An important but understudied characteristic of the information environment involves political information changing across time. This dynamic feature of the environment can make it difficult for voters to possess accurate political knowledge. In this study, we assessed memory for political information using self-report and eye movement methods. We used these metrics to examine how individuals learn facts about policies whose important features have changed across time. We find that eye movements can accurately assess changes in political information even when self-reports fail to do so. Our results highlight the utility of a converging methods approach in the study of dynamic information environments, and specify mechanisms that facilitate or inhibit people’s capacity to recognize changes in political information.

**Keywords:** Information Environment, Eye Movements, Political Knowledge, Political Communication, Memory.

doi:10.1093/hcr/hqy006

The necessity of a politically informed citizenry for a well-functioning democracy is a prominent feature of classic theories of democracy (Dewey, 1927/2016; Mill, 1861/1991). In recent years, there is a growing recognition that people’s information environment plays a critical role in facilitating or impairing their ability to acquire and retain accurate political information (Delli Carpini, 2012; Dimitrova, Shehata, Strömbäck, & Nord, 2014; Eveland, Marton, & Seo, 2004; Mondak, 1995). An important but understudied characteristic of the environment is that political information can change across time: a feature that can make it difficult for voters to possess accurate political knowledge. This dynamic nature of the information environment is pervasive and reflected in instances in which political actors (e.g., politicians, interest groups, media organizations, etc.) disseminate different or contradictory facts at distinct points in time.

Corresponding author: Jason C. Coronel; e-mail: coronel.4@osu.edu
For example, during Donald’s Trump’s presidential campaign, he promised to cut the U.S. military budget when asked about his proposed defense spending policy (Meet the Press, 2015). However, once elected to office, his spending proposals entailed a drastic increase in the military’s budget (Holland, 2017). Likewise, then-presidential candidate Barack Obama’s original health care proposal did not mandate that individuals purchase health insurance, a major and contentious distinction between his platform and Hillary Clinton’s during the 2008 Democratic primary. Shortly following his inauguration, however, his healthcare proposal required individuals to purchase health insurance (Holan, 2009). As highlighted by these real-world examples, the dynamic nature of the information environment raises important questions about the extent to which voters possess accurate political information. For instance, assuming that voters are exposed to information before and after changes have occurred, are voters able to reliably recognize when political information has changed across time?

In the study reported here, we answered this question by using self-report and eye movement assessments of memory for political information. In particular, we employed a novel design using eye movement monitoring as a means of assessing memory for political information as participants evaluated changes in policies across time. A growing body of work in the field of cognitive psychology has shown that memories of specific prior experiences can influence eye movements and, as a consequence, eye movements can be used to indirectly assess memory independent of self-report responses (Hannula et al., 2010). Critical to the design of this study, eye movements have been shown to reveal memory of information despite inaccuracies in self-report responses.

In our study, we exposed individuals to facts about policies at two points in time. At the first time point, individuals were shown an original version of a policy (e.g., purchasing health insurance will be optional). At the second time point, individuals were exposed to a modified version of the policy in which an important feature had been changed (e.g., purchasing health insurance will be mandatory). Individuals were then asked, via self-report, to render a judgment as to whether the policy had or had not changed. Critically, we found that despite rendering an incorrect judgment that a change had not occurred, individuals disproportionately directed their gaze at the modified region (e.g., the word “mandatory”). As we explain in a later section, people’s memory for the policy before any change occurred led them to form an expectation regarding what the policy should state. Violating this expectation, in turn, elicited enhanced attention to the modified region.

Our study contributes to both the political communication and broader message processing literatures in two ways. First, our study advances a theoretical view, rarely considered in the public opinion and political communication literatures, suggesting that remembering is more than the process of retrieving stored information from memory. Individuals also scrutinize the validity of their memories and make subjective judgments regarding the likelihood that a recollected event did, in
fact, occur. This process of validating memories can lead individuals to render incorrect memory judgments (i.e., self-reporting that a change had not occurred) despite possessing accurate memory for political information. Second, our design introduces eye movements as a novel method for assessing memory in the context of message processing.

Our discussion proceeds as follows: the first section discusses the role of memory for political information in a dynamic environment. The second section describes how eye movements can be used to assess memory. We then present our novel design and results. Finally, we discuss the broader implications of our findings for communication and public opinion research.

Dynamic information environments and memory

Over the course of their daily lives, individuals often encounter political information via different forms of media. In some cases, this political information is remarkably stable and does not change across time. For example, rarely do politicians in the United States change their partisan affiliation (Nokken, 2009). This stability is useful, because individuals can rely on the predictability of this information when making political decisions (e.g., a voter can be confident that politicians will likely not switch parties in the near future). However, another critical feature of the information environment is that it is ever-changing. For instance, some politicians often switch policy stances over the course of their careers (Doherty, Dowling, & Miller, 2015).

The ability of individuals to accurately recognize these changes is important in facilitating informed political decisions. Indeed, realizing that a politician has reneged on a campaign promise can be a salient consideration during an election. Furthermore, although we focus on changes in policy proposals, many other characteristics of the information environment can change. Information such as the crime rate, public opinion about LGBTQ rights, or the rate of unemployment can change across time and can be important considerations when individuals render political decisions. For example, knowledge that the crime rate has increased is relevant when assessing the effectiveness of existing anti-crime policies. As shown by these examples, the ability to realize that changes have occurred can be as relevant as an awareness of the political world’s relative stability. Thus, the mechanisms that lead people to successfully recognize (or fail to recognize) changes in the information environment is an important topic to study.

