE)*	*M*	*(SE)*
		Younger Adults		
Low/Low (4/12)	4.82	(.14)	4.41	(.10)
High/Low (9/42)	4.57	(.16)	4.53	(.09)
Low/High (4/44)	4.47	(.14)	4.59	(.10)
High/High (9/40)	4.39	(.17)	4.60	(.09)

Table 17 (continued).

	Older Adults			
Low/Low (4/12)	6.22	(.23)	5.66	(.13)
High/Low (9/42)	6.55	(.25)	5.41	(.12)
Low/High (4/44)	6.43	(.23)	5.59	(.14)
High/High (9/80)	6.54	(.25)	5.57	(.13)

Note. Values represent mean EOL ratings for items that were or were not selected the three grids associated with each of the four experimenter-determined goals. Abbreviations: Low/Low = the low word/ low point goal which required recall of 4 words worth at least 12 points; High/Low = high word/low point goal which required recall of at least 9 words worth 42 or more points; Low/High = the low word/high point goal which required 4 words and 44 points; High/High = the high word/high point goal which required 9 words totaling at least 80 points; EOLs were collected on a 1 to 9 scale for which 1 was easy and 9 was difficult.

Item Selection and Subjective Versus Objective Difficulty

The gammas calculated between each participant's selection order and either item-level EOLs, EOL bins, or normative difficulty provided additional evidence that younger adults' item selection behaviors were more goal-driven than older adults' (see Table 18). Analyses that ignored study time and point order revealed similar gamma values for younger and older adults, with normative difficulty correlating more strongly with selection order than subjective difficulty for individuals in both age groups (as in Experiment 1). However, time-related differences were found when gammas were examined separately as a function of how much time younger adults had been given. Those given less time were likely to have gammas that were similar for subjective and objective (normative) difficulty. In contrast, younger adults given 60 seconds to study items had substantially lower gammas for subjective than for objective difficulty. This

suggests that those given less time utilized both their own impressions of item difficulty as well as the information they were provided about normative difficulty (i.e., which columns contained easy, moderately difficult, and difficult items) when selecting items, whereas 60 second younger adults relied more heavily on the normative information as a basis for item selection.

Of great interest was the impact that point order had on the relationship between subjective and objective difficulty and selection order. Younger adults who received Order 1 (i.e., points that favored recall of easy items first) and 30 seconds to study items still tended to weight subjective difficulty more than younger adults given 60 seconds, although gamma values indicated individuals in both time conditions weighted normative difficulty more than subjective difficulty. However, younger adults given Order 2 (i.e., points that favored recall of difficult items first) combined with 30 seconds to study items had non-existent gammas (e.g., .04 - .07) for both subjective and objective difficulty (see Table 18). Younger adults given longer study times and Order 2 had non-existent (negative) gammas for the subjective EOLs, but higher gammas when based on normative difficulty. Yet, the normative gamma was still substantially lower (.11) than had been found for Order 1 (.53) in both time conditions. This suggests that when individuals were given a point structure that coincided with perceptions of difficulty and normative difficulty that those with less time attended to both types of information whereas those given more time focused more on normative difficulty. Yet when point structure rewarded the selection of difficult rather than easy items, those with less time were more likely to ignore subjective and objective difficulty, whereas those with more time fell back on objective (normative) difficulty to select items, albeit to a lesser extent.

Table 18. Gamma Correlations

	Younger- 30		Younger- 60		Older	
Gamma Type	*M*	*(SE)*	*M*	*(SE)*	*M*	*(SE)*
EOLs/Prestudy Recall	.86	(.02)	.74	(.07)	.68	(.07)
EOLs/Poststudy Recall	.70	(.02)	.61	(.06)	.67	(.04)
DJOLs/Prestudy Recall	.96	(.02)	.93	(.02)	.84	(.07)
DJOLs/Poststudy Recall	.95	(.01)	.91	(.02)	.87	(.04)
Point Order 1						
Item-level EOLs/Order	.34	(.06)	.23	(.11)	.19	(.11)
Bin EOLs/ Order	.38	(.07)	.25	(.13)	.24	(.12)
Norm EOLs/Order	.53	(.08)	.53	(.12)	.29	(.16)
Order 1/ Order 2	.50	(.06)	.33	(.09)	.46	(.08)
Point Order 2						
Item-level EOLs/Order	.05	(.10)	-.01	(.08)	.14	(.07)
Bin EOLs/ Order	.07	(.13)	-.02	(.10)	.16	(.07)
Norm EOLs/Order	.04	(.12)	.12	(.14)	.28	(.09)
Order 1/ Order 2	.43	(.16)	.46	(.10)	.27	(.10)

Note. Entries are means (and standard errors) of Goodman-Kruskal gamma correlations. Item-level EOLs/Order = gamma calculated between selection order and participants' item level EOL judgments; Bin EOLs/ Order = gammas calculated between selection order and EOLS placed into easy, moderate and difficult bins; Norm EOLs/ Order = gamma calculated between selection order and normative item difficulty; Order1/Order 2= gamma calculated between the selection order for the first time items were selected for study and the second time. Point Order 1 = points that initially favored recall of easy items; Point Order 2 = points that initially favored recall of difficult items; Younger-30= younger adults in the 30 second experimental condition; Younger-90= younger adults in the 90 second control condition.

Of interest then was the finding that younger adults' gammas were severely impacted by point order, but older adults' gammas were not and only showed differences of .02 to .07 across the two point orders. Together these findings combine with the EOLs for selected versus ignored items to suggest that younger adults were paying more attention to the points and goals than older adults were. Younger adults shifted item selection based on the point goal demands whereas older adults did to a lesser extent. That older adults' gammas were largely unaffected by point order but younger adults went from having moderate to nonexistent gammas as a function of point order further highlights the discrepancy in what younger and older adults attended to when deciding which items to study.

Reselected items

As in Experiment 1, additional gammas were calculated between the first time individuals selected an item and additional selections of the same item to examine whether gamma values would increase or decrease. These gammas ranged from .27 to .29 for older adults and .33 to .53 for younger adults (see Table 18). Thus, moderate reselection order gammas were obtained despite the younger adults in Order 2 having non-existent gammas between (subjective and objective) difficulty and selection order. This indicates that although individuals given Order 2 were more likely to select items based on points than difficulty level, that once they chose an order in which to select items, they stuck with it.

Additional analyses of participants' reselection behaviors revealed that older adults were more likely than younger adults to reselect items multiple times, in keeping with the results in Experiment 1. Item counts, which tracked how many times each item

was studied by each participant, indicated that across all items younger adults studied items on average 1.10 (.01) times, whereas older adults studied items on average 1.25 (.02) times, a difference that was reliable, F (1, 4643) = 113.34, $p < .001$, partial η^2 = .024. No time-related differences existed in younger adults' reselection behaviors, $F < 1$.

Study Time Allocation

Older adults allocated significantly more time to items within each difficulty level across all grids than did younger adults, F (3, 87) = 33.91, $p < .001$, partial η^2 = .539 (see Table 19). The one exception was the allocation of study time to difficult items within the first grid, for which younger adults actually spent more time than older adults. Despite these differences, participants within both age groups allocated more time to moderately difficult and difficult items than to normatively easy items. This held whether the data were analyzed in terms of the total time spent on items of each type or based on the percentage of time spent on each item type.

