저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.

변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.
The Effect of Test Priming on Subsequent Reading Behavior

읽기 과정에 미치는 시험의 프라이밍 효과

2018년 8월

서울대학교 대학원

교육학과 교육공학 전공

성 가 현
Abstract

To effectively design learning, instructional designers must understand how a specific intervention has what effect on a learner. One common intervention whose effect is multifaceted and often contested is testing. While tests are an important tool in understanding learners and fostering learning, not all its effects are positive. Many studies have investigated the learning gains of testing though another test. However, less attention has been given to why these gains may occur, or what effects it may have on the learning process.

Thus, the current study aims to understand the effect the existence and type of testing has on one crucial learning process, reading. There are research with similar intentions that have depended on learners’ self-reports, but the current study aims to produce precise, quantifiable results through eye-tracking methodologies. In the experiment, eye-tracking, interview and test result data was collected on participants randomly assigned to a memory-based test, higher-order thinking-based test or no test/control condition. The reading patterns of participants before and after testing were recorded and analyzed together with qualitative observations.

Results confirm that tests alter reading strategies for subsequent text. Key findings include a decrease in overall reading time and fixation for the no test condition while no such change was found for both test conditions.
Readers in the memory test condition showed an increased tendency to pay more attention to definitions, keywords, and relations, and more regressive reading behavior. Readers in the higher-order thinking test condition paid more attention to relations. Overall, readers in the memory test condition seem to show the most change in reading behaviors. Possible explanations for this phenomenon were given, including preexisting expectancies, affective reactions to memory-based questions, and the nature of higher-order thinking questions used in the present study. Future directions for research are suggested based on the findings.

Keyword: Testing, global text processing, reading strategy, reading purpose, eye-tracking, expository text

Student number: 2016-21518
# Table of Contents

## I. INTRODUCTION

- Purpose of the Study ................................................................. 1
- Research Question ........................................................................ 4
- Definition of Terms ....................................................................... 4
  - Assessment and Testing ............................................................... 4
  - Reading Strategies ......................................................................... 6
  - Eye Movement and Fixation ......................................................... 6
  - Higher-order thinking ................................................................... 7

## II. REVIEW OF THE LITERATURE

- The Process of Reading ............................................................... 8
  - Information Processing Theory .................................................... 9
  - Bottom-up Model: The Construction-Integration Model .................. 10
  - Top-down model: Constructionist Model ....................................... 12
  - Hybrid Model: The Landscape Model ............................................ 13
- Effect of Learner Elements on Reading ......................................... 15
  - Background Knowledge .............................................................. 16
  - Individual Differences in Working Memory and Linguistic Ability ... 17
  - Reading Purpose ........................................................................... 19
- Priming Effect of Learning Interventions on Reading .................... 21
  - Priming ....................................................................................... 21
  - Learning Activity ........................................................................... 22
  - Tests ............................................................................................. 23
- Eye-tracking Methodology in Educational Research ...................... 25

## III. HYPOTHESES

## IV. METHODS

- Participants .................................................................................... 30
- Experimental Process ...................................................................... 31
- Experimental Tools .......................................................................... 33
  - Text Stimuli ................................................................................ 33
  - Tests ............................................................................................ 36
3.3. Interview .............................................................................................................. 37
4. Data Acquisition and Analysis .............................................................................. 38
V. RESULTS .................................................................................................................. 42
  1. Eye-tracking Results ............................................................................................. 42
     1.1. Memory-based test group .............................................................................. 42
     1.2. Transfer-based test group ............................................................................ 44
     1.3. No test group ................................................................................................ 45
  2. Test Results ........................................................................................................... 46
VI. DISCUSSION AND FUTURE DIRECTIONS ...................................................... 47
VII. SUMMARY AND CONCLUSION ....................................................................... 52
References .................................................................................................................. 54
Appendix ..................................................................................................................... 67

국문 초록 ...................................................................................................................... 86
List of Tables

Table 4.1. Steps of Experiment and Time Spent on Each ............................................. 32
Table 4.2. Parameters of Eye-tracking Data, in Data Pooled across ROI.......................... 41
Table 5.1. Time Spent Reading, Memory Group .............................................................. 43
Table 5.2. Number of Fixations, Memory Group .............................................................. 42
Table 5.3. Time Spent Reading, Transfer Group .............................................................. 47
Table 5.4. Number of Fixations, Transfer Group .............................................................. 44
Table 5.5. Time Spent Reading, No Test Group ............................................................... 45
Table 5.6. Number of Fixations, No Test Group .............................................................. 45
Table 5.7. Results of the Two Tests Across Groups .......................................................... 46
Table 6.1. Summary of Discussions: Key Patterns and Possible Explanations ................. 49

List of Figures

Figure 2.1. Information Processing Theory ..................................................................... 10
Figure 2.2. The Cognitive Landscape during Reading of a Text .................................... 14
Figure 3.1. Sample Pages from the Text Stimuli ............................................................ 36
Figure 3.2. Schematized Eye-tracking Results ............................................................... 42
I. INTRODUCTION

1. Purpose of the Study

In recent years, educational researchers has been giving increasing amounts of attention to how an intervention works, together with how well it works. Determining the effectiveness of a treatment is a necessary focus of fields aiming to improve performance, as is the field of education. But now, noticeably more effort is being devoted to understanding when and how an intervention has an effect on the process of learning (Lai et al., 2013; Schmit, Rotgans & Yew, 2011).

This change belongs to a series of coextensive adjustments the paradigm of learning and teaching has seen since the late twentieth century. Behaviorism, stressing observable behaviors, gave way to cognitivism and constructivism, which concentrates more on the internal and mental process of learning (Ertmer & Newby, 1993; Jonassen, 1990). More effort was accordingly paid to investigating different sub-processes of learning, such as problem solving (Anderson, 1993), feedback cycles (Shute, 2008) and the interaction between learners, instructors, and the learning material (Moore, 1989).

Riding this wave, learner-centered pedagogies then became one of the most extensively researched topics in education, and under this paradigm it
was deemed necessary that an educational intervention be valued and understood from the point of view of the individual learner (Lamber & McCombs, 1998; Smart, Witt & Scott, 2012). Instead of constituting a formal, rigid curriculum to which students must abide by or fail (Zimmerman, 2002), interventions now are expected to mold accordingly to the needs of the learner – and to do this, educators must understand what the individual learner wants, how he/she perceives an intervention, and reacts to it.

There are few interventions that merits more investigation into its workings and reception with learners than testing. Testing is a powerful tool for learning when used properly, as evidenced by the unequivocally positive effect it has on retention and under some circumstances transfer (McDaniel et al., 2007; Carpenter, 2012), the success of the paradigm that assessment must facilitate learning rather than solely testing what has been learned (assessment for learning; Black et al., 2004), and not least, its inclusion into almost every forms of education since antiquity.

Yet testing has often been documented to have unwanted side effects, such as test anxiety (Hill & Wigfield, 1984), negative impact on intrinsic motivation (Harlen & Crick, 2003), and undervaluation of the less measureable aspects of learning from students and educators alike (Kohn, 2000). Thus, role that testing must have in the classroom today is still a fiercely contested subject, and there is a lack of distinct guiding principles as to how and when testing must be employed when designing learning experiences. A better understanding of the context and mechanisms of the
effect that testing has can allow instructional designers to fine-tune test interventions to maximize benefits while minimizing side effects.

The current research looks at one such purported mechanism, the ‘priming’ effect testing has on the reading behavior of learners. That is, a test can induce learners to adjust their metacognitive strategies accordingly. While there has been theoretical suppositions (Wissman, Rawson and Pyc, 2011) related to this mechanism, very few research look at the actual online reading behavior of the learner, and fewer still observe the process with precise, objective and quantifiable means. Thus, this study aspires to help clarify one mechanism in the effect testing has on learning, through the quantitative methodology of eye-tracking.

Eye-tracking is a widely used tool in reading research, due to its usefulness in precisely capturing and quantifying learners’ process of reading. As it is, the current study is also a part of a different line of effort that aims to better understand the learner using new technologies and data-driven methods. New technologies have the capacity to automate some parts of learner observation, or to capture minute details unobservable to the human eye. In this way, they allow an objective, detailed look into each step of the learning process. While still nascent in their form and generality, efforts have been made to harness such technologies in order to acquire newer, real-time, and more detailed information on learners and their learning trajectories.

Work on this front, ultimately, could lead to the practical evaluation of a test by what type of, and how much, learning it induces – and conversely
enable a way of assessing learner progress with less tests, and more with the data they produce.

2. Research Question

The present study aims to understand the way testing has an effect on reading behavior, as well as the mechanisms through which such effects occur. Additionally, it strives to determine whether different types of tests bring about different reading strategies and attitudes in learners, in order to aid the implementation of tests in instructional design. The specific research questions devised to investigate these issues are as follows:

One, do learners demonstrate improved attention and effort in reading after testing?

Two, do different types of test result in different kinds of reading behavior? Namely, do tests in which performance depends solely on rote memory retrieval prompt different reading behavior (i.e. reading pattern, time, and attitude) when compared to tests requiring higher-order thinking (levels 3-6 in Bloom’s taxonomy: application, analysis, evaluation, synthesis)?

3. Definition of Terms

3.1. Assessment and Testing
Assessment and testing are two terms that overlap and often used interchangeably, yet have distinctively different overtones and documented usage in previous research. Assessment is often larger in scope, and deals with all manner of evaluation, for instance those carried out in pen and paper, observation on student behavior, or prompting a verbal response. On the other hand, testing is usually referred to one method of assessment that has formal, specified procedures, most often in the form of a set of questions, and with standardized ways of scoring. The current research focuses on testing.

For a more concrete differentiation, the definitions set forth by the Task Group on Assessment and Testing (TGAT) in 1988 in a seminal report is consulted. In it, assessment is: “[A] general term that embraces all methods customarily used to appraise performance of an individual pupil or group. It may refer to a broad appraisal including many sources of evidence and many aspects of a pupil’s knowledge, understanding, skills and attitudes; or to a particular occasion or instrument. An assessment instrument may be any method or procedure, formal or informal, for producing information about pupils: for example, a written test paper, an interview schedule, a measurement task using equipment, a class quiz.”

On the other hand, testing is: “Strictly, any assessment conducted within formal and specified procedures, designed to ensure comparability of results between different test administrators and between different test occasions. For some it implies a set of written questions, externally prescribed, with written responses marked according to rigid rules; for others, any of a
broad range of assessment instruments with standardised rules of administration and marking which ensure comparability of results.”

3.2. Reading Strategies

Reading strategies include all actions of the reader that aim to “improve some aspect of comprehension” (Graesser, 2007). This action is both cognitive and behavioral, and could be consciously or unconsciously carried out. The term ‘some aspect of comprehension’ is used in place of comprehension because certain types of reading purpose and context dictate reading strategies such as skimming or selective reading. That is, reading strategies also include actions that purposefully lessen the scope or depth of comprehension.

3.3. Eye Movement and Fixation

The eyes, rather than moving along a steady line, make ‘jumps’ and ‘stops’ when reading. That is, due to the rapid decline in visual resolution away from the fovea, the human eye must make short, quick jumps in visual attention, which is called a saccade. During a saccade, the eye cannot register any visual information (Rayner, Foorman, Perfetti, Pesetsky & Seidenberg, 2001). Saccades are intermingled with fixations which is the eye stopping on one point of the text. During fixation, the eye can take in the visual information in its field of attention. Around the fixation point, only four to
five letters can be seen with perfect acuity (Rayner, 1975), yet readers can pick up information such as word-length pattern and paragraph breaks much further outside the immediate point of fixation. For example, a person can recognize word-length patterns at least 12 to 15 character positions to the right of the fixation point (McConkie & Rayner, 1975).

The pattern of saccade and fixation is at the focus of interest in reading studies using eye-tracking methodologies. Individual and group differences in reading manifest themselves in different saccade and fixation patterns. For instance, faster readers make shorter fixations, longer saccades, and less regressions, while slower readers make longer fixations, shorter saccades, and more regressions (Just & Carpenter, 1980). In the eye-tracking literature, regression is a backwards saccade and fixation to a point of the text that already received visual attention, and can indicate confusion or a conscious decision to integrate different parts of the text in memory.

3.4. Higher-order thinking

Higher-order thinking here is defined as the way of thinking at the levels of application, analysis, synthesis, and evaluation in Bloom’s Taxonomy (Bloom et al., 1956). Bloom’s Taxonomy continues to be one of the most influential guide in classifying types of higher-order thinking skills since the 1960’s (Quellmalz, 1985; Imrie, 1995). It divides types of thinking into six categories, knowledge, comprehension, application, analysis, synthesis, and evaluation. The first two constitute lower-order thinking, and
specific cognitive tasks include memorization and recall of information (Zohar & Dori, 2003). Higher-order thinking instances include predicting outcome, analyzing case studies, developing hypotheses, and comparing/contrasting information (Crowe, Dirks & Wenderoth, 2008).

