

Washington University in St. Louis

Washington University Open Scholarship

Arts & Sciences Electronic Theses and
Dissertations

Arts & Sciences

Summer 8-15-2015

How Do Voters Remember Flip-Flopping? Memorial and Social Consequences of Change Recollection

Adam Lewis Putnam
Washington University in St. Louis

Follow this and additional works at: https://openscholarship.wustl.edu/art_sci_etds



Part of the [Psychology Commons](#)

Recommended Citation

Putnam, Adam Lewis, "How Do Voters Remember Flip-Flopping? Memorial and Social Consequences of Change Recollection" (2015). *Arts & Sciences Electronic Theses and Dissertations*. 562.
https://openscholarship.wustl.edu/art_sci_etds/562

This Dissertation is brought to you for free and open access by the Arts & Sciences at Washington University Open Scholarship. It has been accepted for inclusion in Arts & Sciences Electronic Theses and Dissertations by an authorized administrator of Washington University Open Scholarship. For more information, please contact digital@wumail.wustl.edu.

WASHINGTON UNIVERSITY IN ST. LOUIS

Department of Psychology

Dissertation Examination Committee:

Henry L. Roediger, III, Chair

David A. Balota

Ian G. Dobbins

Kathleen B. McDermott

James V. Wertsch

How do Voters Remember Flip-Flopping?
Memorial and Social Consequences of Change Recollection
by
Adam Lewis Putnam

A dissertation presented to the
Graduate School of Arts & Sciences
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

August 2015
Saint Louis, Missouri

© 2015, Adam Lewis Putnam

Table of Contents

List of Figures	v
List of Tables	vi
Acknowledgments.....	viii
Abstract.....	ix
Chapter 1: Introduction.....	1
1.1 Classic Interference Theory	3
1.2 When Change Does Not Lead to Interference	4
1.3 The Recursive Reminders Framework.....	5
1.4 Prior Research: Detecting and Recollecting Change can Reduce Proactive Interference	8
1.5 Prior Research: Recollecting Change With Political Materials.....	11
1.6 How Do People Interpret a Politician’s Change in Position?	16
1.7 How Partisanship Affects Judgments and Memory.....	18
1.8 Overview of The Experiments	20
1.9 Experiment 1: The Intersection of Attitudes With Change Recollection	21
1.10 Experiment 2: Does the News Create a Misinformation Effect?.....	21
1.11 Experiment 3: Can Politicians Use Language to Draw Attention to a Change in Position?.....	22
Chapter 2: Norming Study	24
2.1 Method	24
2.1.1 Subjects	24
2.1.2 Materials	25
2.1.3 Procedure	27
2.1.4 Results.....	29
Chapter 3: Experiment 1	31
3.1 Method	32
3.1.1 Subjects	32
3.1.2 Materials	33
3.1.3 Design and Counterbalancing.....	33

3.2 Procedure	34
3.3 Results.....	36
3.3.1 Recall	37
3.3.2 Change Recollection	39
3.3.3 Recall Conditionalized on Change Recollection	40
3.3.4 Subject Judgment Results	43
3.4 Experiment 1 Discussion	50
Chapter 4: Experiment 2	52
4.1 Method	56
4.1.1 Subjects	56
4.1.2 Design and Procedure	56
4.2 Results.....	58
4.2.1 Recall and Change Recollection	58
4.2.2 Recall Conditionalized on Change Recollection	59
4.2.3 News Phase Responses	61
4.2.4 Did the News Statements Influence Recall and Change Recollection?.....	65
4.2.5 Did News Phase Accuracy Influence Later Recall and Change Recollection?	67
4.3 Discussion.....	70
Chapter 5: Experiment 3	74
5.1 Method	77
5.1.1 Subjects	77
5.1.2 Design, Procedure, and Materials	77
5.2 Results.....	78
5.2.1 Recall and Change Recollection	79
5.2.2 Recall Conditionalized on Change Recollection	79
5.2.3 Recall as a Function of Introductory Phrase	81
5.2.4 Change Recollection as a Function of Introductory Phrase.....	82
5.3 Discussion.....	83
Chapter 6: General Discussion.....	86
6.1 Theoretical Implications	87
6.1.1 The Recursive Reminders Framework.....	87

6.1.2 The Misinformation Effect	90
6.1.3 Attitudes and Memory	91
6.2 Potential Limitations.....	94
6.2.1 Alternatives to the Recursive Reminders Framework	94
6.2.2 Item Selection Effects	95
6.2.3 Artificial Nature of the Paradigm	96
6.3 Applications to Politics.....	98
6.4 Concluding Comments.....	100
References.....	102
Appendix A: Political Materials	110
Appendix B: Photos of the Political Candidates.....	115
Appendix C: Political Expertise Quiz.....	116
Appendix D: Objective Political Orientation Survey	118
Appendix E: Additional Results and Discussion - Experiment 1	119
Appendix F: Additional Results and Discussion - Experiment 2	127
Appendix G: Additional Results and Discussion - Experiment 3.....	139

List of Figures

Figure 1.1: Results of Putnam et al. (2014), Experiments 1 - 3. Cued recall for changed positions conditionalized on change recollection compared to recall for control positions. 13

Figure 1.2: Results from Experiment 3 of Putnam et al. (2014). Cued recall for changed positions, conditionalized on change detection and change recollection in the informed + detection group. The black line indicates recall for the control positions, with standard error. 15

Figure 3.1: Cued recall as a function of self-reported political orientation for the two different candidates. Error bars represent standard error of the means. 38

Figure 3.2: Cued recall for the changed positions conditionalized on change recollection, as compared to the control condition in Experiment 1. Error bars represent standard error of the means. 41

Figure 3.3: Scatterplots showing correlations between change recollection hit rate and subject judgments after the recall test in Experiment 1. * $p \leq .05$ 48

Figure 4.1: Cued recall for the changed positions conditionalized on change recollection, as compared to the control condition, Experiment 2. Error bars represent standard error of the means. 60

Figure 4.2: The correct verification, misleading endorsement, and miss response rates for the news statements during Experiment 2. See the main text for descriptions of how responses were indexed..... 64

Figure 4.3: Probability of recall in Experiment 2 as a function of item type and news statement type. 66

Figure 4.4: Recall by item type conditionalized on news phase verification accuracy. 69

Figure 5.1: Cued recall for the changed positions conditionalized on change recollection, as compared to the control condition, Experiment 3. Error bars represent standard error of the means. 80

Figure 5.2: Probability of recall in Experiment 3 as a function of item type and introductory phrase type. 82

Figure F1: Recall by item type conditionalized on news phase verification response..... 137

List of Tables

Table 1.1: Examples of Positions on an Issue.....	11
Table 2.1: Political Orientation Ratings for Statements and Topics.....	30
Table 3.1: Probabilities of Change Recollection and Signal Detection Measures in Experiment 1.....	39
Table 3.2: Change Recollection Accuracy, Measured by d' for the Different Candidates as a Function of Subject's Self-Reported Political Orientation.....	40
Table 3.3: Proportion of Variance in Recall of Changed Positions Explained by Item Differences, General Memory Ability, and Change Recollection in Experiment 3.....	42
Table 3.4: Voting Results for the Two Candidates After Debate 1 and After the Recall Test by Subject Political Orientation, Experiment 1.....	43
Table 3.5: Subject Opinions About the Candidates as a Function of Their Self-Reported Political Orientation in Experiment 1.....	45
Table 4.1: Probabilities of Change Recollection and Signal Detection Measures in Experiment 2.....	59
Table 4.2: Proportion of Variance in Recall of Changed Positions Explained by Item Differences, General Memory Ability, and Change Recollection in Experiment 2.....	61
Table 4.3: Mean Proportion of Each Type of Response on the News Verification Task Organized by Item Type and News Statement Type in Experiment 2.....	62
Table 4.4: Probabilities of Change Recollection and Signal Detection Measures in Experiment 2, Broken Down by News Statement Accuracy.....	67
Table 4.5: Probabilities of Change Recollection and Signal Detection Measures in Experiment 2, Broken Down by News Verification Accuracy.....	70
Table 5.1: Probabilities of Change Recollection and Signal Detection Measures in Experiment 3.....	79
Table 5.2: Proportion of Variance in Recall of Changed Positions Explained by Item Differences, General Memory Ability, and Change Recollection in Experiment 3.....	81
Table 5.3: Probabilities of Change Recollection and Signal Detection Measures in Experiment 3, Broken Down by Introductory Phrase Type.....	83
Table E1: Recall as a Function of Self-Reported Political Orientation.....	120
Table E2: Recall as a Function of Performance on Political Expertise Quiz.....	121

Table F1: Voting Results for the Two Candidates After Debate 1 and Debate 2 by Subject Political Orientation, Experiment 2	131
Table F2: Subject Opinions About the Candidates as a Function of Their Self-Reported Political Orientation in Experiment 2	133
Table F3: Probabilities of Change Recollection and Signal Detection Measures in Experiment 2, Broken Down by News Verification Accuracy.....	138
Table G1: Introductory Phrases Used in Experiment 3	140
Table G2: Change Detection Rates as a Function of Item Type and Introductory Phrase Type Along With Signal Detection Measures in Pilot Experiment 3A	141
Table G3: Change Detection Rates as a Function of Item Type and Introductory Phrase Type Along With Signal Detection Measures in Pilot Experiment 3B	143

Acknowledgments

First and foremost I would like to thank Roddy Roediger for his support and guidance during my time at Washington University. I feel incredibly lucky to have had the opportunity to work with him. Thanks as well to Dave Balota, Kathleen McDermott, Ian Dobbins, and Jim Wertsch for serving on my committee and helping out on various projects over the years. I am also indebted to Larry Jacoby and Chris Wahlheim for guiding me towards this line of research and for showing me a different intellectual tradition.

The past and current members of the memory lab have been invaluable while I have been here. Thanks to all of the post-doc, graduate students and visitors who spent some time in the lab. Thanks to Angi Senne, Meghan McDoniel, and Allison Obenhaus for coding data, and to Christian Gordon for coding data and helping to create the materials.

Special thanks to my lab mate K. Andy DeSoto for his friendship and advice, for tolerating my interruptions, and for providing some comedy relief during lab meetings and talks. Many thanks to my family for all of their support over the years. Finally, thank you to Sara, for your patience, support, and encouragement during hard times (and helping me with my dissertation formatting). Marrying you is the best graduation present I could ask for.

This research was supported by the James S. McDonnell Foundation 21st Century Science Initiative grant, National Science Foundation Grant No. DGE- 1143954, a Dean's Dissertation Fellowship from Washington University, and an APF/COGDOP Graduate Research Scholarship.

ABSTRACT OF THE DISSERTATION

How do Voters Remember Flip-Flopping?

Memorial and Social Consequences of Change Recollection

by

Adam Lewis Putnam

Doctor of Philosophy in Psychology

Washington University in St. Louis, 2015

Professor Henry L. Roediger III, Chair

This dissertation presents 3 experiments that explore how people notice and remember a politician's change in position. Subjects read position statements made by politicians at two different debates; sometimes the politicians were consistent across debates, sometimes they changed positions, and sometimes they only addressed an issue at Debate 2. Subjects recalled the positions from Debate 2 and reported whether they thought the politician had changed positions on that issue. The results showed that changing positions made it more difficult for people to remember a politician's most recent position; however, recollecting that a change occurred eliminated that memory deficit. Experiment 1 explored how a voter's political orientation influenced their ability to remember a politician's position and whether recollecting change affected voter attitudes towards a politician. Experiment 2 showed that misleading information in the guise of a news report can affect later change recollection, but only if subjects are unable to verify the accuracy of the report. Finally, Experiment 3 showed that politicians can use specific language to make people believe a change occurred when there was actually none. The results of the experiments are discussed in relation to the recursive reminders framework, the misinformation paradigm, and the relationship between memory and attitudes.

Chapter 1: Introduction

During the 2012 presidential election Mitt Romney's public relations aid, Eric Fehrnstrom, was asked how the Romney campaign would shift between the Republican primary and the general election. Fehrnstrom's response was, "Everything changes. It's almost like an Etch A Sketch. You can kind of shake it up and restart all over again," (Shear, 2012). Romney's opponents immediately leaped at the opportunity to argue that Romney was inconsistent, disingenuous, and willing to say anything to win more votes. Indeed, after the election in November many pundits argued that Romney's habit of flip-flopping on important issues like taxes, health insurance, and abortion rights was the main reason he lost to Barack Obama (Trumble, 2012).

It is difficult to know for sure, however, how much Romney's flip-flopping actually affected voters. Did any of them decide to vote for someone else or to not vote as a result of Romney's changes in position? Would voters have even noticed or remembered that Romney had changed position if his opponents and the media had not been constantly accusing him of flip-flopping? Would voters have been more forgiving of Romney if he acknowledged that he had changed positions instead of insisting he had always believed the same thing?

Accusing opponents of flip-flopping has a long tradition in American politics. The term appeared in print as early as 1890 (New York Times, 1890), and today nearly every well-known politician has been accused of flip-flopping at one point. Flip-flopping accusations are intended to disparage a politician by implying that he or she lacks conviction, may not fulfill campaign promises, and is willing to change their platform to appeal to more voters (Safire, 1988). Sometimes such accusations are warranted, and other times they are not. For example, John Kerry was lampooned for changing his stance on an appropriations bill for the Iraq war when he

said “I actually did vote for the \$87 billion before I voted against it,” (e.g., Hummel, 2010), but he later explained that he had changed his vote because of Bush’s mishandling of the war. Politicians change their positions for a variety of reasons, and likewise, those changes can be interpreted in a variety of ways (e.g., was Obama’s reversal in opinion on same sex marriage the result of personal reflection and growth or an attempt to be in sync with popular opinion? See Bowers, 2012).

A more basic set of questions, however, is whether people notice when a politician changes position, whether they remember that change in position later, and whether remembering that change in position affects how the voter remembers what the politician currently believes. Recent research (Putnam, Wahlheim, & Jacoby, 2014) explored exactly those questions by using realistic political materials with a proactive interference design (Anderson & Neely, 1996). In the experiment subjects read position statements made by fictional politicians at two different debates. Sometimes the politicians changed positions and sometimes they did not. Putnam et al., showed that it was more difficult for people to recall what a politician currently believed if the politician had recently changed positions than if he had not recently changed (i.e., proactive interference occurred). However, this difficulty in recalling the changed position disappeared if subjects recalled that the politician had previously changed positions. In other words, recollecting the change in position eliminated proactive interference.

This dissertation reports three experiments that investigated more deeply how people notice and remember change in a political environment. Experiment 1 explored how political orientation affected change recollection and memory for a candidate’s positions and whether recollecting change influenced the subject’s attitudes towards the candidate. Experiment 2 examined how the

media influences what people remember about a politician's flip-flopping behavior by having subjects read potentially misleading information from a "news" organization before attempting to recall what the politician believes. Finally, Experiment 3 explored whether politicians can use specific language to draw attention to a change in position. Before reporting the studies I will briefly review how classic interference theory conceptualizes the effects of change on memory, describe a more nuanced view that argues for the role of reminders in reducing interference in memory (e.g., Wahlheim & Jacoby, 2013), and summarize other research exploring political judgments and memory.

1.1 Classic Interference Theory

The case of a flip-flopping politician is structurally similar to an interference experiment. In memory research interference generally refers to the idea that having two similar items in memory can make it difficult to recall either one of them (Anderson & Neely, 1996). In a typical experiment, subjects study a list of paired associates, such as the word pair "dog – spoon" (designated as an A – B pair). Then, subjects study a second list of paired associates where there are both some new pairs ("car – bottle", a C – D pair) and some pairs where the cue term is repeated from the first list but the target term is different ("dog – chair", an A – D pair). At test subjects are provided with a cue word and asked to recall the target from list 2 ("dog – ???"). Typically, recall for the changed pairs (the A – D items) is worse than for the control pairs (the C – D items). This deficit is known as proactive interference: learning the initial A–B pairs made it harder to learn or retain the A – D pairs. The reverse pattern, where learning a second list makes it difficult to recall what was initially learned, is known as retroactive interference.

Variations of this basic interference design have been used to understand how interference works and why it may contribute to forgetting (Anderson & Neely, 1996). Melton and Irwin's (1940)

two factor theory of interference is one prominent hypothesis that suggests that retroactive interference is caused by both response competition—when a single cue has two possible responses, those responses compete with one another at retrieval—and unlearning, where the process of learning a second list of paired associates causes the association between the shared cue and the first target to weaken. Proactive interference, in contrast, is caused by response competition alone. In comprehensive reviews, Anderson and Neely (1996) and Crowder (1976, chapter 8) provide other conceptions of interference theory and summarize hundreds of experiments that show that changing an item from one list to the next generally results in interference. The bulk of this research uses basic materials, like paired associates, but some exceptions do exist, such as Loftus’s work with the misinformation paradigm that uses pictures of real world events (e.g., Loftus, Miller, & Burns, 1978).

Based on the majority of classic interference research then, a politician (who paid attention during his cognitive psychology class) might expect that changing positions would make it more difficult for voters to remember his current position. As Crowder (1976) noted, however, interference theories became bulkier over time and were less able to specify a mechanism for forgetting that could parsimoniously explain different patterns of data. For example, there were some instances in the literature where change did not harm memory, but instead, enhanced it.

These are reviewed next.

1.2 When Change Does Not Lead to Interference

Several experiments have reported situations where one would expect interference to occur, but the results show that memory for a changed item is equal to or better than memory for the control items (Bruce & Weaver, 1973; E. Martin, 1968; Robbins & Bray, 1974). Barnes and Underwood (1959), for example, had subjects learn two lists of paired associates where the cues were

repeated across lists, but the targets were changed. In one condition the targets changed to an unrelated word, but in another condition the targets changed to a synonym (e.g., insane – crazy). After learning the second list subjects were presented with a cue and asked to recall any responses that came to mind, and to identify which list each response came from (i.e., the modified modified free recall procedure). As expected, when the targets were unrelated retroactive interference occurred; but when the targets were related, retroactive facilitation occurred. Barnes and Underwood suggested that when the two targets were related, subjects could use the first list response as a mediator for learning the second list response (e.g., dog – insane – crazy), and indeed, subjects reported using such mediators. Other research with a similar paradigm (R. B. Martin & Dean, 1964) reinforced the notion that subjects were consciously using the relationship between the two related targets to remember them.

More recently, however, Wahlheim and Jacoby (2013) proposed an alternative explanation for why changing an item could sometimes lead to facilitation, rather than interference. They suggested that interference should be eliminated (and sometimes facilitation could be observed) when subjects notice that a change has occurred (change detection) and are able to recollect that change at test (change recollection). Rather than using a mediational account, which required the items in memory to have a pre-existing association, Wahlheim and Jacoby used an interpretation stemming from the recursive reminders framework.

1.3 The Recursive Reminders Framework

A reminding experience is when one event reminds you of a previous event. Hintzman (2011) argued that reminders are often automatic and occur spontaneously throughout the day. For example, seeing a man in an office building who you have not met might remind you that you saw him yesterday at a meeting. Being reminded in this fashion might lead you to conclude that

the man is a new colleague. Hintzman (2004; 2010; 2011) and others (e.g., Benjamin & Tullis, 2010; Benjamin & Ross, 2010) have used reminders to explain a variety of memory phenomena, including spacing effects, encoding variability, and concept learning.

The role of reminders in memory is perhaps most easily explained in the context of a judgment of recency task, where subjects are asked to recall which of two events happened more recently. Tzeng and Cotton (1980) showed that subjects were much more accurate at recalling which of two words was presented more recently when the words were from the same category (e.g., *apple – peach*) than when they were from different categories (e.g., *knife – sofa*). Hintzman (e.g., 2011) argued that when the words are related, seeing the second word spontaneously reminds subjects of the first word. When this reminding occurs, it causes memory for the first word to become embedded in memory for the second word, along with the conscious experience of noticing the relationship. When making the judgment of recency, subjects who are able to think back to that earlier reminding experience should be able to accurately judge which word occurred more recently, because the temporal order of the words was encoded. Seeing *apple* and being reminded of *peach* is qualitatively different from seeing *peach* and being reminded of *apple* (Hintzman, 2010).

A similar mechanism may be at play in repetition effects, where multiple exposures of an item leads to better memory. Reminders are a spontaneous implicit retrieval of an earlier event, and as such may enhance memory (retrieval practice, see Roediger & Karpicke, 2006). Repetition effects then, may be due to seeing an item and being reminded that it was presented earlier. Asch (1969) described a fascinating experiment where subjects studied two lists of paired associates that were completely different except for one critical pair that was repeated in both lists. At test,

memory for the critical item was better than memory for the control items – but only for subjects who reported noticing the repetition in a post-experiment interview. Subjects likely noticed the repetition during the second list when seeing the critical item reminded them of seeing the item in the first list (i.e., they spontaneously retrieved the first list presentation). In a second experiment, warning subjects to look for the repetition caused nearly all of them to notice the repeated pair, which in turn, led to accurate recall for that item at the test. In other words, the instructions to look for the repetition encouraged subjects to use reminders.

Similarly, reminders have also been used to explain spacing effects in memory (Appleton-Knapp, Bjork, & Wickens, 2005; Wahlheim, Maddox, & Jacoby, 2014). For example, Appleton-Knapp et al. (2005) had subjects study a booklet of advertisements where individual ads were repeated throughout the book at different intervals, and were either identical or slightly varied between presentations. After reading the booklet subjects were asked to recall details about the ads. The results supported the hypothesis that spacing effects may be due to subjects covertly retrieving a first presentation when studying a second presentation (an idea called study-phase retrieval). Spacing repetitions farther apart or increasing the variability of the repetitions made spontaneous reminding less likely to occur, but when those reminders did occur they were more effective at enhancing future recall than if the reminders occurred after a shorter interval (see also Benjamin & Tullis, 2010). Thus, both repetition effects and spacing effects may be due to reminders.

Critically for the current project, reminders may also be important for noticing change (Jacoby, Wahlheim, & Kelly, in press; Jacoby, Wahlheim, & Yonelinas, 2013; Wahlheim & Jacoby, 2013). Just as noticing a repetition involves looking back to an earlier event, noticing that

something has changed requires looking back to the original event and noticing the discrepancy. Detecting change in this fashion may be beneficial to memory. In the misinformation paradigm, for example, subjects who notice inconsistencies between the original event and the misinformation often do not show a memory deficit (Loftus, 1979; Tousignant, Hall, & Loftus, 1986). This may occur because detecting the change requires implicitly retrieving the original event, which, due to both an implicit repetition and implicit retrieval practice, should make the original event more accessible in memory and allow subjects to reject the misinformation. But how could detecting change reduce interference in a proactive interference design where subjects are asked to recall the more recent item? Wahlheim and Jacoby (2013) suggested that it is not enough for subjects to *detect* a change—they must also *recollect* that the change occurred at test.

1.4 Prior Research: Detecting and Recollecting Change can Reduce Proactive Interference

Wahlheim and Jacoby (2013) reported three experiments showing that detecting and later recollecting change can reduce proactive interference. Subjects studied two lists of paired associates. Some pairs appeared in both lists (repetition items, e.g., “apple-core”), some pairs were only presented during the second list (control items, e.g., “lamb-wool”), and some pairs had a target that changed between the lists (changed items, e.g., “knee-bone” became “knee-bend”). While studying the second list subjects pushed a button when they noticed that an item had changed from the first list (a measure of *change detection*). At test subjects were asked to recall the second list target in response to a cue term (“knee - ???”) after which they were asked if any other responses had come to mind (a measure of *change recollection*).

Overall, the results showed that recall for the changed items was equal to the control items, i.e., that there was no proactive interference. However, recall for the changed items depended on

whether change was recollected. When subjects failed to recollect that a change had occurred, recall for the changed items was worse than it was for the control items (proactive interference); In contrast, when subjects did recollect change, recall for the changed items was better than it was for the control items (proactive facilitation). In other words, recall for the changed items was a mixture of proactive interference and proactive facilitation.

Critically, Wahlheim and Jacoby (2013) showed that the improvement to memory for the changed items depended on both change detection and change recollection. When subjects detected a change but failed to recollect that change at the test, then massive interference occurred (recall for those items was much worse than if subjects neither detected nor recollected change). Wahlheim and Jacoby argued that detecting change resulted in a strengthening of the first list response. At the test, failing to recollect the change meant that the increased accessibility of the first list response led to interference. Wahlheim and Jacoby (2013) suggested that detecting change (being reminded of the corresponding first item when seeing the second item) led to memory for both items, along with the experience of detecting the change, to become encoded together in memory. At test, recollecting that a change had occurred allowed subjects to access both responses, which in turn lead to accurate recall. Thus, recollecting that a change had occurred counteracted the increased accessibility of the first list response due to change detection.

One concern with the results of Wahlheim and Jacoby (2013) and other papers (Jacoby & Wahlheim, 2013; Jacoby et al., 2013) is the possibility of item selection effects that occur when relying on conditional analyses. Some items may be easier to remember than others; if subjects have a strong memory for an item, they may also be able to remember that the item changed.

Thus, the reduction in interference could be an artifact of the conditionalized analyses collecting the items that are easier to remember rather than change recollection driving the boost to memory. Indeed, Wahlheim and Jacoby reported hierarchical regression analyses showing that item effects contributed to recall. However, the models also showed that change recollection had a unique contribution to recall after controlling for any individual item effects. Furthermore, several studies now (Jacoby, 1974; Jacoby & Wahlheim, 2013; Jacoby et al., in press; Wahlheim et al., 2014) have shown that reminders can be brought under task control. In these experiments subjects study a set of items; one group is encouraged to look back to a previous event whereas the other group is not. The group that is given looking back instructions shows enhanced change detection, repetition detection, recency judgment, list discrimination, or recall compared to the control group. Thus, two groups of subjects study an identical set of items, but the looking back instructions consistently lead to better memory, suggesting that reminders are driving the boost in performance rather than item selection effects.

In sum, Wahlheim and Jacoby (2013) showed that detecting and recollecting change can reduce proactive interference and can sometimes lead to proactive facilitation. In the introduction to their paper, Wahlheim and Jacoby used the example of a flip-flopping politician to highlight the importance of noticing and remembering change. Indeed, if a politician changed his position from being pro-life to being pro-choice, it may be difficult to remember what the politician currently believed—unless, of course, you remembered that the politician had previously changed positions. Putnam et al. (2014) conducted an experiment to see if noticing and remembering change would transfer to a fictionalized political environment.

1.5 Prior Research: Recollecting Change With Political Materials

Putnam et al. (2014) used the same experimental design as Wahlheim and Jacoby (2013), but framed the experiment as a political campaign. Instead of studying word pairs, subjects read about two fictional politicians (John Baker and Mike Shipman) who expressed their views on current political issues at two different debates. An example of the materials is presented in Table 1. The materials were constructed so that the candidates could change position within a side (liberal or conservative) because pilot work showed that subjects were at ceiling when detecting changes that crossed sides. The first half of each statement served as a cue (along with candidate information) and the second half of each statement served as the target.

Table 1.1: Examples of Positions on an Issue

Platform and Extremeness	Issue: Same Sex Marriage
Liberal	
Strong	Partners of the same sex should be recognized through <i>marriage</i> .
Weak	Partners of the same sex should be recognized through <i>civil unions</i> .
Conservative	
Strong	Traditional marriage should be <i>protected through a constitutional amendment</i> .
Weak	Traditional marriage should be <i>an issue for individual state legislatures</i> .

Note: The positions above appeared without italics in Debates 1 and 2. At test, the first portion of each sentence served as a cue for recall, and subjects' task was to recall the italicized portion. When changing positions candidates changed within a side (liberal or conservative).

Across the two debates the candidates sometimes repeated themselves (repetition items), sometimes addressed an issue only at the second debate (control items) and sometimes changed

positions across debates (changed items). At the test, subjects were cued with the candidate's name and photo along with the topic and first half of the statement. Subjects tried to recall each candidate's Debate 2 position by recalling the second half of the statement, and then indicated whether the candidate had changed position on that topic by pressing a button (a simpler, but still reliable, measure of change recollection; see Jacoby et al., 2013). A measure of change detection was omitted from all the experiments except for one because previous work showed that change recollection rarely occurred without previous detection (Jacoby et al. 2013).

Across three experiments, Putnam et al. (2014) replicated the findings of Wahlheim and Jacoby (2013) in showing that detecting and recollecting change decreased proactive interference.

Overall there was a proactive interference effect – recall for the changed items was worse than for the control items. But examining recall for the changed items as a function of whether change was recollected revealed a more nuanced pattern. In Experiment 1 (see the left panel of Figure 1), for example, recall for the changed items when change was recollected was equivalent to recall for the control items. If change was not recollected, however, recall for the change items was worse than for the control items. Putnam et al. argued that recollecting change meant subjects had previously detected change, and that such detection meant both the first and second positions were encoded together in memory. Having access to this information led to accurate recall, thus eliminating interference.

In Experiment 1, subjects were warned that the candidates might change positions, and that noticing those changes might help their memory for what the candidates believed. In Experiment 2 one group of subjects was warned about the changes as in Experiment 1 (the informed group) whereas a second group of subjects was not warned about the changes (the uninformed group).