Table 19. Mean Study Time Allocation

	Younger- 30		Younger- 60		Older	
Grid/ Item Type	*M*	*(SE)*	*M*	*(SE)*	*M*	*(SE)*
Grid 1						
Easy	1.75	(.28)	1.97	(.55)	6.73	(1.09)
Medium	2.99	(.43)	3.55	(.87)	7.13	(.90)
Hard	6.26	(1.42)	3.89	(.73)	6.96	(1.41)

Table 19 (continued).

Grid 2						
Easy	.82	(.13)	1.05	(.20)	5.10	(.76)
Medium	3.18	(.51)	2.96	(.57)	8.06	(1.19)
Hard	3.72	(.63)	2.55	(.63)	7.00	(1.12)
Grid 3						
Easy	1.21	(.33)	.97	(.20)	3.40	(.54)
Medium	2.38	(.40)	3.22	(.83)	7.43	(1.18)
Hard	3.89	(.69)	4.07	(.77)	8.18	(1.42)
Grid 4						
Easy	1.10	(.22)	1.17	(.21)	4.41	(.73)
Medium	2.28	(.44)	2.55	(.71)	6.82	(1.41)
Hard	3.65	(.77)	3.61	(.93)	7.40	(1.19)
Grid 5						
Easy	1.03	(.47)	1.32	(.43)	4.62	(.95)
Medium	2.10	(.36)	2.66	(.71)	7.11	(1.07)
Hard	2.50	(.60)	3.49	(.94)	7.85	(1.36)
Grid 6						
Easy	1.11	(.26)	1.26	(.53)	5.52	(1.88)
Medium	2.07	(.31)	3.20	(.89)	6.35	(.88)
Hard	3.08	(.60)	2.35	(.63)	8.71	(1.56)

Table 19 (continued).

Grid 7						
Easy	.56	(.12)	1.21	(.38)	3.54	(.84)
Medium	2.52	(.54)	3.40	(.79)	7.05	(.91)
Hard	4.10	(.74)	3.08	(.83)	9.15	(2.17)
Grid 8						
Easy	.95	(.47)	1.35	(.47)	3.99	(.80)
Medium	3.26	(.73)	2.59	(.76)	8.04	(1.16)
Hard	2.49	(.44)	1.97	(.43)	8.51	(1.37)
Grid 9						
Easy	.74	(.21)	.89	(.16)	3.84	(1.06)
Medium	3.22	(.70)	3.39	(.92)	10.02	(1.86)
Hard	3.24	(.54)	2.36	(.51)	8.51	(1.39)
Grid 10						
Easy	.97	(.27)	.57	(.12)	3.64	(.56)
Medium	2.56	(.48)	1.95	(.59)	9.05	(1.37)
Hard	2.89	(.88)	5.12	(1.34)	7.79	(1.09)
Grid 11						
Easy	.91	(.29)	1.16	(.36)	3.06	(.48)
Medium	2.51	(.47)	3.37	(1.15)	8.29	(1.24)
Hard	2.38	(.66)	3.05	(.82)	5.89	(1.34)

Table 19 (continued).

Grid 12						
Easy	.65	(.17)	1.03	(.17)	3.14	(.68)
Medium	2.79	(.62)	3.39	(1.12)	7.91	(1.10)
Hard	2.99	(.71)	3.68	(.93)	5.66	(1.20)

Note. Values represent means and standard errors of time allocated to items of each difficulty level within each of the 12 grids. Younger-30 = younger adults in the 30 second experimental condition; Younger-90= younger adults in the 90 second control condition; Easy = normatively easy items; Medium = moderately difficult items; Hard = normatively difficult items.

Analyses of younger adults' study time allocation behaviors revealed that those given less time actually allocated more time to harder items across grids than those given more time. These differences were not reliable when analyzed based on total time given to items of each type. However, when time allocated to items of each type was analyzed in light of the overall amount of time the younger adults were given (i.e., in terms of percentages), these differences were reliable for moderately difficult, F (1, 58) = 7.60, $p < .01$, partial $\eta^2 = .116$, and difficult F (1, 58) = 13.70, $p < .001$, partial $\eta^2 = .191$, but not easy items, F (1, 58) = 3.34, $p > .05$, partial $\eta^2 = .054$, despite the trend being the same.

No effect of point order was found, either in the analyses of age- or time-related differences.

Strategy Use

As in the first experiment, participants' strategy reports were examined to assess how frequently individuals used each type of strategy. Of interest was whether age-, time-, point-, or goal-related differences would emerge in strategy use. Collapsing across these factors, individuals were found to be more likely to report that they relied on the

similarity between the Spanish and English words to learn easy items, but rote repetition for moderately difficult and difficult items. This held for younger adults regardless of time constraints, but older adults were more likely to report reliance on the similarity of Spanish and English words to learn easy, moderately difficult and difficult items (see Table 20).

Table 20. Reported Strategy Use

	Younger Adults			Older Adults		
Strategy/ Level	E	M	D	E	M	D
Similarity	.78	.19	.19	.66	.20	.21
Rote Repetition	.12	.34	.32	.12	.20	.19
Sentence Gen.	.01	.08	.05	.01	.08	.04
Imagery	.01	.09	.08	.02	.17	.10
Keyword Method	.03	.12	.07	.08	.08	.09
Combination	.01	.05	.05	.01	.01	.01
Not implemented	0	.03	.05	.02	.08	.08
None	.05	.10	.18	.07	.17	.26

Note. Entries are the percentage of time that each strategy was reportedly used as a function of normative item difficulty, where E= Easy, M = moderately difficult, and D = difficult.

Strategy use was also examined as a function of the four experimenter-determined goals participants had been given. Overall, participants relied on the similarity of Spanish and English words to learn the vocabulary terms associated with all four goals.

This held true across both age groups and time conditions. That participants used similarity of Spanish and English words, regardless of which goal they were given, suggests that surface characteristics of the words were a more important determinant of strategy use than either the word or point goals with which the vocabulary terms were associated.

Strategy use was also examined as a function of point values to determine whether individuals were more likely to use normatively effective strategies for higher valued items than lower valued items. Items that had values between 2 and 6 points were considered low value items and those with points above 8 (i.e., 8 to 12) were considered high value items. As may be seen in Figure 11, younger and older adults were more likely to report using normatively effective strategies (i.e., similarity, sentence generation, imagery, and the keyword method) and older adults reported trying to use strategies (i.e., not able to implement) more often for high than lower valued items. In contrast, younger adults were more likely to report not using a strategy for low than higher valued items. Thus younger and older adults showed signs of adjusting their strategy use based on point values.