II. REVIEW OF THE LITERATURE

1. The Process of Reading

Reading skills function as a “gatekeeper” for all other human learning. As it is, the cognitive, biological and neuroscientific process of reading has received overwhelming interest from researchers throughout the past century up till present day. A vast amount of research has been conducted on each and all subprocesses of reading, from orthographical and phonological processing, word recognition, syntax processing, semantic processing, and pragmatics. These steps constitute two different cognitive processes thought to play a part in reading: bottom-up and top-down. Bottom-up processes start from the individual letter and word level, focusing on phonological and orthographical awareness and word recognition. Top-down processes are those that create a macrostructure of the text and integrate the reader’s background knowledge with knowledge presented in the text. It is beyond the scope of this research to describe all models of reading comprehension. Instead, in this review of the literature, I consult and combine existing reviews on the dozens of extant reading comprehension models to choose three seminal, representative
models that each integrate bottom-up and top-down processes. Each are representative in that they have varying emphasis placed on the two processes. The reviews most central to this section were the ones by Sanders, Schilperoord and Spooren (2001), Britton and Graesser (2014), and McNamara (2012). The three models thus chosen are ones that have served as grounds for new text comprehension models, and have accrued substantial empirical evidence since inception.

Most of the extant reading models including the three described in length below take as its premise the structure of the human memory system as posited by the information processing theory. Reading is seen as a process involving both working and long-term memory, two terms that will be briefly introduced to further the understanding of reading comprehension models.

1.1. Information Processing Theory

Information processing theory (Atkinson & Shiffrin, 1968; Broadbent, 1984) serves as an important tool to understand how humans process all information, including textual information. In this theory, memory is divided into three parts. Sensory memory is the receiving port of all incoming information, and can hold information for very limited periods of time. When one pays attention to a certain piece of information, only then is it transferred to working memory. Working memory, also called short-term memory, holds information for slightly longer, yet still short lengths of time, and has a limited capacity. While much contested and elaborated upon since it was first
suggested, a theory by Miller (1956) suggests that the human working memory can hold five to nine bits or chunks of information at a time, the so-called ‘magic number seven, plus or minus two’.

To remember this information for longer than a few seconds, a person must encode it to long-term memory through techniques such as rehearsing or elaboration. Information is lost quickly if it is not encoded to long-term memory. Long-term memory is in essence limitless, but accessibility of information in long-term memory can vary widely depending on how well the information was encoded, or how frequently it was retrieved. The diagram below summarizes the key concepts of information processing theory.

![Information Processing Theory Diagram](image)

*Figure 2.1. Information Processing Theory*

We now turn to concrete reading comprehension models based on information processing theory.

1.2. Bottom-up Model: The Construction-Integration Model

Kintsch’s (1988, 1998) construction-integration model (CI model) is widely acknowledged to be one of the, if not the most, comprehensive and strongly supported models in reading comprehension (Graesser, 2007; Singer
& Kintsch, 2001). It sets forth a naturalistic explanation on how text representations are created, understood, and integrated with the reader’s prior knowledge. It emphasizes bottom-up processes rather than focusing on the role of a top-down expectation or schema in reading. The model later gave rise other important models such as the memory-based resonance model by Myer, O’Brien and their colleagues (1994; 1998), and the capacity-constrained CI model by Goldman, Varma and Cote (1996).

According to the CI model, reading a text triggers the construction phrase, in which an approximate, incoherent mental model is created. This model is largely based on the text but is also influenced by the reader’s goals and knowledge. The initial connection between propositions\(^1\) in this phrase includes ones that are correct but also those that are redundant, contradictory, weak or incomplete. A key assumption in the model is that such connections are created only between nodes that exist in the working memory. Subsequently, in the integration phrase, the inappropriate local constructions are pruned away in order to construct a text representation that is more coherent and accurate. The final product of the integration phrase is the “situation model,” a mental representation of the text which coheres with both logic and prior knowledge of the reader, and can be remembered and used for a specific purpose.

---

\(^1\) A linguistic term for the basic units of meaning found in a text that is constant, regardless of change in such things as voice and illocutionary force of the clause. For instance, in the sentence “The tall, stately building fell,” three propositions are expressed: “The building is tall.” “The building is stately.” and “The building fell.” (Glossary of Linguistic Terms, SIL International)
In the comprehension of longer text, Kintsch underlines the importance of a macro-structure level that works in tandem with the previously explained micro-structure level processing in the sentence or phrase level. In creating this macro-representation of the overall text, the perceived macro-proposition of the text, or the “gist” of the text dominates the meaning-making of the text, to the extent that “material contradictory to what the reader takes to be the text’s macro-structure is simply ignored” (Kintsch, 1998).

1.2. Top-down model: Constructionist Model

The constructionist model proposed by Graesser, Singer and Trabasso (1994) is considered to be a representative strategy-driven model, emphasizing top-down processes instigated by the reader. Three key assumptions are thought most characteristic of this model. First, the explanation assumption states that readers “attempt to explain why actions, events, and states are mentioned in the text.” (Britton & Graesser, 2014) In other words, they make causal links and conjectures on the intent of the author. Second, the coherence assumption states that readers try and create a coherent meaning structure at both local and global levels. If there is a perceived gap in this meaning structure, the reader attempts to “actively think, generate inferences, and reinterpret the text in an effort to fill in, repair, or take note of the coherence gap.” (McNamara, 2012) Third, the reader goal assumption states that readers attend to text in varying ways depending on the goal that
he/she has. For example, news advertisements and science articles would be read in very different ways even by the same reader.

The constructionist model also itemize inferences that the reader generates. Based on the three assumptions above and other contingencies, the model draws a list of subconscious inference production rules that implement the reader’s own active comprehension strategy. These inferences can be created online (i.e. real time; during reading), or offline (i.e. after reading the text).

It is important to note that unlike the CI model, the constructionist model contends that the purpose and goal of the reader is a primary driving force behind both the top-down and bottom-up processing of text, on a both conscious and subconscious level. The reader is seen to possess an innate inclination to creating the deepest, most accurate mental representation of the text, as allowed by their purpose in reading, the extent of their background knowledge, and the clarity and coherence of the text itself.

1.3. Hybrid Model: The Landscape Model

One of the most recent of all the seminal reading comprehension models, the landscape model set forth by Van den Broek et al (1996, 2005) is often called a ‘hybrid model,’ (McNamara, 2012; Britton & Graesser, 2014) in that it places emphasis more evenly on top-down and bottom-up cognitive processes.
The landscape model emphasizes the cyclical process of reading. During reading, the working memory of a reader is used to hold selected information from the text and from prior knowledge. To be more specific, four sources compete for limited attentional resources of working memory: text currently being processed, the immediately preceding reading cycle, earlier reading cycles, and background knowledge. Various factors weigh into determining which information occupy working memory at a certain point in time. This includes location in text, relevance to understanding, availability, text cues and more. Depending on these factors, individual concepts fluctuate in their strength of activation in the working memory, creating with other concepts a mental ‘landscape’ in working memory, as visualized below.

![Figure 2.2. The Cognitive Landscape during Reading of a Text.](image)

We see in the posited cognitive landscape of the reader that the word ‘knight’ has a steady activation rate after its first introduction at cycle 1. As the subject of the entire paragraph, it is retaining its place in the reader’s working memory throughout the text.

The landscape model is significant in that it operationalizes some of the key assumptions of reading proposed by previous models. The activation level of each individual concept can be calculated using the set of rules delineated by the model. Behavioral studies show that the output of such simulations correlated closely with actual human recall (r=.69) (Van den Broek et al., 2002).

2. Effect of Learner Elements on Reading

The conscious and subconscious input of the reader is a key component of most reading models. This input includes the reader’s background knowledge, reading purpose and other individual characteristics. On the other hand, in educational research, more attention has been paid to how elements that can be controlled by the educational provider have an impact on the strategy of the reader. These elements include interventions like learning activities or tests. We first look at the three representative types of learner elements that affect reading behavior.
2.1. Background Knowledge

The background knowledge that a reader has in the content of the text is a critical factor in determining the reader’s idiosyncratic understanding of the text. When a reader faces a situational demand on their background knowledge during reading, they will automatically review and choose the most relevant information they possess. This cumulative action will eventually create a schema or a framework on which readers build their own understanding of the text. A good schema allows the newly learned facts to be retained better, as it successfully links new facts to preexisting nodes. In other words, readers in possession of rich background knowledge only have to fill in a few blanks in his/her schema; readers that are not very knowledgeable of the subject at hand must fill in many more blanks that exist in his/her knowledge base, which can be difficult to both carry out and maintain.

This intuitive mechanism have been validated and elaborated by many empirical studies. For instance, prior knowledge was found to have an effect on performance on questions asking explicitly stated facts but even more so on questions assessing understanding that had to be inferred from the text (Pearson, Hansen & Gordon, 1979), indicating that background knowledge influences depth of understanding. Recall was also found to be influenced by background knowledge (Chi & Ceci, 1987; Kintsch & Franzke, 1995), as was the reported level of interest in the text (Alexander, Kulikowich & Schulze, 1994). Erroneous or inadequate background knowledge resulted in poor memory and failure to make correct inferences (Kendeou & Van den Broek,
Research has also shown that a dearth of prior knowledge can be remedied by a pre-reading intervention. Prepping students in relevant information before reading a text was found to improve performance for both comprehension and recall (Langer, 1984; Guthrie, McRae & Klauda, 2007).

While possessing relevant information is indeed important to understanding, being able to access and use this information another important pillar in successful reading. Various research has demonstrated the existence of cases in which the reader has the relevant knowledge but does not use it (Cain, Oakhill, Barnes & Bryant, 2001; Oakhill & Cain, 2007). Based on such findings, Elbro and Buch-Iversen (2013) showed that an intervention training students to retrieve relevant background knowledge and make inferences using it had a highly significant effect on reading comprehension. Even knowing simply what the context of the text is was found to be beneficial in comprehension (Bransford & Johnson, 1972; Bransford & McCarrell, 1974). This further demonstrates that background knowledge must be both possessed and accessed to be truly useful in reading.

2.2. Individual Differences in Working Memory and Linguistic Ability

In the context of most reading comprehension models, working memory is the cognitive workspace in which information from different parts of text and the reader’s prior knowledge come together under the rules of coherence and reading purpose to create a mental representation of the text. Daneman and Carpenter (1980) most influentially defined the relationship
between working memory and reading comprehension by theorizing that working memory was used to process the text with different strategies and skills, and the remainder of its capacity used to store the intermediate products of reading, such as pronoun referents\(^2\) and information. That is, working memory serves to both process and store.

It is well known fact that working memory capacity varies among individuals (Engle, Kane & Tuholski, 1999). Just and Carpenter (1992) thus theorized that individuals with working memory larger capacities would have the ability to make multiple inferences and meaning-making while storing more syntactic and pragmatic information.

This idea was explored from early on, and working memory and reading has indeed been found to have a steady correlation of about 0.5 to 0.6 in various reading research using standardized tests (Baddeley, 1992). Individuals with high working memory span was found to be especially efficient in processing sentences that required complex processing, such as in garden path sentences\(^3\), which supports the idea that working memory supports information storing and processing during reading.

---

\(^2\) A referent is the concrete object or concept designated by a word or expression. Different words can be used to describe the same referent. For example, George Washington can be referred to as him, he, the president, Mr. Washington, my friend, and so forth. (Glossary of Linguistic Terms, SIL International)

\(^3\) Garden path sentences are sentences that have temporarily difficult to understand because they contain an ambiguity, a word or a phrase that can be interpreted in different ways depending on what section of the text the reader is reading. For example, “While Bill hunted the deer ran into the woods” (Ferreira, Christianson & Hollingworth, 2001) is a garden-path sentence that first lets the reader assume Bill is the subject, then creates confusion as to who the subject of the verb ‘ran’ is as the sentence progresses.
Other research suggests that linguistic and other verbal ability constitutes a separate proponent of reading comprehension from working memory. Vocabulary skill has received high levels of attention in this line of inquiry, with one group of researchers arguing vocabulary skills influence reading directly (Daneman & Green, 1986), and another group proclaiming the existence of a common variable that affects skill acquisition in both reading and vocabulary (Cain, Oakhill & Lemmon, 2004). Cain, Oakhill and Bryant (2004) found, controlling for the effect of working memory and vocabulary/verbal ability, that component skills of comprehension including inference making, comprehension monitoring and story structure monitoring predicted unique variance in reading comprehension.

In an alternative explanation as to how working memory is connected to reading proficiency, some studies find that only working memory measures that require linguistic processing (word, sentence and number-based) can predict reading comprehension, as opposed to those that require manipulation of shapes and patterns (Nation, Adams, Bowyer-Crane, & Snowling, 1999; Seigneuric et al., 2000), suggesting that working memory measures may have at least part of its predictive value through linguistic demands of the task. Waters (1996) notes that most common measures of working memory in fact have predictive value through the “overlap of operations” than by measuring only working memory itself.

2.3. Reading Purpose
Mature readers switch their way of reading effortlessly in everyday life. They are able to read critically when reading an advertising copy, selectively when reading an instruction manual, and integratively when reading a historical text. This effect of a reader’s reading purpose on cognition and behavior during reading has been well documented in empirical research. College students showed differences in reading behavior while reading the same text, depending on whether they thought the text to be a news story or fiction (Zwann, 1994). Newspaper readers skimmed through sports articles more than for news, features and debate articles (Holsanova, Rahn & Holmqvist, 2006). People fixated on text for longer lengths of time when asked to accurately solve a question related to the text, rather than intuitively (Horstmann, Ahlgrimm, & Glöckner, 2009). There are recent findings that even suggest that we can predict the genre of the document being read based on electroencephalogram (EEG) readings of brain activity (Kunze et al., 2013).