What, then, accounts for people’s success or failure at recognizing these changes? Assuming that individuals are exposed to political information before a change has occurred (pre-change information) and after a change has occurred (post-change information), the process of successfully recognizing changes involves at least two stages. First, individuals must have successfully stored the pre-change information in memory. Second, they must compare the pre-change with the post-change information and accurately assess that a change has occurred (Beck,
Failure at either of these two stages will lead to failure to recognize these changes. There are two theories that provide distinct memory-related explanations of why individuals may fail to recognize these changes.

The first account suggests that individuals fail at recognizing changes if they possess no memory of the pre-change information. Hereafter we refer to this as the memory storage account. This lack of memory can be due to several reasons. First, individuals may have failed to store the pre-change information in memory during initial exposure (this process of storing information in memory is often referred to as encoding; Hollingworth & Henderson, 2002). This can occur if individuals were not paying attention when they encountered the pre-change information (Rensink, O’Regan, & Clark, 1997). Alternatively, individuals may not have encoded the information due to interference from other pieces of information in the environment (Baddeley, 1976; Chandler, 1989; Underwood & Postman, 1960). In politics, individuals are continuously bombarded with many distinct pieces of information. They may have redirected their attention to other types of information before successfully encoding the relevant pre-change information. Second, individuals may have successfully encoded the information but, over time, the information has degraded from memory (Baddeley, 1976, 1998). Taken together, these mechanisms suggest that by the time individuals encounter post-change information, they no longer possess memory for the pre-change information. This lack of memory for the pre-change information can lead them to erroneously conclude that a change has not occurred.

A second account suggests that individuals do possess memory of the pre-change information, but errors in judgment occur because individuals mistakenly judge their memories of the pre-change information to be inaccurate. Hereafter we refer to this as the memory monitoring account. According to this perspective, remembering is more than the process of recollecting stored information. Individuals also assess the validity of their memories and make judgments regarding the likelihood that a recollected event did, in fact, occur (Johnson, 2006; Johnson, Hashtroudi, & Lindsay, 1993; Kelley & Jacoby, 1998; Metcalfe & Shimamura, 1994). Furthermore, these attempts to check the validity of one’s memories draw on various types of information, such as general world knowledge, beliefs about the reliability of one’s memories, contextual information, prior information about source cues, social information, and so forth (for reviews, see Koriat, Goldsmith, & Pansky, 2000; Schwartz, 1994). An important implication of this account is that, although individuals may initially retrieve accurate information from memory, the process of validating their memories can lead individuals to render incorrect memory judgments. For example, despite remembering that a change in policy has occurred, a voter’s knowledge that a politician has always kept his or her campaign promise in the past may lead the voter to doubt his or her memory, ultimately concluding that the politician did not change policy positions.

Accordingly, this account is highly relevant in domains in which there is a great deal of conceptual overlap between the pre- and post-change information (Gallo, 2006; Guerin, Robbins, Gilmore, & Schacter, 2012). Consider the statements...
“purchasing health insurance will be optional” (pre-change) and “purchasing health insurance will be mandatory” (post-change). The two statements share overlapping concepts (health insurance, healthcare) and differ only in one: the issue of necessity (optional/mandatory). Although individuals may possess memory for the pre-change information, they may ultimately rely on the large amount of conceptual similarities instead of the finer dissimilarities when rendering their memory judgments, thus leading them to erroneously conclude that a change has not occurred (Guerin et al., 2012).

To summarize, the two accounts specify different processes. Critically, the two accounts make the following distinct predictions:

Memory Storage Account: Individuals possess no memory for the pre-change political information. This leads them to erroneously conclude that no change has occurred.

Memory Monitoring Account: Individuals do possess memory for the pre-change information. Individuals erroneously conclude that no change has occurred because individuals mistakenly judge their memories of the pre-change political information to be inaccurate.

Distinguishing between these two accounts is important. Although the processes specified by each account lead to the same behavioral response—incorrectly judging that political information has not changed across time—the processes that generated these errors could yield different consequences for people’s confidence in the accuracy of their judgments. In particular, confidence judgments are thought to be based on one’s memory for studied information (Kelley & Lindsay, 1993; Robinson, Johnson, & Herndon, 1997; Robinson, Johnson, & Robertson, 2000) and other factors unrelated to memory (Chua, Hannula, & Ranganath, 2012). Under memory monitoring, individuals attempt to reconcile their memory of pre-change information with other pieces of information (e.g., conceptual overlap) that lead them to conclude that no change has occurred. Under the memory storage account, however, this conflict between distinct pieces of information does not occur because individuals do not have memory of the pre-change information. Because of the presence of two conflicting pieces of information, incorrect judgments that are due to the mechanisms specified by the memory monitoring account may elicit lower levels of confidence than judgments generated by failures in memory storage (Chua et al., 2012). These differences in confidence are important given previous work across different domains showing that one’s level of certainty or confidence in one’s beliefs or knowledge can influence behaviors (Geraerts et al., 2008; Locander & Hermann, 1979; Rains & Tukachinsky, 2014). For example, individuals who are uncertain about the validity of their beliefs are more likely to seek out additional information in order to reduce their feelings of uncertainty (Locander & Hermann, 1979).

Finally, it is important to note that in some instances, a failure to recognize changes in political information may be due to failures in memory storage, whereas
other instances are due to failures in memory monitoring. When communication
and public opinion scholars refer to or describe memory processes, they often implicitly subscribe to the mechanisms suggested by the memory storage account. That is, they assume individuals who provide no response in a recall test or an incorrect response in a recognition test do so because they possess no memory of a past event (Chaffee, Zhao, & Leshner, 1994; Eveland & Scheufele, 2000; Hollander, 2014; Price & Czilli, 1996; Prior, 2003). Our primary goal in this study, then, is to examine whether these errors can stem from a different source: memory monitoring.

