STARS

2018

Passive Voice Usage in Undergraduate STEM Textbooks

Huiyuan Luo
University of Central Florida

PASSIVE VOICE USAGE IN UNDERGRADUATE STEM TEXTBOOKS

By

Huiyuan "Tia" Luo

B.A. Capital Normal University, 2011
M.A. University of Kansas, 2014

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the College of Education and Human Performance
at the University of Central Florida
Orlando, Florida

Spring Term

2018

Major Professor: Florin Mihai

ABSTRACT

The present study, a corpus-based quantitative analysis, investigated the use of passive voice in terms of percentage (percentage of total passive usage), constructions (the most commonly used passive forms), and dispersion (the verbs that tended to be associated with passive usage) in college STEM textbooks, more specifically in Science, Technology, Engineering, and Mathematics. The corpus consisted of twenty textbooks with over 1 million running words selected from the textbooks' chapters. By identifying the main verbs and passive verbs used in the textbooks, the percentage of passive voice usage was calculated. The statistical difference of passive constructions between disciplines was examined by conducting a MANOVA test, and an instrument, the Compleat Lexical Tutor was used to generate a word list of the most commonly-used past participles in passive voice.

Results showed that the overall percentage of passive voice usage is over 30%. The MANOVA test revealed a statistically significant difference of passive voice constructions across the four disciplines, i.e. Science, Technology, Engineering, and Mathematics.

Implications for teaching suggest that passive voice was an important grammar feature that needs to be taught tor the English as a second/foreign language students who major in the STEM fields. Additionally, the use of passive constructions should be taught differently in reading and writing based on the needs of the four STEM disciplines, due to the significant difference of passive forms usage in these fields.

Keywords: passive voice usage, passive constructions, past participles in passive voice, corpus linguistics, STEM, college textbooks

ACKNOWLEDGMENTS

Now looking back at my dissertation journey, my emotions were mixed, and the joy of accomplishment has now finally filled my heart over the anxiety of frustration. It was a lonely journey, yet I could not have finished this dissertation without the support of my committee members, family and friends.

First and foremost, I would like to give my sincerest gratitude to my chair Dr. Florin Mihai, for his full support, guidance, and encouragement that truly walked me through this journey. His prompt responses to all my emails and messages, and his firm faith in my capabilities of finishing my dissertation truly gave me courage to keep up my work. In addition, I would like to send my appreciation to Dr. Keith Folse, Dr. Joyce Nutta, and Dr. Haiyan Bai for their contribution and their hard work. Dr. Folse were never tired of checking in with me to make sure I was still breathing, and I was keeping up the good work. His practical advice was not only valuable for writing my dissertation but also very useful for my future professional development. Dr. Nutta also encouraged me with her absolute faith for everything that I was about to accomplish, her feedback was truly helpful in structuring my dissertation and making it flow. Last but not least, I am very grateful for Dr. Bai's constructive advice and feedback on the statistics problems I had. She was very patient, and the way she talked to me and gave me advice made me feel like I was loved by a mom. I know I will have no words to thank them enough for all they have done for me, and I would like them to know that this grateful feeling will last forever in my life.

I must also express my greatest gratitude to my mom, Li Guozhen, who always supports me with her unconditional love. Her company and the comfort food she made while I was in my

worst condition writing my dissertation truly consoled me and encouraged me to continue on doing my work.

Finally, I would like to give special thanks to my friends, who blindly supported me with encouraging words, food, and gifts. Their earnest prayers also meant a lot to me, for I know that I could not have finished my work without these prayers. I am very grateful that I have these friends that I can grow my spiritual life with throughout this journey, and together we made a testimony.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

- *BNC* – British National Corpus

- *C-SMILE* – Corpus of State University of Malang Indonesian Learners' English

- *CANCODE* – Cambridge and Nottingham Corpus of Discourse in English

- *CLT* – Compleat Lexical Tutor

- *EFL* – English as Foreign Language

- *ELL* – English Language Learner

- *ESL* – English as Second Language

- *L1 ACQUISITION* – First Language Acquisition

- *FLOB* – The Freiburg-LOB corpus of British English

- *ICNALE* – International Corpus Network of Asian Learners of English

- *IELP* – Intensive English Language Program

- *IEP* – Intensive English Program

- *L1* – First Language

- *L2* – Second Language

- *NNS* – Nonnative Speaker

- *NS* – Native Speaker

- *NYT* – New York Times corpus

- *PD* – Possessive Determiner

- *POS* – Poverty of the Stimulus

- *RCT* – Randomised Controlled Trials

- *SLA* – Second Language Acquisition

- *SPSS* – Statistical Package for the Social Sciences

- *STEM* – Science, Technology, Engineering and Mathematics

- *UWC* – University Writing Center

CHAPTER 1: INTRODUCTION

When we learn our first language (L1), grammar does not seem to be an essential aspect that merits to be taught separately from reading, writing, speaking and listening; instead, it is embedded into the entire language learning process. If we recorded the language acquisition process of our L1, we would see no evidence of anyone learning any grammatical structure before learning how to use that structure in speaking. In other words, we do not learn our native language, we acquire it (Chomsky, 1969). L1 acquisition is accomplished through a natural process where no systematic grammatical instruction takes place. Evidence can be easily found when a native speaker (NS) is asked by a language learner why a specific grammatical structure is applied in certain situation; most of the times, NSs find it difficult to explain because they do not have the explicit metalinguistic information readily available due to the fact that their L1 is acquired but not learned (Folse, 2004a; Folse, 2004b; Schoonmaker, 2015).

There are obvious facts about L1 acquisition that serve as evidence proving L1 is acquired through a natural path. In contrast, individuals often acquire their second language (L2) through systematic instruction. Grammar is considered an important component in second language acquisition (SLA) if an L2 learner hopes to achieve native-like language proficiency. Just as William Safire (1992) once said, "Only in grammar can you be more than perfect," it may not be as important as learning other language skills (e.g., reading, listening, speaking and writing), but it is emphasized in both young and adult L2 language learning.

Of all the grammatical features in English, passive voice has been one of the controversial topics among researchers, grammar book authors, English teachers, and English as second/foreign language learners. As thousands of books and websites explain, a sentence in passive voice comes from a sentence in active voice that has undergone several changes. In

active voice, the sentence has a subject + verb + direct object (SVO) word order. In passive voice, this word order is reversed, and the original direct object becomes the grammatical subject of the new sentence while the original verb changes to a form of the verb be followed by the past participle of the verb in the active sentence. The action producer (subject) in the active sentence is sometimes omitted or unmentioned in a passive sentence because it is redundant or not new information (Celce-Murcia & Larsen-Freeman, 2015; Folse, 2016).

In passive voice, the original actor is not seen as important, and the new sentence in passive voice focuses on the recipient of the action in the original active voice sentence. Due to the fact that the original actor is being downplayed or even omitted in the new passive voice sentence, it has been claimed that passive voice comes across as ambiguous on the one hand and weakens the writer's scholarly argument on the other hand (Albert 2000; Sheen, 1982; Sigel, 2009; Strunk & White 2000; Taylor 2005). Many non-specialists believe that passive voice is weaker, of less value, or just wrong (King, 2000).

Through the use of corpus linguistics studies on passive voice (Carter & McCarthy, 1999; Folse, 2015; Granger, 2013), researchers have shown that there are actually many circumstances where passive voice is favored or even required (Alvin, 2014; Biondi, 2002; Millar, Budgell & Fuller, 2012). Studies on the suitability of the use of passive voice and its frequency in scientific discourse indicate that, although passive voice use has been decreasing in recent decades, it has most certainly not been completely discarded. Rather, the frequency of passive use is relatively high in certain parts of scientific scholarly writing such as the 'Methods' and 'Results' sections (Alvin, 2014; Ding, 2002; Millar et al., 2012).