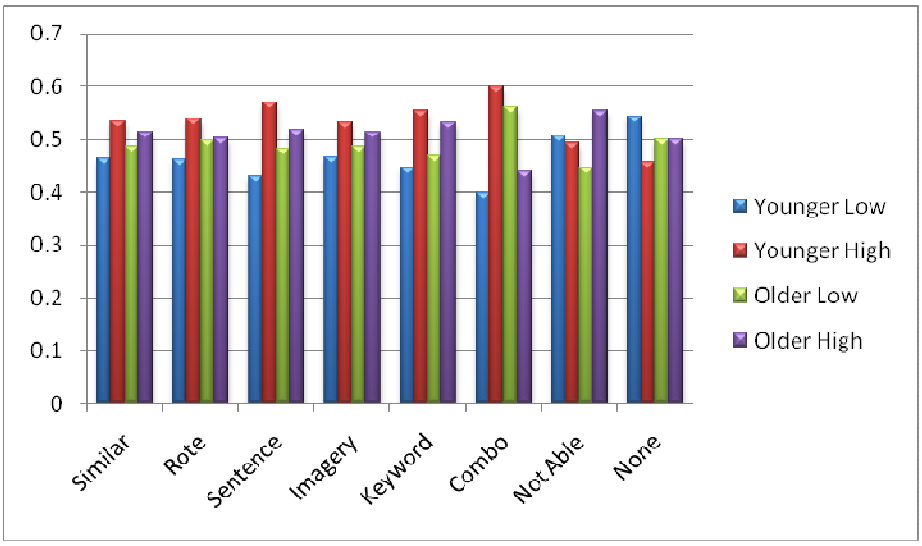

Figure 11. Younger and Older Adults' Strategy Use for High and Low Point Items

Note. Similar = reliance on similarity of Spanish and English words; Rote = rote repetition; Sentence= sentence generation; Imagery = interactive imagery; Keyword= keyword method; Combo = combination of rote repetition and either sentence generation or interactive imagery; Not able = tried but was unable to implement a strategy; None = did not attempt to use a strategy; Low = words with point values between 2 and 6; High = words with point values between 8 and 12.

At first glance the strategy reports in Experiment 2 contradict those obtained in Experiment 1, in which the keyword method was the most popular strategy for all but the most difficult of items. However, the inclusion of a strategy report option that addressed participants' reliance on the similarity of Spanish and English words (i.e., the *similarity* strategy) was inspired by the difficulty participants had in Experiment 1 understanding whether they should report using the keyword method or no strategy when they relied on the overlap between Spanish and English words. Thus it is likely that the strategy reports in Experiment 2 were actually more indicative of what participants were doing while studying items and that strategy reports in Experiment 1 were biased in part by participants' confusion about the nature of the keyword strategy.

Delayed Judgments of Learning

Table 12 contains participants' DJOLs which were significantly higher for younger than older adults for all item types, F (3, 83) = 8.75, $p < .001$, partial $\eta^2 = .240$, indicating that younger adults were more confident than older adults in their ability to recall items.

Recall Performance

Participants' recall was examined as in Experiment 1, but also in light of point values and each of the four goals to see if recall accuracy differed as a function of point value and goal difficulty.

Prestudy Recall Test Performance

Overall, younger adults outscored older adults on the initial test, indicating they either knew more Spanish or were better able to use surface features of the words to correctly guess what the Spanish words meant, F (1, 89) = 32.31, $p < .001$, partial $\eta^2 = .266$ (see Table 21). Analyses of performance for items within each difficulty level indicated that the reason younger adults did better than older adults overall was because they scored higher on both the easy, F (1, 89) = 31.04, $p < .001$, partial $\eta^2 = .259$, and moderately difficult items, F (1, 89) = 5.75, $p < .05$, partial $\eta^2 = .061$, but not the most difficult, $F < 1$. No differences were found in initial performance levels for younger adults in the two time conditions, $F < 1$.

Table 21. Prestudy and Poststudy Recall Test Performance

	Younger- 30	Younger- 60	Older
Test/Level	*M (SE)*	*M (SE)*	*M (SE)*
Prestudy Test			
Easy	.86 (.02)	.87 (.02)	.68 (.04)
Medium	.05 (.01)	.04 (.01)	.02 (.01)
Hard	0	0	0
Overall Actual	.30 (.01)	.31 (.01)	.23 (.01)
Overall Gist	.31 (.01)	.32 (.01)	.25 (.01)
Poststudy Test			
Easy	.95 (.01)	.97 (.01)	.82 (.03)
Medium	.50 (.04)	.45 (.04)	.27 (.04)
Hard	.16 (.04)	.12 (.04)	.04 (.01)
Overall Actual	.54 (.03)	.51 (.02)	.38 (.03)
Overall Gist	.56 (.03)	.53 (.02)	.39 (.03)

Note. Table values are means and standard errors for participants' prestudy and poststudy recall performance as a function of difficulty level and strict versus lenient scoring; Abbreviations: Easy = normatively easy items; Medium = normatively moderately difficult items; Hard = normatively difficulty items; Actual = strictly scored recall performance; Gist = leniently scored recall performance; Younger-30= younger adults in the 30 second experimental condition; Younger-60= younger adults in the 60 second control condition.

Poststudy Recall Test Performance

The final recall test indicated that younger adults learned more both globally and across all 3 difficulty levels than older adults, F (3, 87) = 12.77, $p < .001$, partial η^2 = .306.

Poststudy Recall Test Performance and Strategy Use

The probability of recalling items was examined as a function of which strategy was reportedly used (see Table 22). Which strategies yielded the best recall was again found to differ depending on the normative difficulty of the items. Younger adults experienced success using any of the strategies for easy items, but were slightly less successful if they used rote repetition and were even less likely to recall the easy items if they failed to use any strategy (i.e., reported "none"). For moderately difficult items the younger adults were more likely to recall items if they used one of the normatively effective strategies (i.e., imagery, sentence generation, or keyword method), but also experienced success if they relied on the similarity of the Spanish and English words, which proved to be the most effective strategy for difficult items.

Older adults' recall of easy and difficult items also benefited from capitalizing on the similarity of the words, but the use of either imagery or the keyword method served to increase recall of moderately difficult items. However, the trend was for younger adults to do better with all strategies than older adults. Thus younger adults had significantly higher recall after using the similarity of words, F (1, 73) = 4.11, $p < .05$, partial η^2 = .053, rote, F (1, 74) = 4.65, $p < .05$, partial η^2 = .059, sentence generation, F (1, 49) = 8.28, $p < .01$, partial η^2 = .145, imagery, F (1, 59) = 5.01, $p < .05$, partial η^2 = .078, and no strategy, F (1, 62) = 7.27, $p < .01$, partial η^2 = .105. The keyword method yielded a

similar pattern, but was not reliable, F (1, 50) = 2.92, $p > .05$, partial $\eta^2 = .055$. Only the combination of strategies (i.e., combination strategy report) and inability to implement a strategy did not show any age-related difference, $F < 1$. Thus even when younger adults did not use a strategy they had better recall than older adults, consistent with Dunlosky et al. (2005) in which younger adult participants experienced recall success, even when they were unable to generate a mediator.

Table 22. Conditional Probability of Recall as a Function of Strategy Used

	Younger Adults			Older Adults		
Strategy/ Level	E	M	D	E	M	D
Similarity	.98	.73	.73	.88	.33	.33
Rote Repetition	.89	.58	.25	.56	.35	.04
Sentence Gen.	1.00	.66	.48	.72	.22	.19
Imagery	1.00	.74	.30	.78	.40	.09
Keyword Method	.92	.64	.24	.85	.40	.17
Combination	1.00	.56	.30	.90	.32	.01
Not implemented	1.00	.48	.18	.38	.29	.12
None	.88	.49	.14	.49	.21	.04

Note. Entries are the conditional probabilities that an item was recalled after being encoded with the listed strategy, examined as a function of normative item difficulty, where E= Easy, M = moderately difficult, and D = difficult; Sentence gen = sentence generation; Imagery= interactive imagery; Combination = combination of rote repetition and either sentence generation or interactive imagery; Not implemented = tried to used a strategy but could not.

