



RUTGERS LAW RECORD

The Digital Journal of Rutgers Law School

lawrecord.com

Volume 45

2017-2018

IN THE EYES OF THE LAW STUDENT: DETERMINING READING PATTERNS WITH EYE-TRACKING

TECHNOLOGY

CATHERINE J. CAMERON¹

First-semester law students spend over twenty hours per week reading legal cases in their casebooks for their law school classes.² By the time students have completed their first year of law school, they have spent upwards of six hundred thirty hours reading cases in casebooks.³ And upon

¹ Professor of Law, Stetson University College of Law. I would like to thank Stetson University College of Law for the scholarship grant that supported this study and Stetson University's Center for Excellence in Advocacy for the use of its eye-tracking software. Additionally, I owe many thanks to the research assistants whose assistance in collecting and processing the data for this study was invaluable – Giovanni Giarratana, Samantha Grizzle, and Kelley Parks.

² LAW SCHOOL SURVEY OF STUDENT ENGAGEMENT, ENGAGING LEGAL EDUCATION: MOVING BEYOND THE STATUS QUO 14 (2006), http://lssse.indiana.edu/wp-content/uploads/2016/01/LSSSE_2006_Annual_Report.pdf [hereinafter LSSSE] (summarizing the 2006 Law School Survey of Student Engagement Results). Although there are other types of texts that some professors use to teach first-year law classes, the casebook is the predominant type of text assigned to first-year law students. Stephen M. Johnson, *The Course Source: The Casebook Evolved*, 44 CAP. U. L. REV. 591, 591 (2016); Stephen M. Shepard, *Casebooks, Commentaries, and Curmudgeons: An Introductory History of Law in the Lecture Hall*, 82 IOWA L. REV. 547, 593 (1997). Even though casebook authors do often include commentary and questions to help facilitate learning legal rules, the predominant information in a casebook is case law that has been edited by the author to meet the needs and confines of the casebook format. Johnson, *supra*, at 619-22; RUTH ANN MCKINNEY, *READING LIKE A LAWYER: TIME-SAVING STRATEGIES FOR READING LAW LIKE AN EXPERT* 5-7 (2d ed. 2012).

³ LSSSE, *supra* note 2, at 14. The LSSSE survey indicates that 1Ls average 21 hours a week for reading cases in their assigned casebook reading. This weekly total multiplied by the 30 weeks of classes that the ABA's Program of Legal Education Standard 310 contemplates for the first year of law school equates to 630 hours. See STANDARDS AND RULES OF PROCEDURE FOR APPROVAL OF LAW SCHOOL 21 (AM. BAR ASS'N 2017).

graduation, graduates have slogged through over 1,550 hours of reading cases in casebooks.⁴ Does the human brain develop strategies to make this repetitive task easier and what might that mean for readers and writers of cases? Although researchers have looked into the reading skills of legal readers through an empirical lens since the late 1980s, almost all of their studies have focused on the critical reading and rhetorical skills legal readers use to comprehend text, and the data collection has been achieved through self-reported information from the legal readers themselves.⁵ This article attempts a different approach – to look at the differences between novice legal readers and experienced legal readers through independent empirical evidence.

This article details a study, conducted during the 2015-2016 academic year at a private, suburban law school, which was designed to determine if there are short-cuts that law students develop to make their reading processes more efficient for their purposes. The study found a dramatic change happened to students who had begun the process of reading cases on a daily basis. Using eye-tracking technology to study the reading patterns of thirty-four law students in various points of their law school career, the study results showed that students attending law school (“experienced readers”) read certain parts of a legal opinion at the same speed as students who were not attending law school (“novice readers”). However, there was a statistical difference in the speed at which experienced readers read sections of opinions that were not as critical for understanding the holding of a case, and this difference shaved off over a quarter of the reading time for the experienced readers. Despite the speed difference, there was no statistical difference found in basic

⁴ LSSSE, *supra* note 2, at 14. The LSSSE survey indicates that the average weekly casebook reading totals for a 2L and 3L are 18 hours and 13 hours a week, respectively. When these numbers are multiplied by 30 weeks of classes each and added to the 630 hours previously calculated for 1Ls, the combined total becomes 1,560 hours.

⁵ James F. Stratman, *When Law Students Read Cases: Exploring Relations Between Professional Legal Reasoning Roles and Problem Detection*, DISCOURSE PROCESSES, June 8, 2010, at 57, 58-62.

comprehension between the novice readers and experienced readers. This study suggests that the brain learns how to speed up reading over information in a case that is non-essential to the student's basic understanding of the case, but does not employ these strategies when the information is critical to the basic understanding of the case. These findings verify and explain previous findings that novice legal readers and experienced legal readers in the law school setting read differently but perform similarly on basic comprehension and analysis tests. Verifying this difference opens up avenues for research that could reveal assessment methods that could assist law students with improving their reading skills.

Introduction

There is no shortage of references in legal scholarship regarding the importance of reading skills for law students as well as practicing lawyers. Some of the most important scholars in the legal research and writing field have pointed out that reading forms the basis of all legal knowledge, because it is only through reading cases that students gain the necessary legal knowledge to participate in classroom discussions that broaden and deepen their knowledge of the law.⁶ Although many law professors have undertaken introducing more progressive teaching methods, the traditional structure of legal education involves a cyclical pattern of students reading cases that are then discussed in class.⁷ Many of the popular press books that undergraduates can purchase to

⁶ See MCKINNEY, *supra* note 2, at 3-13.

⁷ See *id.*; WILLIAM M. SULLIVAN ET AL., EDUCATING LAWYERS: PREPARATION FOR THE PROFESSION OF LAW (2007) (Although there has been much innovation in teaching since the 2007 Carnegie Report criticized the teaching methods of law schools as being outdated, most law schools have opted to add interactive components to their otherwise traditional Socratic teaching method, rather than doing away with the Socratic teaching method all together). Jamie R. Abrams, *Reframing the Socratic Method*, 64 J. LEGAL EDUC. 562, 563 (2015) (“The Socratic method persists and endures in law teaching, even while it is increasingly surrounded by innovation and its use is declining. The current approach to legal education is to *add* innovation, such as enrichment and skills opportunities, while simultaneously *retaining* the hallmarks of traditional legal education – the large, lecture-style doctrinal course taught using the Socratic method and the casebook rooted in appellate cases.”).

inform themselves about how to best prepare for law school warn law students of the struggle they may have adjusting to this learning style. One book warns pre-law students that they should not expect law school to be like their undergraduate courses that deliver information in a “relatively straightforward” fashion.⁸ Instead, students should know that “[t]o say that law school is less straightforward is an enormous understatement.”⁹ Another tells students, “[l]earning what the law is from reading a lot of cases is like looking for numbers in those color-blindness tests that have all the dots in them.”¹⁰ These books advise students that the main point of learning by reading cases is to learn how courts apply legal rules so the students can then apply those rules to other fact patterns.¹¹

Despite the importance of reading in the law school curriculum, it wasn’t until the late 1980s that researchers began to look at how legal readers read through empirical studies.¹² And since that time, there have been only a few researchers who have conducted studies that have looked empirically at student reading.¹³ Almost all of these studies collect data through the use of questionnaires that ask readers about their reading experience or the “think aloud” data collection technique.¹⁴ The “think aloud” technique asks participants to read text aloud and intersperse their reading with oral descriptions of their thought processes as they read the text.¹⁵ The students are often aware that they are being recorded during this process. This process is considered to be more

⁸ ANN L. IJIMA, *THE LAW STUDENT’S POCKET MENTOR* 34 (2007).