In addition, knowing the current passive voice conventions in scientific discourse is important for grammar book authors, English teachers, and L2 learners since it helps guide

authors' and teachers' focus on certain grammar features so that L2 learners are better prepared for academic study. However, the conventions of passive voice are somewhat unclear to some L2 writers – and even native writers. For example, Alvin (2014) found though an informal survey on passive voice usage among graduate students at a Singaporean university that approximately 90% of graduate students did not consider passive voice as breaking the flow or diminishing the strength of an academic article. Instead, they believed that passive voice increased objectivity of academic writing. Given that this very important English convention continues to be controversial, there is a need to investigate passive usage so that these stakeholders are to better understand this convention. Therefore, the present study aimed to provide empirical data on passive usage in terms of passive frequency (a percentage of total passive usage), constructions (the most commonly used passive forms), and dispersion (verbs that tend to be associated with passive usage).

Rationale

L2 learners often hear prescriptive grammar rules such as not to end sentences with a preposition, not to use double negatives, and not to use passive voice. Most writing guides discourage students from using passive voice in their writing, asserting that passive voice may weaken writers' statements and cause difficulties in the comprehension of writers' message.

As indicated from the four capital letters, STEM is an acronym that refers to four academic disciplines of Science, Technology, Engineering and Mathematics. According to Xue and Larson (2015), there has been a pressing need for and shortage of STEM professionals in the government sector and private industry, making the job market extremely optimistic for STEM graduates. Additionally, Close (2016) reported that college students who graduate from the STEM majors could earn the highest starting salaries compared to other fields. Thus, it is not

surprising to expect a growing trend in STEM enrollment. The National Science Board (2014) has revealed that one-third or more of first year college students intend to major in Science and Engineering fields. According to Yu, Luo, and Folse (2016) investigating passive voice usage in college introductory level textbooks, passive forms were used more frequently in math (M = 20.7%) and science (M = 16.4%) foundations then in social and historical foundation (M = 3.8%). In addition, by examining college textbooks of General Education Program, Moreb (2016) indicated that the overall percentage of passive voice usage in English composition, history, psychology, and biology was 7.06%, with more occurrence in biology than in the other subjects. Therefore, given that the numbers of international students enrolling in STEM majors has been increasing, there is a need for English instructors to work with L2 learners to meet the language demands in these specific fields.

One of the potential drawbacks of omitting passive voice in English textbooks and instructions is that L2 learners may not even notice passive constructions when they encounter them in their academic readings. The noticing hypothesis has indicated that what we are more likely to learn what we notice (Schmidt, 1990). Therefore, if students do not learn passive voice, they are unlikely to notice the passive structures, with potentially negative consequences on reading comprehension.

It is also important for students who are interested in academic publishing on the one hand and English instructors and grammar book authors on the other hand to understand and acknowledge the norms regarding the use of passive voice. Moreover, undergraduate students need to comprehend their textbooks to acquire skills and knowledge, to complete assignments, to write essays and reports, and to prepare for exams. It is therefore important for them to understand how the passive voice is constructed and how to interpret its meanings. It is also

important for English instructors and grammar book authors to be aware of the general trend regarding the use of passive voice in STEM writings so that they know how to teach this specific grammar point in grammar classes and how to organize content. Therefore, the present study focused specifically on the passive usage in undergraduate STEM textbooks, and explored what exactly the percentage was in passive construction, the difference of passive constructions usage among the four disciplines, and the most frequent verbs that tended to be associated with passive voice.

To teach reading and writing in the STEM fields, teachers need to understand the style of STEM writing and how to guide students through comprehension process when encountering irregular sentence structures in reading. For example, passive voice is used mostly in the Methods and Results section in scientific discourses (Alvin, 2014; Ding, 2002; Millar et al., 2012; Wilkinson, 1992), and many researchers have seen passive voice as a marked stylistic writing (Bazerman, 1988; Ding, 2002), meaning that passive voice is certainly conventionalized in such kind of prose. Moreover, Ding (2002) argues that passive voice in scientific writing is dominant. Despite the fact that passive voice is highly conventionalized, relatively little explanation can be found to guide the L2 writer about when or how a particular voice can be used in academic writing. Additionally, a considerable number of writing guidelines are still setting their position on excluding passive voice. For example, Sigel (2009) conducts a study to prove that by excluding passive voice in academic writing, researchers can convey a more cogent argument and demonstrate less ambiguous articles. Sheen (1982) also pins his faith on the destructiveness passive voice does in medical literature – "Passive voice is the bane of medical writing. It pervades medical literature with the haze and heaviness of stagnant air. Writers sometimes use passive voice in an attempt to sound scholarly and scientific, when actually they are perpetuating

a writing tradition that is fraught with ponderous and obscure language" (Sheen, 1982, p.22). Although some online guidelines for writing instructions tend to present passive voice as an undesirable attribute of written text and suggest switching to active alternative, there seems to have a discrepancy between what is encouraged and what is found in research studies. As such, Hinkel (2004) advocates scrutinizing the gap between theory and reality – "…it is hard to miss incongruities that emerge between the findings of analyses of tense and voice uses in actual academic prose and the presentations of these features of academic writing in instruction" (Hinkel, 2004, p.6). As a matter of fact, the frequency of passive voice in scientific journal articles is relatively high (e.g., Alvin, 2014; Atkinson, 1996; Bazerman, 1988; Ding, 2002; Rodman, 1994), which contradicts many writing guidelines that specifically advocate the avoidance of passive voice in what they consider good and effective academic writing.

Consequently, this concern raised three questions for investigation. The first question regarded the frequency of passive voice usage of present-day norms in undergraduate STEM textbooks. Even though the proportion of passive clauses in American English has dropped from 64.2% in early 20th century to 46.4% in late 20th century (Seoane, 2013), many studies reported that passive voice still accounted for a large portion of scientific writings. Alvin (2014), for example, reported that 30% of all clauses were passive clauses and of all passive clauses in the examined articles, approximately 29% of passive clauses were found in the methodology section. He concluded that it contradicted what had been claimed concerning the disadvantages passive voice presented since passive clauses were still pervasively used in the methodology section. Nevertheless, there has been no study that has investigated the proportion of passive uses in undergraduate STEM textbooks. Therefore, whether the percentage has fallen or risen in recent cases remains unknown.

While many research studies have focused on examining passive voice in academic journals, this study focused on college textbooks instead, for the primary objective L2 learning was to prepare college students to comprehend textbooks rather than to read academic journals. The second and the third question concerned specific details of these norms including typical forms and verbs that were intended to be used in passives. Alvin (2014) also indicated that although many writing guides advocated writing in active voice, they did not illustrate the extent and distribution of passive voice in recent scientific writing.

To discuss these research questions, the present study aimed to analyze academic textbooks across STEM disciplines. The study examined various subject areas of the STEM fields, providing a broader view of current trend respecting passive uses.

Research Questions

The following main research questions were investigated in this study:

1. What was the percentage of passive voice usage in current undergraduate STEM textbooks?

2. Was there a statistically significant difference of seven passive constructions among the four STEM disciplines? If so, what are the differences?