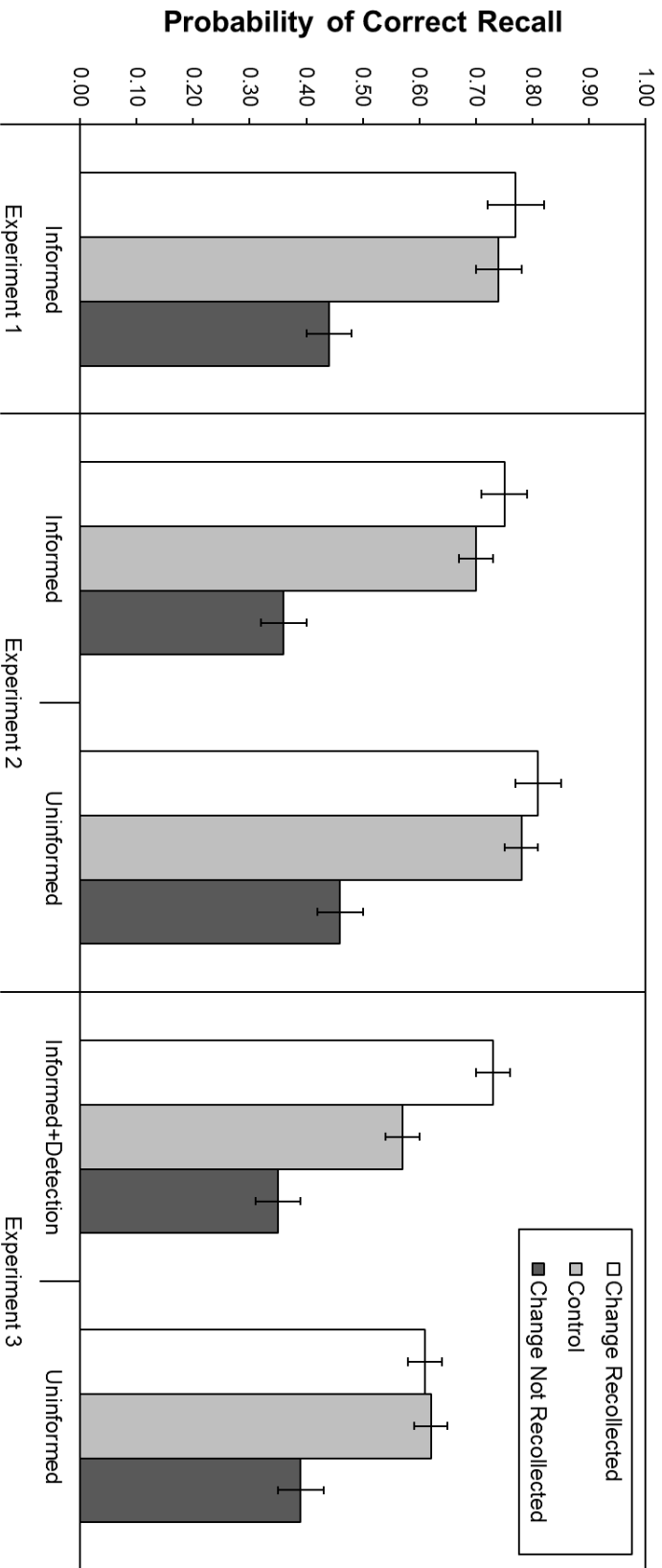


Figure 1.1: Results of Putnam et al. (2014), Experiments 1 - 3. Cued recall for changed positions conditionalized on change recollection compared to recall for control positions.

The results (see middle panel of Figure 1) showed that pre-warning subjects did not influence recall. Putnam et al. (2014) suggested that the rich nature of the materials allowed subjects to detect and recollect change even if they were not specifically instructed to do so.

Finally, in Experiment 3 Putnam et al. (2014) used a stronger manipulation to encourage change recollection. During Debate 2 one group of subjects (the informed + detection group) was asked to push a button when they noticed a candidate changing positions (a measure of change detection) whereas the other group of subjects (the uninformed group) was not warned about change. The results (see the right panel of Figure 1) indicated that while the uninformed group mirrored the pattern from earlier experiments, the informed + detection group showed proactive facilitation: recall for the changed items when change was recollected was better than it was for the control items. Furthermore, Experiment 3 also showed that both change detection and change recollection were required to reduce interference (see Figure 2). Detecting and recollecting change led to facilitation, whereas detecting change but failing to recollect the change later led to massive interference. Putnam et al. (2014) suggested that the inclusion of the change detection task increased the accuracy of the change recollection measure, which is important for finding proactive facilitation. Subjects may guess about change recollection; thus, the conditionalized recall data would underestimate the true accuracy in recall when change is recollected. Any increase in change recollection accuracy would thus provide a better estimate for recall when change is remembered, and may be more likely to show proactive facilitation.

In sum, the results of Putnam et al. (2014) provided further support for the hypothesis that detecting and recollecting change can reduce proactive interference. Importantly, the experiments used realistic political materials that were complex and may have interacted with a

subject's pre-existing knowledge of politics. The experiments in this dissertation addressed new questions by using a design similar to that used in Putnam et al. For example, Experiment 1 examined whether political orientation (both the subject's and the candidate's) influenced recall and change recollection, and began to explore how recollecting change shaped subject attitudes

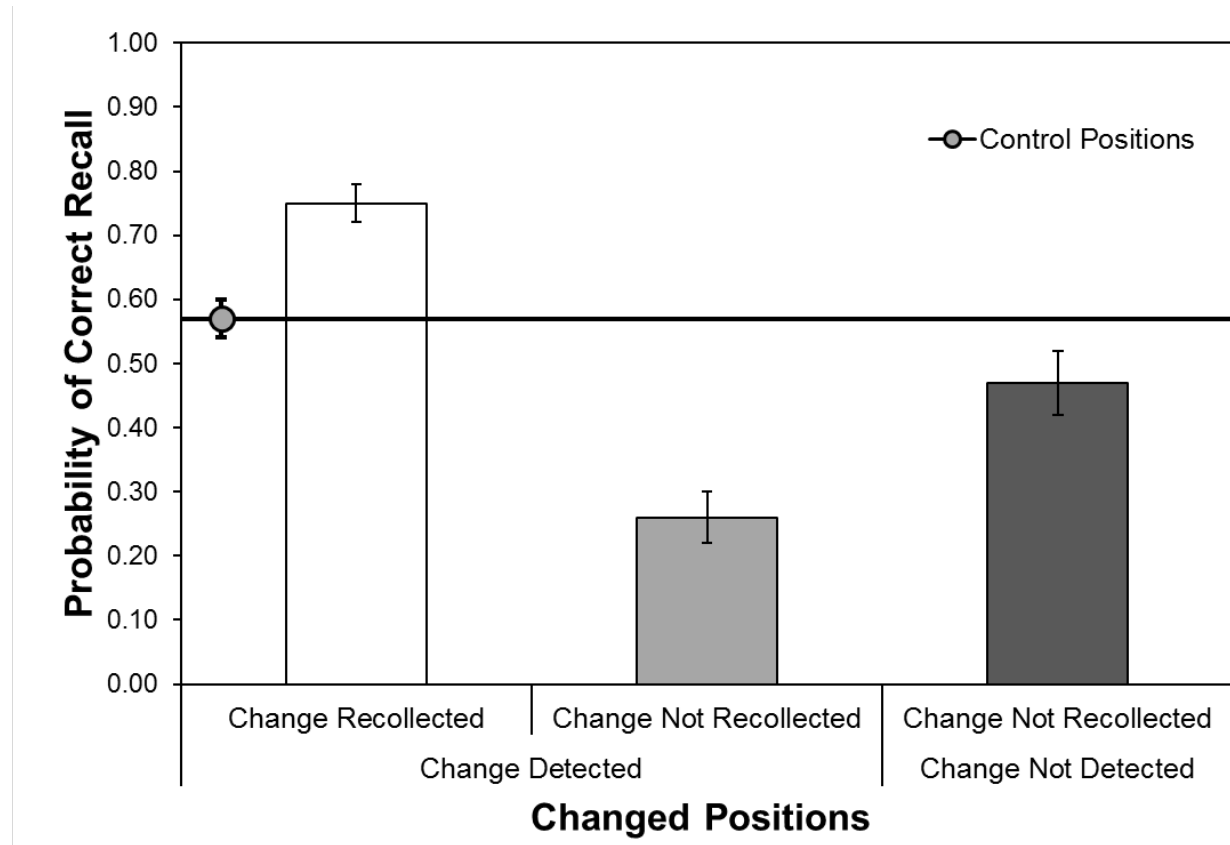


Figure 1.2: Results from Experiment 3 of Putnam et al. (2014). Cued recall for changed positions, conditionalized on change detection and change recollection in the informed + detection group. The black line indicates recall for the control positions, with standard error.

towards the candidates. Before describing the dissertation experiments in detail I will first briefly review some political science and social psychology literature related to how people remember and react to changes in political positions.

1.6 How Do People Interpret a Politician's Change in Position?

Thus far, the discussion has focused on how a politician changing positions affects memory for what the politician believes. But of course when politicians change positions in the real world it changes how the public perceives them. Politicians do occasionally change their position on a current issue. Sometimes there is a strong reason for it (for example, Republican congressman Rob Portman decided to support gay marriage after his son came out of the closet), and sometimes it is simply to move to a more popular position. Any change in position, however, raises the risk of being accused of flip-flopping, even if such accusations mischaracterize what really happened (Croco & Gartner, 2014). The political science literature suggests that changing positions too often or too quickly can have negative consequences, but that changing positions to a more popular stance can sometimes be advantageous.

The obvious downside to changing positions is being accused of flip-flopping (or waffling or being wishy-washy). Such accusations can have two negative consequences. First, someone who flip-flops can become to seem unreliable or unpredictable, which can lead voters to wonder what position a candidate will hold in the future (e.g., Bernhardt & Ingberman, 1985). Second, inconsistency on its own is seen as a negative personality trait. Rosenberg, Nelson, and Vivekananthan (1968), for example, showed that subjects rated the traits *unreliable* and *wavering* about as negatively as *dishonest*, *wasteful*, and *irresponsible*. Thus, flip-flopping may lead voters to question whether a candidate is trustworthy, and whether he or she will follow through with campaign promises (e.g., Hummel, 2010; Tavits, 2007). Given that people value consistency and reliability (e.g., Cialdini, 1993), it is not surprising that flip-flopping can have

negative consequences. Indeed, the results of the 2004 and 2012 presidential elections showed what flip-flopping accusations did to John Kerry and Mitt Romney.

In contrast, changing positions might be a smart strategy in some situations. Some political commentators have suggested that changing a policy position is fine, as long as the candidate is open about the change, is doing it in response to new information (rather than just pandering for votes), and is not breaking a promise (Geraghty, 2008; Safire, 1988; Dickerson, 2008). This view is corroborated by some empirical evidence that suggests that voter opinions of candidates are driven largely by the candidate's current position and not whether he or she has changed positions in the past (Croco & Gartner, 2014; K. D. McCaul, Ployhart, Hinsz, & McCaul, 1995). Voters tend to prefer candidates who currently share their own views and do not appear to take past changes into account. Voters also appear to understand that politicians will shift their platforms between a primary and general election (Hummel, 2010), and appear to be more forgiving of policy changes by the President when a reason is provided for the change (Levendusky & Horowitz, 2012). Thus, changing positions may be an advantageous strategy if the change results in the candidate's position aligning with more voters.

Interestingly, despite the ambiguous outcomes of changing positions, both politicians and voters *believe* that the general public cares a great deal about consistency. In particular, a survey study showed that both college students and state legislators believed that "taking a consistent stand was more important to the average voter than whether a politician agreed with his or her constituents," (McCaul et al., 1995, p. 298). This intuitive assumption about how voters perceive position changes may partially explain why politicians are so fond of accusing each other of flip-flopping.

1.7 How Partisanship Affects Judgments and Memory

To complicate things further, research suggests that people's pre-existing attitudes can influence their judgments, decision-making, and memory. For instance, political orientation (or partisanship) may influence how people interpret a politician's behavior. Barden, Rucker, Petty, and Rios (2014) had subjects read vignettes about politicians who said one thing but did another, and then had the subjects describe the politicians. Subjects were more likely to use the term "hypocrite" to describe a politician when the politician was from an opposing political party than from their own. Similarly, Westen, Blagov, Harenski, Kilts, and Hamann (2006) presented Democratic and Republican subjects with stories about George W. Bush or John Kerry changing his mind on an important issue and asked the subjects to provide an inconsistency rating for each politician. The results showed a crossover interaction based on partisanship: Democrats were more critical of George Bush whereas Republicans were more critical of John Kerry. Westen et al. suggested that this pattern was indicative of motivated reasoning, where subjects implicitly engaged in emotionally-based reasoning to decrease negative affect. Subjects had a harder time accepting their preferred candidate acting in an inconsistent manner, and thus were more likely to find a rational explanation for the change in position.

In another interesting study, Nyhan and Reifler (2010) had subjects read news articles that contained misconceptions about current events (e.g., that Iraq had weapons of mass destruction before the U.S. invaded) and had subjects rate how much they agreed with the claims in the articles. As expected, Nyhan and Reifler found that political orientation affected belief in the misconceptions—subjects were more likely to believe information that was consistent with their worldviews. More interesting, however, is that in some conditions a correction was provided (e.g., that there were no WMDs in Iraq before the U.S. invasion). The corrections not only failed

to reduce misperceptions in some groups, but sometimes led to a backfire effect, where certain groups were *more* likely to believe in the misperception after reading a correction (i.e., in the example above, some conservatives were more likely to believe that Iraq had WMDs *after* reading a correction). Again, motivated reasoning was used to explain the results: subjects were more likely to continue to believe incorrect facts that were congruent with their own political orientation.

In addition to affecting different types of judgments, political orientation may also affect memory for political information. The congeniality hypothesis (Eagly, Chen, Chaiken, Shaw-Barnes, 1999; Eagly, Kulesa, Chen, & Chaiken, 2001; Levine & Murphy, 1943) posits that people may be better at remembering information that is consistent with their attitudes compared to information that is inconsistent with their attitudes. The theory suggests that people typically avoid information that challenges their worldview, and if they do encounter such information, either ignore it, distort it, or fail to store it adequately in memory. In one classic experiment, Pro-Communist and Anti-Communist subjects were asked to remember what they read in a Pro-Soviet Union or Anti-Soviet Union essay (Levine & Murphy, 1943). As expected, subjects recalled more information from the essay that was congruent with their beliefs. Although the congeniality hypothesis has traditionally received strong support, a recent meta-analysis (Eagly et al., 2001) concluded that the evidence for the congeniality hypothesis is mixed, and that there are a variety of ways in which people can react to uncongenial information, some of which lead to a deficit in encoding or later retrieval, and some of which do not.

There are still situations, however, where political orientation can affect memory. A recent false memory study had thousands of people look at photographs of recent political events and report

whether they remembered the events happening (Frenda, Knowles, Saletan, & Loftus, 2013). Critically, in each group of photos there was one digitally altered photo that portrayed an event that never occurred. In some conditions the photo showed President Bush vacationing with a celebrity instead of visiting the aftermath of Hurricane Katrina and in other conditions the photo showed President Obama shaking hands with the president of Iraq. Overall, around 50% of subjects claimed to remember the fabricated events – but conservatives were more likely to falsely remember Obama shaking hands with the president of Iraq, and liberals were more likely to falsely remember Bush vacationing after Katrina. The authors concluded that false memories were more easily formed when they were consistent with a subject’s pre-existing attitudes because the fabricated events seemed more familiar.

In sum, a person’s political orientation can have a measurable impact both on how they make political judgments about a politician’s behavior, and how they remember political information. In general, people are more likely to be critical of a politician from a different political party than their own (and more forgiving of a politician from their own political party) and may find it easier to remember information that is consistent with their existing political beliefs.

1.8 Overview of The Experiments

Three experiments explored change recollection and memory in politics. All three used the design used in Putnam et al. (2014) with a revised set of political materials (described in the norming study below). Subjects were recruited from Amazon’s Mechanical Turk website to ensure political diversity. A brief overview of each experiment is provided below, with more detailed hypotheses in the introduction to each experiment.

1.9 Experiment 1: The Intersection of Attitudes With Change Recollection

Experiment 1 had two main goals. The first was to examine whether partisanship affected memory and change recollection. In particular, the congeniality hypothesis (e.g., Eagly et al., 2001) predicted that subjects should be better at recalling the positions of a candidate from their own political party. Extending that hypothesis to change recollection suggested that subjects would be better at recollecting change when a candidate had similar political views to the subject.

The second goal was to examine whether recollecting change influenced subject attitudes towards the candidates. Experiment 1 added two opportunities for subjects to rate the politicians on several measures (e.g., How trustworthy is this candidate? How likely is this candidate to change positions in the future?). In general, there were two hypotheses. The first was that subjects would support their own candidates, regardless of how much they remembered each candidate changing position. This position would be supported if subjects provided positive ratings to the candidate from their own party, and lower ratings to the candidate from the opposite party. A second hypothesis was that voters might disapprove of politicians who changed positions too often. This position would be supported if the subjects provided low ratings overall for both candidates, or if there was a negative correlation between change recollection and the judgments. Such a correlation would indicate that recollecting change more often was associated with a more negative attitudes towards the candidates.

1.10 Experiment 2: Does the News Create a Misinformation Effect?

Experiment 2 introduced a misinformation element: after Debate 2 but before the final test subjects read statements from a “news” organization that sometimes reported correct information about a candidate (“Mike Shipman was consistent in his view on Same Sex Marriage” when he did not change positions) and sometimes reported misleading information (“John Baker changed position on Gun Control” when he did not change positions). Subjects were asked to verify the truthfulness of each statement. I predicted that for the changed items, accurately verifying a news statement would lead to accurate recall and change recollection at the final test. A correct response during the news phase would indicate that subjects were able to recollect the presence or absence of change (they accessed the recursive trace) and should subsequently lead to better recall and change recollection on the final test for the changed items. In contrast, subjects who endorsed a misleading statement (e.g., agreeing that a candidate had changed positions when he did not) should lead to poor change recollection accuracy during the final test, which in turn should lead to poor recall for the changed items.

1.11 Experiment 3: Can Politicians Use Language to Draw Attention to a Change in Position?

Finally, Experiment 3 explored whether politicians could use specific language to increase the odds of someone remembering a change in position. During the second debate, short phrases were included before each of the position statements. Some of them were neutral (e.g., I’m happy to talk about that issue today...) whereas some of them encouraged subjects to look back to the first debate (e.g., In contrast to my previous position, I now believe...). Previous research has shown that different instructions can influence the degree to which subjects look back to notice repetitions (Wahlheim et al., 2014) or change (Jacoby et al., in press), so the *looking back* statements here should encourage subjects to notice inconsistencies, which in turn should lead to

more accurate change recollection and higher recall for the changed items. However, instructions to look back have always been manipulated between-subjects; here the looking back cue is manipulated within-subjects, appears in a mixed-list format, and is integrated as part of the materials. Thus, an alternative hypothesis, drawn from Loftus's (1979) work exploring blatant change, is that once subjects notice a single change, they will continue to look for other inconsistencies.

Chapter 2: Norming Study

The goal of the norming study was to calibrate a revised set of political materials. The materials in Putnam et al. (2014) suffered from two weaknesses. First, subjects sometimes confused two different topics during recall. For example, the topics *Gun Control* and *Concealed Weapons* resulted in many cross-topic contaminations where subjects recalled a statement from Gun Control in response to the Concealed Weapons cue. Such overlap may cause subjects to be reminded of a similar topic presented in Debate 2, rather than thinking back to the appropriate response in Debate 1. Second, some of the statements within a particular topic were quite similar, raising the concern that they would be interpreted as a change in wording, rather than a true change in position. For the revised materials, I removed topics that showed cross-topic contamination in Putnam et al. (2014), and re-wrote the position statements to make the differences more distinct.

In the norming study, subjects were presented with statements and asked to rate how conservative or liberal they thought each statement was. After a group of 30 - 35 subjects rated the materials, I checked to see whether the materials met the criteria described below. Materials that did not meet the criteria were revised and tested again. Once an item met the criteria in two consecutive waves, it was considered complete and not tested again. A second goal of the norming study was to evaluate methods or assessing the political orientation and political expertise of the subjects.

2.1 Method

2.1.1 Subjects

351 subjects were recruited from Amazon Mechanical Turk, with participation limited to IP addresses originating in the United States. The data from 24 subjects were excluded because they

failed the political expertise quiz or reported low levels of political expertise, leaving a sample of 327 subjects (168 female). The average age of the final sample was 36.90 (SD = 12.22, range = 16 - 68), with 47 different states represented. Subjects were asked to report their level of education (13% reported high school education or less, 37% reported some college education or an associates degree, and 50% reported a bachelor's degree or graduate degree) and twelve subjects reported being born outside the U.S. (their data did not differ significantly from the rest of the sample). Subjects were paid between \$0.25 and \$3.00 for participating, depending on the length of the study (pro-rated at \$3.00 per hour). The Washington University in St. Louis Institutional Review Board approved the norming study and the remaining experiments.

2.1.2 Materials

Political Orientation Surveys. Subjects completed both a self-report and objective measure of their political orientation. For the self-report measure, subjects were asked “Which category best represents your political views?” and selected one of the following responses: Liberal Democrat, Average Democrat, Moderate Democrat, Independent, Moderate Republican, Average Republican, and Conservative Republican.

For the objective measure, subjects rated their agreement with 12 statements about current political issues (e.g. “There need to be stricter laws and regulations to protect the environment” - see Appendix D for list of all statements). The 12 statements were taken from a Pew Research center survey that was administered to a large, representative sample of U.S. citizens, and were the items that most accurately predicted each respondent's political orientation on the seven-point scale described above. This approach to objectively measuring political orientation was used by Zell and Bernstein (2014) who found that young adults often reported being more conservative than an objective measure would suggest.

The population in the norming study replicated the pattern found in Zell and Bernstein (2014). By the self-report measure, the sample leaned left: 53% Democrat, 26% Independent, and 20% Republican. By the objective measure, however, the sample was even more liberal: 70% Democrat, 7% Independent, and 22% Republican, with 56% of respondents falling into the “Liberal Democrat” category. Thus, subjects in the sample were more likely liberal than they believed themselves to be.

Political Expertise Survey. Subjects also completed a self-report and objective measure of their political expertise. Subjects who scored below a set criteria were omitted from data analyses in the norming study and the other experiments. For the self-report measure, subjects were asked “How much do you think you know about American Politics?” and used a slider to report their response. The slider response was converted to a 100 point scale, with 0 representing low knowledge and 100 representing high knowledge; subjects who reported a score less than 10 were excluded. The objective measure consisted of an eight-item political expertise quiz (adapted from Lambert et al. 2010, see Appendix C). Each item was a multiple-choice question, with questions such as “Before his election as Vice President, which state did Joe Biden represent in his role as U.S. senator?” Subjects who answered fewer than three questions correctly were excluded from all analyses.

The average response for the subjective measure of political expertise was 60.02 ($SD = 23.80$) indicating that subjects felt they had a moderate knowledge of politics. For the 8 item political expertise quiz, 17% scored a 3 or 4, 32% scored a 5 or 6 and 50% scored a 7 or 8, with a mean score of 6.26 ($SD = 1.55$). Thus, from a more objective perspective the pool had a fairly good understanding of American politics. Twenty-four subjects were omitted from analysis, 13 for

reporting less than a 10 on the self-report question, and 14 for scoring less than a 3 on the quiz (some failed both).

Political Materials. The materials were based on those used in Putnam et al. (2014) and consisted of 36 current political issues, with 4 position statements associated with each topic. The four positions statements were composed of a strong liberal position, a weak liberal position, a weak conservative position, and a strong conservative position. Within a side (liberal or conservative) the first part of the sentence was identical, which allowed the first half of the sentence to serve as a retrieval cue during the main experiments. The sentence structure allowed candidates to change from a strong position to a weak position or vice-versa. Table 1 shows an example of an individual topic, and Appendix A displays the final set of political materials developed during the norming study.

2.1.3 Procedure

Subjects were recruited from Amazon's Mechanical Turk Website (Mason & Suri, 2012), and completed the experiment on their own computer (presented in Adobe Flash, Weinstein, 2012). Upon starting the experiment subjects read an informed consent letter and reported basic demographic information (age, gender, current state, whether they were born in the United States, and education).

Subjects then read instructions for the study. They were informed that they would be presented with a current political topic and four related position statements. Their job was to rate how conservative or liberal they thought each statement was by using a slider anchored with "liberal" on the left and "conservative" on the right.

There were two versions of the norming study. In both versions the topics appeared in a random order. In the first version subjects saw a topic and a statement, and after 1.5 seconds the slider appeared below. Subjects made their rating, and then pressed “enter” or clicked the submit button to finalize their response. If subjects did not respond within 50 seconds a message appeared urging them to respond, and if they still did not respond after 10 additional seconds then a response was not recorded. After making their response a blank screen appeared for 500 ms and then the next statement appeared. After rating the four statements associated with a topic, subjects moved on to the next topic.

In the second version, subjects rated all four statements associated with a topic at once, rather than seeing them one by one. In this version, a topic appeared along with the four position statements. After 8 seconds, a slider appeared for each statement, and subjects made their ratings while seeing all four statements. After making a rating for each statement the “submit” button appeared that allowed subjects to save their responses. A warning message appeared after 3 minutes and 50 seconds, and if subjects did not submit their response within 4 minutes, their responses were not recorded.

After rating all of the topics, subjects completed the political expertise and political orientation measures and then were taken to a slide that contained a debriefing letter along with instructions for receiving payment.

The experiment was completed in waves. After 30 to 36 subjects rated the materials (the n varied due to the number of subjects who failed the expertise measures), I checked the ratings to see if they met the criteria described below. Any items that failed were revised and tested in the next wave. Each wave used one of the two versions of the experiment. The first version mirrored how

subjects read the statements in the main experiments (seeing one topic at a time) whereas the second version allowed subjects to read all four statements before making their rating, perhaps providing a more accurate estimate. Once a topic met the criteria in two consecutive waves, it was considered complete and omitted from future waves.

2.1.4 Results

Subjects used a slider to rate the materials which was converted to a 100 point scale, with zero representing *liberal* and 100 representing conservative. Each item was revised until it met two criteria: 1) the difference between the strong and weak version of each statement was at least seven and 2) the strong liberal position had a score of less than 25 while the strong conservative position had a score of greater than 75. These criteria ensured that the strong and weak positions were distinct from one another and that the strong positions were indeed more extreme. Table 2 shows the average ratings for each statement in its final wave along with the difference in the scores between the strong and weak position on each issue. The actual items are reported in Appendix A. Subject ratings were unaffected by their political orientation or their political expertise.

Table 2.1: Political Orientation Ratings for Statements and Topics

	Liberal			Conservative		
	Strong	Weak	Difference	Weak	Strong	Difference
Abortion	10.75	34.94	-24.19	71.94	89.00	17.06
Affirmative Action	22.06	42.22	-20.17	61.14	75.97	14.83
Affordable Care Act	23.50	34.50	-11.00	76.28	86.28	10.00
Alternative Energy	16.06	41.69	-25.63	44.00	77.89	33.89
American Patriotism	23.27	57.00	-33.73	72.82	84.07	11.25
Border Security	15.81	34.25	-18.44	57.17	88.00	30.83
Contraception	12.74	23.51	-10.77	63.09	82.34	19.26
Death Penalty	21.19	36.97	-15.78	63.28	77.72	14.44
Don't Ask Don't Tell	16.11	37.83	-21.72	72.92	90.78	17.86
Flag Desecration	21.00	41.70	-20.70	74.52	82.45	7.94
Global Warming	21.94	43.75	-21.81	70.69	84.08	13.39
Government Debt	26.25	47.43	-21.18	60.76	83.21	22.45
Gun Control	19.28	32.84	-13.56	61.94	78.47	16.53
Health Care	9.69	21.89	-12.20	74.43	82.91	8.49
Immigration Benefits	17.96	31.93	-13.96	79.39	90.57	11.18
Labor Unions	11.94	19.00	-7.06	72.03	89.21	17.18
Legal. of Marijuana	14.67	28.50	-13.83	69.53	88.39	18.86
Mitt Romney	15.33	50.97	-35.64	70.00	85.81	15.81
Nuclear Weapons	17.47	44.00	-26.53	71.56	80.86	9.31
Occupy Wall Street	19.82	27.18	-7.36	66.68	88.18	21.50
Off-shore Oil Drilling	15.44	29.50	-14.06	67.28	80.94	13.67
Paying for College	13.07	25.32	-12.25	76.25	86.00	9.75
President Obama	18.39	36.06	-17.67	69.61	88.81	19.19
Role of Government	20.15	37.71	-17.56	48.26	76.53	28.26
Same Sex Marriage	13.75	32.03	-18.28	66.72	78.92	12.19
School Safety	17.94	38.16	-20.22	65.97	87.53	21.56
Public Schools	15.97	32.37	-16.40	69.29	79.71	10.43
Separ. of Church & State	18.92	31.58	-12.67	66.03	85.86	19.83
Sex Education	17.33	31.42	-14.08	64.14	89.47	25.33
Social Security	23.23	37.80	-14.57	59.29	77.00	17.71
Stem Cell Research	17.42	47.92	-30.50	71.36	90.44	19.08
Teaching Creationism	18.71	48.51	-29.80	73.09	86.60	13.51
Trickle Down Econ.	13.76	32.91	-19.15	63.18	85.09	21.91
United Nations	23.36	53.58	-30.22	55.19	76.28	21.08
War on Terror	18.21	44.41	-26.21	58.44	85.47	27.03
Welfare	14.73	23.91	-9.18	67.64	88.55	20.91

Chapter 3: Experiment 1

Experiment 1 had two main goals. The first goal was to evaluate whether partisanship affected recall or change recollection. As described earlier, one prediction, inspired by the congeniality hypothesis (Eagly et al., 2001; Levine & Murphy, 1943) is that people will find it easier to remember information that is consistent with their attitudes. However, evidence for the congeniality hypotheses is mixed (Eagly, Chen, Chaiken, & Shaw-Barnes, 1999; Eagly et al., 2001). An alternative possibility is that recall may be unaffected by partisanship (subjects will recall the Democratic and Republican candidate's positions equally well) – some researchers have suggested that critiquing information that is inconsistent with one's own view is an effective encoding strategy. Both predictions can be extended from overall recall to change recollection. Subjects may be better at noticing and remembering changes in a candidate's positions when the candidate is from their own political party (e.g., an extreme conservative might not realize that supporting sex same marriage might be different from supporting same sex civil unions). In contrast, as reviewed above, subjects are often more critical of politicians from a different political party (e.g., Nyhan & Reifler, 2010), and thus subjects might be better at recollecting change when the candidate is from a different political party. Of course, a final possibility is that change recollection is unaffected by partisanship.

The second goal of Experiment 1 is to determine whether noticing and remembering change influences how voters perceive the candidates. If subjects remember a politician changing positions quite often, would they be less willing to vote for him? Would they trust or like him less? On one hand, voters may distrust a candidate who they perceive as changing positions—it is hard to trust someone if he or she is inconsistent. On the other hand, voters may prefer the candidate from their own party regardless of how often they changed positions. A third

possibility is that voters may be drawn to candidates who change their positions to be more in line with the voter's own views - however, this would be difficult to explore in the current paradigm as candidates never change sides entirely.