In the learning sciences, reading research have paid special attention to how learners read when they are reading to study. Lorch, Lorch and Klusewitz (1993) collected the self-reported reading strategies of college students for different, specific situations (i.e. reading a pamphlet for health information, reading a suspenseful novel, reading a craft magazine for specific direction, etc.) and showed through cluster analysis that readers systematically varied their strategies according to the situation. Reading for study was reported to result in slower reading, more memorization and monitoring of understanding, while in reading for entertainment, students showed either faster or skimming
patterns of reading, and overall paid less careful attention to the text. Van den Broek et al. (2011) also found a difference in readers’ behavior depending on the motivation of reading. Students reading for study produced more coherence-building inferences, while students reading for entertainment were more likely to link the reading to their own experiences.

Reading for study and reading in anticipation of tests can be expected to be similar in nature, but to also each accompany a set of independent motivations, strategies and reading patterns. This is because as in Lorch, Lorch and Klusewitz (1993), studying can be motivated by a variety of reasons other than preparing for tests, which include research, class preparation, to self-educate, or to apply what is learned to a task. On the other hand, reader intent in anticipation of a test will include elements other than the intent to study. For instance, experiencing a certain type of assessment question may encourage the reader to prioritize information that would be useful in answering similar questions; if the assessment was stressful or difficult, it may cause the reader to pay even more attention to the text; a test based on rote memory may cause the reader to focus more on keywords. In extreme cases, a completely unrelated set of questions or questions that can be solved without knowledge from the text may result in a cursory reading behavior.

3. Priming Effect of Learning Interventions on Reading

3.1. Priming
Different learning interventions can have a different effect on subsequent learning without explicit guidance or even intent. This can be explained by the priming effect, a phenomenon and technique where a previous stimulus influences the response to a subsequent stimulus (Bargh & Chartrand, 2000), such as when the list ‘red, green’ invokes the word ‘yellow.’ In the field of education, there has been much research on how different stimulus in the learning material can prime the way students process the material. While there are various types of such stimuli found to be influential, such as illustrations (Choi, 2016), and textual cues (Sohn, 2004), the stimuli most relevant to the current study are learning activities and tests, educational interventions that prompt an output response from the reader. Special attention is given to tests, while other learning activities are discussed together.

3.2. Learning Activity

Different post-reading learning activities, such as drawing diagrams, creating summaries and peer teaching has been observed to have different effects on readers. Here, reading interventions of a clinical or prescriptive nature, such as those administered to dyslexic children or low-level readers, are left out.

Lorch, Lorch and Mogan (1987) finds that tasks that necessitate the construction of topic structure (outlining) had a greater slowing effect than
did a task that did not (verification) on the speed of reading when there was a major shift in topic, for good recallers. Ponce and Mayer (2014) shows that prompting learns to carry out note-taking or graphic organizers result in different cognitive processing of expository text. The note taking group read in a linear fashion, while the graphic organizer group read generatively, switching attention from one part of the text to another nonlinearly in an effort to build an appropriately ordered mental model of the text. Amadieu and colleagues (2015) discovered that reading a text before creating a concept map, as opposed to creating a concept map of prior knowledge before reading the text, fostered better learning and reduced cognitive load.

Overall, prior research in the effect of learning activity show that learners are sensitive to priming effects introduced by different types of learning activity, and that their reaction can be gauged by eye-trackers. The key features given attention in the eye-tracking data seemed to be reading time and linearity of reading.

3.3. Tests

Not surprisingly, many studies find that testing improves retention of learned material (Roediger & Karpicke, 2006; Carpenter & DeLosh, 2006). This line of research has strong ties to what is called the testing effect, a widely researched phenomenon in cognitive psychology. Of special interest to this research are studies that find that the testing effect facilitates the learning of new, subsequent material (Pastötter & Bäuml, 2014; Wissman,
Rawson & Pyc, 2011), in a phenomenon called the ‘interim-test effect’ or the ‘forward effect of testing.’ These studies have shown that test-takers for one learning material do better on the test for the subsequent learning material compared to students who restudy.

The effects of tests on subsequent learning have largely been investigated through the use of standardized tests. Repeated observations have allowed the interim testing effect to become a robust, generalizable and widely accepted phenomenon, valid across different types of material, including authentic text (Wissman et al., 2011) and narratives (Chan et al., 2011). Yet relatively less attention has been paid to the mechanisms of the effect. Discovering what subelements contribute to the effectiveness of testing is important because we have yet to discover what combination of features in a test is most effective. If tests are too frequent, too difficult, bears overly high stakes, or prompt a limited way of learning (i.e. only rote learning), the side effects of testing such as low motivation (Torrance & Coultas, 2004; Harlen & Deakin Crick, 2003) a tendency to guess (Roediger & Marsh, 2005), and uncritical acceptance of what is being taught (Lobascher, 2011) may occur. Thus, it is important to understand in what context, and through what mechanisms tests bring about learning benefits, and fine-tune instructional design to maximize benefits while minimizing side effects.

In Wissman, Rawson and Pyc (2011), the authors sketch a general structure of mechanisms for the interim-test effect, namely enhanced access to topically related information in the first learning material, the use of more
appropriate encoding strategies, and different study behavior. Very recently, several empirical studies have begun an inquiry on this line of research; such studies find that there are several metacognitive factors in play. Testing facilitates a better regulation of studying time (Soderstrom & Bjork, 2014; Yang, Potts & Shank, 2017), enhances induction, or the linking of newly learned facts to a previous experience (Yang & Shanks, 2018), or better self-monitoring and accurate assessment of self-knowledge (Fernandez & Jamet, 2016). This metacognitive nature of the mechanisms involved is further supported by studies that find that the interim-test effect is not found in younger elementary school students (Aslan & Bäuml, 2016).

Existing research have used methods such as think aloud protocols (Fernandez & Jamet, 2016), inductive testing (Yang & Shanks, 2018), or analyzing notes produced during learning (Jing, Szpunar & Schacter). While such methods are meaningful in that they allow a more holistic, intuitive understanding of the cognitive process, it can gain much by being supported by quantifiable, scientific research methodologies. One such method is eye-tracking, where the movement of the eye is recorded with a millisecond-level temporal sensitivity, and thus allow us a precise, objective view of the cognitive process of reading in anticipation of tests. We now look at the method and how it has proved useful for investigating research questions in the field of education and in reading.

4. Eye-tracking Methodology in Educational Research
In educational research, the effectiveness of intervention has most commonly been observed through achievement in tests (Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013; Mason, Tornatora & Pluchino, 2013). While simple and certainly not without its merits, looking only at the result of learning is not an optimal method when the aim is to observe the process of learning – in this case, reading strategy. Test results are also inevitably influenced by factors other than the intervention, such as prior knowledge, test-taking strategies, interest, experience and verbal skills (Anderson, 1991; Gustin & Corazza, 1994; Bereiter & Bird, 1985; McNamara & Kintsch, 1996). Therefore, while effective reading strategies do often coincide with superior test scores (Hyönä, Lorch & Kaakinen, 2002; Dinsmore & Alexander, 2016), online measures are needed to accurately observe how interventions have an influence on learning processes.

Concurrent think-aloud measures have often been utilized to remedy this shortcoming. As readers report a variety of strategies real-time, this method provides a richer understanding of the cognitive processes utilized during reading (Ericsson & Simon, 1993; Long & Bourg, 1996; Pressley & Afflerbach, 1995). Yet the method is not without flaws. Some participants find it difficult to vocalize their thoughts; silence tends to occur in pivotal reading moments (i.e. when encountering parsing difficulties) (Cooke & Cuddihy, 2005). Also, explaining mental processes are likely to take longer than the actual process, causing time lags or omitted observations (Nielson, Clemmensen, Yssing, 2002).
Since the 2000’s, eye-tracking methodologies have received increasing amounts of attention, researchers looking to substitute or complement comprehension tests and think-aloud procedures to objectively observe cognitive processes during reading. In an eye-tracking experiment, the eye position and movement of a participant is recorded then analyzed. It allows a quantitative analysis of reading patterns, and a way to record the process of reading and viewing that sometimes even the participant him/herself may not be aware of.

The use of eye-tracking in this line of research is based on the ‘eye-mind hypothesis’ set forth by Just & Carpenter (1980). This hypothesis posits that there is a strong connection, without any “appreciable lag” (Just & Carpenter, 1980), between human gaze and attention. Since its conception, the hypothesis has accrued strong empirical evidence (Hegarty & Just, 1993; Yang & McConkie, 2001; Van Gog, Kester, Nievelstein, Giesvers & Paas, 2009). It is important, however, to acknowledge the limits of the methodology. A common analogy used to describe the relation of attention and gaze is that of a rubber band (Liversedge & Findlay, 2000). That is, while reading, a covert attentional shift takes place immediately prior to execution of overt eye movement, ‘pulling’ the eye along.

Eye-tracking is still a behavioral method and thus ultimately an indirect measure of a participant’s cognitive processes. For example, a reader could have multiple reasons when they rest their attention on a specific phrase longer. The phrase may simply more difficult to parse, it may have triggered
a personal memory the reader has, or it could be an indication of a conscious decision that the specific sentence is of importance.

While early eye-tracking research on reading has mostly belonged to the field of linguistics, focusing on the lower-level processing of text (i.e. word recognition, sentence parsing), there has been a sharp increase of research from the domain of education looking at the effect of specific learning interventions ranging from design elements such as illustrations (Choi, 2016) to learning activities such as taking notes (Ponce & Mayer, 2014). In a different line of inquiry, others have paved the way for research on global text processing strategies that readers employ to understand authentic expository text (Hyönä, 2010). Successfully using eye-tracking methods on such endeavors can be challenging, as eye-tracking experiments generate large sets of data that cannot intuitively be understood as a pattern of reading. As it is, research of this kind needs a clear set of hypotheses that specify how the predicted patterns are expected to be operationalized in terms of eye-tracking measurements.

III. HYPOTHESES

Based on the well documented effect of reading purpose or learning activity on reading strategy (Hyönä, 2010; Lorch, Lorch and Klusewitz, 1993; Van den Broek et al., 2011), it was believed highly likely that assessment would affect learner purpose and by consequence reading strategy. The current study hypothesized the following:
1. Tests will increase reading time and fixation duration for subsequent text, especially for key parts of text. That is, parts categorized as ‘relation’, ‘definition’, and ‘keyword.’

2. Higher-order thinking tests will elicit a regressive way of reading than before the test. Regressive reading is thought to occur when readers feel the need to process a past part of the text further, or when they want to establish connections between different key parts of text (Rayner et al., 2006). Readers in this group will also focus longer on parts of the text that establish relationships between key ideas, and higher second-pass fixation measures after the test.

3. Memory-based tests will elicit linear reading strategies, where readers focus on specific key words that could be prompted on the test. That is, they will have higher first-pass fixation measures, and lower second-pass fixation measures when compared to reading before the test.

In the discussion of results, differences in the ‘time to nth fixation (n=1, 2, 3, 4)’ measures are disregarded. This is because the distance from the beginning of the text to a specific type of ROI was not controlled between different texts. For example, a definition could appear at any point of the text. In this case, differences on the time it took to fixate on a definition between the two texts could have occurred because definitions tended to come up later in one text compared to the other.

IV. METHODS
1. Participants

Eighteen undergraduate students (10 female, 8 male) from Seoul National University (Seoul, Korea) participated in the study; four additional participants were recruited but excluded from data analysis as the eye-tracking machine failed to correctly register their gazes. All participants were recruited from the school website and screened to ensure that they were not from the Political Science department, as the subject matter of the text dealt with political concepts. All participants underwent a background knowledge test to ensure each experimental group did not diverge too widely in terms of expertise.

One general characteristic worth noting for the participant pool is that they are likely to be skilled readers. Admissions for this particular university is the considered to be the most competitive in Korea, and the student body is mostly consisted of applicants who score at the highest tier (4%) in the national college entrance exam for subjects including language and literature.

Despite this feature of the participant pool, it was expected that comparisons within similarly highly skilled readers would reveal any tendencies and effects that may occur, provided the text and test were sufficiently challenging. Additionally, it is often documented that skilled readers are more aware of the latent purpose of reading and respond more readily to guiding influences (Paris, Lipson & Wixson, 1983; Smith, 1967; Mokhtari & Reichard, 2002). Thus, it was thought that the current population
would have more noticeable responses to the test treatment that may not be as visible in less skilled populations, and so aid the documentation of the cognitive process of reading in anticipation of exams. However, a follow up study determining whether a comparable process occurs for less skilled readers will be necessary for a more comprehensive picture.

2. Experimental Process

Eighteen participants were randomly assigned to three groups depending on the type of test administered: none, rote memory and higher order thinking. The contents of the higher-order thinking and rote memory tests are explained in detail in section 4.1.

Upon arrival, all participants signed a consent form and was informed of the purpose of the experiment. The true purpose of the research was not revealed at this time, and participants were instead told that they were participating in a text design research investigating the effect of a design element on readability. Participants were requested to read in a neutral, natural manner as they would normally do in everyday situations. The eye-tracking, test and interview data was all collected in a one-to-one session with the researcher.