⁹ *Id.*

¹⁰ NANCY B. RAPOPORT & JEFFREY D. VAN NIEL, *LAW SCHOOL SURVIVAL MANUAL: FROM LSAT TO BAR EXAM* 56 (2011).

¹¹ *Id.*; Iijima, *supra* note 8.

¹² Stratman, *supra* note 5 at 57-58.

¹³ *Id.*; see also Leah M. Christensen, *Legal Reading and Success in Law School: An Empirical Study*, 30 SEATTLE U. L. REV. 603, 610-14 (2007) (detailing the four empirical studies conducted on legal readers before Christensen’s study).

¹⁴ Christensen, *supra* note 13.

¹⁵ *Id.* at 617. In Christensen’s study, she “instructed the participants to read the text aloud, stopping every sentence or two to state what they were thinking.” *Id.* at 618.

reliable than questionnaires given to subjects about their reading processes and thoughts after they have completed reading text.¹⁶

Leah Christensen used this technique in a reading study she conducted on 24 first-year law students.¹⁷ She coded their comments about their thought processes while they were “thinking aloud,” and discovered that the students who earned better grades in law school were ones who expressed more sophisticated reading processes while “thinking aloud.”¹⁸ The students receiving the highest grades demonstrated that they were reading critically; that is, they attempted to make connections between concepts in the material.¹⁹ They audibly analyzing what the court communicated through its prose instead of simply regurgitating the court’s text in the order presented.²⁰ Christensen noted that the students in her high-performing study spent much more time critically analyzing the case aloud than students had in a similar study conducted by Dorothy Deegan.²¹ Christensen attributed this discrepancy in part to the direction she gave the students before they began to read the case for the study,²² which was to pretend to be an attorney reading a case to prepare for a meeting with a client.²³ Additionally, unlike Deegan, Christensen had the

¹⁶ *Id.* at 617 (citing Suzanne E. Wade et al., *An Analysis of Spontaneous Study Strategies*, 25 *READING RES. Q.* 147, 150 (1990) and Laurel C. Oates, *Beating the Odds: Reading Strategies of Law Students Admitted Through Alternative Admissions Programs*, 83 *IOWA L. REV.* 139, 140 (1997)).

¹⁷ *Id.* at 616-18.

¹⁸ *Id.* at 625-27. Christensen found that “high-performing” students spent about 50% more of their reading time thinking about “problematizing strategies” or “rhetorical strategies” than the “lower-performing students.” *Id.* at 625. “Problematizing strategies” were defined as thought processes in which “participants actively engaged in the text and responded to the text by ‘drawing a tentative conclusion,’ ‘hypothesizing,’ ‘planning,’ ‘synthesizing,’ or ‘predicting.’” *Id.* at 622. She defined “rhetorical strategies” as thought processes that evidenced “readers examined the text in an ‘evaluative’ way or when readers moved outside the text ‘into the realm of . . . personal knowledge’” through “‘evaluating,’ ‘connecting with prior knowledge or experience,’ and ‘contextualizing.’” *Id.* at 623.

¹⁹ *Id.* at 625-26.

²⁰ *Id.* at 623-26.

²¹ *Id.* at 627-28.

²² *Id.* at 628.

²³ *Id.* at 619.

students read a case rather than a law review article.²⁴ Christensen hypothesized that the students felt the need to do more critical thinking when reading the case because of the prompt they had been given – to prepare for a client meeting.²⁵

Although “talking aloud” gives researchers greater insight into the thought processes of subjects as they read prose, there are some limitations to this mode of data collection. Subjects may not be aware of some of their more automatic or subconscious thought processes and thus may not be able to verbalize them to researchers, and there may be some thoughts processes that are not easily described verbally, so subjects simply may not express them.²⁶ Additionally, psychology scholars have pointed to a host of non-conscious mental processes that occur, which would not be captured by this type of data collection.²⁷ These scholars also point out that there are “reactionary” issues inherent in having subjects describe their own mental processes.²⁸ These issues can involve a disruption to the mental processes, because readers are forced to verbalize what they are thinking, which might actually change what they are thinking.²⁹ Additionally, subjects may feel social pressure to present their thoughts in a certain way to researchers.³⁰ For instance, verbal reports collected from

²⁴ *Id.* at 628 n.156.

²⁵ *Id.* at 628.

²⁶ Many scholars have criticized the “think aloud” method of data collection for these limitations, but the original developers of the “think aloud” method were not shy about acknowledging these limitations in their seminal book on the subject. K. ANDERS ERICSSON & HERBERT A. SIMON, *PROTOCOL ANALYSIS: VERBAL REPORTS AS DATA* 1-30 (MIT Press rev. ed. 1993) (noting that unconscious or automatic actions may be hard for a subject to verbalize).

²⁷ Timothy D. Wilson, *The Proper Protocol: Validity and Completeness of Verbal Reports*, *PSYCHOLOGICAL SCIENCE* 249, 249-50 (1994) (book review) (“In the years since the first edition of *Protocol Analysis* was published, however, it has become increasingly clear that a great deal of information processing occurs outside of awareness . . . One can hardly open a current cognitive or social psychological journal without finding new evidence for nonconscious processing, in its various guises of automaticity, implicit learning, implicit memory, on-line processing, priming, and intuition.”).

²⁸ *Id.* at 250 (“I have found, for example, that asking people to think about the reasons for their attitudes changes these attitudes; people focus their attention on attributes of the attitude object that are easy to put into words and accessible in memory, and their attention is drawn away from information that is difficult to put into words or inaccessible in memory.” (citing studies Wilson conducted with S.D. Hodges in 1992, S.D. Hodges and S.J. LaFleur in 1994, D. Kraft in 1993 and J.W. Schooler in 1991)).