Experiment 1 used the design from Putnam et al. (2014)¹. Subjects were introduced to two political candidates, a Democrat and a Republican, and were told they would be reading excerpts from two debates where the candidates expressed their views on current political issues. Across the two debates the candidates sometimes held the same positions, sometimes changed positions, and sometimes address an issue only at the second debate. Subjects then attempted to recall each candidate's position from Debate 2 and indicated whether the candidate had changed positions. After both Debate 1 and the recall test subjects were asked which candidate they would vote for and made ratings about each candidate's trustworthiness, likability, and likelihood that the candidate would change positions in the future.

3.1 Method

3.1.1 Subjects

The subjects were 126 (63 female) Mechanical Turk workers who had IP addresses originating from the United States. Nine additional subjects participated, but were dropped from the experiment; five for failing the instructional manipulation check (described below), three for reporting low levels of political knowledge, and one for failing the political expertise quiz. The average age of the final sample was 37 (SD = 11.82) with ages ranging from 19 to 66. Subjects

¹ A third goal was to replicate the results of Putnam et al. (2014) with the online subject pool and revised political materials

reported their level of education (9% indicated high school education or less, 40% indicated some college, and 52% indicated a bachelor's or graduate degree), which state they were in (36 states were represented), and whether they were born in the United States (6 were not). Subjects completed the experiment on their own computers and were paid \$3 for participating in the hour-long experiment.

Subjects were asked to report their political orientation (the self-report measure described in the norming study) before starting the experiment. In this way, I was able to recruit an equal number of self-reported Democrats and Republicans (50 each) and a smaller number of Independents (26) (once a bin reached its maximum, other subjects who fit in that category were told they were ineligible for the experiment). The objective political orientation measure revealed that the sample was more liberal than they reported, with 75 Democrats, 10 Independents, and 41 Republicans, replicating Zell and Bernstein's (2014) finding that people often under-estimate how liberal they are. Subjects reported a mean political expertise of 58.74 ($SD = 21.77$) on the self-report question and correctly answered an average of 6.02 out of 8 questions ($SD = 1.66$) on the political expertise quiz (The averages for the expertise measures did not include subjects who were omitted).

3.1.2 Materials

The revised political materials (see Appendix A) and the measures of political expertise and political orientation described in the norming study were used. Additional materials included photos of the two candidates and of the locations where the two "debates" occurred (see Appendix B).

3.1.3 Design and Counterbalancing

The experiment used a 3 (Position Type: repetition, control, changed) x 2 (position strength: strong, weak) x 2 (statement orientation: strong, weak) within-subjects design. The 36 political topics were divided into three groups of 12, and the three groups were rotated through the different conditions (repetition, control, changed). Within each group, half of the topics presented the strong position at Debate 1, and half presented the weak position at Debate 1; this grouping was counterbalanced across subjects. Thus, each candidate addressed 24 topics at Debate 1 (12 to be repeated and 12 to be changed) and 36 topics at Debate 2 (12 from each condition).

In addition, the political affiliation of each candidate was counterbalanced across subjects. For half of the subjects Mike Shipman was a Republican and John Baker was a Democrat, whereas for the other half of the subjects the political parties were reversed.

3.2 Procedure

Before beginning the experiment subjects read an online informed consent letter and agreed to participate. They then answered the demographics questions (described in the norming study) and reported their political orientation. One additional question was an instructional manipulation check to determine whether subjects were reading the instructions (Oppenheimer, Meyvis, & Davidenko, 2009). It appeared to be a question asking subjects to check boxes corresponding to what news sources they followed, but the paragraph of instructions told subjects to ignore the check boxes and to type the word “understand” in a text box. The data from subjects who failed the instructional manipulation check (five in Experiment 1) were excluded from all analyses.

Before beginning the experiment subjects were introduced to the “campaign.” They saw the names, photos, and party affiliations of the two candidates, and were informed that the two politicians debated each other twice over the course of the campaign. Subjects saw a photo of where Debate 1 occurred, and read instructions indicating that they would be reading debate excerpts and trying to form an impression of the candidates.

During Debate 1 each trial presented one statement from a candidate. The candidate’s name, face, party affiliation (just a D or R) and the current topic appeared on screen for 2 seconds. Then the candidate’s position on that topic appeared for 7 seconds, after which there was a 500 ms blank screen before the next trial began. The statements were presented in a random order, with only the repetition and changed items being presented for a total of 48 items (24 for each candidate).

After finishing Debate 1, subjects were asked to vote for one of the candidates. Then, they answered three questions for each candidate: 1) “How much do you like this candidate?” 2) “How trustworthy is this candidate?” and 3) “How likely is this candidate to change positions in the future?” Subjects responded by using a slider, dragging it to the left to indicate dislike, distrust, or a low estimate of future change and dragging it to the right to indicate likability, trust, or a high estimate of future change.

Subjects then spent 3 minutes on a distractor task, writing about what they would have done in October in the two weeks between the debates. Then subjects saw a picture of the college activity center where Debate 2 was held, and read instructions telling them that they would be reading debate excerpts again. Subjects were informed that they would be seeing some old topics

and some new topics, and that for the old topics they should try and notice when the candidates changed positions. Subjects were also told to try and remember each candidate's position.

All of the item types (repetition, changed, control) were presented during Debate 2, for 72 total presentations (36 for each candidate).

Immediately after finishing Debate 2 subjects began the cued recall test, where they were told to try and remember each candidate's most recent position. On a test trial subjects were cued with a candidate's name, photo, topic, and the first half of his position statement, and were asked to type in the remainder of his position from Debate 2. After submitting a response subjects were asked "Did this candidate change positions on this topic between Debate 1 and Debate 2?" and clicked a "yes" or "no" button (the measure of change recollection). There was no time limit for responding, and subjects were tested on all 72 position statements.

After completing the recall test subjects voted again, and answered the same set of questions that was presented after Debate 1. One additional question asked: "This candidate addressed 36 topics across the two debates. How many times do you think he changed positions between Debate 1 and Debate 2?" Subjects used a pull down menu to select a number between 0 and 36. After answering all of the questions, subjects completed the objective political orientation quiz, reported their own level of political expertise, and took the political expertise quiz. Finally, they were taken to a thanks and debriefing page where they were provided with instructions for receiving payment.

3.3 Results

The cued recall responses were coded by two independent raters (inter-rater reliability was reasonably high, as measured by Cohen's Kappa, .75), and any differences were resolved through discussion. The effect sizes for t tests reported below are Cohen's d , and a Bonferroni correction was used for all post-hoc analyses unless otherwise mentioned. Several analyses explored the effects of a subjects' political orientation on other variables; although subjects reported their orientation on a 7-point scale, for simplicity the analyses below used only three categories: Democrat, Independent, and Republican. Additional analyses are included in Appendix E that use the full 7-point scale and the objective measure of political orientation (along with some additional exploratory analyses). In general, however, the self-reported and objective measures of political orientation led to consistent results.

3.3.1 Recall

As expected, proactive interference occurred: the repetition items ($M = .57$, $SEM = .02$) led to better recall than the control items ($M = .49$, $SEM = .02$) which in turn were better than the changed items, ($M = .37$, $SEM = .02$), $F(2, 250) = 116.93$, $p < .001$, $\eta_p^2 = .48$. Planned comparisons revealed that all of the differences were significant, $ts(125) \geq 6.06$, $ps < .001$, $ds \geq 0.40$.

Figure 3 shows recall collapsed across item type for the Democratic and Republican candidate as a function of self-reported political orientation. A 3 (subject political orientation: Democrat, Independent, Republican) x 2 (candidate party: Democrat, Republican) mixed-model ANOVA failed to reveal a main effect of political orientation, $F(2, 123) = 2.81$, $p = .064$, $\eta_p^2 = .04$, or of candidate political party, $F(1, 123) = 0.04$, $p = .837$, $\eta_p^2 = .000$. The interaction, however, was significant, $F(2, 123) = 3.34$, $p = .039$, $\eta_p^2 = .05$, suggesting that a congeniality effect may have occurred. Simple effects analyses revealed that the self-reported Democrats had better recall for

the Democratic candidate's positions than the Republican candidate's positions (.51 vs. .48), $p = .048$, but that the self-reported Independents and Republicans had similar recall for both candidates ($t_s < 1.10$).

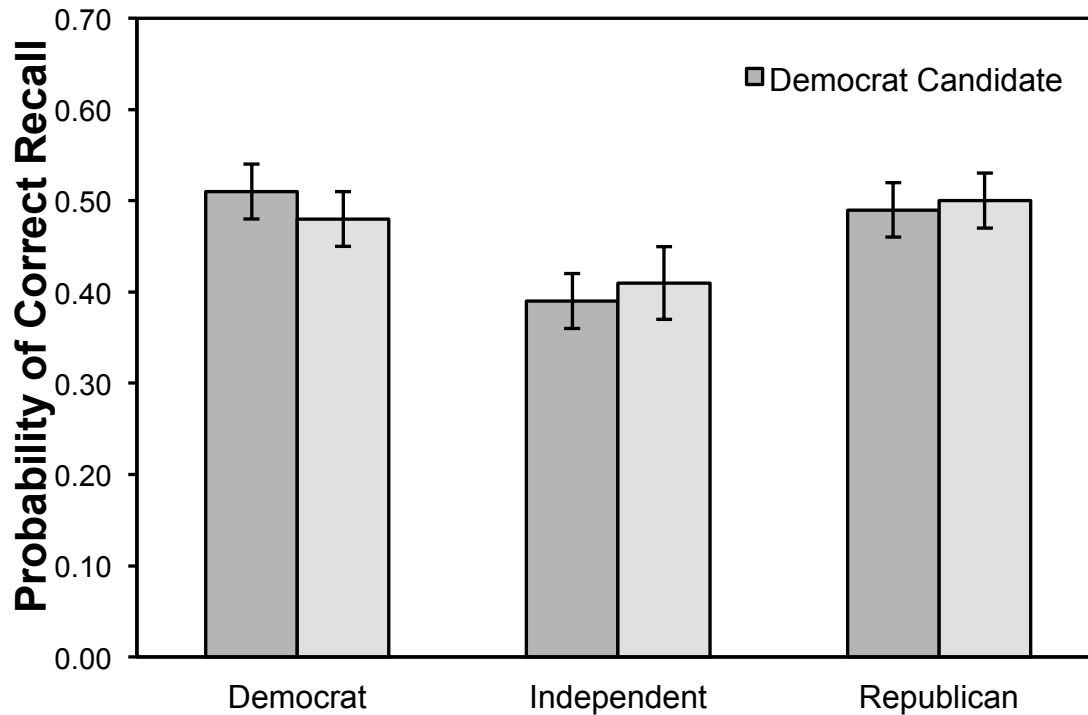


Figure 3.1: Cued recall as a function of self-reported political orientation for the two different candidates. Error bars represent standard error of the means.

To confirm whether a congeniality effect occurred, I compared recall for when a subject's political party was congruent with the candidate's (Democrats recalling Democratic positions and Republicans recalling Republican positions, $M = .51$, $SEM = .02$) to when it was incongruent (Democrats recalling Republican positions and Republicans recalling Democratic positions $M = .49$, $SEM = .02$). Despite the difference between the means only being .02, a paired-samples t -test revealed that recall for the congruent positions was indeed better than for the incongruent positions, $t(99) = 2.22$, $p = .029$, $d = 0.11$, albeit with a very small effect size. This analysis

suggested that subjects may be better at recalling political information when the candidate is from their own political party rather than a different one (i.e., a congeniality effect).

3.3.2 Change Recollection

Table 3 shows how accurate subjects were in recollecting change. Recollecting change for a changed item was treated as a hit and recollecting change for a control item was treated as a false alarm. Table 3 also shows the signal detection measures of d' and c . Subjects successfully discriminated the changed items from the control items: d' was greater than zero, $t(125) = 12.15$, $p < .001$, $d = 1.08$.

Table 3.1: Probabilities of Change Recollection and Signal Detection Measures in Experiment 1

	Item Type		SDT Measure	
	Control (FAs)	Changed (Hits)	d'	c
Experiment 1				
Mean	.13	.37	0.86	0.87
SEM	.01	.02	0.07	0.06

Did political orientation influence change recollection? Table 4 shows change recollection accuracy (measured by d') for the two candidates as a function of subject political orientation. A 3 (subject political orientation: Democrat, Independent, Republican) x 2 (candidate political party: Democrat, Republican) mixed model ANOVA failed to reveal any main effects or interactions, $F_s \leq 2.45$, suggesting that the ability for subjects to recollect change was unaffected by partisanship. I also conducted a congruency analysis (similar to that done with recall above) but failed to show a difference in change recollection for the congruent and incongruent items, $t < 1$.

Table 3.2: Change Recollection Accuracy, Measured by d' for the Different Candidates as a Function of Subject's Self-Reported Political Orientation

	Democratic Candidate	Republican Candidate
Democrat	.88 (.12)	.94 (.11)
Independent	.50 (.13)	.56 (.16)
Republican	.83 (.12)	.80 (.11)

Note. Standard errors presented in parentheses.

3.3.3 Recall Conditionalized on Change Recollection

Figure 4 shows recall for the changed items conditionalized on whether change was recollected compared to recall for the control items. As predicted, failing to recollect change led to proactive interference; recall for the changed items was worse than for the control items, $t(125) = 12.44$, $p < .001$, $d = 1.18$. In contrast, when change was recollected, recall for the changed items was not significantly different from recall for the control items, $t(116) = 1.03$, $p = .31$, $d = 0$ (the difference in degrees of freedom for the conditional analyses represents subjects who did not have an observation in a particular cell). In other words, recollecting change eliminated proactive interference, replicating the results of Putnam et al. (2014).

One concern with using conditional analyses is the possibility of item selection effects and differences in subject memory abilities. In line with previous research (Jacoby et al. in press; Putnam et al. 2014; Wahlheim & Jacoby, 2013), I conducted hierarchical linear regression

analyses at the item and subject level to see if change recollection made a unique contribution to recall of the changed items.

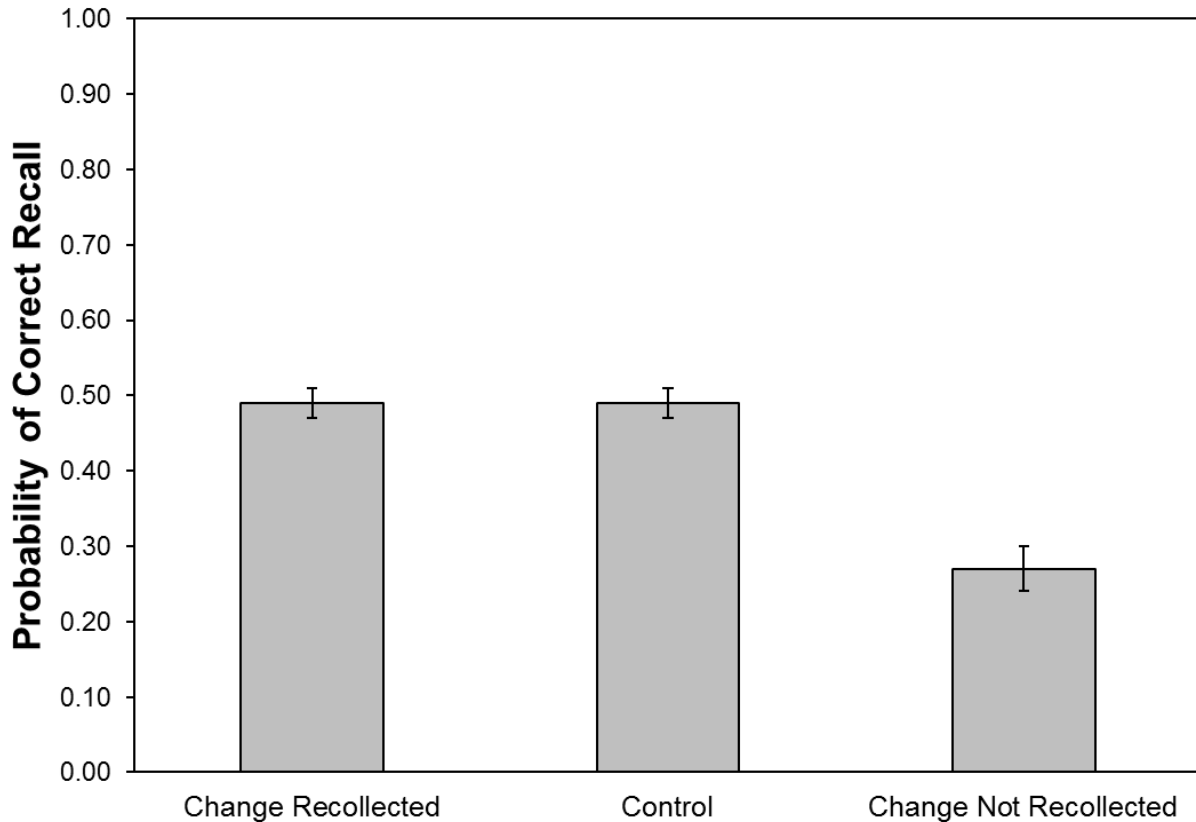


Figure 3.2: Cued recall for the changed positions conditionalized on change recollection, as compared to the control condition in Experiment 1. Error bars represent standard error of the means.

The left column of Table 5 shows the regression model for items. Estimates for each measure were calculated for each of the 144 items. At the first step, recall for the control items was entered; at the second step change recollection (as measured by hits – false alarms) was entered; and at the third step the interaction was entered. Recall for the changed items was the outcome variable. As expected, item differences accounted for significant variance in predicting recall for the changed items (the outcome variable), but change recollection did explain additional variance (.06) above any item differences. The interaction did not improve prediction. The proportion of

variance accounted for by change recollection was similar to previous research (Putnam et al. 2014), suggesting that change recollection did enhance recall of the changed items.

The right column of Table 5 displays a similar analysis, but collapsing across items to get an estimate for individual subjects. Subject differences explained a large portion of the variance, with change recollection only explaining a small amount of additional variance (this additional variance was marginally significant). The interaction did not explain additional variance. The proportion of unique variance explained by change recollection is much smaller than in previous experiments. One potential explanation is that change recollection accuracy was correlated with recall for the control positions, $r = .55, p < .001$, meaning there could be multicollinearity, which can limit the total variance explained (Fields et al., 2012). In addition, individual differences in subject abilities could have been greater in a sample recruited from the internet than in a laboratory-based sample. In sum, however, both models suggest that change recollection is accounting for a unique but small amount of the variance above any item selection or individual subject effects.

Table 3.3: Proportion of Variance in Recall of Changed Positions Explained by Item Differences, General Memory Ability, and Change Recollection in Experiment 3.

	Unit of Analysis	
	Items	Subjects
Step 1		
Item Differences / General Memory	.39 ($p < .001$)	.56 ($p < .001$)
Step 2		
Change Recollection (Hits - FA)	.06 ($p < .001$)	.01 ($p = .056$)
Step 3		
Interaction	.003 ($p = .445$)	.01 ($p = .194$)

Note: Values displayed above are ΔR^2 on each step of the model computed at the item level collapsed across subjects (left) and at the subject level collapsed across items (right). “Item Differences” refers to item differences in control position recall performance, “General Memory” refers to individual differences in control position recall performance, and “Change Recollection (Hits - FA)” refers to individual differences in discriminability of change recollection for changed positions.

3.3.4 Subject Judgment Results

Voting. Table 6 shows the voting choices for the subjects at Debate 1 and after the recall test as a function of their self-reported political orientation. As expected, subjects tended to vote along party lines. Only six subjects changed which candidate they voted for between Debate 1 and the recall test.

Table 3.4: Voting Results for the Two Candidates After Debate 1 and After the Recall Test by Subject Political Orientation, Experiment 1

	Debate 1		Recall Test	
	Democrat	Republican	Democrat	Republican
Total Votes	75	51	79	47
Self-Reported Political Orientation				
Democrats	49	1	50	0
Independents	17	9	19	7
Republicans	9	41	10	40

Estimates of future change. Table 7 displays the ratings that subjects made about the two candidates after Debate 1 and the recall test, broken down by the subjects’ self-reported political orientation. Subjects rated each candidate individually on the probability of their changing positions in the future, their likability, and their trustworthiness twice: once after Debate 1 and once after the recall test. In addition, after the recall test, subjects estimated how often each

candidate changed positions between the two debates. The bottom row represents the average change in scores between Debate 1 and the recall test.

The third and fourth columns of Table 7 display the subject estimates of future change for each candidate. At Debate 1, a 3 (subject political orientation) X 2 (candidate party) mixed-model ANOVA revealed a main effect of candidate party, suggesting that the subjects predicted that the Democratic candidate was more likely to change positions than the Republican candidate, $F(1, 123) = 3.91, p = .05, \eta_p^2 = .03$.

There was no main effect of subject political orientation nor a significant interaction (both F s < 1). In contrast, at the recall test, a 3 X 2 ANOVA failed to show an effect of candidate party, $F(2, 123) = 2.46, p = .09, \eta_p^2 = .04$. (Also, there was no main effect of subject political orientation nor an interaction, both F s < 1). In sum, after Debate 1 subjects predicted that the Democratic candidate would be more likely to change positions in the future than the Republican candidate, but after the recall test made similar estimates for both candidates. Furthermore, it appeared that subjects estimated both candidates as being more likely to change positions in the future after the test than after Debate 1, $ts(125) \geq 5.22, ps < .001, ds \geq 0.61$.

The future change estimates that occurred after the test appeared to be informed by change recollection. The hit rates for recollecting change for the individual candidates was correlated with their respective estimates of future change: $r = .33, p < .001$ for the Democrat and $r = .32, p < .001$ for the Republican (see top row of Figure 5). Furthermore, the hit rates were also correlated with the change in the estimates between Debate 1 and the test: $r = .27, p = .002$ for the Democrat and $r = .27, p = .002$ for the Republican.

Table 3.5: Subject Opinions About the Candidates as a Function of Their Self-Reported Political Orientation in Experiment 1

Subject Political Orientation	Past Change Estimate		Future Change Estimate		Trustworthiness		Likability	
	Democrat	Republican	Democrat	Republican	Democrat	Republican	Democrat	Republican
After Debate 1								
Democrat	-	-	35.50	28.94	70.86	24.60	79.38	14.06
Independents			38.69	36.31	51.35	38.35	52.96	29.12
Republican	-	-	38.90	29.90	40.30	57.22	37.36	66.80
Average			37.51	30.84	54.71	40.38	57.25	38.10
After Recall Test								
Democrat	7.34	7.12	50.48	51.00	64.56	24.38	75.70	13.18
Independents	5.62	5.96	45.08	50.73	50.96	36.54	52.35	25.81
Republican	6.60	6.54	56.06	47.54	37.94	53.06	38.94	62.16
Average	6.69	6.65	51.58	49.57	51.19	38.27	56.29	35.22
Change From D1 to Test								
Average			14.07	18.73	-3.52	-2.11	-0.96	-2.87

For Past Change subjects estimated the number of times between 0 and 36 that a candidate changed. For Future Change subjects estimated how likely a candidate was to change in the future (0 = not very likely, 100 = extremely likely). For Trustworthiness, subjects reported how much they trusted each candidate (0 = not trustworthy at all, 100 = extremely trustworthy). For Likability, subjects were asked how much they liked each subject (0 = don't like at all, 100 = like a lot).

Estimates of Past Change. After finishing the test, subjects were asked to estimate how many times each candidate changed positions; the first two columns of Table 7 shows the estimates. A 3 (subject orientation) X 2 (candidate party) mixed-model ANOVA failed to reveal any significant main effects or an interaction (all F s < 1), suggesting that neither a subject's own political orientation nor the political party of the candidate influenced past estimates of change. How often subjects recollected change for each candidate, however, did appear to influence their estimates of past change. The hit rates for recollecting change for each candidate correlated with their overall estimates of past change: $r = .48, p < .001$ for the Democrat and $r = .51, p = .001$ for the Republican.

Trustworthiness Ratings. As a reminder, there were two predictions about the trustworthiness and likability ratings, 1) that subjects would prefer the candidate from their own party (and dislike the opposite candidate) regardless how often the candidate change positions, and 2) that subjects would provide lower ratings to the candidates who they perceived as changing positions more often. The fifth and sixth columns of Table 7 display the trustworthiness ratings at Debate 1 and after the test for each candidate as a function of subject political orientation. For the Debate 1 ratings a 3 (political orientation) x 2 (candidate party) mixed-model ANOVA revealed a main effect of candidate party indicating that the Democrat was perceived as more trustworthy than the Republican, $F(1, 123) = 21.30, p < .001, \eta_p^2 = .15$, and an interaction, indicating that subjects provided higher trustworthiness ratings to the candidate from their own party, $F(1, 123) = 21.30, p < .001, \eta_p^2 = .15$. There was no main effect of subject political orientation, $F < 1$. The ratings after the recall test yielded an identical pattern: there was a main effect of candidate party, $F(1, 123) = 19.70, p < .001, \eta_p^2 = .14$, a significant interaction, $F(2, 123) = 37.93, p < .001, \eta_p^2 = .38$, and no effect of subject political orientation ($F < 1$). The interaction appeared to be

driven by both the Democratic and Republican subjects. The main effect of the Democratic candidate appearing more trustworthy than the Republican is likely due to the Independents, many of whom would be considered Democrats by an objective measure (Appendix E reports this analysis using the objective measure; the interaction is still significant, but the main effect of candidate political party disappears). Two paired-samples *t*-tests showed that subjects trusted the Democratic candidate more at Debate 1 than after the test, $t(125) = 2.63, p = .010, d = 0.14$, and that they trusted the Republican candidate equally at both times, $t = 1.53$.

Did change recollection influence the trust ratings? After the test, the hit rates for the individual candidates correlated negatively with the trust ratings for the Republican candidate, $r = -.18, p = .045$, but not for the Democratic candidate, $r = -.09, p = .315$ (see middle panels of Figure 5). For the Republican candidate, a higher hit rate was associated with lower trust ratings, supporting the hypothesis that recollecting change led to a decrease in trust. Looking at the change in the trust scores between Debate 1 and the test, however, revealed that the hit rate for change recollection was correlated with the Democratic candidate, $r = -.38, p < .001$, but not for the Republican candidate, $r = -.10, p = .286$. The significant negative correlation for the Democratic candidate suggests that recollecting change led to the lower trust rating after the test. Taken together, these correlations suggest that recollecting change may lead to lower ratings of candidate trustworthiness.

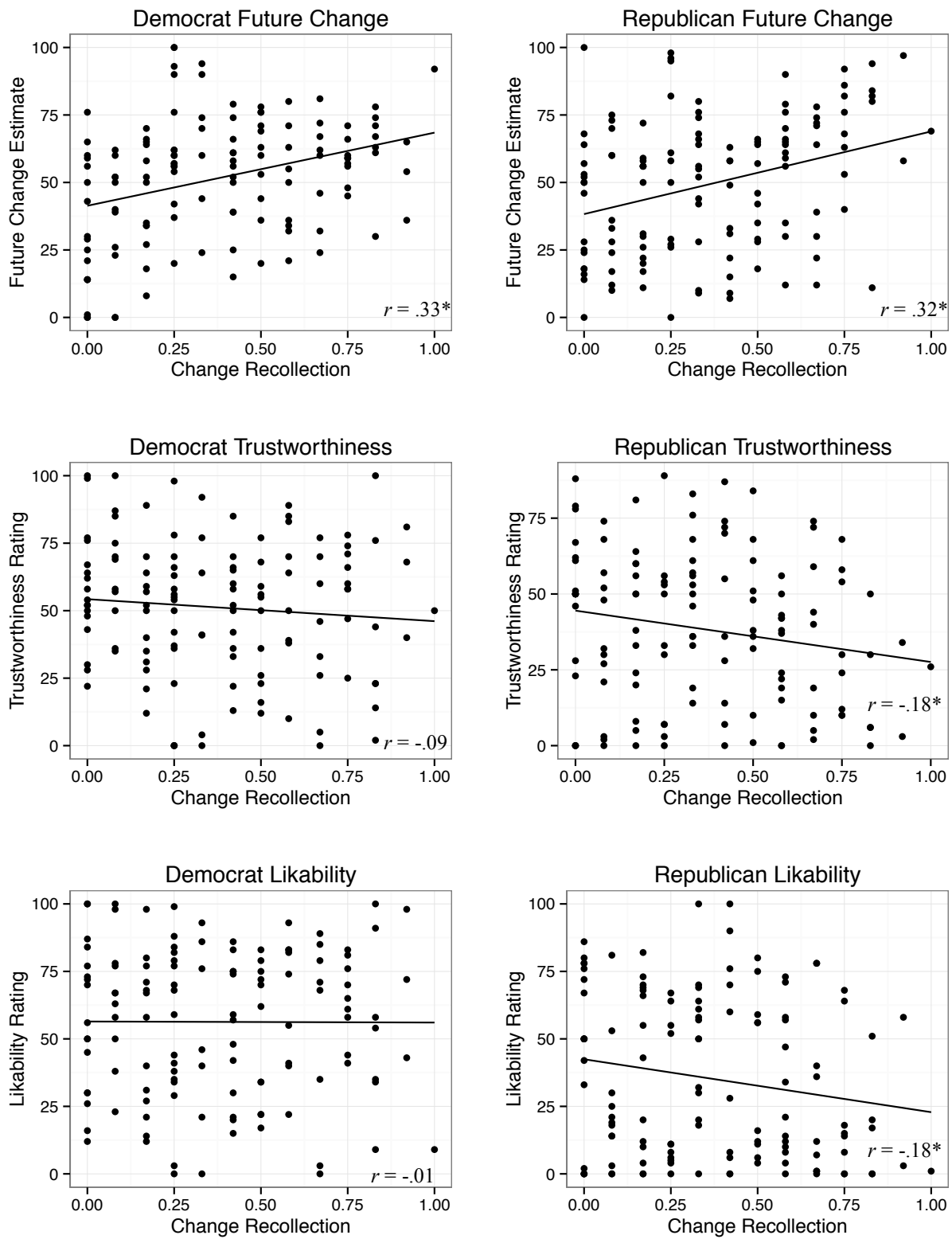


Figure 3.3: Scatterplots showing correlations between change recollection hit rate and subject judgments after the recall test in Experiment 1. * $p \leq .05$

Likability Ratings. The likability ratings were similar to the trust ratings: the last two columns of Table 7 show that subjects liked the candidates from their own party more than the candidates from the opposite party. For the Debate 1 ratings, a 3 (subject political orientation) x 2 (candidate party) mixed-model ANOVA revealed a main effect of candidate party, $F(1, 123) = 38.59, p < .001, \eta_p^2 = .24$, indicating that the Democratic candidate was liked more than the Republican candidate, and a main effect of subject political orientation, $F(2, 123) = 7.14, p = .001, \eta_p^2 = .10$. Post-hoc analyses indicated that the Republican subjects liked both candidates more than the Independents did, $p = .001$. Both main effects, however, were qualified by a significant interaction, $F(2, 123) = 95.57, p = .001, \eta_p^2 = .61$, which suggested that subjects preferred the candidates from their own parties. Similarly, for the post-test ratings the Democratic candidate was liked more than the Republican candidate, $F(1, 123) = 47.60, p < .001, \eta_p^2 = .28$, and the Republican subjects liked both candidates more than the Independent subjects or the Democratic subjects, $F(2, 123) = 8.23, p = .001, \eta_p^2 = .12$ ($ps < .039$ for the post-hoc tests). The interaction was also significant, $F(2, 123) = 79.60, p = .001, \eta_p^2 = .56$, indicating that subjects liked the candidate from their own party more than the candidate from the opposite party.