After calibrating gaze pattern to the eye-tracking system, the participants were first presented with a practice text. They had no knowledge at this point that they would be taking a test later on, and no test followed the
practice text. After the practice text, participants read a short three page text on politics, followed by either a short answer test requiring fact retrieval, a test with essay questions requiring higher-order thinking (i.e. application, analysis, synthesis or evaluation), or no test. They were then presented with the second text. After reading, all participants took a test on the second text. This second test consisted of an even mix of rote memory and higher-order thinking questions. All text was digitally presented on a computer screen, and all tests were given in pen and paper format. The background knowledge test was given later on as to not induce test expectations before reading the first text. To elaborate, the participants of the pilot study, who took the background knowledge test before the other experimental procedures, noted that they fully expected a test to follow while they were reading text 1. Also, the test was not directly related to the contents of the two texts in the experiment, but rather assessed general knowledge levels in the field of politics. Thus, the effect of reading the text on the background knowledge test was deemed minimal, and the test was moved backwards.

Each step of the experiment is summarized below with the approximate time spent on each step specified.

Table 4.1.

Steps of Experiment and Time Spent on Each

<table>
<thead>
<tr>
<th>Step</th>
<th>Process</th>
<th>Time spent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Consent and introduction of experimental process and purpose (with deception technique)</td>
<td>5 mins</td>
</tr>
</tbody>
</table>
3. Experimental Tools

3.1. Text Stimuli

The first text given to participants was a practice text, not included in later analysis. This text dealt with topic different from that of the actual stimuli. It was adapted from a government brochure on different types of copyright infringement cases. It was of similar length as the actual stimulus.

The inclusion of this practice text serves two purposes. One, it familiarizes the participants to the experimental condition. While the eye-tracker used in this study did not have a chin-rest or a head-mounted device,
and so allowed the participant to read in a relatively natural setting, prior experimental studies note that it still can be awkward for users to keep his/head completely still (Zhu & Ji, 2004; Azuma & Bishop, 1994), as many eye tracking systems, including the one used in the present research, require. Indeed, the first attempt at using the eye-tracking system often required the researcher to remind the participant to keep his/her head still several times, which would have interfered with the flow of reading. This problem was noted in the pilot study conducted on two students, and the practice text was incorporated into the actual experiment.

Two, it also aims to alleviate the effect of an anticipated confounding variable – the preexisting test expectancy participants may have had when reading the first actual stimuli. It is well known that in an experimental situation participants are more highly motivated, a phenomenon also known as the Hawthorne effect. Cook (1962) writes of the relevance of the Hawthorne effect in educational research, asserting that learners are more motivated due to the novelty of the intervention, or the fact that they are being experimented on. As the participants of the present study were proficient readers in an experiment they knew was in the domain of education, it is highly likely that they would have had high motivation as well as expectation of a test. It was thus deemed necessary to have a practice text where no test would follow. With it, it was hoped that the participants would come to believe that the purpose of reading was truly to evaluate the design of the text, as they were informed in a pre-briefing session, and read less with the purpose of preparing for a test. It can be said that this measure did achieve what it
intended to do to a certain extent, as many participants noted in the post-experiment interview that they did not know they were going to have a test specifically because of the first practice text.

The actual text stimuli were adapted from a political science text used first by Rha (1988) in English, then translated into Korean and validated by Sohn (2004). Two independent parts of text, dealing with different political concepts, were taken from the original text in Sohn (2004) and edited to ensure the length and layout would be similar. The first text dealt with the concepts of political influence and political progress. The second text was on the concepts of political property and the political decision making process.

The first text was 36 lines or 1784 letters excluding blank space divided into four pages, each with 10, 11, 6, 9 lines. The second text was 36 lines or 1810 letters excluding blank space divided into four pages, each with 11, 10, 8, 7 lines. Two experts with extensive research experience in the general domain of education and two classroom teachers of social science reviewed the text to ensure that the two texts were similar in difficulty, as well as design and length. Sample pages of the text are shown below. The text have wide line spacing to enhance eye-tracking accuracy.
3.2. Tests

A total of four tests were used in this experiment. The first test measured background knowledge and was adapted from Sohn (2004), as it was used to test background knowledge before reading an identical text stimuli. The test was initially created by Sohn with the help of content-matter experts, and its difficulty set to be comparable with the Korean national curriculum for social sciences in a high school level.

Two more tests were used after the reading of the first text. The items used in the memory test were adapted from Sohn (2004). The test consisted entirely of short-answer questions, designed to assess whether learners had retained key terms. The prompted keywords were distributed as evenly as possible throughout the text; one question was asked for every five to ten lines of the text. There were five questions in all.

Higher-order thinking questions were created by the author then revised with the help of four experts, two of which had extensive research
experience in the general domain of education and the other two classroom teachers of social science. Questions were designed to follow the classic definition of higher order thinking by Bloom and his colleagues (1956) as summarized in the “definition” section of this research. The resulting questions dealt with either 1) interpreting or analyzing a new situation through the ideas and principles from the text, or 2) making a judgment on a new situation based on the ideas and principles from the text. The higher-order thinking test had five questions as well.

The last test, to be administered to all participants, was used after the second text. It consisted of five rote memory and five higher-order thinking questions.

3.3. Interview

After the second test, all participants answered questions on self-perceived strategy and reading behavior in an interview. They were asked to retrospectively explain their reading strategies for text one and two, whether there were any perceived differences between the two, their level of understanding for each text, interest and recall for each text, and to also describe the way they usually read when preparing for tests. This was a semi-structured interview (Merriam, 1998), meaning that while there were a set of questions prepared in advance, the interviewees were encouraged to add any relevant information to their direct answer to the question. The interview questionnaire was revised once with the help of experts in the field of
education, and twice with the input from a pilot interview conducted on two students.

4. Data Acquisition and Analysis

Three types of data was collected in this study: The eye-tracking data from the two texts, the test results, and the interview results. Data processing was carried out with excel and NYAN 2.0. software. All statistical analysis was performed with R (version 3.3.3).

A Tobii X60 eye-tracker was utilized to acquire the eye-tracking data. The raw eye-tracking data was in the form of time-stamped fixations (*.cvs). Before statistical analysis, regions of interest (ROIs) were designated using NYAN 2.0 software. Each text was divided into different components, namely ‘all text’, ‘relation’, ‘definition’, ‘keyword’ and ‘non-key.’ All text was a ROI entirely covering the part of the screen with text. Keywords were words or short phrases newly introduced in the text, and served as potential answers for the memory test. Relations dealt with a logical connection with two or more keywords. Definitions were sentences that were crucial to understanding the meaning of a keyword. Non-key parts of text were all remaining text not covered by another ROI. ROIs were designed to be very close in total length between text 1 and 2.

Then, using the same software, different parameters including total fixation time, total fixation count, and first to fourth pass fixation time and
count were extracted for each ROI. Total fixation time, or gaze duration refers
to all fixations that occurred inside a ROI. Total fixation count is the number
of fixations made. Total reading time and fixation count for a ROI is used as
one measure of the amount of cognitive effort that the reader paid to the
region.

Measures for first, second, third and fourth-pass fixations are available
for extraction in the NYAN software. The number and length of fixation tend
to decrease from the first to fourth pass. First-pass fixation time is the amount
of time it took to make the first fixation on an ROI. First-pass fixation duration
is the time spent on an ROI during its first pass, the time between the moment
a gaze first enters an ROI and the moment it leaves. First-pass fixation count
is the number of fixations during the first-pass fixation duration.

Once a gaze leaves a specific ROI, returning to the same ROI
constitutes a *regression*. There can be multiple regressions made to the same
ROI, if the reader returns to it multiple times. Regressions are seen to be
indications of a reader putting in more processing effort, because he/she finds
it difficult to comprehend or because he/she wants to establish connections
between the information in the part of the text being returned to and other
information, which could be from other parts of text or from prior knowledge.

Second-pass fixation time refers to the time it took to regress to an ROI
for the first time, or the time it took until the second pass. Second-pass
fixation duration is the total amount of time a reader spent during the second
pass, when the gaze returns to the ROI for the second time before it leaves
again. Second-pass fixation count is the number of fixations made during second pass. The same holds for third- and fourth-pass fixation time and count.

In the schematized diagram below, the first blue box is a ‘definition’ ROI. Fixations 1, 2, 14, 3, 15, 4 and 5 constitutes total fixation measures. Total fixation duration would be the time spent during fixations 1, 2, 14, 3, 15 and 4. Total fixation count is 6. Time spent on fixations 1, 2, 3, 4 and 5 is first pass fixation duration. First pass fixation count is 5. The eye movement from fixation 13 to 14 is a regression. Fixations 14 and 15 are second pass fixations. Second pass fixation duration would be the time spent on fixations 14 and 15. Second pass fixation count is 2.

\[ \text{Figure 4.2. Schematized Eye-tracking Results} \]

The eye-tracking data was collapsed across ROIs of the same category found in the four pages of a single text. As individual differences in reading strategies are known to be pronounced and consistent (Hyönä, Lorch & Kaakinen, 2002), and the participant sample size in the current research was relatively limited \((n = 18)\), a nonparametric alternative to paired t-tests, namely a Wilcoxon-signed rank test, was conducted on the pooled data. In this pooled data, each parameter signifies the following.
Table 4.2.

*Parameters of Eye-tracking Data, in Data Pooled across ROI*

<table>
<thead>
<tr>
<th>Parameters of Eye-tracking Data</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total fixation time</td>
<td>The total time a reader spent fixating on a specific type of ROI, added across all ROI of the same type found in all four pages of one text.</td>
</tr>
<tr>
<td>Total fixation count</td>
<td>The total number of times a reader made a fixation on a specific type of ROI, added across all ROI of the same type found in all four pages of one text.</td>
</tr>
<tr>
<td>First pass fixation duration</td>
<td>The total duration of time a reader spent fixating on a specific ROI during the <em>first</em> pass, added across all ROI of the same type found in all four pages of one text.</td>
</tr>
<tr>
<td>First pass fixation count</td>
<td>The total number of times a reader made a fixation during the <em>first</em> pass of a certain ROI, added across all ROI of the same type found in all four pages of one text.</td>
</tr>
<tr>
<td>Regressive fixation duration</td>
<td>The total duration of time a reader spent fixating on a specific ROI during the <em>second to fourth</em> passes, added across all ROI of the same type found in all four pages of one text.</td>
</tr>
<tr>
<td>Regressive fixation count</td>
<td>The total number of times a reader made a fixation during the <em>second to fourth</em> passes of a certain ROI, added across all ROI of the same type found in all four pages of one text.</td>
</tr>
</tbody>
</table>

Next, the following investigations were conducted on pen-and-paper test results. Two graders independently scored the tests taken by all participants using a predetermined scoring rubric. Any difference in scoring was discussed until an agreement was reached.

The Shapiro-Wilk normality test was performed on the resulting test score data, and the data was found to overall not follow a normal distribution. Accordingly, the Kruskal-Wallis test was carried out on the data to determine if the outcomes of the second test was statistically different among the three groups. The Kruskal-Wallis test is a non-parametric alternative to ANOVA that is used to assess if the mean of more than two groups are statistically
different. The test was performed to see, specifically, whether there was a difference in higher-order question scores, memory-based question scores and total score for test 2 between the three groups.

Additionally, a Wilcoxon-Mann-Whitney rank sum test was performed to determine whether there was a difference between test 2 scores for both test groups (memory and higher order) and the non-test group. The Wilcoxon-Mann-Whitney rank sum test is a nonparametric alternative to an unpaired two samples t-test.

Lastly, key phrases indicating student attitude (“I didn’t have too much motivation” “I was honestly a little bored”) or strategy (“I first skimmed through the text” “I was looking for key words”) was extracted from the interview recordings to qualitatively support conjectures made based on the eye-tracking and test results.

V. RESULTS

1. Eye-tracking Results

1.1. Memory-based test group

When comparing how reading time is spent during the first and second texts, overall results appear consistent with predictions, with some notable exceptions. It can be said that readers in the memory-based test group (‘memory group’) overall took longer to read the entire text, and made more fixations, which adheres to the first part of the hypothesis made for the
memory group. They especially spent a longer time when they reading the definitions for the first time, and came back to read the keywords more frequently and also took longer to review. Interestingly, the memory group also read the parts of text explaining relations for a longer time, something that the original hypothesis did not anticipate.

Table 5.1.
Time spent reading, memory group

<table>
<thead>
<tr>
<th></th>
<th>Text 1</th>
<th>Text 2 (post-test)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>1st pass</td>
</tr>
<tr>
<td></td>
<td></td>
<td>fixation</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>658.87</td>
<td>784.53</td>
</tr>
<tr>
<td></td>
<td>(125.66)</td>
<td>(+125.66)</td>
</tr>
<tr>
<td><strong>Definition</strong></td>
<td>152.93</td>
<td>240.91</td>
</tr>
<tr>
<td></td>
<td><strong>35.73</strong></td>
<td>(+87.98)</td>
</tr>
<tr>
<td></td>
<td>66.33</td>
<td>115.62</td>
</tr>
<tr>
<td><strong>Keyword</strong></td>
<td>77.36</td>
<td>119.37</td>
</tr>
<tr>
<td></td>
<td>27.98</td>
<td>(+42.01)</td>
</tr>
<tr>
<td></td>
<td><strong>29.10</strong></td>
<td>52.84</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Relations</strong></td>
<td>130.07</td>
<td>280.92*</td>
</tr>
<tr>
<td></td>
<td>55.67</td>
<td>(+150.85)</td>
</tr>
<tr>
<td></td>
<td>52.91</td>
<td>52.91</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note. A Wilcoxon-signed rank test was performed on the before- and after-test metrics for all eye-tracking data. For more details, please refer to “Data Acquisition and Analysis.” Also, likely due to the characteristics of a rank test, p-values under 0.05 were all 0.03125.

* p < .05.