Assessing memory via eye movements

Communication scholars have long assessed memory using self-reports. In recent years, multiple studies in cognitive psychology suggest that people’s memories of specific prior experiences can influence eye movements. As a consequence, eye movements can be used to indirectly assess memory and can complement self-report metrics. Eye movements possess two unique advantages. First, they can be used as an online assessment of memory processes without requiring self-report (Hannula, Baym, Warren, & Cohen, 2012; Hannula, Ryan, Tranel, & Cohen, 2007). Second, and critical to the design of this study, eye movements can accurately assess memory for information despite inaccuracies in self-report responses (for a review, see Hannula et al., 2010).

Indeed, converging evidence over the past 20 years suggests that previous exposure to visual stimuli exerts significant effects on eye movements to subsequent presentations of the same stimuli, revealing memory for past events despite erroneous self-report responses. Of particular interest here are eye movement studies in which participants were presented with photos of realistic scenes (e.g., a photo of two women walking side-by-side across a bridge; see Ryan, Althoff, Whitlow, & Cohen, 2000; Ryan & Cohen, 2004). After studying these photos, participants were then given a memory test in which they were either shown the original scene or a version of the scene in which some element had been changed/manipulated (e.g., the two women have been removed from the bridge). These studies found that participants directed their gaze disproportionately to regions of the scenes that had been changed (e.g., participants looked longer at the region in the scene in which the women were previously located). One explanation for this finding is that elements in the visual world that violate one’s expectations tend to elicit greater levels of attention (Loftus & Mackworth, 1978; Underwood & Foulsham, 2006). Individuals, then, direct greater attention to manipulated areas in scenes because they violate their expectations regarding what the scene should look like (e.g., they expected to see two women on the bridge). Critically, it logically follows that individuals can only form such an expectation if they possess intact memory of the original scene.

This effect of memory on eye movements has been documented even when participants fail to recognize—as measured by self-reports—that changes to the scene have been made (i.e., participants look longer at the modified region, even if they
incorrectly report that the scene has not been changed). Furthermore, some of the strongest evidence demonstrating the causal effects of memory on eye movements have come from studies of individuals with anterograde amnesia (Ryan et al., 2000). Amnesic patients are individuals with severe memory impairments—following damage to a brain region known as the hippocampus—that prevent them from gaining knowledge of facts and events after their brain lesion event (e.g., Cohen & Squire, 1980; Scoville & Milner, 1957). A pertinent study employing the paradigm mentioned above compared the eye movement behavior of amnesic patients with neurologically-intact comparison participants (e.g., matched on age, IQ, sex, etc.; Ryan et al., 2000). This study showed that amnesic patients do not demonstrate the eye movement patterns that are observed in neurologically-intact participants (i.e., increased gaze to the manipulated areas of scenes), suggesting that such eye movement behaviors critically depend on an intact memory system. In addition, this phenomenon, in which eye movements can provide an accurate assessment of memory independent of self-report responses, has been replicated in various contexts and domains (Hannula et al., 2012; Schwedes & Wentura, 2012, 2016). Eye movement metrics can thus provide communication scholars with an additional way of indirectly assessing memory retrieval processes at the point in time when these processes are occurring and, critically, they do not hinge on the (in)accuracy of self-report responses. However, to our knowledge, all existing work that has employed eye movements to assess memory utilized images as stimuli. One important contribution of our study, then, is in examining the extent to which the same eye movement effects observed for images can also be obtained when the stimuli consist of text.

Methods

Participants
We had a total of 137 participants who were recruited from a large public university in the United States and the surrounding community. All participants were compensated with $15 for taking part in the study. We excluded 9 participants who either began to drift asleep during our study, were wearing glasses or contacts that interfered with the calibration of our eye-tracking instruments, or encountered technical problems with the stimuli presentation computer (e.g., in one instance, the stimuli font changed color between study and test). We analyzed data from the remaining 128 participants (66 females; age: $M = 24.40$, $SD = 9.23$, range = 18–74).

Materials
The stimuli consisted of 32 policy statements regarding proposed changes to state policies (see Table 1). We chose this domain of policy facts because information about proposed policies and programs are some of the most important forms of
political knowledge that voters can possess (Delli Carpini & Keeter, 1996). Indeed, individuals often draw upon their specific knowledge of policies when making policy decisions via referenda or when voting for politicians to represent their policy interests (Boudreau & MacKenzie, 2014). Thus, their ability to vote for issues that advance their interests or to select politicians with issue stances that match their own are contingent on their capacity to possess accurate information about policies.

Each of the proposed state laws had two versions that differed in certain facts about the laws (64 policy statements in 32 pairs). These changes reflected a variety of potential key points regarding those laws. For example, changes included numerical values, the government agency affected by the change, qualities of citizens affected by the change, and qualifying words that changed the policy position (e.g., create/abolish; see Table 1). The section of the sentence that differs across the policy statement pairs is the critical region of interest. The policy statement pairs included the same total number of words (range = 11–28) and the same number of words in the critical region of interest (range = 1–8). Finally, the location of the critical region of interest was distributed evenly across the 64 policy issues (i.e.,