Did likability scores change between Debate 1 and the test? Two paired-samples *t*-tests showed that the subjects liked the Republican less after the test than after Debate 1, $t(125) = 2.26, p = .025, d = 0.09$, but that subjects liked the Democrat similarly at both times, $t = 0.74$. The change recollection hit rate did not correlate with the test likability ratings for the Democrat, $r = -.01, p = .966$, but did marginally correlate with the likability ratings for the Republican, $r = -.18, p = .050$. (see bottom panels of Figure 5). The hit rates for the two candidates did not correlate with

the change in likability rating for either candidate, (all $ps \geq .542$). Thus, recollecting change may have resulted in lower likability ratings for the Republican candidate.

3.4 Experiment 1 Discussion

Experiment 1 replicated previous research (Putnam et al., 2014) in showing that recollecting change reduced interference: recall for the changed items was worse than the control items overall, but conditional analyses revealed that recollecting change eliminated interference. This replication is important given the current experiment used a revised set of political materials and a different subject population.

Experiment 1 also addressed whether political orientation affected change recollection or recall. A small congeniality effect occurred, suggesting that subjects were better at recalling information that was consistent with their own political views. The effect size is small ($d = 0.11$), but consistent with other research on the congeniality effect (Eagly et al., 2001). Unlike recall, change recollection appeared to be unaffected by political orientation; it was similar regardless of the subject's or the candidate's party affiliation. In sum, voters may be better able to remember their own candidate's positions, but are equally capable of remembering past changes in position.

Finally, Experiment 1 also explored whether recollecting change influenced subject perceptions of the candidates. Recollecting change was clearly correlated with the estimates of past and future change. Both of those estimates were a form of category level judgment (see Jacoby, Wahlheim, & Coane, 2010), and a correlation between those estimates and change recollection suggests that the category level judgments were driven by the local judgments (e.g., the overall rating of past change was influenced by remembering change on specific items). Thus,

recollecting change at least influenced how the candidates were perceived in terms of their past and future probabilities of changing position.

In contrast, recollecting change appeared to have a much smaller effect (if any) on voting preferences, likability ratings, and trustworthiness ratings. All three measures showed a strong interaction at both times indicating that subjects had more positive views of the candidates from their own party than the opposite. There were, however, some small indications that subjects did update their attitudes towards the candidates after Debate 2. For the Republican candidate recollecting change was negatively correlated with likeability and trustworthiness ratings after the test. In addition, the Democrat was perceived as less likable after the test than after Debate 1. These correlations suggest (albeit in a small way) that recollecting change led to a less positive perception of the candidates.

In sum, Experiment 1 showed that recollecting change reduced interference, that subjects may be slightly better at recalling information that is consistent with their own views, and that recollecting change may have some small impact on how voters perceive candidates. Experiment 2 extended the paradigm to examine how voters integrated information from a third party source.

Chapter 4: Experiment 2

In the context of a real political campaign, voters are probably more likely to learn about a politician changing position from an external source rather than noticing it on their own. For example, some websites keep track of when politicians change positions (PolitiFact.com uses the categories of no flip, half flip, and full flop). However, as noted in the introduction, flip-flopping accusations are not always warranted. Sometimes a politician may be accused of flip-flopping, even though his position has remained essentially the same. Experiment 2 explored how third party information about a candidate's behavior affected later change recollection and recall. Sometimes the information was accurate, and other times it was misleading.

Experiment 2 used the same procedure from Experiment 1 with the addition of a "news analysis" phase inserted between Debate 2 and the final test. During that phase subjects read statements from a "news" source about the consistency of the candidate on a specific position between debates. Half of the time the news statements suggested that the candidates had been consistent and half of the time suggested that the candidate had changed positions. Thus, sometimes the news was accurate (claiming that a candidate changed positions when he did, or claiming that he was consistent when he was), and sometimes the news was inaccurate (claiming that a candidate changed positions when he did not or claiming that he was consistent when he changed).

Subjects were warned that the news statements might not be correct and were asked to verify the accuracy of each statement by pressing a *True*, *False*, or *I Don't Know* button. During the recall test subjects attempted to recall the Debate 2 positions and completed the standard measure of change recollection.

From a reminders perspective, verifying the accuracy of the news statement requires recollecting change: subjects must think back to the candidate's actual behavior across the debates and decide if he actually changed positions or was consistent (or only addressed the topic at Debate 2). Correctly verifying a news statement suggests that subjects were able to remember a previously detected change or repetition. In this way, the news verification task is a novel measure of change recollection. Rather than occurring simultaneously with a recall test (e.g., Wahlheim & Jacoby, 2013) it occurs in separate phase before the memory test. Given that previous works suggests that change recollection is highly correlated with earlier change detection (Jacoby et al. 2013), measuring change recollection before the recall test should yield similar results to measuring change recollection concurrently with recall; recollecting change should ameliorate proactive interference for the changed items.

Additionally, correctly verifying a news statement may provide retrieval practice for recollecting change during the final test (Roediger & Karpicke, 2006). Thus, subjects who correctly verify a news statement should show increased change recollection accuracy in the future.

In contrast, failing to accurately verify a news statement suggests that subjects did not have access to previous information about the candidate's behavior. The subject either failed to detect a change when it occurred, or having detected it, was unable to recollect it during the news phase. Wahlheim and Jacoby (2013) showed that subjects almost never recollect change without initially detecting it, so failing to verify a news statement is a strong predictor that subjects will not recollect change at the final test. Critically, however, if a news statement is untrue, it may also serve as a misinformation—it may suggest that a candidate changed positions when he never did (e.g., Loftus, Miller, & Burns, 1978). A subject who endorses an inaccurate news statement

should show poor change recollection at the final test, which should translate to poor recall for the changed items.

Several predictions can be made about Experiment 2. First, the changed items should show worse recall compared to control items, but recollecting change should reduce or eliminate that proactive interference. This pattern should occur whether the measure of change recollection was the news verification task or the standard task during the final test.

Second, subjects might use the news statements to inform their change recollection judgments at the final test. The news statements that suggest a candidate changed positions should lead to higher change recollection rates across all item types compared to the news statements that suggest a candidate was consistent. This increase would be represented by a higher false alarm rate for the repetition and control items (because no change actually occurred) and a higher hit rate for the changed items (because there was a change).

Third, correctly verifying a news statement should lead to accurate change recollection at the final test. For the repetition and control items, a correct verification should lead to a low false alarm rate and reasonably high levels of recall. For the changed items, a correct verification should also lead to a high hit rate for recollecting change, which consequently should lead to improved recall.

In contrast, if subjects are not able to correctly verify a news statement they should show poor change recollection accuracy during the final test. This may translate to poor recall for some item types. Subjects can fail to make a correct verification in two ways. They can endorse an untrue statement (responding *True* to a flip-flopping claim about a politician who was consistent – a *misinformation* response) or they can fail to verify a true statement (responding *False* to a flip-

flopping statement about a politician who changed – a *miss* response). Both types of errors should harm memory, although a misinformation response may have a more extreme effect, because subjects would actively be endorsing an untrue statement. These errors should have different effects on memory depending on the specific item type. For the repetition items, failing to correctly verify a statement should lead to a higher false alarm rate in change recollection, but only a minor (if any) decrease in recall (because there is no competing position). For the control items, failing to correctly verify a statement should not have a large effect on change recollection. The news statements suggest either that a candidate changed positions or was consistent—both are untrue for the changed items, but the two statements might have contradictory effects on later change recollection. Thus, failing to verify a control statement may not have a large impact on change recollection. Likewise, because there is no competing position from Debate 1, failing to verify a news statement for a control item should not affect recall. Finally, for the changed items, failing to correctly verify a statement should lead to a low change recollection rate which in turn should lead to poor recall of the Debate 2 position. Failing to verify a statement suggests that subjects were unable to access the recursive trace during the news phase, and consequently will be unable to access that trace at the final test. Without the ability to accurately recollect change, recall will suffer from proactive interference.

Finding this predicted pattern of results would be interesting for several reasons. First, it would begin to address the real world question of how third party information, from a news source or political pundit, shapes how voters remember a politician's change in position. Second, the news verification task provides a new approach to measuring change recollection. Third, it would also suggest that retrieving the experience of being reminded (accessing the recursive trace) facilitates future change recollection, showing that covert retrieval practice can enhance the memorability

of a recursive trace (Putnam & Roediger, 2013). Finally, if failing to verify the news statements results in a poorer discrimination for change recollection, it would suggest that subjects incorporate misleading information into their memories, even if they know the source of the information is not to be trusted.

4.1 Method

4.1.1 Subjects

One hundred subjects from the same pool as Experiment 1 participated. The same recruiting strategy was used (having subjects report their political orientation before beginning the experiment) to ensure there were 50 Democrats and 50 Republicans; no self-reported independents were recruited. Sixteen additional subjects participated, but their data were omitted from all analyses: three for failing the instructional manipulation check, seven for reporting low levels of political knowledge, and six for failing the political expertise quiz. The objective political measure revealed that the sample was more liberal than expected from the self-report, with 62 Democrats, 7 Independents, and 31 Republicans. The average age was 36 ($SD = 10.68$), with ages ranging from 21 to 73. Subjects reported their level of education (9% indicated high school education or less, 36% indicated some college, and 55% indicated a bachelor's or graduate degree), which state they were in (32 states were represented), and whether they were born in the United States (2 were not). Subjects reported a mean political expertise of 61.09 ($SD = 22.24$) on the self-report question and answered an average of 5.4 out of 8 questions on the political expertise quiz.

4.1.2 Design and Procedure

The experimental design and procedure was similar to Experiment 1 with the addition of a “news” phase that occurred between Debate 2 and the final test. During this phase subjects read statements from a fictitious news source that claimed that candidates either changed positions or were consistent on individual topics. They were informed that the news was sometimes accurate and sometimes inaccurate, and that their job was to verify, based on their memories, when the news was accurate. Half of the news reports suggested that the candidate changed positions (*flip-flop* statements, e.g., “Mike Shipman changed positions on Same Sex Marriage”); the other half of the news reports suggested that the candidate did not change positions (*straight* statements, e.g., John Baker was consistent in his views on Same Sex Marriage”)². The news statements appeared one at a time in a random order along with a picture of the candidate. After 1.5 seconds, subjects were asked to respond to the statement by pressing a *True*, *False*, or *I Don’t Know* button, after which a blank screen appeared briefly before the next trial began. Subjects were instructed to use the *I Don’t Know* button when there was not enough information to know if the candidate changed position or not (i.e., for the control items). Every item was presented during the news phase, for a total of 72 presentations. Thus the experiment was a 3 (item type: repetition, control, changed) x 2 (news report: straight, flip-flop) design.

A similar counterbalancing scheme to Experiment 1 was used. The 72 topics were split into three groups, with each group assigned to one item type (repetition, control,

² I am using the *flip-flop* label for the news phase to avoid confusion with the changed item type and the *straight* label was used to avoid confusion with the repetition item type.

changed). Within a group half of the items displayed the strong statement at Debate 2 and half displayed the weak statement at Debate 2. Within the sub-groups, half of the items were associated with a *flip-flop* news statement and the other half were associated with a *straight* news statement. These groups were then rotated across conditions, ensuring that each item appeared in every condition equally often across counterbalancing conditions.

4.2 Results

Two independent coders marked the cued recall responses and I resolved any disagreements between the coders (inter-rater reliability, as measured by Cohen's Kappa, was .74). First I will report the overall cued recall and change recollection results, then the results taking the news phase into account. The analyses examining the effects of political orientation on recall and change recollection, along with the voting and judgment data, are reported in Appendix F.

4.2.1 Recall and Change Recollection

Experiment 2 replicated Experiment 1 in showing a proactive interference. The repetition items ($M = .56$, $SEM = .02$) led to better recall than the control items ($M = .43$, $SEM = .02$), which in turn were recalled better than the changed items ($M = .32$, $SEM = .02$), $F(2,198) = 123.82$, $p < .001$, $\eta_p^2 = .56$. Follow up comparisons revealed that all of the differences were significant, $ts(99) \geq 8.08$, $ps < .001$, $ds \geq 0.61$. Table 8 shows the change recollection rates for the control and changed items, along with d' and c . Subjects successfully recollected change: d' was greater than zero, $t(99) = 8.82$, $p < .001$, $d = 0.88$. Accuracy in change recollection was lower in Experiment 2 ($d' = 0.52$) than in Experiment 1 ($d' = 0.86$), $t(224) = 3.53$, $p < .001$, $d = 0.49$, likely because of the misleading news statements.

Table 4.1: Probabilities of Change Recollection and Signal Detection Measures in Experiment 2

	Item Type		SDT Measure	
	Control (FAs)	Changed (Hits)	d'	c
Experiment 2				
Mean	.17	.32	0.52	0.87
SEM	.02	.02	0.06	0.06

4.2.2 Recall Conditionalized on Change Recollection

Figure 6 shows that conditionalizing recall for the changed items on change recollection replicated Experiment 1. When subjects did not recollect change, recall was worse for the changed items than it was for the control items, $t(99) = 10.48$, $p < .001$, $d = 1.18$. But when subjects did recollect change, recall for the changed items was equal to the control items, $t(90) = 1.28$, $p = .206$, $d = 0.14$, showing that recollecting change eliminated proactive interference.

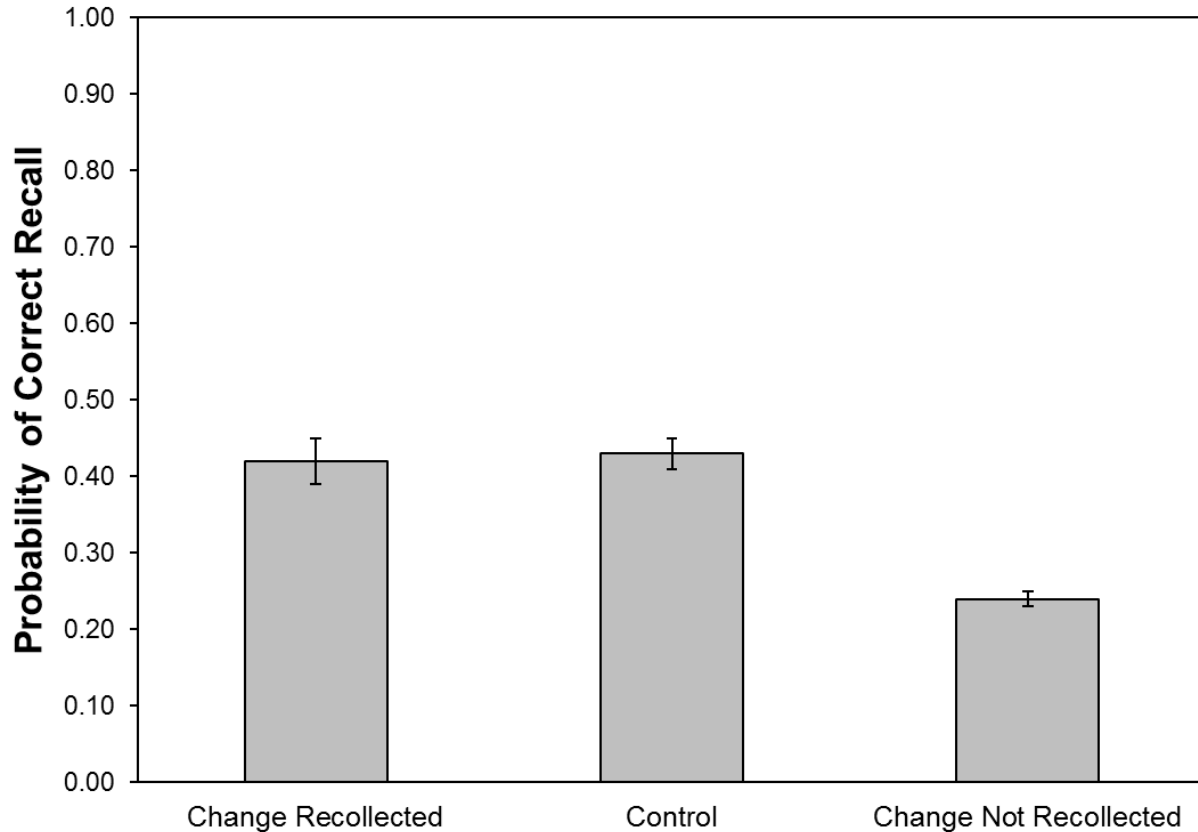


Figure 4.1: Cued recall for the changed positions conditionalized on change recollection, as compared to the control condition, Experiment 2. Error bars represent standard error of the means.

Table 9 displays two hierarchical regression analyses exploring item and subject effects conducted in the same fashion as in Experiment 1. The left column shows that item differences explained a large portion of variance in predicting recall of the change items. Change recollection had a marginally significant contribution, explaining a small proportion of variance. The interaction was not significant. The right column shows that subject differences contributed to the recall of the control items, but that change recollection explained additional variance after controlling for subject differences. Again, the interaction was not significant. The inclusion of the news phase (which affected change recollection) makes it difficult to directly compare

Experiment 2 to Experiment 1. However, the results of these regression analyses suggested that change recollection did make a unique contribution to recall of the changed items.

Table 4.2: Proportion of Variance in Recall of Changed Positions Explained by Item Differences, General Memory Ability, and Change Recollection in Experiment 2.

	Unit of Analysis	
	Items	Subjects
Step 1		
Item Differences / General Memory	.37 ($p < .001$)	.48 ($p < .001$)
Step 2		
Change Recollection (Hits - FA)	.02 ($p = .058$)	.05 ($p = .001$)
Step 3		
Interaction	.003 ($p = .971$)	.007 ($p = .227$)

Note: Values displayed above are ΔR^2 on each step of the model computed at the item level collapsed across subjects (left) and at the subject level collapsed across items (right). “Item Differences” refers to item differences in control position recall performance, “General Memory” refers to individual differences in control position recall performance, and “Change Recollection (Hits - FA)” refers to individual differences in discriminability of change recollection for changed positions.

4.2.3 News Phase Responses

During the news phase subjects saw *straight* statements (suggesting a candidate had been consistent, e.g., Mike Shipman was consistent in his views on Same Sex Marriage) and *flip-flop* statements (suggesting a candidate had changed positions, e.g., John Baker changed positions on Same Sex Marriage), and were asked to push one of three buttons – *True*, *False*, or *I Don’t Know*—based on their memory of the candidate’s past behavior.

Table 10 shows the mean proportion of each response for the different item types and the different news statements. Across all item types (this mean is not shown in Table 10) subjects were more likely to endorse the *straight* news statements as true ($M = .70$, SEM

= .02) rather than false ($M = .18$, $SEM = .01$), $t(99) = 18.44$, $p < .001$, $d = 3.42$, and more likely to endorse the *flip-flop* statements as false ($M = .70$, $SEM = .02$) rather than true ($M = .17$, $SEM = .02$), $t(99) = 18.74$, $p = .001$, $d = 3.43$. Subjects were equally likely to use the *I Don't Know* response for both types of news statements, $t < 1$. This pattern suggests that subjects believed candidates were often consistent between debates despite the news claiming otherwise and the candidates actually changing positions on some topics.

Table 4.3: Mean Proportion of Each Type of Response on the News Verification Task Organized by Item Type and News Statement Type in Experiment 2

Item Type & News Statement	Subject Response		
	True	False	I Don't Know
Repetition			
News: Straight	.82 (.02)	.11 (.01)	.07 (.01)
News: Flip-flop	.10 (.01)	.82 (.01)	.08 (.01)
Control			
News: Straight	.62 (.02)	.15 (.01)	.23 (.02)
News: Flip-flop	.12 (.01)	.65 (.03)	.23 (.02)
Changed			
News: Straight	.65 (.02)	.27 (.02)	.08 (.01)
News: Flip-flop	.28 (.02)	.63 (.02)	.09 (.01)

A more interesting analysis, however, is examining how accurate subjects were in the news verification task. For the news phase, subject responses were grouped into one of three bins *correct*, *misinformation*, or *miss*. *Correct* responses were when subjects accurately verified a news statement. These were indexed as follows: for the repetition items, responding *True* to a straight statement or *False* to a flip-flop statement; for the

control items, responding *I Don't Know* to both the straight and the flip-flop statements; and for the changed items, responding *False* to a straight statement and *True* to a flip-flop statement. *Misinformation* responses were when subjects actively endorsed an untrue statement. These were indexed as follows: for the repetition items, responding *True* to a flip-flop statement; for the control items, responding *True* to either a straight or a flip-flop statement; and for the changed items, responding *True* to a straight statement. Finally, *miss* responses were when subjects did not accurately verify a statement, but also did not endorse an untrue statement. These were indexed as: responding *False* to a straight statement about a repetition item; responding *False* to a flip-flop statement about a changed item; responding *False* to any statement about a control item; or responding with *I Don't Know* to any statement about a repetition or changed item.

Figure 7 shows the response rates for the different item types. The left panel shows the correct response rate. A one-way repeated-measures ANOVA with item type as the independent variable revealed a significant effect, $F(2, 198) = 322.97, p = .001, \eta_p^2 = .77$. Subjects had the highest correct verification rate for the repetition items, a lower rate for the changed items, and the lowest rate for the control items, all $ts(99) \geq 2.15, ps \leq .034, ds \geq 0.25$. A second one-way repeated measures ANOVA with the misinformation responses as the dependent variable (see middle panel) also revealed a significant effect, $F(2, 198) = 305.75, p = .001, \eta_p^2 = .76$. Subjects had the highest misleading endorsement rate for the control items, a slightly lower rate for the changed items, and the lowest rate for the repetition items, all $ts(99) \geq 3.78, ps < .001, ds \geq 0.43$. The high misinformation endorsement rate for the control items is likely because responding *True* to a straight

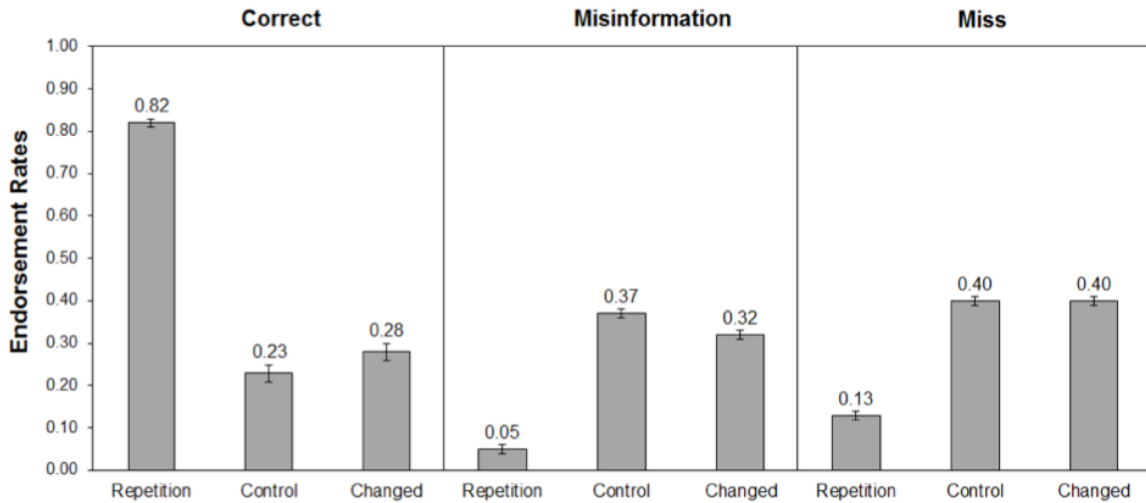


Figure 4.2: The correct verification, misleading endorsement, and miss response rates for the news statements during Experiment 2. See the main text for descriptions of how responses were indexed.

or a flip-flop statement would be considered a misinformation response for a control item, meaning there are two possible ways to make a misleading endorsement. In contrast, for the repetition and changed items there was only one way to make a misleading endorsement—responding *True* to a statement that was inconsistent with the item type. Finally, a third one-way repeated measures ANOVA with the miss responses as the dependent variable (see right panel) revealed a significant effect, $F(2,198) = 177.08, p < .001, \eta_p^2 = .64$. Subjects were less likely to make a miss response for the repetition items compared to the control or changed items, $t_s(99) \geq 14.47, p_s < .001, d_s \geq 2.14$, but were equally likely to make a miss response for the control and changed items, $t < 1$.

In sum, during the news phase subjects had a strong tendency to remember all of the statements as being consistent, regardless of whether they actually were. Subjects were incredibly accurate for the repetition items, showing a high correct endorsement rate and a low misleading

endorsement rate, but were less accurate for the control and changed items. The next section explores how the news statements influenced later recall and change recollection.

4.2.4 Did the News Statements Influence Recall and Change Recollection?

How did the two types of news statements affect later recall and change recollection? Figure 8 shows recall as a function of item type and news statement. A 3 (item type: repetition, control, changed) x 2 (news statement: straight, flip-flop) within-subjects ANOVA revealed a main effect of item type, $F(2, 198) = 123.83, p < .001, \eta_p^2 = .57$, but did not reveal a main effect of news statement or a significant interaction (both $F_s \leq 1.64$). This was surprising—one might expect a significant interaction where news statements that suggested untrue information (e.g., a *straight* statement for a changed item) would lead to worse recall than an accurate news statement. To confirm that no such pattern existed, I compared recall after an accurate news statement (straight statements about repetition items and flip-flop statements about changed items; $M = .45, SEM = .02$) to recall after an inaccurate news statement (flip-flop statements about repetition items and straight statements about changed items; $M = .43, SEM = .02$), but failed to reveal a significant difference, $t(99) = 1.27, p = .208, d = 0.11$. In short, the flip-flop and straight news statements did not influence recall, either as a main effect or as an interaction with item type.

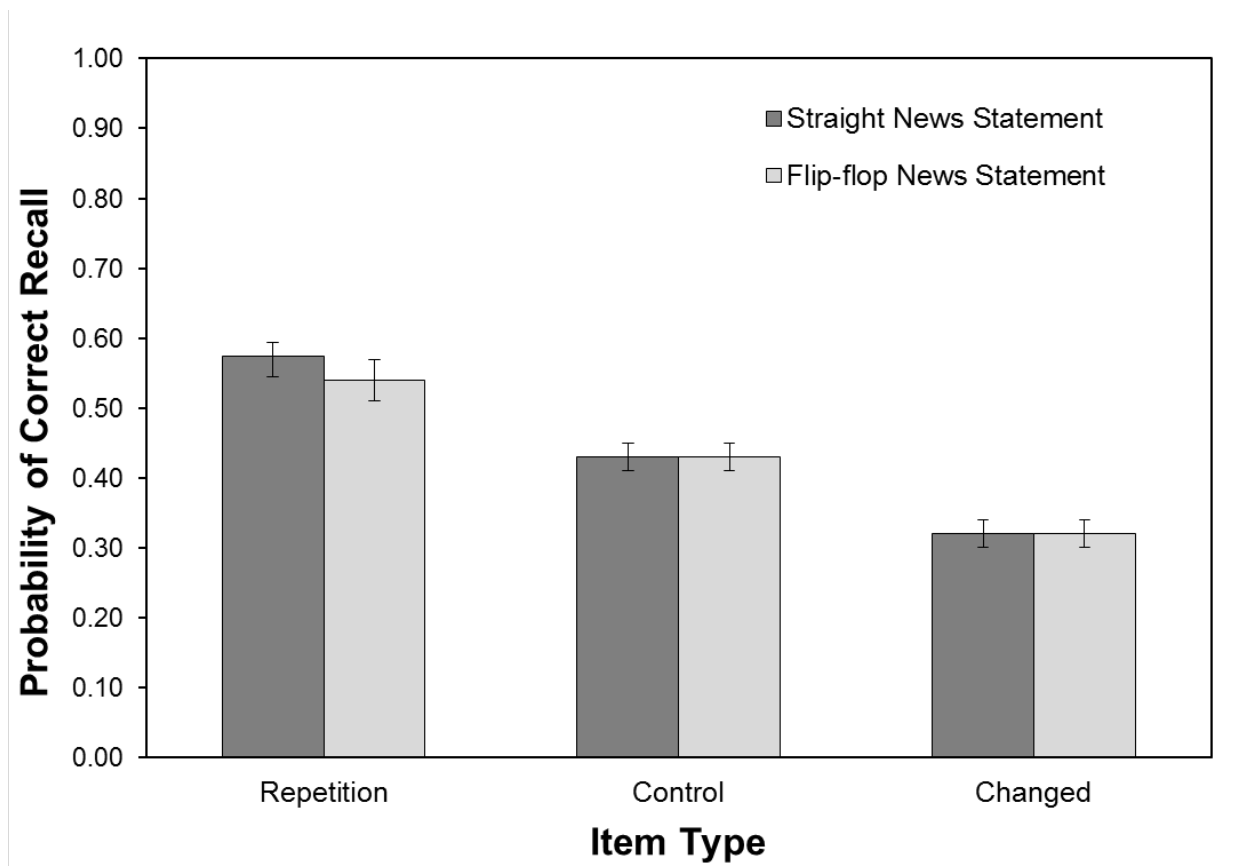


Figure 4.3: Probability of recall in Experiment 2 as a function of item type and news statement type.