3. What were the most frequent verbs that tended to be associated with passive form in STEM textbooks?

Methods of Analysis

To address these research questions, the current study collected data from undergraduate STEM major textbooks used at a metropolitan university in Florida, providing a broad view of a general norm regarding the use of passive voice in the STEM areas. The study followed a

convenience sampling method (Creswell, 2013) to determine courses and textbooks. A random and systematic sampling (Kalton, 2017) was used to select examined texts for each book to create a corpus for the study. Twenty undergraduate STEM textbooks were analyzed to investigate total passive usage in terms of passive frequency, constructions, and dispersion. To investigate the most frequently used verbs in passive forms, an online corpus software, Compleat Lexical Tutor (Cobb, 2017) was used. For the overall quantitative analysis, a non-experimental descriptive design was utilized to answer the three research questions that focused on the frequency, constructions, and dispersion of passive voice constructions in the selected textbooks. For the second research question, a MANOVA test was used for determining whether the frequency of passive constructions differed among disciplines.

Significance of the Study

The primary purposes of research in any language and education-related discipline was to raise recognition of learners' gap in terms of aspects of language knowledge for researchers and practitioners, and to provide suggestions to improve teaching pedagogy and learning conditions. Therefore, the present study was significant for several reasons.

First, there have been many arguments presented about whether passive voice should be avoided in scientific writing. This confuses both L2 learners and English instructors who are unsure if passive voice should be learned and taught purposely. Therefore, it was important to clarify the current trend of passive usage in science and more specifically STEM areas to both teachers and learners. For English teachers and grammar book authors, it was important to have a clear understanding of what the current language trend was to guide students' writing style, as well as to facilitate students' familiarity with STEM writing style in order to improve reading comprehension.

Second, by noticing and learning this grammatical feature, students in Intensive English Programs (IEP) know what to expect from academic textbooks. Instructors in this type of program also know what to teach in order to better prepare students as they step into their future university courses. Hinkel (2004) analyzed and compared the conventionalized use of passive voice in L1 and L2 written academic texts, and she found that even after a long time of L2 learning and practice, non-native speaker (NNS) students who were at the advanced level may still have difficulty in employing passive voice into their academic writing. Swales' influential study of the structure of academic discourse specified that the uses of passive voice were highly conventionalized (Swales, 1990). Interestingly, Swales also indicated that there was a large portion of passive usage throughout academic discourses. Therefore, instead of avoiding passive voice, L2 learners should be taught when and how to use a particular voice in academic writing. Consequently, L2 learners who are majoring in STEM would be able read and write more competently to meet various rhetorical goals of different writing genres.

Moreover, it was important for grammar book authors to be aware of current grammar trends so that they would know how to organize content and how much proportion of passive voice instruction and practice should be included in those books. Accordingly, there would be a need for formative guidance, which would better reflect the reality of conventionalized passive usage.

Limitations of the Study

The present study had several limitations that needed to be addressed. First, the study used convenience and systematic sampling rather than randomized sampling. It aimed to investigate undergraduate STEM textbooks at a metropolitan university in central Florida. Thus, data was limited to the textbooks used at this university. Additionally, due to the limitation of

9

the author's capability of examining all the STEM textbooks that were utilized at the university, only twenty textbooks with selected chapters were examined (five books for each discipline). The limited numbers of examined textbooks could potentially influence the generalizability of findings to the STEM areas. Moreover, the use of convenience sampling was subject to systemic bias, which limited the generalizability of results (Fraenkel & Wallen, 2009). In spite of this potential bias, convenience and systematic sampling methods in this study were appropriate to ensure that the selected texts fulfilled the purpose of the study aiming to examine the passive usage in undergraduate STEM textbooks.

Second, there was no criteria or standard of what proportion of passive use would be considered as high compared with active voice. However, given the highest proportion studies have reported was about average 30% of the use of passive voice in late 20th century's scientific and medical discourses (Alvin, 2014; Millar et al., 2012; Seoane, 2013), the current trend of passive proportion should remain similar rate or even increase. Ding (2002) and Seoane (2013) both agreed that the choice of using passive voice was a sociocultural selection or pattern, and even though writing guides of certain scientific journals clearly specified the rule of avoiding passive voice, passive voice usage was still found to be as high as 68% in published research articles (Millar et al., 2012). In other words, passive voice was hard to avoid in scientific writing even though authors were purposely asked to avoid it.

Organization of the Study

The present study was organized as follows. Chapter One presented an introduction to the study, bringing up the importance of L2 grammar learning. It also touched upon hotly debated controversial issues of passive usage in academic writing. It then illustrated the rationale for conducting the present study. Next, the research questions and a description of the

procedures were presented. Lastly, the chapter contained the significance and limitations of the study, and definitions of key terms for the study.

Chapter Two reviewed and provided theoretical and empirical perspectives related to the research questions being explored. The chapter began with discussing the definition of passive voice, and listing common forms of passive voice. It also reviewed two different standpoints of passive voice usage researchers have argued about and how the controversy has varied. Then the chapter proceeded to review few empirical corpus linguistic studies conducted that have summarized the most common passive constructions and verbs tend to be associated with passive voice.

Chapter Three consisted of the methods and procedures used for data collection of the current study. These included descriptions of the research design, identified subjects, instruments, and data collection procedures. Data analysis procedures were to present in this chapter as well.

Statistical analysis of the collected data was presented in Chapter Four. The results and findings from the research questions were also provided in this chapter.

The last chapter discussed the study's findings in detail. It also provided a summary of key findings, and limitations of the study. Finally, the chapter offered pedagogical implications and recommendations for future studies.

Definitions of Terms for the Study

The following terms and acronyms are defined below for the sake of clarity and understanding.

- *First language (L1)* – the mother tongue acquired from birth
- *Second language (L2)* – an additional language learned after the first language

- *First language acquisition (L1 ACQUISITION)* – the process by which children acquire their L1

- *Second language acquisition (SLA)* – the process by which humans acquire additional languages

- *Native speakers (NSs)* – people who speak a language as their first language

- *Nonnative speaker (NNSs)* – people who use a language they acquired as a second or additional language

- *L2 learners* – second language learners, including English as a second language learner(s) and English as a foreign language learner(s).

- *Active voice* – a grammatical voice when the subject of a sentence is the agent of the verb

- *Passive voice* – a sentence in which the affirmative SVO structure is reversed, the object takes the place of the subject, the subject is placed either in a *by* phrase or omitted, and the main verb also conjugates to its past participle form and follows the verb *be.*

- *Grammar* – the set of systematic and structural rules of a language governing the composition of words to sentences, which includes syntax and morphology and sometimes also phonology and semantics

- *Passive constructions or passive forms* – an individual grammatical construction of the passive voice

- *Verb* – the main part of predicate that expresses an action, an occurrence, or a state of being

- *Possessive determiners* – function words used before nouns to express possession

- *Corpus* – a collection of spoken or written texts (or both) of a particular author or a body of writing on a particular subject

- *Poverty of the Stimulus* – an affirmation that the knowledge of natural language grammar is fully achieved by innate linguistic capability

CHAPTER 2: LITERATURE REVIEW

Unlike L2 grammar learning, grammar in first language acquisition (L1 acquisition) is acquired in a naturalistic way. Research studies also support this standpoint by revealing four facts in L1 acquisition. First, instructive teaching does not really seem to play a major role in L1 acquisition. For instance, an empirical study showed that a mother tried to teach her child the correct sentence structure, but after ten times of attempt of syntactical instruction, the child still consciously uttered an ungrammatical sentence – "Nobody don't likes me." (McNeil, 1966). Despite children making grammatical errors at early age, they end up producing ungrammatical sentences at some point of an early stage. Consequently, children acquire their L1 effortlessly, without any systematic instruction.