²⁹ *Id.*

³⁰ *Id.*

subjects on how much alcohol they drink, often show a discrepancy when compared to the actual sale of alcohol.³¹ Unfortunately, this criticism may undercut some of the previous studies conducted on law students. Christensen's study, in particular, gave students a "purpose" to their reading – the students were asked to pretend to be an attorney preparing for a meeting. To impress the researcher, some students may have been better at verbalizing the advanced critical reading skills expected of attorneys preparing for a meeting, even though other students may have the same exact mental processes as someone not as affected by the social pressure of wanting to impress the researcher. Thus, perhaps, those students who conform to the social pressure are the ones who perform better in law school. This may be the causal effect of the correlation Christensen noted in her study.³²

To correct, or buffer against, any possible error in the "think aloud" protocol for collecting data, scholars suggest having an independent means of assessment that does not rely on subjects' self-analysis of their mental processes.³³ There is one reported empirical study that collected data about legal readers through a testing method that did not require subjects to self-report data.³⁴ In 2008, researchers working for the Law School Admissions Counsel (LSAC) gave a case reading comprehension test to second and third semester students and then repeated a similar comprehension test with 1Ls and 3Ls and found no significant difference in comprehension or

³¹ See, e.g., R.E. Popham and W. Schmidt, *Words and Deeds: the Validity of Self-Report Data on Alcohol Consumption*, 42 J. OF STUD. ON ALCOHOL 355, 355-58 (1981); Lorraine T. Midanik & Alice M. Hines, *'Unstandard' ways of answering standard questions: protocol analysis in alcohol survey research**, 27 DRUG AND ALCOHOL DEPENDENCE 245, 245-52 (1991).

³² Suzanne E. Wade et al., *An Analysis of Spontaneous Study Strategies*, 25 READING RES. Q. 147, 150 (1990) (citing studies the authors felt supported the notion that the inherent weaknesses in the "think aloud" protocol would not change the results of the cognitive processes that were studied) Additionally, other studies have found some weak effects on cognitive processes that may aggregate to affect accurate data collection. K. Anders Ericsson & Herbert A. Simon, *Verbal Reports as Data*, 87 PSYCHOL. REV. 215 (1980).

³³ Christensen, *supra* note at 608.

³⁴ Dorothy H. Evensen, James F. Stratman, Lauren C. Oates & Sarah Zappe, *Developing an Assessment of First-Year Law Students' Critical Case Reading and Reasoning Ability: Phase 2*, LSAC Res. Rep. SERIES, (2008), [https://www.lsac.org/docs/default-source/research-\(lsac-resources\)/gr-08-02.pdf](https://www.lsac.org/docs/default-source/research-(lsac-resources)/gr-08-02.pdf). This study also used "think aloud" protocols on a subset of subjects to determine the reasoning behind the students picking particular answers on the test. *Id.*

reasoning between the two groups in each study.³⁵ To better understand the outcomes of this study and to verify the validity of the multiple choice test questions, the researchers also used “think aloud” protocols on a subset of subjects with a portion of the test questions to determine the reasoning behind the students picking particular answers on the test.³⁶ The researchers indicated that they used this protocol to verify and refine the theories they developed on why students did not improve their reading skills over their law school career.³⁷ Essentially, the researchers argued that a lack of change between the reading skills between these two groups of students should not be looked at pessimistically as evidence that these skills cannot be changed.³⁸ Instead, the researchers argued that law professors should see an opportunity in this results to begin to develop more concrete assessment methods for case reading skills and explicit instruction on how to better those skills.³⁹ Additionally, the authors make this call for further research: “Our last research goal . . . we see as the most difficult, but significantly facilitated if linked with collaborative efforts among law schools: that is, to trace the development of these skills through law school and perhaps into professional practice. We see this as a huge challenge, but one worth pursuing.”⁴⁰

On their face, the results of the LSAC study appear to indicate there is no difference in how first-year law students and third-year law students read. However, I wondered if that was the case or if the process of reading the amount of cases that law students read may, in fact, make students read in a way that may be counter-productive to improving their reading skills. I began a search for a method of reviewing reading habits that would do so without relying on the self-description of

³⁵ *Id.*

³⁶ *Id.* at 21.

³⁷ *Id.* at 39-40.

³⁸ *Id.*

³⁹ *Id.* The authors note that the detailed discussion they offered for the design of their multiple-choice test might be a good starting point for this type of assessment. *Id.* at 40.

⁴⁰ *Id.* at 40.

subjects and would give a more accurate view of the mechanics of reading to see if there were any discernable differences between novice legal readers and legal readers who had progressed some through law school. My attention quickly turned to a technology that has been used to assess reading from a more mechanical approach than the current researchers of legal readers have used – eye-tracking technology.

Assessing reading by monitoring the movement of the eyes, commonly referred to as “eye-tracking,” is not a new concept.⁴¹ In the late 1800s, a French ophthalmologist by the name of Louis Émile Javal noticed that the eyes do not move in a continuous motion when reading.⁴² Instead, the eyes pause every few words, before moving on to reading another set of words.⁴³ The smooth movements in-between pauses later became known as “hops” or “saccades” and the pauses as “fixations.”⁴⁴ Since Javal’s observations, there have been a host of attempts to measure the “saccades” and “fixations” of readers to determine what these mechanical processes mean for the cognitive processes of reading.⁴⁵ In the early 1900s, efforts began to build devices that could measure the “hops” and “fixations.”⁴⁶ These measuring devices evolved from intrusive instruments that required users to wear contact lenses attached to aluminum pointers, into a less intrusive tool that used film and light refracting off of pupils to assess eye movement.⁴⁷ The uses for eye-tracking technology to gauge the effectiveness of advertisements and websites, and the potential for

⁴¹ See David Leggett, *A Brief History of Eye-Tracking*, UX BOOTH (Jan. 19, 2010), <http://www.uxbooth.com/articles/a-brief-history-of-eye-tracking>.

⁴² *Id.*

⁴³ *Id.*

⁴⁴ *Id.*; see also AGA BOJKO, *EYE TRACKING THE USER EXPERIENCE: A PRACTICAL GUIDE TO RESEARCH* 12 (Marta Justak et al. eds., 2013).

⁴⁵ See Leggett, *supra* note 40.

⁴⁶ *Id.*

⁴⁷ Yannjy Yang & Chih-Chien Wang, *Trend of Using Eye Tracking Technology in Business Research*, 3 J. ECON. BUS. & MGMT., 447, 447 (2015).

eye-tracking technology in the world of online gaming drove innovation.⁴⁸ Eye-tracking technology is now relatively unobtrusive; often times resembling nothing more than a small set-top box that can read eye movement from several feet away, and much more affordable than it has been in previous iterations.⁴⁹

Despite the long history of eye-tracking technology's ability to assess reading patterns, I discovered this technology had never been applied to legal readers in any reported study. Yet this technology seemed like a missing link that would allow me to verify the findings of the "think aloud" studies through data that is collected without the subject self-reporting their thoughts. Additionally, this technology seemed to hold the promise of being able to identify some of the mechanical differences in reading patterns that "think aloud" protocols cannot effectively capture.

The Study

This study was conducted over the 2015-2016 academic year. My research assistants and I were able to set up the technology at a few of the campus's incoming-student orientation events, which allowed us to capture volunteers who had not yet started law school. The study continued throughout the academic year and captured experienced legal readers at various points in their law school careers. Both full-time and part-time students volunteered for the study, and some of the participants were LLM students as well as traditional JD students.