In contrast, the news statements did affect change recollection. Table 11 shows the change recollection rates for the three different item types. A 3 (item type: repetition, control, changed) x 2 (news statement: straight, flip-flop) within-subjects ANOVA with change recollection as the dependent variable (and Greenhouse-Geisser adjustment for a violation of sphericity) revealed a main effect of news, $F(1, 99) = 7.78, p = .006, \eta_p^2 = .07$, indicating that across item types change recollection was higher after a flip-flop news statement ($M = .21, SEM = .02$) than after a straight news statement ($M = .19, SEM = .01$). The effect size is small, but this difference makes sense—after all, the flip-flop statements are suggested that the candidate changed positions. There was also a main effect of item type, $F(1.45, 143.72) = 89.06, p = .001, \eta_p^2 = .47$, with post-hoc tests revealing that change recollection was highest for the changed items, lowest for the repetition

items, in between for the control items (all $ps < .001$). The interaction was marginally significant, $F(1.89, 186.62) = 2.46, p = .092, \eta_p^2 = .02$. Change recollection accuracy (measured by d') was numerically higher for the flip-flop statements than the straight statements, but this difference did not reach significance, $t = 1.57$. Likewise, the difference in c between the flip-flop and straight statements was not significant, $t = 1.37$.

Table 4.4: Probabilities of Change Recollection and Signal Detection Measures in Experiment 2, Broken Down by News Statement Accuracy.

News Statement	Item Type			SDT Measure	
	Repetition	Control (FAs)	Changed (Hits)	d'	c
Straight	.10 (.01)	.17 (.02)	.30 (.02)	0.46 (0.07)	0.91 (0.07)
Flip-flop	.12 (.01)	.17 (.02)	.34 (.03)	0.59 (0.08)	0.86 (0.07)

Note: Numbers in parentheses represent standard error.

In sum, the news statements did not influence recall, but did influence change recollection. In particular, the flip-flop news statements made it more likely that subjects would report recollecting change at the final test for all item types (although looking at Table 11 suggests this effect is driven only by the repetition and changed items). This outcome can be thought of as a criterion shift – seeing an item tagged with a flip-flop statement increased the likelihood that subjects would remember change at the final test. C did not show a difference between the two conditions because only control items were used as false alarms in calculating c .

4.2.5 Did News Phase Accuracy Influence Later Recall and Change Recollection?

The main hypothesis for Experiment 2 was that correct verifications during the news phase would lead to accurate recall and change recollection during the final test. Figure 9 shows recall

for the different item types conditionalized on subject responses during the news phase. Responses were grouped into correct and incorrect bins (the *misinformation* and *miss* responses led to similar effects on recall, so those two types of response were grouped together for this analysis – see Appendix F for details). A 3 (item type: repetition, changed, control) x 2 (news verification response: correct, incorrect) within-subjects repeated measures ANOVA revealed a main effect of item type, $F(2, 144) = 33.56, p < .001, \eta_p^2 = .32$, a main effect of news response accuracy, $F(1, 72) = 15.14, p < .001, \eta_p^2 = .17$, and most importantly, a significant interaction, $F(2, 144) = 4.07, p = .019, \eta_p^2 = .05$. Simple effects analysis revealed that the interaction was driven mostly by the changed items: recall was roughly similar for the repetition and control items regardless of whether the news phase verification was correct or incorrect, $t(53) = 1.04, p = .30$ and $t(79) = 0.46, p = .65$ respectively, but recall for the changed items was better after a correct verification than an incorrect verification, $t(94) = 3.40, p < .001, d = 0.45$ (the differences in degrees of freedom were due to some subjects not having an observation in every cell). Thus, it appears that correctly responding during the news phase led to more accurate recall for the changed items, but not for the repetition or control items. This outcome fits with the initial predictions. A correct news verification suggests that subjects had access to a recursive trace and were able to accurately recollect change. For the changed items, this means subjects can recollect change to overcome any competition from the Debate 1 position and show a reduction in interference (as shown). For the repetition and control items, however, recollecting change is less important for recall, because there is no competing response.

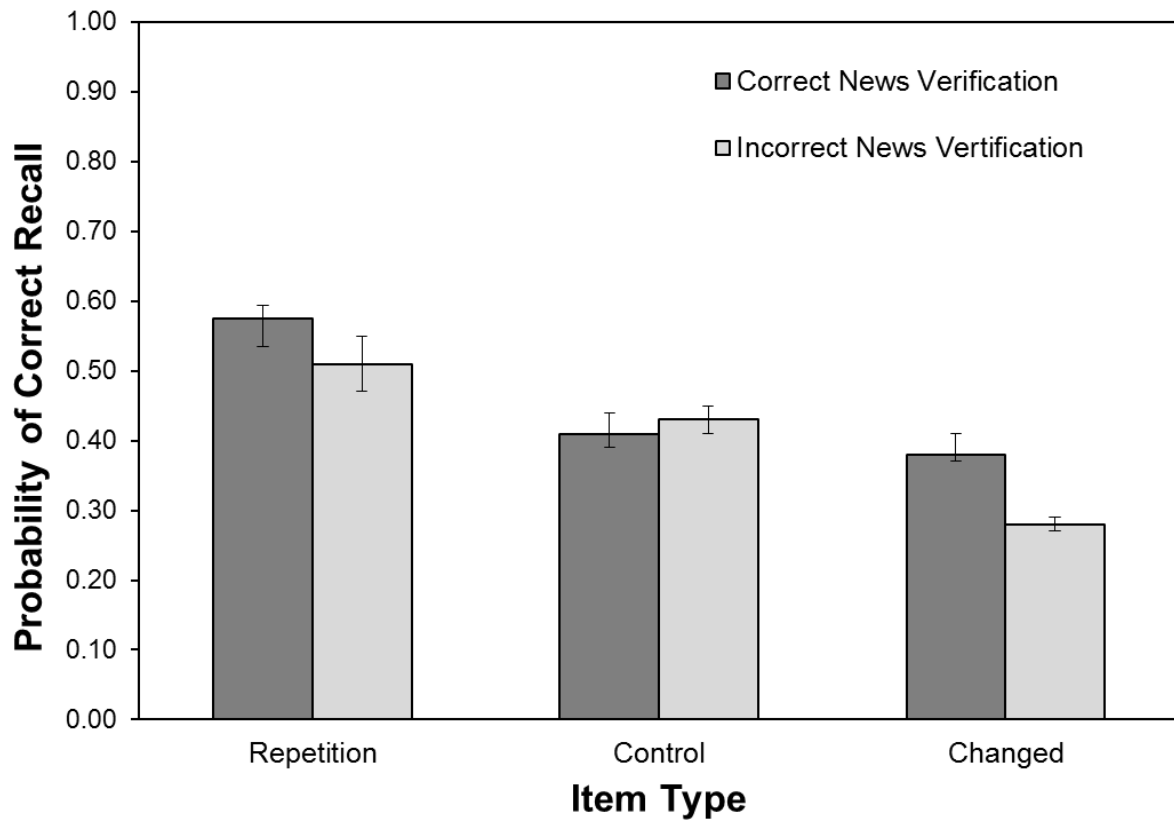


Figure 4.4: Recall by item type conditionalized on news phase verification accuracy.

News verification accuracy also influenced change recollection. Table 12 shows the change recollection rates for the different item types along with signal detection measures conditionalized on whether subjects were correct or incorrect in their news verification (see Appendix F for an analysis that breaks that the incorrect responses into misinformation endorsements and misses). A 3 (item type: repetition, control, changed) x 2 (news verification response: correct, incorrect) within-subjects ANOVA with change recollection as the dependent variable revealed a main effect of item type $F(2, 144) = 91.12, p < .001, \eta_p^2 = .56$, a main effect of news accuracy, $F(1, 72) = 22.23, p < .001, \eta_p^2 = .24$, and a significant interaction, $F(2, 144) = 68.09, p < .001, \eta_p^2 = .49$. Breaking down the interaction revealed that for the repetition items, change recollection was lower after a correct than incorrect news response, $t(94) = 3.37, p <$

.001, $d = 0.43$, that for the control items change recollection was similar after a correct or incorrect response, $t = 1.28$, and that for the changed items, change recollection was higher after an correct response than an incorrect response, $t(94) = 10.60, p < .001, d = 1.14$. Thus, accurately verifying a news statement led to a decrease in erroneous change recollection for the repetition items, an increase in accurate change recollection for the changed items, and no difference for the control items. This difference translated to more accurate change recollection, as measured by d' , after a correct news verification than an incorrect news verification, $t(74) = 7.82, p < .001, d = 1.19$. This outcome suggests that correctly verifying a statement about a candidate's past behavior allowed subjects to more accurately recollect change in the future.

Table 4.5: Probabilities of Change Recollection and Signal Detection Measures in Experiment 2, Broken Down by News Verification Accuracy

News Verification Response	Item Type			SDT Measure	
	Repetition	Control (FAs)	Changed (Hits)	d'	c
Correct	.09 (.01)	.17 (.03)	.54 (.03)	1.47 (0.16)	0.44 (0.09)
Incorrect	.17 (.02)	.18 (.02)	.23 (.02)	0.22 (0.05)	0.99 (0.06)

Note: Numbers in parentheses represent standard error.

4.3 Discussion

Experiment 2 was inspired by the real world question of whether a news organization can cause voters to misremember whether a candidate changed positions. There are several outcomes worth noting. First, Experiment 2 replicated Experiment 1 in showing that proactive interference occurred, but that recollecting change eliminated that interference effect. This finding was expected based on prior studies (e.g., Putnam et al., 2014; Wahlheim & Jacoby, 2013). In addition, Appendix F reports additional analyses examining the effects of partisanship on recall and change recollection along with measures of subject attitudes towards the candidates. In

general, the findings from Experiment 1 were replicated, although some small differences did occur.

Second, during the news verification task, subjects tended to remember the candidates as being consistent on most positions (see Table 10). They labeled a majority of the straight statements as true (regardless of item type) and labeled a majority of the flip-flop statements as false (regardless of item type). As a result, subjects were incredibly accurate in verifying the repetition items, and less so for the control and changed items. The tendency to remember items as consistent may have emerged because the candidates only changed positions on 1/3 of the items. So as a general rule of thumb it may have been advantageous for subjects to assume that the candidates were consistent during the news phase.

Third, the two different types of news statements (straight and flip-flop) did not influence later recall, but did have a small effect on change recollection. Subjects were more likely to remember change for an item following a flip-flop statement than a straight statement. This pattern can be interpreted as a misinformation effect (Loftus, Miller, & Burns, 1978). A subject sees an event, (a candidate changing positions) is exposed to some misleading information, (a news statement saying the candidate was consistent) and then is less likely to recollect change than if no misinformation had been presented. However, the effect is small, and may be driven largely by the changed items.

The fourth and most important result is the conditional analysis showing that subject responses to the news verification task influenced later recall and change recollection. As predicted, a correct verification during the news phase (compared to an incorrect verification) led to slightly higher recall for the repetition items (a numerical, but not significant difference), no difference for the

control items, and higher recall for the changed items (a significant difference). In addition, a correct verification during the news phase also led to increased change recollection accuracy. After a correct verification subjects had lower erroneous change recollection for repetition items and higher accurate change recollection for the changed items.

The results with both change recollection and recall are consistent with the recursive reminders framework. Correctly verifying the news statement required accessing a previous memory of change (or lack of it, depending on the item type). Another way to interpret Figure 9 (which shows recall conditionalized on news verification accuracy) is to think of the news verification responses as a measure of change recollection; rather than happening concurrently with recall, it occurred in a separate phase before the recall test. From this perspective, the recall results for the changed items in Figure 9 are remarkably similar to those shown in Figure 6, which uses the direct measure of change recollection: remembering that a change occurred lead to a decrease in proactive interference.

Previous work has shown that change detection is often correlated with later change recollection (Putnam et al. 2014) and that it is rare for subjects to recollect change without previously detecting it (Wahlheim & Jacoby, 2013). Given that the news verification task requires recollecting change, it is not surprising that accuracy here predicted future change recollection. Furthermore, recollecting change once may serve as retrieval practice for the change recollection measure at the final test (Roediger & Karpicke, 2006). One puzzle is why change recollection for the control items was similar regardless of whether the news verification was accurate or inaccurate. One possibility is that for the control items, responding *True* or *False* to either a straight or a flip-flop statement would be considered incorrect. Because subjects had a tendency to verify straight statements as true and reject flip-flop statements as false, these endorsements

for the control items are encouraging subjects to think of the control statements as repetition items, which may have reduced change recollection at the final test.

In sum, Experiment 2 showed that third party information from a news source can have a small impact on how subjects remember change. Although the statements on their own do not have a large impact on change recollection and recall, subjects who are able to verify the accuracy of those statements show better change recollection at the final test and increased recall for the changed items. The general concept fits with other work (Loftus, 1979; Putnam & Roediger, 2015) that shows that misinformation effects may not occur if subjects notice and remember the inconsistency between the original event and the misinformation. One issue to be aware of in Experiment 2 (and in Experiment 3) is that the news statements do affect the change recollection rate. As a result, the main analyses that rely on conditional analyses may be less sensitive, because change recollection is less accurate. Experiment 3 explores another way to manipulate change recollection – but rather than coming from a third party such as a news group, it comes from the candidates themselves.

Chapter 5: Experiment 3

Experiment 3 explored whether politicians can use specific language to make it more likely that someone would remember a change in position, and whether such language could affect later recall. The same overall design, materials, and procedure from Experiment 1 were used except that during Debate 2, each position statement was presented along with an introductory phrase. Some of the introductory phrases were neutral (e.g., “My stance on this topic is that...”) whereas other phrases encouraged subjects to think back to the first debate by suggesting change (e.g., “In the past I may have said something else, but I will argue that...”). The hypothesis was that the latter type of phrase (a *looking back* phrase) would encourage subjects to think back to Debate 1. In other words, the looking back statements might encourage covert change detection. Such a task may lead to more accurate change recollection at the final test, which would translate to better recall for the changed items.

This prediction is drawn from previous work that has shown different instructions can encourage the degree to which subjects look back to previous events, and that such instructions can impact later memory performance. For example, Putnam et al. (2014, Experiment 3) included an explicit measure of change detection in an experiment that used the political materials. During Debate 2 subjects were asked to push a button when they noticed that an item had changed from Debate 1. Including this change detection task led to more accurate change recollection at the final test which in turn led to proactive facilitation for the changed items (see right panel of Figure 1). Putnam et al. suggested that the explicit change detection task increased the likelihood that subjects would be reminded of a previous event, and that detecting change would lead to more accurate change recollection.

Although Hintzman (2004; 2011) has argued that most reminders are automatic, other work (Jacoby & Wahlheim, 2013; Jacoby et al., in press; Wahlheim et al., 2014) has shown that different instructions can encourage the degree to which subjects look back to previous events. In one experiment, for example, Wahlheim et al. (2014) showed that subjects who were instructed to notice repetitions between lists showed higher cued recall performance on pairs that were repeated across lists than subjects who were only instructed to notice repetitions within a list. Thus, different instructions may encourage the degree to which subjects engage in reminders.

In the current experiment, the looking back phrases may cue subjects to think back to the Debate 1 positions. Doing so should cause the Debate 1 position to become embedded in memory with the Debate 2 position, and accessing that recursive trace at the final test should lead to enhanced recall for the changed items. More specifically, the looking back phrases should lead to more accurate change recollection (both a decrease in the false alarm rate and an increase in the hit rate) compared to the neutral phrases. In addition, the looking back statements should lead to higher recall for the changed items compared to the neutral statements (recall for the control and repetition items should remain unaffected by the looking back phrases).

One caveat, however, is that previous research showing that instructions can influence reminders has exclusively used between-subjects manipulations (Jacoby & Wahlheim, 2013; Jacoby et al., in press; Wahlheim et al., 2014), whereas the current experiment uses a within-subjects manipulation. Will looking back cues work in a within-subjects design? Carryover effects are a concern here. Indeed, Loftus (1979) conducted a misinformation experiment where there was one blatant piece of misinformation (a bold inconsistency that everyone noticed) and four smaller pieces of information. Subjects never showed a misinformation effect for the blatant item (the argument being the change was so obvious people were able to reject it), but more

importantly, once subjects noticed the blatant misinformation, they were much more likely to notice the other pieces of misinformation. Thus, in the current experiment, the looking back statements may not lead to increased change recollection accuracy because subjects could be looking back to Debate 1 for the neutral items as well.

A related concern is that the looking back statements might not cue subjects to look back at all. Reminders occur spontaneously, but as noted above, can be encouraged through different types of instructions (see also Benjamin & Ross, 2010 for the role of reminders in transfer tasks), which suggests that the looking back statements might be effective. If, however, the looking back statements do not encourage reminders, then subjects may show a higher change recollection rate for all item types, because they would be using the information from the looking back phrase to inform their change recollection judgment. For the repetition and control items, this higher rate would be erroneous (because there was no change), meaning the looking back statements would be serving as misinformation.

Finally, one last concern is that subjects may not attend to the introductory phrases. To check whether subjects processed the looking back statements I conducted two pilot experiments (reported in more detail in Appendix G). The first pilot (3A) used the same design from Experiment 3, but included a change detection task during Debate 2 (subjects pushed a button when they notice a position had changed) and omitted the final test. The second pilot (3B) also used the same design from Experiment 3, but the final test was just the change recollection judgment (subjects did not do the cued recall task). The results from both experiments suggested that subjects did process the looking back tasks: change detection in pilot 3A and change recollection in 3B were both reliably greater in the looking back condition than the neutral condition. However, accuracy in change detection and change recollection was similar for both

the looking back and neutral statements. Thus, subjects may attend to the introductory phrases, but it may not lead to more accurate change recollection at the final test.

5.1 Method

5.1.1 Subjects

Fifty workers from Amazon's Mechanical Turk website participated. Thirteen additional subjects participated, but their data were omitted from all analyses: one for failing the instructional manipulation check, four for reporting low levels of political knowledge, seven for failing the political expertise quiz, and one for submitting the same response to every question. The pilot studies indicated that the looking back statements may increase erroneous change recollection, which would have made it difficult to interpret any correlations between change recollection and attitudes towards the candidates. As such, the candidate attitude questions were omitted from Experiment 3, meaning subjects were not recruited based on political orientation (hence the smaller sample size).

Thirty-one of the subjects were self-reported Democrats, eighteen were independents, and one was a Republican. The objective political measure revealed the sample to have 42 Democrats, three independents, and five Republicans. The average age was 34 ($SD = 9.14$) with ages ranging from 19 to 58. Subjects reported their level of education (10% indicated high school education or less, 38% indicated some college, and 52% indicated a bachelor's or graduate degree) and which state they were in (25 states were represented). Subjects reported a mean political expertise of 58.02 ($SD = 23.25$) on the self-report question and answered an average of 5.4 out of 8 questions on the political expertise quiz. All subjects were born in the U.S.

5.1.2 Design, Procedure, and Materials

The basic design, materials, and procedure from Experiment 1 were used with two differences. First, in interest of simplicity the voting and judgment measures were omitted from Experiment 3. After completing Debate 1 subjects moved straight to the distractor task, and after the recall test moved straight to the measurements of political orientation and expertise.

Second, during Debate 2 each statement was presented along with an introductory phrase. There were two types of phrases. One type was the *neutral* phrases, which were essentially verbal filler. For example, here is a neutral phrase paired with a statement about same sex marriage: “I’m glad this topic came up today; partners of the same sex should be recognized through marriage.” The other type was the *looking back* phrases that encouraged subjects to look back to Debate 1 by implying change. Here is a looking back phrase paired with the same statement: “In contrast to my previous position, I now believe that partners of the same sex should be recognized through marriage.” There were six different neutral phrases and six different looking back phrases.

The experiment was a 3 x 2 within-subjects design with *item type* (repetition, control, changed) crossed with *introductory phrase type* (neutral, looking back). A similar counterbalancing scheme to Experiment 2 was used. The 36 topics were split into three different groups (12 topics/group) and rotated through the different item types across subjects. Within each group half of the items presented the strong statement at Debate 2 and the other half presented the weak statement. Within each sub-group, half of the statements were paired with a *neutral* phrase, and half were paired with a *looking back* phrase. Each item appeared in each condition equally often across counterbalancing assignments.

5.2 Results

Responses were coded by two independent raters (Cohen's Kappa = .75), and I resolved any disagreements in coding. First I report the basic recall, change recollection, and conditionalized analyses, and then the results that take the introductory phrases into account.

5.2.1 Recall and Change Recollection

As expected, the repetition items ($M = .56$, $SEM = .03$) led to better recall than the control items ($M = .49$, $SEM = .03$), which in turn were better than the changed items ($M = .39$, $SEM = .03$), $F(2,98) = 29.53$, $p < .001$, $\eta_p^2 = .38$. Follow up comparisons revealed that all of the differences were significant, $t(49) \geq 3.23$, $ps \leq .001$, $ds \geq 0.51$. Table 13 shows change recollection for the control and changed items, along with the signal detection measures of d' and c . Subjects accurately recollected change: d' was greater than zero, $t(49) = 6.27$, $p = .001$, $d = 0.89$. Change recollection in Experiment 3 ($d' = 0.59$) was similar to Experiment 2 ($d' = 0.52$), $t < 1$.

Table 5.1: Probabilities of Change Recollection and Signal Detection Measures in Experiment 3

	Item Type		SDT Measure	
	Control (FAs)	Changed (Hits)	d'	c
Experiment 3				
Mean	.25	.41	0.59	0.55
SEM	.03	.03	0.09	0.09

5.2.2 Recall Conditionalized on Change Recollection

Figure 10 shows that conditionalizing recall for the changed items on change recollection replicated Experiments 1 and 2. When subjects did not recollect change, recall was worse for the changed items than it was for the control items, $t(49) = 7.40$, $p < .001$, $d = 1.18$. But when subjects did recollect change, recall for the changed items was equal to the control items, $t(49) = 1.37$, $p = .177$, $d = 0.16$, showing that recollecting change eliminated proactive interference.

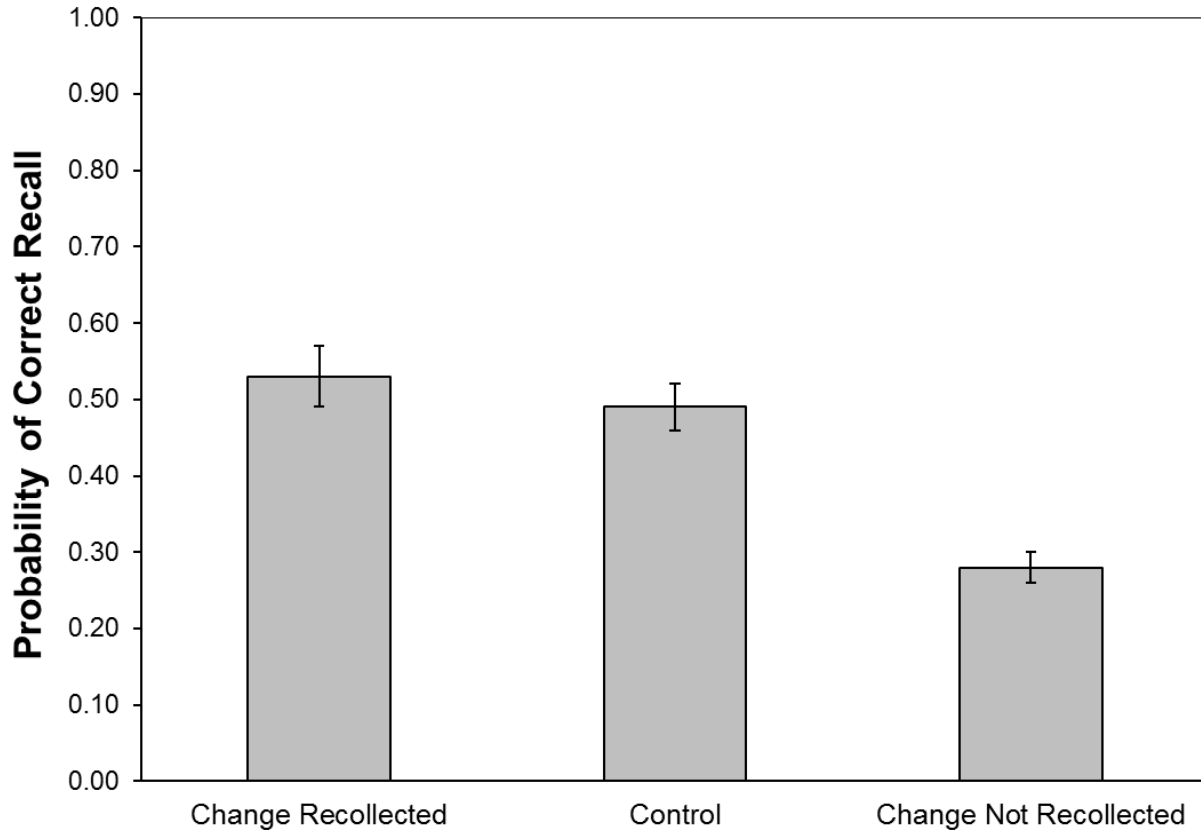


Figure 5.1: Cued recall for the changed positions conditionalized on change recollection, as compared to the control condition, Experiment 3. Error bars represent standard error of the means.

Table 14 shows the results of two hierarchical linear regression analyses done to examine whether change recollection had any effects above item and subject differences. The left column shows that item differences explained a significant proportion of variance. Change recollection and the interaction however, did not explain additional variance. This is a departure from previous experiments where change recollection did contribute to recall of the changed items above any item effects. One difference between Experiment 3 and the previous experiments is the inclusion of the looking back phrases, which may have increased erroneous change recollection (which would be reflected in the hits-false alarms rate). The right column of Table 14 shows that subject memory ability explained a significant proportion of variance and that

change recollection explained additional variance. The interaction did not improve performance. Thus, change recollection contributed to the recall of the changed items above any individual subject differences in memory ability.

Table 5.2: Proportion of Variance in Recall of Changed Positions Explained by Item Differences, General Memory Ability, and Change Recollection in Experiment 3.

	Unit of Analysis	
	Items	Subjects
Step 1		
Item Differences / General Memory	.29 ($p < .001$)	.33 ($p < .001$)
Step 2		
Change Recollection (Hits - FA)	.01 ($p = .301$)	.10 ($p = .007$)
Step 3		
Interaction	.002 ($p = .502$)	.008 ($p = .421$)

Note: Values displayed above are ΔR^2 on each step of the model computed at the item level collapsed across subjects (left) and at the subject level collapsed across items (right). “Item Differences” refers to item differences in control position recall performance, “General Memory” refers to individual differences in control position recall performance, and “Change Recollection (Hits - FA)” refers to individual differences in discriminability of change recollection for changed positions.

5.2.3 Recall as a Function of Introductory Phrase

Did the introductory phrases in Debate 2 affect recall during the final test? Figure 11 shows recall as a function of item type and introductory phrase type. A 3 (item type: repetition, control, changed) x 2 (introductory phrase type: neutral, looking back) repeated measures ANOVA with recall as the dependent variable revealed a main effect of item type, $F(2, 98) = 29.30, p < .001, \eta_p^2 = .37$, and a main effect of introductory phrase, $F(1, 49) = 5.67, p = .021, \eta_p^2 = .10$.

Surprisingly, for the main effect of introductory phrase, recall was higher in the neutral condition

($M = .50$, $SEM = .03$) than in the looking back condition ($M = .46$, $SEM = .03$). The interaction was not significant, $F < 1$.

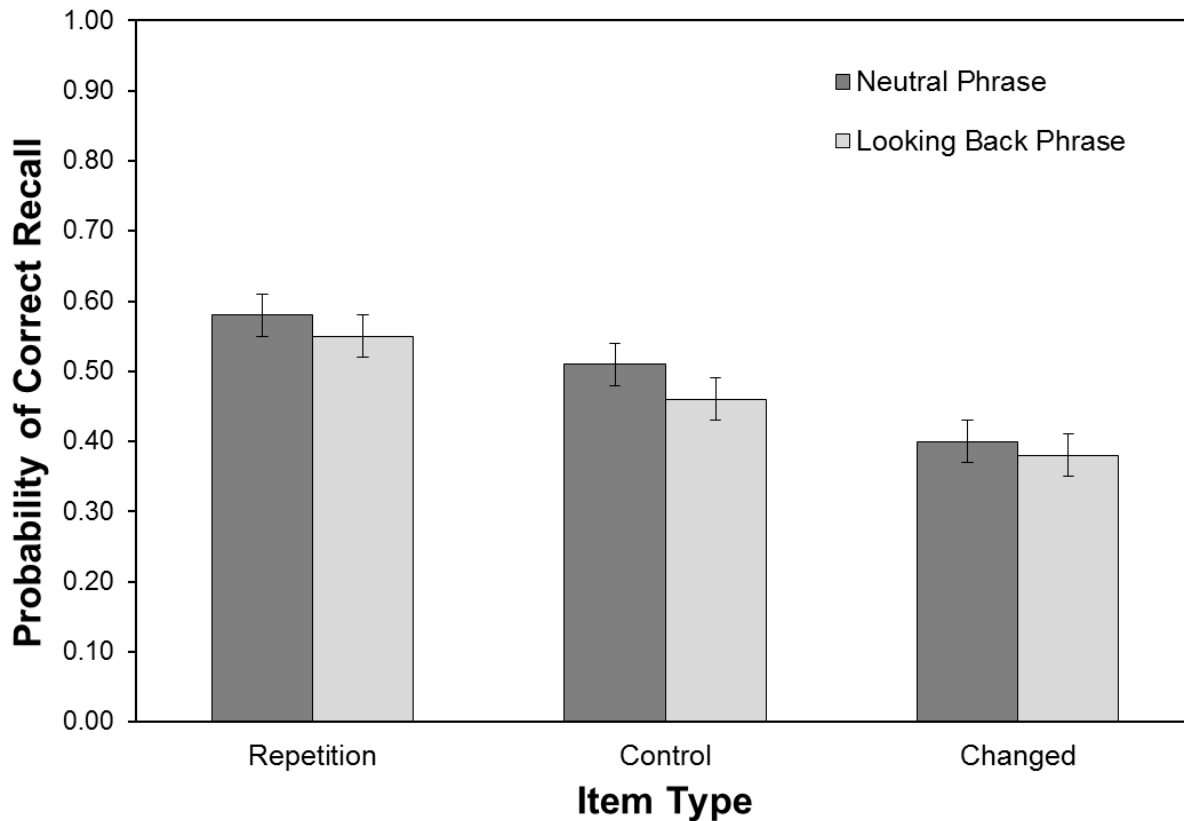


Figure 5.2: Probability of recall in Experiment 3 as a function of item type and introductory phrase type.