Table 1 Examples of Two-Version Policy Statements

<table>
<thead>
<tr>
<th>Version 1</th>
<th>Version 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>The state will allow 25% of public funds to be invested in <strong>private corporate stock</strong>.</td>
<td>The state will allow 25% of public funds to be invested in <strong>federal government bonds</strong>.</td>
</tr>
<tr>
<td>The state Constitution will be amended to limit the voting rights of <strong>incarcerated felons</strong>.</td>
<td>The state Constitution will be amended to limit the voting rights of <strong>homeless persons</strong>.</td>
</tr>
<tr>
<td>Lawmakers will <strong>create</strong> term limits for judges on the state Supreme Court.</td>
<td>Lawmakers will <strong>abolish</strong> term limits for judges on the state Supreme Court.</td>
</tr>
<tr>
<td>Any contributions to elected officials exceeding <strong>250</strong> dollars will have to be posted on the official’s website, along with the donor’s name.</td>
<td>Any contributions to elected officials exceeding <strong>950</strong> dollars will have to be posted on the official’s website, along with the donor’s name.</td>
</tr>
<tr>
<td>Individuals convicted of misdemeanors are required to provide their <strong>DNA to the criminal DNA database</strong>.</td>
<td>Individuals convicted of misdemeanors are required to provide their <strong>criminal history on a public webpage</strong>.</td>
</tr>
<tr>
<td>The state constitution may be amended so that the minimum age for serving in the state Senate is set to <strong>18-years-old</strong>.</td>
<td>The state constitution may be amended so that the minimum age for serving in the state Senate is set to <strong>21-years-old</strong>.</td>
</tr>
<tr>
<td>The state will soon shift taxes from <strong>small business owners on to homeowners</strong>.</td>
<td>The state will soon shift taxes from <strong>homeowners on to large corporate employers</strong>.</td>
</tr>
</tbody>
</table>

Note: Critical regions of interest are denoted in bold and italics. These words were in standard typeface when presented to participants.
approximately one-third each of the critical regions of interest were near the beginning of the sentence, in the middle, and near the end of the sentence).

Finally, we intentionally created fictitious state laws that were low-salience in this foundational study. We avoided high-salience issues (e.g., abortion) in order to increase the likelihood that, when assessing changes to the policy statements, individuals were drawing on information they learned in our study and not on information they learned prior to participating. In recent years, political communication scholars have raised concerns about pre-treatment effects (Druckman & Leeper, 2012). Pre-treatment effects transpire when prior events occurring outside the laboratory (e.g., media coverage on abortion) contaminates one’s experimental treatment. Our use of low-salience issues minimizes this possibility.

Procedure

Participants were tested individually in a quiet room, where they were seated 100 cm away from a 24-inch Asus VG248QE LCD monitor (resolution 1920 × 1080) with a refresh rate of 60 Hz. Before the experiment began, the desktop-mounted SR Research EyeLink 1000 eye tracker was fitted and calibrated for each subject with a 9-point calibration system. A rigid mount was used to keep the chin and forehead from moving. Recordings were taken from the right eye, except for 5 instances in which reflection off the participant’s glasses necessitated left eye recording.

The main experiment employed a straightforward study-test method. More generally, we exposed individuals to facts about policies at two points in time. At the first point in time (study phase), individuals were shown an original version of a policy. At the second time point (test phase), individuals were exposed to either a modified version of the policy, in which an important feature had been changed, or its original version. Individuals were then asked, via self-report, to render a judgment as to whether the policy had or had not changed.

Participants were instructed at the start of the study phase that they would be reading about real policy proposals and were told to pay attention during the first phase of the study, as questions would be asked about them later in the study. Each trial began with a drift-check target, in the form of a fixation cross in the middle of the screen (+; see Figure 1). Participants controlled the time spent on this screen by fixating on the target while pressing the advance button on the left side of the hand-held controller. Participants were then presented with a policy statement in the center of the screen for 6 seconds (see Figure 1). During the study phase, participants were shown a total of 32 unique policy statements. The order of the policies was randomized for each participant. After completion of the study phase, participants were allowed to take a short break (i.e., 1 to 3 minutes).

In the test phase, participants were told that they would be taking a recognition memory test of the policy statements from the previous phase. Participants were shown a total of 32 policy statements during the test phase. Sixteen of the policy statements were old, in that they were exactly the same policy versions participants
were presented during the study phase. The remaining sixteen policy statements were new, in that participants were shown different versions of the policy statements from the ones presented during the study phase (i.e., the content of the critical region of interest had been changed).

We also counterbalanced the policy statements across participants. For example, if one participant was shown the issue “The state will allow 25% of public funds to be invested in private corporate stock” during the study phase and “The state will allow 25% of public funds to be invested in federal government bonds” during the test phase, the participant reporting that the policy had not changed.

Figure 1 (a) Schematic design of the study and test phases. (b) Example of a false alarm trial for one of the policies for a single participant. Dashed squares are the critical regions. The text in the critical region was changed from “250” at the study phase to “950” at the test phase. Eye movements are superimposed on the text. Circles represent fixations. Larger circles represent longer fixation durations. The critical region during the test phase elicited a greater amount of fixations and fixation durations than the critical region during the study phase, despite the participant reporting that the policy had not changed.
during the test phase, a different participant would be shown the statements in reverse, with “The state will allow 25% of public funds to be invested in federal government bonds” displayed during the study phase and “The state will allow 25% of public funds to be invested in private corporate stock” during the test phase. We incorporated this feature in our design in order to ensure that any observed differences in gaze to the critical regions between the study and test phases were not due to the differences in the size of the text across the two phases.

As shown in Figure 1, each trial began with a policy statement. Afterwards, participants were instructed to report, via a button press, whether the policy statement was old or new: that is, whether or not they had been shown that exact policy statement during the study phase. Immediately following each old/new decision, participants also made a confidence judgment (“somewhat sure” or “very sure”) on their recognition decision. The button used for the old/new and confidence judgments were counterbalanced between participants: for example, some were instructed to use the right button for old and some to use the left.

After the eye movement task, participants answered a series of demographic questions (age, gender, and political affiliation). They also answered a 20-item political knowledge questionnaire, which tested the participants’ knowledge about general and policy-specific facts (e.g., names of political figures currently holding office); these items were inspired by ones developed by Delli Carpini and Keeter (1996). We used this as a measure of political sophistication. Politically sophisticated citizens are commonly viewed as ones who possess a large and wide-ranging base of knowledge about politics and are engaged with, and motivated to learn about, the political system (Converse, 1964; Delli Carpini & Keeter, 1996; Luskin, 1987).