Before the study began, subjects were asked to fill out an informational questionnaire. Subjects were asked which of the twelve degree programs offered by the law school they were

⁴⁸ Tom Simonite, *CES 2014: Eye-Tracking Game Controller for PCs Launching This Summer*, MIT TECH. REV. (Jan. 6, 2014), <https://www.technologyreview.com/s/523281/ces-2014-eye-tracking-game-controller-for-pcs-launching-this-summer/>.

⁴⁹ Yang & Wang, *supra* note 45.

enrolled in.⁵⁰ Subjects were also asked how many credit hours they had completed towards those degrees and whether they had read cases on their own prior to attending law school. Subjects were also asked if English was the language they were most proficient in.

There are many manufacturers of eye-tracking technology, but I chose to use a technology suite produced by Gazepoint Research, Inc. The hardware components for the Gazepoint technology purchased for this study were housed in a small rectangular box that could be placed a few feet in front of a subject reading on a computer screen. Inside the box was a camera that read the reflection from the pupil of an infrared light that was also housed in the box. The hardware was bundled with a software package that assisted test administrators for this study in calibrating the system for each new subject's eyes.

Once the system was calibrated for each new subject, the subjects were then asked to read through *People v. Kay*.⁵¹ I chose the *People v. Kay* opinion because it is a six-page, one thousand five hundred and four -word opinion, that could be read in under ten minutes. Additionally, the case was on an obscure topic in a foreign jurisdiction – whether a dog can be considered a “dangerous weapon” under the Michigan Assault with a Dangerous Weapon statute –⁵² and my hope was that no subject would be familiar with the case and have a reading advantage over other subjects. Finally, the *People v. Kay* case came in at a readability score that is typical for a judicial opinion.⁵³

⁵⁰ These are a full-time JD program; a part-time JD program; LLM programs in Advocacy, Elder Law, and International Law program; and accelerated degrees that allow students to obtain their JD degrees at the same time as various Bachelor's degrees, LLM degrees, and Masters degrees.

⁵¹ 121 Mich. App. 438 (Mich. App. 1982).

⁵² *Id.* at 439.

⁵³ In their article, *Does the Readability of Your Brief Affect Your Chance of Winning an Appeal? An Analysis of Readability in Appellate Briefs and Its Correlation with Success on Appeal*, Lance Long and William Christensen found that a random selection of the majority opinions from ninety state supreme court cases had an average Flesch Reading Ease score (the test Microsoft Word uses to gauge the “readability” of selected text) of 34.04 and ranged from a Flesch Reading Ease score of below twenty to one over forty. The *People v. Kay* opinion had a 36.6 Flesch Reading Ease score, making it squarely in the middle of that range and close to the average of that range. See Lance Long & William Christensen, *Does the*

At the completion of the study, students were asked to answer three multiple choice questions designed to ensure they fully read and processed the *People v. Kay* opinion. The first question quizzed the subjects on the issue in the case.⁵⁴ The second required the subjects to identify the defendant's argument in the case.⁵⁵ The final question asked subjects to identify the holding of the case.⁵⁶

After the study was complete, the computer software produced a real-time recording of the subjects' eyes and graphical displays of the portions of the text they read as they progressed through their assigned document, along with a comprehensive spreadsheet that logged the time, location, and duration of every eye fixation the subject exhibited during the study.

Results

A total of sixty-three students participated in the study. There were technology errors with seventeen of the studies, which had to be discarded from the overall data tabulation because the data was not readable or accurate. These issues were likely due to insufficient lighting conditions or

Readability of Your Brief Affect Your Chance of Winning an Appeal? An Analysis of Readability in Appellate Briefs and Its Correlation with Success on Appeal, 12 J. APP. PRAC. & PROCESS 145, 158 (2011).

⁵⁴ This question read:

What was the issue the court had to decide in this case?

- A. Whether a dog is a "dangerous weapon" within the Washington assault with a deadly weapon statute.
- B. Whether assaulting a person with a 'dangerous weapon' is assault with a deadly weapon.
- C. Whether a dog is a 'dangerous weapon' within a Michigan statute.

The correct answer was "C."

⁵⁵ This question read:

What was the defendant arguing in this case?

- A. A dog is not a 'dangerous weapon' under the terms of the statute.
- B. A dog is an inanimate object.
- C. The lower court was correct in finding that the dog could not be considered a 'dangerous weapon.'

The correct answer was "A."

⁵⁶ This question read:

What was the court's decision in this case?

- A. Errors regarding jury instructions cannot be reviewed by appellate courts.
- B. A dog can be a 'dangerous weapon' within the terms of the statute.
- C. The statute in question precludes animate objects.

The correct answer was "B."

computer processing speeds for the eye-tracking technology to be able to effectively capture eye movements.⁵⁷ Additionally, there were twelve studies with unusable data due to test administration errors.⁵⁸ Once these tests were removed, there were thirty-four viable tests used to conduct the results of this study.

Of the thirty-four viable tests, fourteen were from students who had yet to begin even the first day of law school.⁵⁹ The other twenty participants had begun law school at the time they participated in the study. Of these twenty participants, three were in their first year of law school, eight were in their second year of law school and nine were in their third year of law school. The subjects who had not begun law school – the “novice reader” group – took an average of four minutes and forty-three seconds to read the opinion. The subjects in their first year of law school took an average of two minutes forty-eight seconds to read the opinion, and the second-year law students spent an average of three minutes five seconds reading the opinion. Finally, the third-year law students averaged three minutes thirty-one seconds.

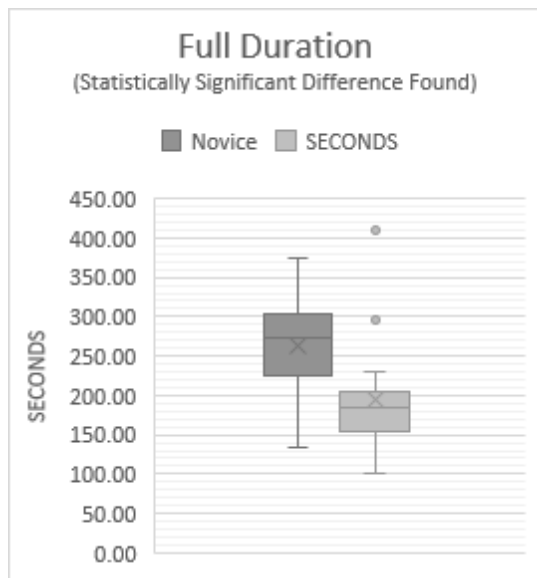
When the times for the 1Ls, 2Ls, and 3Ls were aggregated into a group labeled “experienced readers,” their combined average reading time was three minutes fourteen seconds. Aggregating these three classes allowed for statistical comparisons to be made that could not be made on the low numbers of subjects enrolled in an individual year of study. To determine if the difference in average time between the “novice” and “experienced” readers was significant, the widely accepted two-tailed

⁵⁷ In eleven instances, the system couldn't detect the movement of the subject's eyes due to an overabundance of light, which confused the system's sensors, or the system read the reflection off the subject's glasses as a retinal movement. Additionally, the computer did not process the results of six studies due to insufficient computer processing speeds.