5.2.4 Change Recollection as a Function of Introductory Phrase

Did the introductory phrases influence the overall change recollection rates or change

recollection accuracy? Table 15 shows change recollection broken down by item type and

introductory phrase type. A 3 (item type: repetition, control, changed) x 2 (introductory phrase

type: neutral, looking back) repeated measures ANOVA with change recollection as the

dependent variable revealed a main effect of item type, $F(2, 98) = 47.48$, $p < .001$, $\eta_p^2 = .49$.

Collapsing across introductory phrase type, post-hoc analyses revealed that the repetition items

led to the lowest change recollection rate, followed by the control items, with changed items

leading to the highest change recollection rate (all $ps \leq .001$). There was also a main effect of introductory phrase, $F(1, 49) = 5.57, p = .022, \eta_p^2 = .10$, indicating that change recollection was higher in the looking back condition ($M = .30, SEM = .03$) than the neutral condition ($M = .26, SEM = .03$). The interaction was not significant, $F = 1.45$.

Table 5.3: Probabilities of Change Recollection and Signal Detection Measures in Experiment 3, Broken Down by Introductory Phrase Type

Introductory Phrase Type	Item Type			SDT Measure	
	Repetition	Control (FAs)	Changed (Hits)	d'	c
Neutral	.15 (.02)	.23 (.03)	.41 (.03)	0.61 (0.10)	0.61 (0.10)
Looking Back	.21 (.03)	.27 (.04)	.42 (.03)	0.59 (0.13)	0.54 (0.10)

Note: Numbers in parentheses represent standard error.

Perhaps the most important analysis in Experiment 3 was whether change recollection accuracy was better in the looking back condition than the neutral condition. Both conditions, however, had similar change recollection accuracy as measured by d' , $t(49) = 0.19, p = .848$. C was similar in both conditions as well, $t = 1.01$. Thus, it appears that the looking back statements did not lead to more accurate change recollection.

5.3 Discussion

The main question in Experiment 3 was whether the introductory phrases would affect later change recollection and recall. Overall, Experiment 3 replicated the main results seen in previous experiments: interference occurred, but recollecting change reduced that interference. The introductory phrases did influence change recollection and recall, but not quite as predicted.

First, across all item types the neutral statements led to better recall than the looking back statements. However, an interaction was predicted. If the looking back statements really encouraged subjects to think back to the Debate 1 positions (leading to covert change detection), then it should have led to higher recall for the changed items because subjects would have had more accurate change recollection. In contrast, the looking back statements should not have had an effect on recall for the control items (because there was nothing to look back to) and should have either enhanced recall or had no effect for the repetition items (a looking back statement, even though it implied change, may still have caused subjects to notice the repetition, which could lead to a benefit to recall regardless of change recollection). Instead, what occurred is an advantage in recall for the statements paired with the neutral statements over the looking back statements. One explanation for why this occurred is that—as discussed in more detail below—the looking back statements increased change recollection across all item types. Thus subjects may have been trying to remember changed statements (for the repetition and control items) that did not actually exist.

Similarly, the introductory phrases influenced change recollection, but not as predicted. One hypothesis, drawn from previous work showing that instructions could influence reminders (Jacoby & Wahlheim, 2013; Jacoby et al., in press; Wahlheim et al., 2014), was that the looking back statements would encourage subjects to think back to Debate 1 to confirm that the candidate actually changed. Doing so would lead to an increase in accurate change recollection for the changed items, and a decrease in erroneous change recollection for the control items and especially the repetition items, because looking back would confirm the repetition (even if the looking back statement suggested change). However, this pattern did not occur. Instead, the results showed that change recollection accuracy was similar in the neutral and looking back

conditions. The introductory statements did influence change recollection, but as a main effect – change recollection was higher after the looking back statements than the neutral statements. Although there was not a significant interaction between the introductory statements and the item types on change recollection, examining Table 15 clearly shows that the repetition and control items are driving the increase in change recollection. Thus, the looking back statements may only be increasing erroneous change recollection.

The looking back statements did not have the expected effect of encouraging contact with the Debate 1 positions. Instead, it appears that the looking back statements led to an increase in change recollection across all item types, but primarily as an increase in erroneous change recollection for the repetition and control items. Instead of looking back to Debate 1, subjects were using the suggestion of change to inform their change recollection judgments. In other words, subjects trusted the politicians when they said they changed positions. For the repetition and control items, the looking back phrases served as misinformation. Granted, in the real world, a politician would never claim to change positions when they were consistent.

If the looking back statements led to higher levels of change recollection, why did this not translate to increased recall for the changed items? Examining Table 15 reveals that the change recollection rate for the changed items was similar for the neutral and looking back conditions. Even if change recollection was higher in the looking back condition it may not translate to increased recall for the changed items, because the recollection of change might be artificially inflated. A subject who claims to remember change could actually be recollecting it, or he could be responding based on the suggestion of change from the looking back statements. Thus, an increased changed recollection rate would not necessarily translate to higher recall for the

changed items because the measure of change recollection would not represent true change recollection.

Chapter 6: General Discussion

This dissertation explored proactive interference with realistic political materials. Three experiments confirmed earlier work in showing that it is more difficult to remember a politician's most recent position if he recently changed positions on that issue—but that the difficulty in remembering can be reduced or eliminated if people recollect that a change occurred. This result is consistent with a recursive reminders framework that suggests that noticing and remembering change through reminders can reduce interference (e.g., Jacoby et al., in press). In addition, the three experiments addressed other questions. By using a politically diverse sample, Experiment 1 showed that political orientation had only a small (if any) effect on recall and no effect on change recollection. Experiment 1 also showed that subjects generally trusted, liked, and voted for the candidates from their own political party, but that remembering change (i.e., flip-flops) may lead to a more negative attitude towards those candidates. Experiment 2 showed that third party information (such as a news report) can influence people's ability to remember change—but only if people do not notice inconsistent statements. Accurately verifying news statements, by endorsing true ones and rejecting false ones, led to more accurate change recollection at the final test which in turn led to increased recall for the instances where candidates actually changed positions. Finally, Experiment 3 explored whether politicians can use specific language to influence a voter's ability to remember change. Claiming to change positions did not increase change recollection accuracy, as predicted, but rather served as a source of misleading information, biasing subjects to remember change even if no such change occurred.

In the remainder of the discussion I will discuss the theoretical implications of these experiments for the recursive reminders framework, the misinformation effect, and the relationship between memory and attitudes. In doing so I will suggest a few directions for future research. Then, I will note a few limitations of the current research and close with some suggestions for applications to politics.

6.1 Theoretical Implications

6.1.1 The Recursive Reminders Framework

The finding that recollecting change reduced proactive interference is consistent with previous work using both political materials (Putnam et al. 2014) and word pairs (Wahlheim & Jacoby, 2013). This result reinforces the importance of recursive reminders in understanding interference. As outlined in the introduction, detecting a change during a second event requires covert retrieval of the first event. Doing so results in both events becoming embedded in memory along with the experience of detecting the change. Later, recollecting this change (or accessing the recursive trace) should lead to accurate recall because doing so provides access to both the first and second positions along with their temporal order. The current experiments corroborate this interpretation with a revised set of realistic political materials and with a more diverse subject population.

Experiment 2 provided a new approach to measuring change recollection. Previous methods have included directly asking subjects if they remember change (e.g., Jacoby et al. 2013; Putnam et al. 2014) and by using a reminders report procedure, where subjects are asked to recall the more recent item, and then are asked to recall any other items that came to mind (Wahlheim & Jacoby, 2013). Both of these approaches occurred immediately after recalling the target item. In Experiment 2, change recollection was required to accurately respond in the news verification

task, which occurred in a separate phase before the recall test. Conditionalizing cued recall on the news verification task yielded a similar outcome to conditionalizing recall on the standard change recollection task: proactive interference was eliminated for the changed items when subjects recollected change. This finding suggests that the news verification task is a reliable method of measuring change recollection.

In addition, the news verification task may have provided retrieval practice for remembering change. Correctly verifying a news statement in Experiment 2 led to the most accurate change recollection at the final test across all experiments in this dissertation: d' was 1.47 whereas the next highest (in Experiment 1) was 0.86. This is interesting, because it suggests that retrieval practice can enhance memory for change. This concept could be explored in more detail in an experiment where some items are not exposed in the news phase, some items are presented without requiring subject verification, and some items are presented but require subjects to verify their accuracy. This design would parallel a testing effect experiment with a test, restudy, and no re-exposure condition (see Roediger, Putnam, & Smith, 2011 pp. 5-7 for a discussion of control conditions in testing effect experiments).

A key idea in the recursive reminders framework is that change recollection depends on earlier detection (Jacoby et al. 2013). Thus, any manipulation that can encourage accurate change detection should lead to more accurate change recollection. As noted earlier different instructions can guide subjects to look back to previous events in different ways (Jacoby & Wahlheim, 2013; Jacoby et al., in press; Wahlheim et al., 2014). The hypothesis in Experiment 3 was that the looking back phrases (e.g. “In contrast to my previous position I now believe...”) would encourage subjects to look back to the first debate. Doing so would increase covert change

detection which should lead to more accurate change recollection at the test. The looking back statements, however, did not improve change recollection accuracy, but instead, made it more likely that subjects would report remembering change for the repetition and control items.

Why did the looking back statements not lead to more accurate change recollection? It could have been a carryover effect—using a within-subjects manipulation might have led to subjects to look back for every item—but looking back in such a fashion should have led to more accurate change recollection, which did not occur (change recollection accuracy in Experiment 3 was similar to Experiment 2, and both were lower than Experiment 1). A more likely explanation is that subjects trusted the politicians. When the candidates said they changed positions, the subjects believed them, and were more likely to report remembering change at the final test. For the repetition and control items, this increase in the change recollection rate was erroneous, meaning the looking back statements were serving as misinformation.

Several variations of Experiment 3 could be interesting. For example, all of the introductory phrases could be accurate (e.g., the repetition items could be paired with consistent statements, e.g., “I’ve said this before and I’ll say it again...”). Or a between-subjects manipulation could be used where one group gets no introductory statements and another gets accurate introductory phrases. In both cases change recollection accuracy should be improved at the final test.

However it would be difficult to determine whether that increased accuracy is the result of actual change recollection or subjects relying on the suggestion of the introductory phrases. Pulling those two factors apart would be difficult. In sum, methodological problems in Experiment 3 prevented any new conclusions about the recursive reminders framework. However, the looking back statements did increase erroneous change recollection, which is an interesting finding from

the perspective of the misinformation effect (although it would be quite unusual for a politician to claim to change positions when he actually has been consistent).

6.1.2 The Misinformation Effect

Inaccurate information was presented to subjects in both Experiment 2 and Experiment 3. In both cases, the misleading information affected later change recollection. Subjects were more likely to recollect change after seeing a news statement that suggested change in Experiment 2 and after seeing an introductory statement from a politician where he claimed to change positions in Experiment 3. Change recollection accuracy did not improve, nor was recall for the changed items strongly affected. This pattern of results indicates that the suggestion of change led to an increase in the change recollection rate, but that the increase did not reflect true change recollection; subjects were reporting change on the basis of the suggestion from the news statement or introductory phrase rather than actually recollecting change.

Why did the potentially misleading information not affect recall? One answer is that both the news statements and the introductory phrases were about the candidate's consistency between debates, rather than his actual positions. If the news statements had made concrete but misleading suggestions about the candidate's current position (e.g., "Mike Shipman said he supported same sex marriage"), then recall might have suffered regardless of change recollection.

In both Experiment 2 and 3, the misleading information had small effect sizes ($\eta_p^2 = .07$ and $\eta_p^2 = .10$ respectively), raising the question of why the misleading statements did not have a larger effect on change recollection. Previous work has suggested that misinformation is ineffective when subjects notice the inconsistency between the original event and the post-event narrative

(Loftus, 1979; Tousignant et al. 1986). Along these lines, recent work on our lab (Putnam & Roediger, 2015) conducted a misinformation experiment with measures of change detection and change recollection. Recollecting change eliminated the misinformation effect, and sometimes even led to improved memory compared to the control condition. Thus, there is strong evidence that misinformation is ineffective when subjects notice the inconsistency *and* remember it at the final test. Experiment 2 provided further support for this view. Subjects who correctly verified the news statements showed incredibly accurate change recollection, and consequently showed improved recall for the changed items. In contrast, subjects who did not correctly verify the news statements showed poor change recollection accuracy.

The consequences of failing to reject misinformation can be seen clearly in the change recollection rates for the repetition items in Table F3 (in Appendix F). If a subject correctly verified a statement (e.g., responding *True* to the news statement “Mike Shipman was consistent in his views on Same Sex Marriage” in reference to a repetition item), the erroneous change recollection rate was .09. In contrast, if a subject endorsed a misleading statement (e.g., responding *True* to the news statement “Mike Shipman changed positions on Same Sex Marriage” in reference to a changed item), then the erroneous change recollection rate was .27. Thus, subjects were three times more likely to remember change when none occurred if they endorsed a misleading statement than if they correctly verified a statement. In sum, noticing and remembering a change may prevent misinformation.

6.1.3 Attitudes and Memory

Collecting data on the internet allowed me to recruit a politically diverse group of subjects. In doing so I was able to begin to address some questions that explored the relationship between memory and attitudes. The relationship is a complex one. Hastie and Park (1986) outlined five

different models from the social cognition literature, with some models suggesting that memories shape judgments and attitudes (as in the case of the availability heuristic) and other models suggesting that judgments and attitudes shape memories (e.g., through biased encoding or retrieval). The current experiments (in particular, Experiment 1) addressed both the question of whether attitudes could shape memory (through a congeniality effect) and whether memory can shape attitudes (via change recollection).

Experiment 1 showed a small congeniality effect in recall (Cohen's $d = 0.11$) suggesting that subjects had better memory for information that was consistent with their self-reported political orientation. This small effect size is not unusual when compared to previous work; a meta-analysis examining 70 congeniality effect experiments found an average effect size of 0.23 with a 95% confidence interval of 0.18 to 0.27, leading the authors to conclude that there may be no distinct advantage for remembering attitude-consistent information (Eagly et al. 1999). Along those lines, no congeniality effect was found in Experiment 2. Eagly et al. (2001) argued that congeniality effects might be hard to find because actively defending one's viewpoint against challenging, attitude-inconsistent information may be an effective encoding strategy. Indeed, recent research has failed to show strong evidence of a congeniality effect (Eagly, Kulesa, Brannon, Shaw, Huston-Comeaux, 2000). Eagly et al. (2000), however, did show that many other variables, such as attitude strength, may moderate how well relevant information is remembered. Thus, the current experiments do not contribute much to an understanding of the congeniality effect. The hypothesis that change recollection might also be affected by the congruency between a subject's attitudes and the current topic was rejected in both Experiment 1 and Experiment 2.

As Eagly et al. (2000) noted, many variables could potentially affect recall for attitudinally relevant information, including how strong someone's attitudes are. A future project could have subjects rate how important each political topic is to them personally, or rate how familiar they are with the issue. The expectation is that subjects would have better recall and change recollection for the topics for which they hold strong opinions or know well. Such a finding would provide a concrete example of how individuals differ in change recollection, something previous work has suggested. (Jacoby & Wahlheim, 2013).

On a side note, one curious pattern that occurred in both Experiment 1 and 2 was that subjects had better recall for the changed positions when the candidates were shifting *away* from the subject's self-reported political orientation, rather than *towards* the subject's political orientation. This pattern occurred despite change recollection being equal in both directions. One possibility is that subjects felt betrayed by the politicians when the candidates moved away from the subjects' own views, which may have made those positions more memorable.

Experiment 1 also addressed the question of whether memory can shape attitudes. In both Experiment 1 and 2 subjects preferred the candidates from their own parties. They were more likely to vote for them and provided higher rankings of trustworthiness and likability. Despite that strong tendency, there was also some indication that subjects' experiences with the candidates during Debate 2 shaped the post-test judgments. The change recollection rates for each candidate were correlated with the estimates of how candidates would change in the future, suggesting that noticing and remembering change shaped the future change judgment. In addition, in Experiment 1 change recollection and the candidates' trustworthiness and likability ratings (either in the second judgments or in the change between the two judgments) were

negatively correlated, indicating that remembering more change was associated with a more negative attitude towards the candidates. (Experiment 2 did not show that pattern, but the news statements may have added noise to the change recollection judgments.) Thus, subjects did appear to base their judgments on their memories of the candidate, although the exact mechanism is unclear. Subjects could be updating an on-line tally of how often a candidate changes position and then retrieving the most recent tally to make their judgment; or subjects could be thinking back to the actual debates and making an estimate of how often the candidate changed positions.

6.2 Potential Limitations

6.2.1 Alternatives to the Recursive Reminders Framework

The recursive reminders framework does have some similarity to other models of memory.

Karpicke, Lehman, and Rue (2014), for example, recently proposed an episodic context account of retrieval practice, arguing that the memory benefits of retrieval practice accrue from updating prior contexts. Their account was heavily influenced by a temporal context model approach to memory (Estes, 1955; Howard & Kahana, 2002; Kahana, Howard, & Polyn, 2008; Siegel & Kahana, 2014). Such a theory suggests that people encode information about events as it occurs in a given context. Context changes slowly over time, and retrieving information involves partially re-instating the prior context. Doing so results in an updating process where elements of both the original and the new context are combined. On a later test, retrieval is easier because the updated context provides a means for restricting a search set.

The episodic context account of retrieval practice (and the temporal context model) bear some similarity to the recursive reminders framework in that successful recall is the result of connecting two separate events in memory. They are different in that the recursive reminders

framework emphasizes the target events (or items) whereas the episodic context account emphasizes the context in which those events occur. The engine in the recursive reminders framework is spontaneous study-phase retrieval whereas the engine in the episodic context account is intentional retrieval practice. However, the two approaches can co-exist: Reminders are a form of covert retrieval, so the episodic contextual account may partially explain why reminders enhance memory.

6.2.2 Item Selection Effects

As outlined in the introduction, item selection effects are a concern when using conditional analyses. A critic might argue that the reduction in proactive interference seen when change is recollected is not driven by change recollection, but rather is an artifact of using conditional analyses. Some items may be easier to recall than others, and the conditional analyses may simply collect those items (for which both recall and change recollection is accurate). However, previous experiments have shown that reminders can lead to improved memory without the need to rely on conditionalized analyses, reducing the concern of item selection effects (Jacoby, 1974; 2013; Jacoby & Wahlheim, 2013; Jacoby et al., in press; Wahlheim et al., 2014). In addition, the current project included hierarchical regression analyses that suggested that change recollection often explained unique variance in predicting recall after controlling for item effects and differences in subject memory abilities, replicating previous work (Jacoby & Wahlheim, 2013; Putnam et al. 2014; Wahlheim & Jacoby, 2013). Experiment 3 did not show a unique contribution of change recollection, but the looking back statements in that experiment may have biased the change recollection measure. Taken together, these results suggest that reminders are driving the reduction in proactive interference, rather than any item selection effects.

A more nuanced concern is that item selection effects do exist, but ones that are unique to individual subjects. Rather than some items being more memorable overall, each subject could find some items more memorable than others, raising a similar concern that the reduction in proactive interference is due to the change recollection measure collecting the items that are easier to remember rather than change recollection driving the reduction in interference. Addressing this question statistically, however, would require knowing the baseline level of recall for each item for each subject (recall for a control item) and comparing that to recall for a changed item when change or was not recollected. This comparison is not possible in the current experiments because items are never in both the control and the changed condition for a single subject. This concern could be ameliorated in an experiment that uses a between-subjects design where some subjects are encouraged to look back to the earlier debate and others are not. If the looking back group shows enhanced recall for the changed items compared to the control group, this would suggest that reminders are reducing interference without the need to rely on conditional analyses. Such a finding was reported in Putnam et al. (2014, Experiment 3), where a group that was given an explicit change detection task showed better recall for the changed items than a control group. Thus, it is unlikely that a subject by item interaction is driving the reduction in proactive interference.

6.2.3 Artificial Nature of the Paradigm

Using political materials to explore interference is exciting for many reasons - it provides a conceptual replication of previous work, suggests new directions for research, and even addresses new questions that cannot be answered with basic materials. However, straddling the line between basic and applied research does raise some challenges, particularly with applications to politics.

For example, in the current experiments the candidates change positions more often than would be expected in a real campaign. Currently, the candidates change positions on 1/3 of the issues. This rate was chosen to provide a large number of observations and to ensure a balanced design. Future work, however, could and should explore how people notice and remember a politician's single change in position amidst an otherwise consistent platform.

A related criticism is that the changes the politicians make in the experiments are not true flip-flops. Rather, the candidates are shifting within their own camp and never changing to the opposite side. Although complete flip-flops are what often grab headlines (e.g., John Kerry's comment on Iraq; Hummel, 2010), smaller changes in position—perhaps changing the implementation of a policy without changing the underlying principle—are perhaps more realistic in politics today. It is possible that complete flip-flops are more memorable or noteworthy, and almost certainly they would have larger impact on voter attitudes. Full flip-flops are certainly worth exploring in future research.

Another concern is that the assessment of political orientation assumed a single continuum, ranging from liberal Democrats on the left to conservative Republicans on the right. However, some political affiliations do not fit neatly into that continuum. Libertarians, for example, typically endorse liberal views on social issues and conservative views on fiscal issues. A more nuanced approach to measuring political orientation could be used (e.g., by having subjects report a separate category for social and economic issues) but most Americans can easily place themselves in the left-right continuum (Jost, 2006).

One final challenge is the number of ways a voter can interpret a politician's change in position. To provide a concrete example, in 2012 President Obama reversed his position on same sex

marriage, saying he now supported it. Voters could have reacted negatively, either because they did not support gay marriage *or* because Obama's change made him seem untrustworthy and opportunistic. In contrast, voters could have reacted positively, either because they supported same sex marriage or because Obama's change made him appear open-minded and in touch with the rest of society. Indeed, how a politician frames a change may influence how it is interpreted. A recent newspaper article noted that many politicians (including Obama) have started using the term "evolve" to describe a change in position, both because it implies personal growth and because it suggests external forces are responsible for the change (Leibovich, 2015). Thus, any researcher who wishes to address voter reactions to flip-flopping needs to consider many factors, such as the voter's views on the issue, what position the politician holds, and how the change is framed. The current experiments had politicians changing positions on a variety of issues, and framed them all similarly, but despite that, managed to show some negative correlations between change recollection and attitudes towards the candidates.

6.3 Applications to Politics

What implications do the results of these experiments have for politics? Despite being grounded in an artificial paradigm there might still be a few suggestions for voters and politicians.

Political scientists have long discussed how voters make decisions about who to vote for. One theory suggests an on-line model where voters continually update their evaluation of candidates as they process campaign information. When it is time to vote, people base their decision on the summary evaluation of the candidates rather than thinking back to the campaign (Lodge, Steenbergen, & Brau, 1995). Another theory suggests a memory-based model where people do think back to campaign events to shape their voting choices (Redlawsk, 2001).

In either case, the current results may be relevant. If voters are thinking back to campaign events, then clearly reminders are important—any event that is well remembered (perhaps because of noticing a repetition or detecting change) may have a larger impact on a voting decision than something that is less well remembered. If voters are updating an on-line tally it could still be beneficial to notice and remember the change. In particular, campaigns are filled with half-truths, misconstrued stories, and political spin. Having a clear recollection of when a politician has changed position and when he has not will help prevent a voter from falling prey to misinformation. Also, voters may have certain issues that they feel strongly about; if a politician changes position on that issue it could have a large impact on how that voter views the politician.

Politicians may also learn something from these experiments. In most cases, subjects continued to support the candidates from their own party, despite the politicians changing positions quite often. This suggests that politicians should have some flexibility to change positions (at least within their own camp), but doing so too often may result in some negative consequences, such as developing a reputation as a flip-flopper. In many ways the results here affirm what political scientists have previously believed: changing positions is often advantageous, but doing it too quickly or too often can backfire.

If a politician is changing position and wants people to remember what his or her new position is, then it would be beneficial for people to notice and remember that the change occurred. Having third party sources repeat the new position and reaffirm that it is a change may help. Other research suggests that the politician may want to provide a justification for the change, to reduce any negative perceptions associated with changing (Levendusky & Horowitz, 2012). Simply

acknowledging the change would make it more likely that people would remember the change in the future, but it is unclear whether that would translate to better memory for the new position.

In contrast, how could a politician change positions without having anyone remember it in the future? Saying that he has always held the same position might work, as would having a third party reinforce his consistency. But in either case such tactics might backfire if subjects have a clear memory for change. Returning to the quote from the introduction, in response to the Etch A Sketch comment, Romney said “The issues I’m running on will be exactly the same. I’m running as a conservative Republican, and I was a conservative Republican governor ... The policies and positions are the same,” (Parker, 2012). Despite claiming to be consistent, Romney was unable to avoid being labeled a flip-flopper.

6.4 Concluding Comments

The three experiments in this dissertation explored proactive interference with realistic political materials. Neisser (1976) might have approved. He wrote:

“...cognitive psychologists must make a greater effort to understand cognition as it occurs in the ordinary environment and in the context of natural purposeful activity. This would not mean an end to laboratory experiments, but a commitment to the study of variables that are ecologically important rather than those that are easily manageable” (p. 7).

Using the political materials provided a way to explore proactive interference with in a context that at least has some parallel to the real world. Engaged voters really do want to remember what political candidates claim to believe. Using the political materials also provided more support for the recursive reminders framework and allowed the examination of new questions that would not be possible with basic materials, such as understanding the relationship between attitudes and memory. Flip-flopping accusations will likely continue to be a part of American politics for a

long time, providing a rich arena for exploring questions about memory, attitudes, and decision-making. This research is one step towards understanding the cognitive and social implications of noticing and remembering changes in position.

References

- Anderson, J. R. & Neely, J. H. (1996). Interference and inhibition in memory retrieval. In E. L. Bjork & R. A. Bjork (Eds.), *Memory. Handbook of perception and cognition (2nd Ed.)*, pp. 237-313. San Diego, CA: Academic Press.
- Appleton-Knapp, S. L., Bjork, R. A., & Wickens, T. D. (2005). Examining the spacing effect in advertising: Encoding variability, retrieval processes, and their interaction. *Journal of Consumer Research*, *32*, 266–276.
- Asch, S. E. (1969). A reformulation of the problem of associations. *American Psychologist*, *24*, 92–102.
- Barden, J., Rucker, D. D., Petty, R. E., & Rios, K. (2014). Order of actions mitigates hypocrisy judgments for ingroup more than outgroup members. *Group Processes & Intergroup Relations*, *17*, 590–601. doi:10.1177/1368430213510192
- Barnes, J. M., & Underwood, B. J. (1959). “Fate” of first-list associations in transfer theory. *Journal of Experimental Psychology*, *58*, 97–105.
- Benjamin, A. S., & Ross, B. H. (2010). The causes and consequences of reminding. In A. S. Benjamin (Ed.), *Successful remembering and successful forgetting: A Festschrift in honor of Robert A. Bjork* (pp. 71–87). New York, NY: Psychology Press.
- Benjamin, A. S., & Tullis, J. (2010). What makes distributed practice effective? *Cognitive Psychology*, *61*, 228–247. doi:10.1016/j.cogpsych.2010.05.004
- Bernhardt, M. D., & Ingberman, D. E. (1985). Candidate reputations and the “incumbency effect.” *Journal of Public Economics*, *27*, 47–67.
- Bruce, D., & Weaver, G. E. (1973). Retroactive facilitation in short-term retention of minimally learned paired associates. *Journal of Experimental Psychology*, *100*, 9–17.

Bowers, B. (2012). President Barack Obama's shifting stance on gay marriage. PolitiFact.com.

Retrieved from www.politifact.com

Cialdini, R. B. (1993). *Influence*. New York, NY: Quill William Morrow.

Croco, S. E., & Gartner, S. S. (2014). Flip-Flops and High Heels: An Experimental Analysis of Elite Position Change and Gender on Wartime Public Support. *International Interactions*, 40, 1–24. doi:10.1080/03050629.2013.863195

Crowder, R. G. (1976). *Principles of Learning and Memory*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Dickerson, J. (2008, June 20). The flip-flop brothers: Both McCain and Obama wobbled. Who will pay the price? *Slate Magazine*. Retrieved from www.slate.com

Eagly, A. H., Chen, S., Chaiken, S., & Shaw-Barnes, K. (1999). The impact of attitudes on memory: An affair to remember. *Psychological Bulletin*, 125, 64-89.

Eagly, A. H., Kulesa, P., Brannon, L. A., Shaw, K., & Hutson-Comeaux, S. (2000). Why Counterattitudinal Messages Are as Memorable as Proattitudinal Messages: The Importance of Active Defense Against Attack. *Personality and Social Psychology Bulletin*, 26, 1392–1408. doi:10.1177/0146167200263007

Eagly, A. H., Kulesa, P., Chen, S., & Chaiken, S. (2001). Do attitudes affect memory? Tests of the congeniality hypothesis. *Current Directions in Psychological Science*, 10, 5–9.

Estes, W. K. (1955). Statistical theory of spontaneous recovery and regression. *The Psychological Review*, 62, 145–154.

Field, A., Miles, J., & Field, Z. (2012). *Discovering Statistics Using R*. Thousand Oaks, CA: Sage.