Analysis strategy
The recognition memory test generated four types of trials: hits (old/repeated policies correctly classified as old), misses (old/repeated policies incorrectly classified as new), correct rejections (new/modified policies correctly classified as new), and false alarms (new/modified policies incorrectly classified as old; these are terms typically used in the memory literature). Given the study’s substantive focus on dynamic political environments, of critical interest here are the false alarms. These are instances in which a policy had been modified and participants incorrectly judged, as reflected in self-reports, that the policy had not been changed. The memory storage account predicts that this failure is due to participants not possessing memory of the pre-change policy information. In contrast, the memory monitoring account predicts that participants do possess memory of the pre-change policy information, and that failure occurs because individuals mistakenly judge their memories of the pre-change information to be inaccurate.

To distinguish between these two processes, we operationalized gaze to the critical region of interest in two ways: (a) as the number of total fixations and (b) as the total duration of fixations directed to that critical region in both the study and test phases of the false alarm trials. One of the functions of fixations is to place
information in the environment within our foveal vision (which corresponds to the center of our gaze) where visual acuity is the highest (Rayner, 1998). Fixation duration corresponds to the amount of time that the fovea is directed at a specific location in the visual world. Although these concepts of number of fixations and fixation duration are conceptually distinct, they can be highly related, in that specific elements of our visual world that violate our expectations (e.g., expectations based on our prior experiences or memory) can attract disproportionate attention either by (a) increasing sampling behavior of these regions via an increase in the number of fixations directed to these locations or (b) an increase in the amount of time that a fixation is directed to a specific element in the visual world. Indeed, the number of total fixations and total fixation durations are common operationalizations of gaze in eye movement studies of memory and are often highly correlated and yield similar results (Hannula et al., 2007, 2010; Ryan et al., 2000; Ryan & Cohen, 2004). Similar to other eye movement studies on text comprehension (Stites & Federmeier, 2015), fixations which lasted under 80 ms or over 800 ms were discarded. In the context of reading, single fixations under 80 ms are unlikely to represent meaningful cognitive processing (Rayner, 1998), and fixations greater than 800 ms often do not represent normal acquisition of information from text (Rayner, Sereno, Morris, Schmauder, & Clifton, 2007). On average, 5% of fixations were excluded from the analysis.

If individuals do not possess memory for the pre-change policy information (as predicted by the memory storage account), then we expect the amount of gaze directed to the critical region during testing to be equal to the amount of gaze directed to the critical region during the study phase. In this case, given that they possess no memory of the critical region during the study phase, individuals do not have an expectation regarding what information should be contained in the critical region presented in the test phase. Thus, there is no reason why they should direct greater attention to this region during the test phase.

In contrast, if individuals possess memory of the pre-change policy information (as predicted by the memory monitoring account), then given previous findings that eye movements can accurately represent memory despite erroneous self-report responses (Hannula et al., 2012; Schwedes & Wentura, 2012, 2016), we expect the amount of gaze directed to the critical region during the test phase to be greater than the amount of gaze directed to the critical region during the study phase. This result would suggest that individuals formed an expectation regarding what information should be contained in the critical region in the test phase. This expectation is based on their memory of the information contained in the critical region during the study phase. Violating this expectation by changing the text in the critical region during the test phase elicits greater attention to that region. By comparing gaze directed to the critical region in both the test and study phases, our design allows us to determine whether the failures to recognize changes (false alarms) are due primarily to the mechanisms specified by the memory storage or the memory monitoring account.
Results

Recognition memory
Our first set of analyses examined the extent to which individuals were able to reliably recognize when policy changes occurred and to determine the prevalence of false alarms (our main trials of interest). Overall, participants demonstrated robust recognition memory performance, given that their hit rate (proportion of old items correctly classified as old, $M = 71\%$) was greater than their false alarm rate (proportion of new items incorrectly classified as old, $M = 20\%$). In the context of overall good memory for the issues, however, we nevertheless observed relatively high rates of false alarms for the new/modified policies ($M = 20\%, SD = 14\%$). Our primary interest was to determine the extent to which the false alarms might have been generated by the mechanisms specified by the memory monitoring account.

Eye movements to false alarms
To test whether the preponderance of false alarms in our study were generated largely by the mechanisms specified by the memory storage or memory monitoring accounts, we used the test phase as an independent variable (test phase = 1, study phase = 0) and total fixations and total fixation duration to the critical region as dependent variables. We estimated mixed-effects models in which the test phase was treated as a fixed effect and participants and policy issues were treated as random effects. A significant and positive effect of the test phase across our two operationalizations of gaze (see Table 2) suggests that participants were more likely to direct their gaze to the critical region in the test phase than in the study phase (also see Figure 1 for an example). This outcome is consistent with the predictions of the memory monitoring account.