⁵⁸ Unfortunately, the informational sheets for five subjects were lost or incomplete, making the results from those studies unusable. Additionally, seven subjects physically moved out of the field of vision of the set-top box so the technology could not track their eye movement.

⁵⁹ These students were on campus to participate in orientation for incoming students.

t-test for independent means was used.⁶⁰ This test determines the probability or “p-value,” by comparing the data from each group, and if that p-value is less than .05, then there is a greater than 95% chance that the difference is not due to chance alone, and is, therefore, generally considered “statistically significant.”⁶¹ When comparing the reading times of the novice and experienced readers, the t-test returned a p-value of .003684, which means that there was a 99.7% chance the difference was not due to chance along. Under generally accepted principles of statistics, this probability is “statistically significant.”



⁶⁰ See FRANS L. LEEUW WITH HANS SCHMEETS, *EMPIRICAL LEGAL RESEARCH: A GUIDANCE BOOK FOR LAWYERS, LEGISLATORS, AND REGULATORS* 168-70 (2016).

⁶¹ See *id.* at 169; Jean-Baptist du Prel, et al., *Confidence Interval or P-Value? Part 4 of a Series on Evaluation of Scientific Publications*, *DEUTSCHES ARZTEBLATT INTERNATIONAL* 335-338 (2009), <https://www.aerzteblatt.de/int/archive/article?id=64639>. Admittedly, there are criticisms of using p-values to answer research questions, especially when trying to divine causation for an observed difference between groups of similar groups; Regina Nuzzo, *Scientific method: Statistical errors*, *NATURE* (Feb. 12, 2014), <https://www.nature.com/news/scientific-method-statistical-errors-1.14700>. However, because the analysis of this study only conjectures at the reasons for any noted statistically significant differences, and intends to be “part of a fluid, non-numerical process that blended data and background knowledge to lead to scientific conclusions,” the p-value appears to be an appropriate measure of the data collected in this study. See *id.*

Two-tailed t-tests for two independent means were run on both groups to determine if there was a statistically significant difference in the backgrounds of the subjects in each group that might account for the differences in the time it took to read the case. The p-value did not indicate a statistically significant difference in correct answers to the comprehension questions between the two groups of students.⁶² Notably, this finding is in-line with the finding of the 2008 LSAC study that found no comprehension or reasoning and analytical differences between second and third semester law students or 1Ls and 3Ls.⁶³ Additionally, there was no statistically significant difference found between the average response to the amount of cases read before law school between the novice and experienced readers.⁶⁴ There was also not a statistically significant difference found in the responses to the question whether English was the most proficient language for each subject as only one subject answered that in the negative.⁶⁵ Finally, a statistically significant difference was not found between the groups in the amount of students enrolled in various programs within the school's curricular offerings.⁶⁶ Because there was no difference in the backgrounds of the two groups,

⁶² Three out of the fourteen novice legal readers incorrectly answered at least one question, and 8 out of the twenty experienced legal readers incorrectly answered at least one question. A two-tailed t-test for two independent means was conducted on the number of incorrect questions for each of these groups, and the p-value was .77614, which means the difference was not statistically significant between the two groups.

⁶³ Evensen et al., *supra* note 33.

⁶⁴ Subjects were given five optional responses to the question on the initial questionnaire about the number of cases the subjects had read prior to beginning their law school studies: "a. No experience," "b. Fewer than ten cases," "c. Between eleven and fifty cases," "d. Between fifty-one and one hundred cases" and "e. More than one hundred cases." Six subjects in the novice group chose "b.," four chose "c.," and one chose "e." Eleven in the experienced group chose "b.," Three chose "c.," one chose "c.," and two chose "d." Three subjects in the novice group chose "a.," six chose "b.," four chose "c.," and one chose "e." Three subjects in the experienced group chose "a.," eleven chose "b.," three chose "c.," one chose "d.," and two chose "e." When a two-tailed t-test for two independent means was run on the responses to this question for the novice group and the experienced group, the test returned a p-value of .770196, which means there was not a statically significant difference in the amount of cases each group had read before beginning law school.

⁶⁵ Just for thoroughness, a t-test was run on this question and the p-value was -.83, so there was no statistical significance between the two groups based on the one negative answer to this question in the experienced group.

⁶⁶ The bulk of respondents were enrolled in a traditional full-time program. The novice group contained three students enrolled in the part-time JD program and two students enrolled in the part-time JD/MBA program. The experienced group contained two students enrolled in a JD/LLM program. A t-test that compared the number of subjects not enrolled in the traditional full-time program in the novice group to the number of subjects not enrolled in a traditional

presumably, any statistical differences are due to changes that occur after the subjects started law school.

Because the time it took to read the opinion appeared to be the only difference between the two groups of readers, the data was then processed to determine where that time was spent. The opinion was broken up into twenty sections based on the information described in each section. The following chart identifies each section and describes the information the court described in each section:

1.	Heading	The caption of the case and the description of the attorneys and law firms.
2.	Overview	The first paragraph of the case that gives an overview of the legal question the court was analyzing.
3.	Facts	A recitation of the underlying facts of the case.
4.	Procedure	A detailed procedural history of the case.
5.	Statute	A quotation of the assault with a dangerous weapon statute.
6.	Defense arg	A paragraph that sets forth the defense’s argument.
7.	FN1	A footnote in the defense argument that string cites several cases that found inanimate objects to be “dangerous weapons.” If a subject read this footnote at the point the footnote number appeared in the main text of the article, that time was logged as time for this section.

full-time program in the experienced group. The result was a p-value of 0.071495, which is above the .05 threshold that is generally accepted as statistically significant.

8.	ALR quote ⁶⁷	A long quotation from an ALR article about the use of dogs as “dangerous weapons”
9.	<i>Tarrant</i>	An explanation of a precedent case – <i>Commonwealth v. Tarrant</i>
10.	FN2	A footnote in the <i>Tarrant</i> explanation that includes a definition of a term from the Michigan criminal code. If a subject read this footnote at the point the footnote number appeared in the main text of the article, that time was logged as time for this section.
11.	<i>JR</i>	An explanation of a precedent case – <i>State In the Interest of J.R.</i>
12.	<i>JR</i> quote	A quotation out of the <i>JR</i> opinion
13.	<i>Torrez</i>	An explanation of a precedent case – <i>People v. Torrez</i>
14.	<i>Torrez</i> quote	A quotation out of the <i>Torrez</i> opinion
15.	<i>Torrez</i> 2	More explanation of the <i>Torrez</i> case
16.	Analysis	The court’s analysis of the precedent cases, which does explain another case – <i>People v. Goolsby</i> – in the middle of the analysis.
17.	<i>Goolsby</i> quote	A quote from the <i>Goolsby</i> case.
18.	Analysis 2	Additional analysis of the precedent cases and how they apply to the facts before the court
19.	End	Disposes of an issue that was not preserved for appeal and states the lower court was affirmed.