- Frenda, S. J., Knowles, E. D., Saletan, W., & Loftus, E. F. (2013). False memories of fabricated political events. *Journal of Experimental Social Psychology, 49*, 280–286.
doi:10.1016/j.jesp.2012.10.013
- Geraghty, J. (2008). Measuring the flip-flops. National Review Online. Retrieved from
www.nationalreview.com
- Hastie, R., & Park, B. (1986). The relationship between memory and judgment depends on whether the judgment task is memory-based or on-line. *Psychological Review, 93*, 258-268.
- Hintzman, D. L. (2004). Judgment of frequency versus recognition confidence: Repetition and recursive reminding. *Memory and Cognition, 32*, 336–350.
- Hintzman, D. L. (2010). How does repetition affect memory? Evidence from judgments of recency. *Memory and Cognition, 38*, 102–115. doi:10.3758/MC.38.1.102
- Hintzman, D. L. (2011). Research strategy in the study of memory: Fads, fallacies, and the search for the “coordinates of truth.” *Perspectives on Psychological Science, 6*, 253–271.
doi:10.1177/1745691611406924
- Howard, M. W., & Kahana, M. J. (2002). A Distributed Representation of Temporal Context. *Journal of Mathematical Psychology, 46*, 269–299. doi:10.1006/jmps.2001.1388
- Hummel, P. (2010). Flip-flopping from primaries to general elections. *Journal of Public Economics, 94*, 1020–1027. doi:10.1016/j.jpubeco.2010.08.006
- Jacoby, L. L. (1974). The role of mental contiguity in memory: Registration and retrieval effects. *Journal of Verbal Learning and Verbal Behavior, 13*, 483–496.
- Jacoby, L. L., & Wahlheim, C. N. (2013). On the importance of looking back: The role of recursive reminders in recency judgments and cued recall. *Memory and Cognition, 41*, 625–637. doi:10.3758/s13421-013-0298-5

- Jacoby, L. L., Wahlheim, C. N., & Coane, J. H. (2010). Test-enhanced learning of natural concepts: Effects on recognition memory, classification, and metacognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 1441–1451.
doi:10.1037/a0020636
- Jacoby, L. L., Wahlheim, C. N., & Kelley, C. M. (in press). Memory consequences of looking back to notice change: Retroactive and proactive facilitation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.
- Jacoby, L. L., Wahlheim, C. N., & Yonelinas, A. P. (2013). The role of detection and recollection of change in list discrimination. *Memory and Cognition*, *41*, 638–649.
doi:10.3758/s13421-013-0313-x
- Jost, J. T. (2006). The end of the end of ideology. *American Psychologist*, *61*, 651–670.
doi:10.1037/0003-066X.61.7.651
- Kahana, M. J., Howard, M. W., & Polyn, S. M. (2008). Associative retrieval processes in episodic memory. In H. L. Roediger, *Cognitive psychology of memory. Vol. 2 of Learning and memory: A comprehensive reference, 4 vols. (J. Byrne, Editor)* (pp. 467–490). Oxford: Elsevier.
- Karpicke, J. D., Lehman, M., & Aue, W. R. (2014). Retrieval-Based Learning. In *Psychology of Learning and Motivation* (Vol. 61, pp. 237–284). Elsevier. doi:10.1016/B978-0-12-800283-4.00007-1
- Lambert, A. J., Scherer, L. D., Schott, J. P., Olson, K. R., Andrews, R. K., O'Brien, T. C., & Zisser, A. R. (2010). Rally effects, threat, and attitude change: An integrative approach to understanding the role of emotion. *Journal of Personality and Social Psychology*, *98*, 886–903. doi:10.1037/a0019086

- Leibovich, M. (2015, March 10). You and I Change Our Minds. Politicians 'Evolve'. *The New York Times*. Retrieved from www.nytimes.com
- Levendusky, M. S., & Horowitz, M. C. (2012). When Backing Down Is the Right Decision: Partisanship, New Information, and Audience Costs. *The Journal of Politics*, *74*, 323–338. doi:10.1017/S002238161100154X
- Levine, J., & Murphy, G. (1943). The learning and forgetting of controversial material. *The Journal of Abnormal and Social Psychology*, *38*, 507-517.
- Lodge, M., Steenbergen, M. R., & Brau, S. (1995). The responsive voter: Campaign information and the dynamics of candidate evaluation. *The American Political Science Review*, *89*, 309–326.
- Loftus, E. F., Miller, D., & Burns, H. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology*, *4*, 19–31.
- Loftus, E. F. (1979). Reactions to blatantly contradictory information. *Memory and Cognition*, *7*, 368–374.
- Martin, E. (1968). Stimulus meaningfulness and paired-associate transfer: An encoding variability hypothesis. *Psychological Review*, *75*, 421–441.
- Martin, R. B., & Dean, S. J. (1964). Implicit and explicit mediation in paired-associate learning. *Journal of Experimental Psychology*, *68*, 21.
- Mason, W., & Suri, S. (2011). Conducting behavioral research on Amazon's Mechanical Turk. *Behavior Research Methods*, *44*, 1–23. doi:10.3758/s13428-011-0124-6
- McCaul, K. D., Ployhart, R. E., Hinsz, V. B., & McCaul, H. S. (1995). Appraisals of a consistent versus a similar politician: Voter preferences and intuitive judgments. *Journal of Personality and Social Psychology*, *68*, 292.

- Melton, A. W. & Irwin, J. M. (1940). The influence of degree of interpolated learning on retroactive inhibition and the overt transfer of specific responses. *American Journal of Psychology*, 3, 173-203.
- Neisser, U. (1976). *Cognition and Reality*. San Francisco: W H Freeman & Company.
- New York Times (October 23, 1890) Mighty Cheers for Scott: A great meeting in Cooper Union last night. *The New York Times*. Retrieved from www.nytimes.com
- Nyhan, B., & Reifler, J. (2010). When corrections fail: The persistence of political misperceptions. *Political Behavior*, 32, 303–330. doi:10.1007/s11109-010-912-2
- Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology*, 45, 867–872. doi:10.1016/j.jesp.2009.03.009
- Parker, A. (2012, March 21). Romney tries to move spotlight off Etch a Sketch. *The New York Times*. Retrieved from www.nytimes.com
- Putnam, A. L., & Roediger, H. L. (2013). Does response mode affect amount recalled or the magnitude of the testing effect? *Memory and Cognition*, 41, 36–48. doi:10.3758/s13421-012-0245-x
- Putnam, A. L., & Roediger, H. L. (2015). *Being reminded of misinformation: Change detection and recollection are both important for reducing false memories*. Manuscript in preparation.
- Putnam, A. L., Wahlheim, C. N., & Jacoby, L. L. (2014). Memory for flip-flopping: Detection and recollection of political contradictions. *Memory and Cognition*, 42, 1198–1210. doi:10.3758/s13421-014-0419-9
- Redlawsk, D. P. (2001). You Must Remember This: A Test of the On-Line Model of Voting. *The Journal of Politics*, 63, 29–58.

- Robbins, D., & Bray, J. F. (1974). The spacing effect and the A-B, A-C paradigm: Evidence for retroactive facilitation. *Journal of Experimental Psychology*, *103*, 420–425.
- Roediger, H. L., & Karpicke, J. D. (2006). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science*, *1*, 181–210.
- Roediger, H. L., Putnam, A. L., & Smith, M. (2011). Ten Benefits of Testing and Their Applications to Educational Practice. In J. Mestre & B. H. Ross, *Psychology of learning and motivation: Advances in research and theory* (Vol. 55, pp. 1–36). Oxford: Elsevier.
- Rosenberg, S., Nelson, C., & Vivekananthan, P. S. (1968). A multidimensional approach to the structure of personality impressions. *Journal of Personality and Social Psychology; Journal of Personality and Social Psychology*, *9*, 283-294.
- Safire, W. (1988, March 13). On language; Phantom of the phrases. *The New York Times*. Retrieved from www.nytimes.com
- Siegel, L. L., & Kahana, M. J. (2014). A retrieved context account of spacing and repetition effects in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 755–764. doi:10.1037/a0035585
- Tavits, M. (2007). Principle vs. pragmatism: Policy shifts and political competition. *American Journal of Political Science*, *51*, 151–165.
- Tousignant, J. P., Hall, D., & Loftus, E. F. (1986). Discrepancy detection and vulnerability to misleading postevent information. *Memory and Cognition*, *14*, 329–338.
- Trumble, D. (2012, November 7). Romney lost because he was too damn good at running... against himself. *The Huffington Post*. Retrieved from www.huffingtonpost.com
- Tzeng, O. J., & Cotton, B. (1980). A study-phase retrieval model of temporal coding. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 705-716.

- Wahlheim, C. N., & Jacoby, L. L. (2013). Remembering change: The critical role of recursive reminders in proactive effects of memory. *Memory and Cognition, 41*, 1–15.
doi:10.3758/s13421-012-0246-9
- Wahlheim, C. N., Maddox, G. B., & Jacoby, L. L. (2014). The role of reminding in the effects of spaced repetitions on cued recall: Sufficient but not necessary. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*, 94–105. doi:10.1037/a0034055
- Weinstein, Y. (2012). Flash programming for the social & behavioral sciences: A simple guide to sophisticated online surveys and experiments. Thousand Oaks, California: SAGE.
- Westen, D., Blagov, P. S., Harenski, K., Kilts, C., & Hamann, S. (2006). Neural Bases of Motivated Reasoning: An fMRI Study of Emotional Constraints on Partisan Political Judgment in the 2004 U.S. Presidential Election. *Journal of Cognitive Neuroscience, 18*, 1947-1958.
- Zell, E., & Bernstein, M. J. (2014). You may think you're right ... Young adults are more liberal than they realize. *Social Psychological and Personality Science, 5*(3), 326–333.
doi:10.1177/19485506134928251947–1958.

Appendix A: Political Materials

Topic	Liberal Strong	Liberal Weak	Conservative Weak	Conservative Strong
Abortion	Abortions should be available and legal to women of all ages.	Abortions should be safe, legal, and rare.	Abortions should be available only in cases of rape.	Abortions should never be available to anybody.
Affirmative Action	When making hiring decisions, it is important to help applicants from disadvantaged groups, such as minorities.	When making hiring decisions, it is important to consider a variety of factors.	Hiring decisions should be made according to the skills and qualifications of the candidate.	Hiring decisions should be made according to merit; affirmative action is just another form of racism.
Affordable Care Act	The Affordable Care Act is the perfect solution to increasing health care coverage.	The Affordable Care Act needs a little bit of tweaking in its implementation.	The Affordable Care Act will lead us to impede on the freedom of American citizens.	The Affordable Care Act will lead us to death panels and financial ruin.
Alternative Energy	Improving alternative energy options, such as wind and solar, should be our nation's biggest priority.	Improving alternative energy options, such as wind and solar, is tentatively worth exploring.	Wind and solar energy programs are a nice supplement to fossil fuels.	Wind and solar energy programs are untested, unreliable, and expensive substitutes for oil.
American Patriotism	Our American country is a great one, but we are no better than any other country.	Our American country is a great one, and we should be honored to be a leader in the global community.	America's freedom, equality, and laissez-fair economics make us a major world superpower.	America's freedom, equality, and laissez-fair economics are a sign that God wants us to lead the world.
Border Security	We don't need to build a wall along the border, we need to welcome more immigrants.	We don't need to build a wall along the border, we need to focus on immigration policies.	The best way to secure the US-Mexico border is by improving both security and immigration policies.	The best way to secure the US-Mexico border is to build a 1,950 mile fence.
Contraception	Contraception should be available to women of all ages for free.	Contraception should be available to all adult women through their health insurance.	Forcing health insurance companies to pay for contraception is a complicated issue.	Forcing health insurance companies to pay for contraception violates religious freedom.

Death Penalty	Wrongful convictions mean the death penalty should never be used as punishment.	Wrongful convictions mean the death penalty should only be used in extreme cases.	In a case of first degree murder the death penalty should be considered as an option.	In a case of first degree murder the death penalty should always be used.
Don't Ask Don't Tell	Don't Ask Don't Tell needed to be repealed because it was a homophobic and discriminatory policy.	Don't Ask Don't Tell needed to be repealed because it was an ineffective policy.	Don't Ask Don't Tell was a good policy that maintained military morale.	Don't Ask Don't Tell was a good policy that did not go far enough to keep gays out of the military.
Flag Desecration	The right to burn a flag must be protected as a form of free speech.	The right to burn a flag is a complicated issue.	Flag burning should be discouraged because it is unpatriotic.	Flag burning should be prohibited because it is treasonous.
Global Warming	According to scientific research, global climate change is man-made, and is the largest threat to human society today.	According to scientific research, global climate change exists, but might not be due to human activity alone.	Global climate change is not a big problem, because the Earth could be entering a natural warmer phase.	Global climate change is not a problem, because we are not sure if it even exists.
Government Debt	Running a government is expensive; to maintain programs without increasing debt we will need to raise taxes.	Running a government is expensive; we will need to consider both raising taxes and trimming the budget.	The government should cut spending to reduce our debt responsibly.	The government should cut spending across the entire budget, except for defense, to eliminate the deficit.
Gun Control	The right to own guns applies to militias; individuals don't need weapons.	The right to own guns should be subject to some restrictions.	The second amendment provides a right to bear arms; law abiding citizens should be able to own weapons.	The second amendment provides a right to bear arms; every citizen should own a firearm.
Health Care	The federal government should provide universal health care for all Americans.	The federal government should provide universal health care for all children under 18.	The federal government should not be providing health care; that should be left to the states.	The federal government should not be providing health care; that should be left to the free market.
Immigration Benefits	Illegal immigrants in good standing should be granted citizenship.	Illegal immigrants in good standing should be given work permits after paying a fee.	Living in a state illegally, even for a long time, should not allow immigrants to take advantage of American resources.	Living in a state illegally for any amount of time should result in immediate deportation.

Labor Unions	Labor unions protect the basic human rights of workers from greedy corporate exploitation.	Labor unions protect workers by providing collective bargaining rights.	Labor unions need limits on collective bargaining rights.	Labor unions inflate costs and are a front for the liberal agenda.
Legalization of Marijuana	Marijuana should be legal for all adults without restriction.	Marijuana should be legal for all adults for medicinal purposes.	Marijuana should be illegal in this country, but some states have decided it is OK.	Marijuana should be illegal in this country, because it is unsafe, destroys lives, and causes crime.
Mitt Romney	Mitt Romney made his wealth by lying, swindling, and cheating to get ahead.	Mitt Romney made his wealth by being a shrewd businessman.	As a presidential candidate, Governor Romney had some issues but would have been a fine president.	As a presidential candidate, Governor Romney had the potential to be the best leader our country has ever seen.
Nuclear Weapons	Nuclear weapons are dangerous regardless of who holds them; we must eradicate our entire nuclear arsenal.	Nuclear Weapons are dangerous regardless of who holds them; we need to ensure our arsenal is secure.	To deter terrorists from acquiring nuclear weapons we must maintain our nuclear program.	To deter terrorists from acquiring nuclear weapons we must have the strongest nuclear weapons program in the world.
Occupy Wall Street	The occupy wall street movement is about frustration with income inequality.	The occupy wall street movement is about frustration with the economy.	The occupy wall street protestors do not have a clear agenda about what they want.	The occupy wall street protestors are just loud complainers who are too lazy to work.
Off-shore Oil Drilling	Off-shore oil drilling is dangerous for our environment and should be completely banned.	Off-shore oil drilling is dangerous for our environment and should be regulated.	Domestic oil is vital for our future; we should carefully and slowly expand off-shore drilling.	Domestic oil is vital for our future; we should immediately lift all restrictions and invest in new sites.
Paying For College	The federal government should forgive all student loan debt.	The federal government should reduce the interest rate on student loans.	College students should work hard to pay for college, getting a part-time job if they need it.	College students should work hard to pay for college, not rely on government handouts.
President Obama	President Obama has done a fabulous job leading the country in both foreign and domestic areas.	President Obama has done a fabulous job leading the country on foreign issues, but has faltered on some social issues.	President Obama's policies have not been able to solve our country's problems.	President Obama's policies are undermining American freedom and leading us towards a socialist society.

Public Schools	To continue educating our youth we must invest more money into public schools.	To continue educating our youth, we must invest more money into public and private schools.	The public school system has many problems; students should be able to go to private schools.	The public school system has many problems; we should move all teaching to private schools.
Role of Government	The role of government is to ensure that every person has a chance to succeed.	The role of government is to secure the rights and freedoms of individual citizens.	The ideal federal government should work in tandem with State governments and the private sector.	The ideal federal government should stay small and out of the lives of its citizens.
Same Sex Marriage	Partners of the same sex should be recognized through marriage.	Partners of the same sex should be recognized through civil unions.	Traditional marriage should be an issue for individual state legislatures.	Traditional marriage should be protected through a constitutional amendment.
School Safety	Arming teachers and security guards will contribute to a culture of violence in our society.	Arming teachers and security guards might help, but we also need better mental health resources.	In order to protect classrooms, we should increase the number of school security guards.	In order to protect classrooms, teachers and school administrators should be allowed to carry firearms.
Separation of Church and State	Religious symbols in government spaces should be removed, because our country has many religious traditions.	Religious symbols in government spaces should not be added, even though we are a mostly Christian country.	Within public and government spaces, religious symbols should be allowed.	Within public and government spaces, we should honor our Christian heritage.
Sex Education	The best way to combat teen sex issues is through comprehensive sex education programs starting in kindergarten.	The best way to combat teen sex issues is through exposure to educational resources.	Teen sex education programs should be abstinence-plus.	Teen sex education programs should be abstinence-only.
Social Security	We cannot cut back social security now, we need to expand it for future generations.	We cannot cut back social security now, we just need to make some adjustments.	Social security will be bankrupt in a few years - we need to overhaul the program.	Social security will be bankrupt in a few years - seniors should manage their own savings in the private market.
Stem Cell Research	Stem cell research is vital to curing many horrible genetic diseases.	Stem cell research is vital to medical science but is also ethically questionable.	Stem cell research is probably not worth it, as major medical breakthroughs are rare.	Stem cell research is just not worth it; taking an innocent life never is.
Teaching Creationism in Schools	Creationism should never be taught in public schools.	Creationism should only be explored in schools as a historical alternative to evolution.	Creationism should be taught in schools, as the view that God created the world is plausible.	Creationism should be taught in schools, as the view that God created the world is correct.

Trickle Down Economics	Giving tax cuts to the rich only makes them richer when we should help the poor directly.	Giving tax cuts to the rich is probably not the best way to grow the economy.	The route to economic recovery is through tax breaks for businesses and the middle class.	The route to economic recovery is through tax breaks for corporations and wealthy Americans.
United Nations	The US should support the UN in every way possible, including delegating our resources to their leadership.	The US should support the UN as long as their goals are compatible with ours.	As an organization, the UN is ineffective and relatively harmless.	As an organization the UN is corrupt, incompetent, and often anti-American.
War on Terror	Terrorism must be addressed through diplomacy as more war leads to more terrorists.	Terrorism must be addressed by a combination of the careful use of force and diplomacy.	Radical Islamists pose a great threat to our nation; we must deal with them strategically and carefully.	Radical Islamists pose a great threat to our nation; we need to continue the war on terror at all costs.
Welfare	Welfare programs are essential; they provide a safety net for all Americans.	Welfare programs are expensive but useful; they support people in times of need.	Welfare programs, although sometimes helpful, also discourage people from looking for work.	Welfare programs are a huge drain on our economy and reward the least deserving of our citizens.

Appendix B: Photos of the Political Candidates

Photos of the political candidates, John Baker and Mike Shipman, and of the locations where the two debates occurred.



Appendix C: Political Expertise Quiz

Political Expertise Quiz. Correct answers are bolded (order of responses was randomized).

Of the two, which past presidential candidate is most likely to oppose the legalization of same-sex marriages?

- Barack Obama
- **Mitt Romney**
- not sure

Before his election as Vice President, which state did Joe Biden represent in his role as U.S. senator?

- Connecticut
- **Delaware**
- New York
- Pennsylvania
- not sure

How many senators are in the U.S. Senate?

- 50
- **100**
- 535
- 538
- not sure

Who is the current Speaker of the House?

- Newt Gingrich
- **John Boehner**
- Nancy Pelosi
- Harry Reid
- not sure

Which party currently controls the US Senate?

- Democrats
- **Republicans**
- not sure

Who is the current Secretary of State for the US Government?

- Hillary Clinton

- **John Kerry**
- Condoleeza Rice
- Colin Powell

Out of the following states, which has historically voted strongly Republican in national elections?

- Minnesota
- Delaware
- **Oklahoma**
- Pennsylvania
- Washington

Of the two past candidates, which is most likely to restrict the circumstances under which women will be able to legally obtain an abortion?

- Barack Obama
- **Mitt Romney**
- not sure

Appendix D: Objective Political Orientation Survey

Objective Political Orientation Survey. Subjects responded on a 4 point scale with the following options for each statement: *Completely Agree, Somewhat Agree, Somewhat Disagree, Completely Disagree.* Survey adapted from Zell and Bernstein (2014).

- 1) There need to be stricter laws and regulations to protect the environment.
- 2) The government should help more needy people even if it means going deeper into debt.
- 3) The growing number of newcomers from other countries threaten traditional American customs and values.
- 4) I never doubt the existence of God.
- 5) Business corporations make too much profit.
- 6) Gays and lesbians should be allowed to marry legally.
- 7) The government needs to do more to make health care affordable and accessible.
- 8) One parent can bring up a child as well as two parents together.
- 9) Government regulation of business usually does more harm than good.
- 10) Abortion should be illegal in all or most cases.
- 11) Labor unions are necessary to protect the working person.
- 12) Poor people have become too dependent on government assistance programs.

Appendix E: Additional Results and Discussion - Experiment 1

This appendix contains additional analyses not reported in the results section of Experiment 1.

All post-hoc comparisons used a Bonferroni correction.

Recall

I conducted several other analyses with the recall results. First, subjects were better at recalling the strong version of a statement ($M = .56$, $SEM = .02$) than the weak version ($M = .40$, $SEM = .02$), $t(125) = 16.40$, $p < .001$, $d = 0.82$. This pattern is inconsistent with previous results (Putnam et al., 2014) that showed no difference in recall for the different item strengths. One explanation could be the revised materials, where the strong statements were more extreme, which may have made them more salient.

Second, a one-way ANOVA with counterbalancing order as a between-subjects factor suggested there was a significant effect, $F(5, 120) = 2.44$, $p = .04$, $\eta_p^2 = 0.09$, but post-hoc comparisons revealed that there was only a marginal effect of one comparison (CB order 1 vs. CB order 2), $p = .06$.

Third, several analyses in the main results section reported outcomes as a function of the subjects' self-reported political orientation. Subjects reported their political orientation on a 7-point scale, but in interest of simplicity, political orientation was collapsed into three groups (Democrats, Independents, and Republicans). Additionally, subjects completed an objective measure of political orientation which also assigned them to one of seven political orientations.

Table E1 shows recall collapsed across item type as a function of self-reported political orientation. A one-way ANOVA showed that self-reported political orientation affected recall, $F(6, 119) = 2.91, p = .01, \eta_p^2 = 0.13$, with post-hoc comparisons revealing that the Moderate Republicans had better recall than the Independents and Average Republicans, $p = .040$ and $p = .047$ respectively. In contrast, objective political orientation did not influence recall, $F = 1.41$. Taken together, political orientation did not appear to affect overall recall.

Table E1: Recall as a Function of Self-Reported Political Orientation

Self-Reported Political Orientation	Mean	SEM	n
Liberal Democrat	0.51	0.03	25
Average Democrat	0.45	0.06	12
Moderate Democrat	0.51	0.06	13
Independent	0.40	0.03	26
Moderate Republican	0.56	0.04	23
Average Republican	0.38	0.04	16
Conservative Republican	0.53	0.04	11

Finally, I checked whether any subject-level variables affected recall. Table E2 shows that recall (collapsed across item type) was correlated with performance on the political expertise quiz, $r_{\tau} = .24$, $p < .001$, suggesting that subjects who knew more about politics may have had better recall in the experiment. The following variables, however, did not influence recall: Self-reported political expertise, subject education, and age, all $ps \geq .18$. Subjects who were born outside of the U.S. may have had better recall than subjects born inside the U.S. ($M = .54$ vs. $M = .47$) but having only six subjects born outside the U.S. made a direct comparison untenable.

Table E2: Recall as a Function of Performance on Political Expertise Quiz

Score On Expertise Quiz	Mean Recall	SEM	n
3	0.39	0.05	10
4	0.44	0.04	18
5	0.36	0.04	22
6	0.51	0.04	19
7	0.51	0.04	22
8	0.55	0.03	35

Change Recollection

For the additional change recollection analyses, the primary dependent variable was change recollection accuracy as measured by Hits - False Alarms (with Control items serving as the false alarms). Counterbalancing order significantly affected change recollection, $F(5, 120) = 3.68, p = .004, \eta_p^2 = 0.13$, with post-hoc tests suggesting that counterbalancing order 0 ($M = .37, SEM = .06$) led to better change recollection than counterbalancing order 4 ($M = 0.16, SEM = .03$). Examining the items in the different counterbalancing groups failed to reveal any explanation for why change recollection was better in order 0. After noting this pattern I re-assigned items to different counterbalancing groups for Experiments 2 and 3.

Subjects were equally accurate in recollecting change for the Democratic and Republican candidate (0.23 and 0.24 respectively), $t < 1$, and had similar change recollection for the strong and weak items (both 0.24), $t < 1$. Furthermore, subjects from all political orientations had similar change recollection levels, regardless of whether political orientation was measured by self-report or the objective measure ($F_s \leq 1.74$). Additionally, I checked to see whether subjects were better at recollecting change when the candidates were from their own political party. However, two separate 3 (subject political orientation) x 2 (candidate party) ANOVAs failed to reveal any main effects or interactions, regardless of whether subject orientation was measured by self-report or the objective quiz (all $F_s \leq 2.45$). In addition, I conducted a paired samples t -test comparing change recollection for congruent items (e.g., when Democrats were attempting to recollect change for a liberal position) to change recollection for incongruent items (e.g., when

Republicans were attempting to recollect change for a liberal position). The means were identical in both bins ($M_s = 0.21$, $t < 1$).

Finally, none of the subject level variables (objective political expertise, subjective political expertise, age, and education) influenced change recollection accuracy, all $ps \geq .09$.

Response Time Results

Response times were measured in seconds, and represented the time from the presentation of the cue to the submission of the subject response. Response times were similar for recalling the different item types, repetition = 16.46, control = 17.98, and changed = 17.26, $F = 1.05$. The response times for recollecting change were also similar, repetition = 1.66, control = 1.86, and changed = 1.95, $F = 1.92$. Furthermore, recollecting change did not influence the reaction times for recalling a changed item, (change recollected = 17.48, change not recollected = 17.24, $t = 1.06$) or for the change recollection judgment itself (change recollected = 1.75, change not recollected = 2.05, $t < 1$).

Changing Towards and Away

Is it easier for a voter to remember a change in position if a politician makes a shift to be more in line with the voter's own views? Or is it easier when the politician shifts to a position that is more different from the voter's position? To address the question I grouped subjects by their self-reported political orientation, and omitted all of the self-reported independents. Then I calculated the average recall for the changed items when the change was towards or away from the subjects' own orientation. A *change towards* item would be a self-reported Democrat recalling an item where the Republican candidate shifted from a strong position to a weak

position (or the Democrat shifting from a weak position to a strong position), whereas a *change away item* would be a self-reported Democrat recalling an item where the Republican shifted from a weak positions to a strong position (or the Democrat shifting from a strong position to a weak position). Subjects had better recall when candidates changed away from the subjects' own view ($M = .41$, $SEM = .02$) than when candidates changed towards the subjects' own view ($M = .37$, $SEM = .02$), $t(99) = 2.26$, $p = .03$, $d = 0.18$. The same pattern occurred when subjects were grouped by their objective political orientation, rather than their self-reported orientation. In contrast, change recollection did not appear to be affected by the direction of the change. The hit rate for the change items was similar regardless of whether the candidate was changing away ($M = .39$, $SEM = .03$) or changing towards ($M = .40$, $SEM = .02$) the subject's own position, $t < 1$. Again, an identical pattern occurred when grouping subjects by their objective political orientation. In sum, subjects may be better at recalling changed positions when the candidates are shifting away from their own political views than when they are shifting towards their own political views.

Subject Judgment Results

In general, the subject judgments results (i.e., voting choices, estimates of future flip flopping etc.) were similar regardless of whether the self-reported or objective measure of political orientation was used to group subjects. However, there were a few exceptions for the trustworthiness and likability ratings—some main effects that occurred when using the self-reported measure did not replicate when using the objective measure. Critically, the interactions

(suggesting that subjects preferred the candidates from their own parties) were still significant, meaning the main effects should not be granted too much weight (Field, Miles, & Field, 2012).

For the trustworthiness ratings, using self-reported political orientation as an independent variable suggested a significant interaction, but also a main effect of candidate party, where the Democratic candidate was trusted more than the Republican at Debate 1. Using the objective measure of political orientation replicated the interaction, $F(2, 123) = 48.02, p = .001, \eta_p^2 = .44$, but failed to reveal the main effect of candidate party, $F = 1.16$. Collapsing across subject political orientation entirely and looking only at the average trust ratings for the two candidates suggested that the Democrat was indeed trusted more than the Republican, $t(125) = 3.73, p = .001, d = 0.55$. An identical pattern occurred after the recall test where the interaction still occurred, $F(2, 123) = 48.02, p = .001, \eta_p^2 = .44$, but the main effect of candidate party disappeared when objective orientation was used to group subjects, $F = 1.16$. Again, collapsing across political orientation suggested that the Democrat was trusted more than the Republican, $t(125) = 3.62, p = .001, d = 0.52$ after the recall test.

The likability ratings yielded a similar pattern at both Debate 1 and after the test. Using self-reported orientation led to a significant interaction (subjects liked the candidates from their own party), a main effect of candidate party (the Democrat was liked more than the Republican), and a main effect of self-reported political orientation (the Republicans tended to like both candidates more). With objective political orientation, the interaction occurred at Debate 1 and after the test, $F(2, 123) = 99.26, p < .001, \eta_p^2 = .62$ and $F(2, 123) = 95.94, p < .001, \eta_p^2 = .61$ respectively, but the main effects disappeared. At Debate 1 there was no main effect of candidate party, $F < 1$, or

subject orientation, $F < 1$, whereas at the recall tests there was only a marginal effect of candidate party, $F(1, 123) = 3.63, p = .059, \eta_p^2 = .03$, and no effect of subject orientation, $F < 1$. Collapsing across all subjects revealed that Democrat was liked more than the Republican at Debate 1, $t(125) = 3.95, p = .001, d = 0.64$, and at Debate 2, $t(125) = 4.62, p = .001, d = 0.74$.