Second, children are constantly surrounded by adults' speech, which serves as linguistic input. However, when carefully examining the nature of adults' speech, people have noticed that their utterances may sometimes contain speech errors and ungrammatical sentences. Additionally, even though children are exposed to those speech errors, adults appear to keep silent about why the misspoken expression is ungrammatical to children. Interestingly, children also end up being able to distinguish grammatical sentences and ungrammatical ones. Third, even if children are exposed to only a finite number of sentences, they come to acquire the ability to generate an infinite number of sentences based on limited amount of input. Lastly, as learners of a language, children often present non-adult-like linguistic performances by uttering sentences, which an adult speaker of a language would not produce. Nevertheless, studies indicate that children's "errors" do not seem to be simply messing up, but the pattern of their "errors" appears to be selective or limited. Stromswold (1990) found that although children are known to overgeneralize the past morphological attachment –*ed* to the wrong verbs (e.g. do –

doed, have – haved), they do not overgeneralize the *–ed* attachment to *do* and *have* when they function as auxiliary verbs. For example, they would make mistakes like "I doed it" and "I haved it", but they would not generate any sentences like "Doed you come?" or "I haved eaten". Despite the fact that children are not exposed to all structures of language, they are able to fully acquire the fundamental linguistic knowledge at their early stage. This concerned with a logical problem of L1 acquisition, which is called Poverty of the Stimulus (POS) (Chomsky, 1980). However, SLA is usually achieved through systemic instruction, which requires careful attention and efforts. Therefore, as discussed in Chapter 1, grammar learning and grammar instruction are crucial and often emphasized in second or foreign language learning.

Given that passive voice is one of the important and the most disputed grammatical features in English learning and academic writing, this chapter reviewed literature relating to the concern of whether passive voice should appear in academic writing, and whether it influenced L2 learners' academic performance. It first presented the definition of passive voice and its common constructions. Then, the chapter reviewed the controversial debate upon the suitability between active and passive voice usage in academic scientific writing. Next, this chapter discussed the importance of passive voice in language learning and with its further impact on academic performance. Furthermore, the chapter explained why this grammatical feature was important for L2 learners. Finally, this chapter ended with a review of corpus-based grammar list specifically related to passive voice constructions and verbs that tended to be associated with passive forms.

<center>**Defining Passive Voice**</center>

Definition of Passive Voice

In English grammar, voice refers to the active and passive use of a verb. The active voice is used to make a direct statement about an action; the executor of the action is the subject of the sentence, such as in *A technician used an oscilloscope in the experiment*. The executor of the action is "a technician." However, passive voice is used when an indirect statement is made about some action, e.g., *An oscilloscope was used in the experiment by the technician*. The subject of the sentence is the receiver while the original executor of the action is the object of the sentence preceded by the preposition *by*. In other words, distinguishing between active and passive voice differs in the relationship of the verb to the subject; the subject does the acting in an active sentence while the subject is acted upon, or neglected in a passive sentence. Only transitive verbs can be used in the passive voice because these verbs can be followed by a direct object, and since intransitive verbs are never followed by a direct object, thus they cannot be used in passive sentence (Folse, 2016).

There are two common types of passives, adjectival and verbal. The characteristics of the past participle is the symbol of distinguishing these two passives, where verbal passive is defined when a verb characteristic was found, and likewise adjectival passive is defined when adjectival characteristic was found (Toyota, 2009). For example, adjectives can have a negative prefix *un-*, but verbs cannot; and there are comparative and superlative forms in adjectives (Toyota, 2009). However, Wasow (1997) has opposed the statement that some passive participles are adjectives and illustrated why adjectival participles rarely appear. First, Wasow argued upon double-object verbs, which had several problems in terms of grammar learning on the passive voice perspective. He pointed out that it was hard to address subcategorization requirements in adjectival analysis. For example, the selectional restrictions of the participle in the following

<center>16</center>

sentences *$10 was given (to the United Fund)* and *someone gave $10 (to the United Fund)* were the same as those of the verb. Therefore, he indicated that the participle was a verbal passive in this case. Second, Wasow disputed the complement-taking status of certain participles. He claimed that if the word *considered* in *John is considered a fool* was an adjective, then the element that followed *considered* could be an adjective as well. Yet noun phrases come afterward verbs, hence the participle should be a verb. He further argued that canonical adjectives like *obvious* in **John is obvious a fool* did not take complements like *a fool*. Therefore, *considered* cannot be an adjective because it took that kind of complement. Wasow claimed that there were no passive participles that were adjectives (1997). However, despite Wasow's argument, adjectival passive is still categorized as one of the passive constructions, discussed in the next section.

Passive Constructions

In a passive structure, the affirmative SVO structure is reversed, the object takes the place of the subject, the subject is placed either in a *by* phrase or omitted, and the main verb also conjugates to its past participle form and follows the verb *be*. In addition to the basic structure of passive sentence, there are some other passive constructions seen in English writings. Granger (2013) pointed out that passive structures could be mainly centered on the structural aspects: active-passive transformation (e.g., *he took the picture/ the picture was taken by him*), the variety of verbal passive tenses (e.g., present or past ones like *is/was washed by*, present progressive or past perfect ones like *is being made/had been made*, and even the complex ones like *will be being made* or *should have been being made*), or could be categorized in different structures, namely monotransitive, ditransitive, prepositional and impersonal (see Table 3 for examples).

In a broad sense, passive constructions can also be categorized into short passive and long passive (also known as full passive) in terms of whether a *by*-phrase is presented to specify

17

the agent of the action. For example, the passive sentence *Pride and Prejudice was written by Jane Austen in the 19th century* is a long passive while its *by*-phrase reduced version *Pride and Prejudice was written in the 19th century* is a short passive. Nonetheless, exceptions also exist in the use of *by* to indicate the agent, such as *he is well known to the world* and *the chicken is filled with spices*. That is, when the agent of a passive sentence is not specified either because it is unknown, unimportant or the information is common sense, short passive is presented.

In general, long passive can be replaced by an active clause conveying the same meaning, but sometimes using passive voice seems to be more appropriate in certain circumstances (Biber, Concord, & Leech, 2003). Interestingly, although a short passive looks like omission of the agent from a long passive, it should not be treated as a short version of long passive but as an agentless sentence (Coetzee, 1980). Then in what conditions should long and short passives be used in writing? Yannuar, Shitadevi, Basthomi, and Widiati (2014) summarized the principles for utilizing long passives in writing:

a) the information-flow, in which preference for presenting new information at the end of a clause;

b) end-weight, in this case the agent does not hold up the processing of the rest of the clause;

c) the long passive place initial emphasis on an element of the clause, which is the topic or theme (pp. 1401-1402).

Meanwhile, short passive is helpful and occurs frequently in scientific writing because it gives the status of topic to the receiver (Biber et al., 2003) and allows the readers and authors to rediscover the hidden agent (Coetzee, 1980). In fact, agentless passives are frequently seen in

English prose (e.g., Carter & McCarthy, 2006), and the first half of this very sentence is a good example.

There are many ways to classify passive constructions, some classify them according to verb tenses, some split them into clauses, some analyze them in a psycholinguistics perspective, and some categorize them based on their semantic and syntactic functions. The most commonly seen passive constructions are demonstrated as follows.

***Be*-passive.** The basic passive constructions, which subject to the *be + past participle* structure. In a *be*-passive sentence, the subject is being acted upon while the object is the executor of an action if presented, or unmentioned. For example, *this painting was painted by Da Vinci.*

***Get*-passive.** The passive constructed with auxiliary *get* is called the *get*-passive. The syntactic structures of *get-* and *be-* passives seem very close to each other (followed by a past participle), yet they have acknowledged differences in terms of semantic and/or pragmatic implications. The *get*-passive implies a sudden change, while the *be*-passive indicates a result (Folse, 2016). Therefore, *he was beaten* depicts a consequence over a fight, whereas *he got beaten* indicates that someone suddenly beat him, and he was caught off guard. Additionally, a number of past participles used with *get* normally imply an adversative or precarious situation, for example, *get arrested, get killed, get sued, get hurt, get beaten, get picked on, get criticized, etc.* (Carter & McCarthy, 1999; Folse, 2016). However, positive examples also exist, such as *get hired, get married, get promoted*. Even though *get*-passive is a common passive construction, the use of *get* is considered informal and thus normally used in conversation other than in formal language settings. The synonym *become* often substitutes for *get* in formal writing.