Poststudy Recall Test Performance and Point Value

Because Castel et al. (2002) found less discrepancy in younger and older adults' recall for high as opposed to lower valued items, participants' recall performance was examined as a function of how many points items were worth. If participants in either age group were selectively learning and recalling higher valued items then their recall for high valued items should be better than for lower valued items. To examine this issue, recall was compared for items worth 2, 4, or 6 points (i.e., low point value) with items worth 8, 10 or 12 points (i.e., high point value). Consistent with the previously discussed recall results, older adults had reliably lower recall than younger adults for both low, F (1, 89) = 23.55, $p < .001$, partial $\eta^2 = .209$, and high valued items, F (1, 89) = 30.58, $p < .001$, partial $\eta^2 = .256$. Yet, younger adults had better recall for high valued items (M = .55, SE = .02), than lower valued items (M = .51, SE = .01), whereas older adults showed no discrepancy in recall as a function of point value (M = .38, SE = .03 for high and M = .37, SE = .03 for low). Thus although older adults' strategy use differed for high and low valued items, suggesting they were aware of the differences in point values, their recall did not differ as a function of points, whereas younger adults' recall did.

Additional analyses indicated that point order influenced the impact of point value on recall performance. Those given point Order 1 (M = .53, SE = .02 for high and M = .50, SE = .02 for low) recalled significantly more than those given point Order 2 (M = .45, SE = .02 for high and M = .42, SE = .02 for low) for both high, F (1, 89) = 6.33, $p < .05$, partial $\eta^2 = .066$, and lower valued items, F (1, 89) = 6.61, $p < .05$, partial $\eta^2 = .069$. Thus receiving grids with point values that favored recall of easy items earlier (i.e., Order 1) rather than later (i.e., Order 2) benefited recall of both low and high valued items.

Separate analyses of younger adults' data revealed no time-related differences, but reliable order differences such that recall was 7% higher in Order 1 than Order 2 for both low and high valued items. Thus although the trend was for recall to be better for higher than lower valued items across both orders, younger adults that received Order 1 (M = .58, SE = .02 for high and M = .54, SE = .02 for low) had reliably higher recall than Order 2 (M = .51, SE = .02 for high and M = .47, SE = .02 for low) for both high, F (1, 58) = 4.35, $p < .05$, partial $\eta^2 = .070$, and low point items, F (1, 58) = 5.95, $p < .05$, partial $\eta^2 = .093$.

Point order effects were found in older adults' recall as well, but were of a much smaller magnitude. Thus older adults that received Order 1 (M = .41, SE = .05 for high and M = .39, SE = .03 for low) recalled more than those that received Order 2 (M = .36, SE = .03 for high and M = .36, SE = .04 for low), but whereas Order 1 older adults showed slight differentiation in their recall for low and high point items, those in Order 2 did not. These differences were not reliable, $F < 1$.

Thus point orders that first favored recall of difficult items hurt both younger and older adults' recall performance, relative to those that received a point order that first favored recall of easy items. However, younger adults showed differentiation between recall of high and low valued items regardless of order, whereas older adults did not.

Poststudy Recall Test Performance and Goals

Recall performance for items associated with each goal was examined to determine the impact of the goals, if any, on the level of performance achieved as well as to assess whether there were age- or time-related differences in participants' recall for each of the four goals. Age-related differences existed in recall for all four goals because

older adults recalled significantly fewer items than did younger adults, F (4, 86) = 9.03, p < .001, partial η^2 = .296. Younger adults' had similar recall performance in both time conditions across all four goals, despite a trend for those given less time to have higher recall for the Low word/ High point, F (1, 58) = 1.14, p > .05, partial η^2 = .019, and High word/ High point goals, F (1, 58) = 2.29, p > .05, partial η^2 = .038. However, reliable point order effects were found for the low word/low point goal, F (1, 89) = 9.01, p < .01, partial η^2 = .092, and the high word/ low point goal, F (1, 89) = 6.98, p < .01, partial η^2 = .073, due to Order 1 participants having higher recall than Order 2 participants. That point order effects were found in recall performance was not surprising, given the earlier evidence of point order effects in goal achievement.

Postdictions

Participants' postdictions indicated that older adults believed they correctly recalled a similar number of items as younger adults, overall, F (1, 89) = 1.10, p > .05, partial η^2 = .012, and for moderately difficult items, F (1, 89) = 1.37, p > .05, partial η^2 = .015. The difference in postdictions for difficult items was not reliable despite the trend for younger adults to postdict higher recall than older adults, F (1, 89) = 3.86, p > .05, partial η^2 = .042. Surprisingly, it was only for easy items that older adults were reliably less confident in their recall than younger adults, F (1, 89) = 13.31, p < .001, partial η^2 = .130 (see Table 12).

Relative Accuracy of Metamemory Judgments

The accuracy of participants' judgments was examined in terms of relative rather than absolute accuracy, as in Experiment 1. However, the absolute accuracy means may be found in Appendix C.

EOLs

Younger adults' resolution was slightly, but not reliably higher than older adults' on the pretest, F (1, 89) = 3.22, $p > .05$, partial $\eta^2 = .035$, but not the posttest for which there was no difference, $F < 1$. However, the magnitude of the gamma values was higher for prestudy than poststudy recall, in keeping with Experiment 1 findings (see Table 18).

DJOLs

Older adults had lower resolution than younger adults on both the pretest, F (1, 89) = 3.82, $p > .05$, partial $\eta^2 = .041$, and posttest, F (1, 89) = 3.74, $p > .05$, partial $\eta^2 = .040$, but the differences were not reliable. However, younger adults in the 30 second condition had better resolution on average than those in the 60 second condition, F (1, 58) = 5.05, $p < .05$, partial $\eta^2 = .080$. Thus having more time to study items did not improve 60 second younger adults' ability to differentiate items they would and would not recall.

Goal Setting for the Hypothetical Trial 2

Older adults' self-determined goals for the hypothetical second study opportunity were reliably lower than younger adults', both overall and across all difficulty levels, F (4, 86) = 3.93, $p < .01$, partial $\eta^2 = .104$ (see Table 12). Participants' goals were significantly and positively correlated with recall performance, suggesting individuals' performance with experimenter-determined goals predicted their self-determined goals.