Table 5.2.
Number of fixations, memory group

<table>
<thead>
<tr>
<th></th>
<th>Text 1</th>
<th>Text 2 (post-test)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>1st pass</td>
</tr>
<tr>
<td></td>
<td></td>
<td>fixation</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>3721</td>
<td>4415</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(+694)</td>
</tr>
<tr>
<td><strong>Definition</strong></td>
<td>779</td>
<td>1235</td>
</tr>
<tr>
<td></td>
<td><strong>177</strong></td>
<td>(+456)</td>
</tr>
<tr>
<td></td>
<td>328</td>
<td>365*</td>
</tr>
<tr>
<td><strong>Keyword</strong></td>
<td>389</td>
<td>612</td>
</tr>
<tr>
<td></td>
<td><strong>134</strong></td>
<td>(+223)</td>
</tr>
<tr>
<td></td>
<td><strong>150</strong></td>
<td>147*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(+13)</td>
</tr>
<tr>
<td><strong>Relations</strong></td>
<td>789</td>
<td>203</td>
</tr>
<tr>
<td></td>
<td>303</td>
<td>1560*</td>
</tr>
<tr>
<td></td>
<td>331</td>
<td>203</td>
</tr>
</tbody>
</table>

* p < .05.
1.2. Transfer-based test group

The transfer-based test group (transfer group) did not show as markedly and consistently significant differences as compared to the memory group. They did, however, take longer to read the parts of text explaining the relation between ideas. They showed several behavior features inconsistent with the original hypothesis. For one, while this was determined statistically insignificant, there is a not-too-small decrease in total reading time for the transfer group. Also, they looked at keywords less.

Table 5.3.
Time spent reading, transfer group

| Time spent reading (sec) | Text 1 |  | Text 2 (post-test) |  |
|-------------------------|--------|------------------|-------------------|
|                         | Total  | 1st pass fixation | Regressive fixation | Total | 1st pass fixation | Regressive fixation |
| All text                | 753.72 | 603.78 (-149.94) |  | 95.36 (+39.98) | 59.58 (-18.23) |
| Definition              | 165.34 | 155.40 (-9.94) | 77.81 | 75.36 (+39.98) | 59.58 (-18.23) |
| Keyword                 | 99.60  | 81.35 (-18.25) | 40.30 | 28.57 (-3.87) | 39.69 (-0.60) |
| Relations               | 150.27 | 247.22 (+96.96) | 58.46 | 113.44 |  |

* p < .05.

Table 5.4.
Number of fixations, transfer group

| Fixation count (times) | Text 1 |  | Text 2 (post-test) |  |
|------------------------|--------|------------------|-------------------|
|                         | Total  | 1st pass fixation | Regressive fixation | Total | 1st pass fixation | Regressive fixation |
| All text               | 3890   | 3077 (-813) |  |  |  |
| Definition             | 806    | 767 (-39) | 371 | 360 (+188) | 301 (-70) |
| Keyword                | 456    | 142 (-5) | 183 | 194 (+11) |  |
| Relations              | 789    | 1250 (+452) | 368 | 289 (+19) | 577 (+209) |

* p < .05.
1.3. No test group

For the no test group, the only statistically significant difference was in reading the entire text. The group took less time to read, and made less fixations. Interestingly, a large chunk of the reduced time and fixations made in the entire text are not even halfway accounted by the key parts of text – which suggests that these skilled readers may be engaging in a type of skimming reading behavior, where they skip most the parts of text deemed unimportant.

Table 5.5. Time spent reading, no test group

<table>
<thead>
<tr>
<th>Time spent reading (sec)</th>
<th>Text 1</th>
<th></th>
<th></th>
<th>Text 2 (post-test)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>1st pass fixation</td>
<td>Regressive fixation</td>
<td>Total</td>
<td>1st pass fixation</td>
<td>Regressive fixation</td>
</tr>
<tr>
<td>All text</td>
<td>796.02</td>
<td></td>
<td></td>
<td>615.26*</td>
<td>(-180.76)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Definition</td>
<td>195.06</td>
<td>52.60</td>
<td>101.86</td>
<td>154.59</td>
<td>(+18.09)</td>
<td>63.78</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Keyword</td>
<td>92.79</td>
<td>27.84</td>
<td>34.67</td>
<td>82.89</td>
<td>(+0.98)</td>
<td>36.34</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relations</td>
<td>231.39</td>
<td>66.50</td>
<td>101.35</td>
<td>199.19</td>
<td>(+16.30)</td>
<td>89.63</td>
</tr>
</tbody>
</table>

* $p < .05$. 
### Table 5.6.
Number of fixations, no test group

<table>
<thead>
<tr>
<th>Fixation count (times)</th>
<th>Text 1</th>
<th></th>
<th></th>
<th>Text 2 (post-test)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>1st pass fixation</td>
<td>Regressive fixation</td>
<td>Total</td>
<td>1st pass fixation</td>
<td>Regressive fixation</td>
</tr>
<tr>
<td>All text</td>
<td>3982</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Definition</td>
<td>928</td>
<td>250</td>
<td>465</td>
<td>796</td>
<td>354</td>
<td>340</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-132)</td>
<td></td>
<td></td>
<td>(+104)</td>
<td>(-125)</td>
</tr>
<tr>
<td>Keyword</td>
<td>429</td>
<td>124</td>
<td>169</td>
<td>426</td>
<td>147</td>
<td>187</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-3)</td>
<td></td>
<td></td>
<td>(+23)</td>
<td>(+18)</td>
</tr>
<tr>
<td>Relations</td>
<td>1157</td>
<td>331</td>
<td>532</td>
<td>996</td>
<td>256</td>
<td>457</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-161)</td>
<td></td>
<td></td>
<td>(-75)</td>
<td>(-75)</td>
</tr>
</tbody>
</table>

* *p < .05.

### 2. Test Results

While the two test groups outperformed the no test group by a slight margin, this difference was not statistically significant. Looking at the components of test two, the memory group compared to the higher-order thinking group do show slightly higher scores in the memory-based portions of the test, as does the higher order thinking group in their corresponding section. However, both these results were statistically insignificant.

### Table 5.7.
Results of the two tests across groups

<table>
<thead>
<tr>
<th>Scores (score/10 ; Total: score/20)</th>
<th>Test 1</th>
<th></th>
<th></th>
<th>Test 2</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Memory</td>
<td>Higher order thinking</td>
<td>Memory</td>
<td>Higher order thinking</td>
<td>Total</td>
<td></td>
</tr>
<tr>
<td>Memory group</td>
<td>3.50</td>
<td></td>
<td></td>
<td>6.83</td>
<td>6.17</td>
<td>13</td>
</tr>
<tr>
<td>Higher-order thinking group</td>
<td></td>
<td>7.21</td>
<td></td>
<td>6.33</td>
<td>6.67</td>
<td>13</td>
</tr>
<tr>
<td>No test group</td>
<td></td>
<td></td>
<td></td>
<td>7.17</td>
<td>5.67</td>
<td>12.83</td>
</tr>
</tbody>
</table>
VI. DISCUSSION AND FUTURE DIRECTIONS

Contrary to the initial hypothesis, there were no increases in reading time or fixation for the entire text for the test groups (memory and higher-order thinking). However, seeing as the no test group show marked decrease in reading time and fixation count, this lack of statistically significant change in the test groups could be caused by a practice effect, or because the second text was easier to process compared to the first text. Indeed, in the post-test interview, several participants mentioned that the second text allowed for a slightly easier reading experience. Thus, it can be said that participants who were subject to testing showed longer gaze durations for the second text when compared to the no-test group, which suggest that they allocated higher levels of attention after assessment. The no test group showed decrease in time and fixation especially for non-key parts of text (that is, parts of body text not included in ‘keyword’ ‘relation’ and ‘definition’ ROIs), suggesting a skimming type of reading behavior.

Overall, the memory test group showed greater changes in performance before and after testing, compared to the higher-order thinking test group. While no changes were observable in the test results, one possible explanation may be that the memory-based test induced the largest amount of affective response from the learners. Participants of the study were often flustered when faced with a test they did not expect, and this tendency was especially
more marked for the memory-based test. This may be explained by the interviews of two participants in the higher-order thinking group.

“I didn’t remember […] but it was doable, because I wrote as much as I knew.”

(Participant 2)

“Honestly I couldn’t remember the exact keywords… so [when I read for the second time] I tried to keep that into mind. I just answered [the first test] with what I already knew.”

(Participant 8)

Accordingly, there were less questions left blank in the transfer tests, and one participant (participant 1) even reported that the first test had been ‘interesting to solve.’ Another in the same group (participant 8) appraised the experimental process as being ‘enjoyable,’ a sentiment that was missing from the memory test group.

On the other hand, answering based on prior knowledge was an option that was less available to participants in the memory test condition, as they had to answer questions with clear-cut answers. One participant in the memory test group noted that he had expected a test from the first text, but still “felt spaced out when [he] saw the questions.”

Alternatively, it could be that college students are more accustomed to rote memory quizzes, and thus had readily accessible reading strategies for preparing for memory-based tests than they did for higher-order thinking-
based tests. This is supported by the way the majority of students described how they usually read to prepare for tests. Out of the 18 students, 12 students counted some form of memorization in their test preparation strategies. Even among the 6 who did not, three replied that they read the text, or a summarized version of the text, repeatedly until they learned the text completely.

Another possible explanation could be that the higher-order thinking questions for in the first test did not sufficiently necessitate inference-making inside the text. As in Ponce and Mayer (2014), regressive, nonlinear patterns of reading seem to occur when there is a strong need to tie together different information in the text. However, the higher-order questions asked in test 1 usually asked to evaluate, make predictions on, and analyze a situation using one concept from the text. This may have necessitated the linking of prior background knowledge and knowledge from the text, or knowledge from the text with the knowledge in the questions, but not between two or more knowledge from the text. If this is the case, there may have been other meaningful cognitive conditions that were not captured or discovered in the eye-tracking data during text reading. For instance, could a parameter in eye-tracking or in any other data-driven methodology pinpoint the cognitive process of linking prior knowledge to in-text knowledge? This is an interesting question that merits further investigation.

<table>
<thead>
<tr>
<th>Memory group</th>
<th>Summary of discussions: key patterns and possible explanations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Results suggest marked increase of attention in reading most</td>
</tr>
<tr>
<td></td>
<td>key parts of text. Especially for keywords, which they would</td>
</tr>
<tr>
<td></td>
<td>have known they would be prompted on, regressive reading</td>
</tr>
<tr>
<td></td>
<td>behavior is observed.</td>
</tr>
</tbody>
</table>

Table 6.1
| Possible explanations | i) Readers are more flustered, because there is a clear right/wrong answer. They accordingly put extra effort into subsequent reading.  

ii) They are more accustomed to this type of testing, and fall readily into established, effortful rote studying habits that the majority of them report using during actual reading for tests.  

| Transfer group | Overall reading effort do not show clear and unilateral changes; however, readers look at relations longer. They take less time to read keywords.  

Possible explanations | i) Readers are less stressed about the test, because they could write as much as they knew in the first test, even if they did not know the exact answer. One participant even report specifically the first test as being ‘interesting, while another in the group appraised the experiment as being ‘enjoyable.’  

ii) The transfer test tested transfer across prior knowledge and knowledge in the text, but not between knowledge in the text. Ergo, there was no regressive reading behavior.  

| No test group | Overall reading efforts decline.  

Possible explanations | i) Being highly skilled readers, the participants adjusted to the reading conditions (topic, eye-tracking setup) even after the practice text. This was offset by more effort in reading for the test groups.  

ii) The second text was slightly easier than the first text.  

Looking at the test results, there were no statistically significant differences found in the higher-order question scores, memory-based question scores and total score for test 2 between the three groups. There were no differences between test 2 scores for the test groups (memory and higher-order) and non-tested group. This does not correspond with the results of prior research on the interim test effect (Wissman, Rawson & Pyc, 2011; Carpenter, 2012; Brame & Biel, 2015). This is most likely to have occurred due to the size of the participant pool being too small to overcome individual variance and show a clear, unilateral change in test scores. While a larger number of
participants may have eliminated such incongruences, the research design was determined thusly nevertheless, as the primary purpose of the current paper was not to replicate prior research on the interim testing effect, but to observe differences in reading behavior after testing through eye-tracking. Under such purpose, it can be said that the test served more as a priming device than a standardized measure.

Additionally, it is interesting to note that while there were no statistically relevant differences in test scores, the eye-tracking measures were still able to highlight meaningful changes in the way participants read before and after testing. This is further evidence that more knowledge could be gained on the effect of testing on learning when observing the process of learning together with the results of tests.

Implications that results might have, taking into consideration the previous possible explanations, must be made cautiously. The small size of the population, as well as the descriptive, qualitative nature of the data makes this more of a departure line from which more practical research may arise. Nevertheless, there are several interesting issues that can be raised in this context of highly skilled readers in Korea. Such readers seem to be very test-oriented, and it may be said that their learning is tied closer than most to what is being tested. In a result-focused way of thinking, it could then be said that to make such readers put more effort into what they are reading, memory tests are the best tool. However, looking at the way many students noted their high levels stress in the memory test, while some in the transfer test said they
enjoyed the experience, it is clear that instructional designers must heed the
demotivating effects memory tests could bring. If the learning material needs
to be enjoyable, for instance due to age or the goal of education, caution must
be employed in using memory tests.