Since there are various structural configurations *get*-passive can function with, it is hard to simply identify the *get*-passive in its narrow sense. Carter and McCarthy (1999) summarized

a range of *get*-passive forms that contain closely related meanings with the canonical *be*-passives. Table 1 shows both the related forms and examples.

In addition to the semantic distinction of *get*- and *be*- passives discussed above, Carter and McCarthy (1999) also illustrated the semantic and pragmatic distinctions of these two passive constructions (see Table 2 for examples):

> (1') neutralizes the focus on the patient in (1). (2') removes the marking of agency/implied responsibility of the grammatical subject in (2). (3') retains implicit agency and seems to differ from (3) only in degree of formality, while (3″) neutralizes agency and is ambiguous between description of a state and reporting of an event. (4') is like (3'), apparently affecting degree of formality only. (5) is ambiguous between speaker as agent and some other party as agent; (5') retains this ambiguity, (5″) removes it, with speaker clearly as agent. (6') likewise affects formality, but (6″) removes the emphasis on change of state. (p. 48).

Table 1.

Structural Configurations of Get-passives (Carter & McCarthy, 1999, pp.46-47)

Structure	Example
a. X *get* V-en (by Z) (where X is patient)	(1) He got killed trying to save some other man.
b. X *get* reflexive pronoun V-en (where X is patient but with overtones of agency)	(2) You see, if ever you get yourself locked out… I showed her how to get in.
c. X *get* Y V-en (where X is patient and indirect agent)	(3) Rian got his nipple pierced and it was so gross.
d. X *get* Y V-inf (where X is indirect agent and beneficiary)	(4) She got me to do a job for her, fencing.
e. X *get* Y adjectival past participle (where X is agent, either direct or indirect)	(5) Right we've got to *get* you kitted out cos you can't go in like that.
(f) X *get* adjectival past participle (where X is patient)	(6) The tape seems to have got stuck.

Table 2.

Comparison of Get-passive structures and the Corresponding Be-passive Alternatives (Carter & McCarthy, 1999, p.48)

Type	*Get*-passive	*Be*-passive
a	(1) He got killed trying to save some other man.	(1′) He was killed trying to save some other man.
b	(2) You see, if ever you get yourself locked out	(2′) You see, if ever you are locked out
c	(3) Rian got his nipple pierced and it was so gross.	(3′) Rian had his nipple pierced and it was so gross. (3″) Rian's nipple was pierced and it was so gross.
d	(4) She got me to do a job for her, fencing.	(4′) She had me (to) do a job for her, fencing.
e	(5) Right we've got to get you kitted out	(5′) Right we've got to have you kitted out (5″) Right we've got to kit you out
f	(6) The tape seems to have got stuck.	(6′) The tape seems to have become stuck. (6″) The tape seems to be stuck.

Collins (1996) further identified *get*-passive constructions as central *get*-passives (consist with propositionally equivalent active alternatives), psychological *get*-passives (past participles exhibit a mixture of verbal and adjectival properties, in which *frustrated* conveys the fact that someone is being frustrated by and also indicates an anxious feeling), reciprocal/reflexive *get*-passives (as in the second construction in Table 1), adjectival *get*-passives (past participles are adjectives and *get* functions as a copula rather than a passive auxiliary), and formulaic *get*-passives (idiomatic expressions such as *get used to, get stuck into, get accustomed to, etc.*).

Phrasal Verb Passive. Phrasal verb passives are regarded as where the subject in a passive construction corresponds to the object of a preposition in its active alternation (Tseng, 2007). For instance, the active structure *the board called off today's meeting* has a passive structure *today's meeting was called off by the board*. Phrasal verb passives may be long passive, or short passive, as in *today's meeting was called off*.

Bare Passive. English allows a passive structure to include just a subject and the past participle of a verb, and this is called bare passive. Bare passive can occur as adjuncts (e.g., *one of its ads shows a washed-out manager, arms folded, sitting in a corner*), on its own in newspaper headings (e.g., *28 injured in accidental detonation*), and as independent clauses (e.g., *an expert like him duped by an email scam*) (Pullum, 2014).

Stative Passive. In some instances, the *be + past participle* structure describes a state or condition instead of an action. This characteristic of a passive construction is called stative passive. For example, *the account is unlocked* describes the condition of the status of the account now, not an action. Therefore, in stative passive, the past participle functions more as an adjective than a verb (Folse, 2016).

Adjectival Passive. As discussed previously, the *be* passive and adjectival passive are often ambiguously explicated. Additionally, take the example in the discussion of stative passive above, one can argue that as a *be* passive it reveals that someone took an action and unlocked the account at some point, and others can also dispute that the account remains an unlocked status since a point of time. To solve this ambiguity, Pullum (2014) provided two rules to test whether this kind of clause is an adjectival passive: 1) verb *be* can be replaced, e.g., *the account seems unlocked*, 2) adjectives can be derived with the negative prefix *un-,* e.g., *"the island was uninhabited by humans"* (p.63). Moreover, Toyota proposed three principles to test adjectival passives, namely affix, comparative and quasi-copula. Since the negative prefix *un-* can be attached to adjectives originate from a verb but not to a verb, we may assume that the past participles are originated from an adjectival participle but not from an original verb form. Furthermore, adjectives can occur in comparative and superlative forms but verbs cannot, and those adjectival passives can be graded with *very* as in *I am very confused*. Finally, quasi-copula refers to the copula with real semantic content that appear in the place of *be* (e.g. look, seem, etc.) (Toyota, 2009).

On the whole, as discussed in the first section of this chapter, the opponent opinion on adjectival passive also brings up good discussion upon the existence of such construction.

Embedded Passive. English also allows passive clauses embedded in active clauses. For example, the main clause in *the government had the case investigated by the police* is active with the subject *the government* and the verb *had*, while the subordinate clause is passive because *the government* is not the direct executor of the investigation (Pullum, 2014).

Concealed Passive. Concealed passive does not have a past participle as head but a gerund-participle, thus the prominent symbol (past participle) of a passive construction is missing in a

concealed passive structure. Take this sentence as an instance, "*this rug badly needs washing*" (Pullum, 2014, p. 63), although there is not any passive symbol presented in the sentence, it is obvious that the rug cannot wash itself but only to be washed by human beings (an agent), indicating that this sentence is a passive sentence.

All in all, Table 3 presents a summary of commonly seen passive constructions. Examples are also provided for each construction.

Table 3.