Discussion

Research in the metacognitive domain (e.g., Dunlosky & Connor, 1997) has suggested older adults are less efficient at regulating their study time than younger adults. In contrast, the memory research conducted by Castel et al. (2002) suggested points could

reduce age-related differences in recall performance because older adults were equally, if not more, adept at controlling their selection and recall behaviors so as to maximize points earned. However, the results in Experiment 2 were more consistent with the Dunlosky and Connor than the Castel et al. results because older adults had reliably lower recall performance than younger adults for low and high valued items despite allocating significantly more time to items. Thus unlike the Castel et al. findings, the recall of older adults in the present study did not differ as a function of point values. However, younger and older adults' recall performance did differ depending on point order (i.e., whether individuals received Point Order 1 or 2), and strategy use varied based on point values, suggesting participants in both age groups were aware of the points. Yet EOLs for selected versus ignored items, the selection order gammas (with subjective or objective difficulty), and study time allocation behaviors all combined to suggest that younger adults' selections were more influenced by point values than older adults'. Thus younger adults were more likely to ignore high EOL (i.e., difficult) and to select low EOL (i.e., easy) items for low point goals, but to do the opposite for high point goals. In contrast, older adults' selections were characterized by emphasizing easy and moderately difficult items at the expense of difficult items, regardless of experimenter-determined goals.

Because younger adults were more likely to select items based on points whereas older adults focused on selecting the easiest items, regardless of points, younger adults' selections were more reactive to point order than older adults'. Yet younger adults were able to obtain higher recall levels for higher than lower valued items across both point orders, whereas older adults only showed differentiation in high and low point recall when they received Order 1. This raises concerns that the point order manipulation,

which was included solely to test for order effects, may have prevented older adults (and younger adults to some extent) from capitalizing on point values. Thus although point values were found to have a motivating effect, in the sense that participants in both age groups were more likely to pursue the high word goal when it was paired with a high point than a low point goal, older adults were less likely to capitalize on points to achieve the goal, perhaps because point order hurt their ability to form and implement a selection plan more than it did younger adults.

Of additional interest was the finding that point order affected selection order and the mean level of difficulty of each selection. Thus although individuals in both age groups selected items in a manner consistent with RPL model predictions (i.e., easier first), point order influenced the mean level of difficulty of items initially selected. That is, those given Order 1 were more likely to select easier items first and Order 2 participants were more likely to select slightly more difficult items initially, in keeping with what the points were rewarding in each order. Thus, the structure of point values was sufficient to weaken the RPL-consistent selection order effects found, both in Experiment 1 and for those given point Order 1 in Experiment 2.

Point order was also found to interact with time constraints to influence whether younger adults selected items based on normative or subjective difficulty, but did not impact the basis of older adults' selections, which were similar across both point orders. Yet point order affected both younger and older adults' recall performance, which suggests that the influence of points and point order on item selection behaviors may be distinct from how they affect recall. That point order affected selection behaviors and recall differently could have occurred because participants had no way of knowing how

their item selection and study time allocation behaviors would influence their ability to recall items, if at all. It is possible that if participants were given test experience in between grids which allowed them to monitor whether their study behaviors were translating into effective recall that this dissociation in study behaviors and recall would disappear. Unfortunately the limited number of subjects within each age group and point order condition made fully disentangling the impact of point order and point values impossible. However, because point order and point values changed the basis of item selection and served to reduce RPL selection order effects, it would seem worthwhile to further examine the impact of point values and point order on participants' study behaviors and recall in the context of multiple study-test trials.

Experiment 2 provided evidence that the addition of points, which had been included as extrinsic motivation to examine whether they would increase goal pursuit, goal achievement, and learning (i.e., recall performance), served to only increase goal pursuit, and then only when presented in a particular order (Order 1), and instead seemed to hurt recall performance, when compared with recall rates obtained in Experiment 1. Together, the results indicated that points might have increased participants' motivation, but hurt their ability to execute a selection plan, depending on point order. More important, the impact of points and point order highlighted that RPL selection effects are malleable, a surprising finding, given the consistency with which Metcalfe and colleagues (Kornell & Metcalfe, 2006; Metcalfe, 2002; Metcalfe & Kornell, 2003, 2005) have found the easy to difficult selection order pattern.

CHAPTER 4

GENERAL DISCUSSION

Several key results were obtained across the two experiments, which served to both expand our understanding of younger and older adults' self-regulated learning behaviors and raise questions for future research. First, older adults, like younger adults, were found to select items in an easy to difficult order, thereby providing the first evidence to date that RPL effects would extend to older adults' item selection behaviors. These findings contrast those obtained in experiments that have used items with little variability in difficulty level in which younger and older adults have been found to select items judged to be difficult (those given low JOLs) rather than easier items first (Dunlosky & Hertzog, 1997; Nelson & Leonesio, 1988). Thus whether individuals are likely to first select difficult or easier items seems to critically depend on whether the items vary in difficulty. When items are all of similar difficulty level, individuals tend to select items in a manner that is more consistent with the DRM (Nelson & Leonesio, 1988). However, when items are strongly graded by difficulty level, and individuals are novices, selection order is more consistent with RPL model predictions.

The heterogeneity in item type not only influenced item selection order, but also the basis of item selection. Thus, the second key finding was that younger and older adults in Experiment 1 were both more likely to base their selections on normative (objective) than subjective item difficulty. Yet when individuals utilized subjective difficulty to select items, they did so at a global rather than fine-tuned (item) level, as indicated by the binned EOLs having a higher gamma than the item-level EOLs. A

similar reliance on normative difficulty was found in Experiment 2, but only for those given a point order that first favored recall of easy items (i.e., Order 1) and more so for younger adults in the 60 second than for those in the 30 second condition. For younger adults that received Order 2, the gammas for normative and subjective difficulty were non-existent, especially for those given less study time. Yet older adults' gammas were unaffected by point order, indicating they continued to rely on normative difficulty as a basis for item selection, even when point values and point order would have made it better to have deviated. Thus younger adults were better able to adapt their selections and basis of selections than older adults so as to increase the likelihood of goal achievement.

Although participants' item selection behaviors were consistent with RPL model predictions, the third important finding was that participants' study time allocation behaviors provided support for both the DRM and RPL model's predictions. Across both experiments younger and older adults allocated more time to the more difficult items than the easier items. This would seem more consistent with the DRM prediction that individuals will focus the most time on the least well learned items so as to reduce the discrepancy between the current and desired levels of knowing. However, the fact that younger adults in the shorter time conditions allocated proportionally more time to the moderately difficult items and less time to the difficult items than younger adults given more time would seem to support the RPL prediction that individuals will focus and allocate time to items they feel they are most likely to be able to learn when placed under time constraints. Thus study time allocation behaviors were consistent with both models of self-regulated learning, whereas item selection behaviors were consistent with only RPL predictions.

The fact that each model accounted for different aspects of self-regulated learning (i.e., study time allocation and item selection) would seem consistent with Thiede and Dunlosky's (1999) notion of a hierarchical model in which superordinate goals, such as selecting easier items first or as few items as possible to achieve self- or experimenter-determined goals, may be attained via a discrepancy reduction mechanism in which individuals attempt to reduce the discrepancy between the number of items already learned and the number of items they want to learn, or study more difficult items longer in an attempt to achieve their goals. Thus as their hierarchical model implies, it seems unlikely that either the DRM or RPL model will adequately account for all self-regulated learning behaviors under all conditions. Instead, it seems more probable that each model will continue to account for different aspects of self-regulated learning, with task demands and person characteristics (e.g., motivation, memory beliefs, goals) influencing which model is most likely to be supported by the data.