Table 2  Gaze to Critical Region in False Alarm Items

<table>
<thead>
<tr>
<th></th>
<th>Total Fixations</th>
<th>Total Fixation Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1</td>
<td>Model 2</td>
</tr>
<tr>
<td>Fixed Effects</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test Phase</td>
<td>0.33* (0.13)</td>
<td>0.33* (0.13)</td>
</tr>
<tr>
<td>Age</td>
<td>−0.01 (0.01)</td>
<td></td>
</tr>
<tr>
<td>Political Knowledge</td>
<td>0.04 (0.02)</td>
<td></td>
</tr>
<tr>
<td>Random Effects</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participants</td>
<td>0.55</td>
<td>0.54</td>
</tr>
<tr>
<td>Items</td>
<td>2.54</td>
<td>2.54</td>
</tr>
</tbody>
</table>

*Note:* "Test Phase" variable: study phase = 0; test phase = 1. Regression coefficients are shown with standard errors in parentheses. We had 117 participants, given that 11 participants had a zero false alarm rate. Values in the random effects columns are estimated standard deviations of the intercepts for participants and items. *$p < .05$. 
We estimated a second model in which we included several control variables. Because the integrity of memory function may change as people get older, we also included age in our analysis (Ryan, Hannula, & Cohen, 2007). We also included political sophistication as a control variable. Inclusion of these additional variables in the model produced substantively similar effects for the test phase variable (Table 2). Our results, therefore, are robust across different model specifications and operationalizations of gaze.11

Eye movement monitoring as a valid method for assessing memory
We also conducted additional analyses in order to determine the extent to which the amount of gaze to the critical region during the test phase was a valid assessment of memory for political information contained in the critical region during the study phase. More specifically, we examined eye movement responses to correct rejections and hits.

Correct rejections are instances in which individuals correctly classified a modified policy as having been changed. This suggests that individuals possess an expectation of what information should be contained in the critical region during the test phase, given that they possess memory of the critical region during the study phase. Violating this expectation in the critical region during the test phase should elicit greater attention to that region. Thus, for correct rejection trials, we expect gaze to the critical region during the test phase to be greater than gaze to the critical region during the study phase.

For correct rejection trials, we used the test phase as an independent variable (test phase = 1, study phase = 0) and total fixations and total fixation duration to the critical region as dependent variables. A significant and positive effect of test phase (total fixations as dependent variable: $B = 0.88$, $SE = 0.07$, $p < .001$; total fixation duration as dependent variable: $B = 217.29$, $SE = 18.56$, $p < .001$) indicates that participants, as expected, directed a greater amount of attention to the critical region during the test phase, compared to the study phase, for correct rejections.

Next, hits are instances in which individuals correctly classified a repeated/old policy as unchanged. In this case, their expectation of the information contained in the critical region during the test phase is not violated, as the repeated information comports with their memory for the information during the study phase. Thus, there is no reason why they should direct greater attention to the critical region during the test phase. For hit trials, then, we expect gaze to the critical region during the test phase to be equal to the critical region during the study phase.

For hit trials, we used the test phase as an independent variable (test phase = 1, study phase = 0) and total fixations and total fixation duration to the critical region as dependent variables. As expected, a non-significant effect of test phase (total fixations as dependent variable: $B = -0.03$, $SE = 0.07$, $p = .68$; total fixation duration as dependent variable: $B = -24.04$, $SE = 17.62$, $p = .17$) indicates that participants directed an equal amount of gaze to the critical region for the test and study phases. Taken together, these results are consistent with the claim that the amount
of gaze directed to the critical region during the test phase is an accurate assessment of people’s memory of the information contained in the critical region during the study phase.

**Effects on confidence**

Finally, we conducted exploratory analyses in order to examine the effects of memory storage and monitoring processes on people’s confidence in their recognition judgments. As mentioned previously, incorrect recognition judgments that are due to the mechanisms specified by the memory monitoring account may elicit lower levels of confidence than judgments generated by failures in memory storage. In the former, individuals are attempting to reconcile two conflicting pieces of information (e.g., memory for the pre-change information and other factors, such as conceptual overlap), whereas this conflict does not exist in the latter. Given our results, it is likely the case that a preponderance of false alarms in our study were due to the mechanisms specified by the memory monitoring account. However, some of the false alarms may have been generated due to failures in memory storage.

In our design, then, we assumed that gaze to the critical region during the test phase for the false alarm trials can serve as an indicator of the presence or absence of memory for information contained in the critical region during the study phase. As gaze to the critical region in the test phase increased, we assumed that individuals were more likely to possess memory for the information contained in the critical region during the study phase (i.e., memory for pre-change information leads to a violation of expectation, which in turn increases attention to the critical region). If incorrect recognition judgments due to the mechanisms specified by the memory monitoring account elicit lower levels of confidence than judgments generated by failures in memory storage, then we expected an increase in gazes to the critical region during the test phase to be negatively associated with people’s level of confidence for false alarm trials. Critically, correct rejections serve as an important comparison, given that these were trials in which individuals were able to render accurate recognition judgments because individuals were primarily relying on their memory of the pre-change information (instead of other pieces of information) to inform their recognition judgments. Thus, for correct rejections, an increase in gaze to the critical region during the test phase should be positively associated with people’s level of confidence.

To explore these possibilities, we estimated logistic mixed-effects models in which we modeled gaze (operationalized as the total number of fixations and total fixation duration) and trial type (correct rejections = 0, false alarms = 1) and the interaction between the two as our independent variables. Our dependent variable was the level of confidence for participants’ old/new judgments (0 = Somewhat Confident, 1 = Very Confident). We obtained a negative and significant gaze by trial type interaction (total fixations as dependent variable: \( B = -0.16, SE = 0.04, p < .001 \); total fixation duration as dependent variable: \( B = -0.001, SE = 0.0002, \)
To interpret this interaction, we created predicted probability plots (Figure 2). As expected, an increase in gaze to the critical region during the test phase was associated with lower levels of confidence for recognition judgments among false alarms. In contrast, an increase in gaze to the critical region during the test phase was associated with higher levels of confidence for recognition judgments among correct rejections.

Discussion

A critical and constant feature of the information environment is its dynamic nature. Arguably, important forms of political information that voters need to know are proposed policies and programs, which often change across time. Such shifts in political information make it difficult for individuals to possess accurate political knowledge.