The discrepancy between the self-reported and objective measures of political orientation likely results from the finding that subjects are more likely more liberal than they claim to be (Zell & Bernstein, 2014). Regardless, the interaction was significant in both cases, and should trump concerns over whether a main effect was present.

Appendix F: Additional Results and Discussion - Experiment 2

Appendix F contains additional analyses not reported in the results section of Experiment 2.

Some of these analyses were basic checks (e.g., to see if there were any effects of counterbalancing order) whereas other analyses addressed the main questions from Experiment 1. All post-hoc comparisons used a Bonferroni correction unless otherwise noted.

Recall

There were several additional recall analyses not reported in the main text. First, counterbalancing condition did not influence overall recall ($F < 1$). Second, recall for the strong items ($M = .50$, $SEM = .02$) was better than recall for the weak items ($M = .37$, $SEM = .02$), $t(99) = 11.14$, $p < .001$, $d = 0.72$, across all item types. This pattern was consistent with Experiment 1, but inconsistent with previous work using similar materials. (Putnam et al. 2014).

Third, I checked to see whether subject political orientation or the candidate party affiliation affected recall. Subjects had similar levels of recall for the Democratic ($M = .43$, $SEM = .02$) and Republican ($M = .44$, $SEM = .02$) candidates ($t < 1$). Furthermore, self-reported Democratic subjects had similar recall ($M = .44$, $SEM = .03$) to self-reported Republican subjects ($M = .43$, $SEM = .02$) ($t < 1$). Using the full 7-point political orientation scale or the objective measure of subject political orientation also failed to reveal any effects of political orientation on recall.

Fourth, although a small congeniality effect occurred in Experiment 1 (subjects were better at recalling attitude-consistent information than attitude-inconsistent information), this pattern did

not replicate in Experiment 2. A 2 (subject self-reported political orientation: Democrat, Republican) x 2 (candidate political party: Democrat, Republican) mixed-model ANOVA with recall as the dependent variable failed to reveal any significant main effects or interactions ($F_s \leq 1.99$). In addition, recall for the congruent items (e.g., Democrats recalling liberal positions; $M = .44$, $SEM = .02$) was similar to recall for the incongruent items (e.g., Democrats recalling conservative positions; $M = .43$, $SEM = .02$), $t = 1.42$. Similar outcomes occurred if the objective measure of political orientation was used instead of the self-report. Thus, recall was similar regardless of whether there was a match between the subject's political orientation and the content of the statement they were recalling, failing to replicate the congeniality effect that occurred in Experiment 1.

Finally, I also checked to see whether any subject demographic variables affected recall. Objective political expertise was modestly correlated with recall ($r_\tau = .21$, $p = .006$) suggesting that subjects who scored higher on the political expertise quiz also recalled more items in the experiment. Recall did not correlate with self-reported political expertise, being born in the U.S., age, or education ($ps \geq .067$).

Change Recollection

For the additional change recollection analyses, the dependent variable was Hits – False Alarms unless otherwise mentioned. Several analyses were conducted. First, counterbalancing order failed to influence change recollection ($F = 1.22$). Second, change recollection for the strong items ($M = .18$, $SEM = .02$) was just barely better than for the weak items ($M = .13$, $SEM = .02$), $t(99) = 2.00$, $p = .049$, $d = 0.24$. This pattern did not occur in Experiment 1.

Third, political orientation did not influence change recollection accuracy. Subjects were equally accurate at recollecting change for the Democratic and the Republican candidates, ($t = 1.08$).

Furthermore, subject political orientation did not influence change recollection accuracy, either as measured by self-report ($F < 1$) or the objective measure ($F < 1$). Thus, just as in Experiment 1, subject political orientation and the candidate's political party failed to affect change recollection.

Fourth, a 2 (self-reported subject political orientation: Democrat, Republican) x 2 (candidate political party: Democrat, Republican) mixed-model ANOVA with change recollection accuracy as the dependent variable failed to reveal any significant main effects or interactions, ($F_s \leq 1.99$). A similar pattern occurred when the objective measure of political orientation was used. The lack of an interaction suggests that subjects were equally capable of recollecting change, regardless of whether the position was consistent or inconsistent with their own political views. Experiment 1 showed the same pattern. This suggests that there was no congeniality effect in recollecting change.

Finally, none of the subject-demographic variables affected change recollection. Objective political expertise, self-reported political expertise, being born in the US, age, and education all failed to influence change recollection ($p_s \geq .181$).

Response Time Results

The response times for recalling the different item types were similar: repetition = 14.06, control = 14.55, changed = 14.76, $F = 1.92$. As were the reaction times for recollection change: repetition = 1.63, control = 1.60, changed = 1.70, $F = 1.63$. Recollecting change did not

influence reaction times for recalling a changed item (change recollected = 14.39, change not recollected = 14.65, $t < 1$), or for the change recollection judgment itself (change recollected = 1.75, change not recollected = 1.70, $t = 1.30$).

Changing Towards and Away

As in Experiment 1 (reported in Appendix E) subjects in Experiment 2 had better recall for changed positions when the candidates were moving away from the subject's own political views ($M = .36$, $SEM = .02$) than when the candidates moved towards the subject's own views ($M = .28$, $SEM = .02$), $t(99) = 3.87$, $p < .001$, $d = 0.42$. Thus, subjects may have more accurate recall for a candidate's position when he is moving away from the subject's own political views (as measured by the subject's self-reported political orientation) compared to a candidate moving towards the subject's own political views. This advantage, however, did not extend to change recollection; subjects were equally likely to recollect change when the change was away from their own position ($M = .31$, $SEM = .02$) than when it was towards their own position ($M = .33$, $SEM = .02$), $t < 1$.

Subject Judgments

Table F1 displays the voting results from Experiment 2. Table F2 displays the other judgments that subjects made after Debate 1 and after the recall test as a function of subjects' self-reported political orientation. In general, similar patterns emerged compared to Experiment 1: voting preferences were largely along party lines, estimates of future change went up after the recall test, and subjects had more positive attitudes towards the candidate from their own political party. One major difference was that in Experiment 1 several measures correlated with the

change recollection hit rate for the subjects, but this did not occur in Experiment 2. One explanation for this difference is the inclusion of the news phase in Experiment 2, which influenced the change recollection rates.

Table F1: Voting Results for the Two Candidates After Debate 1 and Debate 2 by Subject Political Orientation, Experiment 2

Subject Political Orientation	Debate 1		Debate 2	
	Democrat	Republican	Democrat	Republican
Total Votes	64	36	64	36
Self-Reported				
Democrats	49	1	50	0
Republicans	15	35	14	36

Voting. Table F1 shows the voting results after Debate 1 and the recall test as a function of the subjects’ self-reported political orientation. The Democratic subjects stuck to party lines, whereas the Republican subjects had a mixture of voting preferences. This discrepancy may be due to several of the self-reported Republicans actually having attitudes more consistent with independents or moderate Democrat as indicated by the objective measure. Eight subjects changed their votes between Debate 1 and the recall test.

Estimates of Future Change. Table F2 displays the results of the subject judgments broken down by self-reported political orientation. The third and fourth columns display the estimates of future change for each candidate at Debate 1 and after the recall test. At Debate 1, a 2 (subject political orientation: Democrat, Republican) x 2 (candidate party: Democrat, Republican) mixed-model

ANOVA revealed a main effect of candidate party, suggesting that subjects predicted that the Democratic candidate would be more likely to change positions in the future than the Republican candidate, $F(1, 98) = 11.81, p < .001, \eta_p^2 = .11$. There was no main effect of subject political orientation or a significant interaction (both $F_s \leq 1.98$). At the recall test a different pattern emerged. A 2 x 2 ANOVA failed to show a main effect of candidate party or an interaction (both $F_s \leq 2.15$), but did show a main effect of subject political orientation. This main effect indicated that the self-reported Democrats ($M = 47.03, SEM = 2.71$) provided lower future change estimates than the self-reported Republicans ($M = 55.02, SEM = 2.54$), $F(1, 98) = 4.62, p = .03, \eta_p^2 = .05$. Thus, at Debate 1, subjects thought the Democrat was more likely to change positions in the future than the Republican (replicating Experiment 1), but after the recall test, provided similar future change estimates for both candidates. However, after the recall test, the Republican subjects provided higher future change estimates than the Democrats (which did not occur in Experiment 1). Finally, two paired-sample *t*-tests confirmed that subjects provided higher ratings for both candidates after the recall test than after Debate 1, $t_s(99) \geq 4.18, p_s < .001, d_s \geq 0.48$, suggesting that experiencing the candidates change positions led to the higher estimates of future change.

Table F2: Subject Opinions About the Candidates as a Function of Their Self-Reported Political Orientation in Experiment 2

Subject Political Orientation	Past Change Estimate		Future Change Estimate		Trustworthiness		Likability	
	Democrat	Republican	Democrat	Republican	Democrat	Republican	Democrat	Republican
After Debate 1								
Democrat	-	-	36.36	29.00	64.72	33.24	79.14	16.56
Republican	-	-	46.02	31.76	42.30	60.78	38.00	62.96
Average			41.19	30.38	53.51	47.01	58.57	39.76
After Recall Test								
Democrat	6.68	7.00	46.96	47.10	60.00	29.84	74.24	14.74
Republican	7.18	6.80	58.90	51.14	37.80	54.88	35.84	61.22
Average	6.93	6.90	52.93	49.12	48.90	42.36	55.04	37.98
Change From D1 to Test								
Average			11.74	18.74	-4.61	-4.65	-3.53	-1.78

Note. For Past Change subjects estimated the number of times between 0 and 36 that a candidate changed. For Future Change subjects estimated how likely a candidate was to change in the future (0 = not very likely, 100 = extremely likely). For Trustworthiness, subjects reported how much they trusted each candidate (0 = not trustworthy at all, 100 = extremely trustworthy). For Likability, subjects were asked how much they liked each subject (0 = don't like at all, 100 = like a lot)

In contrast to Experiment 1, the recollection of change hit rates for the individual candidates did not correlate with the estimates of future change for those candidates after the recall test ($ps \geq .409$). However, the change in estimates between Debate 1 and the recall test was correlated with the change recollection rate for the Republican candidate ($r = .24, p = .018$), and marginally correlated for the Democratic candidate ($r = .19, p = .059$), generally reaffirming what was found in Experiment 1: higher change recollection rates were associated with an increase in the future change estimates from Debate 1 to the recall test.

Estimates of Past Change. After the recall test, subjects were asked to estimate how often each candidate had changed positions across debates. The first two columns of Table F2 show those estimates for each candidate, broken down by the subjects' self-reported political orientation. A 2 (subject political orientation: Democrat, Republican) x 2 (candidate party: Democrat, Republican) mixed-model ANOVA failed to reveal any main effects or interactions ($F_s \leq 1.79$) indicating that neither the political orientation of the subjects nor the candidate's political party influenced the estimates of past change. Curiously, in contrast to Experiment 1, the hit rates for the candidates did not correlate with the estimates of past change ($ps \geq .288$).

Trustworthiness Ratings. The fifth and sixth columns of Table F2 display the trustworthiness ratings after Debate 1 and the recall test. At Debate 1 there was a main effect of candidate political party, $F(1, 98) = 4.11, p = .045, \eta_p^2 = .04$, indicating that the Democratic candidate was deemed more trustworthy than the Republican candidate, and an interaction, $F(1, 98) = 60.62, p < .001, \eta_p^2 = .38$, suggesting that subjects trusted the candidate from their own political party more than the candidate from the opposing party. There was no main effect of subject political

orientation ($F < 1$). After the recall test, an identical pattern emerged: the Democratic candidate was deemed more trustworthy than the Republican, $F(1, 98) = 4.30, p = .041, \eta_p^2 = .04$, and a significant interaction suggested subjects trusted the candidates from their own party, $F(1, 98) = 56.09, p < .001, \eta_p^2 = .364$ (again there was no main effect of subject political orientation, $F < 1$). In addition, two paired-samples t -test indicated that subjects trusted both candidates more after Debate 1 than after the recall test, $ts(99) \geq 2.65, ps \leq .009, ds \geq 0.18$. This pattern of results generally replicates what was found in Experiment 1.

Experiment 1 showed that the trustworthiness ratings after the recall test were negatively correlated with the change recollection rates. In contrast, in Experiment 2, there were no significant correlations for either candidate, either for the trustworthiness judgments that occurred after the recall test ($ps \geq .144$), or for the change in trustworthiness judgment between Debate 1 and the recall test ($ps \geq .067$).

Likability Ratings. The last two columns of Table F2 display the likability ratings for the two candidates after Debate 1 and after the recall test broken down by subject political orientation. At Debate 1 there was a main effect of candidate party, $F(1, 98) = 25.65, p < .001, \eta_p^2 = .21$, indicating that the Democrat was liked more than the Republican, no main effect of subject political orientation, $F = 1.10$, and a significant interaction, $F(1, 98) = 138.88, p < .001, \eta_p^2 = .59$, indicating that candidates preferred the candidate from their own political party. An identical pattern occurred after the recall test: there was a main effect of candidate party, $F(1, 98) = 23.17, p = .001, \eta_p^2 = .19$, no main effect of subject political orientation, $F = 2.63, p = .108, \eta_p^2 = .03$, and an interaction, $F(1, 98) = 143.39, p < .001, \eta_p^2 = .59$. The interaction replicated what was

found in Experiment 1. Two paired-samples t -tests showed that subjects like the Democratic subject more at Debate 1 than after the recall test, $t(99) = 2.64, p = .010, d = 0.12$, but that they liked the Republican candidate similarly at both times, $t(99) = 1.25, p = .213$ (Experiment 1 showed an opposite pattern, with subjects liking the Republican candidate less after the recall test and liking the Democrat equally at both times).

Finally, the change recollection hit rates for both candidates did not correlate with either the recall test likability ratings ($ps \geq .678$) or the change in likability ratings between Debate 1 and the recall test ($ps \geq .075$).

Effects of News Statements Response on Recall and Change Recollection

The focal analysis of Experiment 2 examined how making an accurate or inaccurate news verification response influenced later recall and change recollection for the different item types. As discussed in the main text, during the news phase subject responses could be categorized into one of three groups: a correct endorsement (accurately verifying a statement, e.g., responding *True* to a flip-flop statement about a changed item), a misinformation endorsement (endorsing an untrue statement, e.g., responding *True* to a flip-flop statement about a repetition item), or a miss response (failing to accurately verify a statement but not endorsing an untrue statement, e.g., responding *False* to a flip-flop statement about a changed item). In the main text the analyses grouped the misinformation and miss categories together for ease of interpretation. As shown here, those two categories of responses did not influence recall, but did have an effect on change recollection.

Figure F1 shows recall for the different item types as a function of the different news phase responses (the correct responses are included here for comparison). A 3 (item type: repetition, control, changed) x 2 (news accuracy: misinformation, miss) within-subjects ANVOA failed to reveal a main effect of news accuracy, $F(1, 45) = 0.76, p = .389, \eta_p^2 = .02$, suggesting that misinformation endorsements and misses had similar effects on subsequent recall.

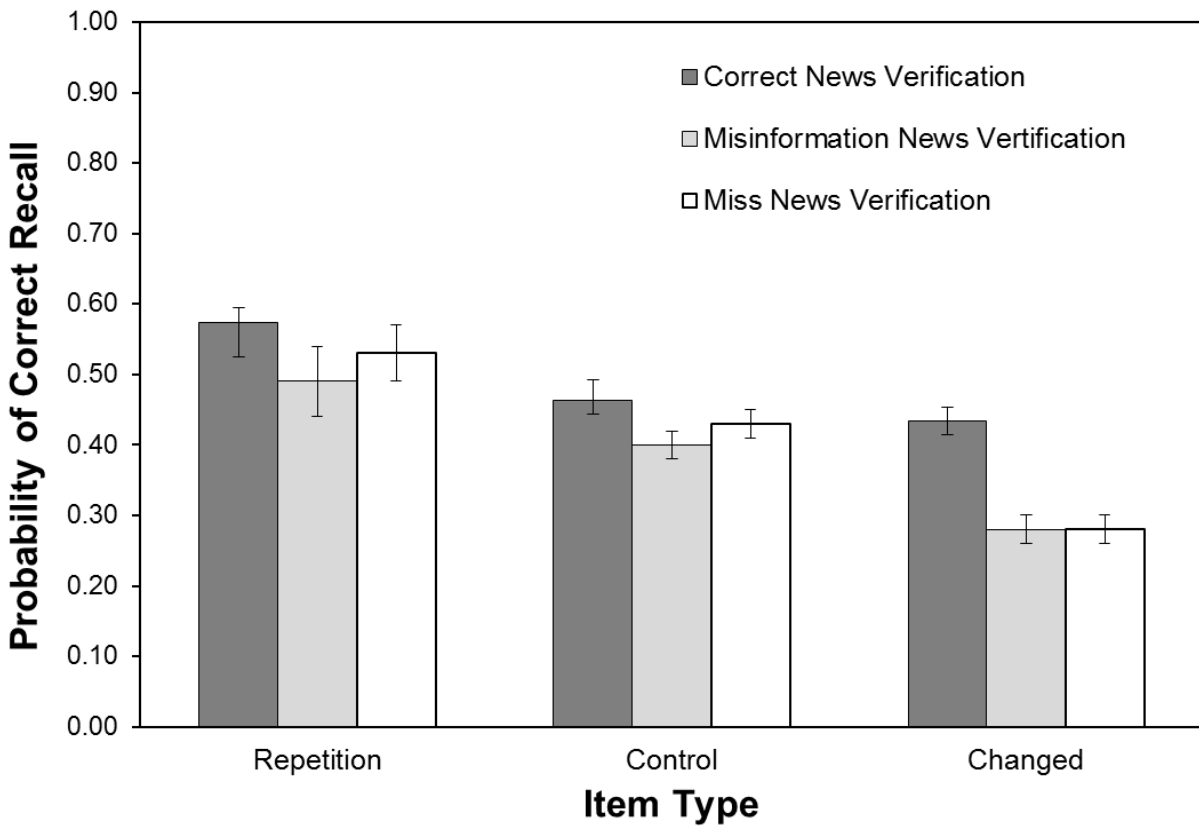


Figure F1: Recall by item type conditionalized on news phase verification response.

Table F3 shows change recollection for the different item types as a function of the different news responses (again, the accurate responses are included here for comparison, but omitted

from the analysis). A 3 (item type: repetition, control, changed) x 2 (news accuracy: misinformation, miss) within-subjects ANVOA failed to reveal a main effect of news accuracy ($F < 1$) but did reveal a significant interaction, $F(2, 90) = 4.12, p = .019, \eta_p^2 = .08$. Simple effects analysis revealed that the interaction was driven by the repetition items, where there was more erroneous change recollection following a misinformation endorsement than a miss, $t(45) = 2.47, p = .017, d = 0.42$. For both the control and the changed items, change recollection was similar following a misinformation endorsement or a miss, $ts \leq 1.60$. Thus, if subjects make a misinformation endorsement during the news phase, they may be more likely to erroneously recollect change for the repetition items. Again, for the control items, a misinformation response could be characterized by responding *true* to a straight statement *or* a flip-flop statement, thus any effects of making a misinformation response on change recollection could be cancelling each other out.

Table F3: Probabilities of Change Recollection and Signal Detection Measures in Experiment 2, Broken Down by News Verification Accuracy

News Verification Response	Item Type			SDT Measure	
	Repetition	Control (FAs)	Changed (Hits)	d'	c
Accurate	.09 (.01)	.17 (.03)	.54 (.03)	1.47 (0.16)	0.44 (0.09)
Misinformation	.27 (.05)	.17 (.02)	.21 (.02)	0.17 (0.08)	1.12 (0.08)
Miss	.14 (.03)	.19 (.02)	.25 (.02)	0.23 (0.08)	0.99 (0.07)

Note: Numbers in parentheses represent standard error.

Appendix G: Additional Results and Discussion - Experiment 3

Appendix G contains brief descriptions of two pilot experiments along with additional analyses not reported in the results section of Experiment 3. Because Experiment 3 did not have a politically diverse sample (there was only one Republican) some analyses that were reported in earlier experiments are not reported here.

Pilot Experiment 3A: Change Detection

One concern with the design of Experiment 3 was that subjects would ignore the introductory statements and focus only on the position statements. The goal of Pilot Experiment 3A was to show that the introductory statements could affect memory. The design was a variation of Experiment 3. Subjects studied statements at Debate 1 and Debate 2 as in Experiment 3. During Debate 2, the statements were presented along with the introductory phrases, as described in the methods section of Experiment 3. Table G1 presents the full list of introductory phrases.

However, in the pilot experiment subjects completed a measure of change detection during the Debate 2 presentations, which was the only dependent variable in the experiment. Subjects were instructed to read the Debate 2 positions, but to press a “changed” button if they detected that a candidate had changed positions. If they did not notice a change in position, they pressed the “next” button to advance to the next item (presentation was self-paced). Subjects were warned ahead of time that the candidates might not always be telling the truth when they claimed to change positions. As in Experiment 3, every item type had an introductory phrase. Thus, the

experiment was a 3 (item type: repetition, control, changed) x 2 (introductory phrase type: neutral, looking back) within-subjects design with change detection as the dependent variable.

Table G1: Introductory Phrases Used in Experiment 3

Introductory Phrase Type	
Neutral	Looking Back
I'm glad this topic came up today	In contrast to my previous position, I now believe that
Let me take this opportunity to say that	Since the last debate I've shifted on this - I think that
My message is that	I recently changed my view on this topic - I now believe that
I hope everyone in the room understands my position – that	In the past I may have said something else, but I will argue that
My stance on this topic is that	I've changed my mind. I think that
Its certainly that case that I believe that	I have altered my stance on this issue. I am of the opinion that

Twenty-three subjects from the same pool as Experiment 1 participated. Table G2 shows the change detection rates, broken down by item type and introductory phrase type. A 3 (item type: repetition, control, changed) x 2 (introductory phrase type: neutral, looking back) repeated measures ANOVA revealed a main effect of item type, $F(2, 44) = 34.93, p < .001, \eta_p^2 = .61$, a main effect of introductory phrase, $F(1, 22) = 18.14, p < .001, \eta_p^2 = .45$, but no interaction, $F < 1$. As seen in the table, change detection was highest for the changed items, moderate for the control items, and lowest for the repetition items; post-hoc analyses confirmed that all comparisons were significant, ($ps < .001$). The looking back statements ($M = .32, SEM = .04$) led

to higher change detection rates than the neutral statements ($M = .18$, $SEM = .02$). The lack of an interaction suggested that the looking back phrases did not lead to more accurate change detection. Indeed, d' was similar in both conditions ($t < 1$). However, the introductory phrases did lead to a criterion shift; As measured by c , the looking back statements led to a more liberal response bias for change detection than the neutral statements, $t(22) = 3.13$, $p = .005$, $d = 0.62$.

Table G2: Change Detection Rates as a Function of Item Type and Introductory Phrase Type Along With Signal Detection Measures in Pilot Experiment 3A

Introductory Phrase Type	Item Type			SDT Measure	
	Repetition	Control	Changed	d'	c
Neutral	.05 (.02)	.16 (.03)	.34 (.04)	0.67 (.19)	.82 (.09)
Looking Back	.18 (.04)	.30 (.05)	.49 (.06)	0.63 (.17)	.42 (.17)

Note: Change detection occurred during Debate 2. Signal Detection measures of accuracy (d') and criterion (c) were calculated using the change detection rate for changed items as hits and the change detection rates for control items as false alarms. Standard errors are in parentheses.

The results of Pilot 3A suggested that subjects did attend to the introductory phrases. Subjects were more likely to report detecting change when a statement was paired with a looking back phrase than a neutral phase. This finding was reassuring for the main goal of Experiment 3, in showing that subjects did attend to the statements. But the finding was also the first indication that the looking back phrases might not be effectively cueing subjects to look back to the Debate 1 positions. Instead, the looking back statements appeared to simply cue subjects that a change had occurred.

Pilot Experiment 3B: Change Recollection

Pilot Experiment 3B was similar to Pilot Experiment 3A except that rather than completing a measure of change detection, subjects completed a measure of change recollection. The goal of this pilot experiment was to ensure that the introductory statements influenced later change recollection. During Debate 2, the position statements were presented along with the introductory phrases as in Experiment 3. No change detection measure was collected. Also, subjects were not warned that the candidates might be lying when they said they changed positions (paralleling Experiment 3). For the final test subjects simply completed the change recollection measure from Experiment 3 without attempting to recall each position first. Subjects were presented with candidate information and a topic and were asked if the candidate changed positions on that topic. Subjects then pressed a “yes” or “no” button before moving on to the next item. Twenty-three subjects participated.

The results paralleled those of Pilot Experiment 3A. Table G3 shows the change recollection rates broken down by item type and whether the statements were paired with a neutral or looking back phrase. A 3 (item type: repetition, control, changed) x 2 (introductory phrase type: neutral, looking back) revealed a main effect of item type, $F(2, 44) = 9.91, p < .001, \eta_p^2 = .31$, and a main effect of introductory phrase type, $F(1, 22) = 12.84, p = .002, \eta_p^2 = .37$, but no interaction ($F = 1.07$). Post-hoc comparisons revealed that change recollection was higher for the changed items than the repetition or control items, $ts(22) > 3.10, ps < .005, ds > 0.52$, but that the repetition and control items led to similar change recollection rates, $t = 1.40$. The looking back statements ($M = .33, SEM = .04$) led to higher change recollection rates than the neutral statements ($M = .26, SEM = .04$), $t(22) = 3.55, p = .002, d = 0.39$. The lack of an interaction suggested that the looking

back phrases did not lead to more accurate change recollection. Indeed, d' was similar in both conditions ($t < 1$). However, the introductory phrases did lead to a criterion shift. As measured by c , the looking back statements led to a more liberal response bias for change recollection than the neutral statements, $t(22) = 4.42, p < .001, d = 0.48$.

Table G3: Change Detection Rates as a Function of Item Type and Introductory Phrase Type Along With Signal Detection Measures in Pilot Experiment 3B

Introductory Phrase Type	Item Type			SDT Measure	
	Repetition	Control	Changed	d'	c
Neutral	.23 (.05)	.22 (.04)	.32 (.05)	0.30 (.13)	.75 (.13)
Looking Back	.26 (.04)	.31 (.04)	.41 (.04)	0.31 (.15)	.45 (.12)

Note: Change recollection occurred after Debate 2. Signal Detection measures of accuracy (d') and criterion (c) were calculated using the change recollection rate for changed items as hits and the change recollection rates for control items as false alarms. Standard errors are in parentheses.

Thus, just as with the change detection pilot, the looking back phrases did not lead to more accurate change recollection. Instead, the looking back phrases led to a criterion shift, where subjects were more likely to report remembering change after a looking back statements. Again, this finding was reassuring for Experiment 3, because the introductory statements did affect later change recollection, but worrisome in that they did not encourage more accurate change recollection.

Finally, I ran a third pilot experiment that was identical to Pilot Experiment 3B except that subjects were explicitly warned that the politicians might not be telling the truth when they

claimed to change positions. The results, however, were nearly identical to Experiment 3B, so a warning was not included in Experiment 3.

Experiment 3: Additional Results and Discussion

Recall

There were several additional recall analyses not reported in the main text. First, counterbalancing condition did not influence overall recall ($F < 1$). Second, recall for the strong items ($M = .55$, $SEM = .03$) was better than recall for the weak items ($M = .41$, $SEM = .03$), $t(49) = 8.00$, $p < .001$, $d = 0.74$. Third, I checked to see whether the political orientation of the subjects or the candidates affected recall. Subjects had similar levels of recall for the Democratic ($M = .49$, $SEM = .03$) and Republican ($M = .47$, $SEM = .03$) candidates ($t = 1.46$). Furthermore, self-reported Democratic subjects had similar recall ($M = .50$, $SEM = .03$) to the self-reported independents, ($M = .46$, $SEM = .04$), as did the one self-reported Republican subject ($M = .42$, $SEM = .03$) ($F < 1$). Using the full 7-point political orientation scale or the objective measure of subject political orientation similarly failed to reveal any effects of political orientation on recall. Given the sample was not politically diverse, I did not examine whether a congeniality effect occurred in recall or change recollection. Finally, recall was not correlated with objective political expertise, self-reported political expertise, age, or education ($ps \geq .126$).

Change Recollection

For the additional change recollection analyses, the dependent variable was Hits - False Alarms unless otherwise mentioned. Several analyses were conducted. First, counterbalancing order did not influence change recollection ($F < 1$). Second, change recollection was similar for the strong ($M = .15$, $SEM = .03$) and weak items ($M = .17$, $SEM = .03$), $t < 1$. This pattern replicated

Experiment 1; Experiment 2 showed a marginal advantage in recollecting change for the strong items.

Third, change recollection was similar for the Democratic and Republican candidates ($t < 1$). Change recollection accuracy was also similar regardless of the subject's political orientation, whether it was measured by self-report ($F < 1$) or the objective measure ($F = 1.13$). Thus, just as in the previous experiments, subject political orientation and the candidate's political party failed to influence change recollection.

Finally, none of the subject-demographic variables affected change recollection. Objective political expertise, self-reported political expertise, being born in the US, age, and education all failed to influence change recollection ($ps \geq .082$).

Reaction Time Results

The reaction times for recalling the different item types were similar: repetition = 14.21, control = 14.45, changed = 15.17, $F = 1.31$. As were the reaction times for recollection change: repetition = 1.63, control = 1.77, changed = 1.60, $F = 1.53$. However, recollecting change ($M = 16.35$, $SEM = 1.23$) led to a slower response time for recalling a changed item than not recollecting change ($M = 14.79$, $SEM = 0.92$). Subjects may have had more detailed responses for items where they remembered change, which may have led to the longer response times (response time was measured when subjects submitted their response). Recollecting change for a changed item did not influence reaction time for the change recollection judgment itself (change recollected = 1.93, change not recollected = 1.77, $t = 1.07$).