Although all participants allocated more time to difficult than easier items, the fourth key finding was that younger adults were more effective at allocating study time than older adults, ironically more so when given less time. Thus although younger adults given more study time allocated similar amounts of time to moderately difficult and difficult items, neither older nor younger adults' recall benefited from this additional time. In fact, in both experiments the additional time that older and (60 second) younger adults spent on difficult items could be considered a labor in vain because they had similar (younger adults) or worse (older adults) recall rates than those younger adults given less time. It remains possible that the additional study time failed to benefit recall of more difficult items because of differences in the information uptake functions for

these items, because participants' perceived rates of learning (i.e., jROLs) caused them to cease studying items before they achieved additional learning gains, or because of some combination of these or other factors (Metcalfe & Kornell, 2005). Yet the fact that older adults were more likely to reselect and allocate additional time to items suggests they were trying to maximize the likelihood of recalling items, either through spaced practice or maintenance rehearsal, but these behaviors did not reduce the age-related differences in recall.

The fifth key finding was that differences in study time allocation and recall did not seem to be due to either differences in memory beliefs (i.e., memory self-efficacy) or choice of encoding strategy. Memory beliefs ratings did not differ reliably across age groups, and both younger and older adults were found to use normatively effective strategies more than rote repetition, except on the most difficult of items. A similar pattern was found in the second experiment if one assumes that many instances of what Experiment 1 participants called "keyword method" were instead reliance on the similarity of Spanish and English words. However, Experiment 2 revealed that younger adults adapted their strategy use as a function of point value more than older adults did, again suggesting greater metacognitive control in younger adults. That strategy use did not differ as a function of goals suggests that strategy use decisions were occurring at the item (point) level rather than at the goal level. Thus unless older adults were unable to decode the strategies they used (Dunlosky, et al., 2005), strategy use would not seem to account for the age-related differences that were found.

Instead, it seems likely that the age-related differences found in item selection, study time allocation, and recall performance were due to age-related declines in either

metacognitive control, episodic memory, or some combination of these. If older adults experienced an age-related decrease in metacognitive control, it could affect their ability to effectively manage their study time allocation and item selection behaviors, which could in turn impact their ability to achieve the goals. If older adults experienced age-related declines in episodic memory then this could prevent goal achievement because they would be unable to recall enough items to achieve the goals, irrespective of their study behaviors. Thus older adults experiencing declines in one or both of these would be expected to have great difficulty effectively pursuing and achieving the goals.

Consistent with the episodic memory explanation is the fact that older adults had reliably lower recall than younger adults for all item types in both experiments. Yet recall was better for individuals in both age groups in Experiment 1 than in Experiment 2, which suggests the points and goals may have overwhelmed participants' ability to exert metacognitive control, especially when given point Order 2, a finding consistent with both a metacognitive control and episodic memory explanation. Ironically, then, the inclusion of points as extrinsic motivation served to reduce rather than increase performance, relative to when performance was based solely on intrinsic motivation.

Additional evidence that points may have overwhelmed participants is derived from a comparison of the items selected and ignored by participants across the two experiments. In Experiment 1 participants tended to ignore easier and select more difficult items, whereas participants in Experiment 2 showed an overall tendency to select easier and avoid more difficult items, as measured by their EOLs. Moreover, older adults selected easier items and ignored more difficult items than younger adults. Whether participants in Experiment 2 showed different EOL select/ignore patterns than those in

Experiment 1 because they were overwhelmed by the points or because the points were not sufficiently motivating remains an open question and raises questions about how intrinsic and extrinsic motivation combine to influence task performance and whether they do so differently at different ages.

These findings would seem to cloud the issue of whether the age-related differences were a result of differences in motivation or metacognitive control. However, inconsistent with a motivational explanation for the observed differences is the fact that older adults were more likely to repeatedly study items than younger adults in both experiments and to select more items for study in Experiment 2. If older adults were concerned only about goal achievement, they should have attended to fewer rather than more items, especially for the lower goals that afforded it. Thus it seems that the observed age-related differences in self-regulated learning behaviors were due to age-related differences in metacognitive control, not to age-related differences in motivation. If anything, the data suggested greater motivation in older than younger adults.

The two experiments thus provided interesting insight into what influenced younger and older adults' self-regulated learning behaviors. However, the results also raised questions worthy of future empirical research. For example, individuals in both experiments relied on normative rather than subjective difficulty as a basis for item selection decisions (at least in Order 1), but would normative difficulty still be more predictive of selection order if participants were *not* told about normative difficulty? It seems likely that item selection behaviors would be based more on EOLs if people were not told which items were easy, moderately difficult, and difficult. This of course remains an empirical question.

Heterogeneity versus homogeneity of items has been found to influence whether item selection order is more consistent with RPL model or DRM predictions. Yet if individuals are truly inspired to select items within their range of proximal learning, as the RPL model predicts, then it should not require telling participants which items are easiest, nor should it require that items be arranged in an order based on normative difficulty (i.e., easy in the left column, moderately difficult in the middle, and difficult in the right column). Perhaps a stronger test of RPL predictions would be provided by scrambling presentation order within a grid so that items of each difficulty were no longer contained within single columns, but instead scattered throughout. The test would become even stronger if individuals were forced to make selection decisions based on their own subjective impressions of difficulty (i.e., rely on their EOLs) as opposed to being told that items varied in normative difficulty and which items were of each type. That is, telling individuals that certain items are easier to recall may combine with the fact that they appear easier to make individuals more likely to study them and therefore more likely to recall them (i.e., create a self-fulfilling prophecy). Seeing what individuals choose to study first when they have not been told about normative difficulty and when items are no longer arranged as a function of difficulty level would provide a much stronger test of the RPL model.

However, because the surface characteristics of the Spanish words (e.g., word length, similarity to English words) make the easier seem easier, the moderately difficult seem moderately difficult and the difficult appear difficult, it would be interesting to present grids that only contained items from a single difficulty level to see if individuals would then be more likely to select items based on fine-tuned item-level EOLs (as

opposed to global judgments of difficulty, which should no longer matter) and whether selection order would then be more consistent with DRM (since items would be heterogeneous with regard to difficulty level). What if individuals were instead asked to study 3 X 3 grids that only contained easy, moderately difficult, or difficult items (i.e., each grid was homogeneous with regard to difficulty level), but participants could choose the order in which they studied the grids (e.g., easy grid first or difficult grid first)--would they still choose easier before difficult items as the RPL posits?

The fact that point order made previously high gammas between subjective and objective difficulty and selection order negligible suggests another interesting experiment would be to simply reverse the column order so that difficult items appeared in column 1 and easy items in column 3. If this new order were combined with the point orders used in Experiment 2 it seems likely that the point orders would affect selection behaviors and recall results in exactly the opposite way. That is, gammas would likely be reduced for those in Order 1, but be higher for those given point Order 2. Changing the order of columns, while maintaining the two point orders, would allow examination of whether gammas were nonexistent for those given Order 2 in Experiment 2 because of a focus on points or difficulty level. It would also test whether RPL selection effects occur simply based on the order of columns.