⁶⁷ Large quotes were separated out from the text to determine if there was a difference in how novice and experienced readers processed quotes.

20.	FN End	The two footnotes were physically located at the end of the opinion, so subjects who read the footnotes after reading the “End” section, had that time logged under this section.
-----	--------	---

The time each subject spent on each of these nineteen sections was then logged. The time the novice legal readers spent on each section were compared against the time the experienced legal readers spent on each section. Statistically significant differences were not found for the following sections: Heading,⁶⁸ Facts,⁶⁹ Procedure,⁷⁰ Statute,⁷¹ ALR quote,⁷² Analysis,⁷³ Analysis 2,⁷⁴ or End.⁷⁵ Statistically significant differences were found for the Overview,⁷⁶ Defense arg,⁷⁷ *Tarrant*,⁷⁸ *JR*,⁷⁹ *JR* quote,⁸⁰ *Torrez*,⁸¹ *Torrez* quote,⁸² *Torrez* 2,⁸³ and *Goolsby* quote.⁸⁴ The text of the two footnotes were physically located at the end of the opinion and only one subject – an experienced legal reader – jumped to the end of the opinion to read a footnote at the point the footnote number appeared in the main text. Only three novice readers and two experienced readers read through the footnotes at the end of the opinion. A two-tailed t-test for independent means did not calculate a p-value that

⁶⁸ The two-tailed t-test for independent means was used to test the statistical significance of the difference between the reading time for the novice group and the experienced group for each section of the opinion. For the Heading section, the returned p-value from the t-test was .070679.

⁶⁹ The t-test returned a p-value of .086822 for the Facts section.

⁷⁰ The t-test returned a p-value of .370564 for the Procedure section.

⁷¹ The t-test returned a p-value of .132981 for the Statute section.

⁷² The t-test returned a p-value of .052596 for the ALR quote section.

⁷³ The t-test returned a p-value of .437772 for the Analysis section.

⁷⁴ The t-test returned a p-value of .577386 for the Analysis 2 section.

⁷⁵ The t-test returned a p-value of .252929 for the End section.

⁷⁶ The t-test returned a p-value of .003351 for the Overview section.

⁷⁷ The t-test returned a p-value of .024132 for the Defense arg section.

⁷⁸ The t-test returned a p-value of .017287 for the *Tarrant* section.

⁷⁹ The t-test returned a p-value of .031384 for the *JR* section.

⁸⁰ The t-test returned a p-value of .006151 for the *JR* quote section.

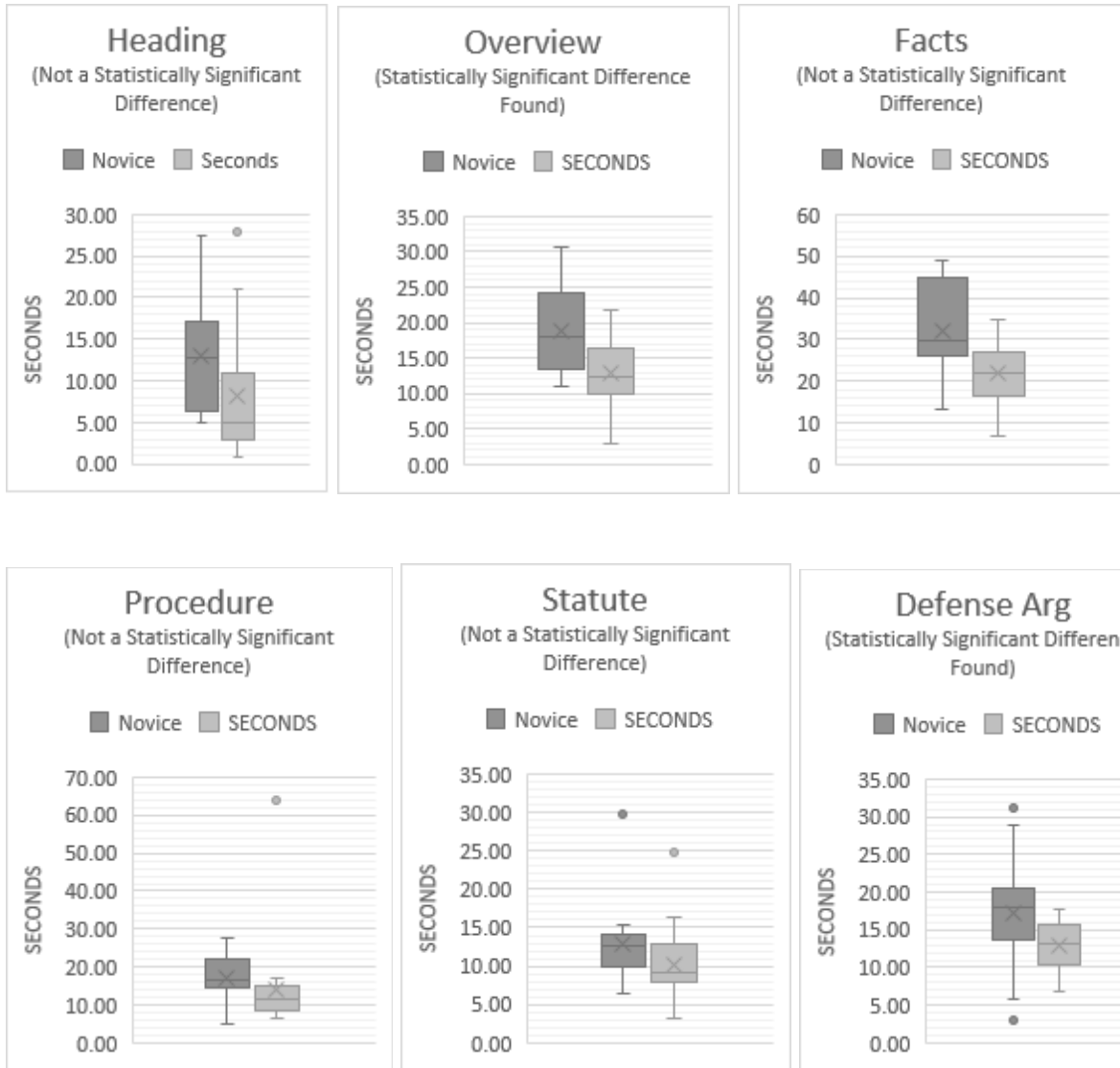
⁸¹ The t-test returned a p-value of .013085 for the *Torrez* section.

⁸² The t-test returned a p-value of .004456 for the *Torrez* quote section.