New questions arise from the implications and results that could be
addressed by future research. One, are there qualitatively different cognitive
processes that follow transfer tests that are not captured in eye-tracking? Two, why do memory tests elicit a regressive type of reading behavior? Three, how
could the cognitive process of linking prior knowledge with knowledge from
the text be observed? In all of these questions, using other qualitative methods
of online observation together with eye-tracking measures could offer more
insight, such as video analysis or concurrent think-aloud techniques. Lastly,
would a comparable process occur in real classroom situations? The results
of the study indicate that there may have been preexisting tendencies in the
way students read because of their awareness that this was an experiment. It
would be beneficial to investigate whether this effect of testing on reading
occurs in actual classrooms, on a larger population.

VII. SUMMARY AND CONCLUSION

All in all, the results of the current study suggest that there are
observable differences in reading after tests have been administered, and that
it differs between different types of tests. Decreased reading time and overall
fixations in the no test condition suggest that tests offset decreases in effort occurring in successive reading.

The memory test prompted readers to regress back more and longer to keywords that served as potential answers to the memory test. Readers in this group also paid more attention to definitions, and parts of text explaining the relationship between different notions.

The transfer test did not rouse so strong a response in reading but still caused readers to take more time to read parts of text explaining relationships.

While this study had a small population and could only show tendencies, the results still offer empirical support to the hypothesis that there truly is a metacognitive element to the effect testing has on subsequent learning. That is, it showed through the quantitative and objective results of eye-tracking that students switched their way of reading learning material to serve their purpose as primed by the existence and type of testing.

This study supports the common wisdom that tests, when administered in a right way, is an effective way to promote learning. At the same time, it indicates that learners react sensitively to the differences in test interventions, and that instructional designers must pay more effort to understand and take into consideration exactly how a test influences learning.
References


Kunze, K., Shiga, Y., Ishimaru, S., & Kise, K. (2013, August). Reading Activity Recognition Using an Off-the-Shelf EEG--Detecting Reading Activities and Distinguishing Genres of Documents. 12th International Conference on Document Analysis and Recognition (ICDAR), (pp. 96-100). IEEE.


Lobascher, S. (2011). What are the potential impacts of high-stakes testing on literacy education in Australia?. Literacy learning: The middle years, 19(2), 9.


Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological review, 63(2), 81.


Torrance, H., & Coultas, J. (2004). Do summative assessment and testing have a positive or negative effect on post-16 learners’ motivation for learning in the learning and skills sector? A review of the research literature on assessment in post-compulsory education in the UK.


Appendix

Appendix 1. Background Knowledge Test
Appendix 2. Practice Text
Appendix 3. Text Stimuli
Appendix 4. Test 1 - Memory
Appendix 5. Test 1 - Higher-order thinking
Appendix 6. Test 2
Appendix 7. Interview questionnaire
선수지식 확인문제

본 평가는 학습 내용에 대해 어느 정도 알고 있는지에 대해 파악하기 위한 것이므로, 정확히 아는 것만 답을 적어 주시기 바랍니다. 풀이 시간은 5분입니다.

1. 한 사회의 구성원들이 일반적으로 공유하는 정치적 규범이나 관습 및 사고방식으로 정의되며 지방형, 신민형, 참여형으로 나눠질 수 있는 것은 무엇인가?

2. 한 선거구에서 2명에서 10명까지의 의원을 선출함으로써 후보 선택이 제한되고 사표가 초래되는 단점을 최소화 하는 선거제도를 무엇이라고 하나?

3. 분쟁 당사자 사이에 제3자가 중개하여 합의를 성립시킴으로써 분쟁의 해결을 도모하는 제도는 무엇인가?

4. 행위를 규제하는 공통의 가치나 도덕기준을 잃은 혼돈 상태로 정의되며, 규범이 혼재되어 있어 어느 규범도 제 기능을 발휘하지 못함으로써 나타나는 무규범 상태를 무엇이라고 하나?

5. 특수 이익을 보호하기 위하여 입법이나 정책결정 과정에 영향력을 행사하는 집단을 무엇이라고 하나?

6. 시민이 그의 대표자인 국회의원, 기타의 대표자를 통하여 정치에 참여하는 것을 원칙으로 하는 민주정치를 무엇이라고 하나?

7. 우리나라라는 지역 대표성을 고려하여 선거구를 확정하는데, 이같이 행정구역을 기준으로 선거구를 확정하는 것은 집권당에 의한 _________________을 방지하기 위한 것이다.

8. 개인이 제도화된 규범이나 가치, 절차를 내면화 함으로써 정치 체계에 적응해 가는 과정을 정치 _________________라고 한다.

9. _________________주의는 상대적으로 자율적인 다수 집단을 상정하여 이러한 집단들이 정치 과정에 어느 정도 균등한 영향력을 행사할 수 있는 것으로 간주하는 것이다.

10. _________________는 정치적 갤등의 한 원인으로서 자기 지역의 이익만을 고집하는 형상을 일컫는다.

[Appendix 2: Practice Text]
실험 안내

실험이 참가해주셔서 감사합니다.

본 실험은 성인 학습자의 자연스러운 임기 과정에서, 사용되는 강조기법에 따라 어떠한 임
기패턴이 나타나는지를 관찰하고자 하는 것입니다.

간장은 들고, 최대한 평소에 금을 잃는 것과 유사하게 모니터의 글을 빵 주십시오.

여러분은 짧은 세 개의 글을 읽게 됩니다. 페이지 별 시간 제한은 없지만, 한번 페이지를 넘
긴 후에는 다시 뒤로 돌아가실 수 없습니다.

이 글을 모두 이해했다면 스페이스 바를 눌러주세요.

<Space Bar>

글 1

< space >
1. 드라마성 캐릭터를 설립 중에 이름

우리 관심의 단어와 단어의 조합 또는 단운 빠르고 캐릭터 이름 등의 창작물원을 부인하고 있다. 따라서 캐릭터 이름 자체를 이용하는 것은 저작권상 문제의 사례로 폐기할 필요가 있다. 또한 저작권에 보호대상인 창작물로서 아이디어 자체는 보호하지 않으므로 드라마가 등장하는 인물들을 독점적으로 캐릭터로 표현하는 것은 가능하다.

그러나 이때 표현된 캐릭터가 영화나 TV 드라마 등에 출연하는 배우들이 오랫동안 특정한 극중 역할을 맡음에 따라 그들의 독특한 이미지를 형성하게 되어 당해 인물이 연출될 수 있는 캐릭터를 상업적으로 이용할 때에는 '복잡리사리권' 문제가 발생할 수 있다.

드라마의 주인공과 같은 인물들이 해당 연예인의 하락 없이 드라마가 완전 인기를 얻고 있던 시점에 귀하의 드라마로 채용된 배우가 된 시점에서도 비록 이 사건 인물들은 별도의 실제 모습과 다르게 수용이나 보복도 해치지 못한 동일한 것은 아니지만 드라마 특성으로 인해 많은 판독의 특장점 부분들이 대부분 확정되어 있고, 오히려 아이디어를 보였거나 보복을 맡을 수 있는 사항이라면 누구나 이 사안 인물들과 보고 드라마의 영향력을 분명히 갖고 모욕할 때로는 무죄하다가 여 복잡리사리권 문제를 인정하였다.

2. 지도와 저작권

저작권법상 보호를 받는 창작물로서 상업화하기 위해서는 창작성의 요구한다. 저작권법은 도형저작물에 예시로 지도 도표 상세도 임도도보험 그 외의 도형 등을 열거하고 있는데, 이러한 도형저작물 중 특히 지도의 창작성에 관한 문제가 종종 제기되고 있다.

지도에서 창작성이 문제가 되는 이유는 지도는 주요성의 자연적 또는 인간적상의 형상에 전부 또는 일부를 말할 수 있는 적절한 기술로 사용하여 근본적으로 표현하는 것으로, 지도상에 나타나는 형상은 사실 그 자체에 허나지 아니하여 저작권의 보호대상이 되지 아니함은 물론이면 그 표현방식, 독특한 기호를 사용하며는 동 창작성이 벌어질 여지가 적고, 일시 잠재화하더라도 그 벌어가 매우 제한적일 수밖에 없기 때문이다.

따라서 지도가 저작권법상 보호되는 창작물로 상업화하기 위해서는 지도의 내용에 있는 자연적 형상과 인간적 형상의 총체를 다른 다른 지도들과는 다르게 새로운 방식으로 표현해야 할 것이다. 즉, 외부로 나타난 표현방식을 다자인이나 색상 등에 화학적인 요소가 없어지거나 사실정보는 달리의도적으로 크기 등을 이용하여 표현하는 경우 등이 저작물로 인정될 수 있을 것이다.

또한 근래 들어 지도를 3D형태로 재작해야 되어 3D지도의 저작물상에 대한 논의가 증증 있는데, 앞서의 사례에서 착한 바와 같이 저작권법은 아이디어가 아닌 표현물을 보호하므로, 3D형식의 지도저작법적인 아이디어에 불과하여 총론에 다자인이 화학적인 요소의 추가없이 기존 지도를 3D로 변환한 것에 불과하다면 그 자체로서 저작물로 성립한다고 볼 수 없는 것이다.
3. 아이돌 그룹의 안무와 작곡

작곡법은 '연극 및 무용 무용극 그 밖의 연극작품'을 저작물로 한 분류로 예시하고 있다. 따라서, 무용 등의 경우는 안무가에 의하여 창작적으로 만들어진 '동작의 변(形)'으로 저작물로 보호될 수 있음을 것이다. 하지만, 작곡법은 '연극의 사상이나 감정을 창작적으로 표현한' 부분을 저작물로 보호하고 있기에, 단순한 불통적 하나하나가 무용작품으로 보호되기에는 미루었을 것이고, 연속 동작을 통하여 사상이나 감정이 표현된 하나의 국물 표현하는 전체적인 부분이 보호받을 수 있을 것이다. 따라서, 연극에 있어서 하나의 국물 이하는 연속적인 안무 전체는 저작권법상 무용작품으로 인정될 수 있으므로, 이를 안무 교습소에서 고스란히 받고 갑작하거나, 공중 전에 그대로 이용하는 것은 저작의 문제가 발생할 수 있을 것이다.

이 사건은 유명 인공물의 저작권의 안무가인 원고가 자신의 연극을 단독교육학원에서 강습하고 이를 활영하여 홈메이저에 올려놓은 것에 대하여 작곡 재작을 주장한 사안으로, 법원은 단독의 창작적인 호흡, 가사의 내용 및 가사들에 적합한 일반의 신체적 동작 및 음조를 조합해낸 것으로 원고의 사상 또는 감정을 표현한 창작물이라 불리 상당하고, 피고들의 주장과 같이 대중에게 널리 알려진 소문난소의 기본적인 스타일의 부분적으로 포함되어 있다고 하더라도 이것은 사건 연극의 창작성을 주장하는 근거로 되지 않는다. 물론, 연극이나 안무가 재작권으로 인정받으려면 연극 자체의 독창성과 연무의 독특함이 나 매번 동작에 있어 반복작성이 가능하도록 일관성이 있어야 한다고 판단하였다.

<급 1을 모두 읽으셨습니다.>
글 2

< space >

1. 정치적 영향력:

모든 사회 체계에서 사람들은 정치적 목적을 이루려고 한다. 정치적 목적은 때에 따라서 정치적 변화를 포함하기도 하고 정치적 유지를 포함하기도 한다. 이러한 목적을 이루기 위해서는 정치적 영향력이 필요하다. 이에따라 다음 두 주장이나 단체들이 항상 정치적 변화와 유지를 위해 활동하기 때문이다.

정치적 영향력은 정치적 의사 결정 과정에서 다른 사람들보다 우위에 있어서 영향력을 행사하는 사람이나 단체의 힘을 말한다. 이러한 관계는 어떤 이 수에 대해 논쟁적이 작은 단체가 자신의 입장과 전략을 포기할 때 존재하게 된다. 정치적 의사결정 과정에서 큰 영향력을 가지고 있는 단체는 다른 단체에게 자신의 입장과 힘으로 할 수 있다. 그리고 정치적 영향력을 쉽게 가진 단체는 결국 더 영향력을 가진 단체를 파괴하게 될 것이다.

정치적 영향력이 어떻게 생기는 것인가? 정치적 영향력은 의사결정에 영향을 끼치는 정치 자율을 사용할 때 생긴다고 말할 수 있다. 정치적 영향력으로 바뀔 수 있는 정치 자율는 두 가지 종류가 있다. 첫째는 정치적 재산이고, 둘째는 정치적 이데올로기이다. 정치적 재산은 영향 자율과 직접 자율을 포함한다.

다. 만약 자율은 공중체에서 개인의 자율이나 정성, 분만, 특별한 논리, 조직화 기술과 같은 것을 말한다면. 그리고 물적 자원은 단체 구성원들이 가부한 돈, 장비와 사물들에 있다. 만약 어떤 사람에게 특정 이념에 대한 생물학적 변화를 한다면, 아마도 그 사람들은 자신의 이데올로기를 설명할 것이다.
정치적 연합적

정치적 계산과 정치적 이해와는 별개로, 어떤 단체가 강한 정치적 자원을 가지고 있을 수는 있지만, 항상의 좋은 이해용로는 가지고 있지 않을 수 있다. 또 다른 단체는 정치적 계산을 가지고 있지 않아 이해용로로서 매우 강한 성공이나 가치를 가지고 있을 수 있다. 정치적 계산과 이해를
모기 등에 해당한 계층이 변화하는 것이 경제적 변화와 포괄의 변화를 막는 하락한 계층 변화로 한다. 정치적 변화는 이러한 변화 과정에서 도출될 수 있다.