Finally, it would seem worthwhile to further examine the impact of intrinsic and extrinsic motivation on participants' self-regulated learning behaviors. Participants were told to try and achieve the four experimenter-determined goals in Experiment 2, but were never told whether they succeeded in achieving the goals. The fact that Castel et al. (2002) provided performance feedback to their participants raises the question whether

the feedback allowed older adults to better control their study behaviors, thereby reducing the age-related differences in recall of high but not lower valued items. Moreover, West and Thorn (2001) found that providing (positive) feedback influenced younger and older adults' tendency to pursue and raise goals, regardless of whether the goals were self-initiated or experimenter initiated. The Castel et al. and West and Thorn results suggest it would be interesting to actually tell participants how they did (e.g., how many points they earned and words they recalled) before exposing them to a second learning trial. Would providing feedback to participants serve to increase or decrease their motivation? Would performance increase as a result of having been tested? Would the impact of feedback differ for younger and older adults since younger adults would likely receive positive and older adults would likely receive negative feedback? These questions address the increased focus that self-regulated learning (SRL) researchers have placed on the impact that motivation has on SRL (Boekaerts, 1995). Because models of SRL are now incorporating motivation and metacognition, it would seem important to simultaneously test these in future empirical research.

Future research should examine the conditions under which RPL consistent selection behaviors are and are not found. That is, what are the boundary conditions under which RPL consistent effects are more or less likely to be found? These experiments have provided an initial basis for examining the constraints that task manipulations (e.g., points, goals, point order, etc.) might have on the likelihood of RPL model predictions holding, based on the finding that selection order was somewhat malleable to points and goals. Yet many questions still remain as to what drives self-regulated learning. Now that there is evidence that older adults select items in an RPL

consistent fashion when the standard Metcalfe and Kornell (2005) presentation method is used, it opens the door to examine whether task manipulations (e.g., presentation methods, feedback, etc.) influence younger and older adults' self-regulated learning behaviors in similar ways. Given larger sample sizes it may also be possible to evaluate whether age-related differences in self-regulated learning behaviors, if they do exist, are due to differences in motivation, metacognitive control, cognitive ability, or some combination of these.

FOOTNOTES

1. Actual time spent studying all items rather than allotted time (i.e., either 45 or 90 seconds per grid) was used as the denominator because not all participants chose to use all of their allotted time. Younger adults given 45 seconds per grid used on average 81% (4.0) of their allotted time, whereas younger adults given 90 seconds only used 59% (5.0) of their time and older adults used 69% (5.0). The analyses therefore focused on the time participants did use and how it was divided between studying easy, moderately difficult, and difficult items.
2. The counterbalancing of goal orders was initially balanced across participants, but became uneven when several older adults failed to complete the experiment and data analyses revealed that additional individuals' data could not be kept and analyzed (see footnote 3).
3. Two younger and one older adult used the "finished studying this grid" option to avoid studying any items. These three individuals were excluded from the reported analyses. Another older adult avoided studying items associated with one of the goals (i.e., did not select items within any of the three grids for that goal). Because the participant had studied 52 of the 72 items, which was within the range of others who had not skipped entire grids, his data were left in the analyses and selections for the three skipped grids were recorded as missing values.
4. A computer problem resulted in the loss of PBMI data for 11 participants in the second experiment. Four of the lost files were for older adults, 1 lost file was for a younger adult in the 30 second time (experimental) condition and the other 6 files

were for younger adults in the 60 second (control) condition. It remains possible that the loss of these data impacted the analyses of participants' memory beliefs. However, similar patterns existed in participants' ratings across both experiments.

5. Fewer selections were examined in the second experiment because the grids in Experiment 2 only contained six items as compared to the nine item grids in Experiment 1. The mean level of difficulty was therefore examined for only the first six, rather than nine, items selected.

APPENDIX A

SPANISH-ENGLISH VOCABULARY PAIRS

Easy Items
airport – aeropuerto
aspirin – aspirina
asteroid – asteroide
battery – bateria
biography – biografia
buffalo – bufalo
captain – capitan
defensive – defensivo
deposit – deposito
ecology – ecologia
education – educacion
encyclopedia – enciclopedia
family – familia
fantastic – fantastico
fortune – fortuna
fossil – fosil
fugitive – fugitive
future – futuro
galaxy – galaxia
gasoline – gasoline
graphic – grafica
lion – leon
photograph – fotografia
phrase – frase
tomato – tomate

Moderately Difficult Items
ancient – antiguo
blood – sangre
blush – rubor
brightness – brillantez
dishwater – lavazaas
doorlatch – picaporte
dove – paloma
early – temprano
earthquake – terremoto
enrage – enrabiar
farmer – labrador
food – alimento

foundation – cimentacion
lately – ultimamente
motorboat – lancha
music hall – vodevil
pen – pluma
penholder – portaplumas
pregnancy – embarazo
sleepwalking – noctambulo
subhuman – infrahumano
temptation – tentacion
tragedy – desgracia
turn – volver
wind – viento

Difficult Items
blowpipe – cerbatana
clawhammer – arrancaclavos
crack – requebrajadura
cranberry – arandano
drain – alcantarilla
fascination – embobamiento
gibberish – jerigonza
gluttony – lambisconeria
handkerchief – paliacate
harvester – cosechadora
insecticide – garrapaticida
locksmith – cerrajero
masonry – mamposteria
merry – jacarandoso
migraine – zangarriana
mischief – barrabasada
mower – guadanadora
sarcasm – socarroneria
shredder - trituradora
sorcery – jorguineria
stagnation – anquilasamiento
stain – chafarrinada
stutter – tartamudez
swallow – golondrina
weak – gurrumino

APPENDIX B
EXPERIMENT 1 ABSOLUTE ACCURACY OF JUDGMENTS

	Younger		Older	
Judgment Type	*M*	*(SE)*	*M*	*(SE)*
Prestudy Prediction				
Global	-11.84	(1.46)	-6.50	(2.61)
Easy	-4.33	(.49)	-7.82	(.87)
Medium	-4.46	(.70)	-0.33	(1.31)
Hard	0.07	(.81)	4.16	(.87)
Poststudy Prediction				
Global	-6.20	(1.70)	-2.77	(2.05)
Easy	-2.61	(.53)	-3.81	(.90)
Medium	-2.20	(.63)	1.92	(.98)
Hard	2.11	(.79)	4.85	(.75)
DJOLs				
Global	1.78	(1.14)	3.23	(1.66)
Easy	-1.02	(.51)	-0.73	(.75)
Medium	1.03	(.50)	2.83	(.73)
Hard	1.77	(.37)	1.11	(.55)
Postdictions				
Global	-6.70	(1.18)	0.10	(1.92)
Easy	-2.63	(.39)	-3.71	(.69)
Medium	-3.47	(.55)	1.66	(1.15)
Hard	-0.35	(.68)	4.70	(.88)

Note. Values reflect the discrepancy between participants' mean recall performance and what they predicted/postdicted they would recall, thus positive values reflect overconfidence and negative values underconfidence; Easy = normatively easy items; Medium = moderately difficult items; Hard = normatively difficult items; DJOLs = delayed judgments of learning.