This study advances our understanding of how people process changes in political information by examining instances in which important features of policy have been altered upon subsequent exposure and, critically, investigating why individuals fail to recognize these changes in the political environment. We reasoned
that individuals fail to recognize changes in dynamic environments if they possess no memory of the pre-change policy information (memory storage account) or, if they do possess memory for pre-change policy information, they mistakenly judge their memories of the pre-change information to be inaccurate (memory monitoring account). We examined instances in which changes in policies had occurred but individuals erroneously reported that no changes had occurred (false alarm trials). We used eye movements to determine the extent to which individuals did or did not possess memory for pre-change information, despite reporting that no change had occurred.

Critically, we found that participants were more likely to direct a greater amount of gaze to the critical region in the test phase than in the study phase among these false alarm items. This outcome is consistent with the predictions of the memory monitoring account. That is, individuals did possess memory for the pre-change political information, but were likely influenced by other types of information (e.g., the conceptual overlap between the pre- and post-change information) that led them to an incorrect memory judgment (as reflected in their self-reports).

Given our findings, our study has several substantive and methodological contributions. First, our study suggests that people’s inability to detect changes in political information (as measured by self-report) can also be due to failures in people’s ability to accurately assess the validity of their retrieved memories. As mentioned previously, communication and public opinion scholars often implicitly subscribe to the mechanisms suggested by the memory storage account when describing memory processes. Our findings suggest that lack of memory is not the only explanation for these failures in self-reported memory judgments. Our study advances the literature by highlighting the role of memory monitoring when individuals remember political information.

Furthermore, incorrect recognition judgments that are due to the mechanisms specified by the memory monitoring account can, as indicated by our exploratory analyses, elicit lower levels of confidence than judgments generated by failures in memory storage. These differences in confidence are important given previous work showing that confidence judgments can influence behaviors (Geraerts et al., 2008; Locander & Hermann, 1979; Rains & Tukachinsky, 2014). For example, individuals may be more likely to engage in information searches (e.g., online) if they are not confident in the accuracy of their judgment that political information has not changed across time (Locander & Hermann, 1979).

The memory monitoring account described here comports with theoretical views in the memory literature (e.g., Moscovitch, 2008) suggesting that memory retrieval occurs in two stages. According to this perspective, during the first stage of memory retrieval, stored information is initially retrieved quickly and automatically from memory (i.e., without deliberate intention). However, in the second stage, memory retrieval involves more controlled and deliberative/intentional processing, in which individuals make selective and strategic decisions during memory
search and retrieval. It is in this second, slower stage, however, that other cognitive/affective processes intervene (e.g., people begin to use other types of information to inform their memory judgments), steering participants—despite initially retrieving an accurate memory representation—into making an erroneous memory judgment.

One such intervening process that future work can explore is motivated reasoning, or people’s tendency to maintain their existing attitudes in the face of information that challenges their beliefs (Kunda, 1990). In the real world, political information that changes across time often pertains to a political actor (e.g., politician, government agency, etc.). Furthermore, it is likely the case that individuals have prior affective reactions to these political actors. For example, we expect that people’s positive emotional associations toward a political actor (e.g., a member of the U.S. Congress) can lead them to be more likely to doubt their memories of negative information disseminated about the politician via processes related to defensive reactance (i.e., counterarguing). Such salient emotional associations can cause people to doubt their memories of pre-change information, potentially leading to more instances of failures of memory monitoring.

In terms of methodology, our novel contribution is the use of eye movement assessments of memory to offer unique leverage when used in conjunction with self-reports. Critically, our study shows that eye movements can reveal memory for political information despite erroneous self-report responses. To our knowledge, communication scholars have yet to consider the effects of memory on eye movements or to use eye movements as a method for indirectly assessing memory. Indeed, political communication scholars have primarily used eye movements to examine how factors such as personal preferences (e.g., political attitudes, satisfaction from reading the news, etc.) influence attention to content and cues in the information environment (Bode, Vraga, & Troller-Renfree, 2017; Dvir-Gvirsman, 2017; Marquart, Matthes, & Rapp, 2016; Vraga, Bode, & Troller-Renfree, 2016). Our study, then, follows a long tradition in communication research of developing tools that do not rely entirely on the accuracy of self-report responses (Lang, Potter, & Bolls, 2009).

Furthermore, to our knowledge, all existing work that has employed eye movements to assess memory has utilized images as stimuli (for a review, see Hannula et al., 2010). Given that much of the information in the political environment is conveyed via text (e.g., websites, newspapers, etc.), our study contributes to the literature by establishing that the same eye movement effects observed for images can also be obtained for text.

Although this study is a useful first step, caution is warranted in terms of generalizing some of the study’s findings. Our study does not directly examine extended periods of learning. In this study, participants were tested a few minutes after they first learned about the issues. In the real world, participants may be exposed to political information from the media environment and not need to retrieve that stored information for weeks, or perhaps longer. In addition, individuals are often
exposed to political information at multiple points in time. Further research should determine the capacity of eye movements to assess memory for political information when a longer timescale is taken into account or when individuals receive multiple exposures to the same (or related) information. Future work should also examine the extent to which our eye movement metrics are assessing either explicit/declarative (memories that are accessible to conscious awareness) or implicit/non-declarative memories (memories that are not accessible to conscious awareness). This question is still under debate in the eye movement and memory literature (Hannula et al., 2010; Kumaran & Wagner, 2009). Irrespective of whether conscious awareness plays a role, our exploratory analyses indicate that the absence and presence of memories for pre-change information can influence confidence judgments.

Furthermore, open questions remain regarding whether the effects we observed here can be obtained in more ecologically valid settings in which individuals are exposed to more than one or two sentences of text. For example, government officials (e.g., congressional staffers) have been known to strategically alter Wikipedia pages of members of Congress (Anderson, 2006). Additionally, the policy issues used in this initial study were constructed to be of low salience. We intentionally made this decision in order to avoid problems associated with pre-treatment effects (Druckman & Leeper, 2012). As a consequence, the messages are not representative of the broad array of messages that populate the current political information environment. Future work should investigate the extent to which the effects we observe here are also obtained for highly salient, partisan, and affectively-charged policy issues. In addition, future research should examine how individuals evaluate political information in dynamic environments via other modalities (e.g., images, video). Our participants are also not a nationally representative sample.