정치적 이해와는 다른 사람들의 어떤 관점에 따르도록 정치적 자원을 사용하는 것이 항상 포함된다. 정치적 이해가 정치적 변화를 나타내는 것은 아니다. 정치적 이해는 어떤 단체를 들는 방식으로 사용되어야 한다. 그리고 이러한 과정에 어느 정도의 경로를 따르는 데에 의해 정치적 변화를 공극적으로 파악할 수 있다.

여러분은 불필요해서 나타나는 경향적 이해는 공극적 변화를 줄일 수 있다. 변동성을 가진 사람이 자신의 지위를 사용하여 경향적 변화를 광범위하게 한다. 또한 노동 구조는 대중들의 이해와 이해를 배출하기 위해 자신의 이해로부터 경향적 변화를 줄일 것이다. 이것은 강력한 자원을 가지고 있는 정치 단체가 그 자원을 활용하여 여성 결합에 경향적 변화를 방해할 수가 있다. 그들은 또한 정치적 자원을 부작용한 방식으로 사용하겠다고 주장함으로써 같은 결과를 얻을 수 있을 것이다.

2. 정치적 발전

모든 사회 체제 안에 있는 사람들은 두 가지 방법으로 이용하는 경향이 있다. 그 방법 중 하나는 정치적 변화하고 다른 하나는 정치적 유지이다. 여기에서
는 소위 정치적 발전이라고 불리는 정치적 변화에 대해 살펴보겠다.

정치체제는 여러 방향으로 변화할 수 있다. 그 중 전 가치 방법은 특정 기간 동안 정치에 필요한 여러 자원들이 각각 다르게 유용하려는 것이다. 어떤 집단
은 더 부유해지고 더 높은 위치로 올라갈 수 있으며, 그들이 복잡한 것을 이루기 위해 더욱 효율적인 안전과 자원을 차지할 수 있다. 자원이 밝혀에서의 변화는
공동체의 정치 참여 패턴에 영향을 끼치고 개인과 집단의 정치적 영향력을 변화시킨다. 이러한 변화 과정에 있어서 공극적으로 이루어지는 것이 정치적
의사결정이다.
정치적 발전은 정치 체제 전반에 걸쳐 일어난 변화들을 하나로 통합시키는 과정이다. 통합 과정 안에서는 정치 참가, 지도력, 참여, 의사결정의 모든 변화들이 하나의 발판으로 작용한다. 이러한 변화의 과정이 전반적이라면 하나의 발판으로 동작한다면 우리는 정치적 발전이라고 말할 수 있다.

정치적 발전은 자원, 지도력, 참여, 의사결정 등의 요소들이 동시에 다른 방향으로 변화하는 경향을 말하는 것이다. 정치적 발전 과정에서 이러한 요소들의 변화가 서로 다르고 매료는 고립된 의사결정에 요구한다는 것을 개선할 수 있다. 매료하면 사회 공통체는 동시에 서로 다른 정치 체제와 복잡성을 함께 나아가고 있기 때문이다.

정치적 발전과 정치적 발전은 두 가지 쪽면에서 다르다. 첫째는 발전과 발전이 일어나는 시간 차원에서 다르다. 정치적 발전은 진전되는 과정이지만, 정치적 발전은 일정한 시간에 제한을 받기 때문에 주요한 과정이라 할 수 있다. 둘째는 정치 변환의 통합과 관련이 있다. 정치 변환들의 통합의 확실히 이루어지지 않는다면 그 과정은 정치적 발전이라 할 수 없다. 반면에 변화들이 전반적이고 같은 방향으로 동일하다면 정치적 발전이 이루어진 것이다.

<글 2를 모두 읽으셨습니까.>
3. 정치적 재산

우리의 정치 체제는 정치 활동을 통해 유지되거나 변화한다. 만약 구성원들이 새로운 체제를 원한다면 기존의 체제에 대항하여 반대할 것이고, 반면에는 체제에 만족한다면 그것을 유지하라고 할 것이다. 정치적 재산은 이 과정에서 평형적을 형성하기 위해 사용하는 정치 자원이다. 그리고 정치 활동은 이러한 정치 자원의 문제에 영향을 준다. 앞으로 무엇이 정치적 재산을 구성하는지, 그리고 의사결정에 직접적으로 영향을 미치기 위해 정치적 재산이 어떻게 사용되는지를 함께 알아보도록 하자.

정치적 재산은 물질 자원과 인적 자원의 결합이다. 사람들은 보통 의사결정에 영향력을 미치기 위해 돕거나 사무실 설비 등의 물질을 사용한다. 이렇게 사용되거나 가동되는 모든 물질 자원은 정치적 재산이 되며, 이것은 항상 사회에 의해 가치가 평가된다.

정치적 재산을 나타내는 또 하나는 인적 자원이다. 이것은 사실상 개인적이고, 정치적 의사결정에 영향력을 미치는 매우 중요한 자원이다. 예를 들어, 어떤 사람들끼리 공동체 내에서 높은 위치에 있다고 할 때, 낮은 위치에 있는 사람들보다 더욱 강력한 영향력을 행사할 수 있고, 다른 사람들로부터 이의를 의견을 더욱 경청할 수 있게 할 수 있다. 우리 중 대부분은 많은 종류의 인적 재산을 소유하고 있다. 사회적으로 높은 위치에 있지 않는 사람들도 전문 지식, 조직 능력, 정보라고 하는 지식 등이 다른 재산의 인적 자원을 가지고 있을 수 있다.
정치적 평가

한 악에 대한 모든 개인적 자선은 단지 자율적으로 일어날 수 있다. 만약 한 악에 대해 충분한 연락이 없다면 그 단체를 아우고, 조직하고, 단체 내의 의견을 사회에 전달할 수 있는 목적을 위해 적절한 사람들을 찾아 모아낼 수 있다. 정치 자원은 개인의 자산을 공동의 목표를 위해 내아 놓는 사람들에 인위적 영기를 얻는 것이고, 활동 가능한 민족의 많은 개인적 자선을 내어놓는 사람의 수에 달한다. 정치적 자선과 단체 자선 모두를 합한 것은 한 집단의 정치적 총체성을 나타낸다.

물질 자원과 정치 자원은 본질적으로 다르다. 모두 그것은 사람들이 행동에 требуется 물질 자원과 정치 자원을 사용한다. 물질 자원은 사람들을 위해 거대하거나 특별히 위급한 상황에 의해 기부된다. 그러나 정치 자원은 기부될 수 없다. 보호 그 능력을 가지는 사람이 기부할 수 있는 것이다.

정치 집단은 국가 결정에 영향력을 발생하기 위해 자신들의 물질 자원과 정치 자원을 사용한다. 그러나 그러한 자원들이 영향력에 영향을 미치는 형태는 매우 따라 변화하는 방향이다. 특정 선거에서 성공적이었던 자원이 다른 선거에서는 효과를 못 보 수도 있다.

상황, 다 자의 3차의 이론에 의문이 있는 정치 자원의 영향에 따라 달라진다. 첫째, 정치적 자원의 숫자가 점차로 증가하지 않아서 주어진 정치적 자원의 영향에 따라 다른 영향을 미치는 영향이다. 둘째, 사람들이 연계나 같은 목적으로 같은 자원을 기부하는 데가 때문이다. 셋째, 정치 활동이 영향에 따라 정치적 자원의 공통적 채권에서 다르게 분배되기 때문이다.

4. 정치적 의사결정 원칙

정치적 대안을 마련할지 아니면 거부할지 결정할 때, 두 가지 기본적인 의사결정 원칙을 적용할 수 있다. 첫 번째 원칙은 에릴리스 원칙으로, 이것은 의사결정 핵심의 단일이 순전히 법칙이다. 두 번째는 검과 연속으로 의사결정에 있어서 100% 지지를 둔 것이 요구되는 원칙이다.

그러나 이 두 가지의 의사결정 원칙은 실제 주변에 적용되기 매우 어렵다. 매년 야외로 통합한tri에서 얻어진 의사결정은 집단 전체의 의견을 대변하게 될 수가 있고, 협의의 끝에서 역시 모든 사람의 동의를 얻어야 할 수 있기 때문이다. 이러한 이유로 의사결정 원칙은 계속 수립되었다. 수립된 원칙으로 다음의 두 가지를 볼 수 있다.

첫 번째는 수립한 야외리스 원칙으로, 수립한 야외리스 원칙은 개인의 아닌 소집단에게 의사결정 책임을 부여하는 것이다. 이러한 의사결정의 방식은 최종 결정된 사람이 집단 사람들과의 의견을 더욱 잘 반영하도록 설계된다. 그러나 여전히 소수의 야외리스가 모여 수권, 수권 소수의 사람들에게 영향을 미치는 의사 결정을 한다는 문제를 안고 있다.
두 번째는 수험원 합의 원칙을 들 수 있다. 이것은 기존의 합의 원칙보다 실제 현장에 쉽게 적용할 수 있기 때문에 많이 사용된다. 수험원 합의 원칙에서는 모든 사항을 100% 동의를 얻을 필요가 있다. 그 대신, 의사결정의 성장을 나타내기 위해 충분한 비율의 동의만 얻으면 되는데, 이것은 보통 50% 수준이다.

엘리트 원칙과 합의 원칙은 근거한 의사결정에서 중요할 수 있는 이러한 틀들의 각각 결과에 다른 영향을 미칠 수 있다는 것이다. 합의 원칙이나 수험원
합의 원칙 하에서는 전체 동의 비율이 의사결정에 영향을 미칠 경우를 무시해 흐른다. 그러나, 엘리트 원칙이나 수험원 엘리트 원칙 하에서는 다른 정적 대안을 지지하는 대수의 세칙이 존재하더라도, 의사결정 책임이 부여된 개인이나 집단이 다수의 요구를 고려하여 다수의 의지를 거부할 수도 있다.

<글 3을 모두 읽으셨습니다>
다음의 문제에 답하시오.

1. 정치적 의사결정은 두 가지 목적을 위해 사용될 수 있다. 하나는 정치적 변화이고 다른 하나는 정치적 ________________(이)다.

2. 정치적 재산에는 두 가지 종류가 있는데, 이 중에서 지위, 명성, 사회적 능력은 ________________자원이라고 불려진다.

3. 정치적 영향력은 ________________과정에 어느 정도로 영향을 미치는가에 의해 궤극적으로 파악할 수 있다.

4. 정치에서 ________________분배 패턴의 변화는 개인과 집단의 정치적 영향력을 변화시킨다.

5. 정치적 발전과 정치적 갈등은 두 가지 차원에서 중요한 차이점을 가진다. 한 가지 차원은 시간이고 다른 한 차원은 ________(이)다.

답안:

1. 유지
2. 인적
3. (정치적) 의사결정
4. 자원
5. 통합 / 방향
다음의 문제에 답하시오.

1. 정치적 영향력이라는 개념을 사용하여 대기업의 기술 탈취가 성립하는 이유를 설명하시오. (100자 이하)

대기업의 중소기업 기술 탈취는 해마다 거론되는 고질적인 문제다. 중소기업은 가진 신기술로 대기업과 거래하기 위해 특허출원 명세서와 각종 실험 데이터 등을 제출하는데, 대기업이 이 자료를 활용해 새로운 특허를 출원하거나, 협력업체에 기술을 넘겨 납품단가를 깎는 데 활용하는 경우가 발생하고 있다. 중소기업으로서는 시간과 비용 때문에 승률이 낮은 법적 소송을 포기하는 것이 보통이다.

2. 개인적 자산, 물적 자산, 이데올로기 측면에서 다음의 정치적 자원을 분석하시오. (100 자 이하)

① 로펌 변호사의 전직 재판관이라는 경력
② “아이들이 행복한 세상”이라는 유니세프의 모토
③ 특정 정당에 속한 정치인의 대중 호감도
④ 기독교 기업인이 기부한 교회 부지
⑤ 노동 조합의 조합원이 가지고 있는 연대라는 가치
⑥ 기업의 대표가 선행을 통해 구축한 좋은 평판
⑦ 대통령 출마 후보의 선거본부에서 사용하는 컴퓨터
⑧ 초대 회장이 정립한 모 회사의 인본주의 사상
3. 다음 시나리오에 등장하는 A 기업 노동조합이 보유한 정치 자원이 어떻게 변화하였는지를 진단하시오. (100 자 이하)

A 기업은 업계 전반에 걸친 수주 가뭄에 대규모 구조조정을 추진하였지만 노동조합의 반발에 부딪혔다. 노조는 강력한 지도력을 통해 8 주간의 총파업과 역대 최고 인원을 동원한 점거 농성을 성공시켰다. 그러나 하반기 매출이 역대 최약임이 밝혀지고, 정부의 구조조정 압박이 심화되면서 사측은 구조조정 대신 15% 임금 반납을 제안했다. 이 제안에 노조 지도부는 파열이 나뉘어 대립했고 결국 협상은 결렬되었다. 파업은 장기화되었고, 급여 지급이 늦어지자 노조원들은 경제적인 어려움을 호소하였다. 노조원들은 점차 집행부에 대한 지지를 철회하고 집회에 불참하기에 이르렀다.

4. 다음은 정치 체제가 변화할 수 있는 한 가지 방향이다. (2)~(5)에 해당하는 단계를 쓰시오.