Passive Voice Constructions Classification

Verb tenses	Passive constructions	Other passive structures
Present, e.g., *my bathroom is cleaned every weekend.*	*Be*-passive, e.g., *the house was built by his farther*	Monotransitive, e.g., *the company was established in 1987*
Past, e.g., *King Lear is written by Shakespeare.*	*Get*-passive, e.g., *he got fired yesterday*	Ditransitive, e.g., *they were given $20*
Present/past progressive, e.g., *our travel plan is/was being made.* Aka non-finite – ing construction.	Bare passive, e.g., "*28 injured in accidental detonation*" (*Los Angeles Times,* 5 July 2013)	Phrasal verb, e.g., *the meeting was called off by the board*
Be + going to, e.g., *the main campus is going to be expanded.* Aka to-infinitive construction.	Phrasal verb passive, e.g., *the meeting was called off by the board*	Impersonal, e.g., *it is generally believed that eating habit has essential influence on physical health*
Present perfect, e.g., *the Nobel Peace Prize ceremony has been held at three different places.*	Adjectival passive, e.g., *his stomach is remained full* Stative passive, e.g., *the window is closed*	
Modals, e.g., *how the prizes should be awarded was discussed by the academy.*	Concealed passive, e.g., "*this rug badly needs washing*" (Pullum, 2014, p. 63)	

Controversy on the Use of Passives

Safire (1992) has strongly underlined that language is not complete without grammar and that grammar is one of the crucial symbols of accuracy and completeness of language expression. History has witnessed changes in the use of passive voice. From 18th century till the end of 19th century, early scientific prose in fact favored active voice (Ding, 1998). Then a prevalent statement in 20th century suggested that scientific articles ought to expound objective matters, thus they should be more objective when conveyed opinions. To sound more objective, passive voice appeared to do a better job since the object was put in the subject position (i.e. the agent was not prominent in action). Nearly all researchers of written academic corpora comment on the fact that passive forms are the most prevailing in the academic genre than in any other (Hinkel, 2004; Quirk et al., 1985; Swales, 1990). However, the scientific writing style nowadays seems to come back full circle with regained favor for the active voice. Nevertheless, since the present study focused on the current situation and the controversy of passive usage in terms of frequency (a percentage of total passive usage), constructions (most commonly used passive forms) and dispersion (verbs tend to be associated with passive usage) in undergraduate STEM textbooks, it did not examine the history of voice change in scientific writing in detail.

Shortcomings of Passive Voice

The biggest judgment of using passive in academic writing would be on a dull and vague style passive does to the expression (Beason & Lester, 2003; Matthews, Bowen, & Matthews, 1996; Sigel, 2009). Passive has also been accused of being indirect because the use of passives enables author to put reader's attention on the information (e.g., "*All analyses were performed with SAS…*" and "*… no significant effect was observed in the VIA group…*"), rather than on the authors or doers (e.g., "*We performed all analyses by using SAS software…*" and "*We observed no adverse events in either group…*") (Millar et al., 2012, pp. 402-403).

Sigel (2009) advocated for avoiding passive constructions in scholarly writing because passive voice weakened scholarly argument while active voice presented a more thorough understanding of the writing, hence strengthening arguments and increasing clarity of an article. Moreover, passive voice was considered as a bane in medical writing (Sheen, 1982), for Sheen claimed that passive voice "pervades medical literature with the haze and heaviness of stagnant air" and resulted in a perpetual writing tradition that was "fraught with ponderous and obscure language" (as cited in Millar et al., 2012, p. 394). First, using the passive voice would result in a more complex structure (more verbose) than its active counterpart and thus harder to understand. For example, comparing the active sentence "The dog chased the boy" versus its passive form *the boy was chased by the dog*, although both sentences convey the same meaning, the passive voice required 40 per cent more words, and in some other cases, the percentage was even higher.

Nevertheless, passives can be less verbose when short passive is used instead of long passive. Evidential support was found in medical journals where more than 90 per cent of all passives were presented in the short passive structure (Amdur, Kirwan, & Morris, 2010). By omitting the producer/initiator of an action, passive voice allows the writer to leave out information that is redundant or of little informational value (Biber, Johansson, Leech, Conrad, & Finegan, 1999; Folse, 2016). Millar et al. (2012) also indicated that it might be acceptable to omit the producer of an action in the conventionalized discourse structure of scientific writing. For example, subject *we* in the description of methodology could be omitted in a passive structure without any loss of meaning (e.g., "*proportions were analyzed with a stratified Mantel–Haenszel approach' (Lancet-1-M)"* vs. "*we analyzed proportions with a stratified Mantel–Haenszel approach (Lancet-1-M)"* (Millar et al., 2012, p. 395). In this example, sentences in

active voice and its passive alternative contained an equivalent number of words. Therefore, the use of the passive voice did not necessarily result in verbose writing.

The second claim argues that the meaning of a passive sentence is indirect and/or less forceful than an active alternative (Day 1989; Matthews et al. 1996; Sheen 1982; Strunk & White 2000; Taylor 2005). However, arguments that the passive voice will necessarily result in writing which is less forceful are also open to scrutiny. The passives are prejudicially opposed by actives advocators, where the criticisms of the passives are alleged to be dull, static, sneaky, evasive, feeble, weak and ineffective. However, it is important for writers, especially novice writers, to know what exactly justifies the accusation of stylistic badness; yet all the accusations are unsubstantiated, and somehow clearly wrong. Pullum (2014) provided real passive instances and syntactic description of passive clauses as very strong arguments against each of the alleged adjective above. Finally, he pointed out a harsh yet realistic fact, which had been often denied by actives supporters, that

> The claims about why you should avoid passives – the allegations about why they are bad – are all bogus, and the interesting point (the discourse condition) is always missed. The advice is often supplied by advice-givers who don't respect their own counsel – though they are unaware of that because they are commonly hopeless at distinguishing passives from actives. But the recipients of the advice can't identify passives either, so they are powerless to spot the blunders of their teachers. Even if they managed to follow the advice rigorously (which they can hardly do if it is not clear to them what a passive is), it would usually not improve their writing one whit. It would certainly make them write less like great writers of the past – and more a little child. (p. 73)

Pullum's standpoint is reasonable, we should not blindly shun the passives without understanding what they really are and what benefits they may bring to the beauty of our writings. Therefore, it is also important to teach students grammar to allow them to recognize passive voice in what they read and what they write. The importance of knowing passive voice will be discussed in later section of this chapter.

Several scholars, educators and grammar guidelines seem to continuously view the passive voice adversely and inherit the legacy of prescriptive rule about the passive voice: to avoid it whenever possible. Others are suspicious of the efficiency the passive voice display in presenting a cogent and coherent writing. However, it is worth noting that most articles that suppress the use of passive voice, such as Sigel's (2009) do not have absolute rigid rules of prohibiting the use of passives. Even though they strongly suggest using active voice in academic writing, they are still aware that passive voice can be used in appropriate contexts. Thus, exploring what the current norm of the passive voice usage was needed.

What Empirical Studies Have Found on Passive Uses

Despite the outcry of using passives in academic writing, it should be noted that a great number of studies disclose the actual desire and frequency of passive voice used in scholarly writing. If passive voice were unnecessary and loathed by scientific writing guidelines and scholarly writers as Pullum (2014) pointed out, the question then will be asked – should passive voice not appear as frequent as what the research has found, and should it completely disappear in any kind of writing? Given that there has been a controversy on the uses of the passive voice, more evidence is needed in order to draw an objective conclusion. Bazerman (1988) examined 23 articles from *Physical Review* published between 1893 and 1980 and found that at least 75 per cent of all main transitive verbs were in passive forms in the examined articles throughout the period. Moreover, Rodman (1994) investigated a corpus of 16 scientific articles containing

66,500 words and revealed that 66 per cent of transitive verbs were in the passive forms in a corpus of 2,215 transitive verbs, which was close to Bazerman's findings. Both studies indicated that passive constructions did not only occasionally appear but indeed were frequently used in scientific discourse.

Additionally, Harmon (1992) studied 50 most-cited scientific literature and the result showed that 53 percent of the main verbs were in the passive structures. However, Harmon's corpus of the main verbs contained linking verbs and intransitive verbs, which might have influenced the proportion of passive frequency in his study due to a bigger number of total main verb count in the denominator, thereby reducing the proportion of passive frequency within the tested texts. Ding (2002) also suspected the proportion to be higher and argued that if linking verbs and intransitive verbs had been removed from the study, the proportion of passive frequency would have increased, and the percentage would have been closer to the results in Bazerman's and Rodman's studies.