APPENDIX C

EXPERIMENT 2 ABSOLUTE ACCURACY OF JUDGMENTS

	Younger		Older	
Judgment Type	*M*	*(SE)*	*M*	*(SE)*
DJOLs				
Global	4.84	(.77)	3.96	(2.13)
Easy	-.17	(.16)	-1.79	(.97)
Medium	2.25	(.43)	3.49	(.81)
Hard	2.77	(.45)	2.27	(.70)
Postdictions				
Global	-8.07	(1.25)	-.42	(2.60)
Easy	-4.52	(.62)	-5.81	(1.04)
Medium	-1.35	(.74)	2.10	(.97)
Hard	2.65	(.65)	2.90	(.52)

Note. Values reflect the discrepancy between participants' mean recall performance and what they predicted/postdicted they would recall, thus positive values reflect overconfidence and negative values underconfidence; Easy = normatively easy items; Medium = moderately difficult items; Hard = normatively difficult items; DJOLs = delayed judgments of learning.

REFERENCES

Atkinson, R. C. (1972). Optimizing the learning of a second-language vocabulary. *Journal of Experimental Psychology, 96,* 124-129.

Boekaerts, M. (1995). Self-regulated learning: Bridging the gap between metacognitive and metamotivation theories. *Educational Psychologist, 30,* 195-200.

Brigham, M. C., & Pressley, M. (1988). Cognitive monitoring and strategy choice in younger and older adults. *Psychology and Aging, 3,* 249-257.

Castel, A. D., Benjamin, A. S., Craik, F. I. M., & Watkins, M. J. (2002). The effects of aging on selectivity and control in short-term recall. *Memory & Cognition, 30,* 1078-1085.

Dunlosky, J., & Connor, L. T. (1997). Age differences in the allocation of study time account for age differences in memory performance. *Memory and Cognition, 25,* 691-700.

Dunlosky, J., & Hertzog, C. (1997). Older and younger adults use a functionally identical algorithm to select items for restudy during multi-trial learning. *Journal of Gerontology: Psychological Sciences, 52,* P178-P186.

Dunlosky, J., & Hertzog, C. (1998). Training programs to improve learning in later adulthood: Helping older adults educate themselves. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 249-275). Mahwah, NJ: Erlbaum.

Dunlosky, J., & Hertzog, C. (2000). Updating knowledge about encoding strategies: A componential analysis of learning about strategy effectiveness from task experience. *Psychology and Aging, 15,* 462-474.

Dunlosky, J., Hertzog, C., & Powell-Moman, A. (2005). The contribution of mediator-based deficiencies to age-related differences in associative learning. *Developmental Psychology, 41,* 389-400.

Dunlosky, J., & Thiede, K. W. (2004). Causes and constraints of the shift-to-easier materials effect in the control of study. *Memory & Cognition, 32,* 779-788.

Ekstrom, R. B., French, J. W. and Harman, H. H. (1976) *Kit of factor-referenced cognitive tests,* Educational Testing Service, Princeton.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (in press). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavioral Research Methods.*

Gruneberg, M. M, & Pascoe, K. (1996). The effectiveness of the keyword method for receptive and productive foreign vocabulary learning in the elderly. *Contemporary Educational Psychology, 21,* 102-109.

Inquisit 3.0.0.0 [Computer software]. (2007). Seattle, WA: Millisecond Software, LLC.

Koriat, A., Ma'ayan, H., & Nussinson, R. (2006). The intricate relationships between monitoring and control in metacognition: Lessons for the cause-and-effect relation between subjective experience and behavior. *Journal of Experimental Psychology: General, 135,* 36-69.

Koriat, A., Ma'ayan, H., Sheffer, L., & Bjork, R. A. (2006). Exploring a mnemonic debiasing account of the underconfidence-with-practice effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 595-608.

Koriat, A., Sheffer, L., & Ma'ayan, H. (2002). Comparing objective and subjective learning curves: Judgments of learning exhibit increased underconfidence with practice. *Journal of Experimental Psychology: General, 131,*147-162.

Kornell, N., & Metcalfe, J. (2006). Study efficacy and the region of proximal learning framework. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 609-622.

Lachman, M. E., Bandura, M., Weaver, S. L., & Elliott, E. (1995). Assessing memory control beliefs: The memory controllability inventory. *Aging and Cognition, 2,* 67-84.

Lawson, M. J., & Hogben, D. (1998). Learning and recall of foreign-language vocabulary: Effects of a keyword strategy for immediate and delayed recall. *Learning and Instruction, 8,* 179-194.

Le Ny, J. F., Denhiere, G., & Le Taillanter, D. (1972). Regulation of study-time and interstimulus similarity in self-paced learning conditions. *Acta Psychologica, 36*, 280-289.

Lineweaver, T. T., & Hertzog, C. (1998). Adults' efficacy and control beliefs regarding memory and aging: Separating general from personal beliefs. *Aging, Neuropsychology, & Cognition, 5,* 264-296.

Metcalfe, J. (2002). Is study time allocated selectively to a region of proximal learning. *Journal of Experimental Psychology: General, 131,* 349-363.

Metcalfe, J., & Kornell, N. (2003). The dynamics of learning and allocation of study time to a region of proximal learning. *Journal of Experimental Psychology: General, 132,* 532-542.

Metcalfe, J., & Kornell, N. (2005). A region of proximal learning model of study time allocation. *Journal of Memory and Language, 52,* 463-477.

Microsoft Visual C#, Version 2.0 [computer software]. (2005). Microsoft Corporation.

Nelson, T. O., Dunlosky, J., Graf, A., & Narens, L. (1994). Utilization of metacognitive judgments in the allocation of study during multitrial learning. *Psychological Science, 5,* 207-213.

Nelson, T. O., & Leonesio, R. J. (1988). Allocation of self-paced study time and the "labor-in-vain effect". *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 676-686.

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 26, pp. 125-141). New York: Academic Press.

Peligrina, S., Bajo, M. T., & Justicia, F. (2000). Differential allocation of study time: Incomplete compensation for the difficulty of the materials. *Memory, 8*, 377-392.

Sagarra, N., & Alba, M. (2006). The key is in the keyword: L2 vocabulary learning methods with beginning learners of Spanish. *The Modern Language Journal, 90,* 228-243.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review, 103*, 403-428.

Salthouse, T. A., & Babcock, (1991). Decomposing adult age differences in working memory. *Developmental Psychology, 27,* 763-777.

Son, L. K., & Metcalfe, J. (2000). Metacognitive and control strategies in study-time allocation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 204-221.

Sparks, R., & Ganschow, L. (1993). Searching for the cognitive locus of foreign language learning difficulties: Linking first and second language learning. *Modern Language Journal, 77,* 289-302.

Thiede, K. W., & 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,* 1024-1037.

Underwood, B. J. (1966). Individual and group predictions of item difficulty for free learning. *Journal of Experimental Psychology, 71,* 673-679.

West, R. L., Bagwell, D. K., & Dark-Freudeman, A. (2005). Memory and goal setting: The response of older and younger adults to positive and objective feedback. *Psychology and Aging, 20,* 195-201.

West, R. L., & Thorn, R. M. (2001). Goal-setting, self-efficacy, and memory performance in older and younger adults. *Experimental Aging Research, 27,* 41-65.

West, R. L., Thorn, R. M., & Bagwell, D. K. (2003). Memory performance and beliefs as a function of goal setting and aging. *Psychology and Aging, 18,* 111-125.

West, R. L., Welch, D. C., & Thorn, R. M. (2001). Effects of goal-setting and feedback on memory performance and beliefs among older and younger adults. *Psychology and Aging, 16,* 240-250.