<table>
<thead>
<tr>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>정치에 필요한 물적 자원이 A 협회에게 많이 분배됨</td>
<td>공동체의 구성원이 A 이슈에 우호적인 정치인을 더 뽑게 됨</td>
<td>A 협회가 부유해짐</td>
<td>정치적 의사결정이 A 이슈에 유리하게 이루어짐</td>
<td>A 협회가 광범위하고 공격적인 캠페인을 펼침</td>
</tr>
</tbody>
</table>
5. 다음이 정치적 발전인지 갈등인지를 판단하고, 그 이유를 서술하시오. (100 자 이하)

여당인 A 당은 야당 B 당에게 노동 개정안에 동의하지 않을 경우 내년도 예산안을 단독 처리하겠다는 최후 통첩을 지난 달 보냈다. 원내 과반을 차지한 A 당은 부정부에 자신의 입장만을 추가한 수정안을 올려 단독 처리가 가능하다. 야당은 여당과의 밀접 협상 끝에 합의처리 및 본회의 개최를 약속하게 되었다. 그러나 이에 불만을 품은 일부 야당 의원은 의총을 열어 본회의 참석 거부 의사를 밝혔고, 여당은 직권상정을 추진했으나 선거로 바빠진 의원들이 대거 불참하여 이마저 무산되었다.

채점 루브릭

<table>
<thead>
<tr>
<th>기준</th>
<th>점수</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>정치적 영향력의 효과가 무엇인지지를 개념적으로 설명한다. Ex) 큰 정치적 영향력을 가진 단체는 다른 단체에게 입장을 따르게 할 수 있다.</td>
</tr>
<tr>
<td></td>
<td>이것을 주어진 예시에 대입하여 설명한다. Ex) 즉, 대기업은 자본과 인력 등 중소기업에 비해 정치 자원이 더 많고 이것을 활용하여 막강한 정치적 영향력을 행사하기에, 중소기업은 자신의 이권을 포기하게 되는 것이다.</td>
</tr>
<tr>
<td></td>
<td>맞게 분류한 한 항목마다 0.25 점 (답: ①: 1,3,6 / ②: 4,7 / ③: 2,5,8) 각 025</td>
</tr>
<tr>
<td>2</td>
<td>노조의 초기 정치 자원을 올게 설명한다. ➔ 강력한 지도력을 통해 많은 인적 자원을 파업 및 농성에 동원할 수 있었다.</td>
</tr>
<tr>
<td></td>
<td>노조가 초기에 보유했던 정치 자원을 잃었음을 올게 설명한다 ➔ 그러나 지도부가 파벌로 나뉘며 구성원들의 신망을 잃게 되었다. 집회에 참여하는 인원이 줄었다는 것은 인적 자원이 소실된 것이다.</td>
</tr>
<tr>
<td>4-2-5-3 (각 0.5 점)</td>
<td>2점</td>
</tr>
<tr>
<td>4</td>
<td>갈등을 올게 고른다.</td>
</tr>
<tr>
<td>5</td>
<td>첫 번째 이유를 설명한다. ➔ 시간적 측면에서 살펴보았을 때 진전이 없고, 일정한 시간의 제한을 받으며 멈추어 있다.</td>
</tr>
<tr>
<td>6</td>
<td>두 번째 이유를 설명한다. ➔ 서로 반대되는 방향의 정치적 변화가 서로 대치되듯이 일어나고 있다. 정치적 발전이 되려면 전반적이고 같은 방향으로 변화가 통합되어야 한다.</td>
</tr>
</tbody>
</table>
Appendix 6. Test 2

다음의 문제에 답하시오.

1. 정치적 재산은 정치적 __________________ 에 가장 직접적으로 영향을 미친다.

2. 한 단체 내의 모든 개인적 자산은 _____________ 자원이라고 말할 수 있다.

3. 자신이 가지고 있는 개인적 재산을 한 가지 들고, 이것이 자신이 속한 공동체 속에서 정치적 의사결정에 어떠한 영향력을 미칠 수 있을지를 서술하시오.
(100 자 내)

4. 다음 시나리오에 등장하는 A 씨의 자원이 정치적 재산이 될 수 있는 두 가지 방법을 주어진 맥락에서 설명하시오. (100 자 내)

A 씨는 신재생 에너지에 관심이 많은 시민이다. 그는 서울 00구에 건물을 한 채 보유하고 있다. 돌아오는 시장 선거에서 그는 B 후보와 C 후보 중, 신재생 에너지로 생산된 전력량에 대해 보조금을 지급하는 발전차액지원제도에 찬성하는 C 후보를 지지하고자 한다.

5. 정치적 재산의 두 유형은 본질적으로 서로 다르는데, 한 가지 자원만이 _____________가 가능하다는 점이 그러하다.

6. 의사결정 과정에서 사용될 수 있는 두 가지 원칙은 _______ 원칙과 엘리트 원칙이다.

7. 정치적 의사결정의 두 기본적인 원칙 중 어떤 것이 국가 차원에서는 더욱 현실적으로 적용이 어려울 것으로 예상되는가? 그 이유를 밝혀시오.
8. 수정된 엘리트 원칙은 개인이 아닌 ______________에게 의사결정 책임을 부여한다.


미 대선에서 기업인 A 후보는 재계 및 자신 개인의 자금을 투입하여 선거자금 4억 달러를 모았다. 반면 B 후보는 소액 기부자들이 모금액의 80%를 차지했지만 총 금액은 1억 5천 달러에 그쳤다. 얼마나 선거 자금을 많이 모으느냐에 따라 승패가 달라진다고 까지 하는 미 대선이지만, 신문들은 앞다투어 B 후보의 독특한 행보와 ‘B 열풍’에 대해 보도했다. 결국 B 후보는 대선에서 승리를 거두었다.

10. 동, 서양을 막론하고 국가의 정치 체제는 절대 왕정에서 민주주의로 변화하여 왔다. 이러한 과정이 일어났을 이유를 의사결정 원칙의 관점에서 설명해 보시오.
## 채점 루브릭

<table>
<thead>
<tr>
<th>답</th>
<th>점수</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1</strong> 의사 결정</td>
<td>2점</td>
</tr>
<tr>
<td><strong>2</strong> 인적</td>
<td>2점</td>
</tr>
<tr>
<td><strong>3</strong> 자신이 가지고 있는 개인적 재산을 올게 돼 있다. Ex) 00 분야에서의 전문성, 공부하고자 하는 의지, 리더십, 평판 등</td>
<td>1점</td>
</tr>
<tr>
<td><strong>4</strong> 이것이 공동체의 의사결정에 어떠한 영향력을 미칠 수 있을지를 타당하게 적는다. Ex) 원하는 방향으로 동아리/학과/소모임 등을 이끌어 나간다.</td>
<td>1점</td>
</tr>
<tr>
<td><strong>5</strong> 한 가지 방법을 올게 설명한다. 건물을 C 후보가 원하는 대로 선거에 활용할 수 있도록 선거기간 동안 제공한다. (선거본부, 옥외광고 부착 등)</td>
<td>1점</td>
</tr>
<tr>
<td><strong>6</strong> 다른 한 가지 방법을 올게 설명한다. 건물을 기증하거나 매각하여 대금을 기부한다.</td>
<td>1점</td>
</tr>
<tr>
<td><strong>7</strong> 거래 (기부)</td>
<td>2점</td>
</tr>
<tr>
<td><strong>8</strong> 합의 (합의 의사 결정)</td>
<td>2점</td>
</tr>
<tr>
<td><strong>9</strong> 의사 결정 원칙이 더 어려울 것임을 올게 고른다.</td>
<td>1점</td>
</tr>
<tr>
<td><strong>10</strong> 이론을 타당하게 든다. 국가 차원에서의 의사결정에서는 모든 국민의 동의를 받아야 할 것이며, 이것은 사실상 불가능하다.</td>
<td>1점</td>
</tr>
<tr>
<td><strong>11</strong> 소집단 (소수, 소수의 엘리트, 소그룹, 소수 집단)</td>
<td>2점</td>
</tr>
<tr>
<td><strong>12</strong> 정책자원이 의사결정에 미치는 영향은 상황에 따라 달라질 수 있음을 밝힌다.</td>
<td>0.5점</td>
</tr>
<tr>
<td><strong>13</strong> 3 가지 원인 중 해당 상황에 부합하는 세부적인 이유를 밝힌다. 정책자원의 가치는 항상 목적이 달성되거나 지속되지 않는다 (공동체에 따라 다른 영향을 미친다.)</td>
<td>0.5점</td>
</tr>
<tr>
<td><strong>14</strong> 해당 예시에 개념을 대입하여 설명한다. Ex) 즉, 4억 달러와 1억 5천 달러 중, 해당 국가의 여론에서는 후자가 더 가치가 높은 것으로 판단된 것이다.</td>
<td>1점</td>
</tr>
<tr>
<td><strong>15</strong> 국가의 정치 체제는 정책 엘리트 원칙에서 합의 원칙으로 변화하여 왔음을 올바르게 선택한다.</td>
<td>1점</td>
</tr>
<tr>
<td><strong>16</strong> 합의 원칙과 엘리트 원칙의 특성에 바탕을 둔 합리적인 이유를 든다. Ex) 정책 대중의 빈곤에 의한 의무결정권을 요구하게 되었을 것이다. 엘리트 원칙은 대중의 의사를 반영할 수 없으므로 반발이 컸을 것이다. 합의 원칙이 더 많은 의사 결정의 정당성을 갖기 때문이다.</td>
<td>1점</td>
</tr>
</tbody>
</table>
Appendix 7. Interview questionnaire

인터뷰 질문지

1. 글 1을 읽을 때 어떠한 전략을 사용했는가?

2. 첫 번째, 두 번째 시험을 볼 때의 사고는 각각 어떠하였는가?

3. 글 1과 2를 읽을 때 스스로 판단하기에 읽기 방식이 변화한 부분이 있는가?

4. (3에 이어) 있다면 무엇인가? / 없다면 그 이유는 무엇인가?

5. 글 1과 2의 이해도/기억한 정도/ 흥미도를 각각 1 부터 10 까지 매긴다면?

6. 평소에 시험 준비를 할 때 어떠한 식으로 시험 범위의 글을 읽는가? 전략, 태도 위주로 설명한다면?
국문 초록

시험이 이후 읽기 행동에 미치는 영향

서울대학교 대학원
교육학과 교육공학 전공
성 가 현

효과적인 학습 설계를 위하여 교수 설계자는 특정한 교육적 개입이 학습자에게 어떠한 효과를 불러일으킬지를 충분히 이해해야 한다. 널리 사용되고 있으면서도 다양적인 효과를 불러일으켜 논란이 끝이지 않는 교육적 개입으로 시험이 있다. 시험은 학습자를 이해하고 학습을 촉진시키는 중요한 도구이지만, 시험의 부정적인 효과 또한 존재한다는 것은 널리 알려진 사실이다. 시험이 학습에 미치는 영향에 관한 연구 대부분은 시험의 효과를 또 다른 시험으로 측정하여 왔다. 이러한 효과가 어떠한 기저 원리를 통해 발생하는지, 또 시험이 학습의 결과가 아닌 과정에는 어떠한 영향을 미치는지에 대한 연구는 상대적으로 부족하다고 할 수 있다.

따라서 해당 연구에서는, 시험의 유무와 종류가 중요한 학습 과정인 읽기에 어떠한 영향을 미치는지를 탐구하고자 한다. 이러한 탐구는 기존에는 학습자의 자기 보고에 의존하여 왔으나, 해당 연구에서는 보다 정밀하고 정량화 가능한 결과를 위해 아이트래킹 기법을 활용하였다. 실험이라는 참가자를 파지 기반, 고차적 사고 기반 시험 집단과 무시험 통제 집단으로 나누어 아이트래킹 데이터와 인터뷰, 시험 결과 데이터를 수집하였다. 실험을 보기 전과 보고 난 후, 아이트래킹에서 나타난 읽기 패턴을 질적 데이터와 함께 해석하였다.

연구 결과, 시험이 이후 읽기 행동에 영향을 미친다는 사실이 아이트래킹 데이터를 통해 확인되었다. 무시험 집단의 경우 읽기에 걸리는 시간과 시선 고정 횟수가 전반적으로 줄어들었지만, 시험을 친 집단의 경우 이러한 현상이 나타나지 않았다. 또한, 파지 기반 시험의 경우, 글의 내용 중 정의와 키워드, 관계 설명구에 더 많은 주의를 기울이는 경향을 보였고, 그 중 키워드의 경우 한 차례 읽은 후에도 다시 돌아와 읽는
역행성(regressive) 읽기 행동을 보였다. 고차적 사고 기반 시험 집단의 경우 관계 설명구에 더 많은 주의를 기울이는 경향을 보였다. 전반적으로 시험에 대한 반응이 파지 기반 시험에서 더욱 확연히 드러난 것은, 기존에 학습자들이 가지고 있는 시험에 대한 기대 및 인식, 파지 기반 시험에서 학습자들이 느낀 스트레스, 그리고 해당 연구에서 사용된 고차적 사고 기반 시험 문제의 특징 등에 의한 것으로 추정되었다. 이러한 결과를 바탕으로 추후 연구를 위한 몇 가지 연구 문제가 제안되었다.

주요어: 시험, 텍스트 이해, 읽기 전략, 읽기 목적, 아이트래킹, 설명문
학번: 2016-21518