Atkinson (1996) examined 20 articles from *Philosophical Transactions of the Royal Society of London* between 1675 to 1975 and found the passive voice was truly undesired in the 16th century, but it had begun to become more and more prevalent in the 17th century until it even reached an "impersonal" phase in the 20th century. Based on his research, Atkinson came to the conclusion that the passives were "a high-frequency feature of modern scientific research writing" (p. 359).

Ding (2002) claimed that two social values in science – "falsifiability of science and cooperation among scientists" – determined the use of passives in scientific communication (p.137). He indicated not only passive voice was a high-frequency character of scientific writing, it was also a dominant stylistic characteristic in scientific writing. He further argued that

objective matters in this world were the representation of science, thus scientific writing should be thing-centered rather than human-centered. Since it was thing-centered, the subject of a sentence should refer more to objective matters than to humans (researchers or authors), because subjects emphasized topics of the sentences. Therefore, in the case of scientific writing, passive voice met the grammatical requirement of thing-centered scientific work. It was also a scientific responsible of concentrating on things being acted upon or experiments being done instead of focusing on researchers or authors who acted upon things or completed experiments. Without human (researcher) intervening between readers and the study, the reader felt more connected to the experimental process.

Here one may argue that human-centered discourse does not contradict scientific writing because there was a certain period when the Royal Society had favored author-centered discourse and author-centered discourse had reached its climax in 1775 (Atkinson, 1996). However, even though human-centered discourse represented the typical norm of scientific writing in 1775, the stylistic thing-centered prose was also developing progressively, and passive voice was mostly seen in the Methodology section as well at that time. After one hundred years, only 4 out of 22 articles in Philosophical Transactions were written in an author-centered approach, while the rest of the articles had switched to the norm of an object-centered way (Atkinson, 1996).

In Ding's article, he summarized three merits of passive voice as it was an integral part in scientific writing. First, he noted that using passives to concentrate on the physical world helped "de-emphasize discreteness of scientific experiments" (2002, p. 137). Additionally, it removed "personal qualifications of observing experimental results" (2002, p. 137). Finally, by presenting objects of scientific work, passives improved collaboration among scientific researchers. To

fulfil the two scientific social values, the passive voice played a role as objective practices of scientific studies rather than choices of personal preference among independent scientists. He pointed out that scientific theories may sometimes be invalid, thus scientific experiments needed to be even objective. Scientists could examine them critically so that a valid theory could be replaced, only when the theory was falsifiable and refutable. On the other hand, since valid scientific theories must be testable and repeatable with the same outcome, the primary attention should not be focusing on who conducted the experiment and who proposed the theory but focusing on the experimental procedures and the results. In addition, the second social value manifests in Ding's article indicated the integrality of scientific work, suggesting a close relevance and connection among theoretical foundations and experimental practices. Therefore, he saw a general trend in scientific communities, which embraced cooperation among scientists including scientific writing. To think in this way, the use of passive voice provided a common knowledge foundation among working scientists in all fields, which fostered collaboration (Ding, 2002).

Historically, passive voice used to be taught as the standard in major universities in the United States from the early 1820s. In the early 2000s, however, some progressive teachers and writers have begun to go on an opposite path by advocating a more personal "active" voice to novice writers in writing lab reports and research manuscripts (Stone, 2014). As a result, recent studies on the use of passives have revealed a less but still relatively high frequency of passive proportion in scientific writing. For instance, Alvin (2014) demonstrates that passive clauses constituted 30.02 per cent of the total number of clauses analyzed in 60 scientific articles published in 2013. According to his findings, Alvin brought up a pivotal point of what really happened in scientific writing versus what had been advocated in writing guidelines:

…however, the use of the passive voice may at times be more a matter of author choice rather than editorial preference. The varied proportions in Sci lend some support to this suspicion. …the anomaly could have been due to the vague nature of the advice given to authors (p. 9).

In fact, the passive voice is hard to avoid. The noted interference effects of single word production experiment prove that semantically similar distractor words influence the speed of naming a picture. The most well-known result is that when the distractor word is semantically related (of the same semantic category) to the target word (e.g., cat and dog vs. cat and clock), the picture naming speed is slower than when the word is semantically unrelated (Levelt, Schriefers, Vorberg, Meyer, Pechmann, & Havinga, 1991; Mahon, Costa, Peterson, Vargas, & Caramazza, 2007; Rosinski, Golinkoff, & Kukish, 1975). Later research studies started to look into the aftereffects of this interference in sentence production and found that producers tended to omit one of the interfering elements when there was low interference (i.e., low similarity-based interference). For example, Christianson and Ferreira (2005), Gennari, Mirkovi´c, and MacDonald (2012), and Hsiao, Gao, and MacDonald (2014) concluded that speakers were inclined to omit subjects when both the subject and object had low similarity-based interference. In their studies, participants tended to utter more agentless passives when the agent was semantically or conceptually similar than when the agent was disparate. In Gennari et al.'s study, they examined the interference effects in English, Spanish and Serbian while the other two studies investigated the effects in Mandarin (Hsiao et al., 2014) and Odawa (Christianson & Ferreira, 2005). The results reconciled with each other revealing the fact that passives tended to be produced by participants when the subject and object were semantically or conceptually similar even though active structures could also make perfect compositions. Furthermore, many

studies also showed that before individuals produced sentences, their brain would go through an utterance planning process where they would strategically choose an utterance form that reduced the difficulty of the whole planning process, i.e. reduce interfering elements of a sentence (MacDonald, 2013). One way of doing so was to choose a form where the interfering elements were placed in a selected syntactic position in which the interfering elements were at a distance from each other (such as in passive voice *The boy was chased by the dog*), or in which one element was omitted (Hsiao et al., 2014). MacDonald argued that this strategic utterance planning mechanism emerged from non-linguistic action and motor planning, where easier motor plans (i.e., utterances that are simpler, practiced more, used recently, etc.) captured internal attentional resources over more intricate plans, so that it was more likely to be carried out than the more intricate alternatives (MacDonald, 2013).

Studies showed clear effects of lexical accessibility on sentence structure and word order in production, so that more conceptually salient (accessible) nouns tended to be placed earlier in the utterance or at a syntactically more prominent position (e.g., Bock and Warren, 1985; Tanaka et al., 2011), such as the grammatical subject. For example, animate nouns, thought to be more salient and recalled more rapidly from long-term memory, were more likely than inanimate nouns to be uttered early and assume the surface subject position even when they are not agents of the event, resulting in the production of passive sentences like *the boy was hit by the ball* rather than the active form *the ball hit the boy* (Bock etal.,1992).

In addition, Gennari et al. (2012) investigated active (*the baby that the woman is holding*) vs. passive (*the baby that is being held by the woman*) relative clause utterance in English, Spanish and Serbian. Unlike simple active and passive sentences that have completely reversed noun orders, relative clauses in these three languages fixed the position of the modified noun

"baby" in the clause-initial position and thereby presented better comparison result between the active and passive forms. Therefore, in the active form *the baby that the woman is holding*, the modified noun *baby* and the embedded subject *woman* were close to each other and were both in salient grammatical positions (i.e., the baby: main clause subject, the woman: relative clause subject). The results indicated that the passives were frequently used when both entities were salient (animate, in this case, such as a woman holding a baby) than when an inanimate entity was being acted upon (a woman holding a vase). Therefore, it appeared that the choice of using passive voice was a speech and brain mechanism containing complex interrelated linguistic factors, which was impossible to discard.