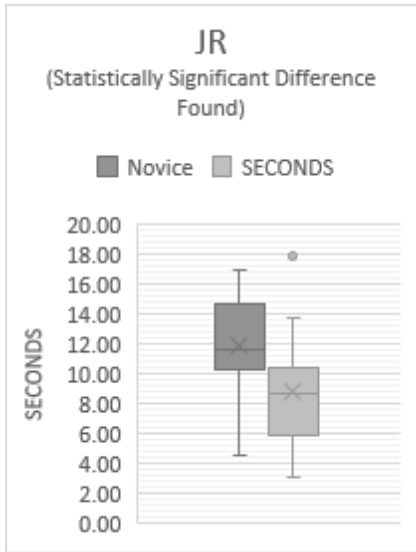
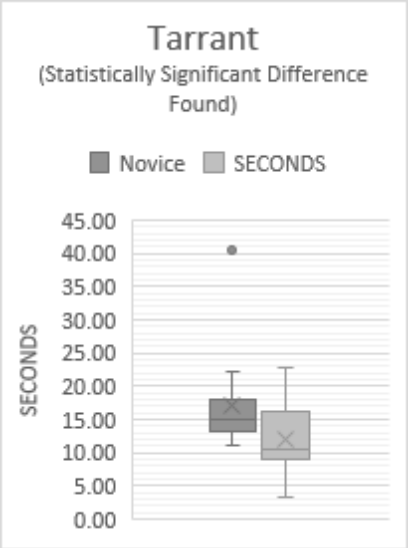
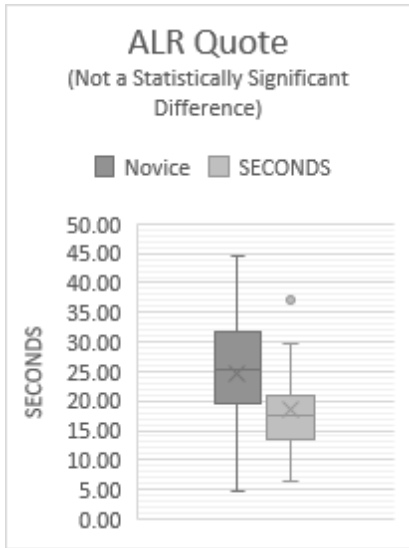
⁸³ The t-test returned a p-value of .011248 for the *Torrez* 2 section.

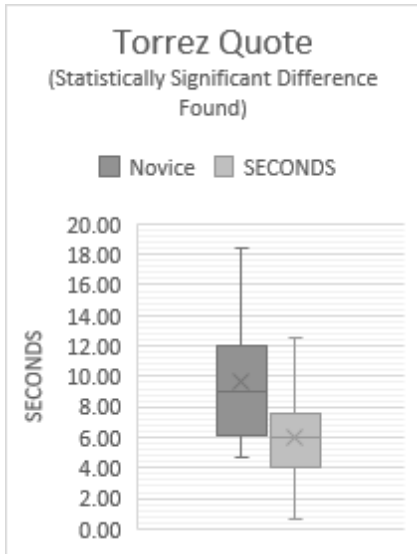
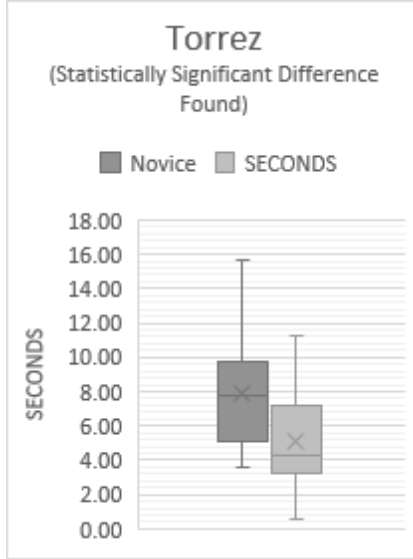
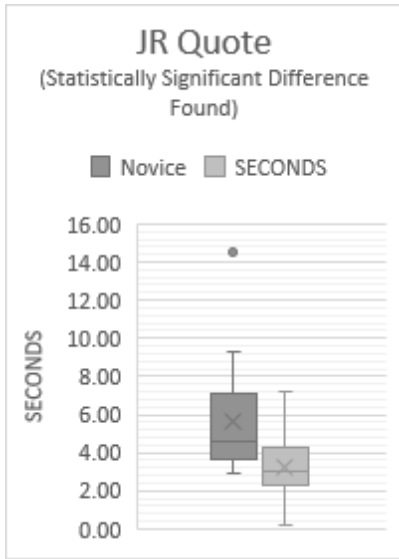
⁸⁴ The t-test returned a p-value of .005996 for the *Goolsby* quote section.

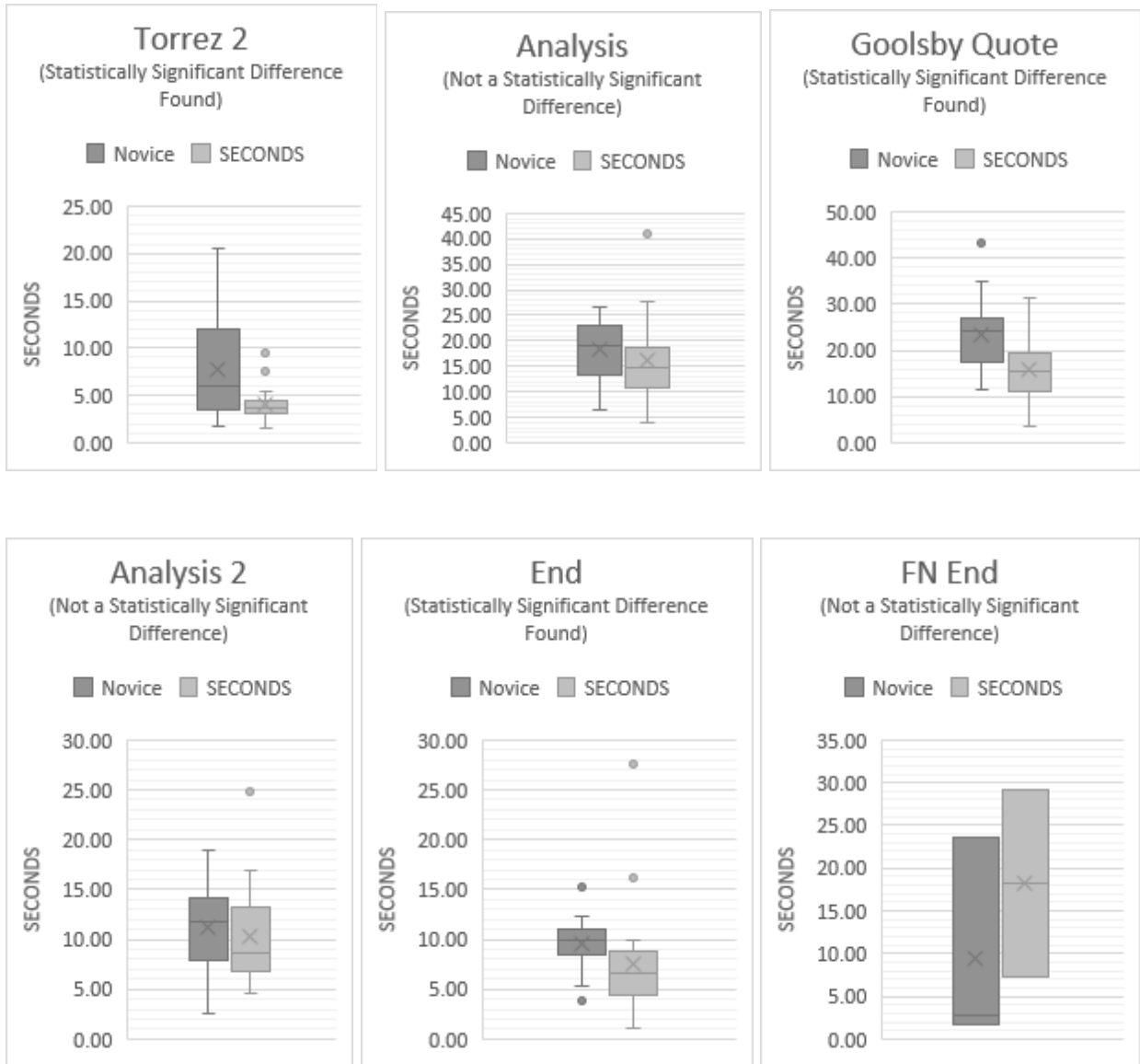
indicated a statistical difference between the amount of time the few students who did look at footnotes spent reading through them.⁸⁵ Not surprisingly, all of these differences indicated that the experienced legal readers read at a significantly faster rate than the novice legal readers. Therefore, these sections account for the significantly different overall time spent on reading the opinion.