Despite these limitations, our study further highlights the utility of a converging methods approach for investigations of message processing in dynamic information environments. As shown here, eye movements can be used by communication scholars to assess memory in conjunction with self-report measures. More broadly, there are two developments in eye movement monitoring research that can make the technology a cost-efficient and scalable method in the future: (a) the widespread use of smartphones and (b) the ongoing miniaturization of eye movement monitoring devices. For example, computer scientists are currently developing software that can convert a smartphone into an eye movement monitoring device (Hardesty, 2016). As more individuals use their smartphones to search for and acquire political information, eye movement monitoring can potentially become a scalable method that can be used to determine how large groups of individuals consume political information (i.e., determine what types of information elicits attention, etc.) transmitted over their smartphones.

Finally, the utility of eye movements goes beyond the fields of political communication and public opinion. Eye movements can be used by communication scholars interested in determining the types of information individuals retrieve from...
memory in dynamic information environments. They are potentially useful across a wide range of domains, providing unique information on the extent to which individuals possess intact memories for message information that changes across time. Examples of such domains include health communication (e.g., health guidelines that change across time; Allen et al., 2013) and science communication (e.g., information about risks associated with technology that change across time; Loewenstein & Mather, 1990).

In summary, our results suggest that individuals can make erroneous memory judgments despite possessing memories of relevant political information. People's ability to accurately determine whether political information has changed across time depends not only on their actual memories, but also on their assessment of the validity of such memories. These findings are pertinent to the tidal wave of information that voters receive in their everyday lives; the political environment is characterized by a constant dynamic and ever-changing nature. The ability of voters to accurately recognize these changes and their confidence in their judgments determine the extent to which they are informed or misinformed about the political world.

Notes

1 Validation of retrieved memories is often referred to as “metamemory” (Schwartz, 1994) processes in the cognitive psychology literature.

2 Some researchers have proposed that the deployment of attention to novel and unexpected events in the environment evolved as part of a motivational and defensive system meant to increase an organism’s chances of survival (Bradley, 2009).

3 Participants were recruited by posting flyers around the campus of a large university. Additionally, flyers were emailed to instructors of classes which met in-person on campus, to increase our visibility. These flyers contained the email address of the lab as well as the location where the study was being conducted. Included in the flyer was information pertaining to eligibility requirements, compensation, and the anticipated duration of the study. The only phrase on the flyer which pertained to the purpose of the study was the phrase “The purpose of this study is to examine how people make political decisions.” Data collection was conducted between March and September, 2016. Participants were scheduled during normal business hours (9 a.m. to 5 p.m.).

4 To clarify, it was not the case that participants actually slept through a large portion of the study, but rather began “drifting” asleep to the point where their eyes could no longer be tracked accurately by the eye movement monitoring system.

5 We manually defined regions of interest around words that constitute our critical regions of interest prior to data collection. A better estimate of the size of our areas of interest would be the number of characters in our regions of interest (as opposed to the number of words). We include spaces and periods in our character count. Overall, our average character count for our critical regions is 24.23 (SD = 10.74). Version 1’s character count (M = 23.66, SD = 10.67) did not differ significantly from version 2’s character count (M = 24.81, SD = 10.94; t[62] = -0.43, p = .67).

6 Note that we assessed the amount of time it took participants to read the entire sentence by measuring how long it took participants to direct their gaze to the final word in the sentence.
On average, it took participants 3.68 seconds (SD = 468 milliseconds) until their gaze reached the final word, suggesting that participants had enough time to read the entire sentence.

None of the political knowledge questionnaires were related to any of the policies we used during the eye movement task.

Recognition memory was assessed using the discriminability index d-prime, a statistic commonly used in the memory literature to assess recognition memory (MacMillan & Creelman, 2005). As a group, participants demonstrated above-chance performance (a d-prime score above zero) on recognition memory (M d’ = 1.55, SD = 0.75, range −0.5–3.7), t[127] = 23.28, p < .001.

Mixed-effects models are useful when observations are characterized by two types of dependencies (Baayen, Davidson, & Bates, 2008). First, observations within a participant are not independent. Second, if the participants are exposed to the same stimuli (e.g., our 64 policies), then the observations within a stimulus are also not independent. Failure to account for this interdependence in observations within participants and stimuli can lead to spurious results (Quené & van den Bergh, 2004). Mixed-effects models account for the interdependence of observations using random effects that correspond to the grouping variables (participants and stimuli; Baayen et al., 2008).

Note that the character count in our critical region of interest during the study phase (M = 24.23, SD = 10.65) did not significantly differ from the character count during the test phase (M = 24.24, SD = 10.65 t[8166] = 0.04, p = .97).

Another benefit of using mixed-effects models is that they allow one to model both by-participant and by-item (i.e., stimuli) variations within a single statistical model by modeling random intercepts for each participant and item (Baayen et al., 2008). As can be seen in Table 2, values in the random effects columns are estimated standard deviations of the intercepts for participants and items, indicating how much of the variability in the dependent measure is due to participants and items. When examining these values, it can be seen that items have more variability than participants. This suggests that our effects are likely sensitive to the type of change. Future work should systematically examine the effect of one type of change on people’s ability to recognize changes in political information.

References


Kumaran, D., & Wagner, A. D. (2009). It’s my eyes, but it doesn’t look that way to me. Neuron, 63, 561–563. doi:10.1016/j.neuron.2009.08.027