⁸⁵ The t-test returned a p-value of .935554 for the average time it took each group to read the footnotes, whether the footnotes were read in-line with the text or at the end of the opinion.







Analysis

In looking through the sections that experienced legal readers read at speeds that were faster than novice legal readers in a statistically significant degree, it is noticeable that these are all sections that involve simple background explanations of law or facts that are not as critical to understand in a nuanced fashion to get what a law student needs from a case – namely, an understanding of the rule of law that the court developed in its holding. Experienced readers read through recitations of case

descriptions by the court quicker than novice legal readers at a level of statistical significance (the *Tarrant, JR, JR* quote, *Torrez, Torrez* quote, *Torrez 2*, and *Goolsby* quote sections all explained precedent case law). The other two sections that experienced legal readers read at statistically-significant faster pace than novice legal readers were the overview section and the section that detailed the argument the defense made. Although it was beyond the confines of this study to determine the reasons for these differences, it certainly makes intuitive sense that a student who is focused on determining what the court held in a case, would find these sections not to be helpful to meet that end goal because the overview will be reiterated later in the opinion and the defense argument may very well be completely rejected by the court.

Unlike the sections that the experienced readers read at a rate that was faster than novice readers to a statistically significant degree, the sections that experienced readers read at a pace that was not a statistically significant difference from the pace of novice readers do appear to be those sections that law students perceive as being critical to their understanding of the holding. The Heading, Facts, Procedure, and Statute sections all detail the core facts of the case. The Heading detailed the parties and the level and type of court deciding the case, the Facts section explained all of the legally relevant facts, the Procedure section set forth the relevant procedural history, and the Statute section contained a quote of the core statute being analyzed in the case. Without understanding these underlying facts in detail, a legal reader would have difficulty grasping the full import of the rule the court develops. Additionally, the experienced readers read the two sections that contained the core of the court's analysis of how the precedent case law applied to the facts – Analysis and Analysis 2 – at a pace that was not a statistically significant difference from the pace of the novice readers. Because these sections are the heart of the explanation of the court's rule, the

similar reading pattern between experienced readers and novice readers fits the theory that the experienced readers do not read faster while reading information critical to understanding the rule from the case. At first blush, it may appear that the ALR category and the End category do not fit with the working theory that experienced readers skim material not critical to their understanding of the rule from the case. However, the ALR quote section detailed an analysis by an ALR editor of how courts across the country have analyzed similar statutes, which experienced readers may have seen as important to their understanding of the eventual holding. Additionally, the p-value of the comparison was 0.052596, which is extremely close to the “statistically significant” threshold of 0.05. In essence, this means that many experienced readers were breezing through the ALR quote section as they did the other “background” sections where a 0.05 significance level was found. Finally, the “End” section was found to have been covered at the same rate by both groups. This could be simply because the section was comprised of the last three sentences in the entire opinion. So even if the experienced group had sped through this section, perhaps the novice group did as well because they wanted to be done with reading. Or perhaps the experienced group did not choose to skim through this section because it did include a recitation of the ultimate holding.

Another piece of data that seemed to be consistent among all readers in this study is the inattention given to footnotes. Only one subject looked at the footnote in-line with the text. The combined time it took the subject to read the footnotes was 24.52 seconds. This may have been affected by the fact that the footnotes were located at the end of the entire document. However, even when subjects reached the end of the document, very few took the time to read the footnotes. Only five out of thirty-four subjects read through the footnotes and two of those looked at the footnotes for under three seconds. The other two subjects took twenty-three seconds to read

through the footnotes. This disparity indicates that only two out of thirty-four subjects read the footnotes thoroughly at the end. Even coupled with the one subject that read the footnotes in-line with the text, this means less than 9% of readers read the footnotes. It is beyond the confines of this study to determine whether this inattention to footnotes is typical among all experienced legal readers – even those who are reading for purposes other than learning the law in law school – but if this low level of attention is typical, this finding certainly calls into question whether footnotes are useful at communicating information to a reader.

These results may help to explain the similarities in comprehension and analysis of cases between novice and experienced readers in this study as well as the LSAC study. If experienced legal readers who are reading simply for the rule of a case skim over “non-essential” portions of the case, they may be missing portions of the case that would allow them to conduct more significant analysis of the case, yet they would still maintain the basic comprehension that novice readers can also obtain when reading carefully through the case. If the experienced readers are becoming “lazier” readers as they progress through law school, this is certainly something that can be remedied through adequate assessments that reward careful reading, as the LSAC study authors suggested. Further research is needed to test these theories.

Additionally, the results of this study are simply the beginning of a verification process of the “talk aloud” research studies. Eye-tracking technology proved to be an effective tool in this study for clarifying that there is a difference in how novice and experienced law school readers read, but that the difference may be one that does not result in performance differences on basic comprehension tests. The next step is to determine if eye-tracking technology can verify and explain

the findings of the “talk aloud” studies that readers enhance their reading skills when reading for particular purposes.