Ferreira (1994) also disclosed consistent findings regarding whether animate or inanimate nouns were to stimulate the passive structures, and adding to the results for the type of verbs that tended to elicit more passives as well. By employing verbal input in her study, each participant (all participants were English NSs) was asked to generate a sentence using prompts (two nouns and one verb). The primary verbal input of verbs was manipulated as experiencer-theme verb (e.g., avoided) and as theme-experiencer verb (e.g., challenged). The results indicated that passive constructions occurred more frequently when theme-experiencer verbs were presented than when experiencer-theme verbs were presented. As for observed nouns in her study, Ferreira found that passive constructions also occurred more frequently when one of the nouns in a sentence was animate while the other was inanimate than when both nouns were animate. Even in her last experiment, she controlled the independent variable by removing the morphological marking for tense from all verbs (e.g., challenged was switched to challenge) so that participants' intention of constructing passive sentence was diminished. However, the results from the last experiment were substantially the same as previous experiments when morphological cue was

controlled, reinforcing that participants tended to generate more passive constructions with theme-experiencer verbs and with animate and inanimate nouns combination within one sentence.

In Hundt, Schneider, and Seoane's (2016) study, they examined the use of *be*-passives in academic Englishes from the ICE corpora (the ICE corpora included in their study contain academic English texts from 15 countries where English is or is not the native language). The results demonstrated that, except for American English, *be*-passives were about equally frequent in both English as L1 and regional contexts, and they were distributed similarly across all investigated disciplines (i.e., humanities, social sciences, natural sciences and technology) (Hundt et al., 2016). Literature discussed so far has aimed to explore the frequency of passive voice in scientific discourse, opening the question whether the passive usage is equally prevalent in other areas of study in STEM.

All in all, what has been claimed about the use of passive voice is not in accordance with what really manifests in academic writing, especially in scientific discourse. Passives are in fact necessary in certain contexts and have essential functions in fulfilling important roles in scientific writing. Therefore, they need to be critically inspected for their grammatical as well as contextual roles in scholarly writing, rather than being avoided in general. In the following section, the important roles passives play in English language learning as well as in scholarly writing was discussed.

The Importance of Passive Voice in English Learning and Academic Study

There are particular times "when the passive is not only appropriate but actually the best choice. It depends on what the writer or speaker wants to accomplish with any given sentence and how that sentence fits into the paragraph" (Folse, 2016, p. 293). Passive is not just a regular

grammatical feature that fulfills the needs of language expression, but it essentially enriches

literary diversity, emotional expression and standard formulation in literature. Additionally,

passive constructions are used to perform special discourse functions because of its marked voice

(Biber et al., 2003). It is equally important for L2 learners in learning the target language, for

they can better understand what they receive as well as precisely express what they produce.

The Importance of Passive Voice in Reading

Reading is an important language skill, and it is crucial to every aspect of learning.

Students learn new vocabulary through reading, they learn grammar through reading, and they

start composing because they can convert receptive knowledge (reading) to productive

knowledge (writing). Reading can not only improve students' language proficiency, but it opens

a door to all knowledge. Being able to read and to understand what we read is one crucial way to

study languages and content knowledge in schools. As Zygouris-Coe noted in her study,

"without comprehension, reading can be frustrating and at times even painful. A major goal of

reading comprehension instruction is to help students develop the knowledge, skills, and

experiences they need to become independent readers and lifelong learners" (Zygouris-Coe,

2009, p. 2).

Moreover, genres and text types, which are distinguished by different disciplines, should

also be considered as important elements in students' reading comprehension. Genres are

important socio-cognitive schemas for writers to produce texts to meet the demands of various

contexts (Bazerman, 1997; Grabe & Kaplan, 1996; Johns, 2008). For example, a writer might

have a genre schema for writing an academic article, and academic textbook, or a conference

proceeding. Therefore, from this perspective, genres are also purposeful, and writers should be

aware of using appropriate textual approaches for specific rhetorical demands. However, Johns

(2008) argued that novice academic readers were not ready to be exposed to various genres unless they were trained; in another word, they had not developed genre awareness yet.

In addition, disciplinary literacy approach highlighted that disciplines did not just have content differences but also differed in the ways how content was produced and communicated (Fang & Coatoam, 2013). Students' awareness about how language constructs knowledge variously in different disciplines could be raised by building up their ability to recognize conventional language patterns in different fields of study. In addition, students could better acquire school knowledge and develop more effective disciplinary literacies across academic subjects when taught in a language explicit and discipline-specific way (Fang & Schleppegrell, 2010).

With regard to why the language used in the study of science need to be paid more attention in reading instruction, Fang (2007) presented some salient features of the language of school science and its difference from children's everyday language and discussed why the expository language is challenging to middle-school students in terms of reading comprehension. He listed eleven features, such as technical terms (with multi-morphemic and Latin and/or Greek origins) that rarely occurred in everyday conversation, and ordinary words with non-vernacular meanings or usages (e.g., schools as in groups and fault as in an extended break in rock formation). It was also worth noting that passive voice was one of the salient features proposed in his study. By providing some example passive sentences, he suggested that passive voice "be a useful grammatical resource for obscuring responsibility and agency" (Fang, 2007, p. 505). For instance, he suggested by saying *photosynthesis and respiration can be thought of as opposite processes* over *scientists think that photosynthesis and respiration are opposite processes*, the doer agency was hidden through the passive form, which created an objective

author stance and at the same time "makes the text less involving and more alienating" (Fang, 2007, p. 505). He further argued that readers who did not understand the essential role passive voice plays in "construing values and ideology are less likely to become critical readers and aware of the environmental and ecological issues implicated in the text." (Fang, 2007, p. 505).

Since passive structures appear commonly in science and technical texts (Alvin, 2014; Krisch & Houdek, 2015), and L2 learners have difficulty in comprehending the passives in readings, it is necessary to teach L2 learners, especially those students who study in the STEM field, the passive structures in order to help enhance their reading comprehension. The experiments of Nation and Snowling (2000) reinforced the viewpoint that children's syntactic awareness skills and their reading skills were related to each other, which suggested that being able to consciously aware of the passive constructions may do good to readers' reading comprehension when passive forms encountered and explicitly taught.

Given that genres and text types are important factors that might influence L2 learners' ability to comprehension different contexts across disciplines, explicit reading instruction is needed. By emphasizing the important role reading comprehension plays in academic study, Zygouris-Coe urged effective teachers to employ appropriate teaching strategies to maximize students' learning outcome by improving their reading comprehension (Zygouris-Coe, 2009). Researchers have also advocated explicit literacy instruction to "discipline-specific cognitive strategies, language skills, literate practices, and habits of mind" (Fang & Coatoam, 2013, p. 628).

Research study has indicated that there was a need to explicitly instruct middle school students for the purpose of helping them engage with complex texts in science (Fang, 2005), and it was suggested that explicit instruction of the science-based language be embedded in science

literacy pedagogy (Fang, 2007). In fact, based on empirical evidence, explicit reading

instruction has been suggested by numerous scholars as a way to enhance L2 students' reading

comprehension (Fang & Wei, 2010; Graves, Juel, & Graves, 2001; Kern, 1989). For example, in

Fang and Wei's (2010) study, they divided a total of 233 Grade 6 students into control group (n

$= 93$) and experimental group ($n = 140$) to examine if a reading infusion program could have

positive effect on students' reading comprehension. Only the inquiry-based curriculum was used

to the control group, and the inquiry-based curriculum plus reading infusion (combined with

explicit instruction of reading strategies and home reading program) was utilized to the

experimental group. The reading strategies employed in their study including predicting, thick

and thin questioning, concept mapping, morphemic analysis, recognizing genre feature,

paraphrasing, note taking, and think-pair-share. The home reading program required students to

read one science book per week, share the book with their family members, and complete a

Reading Response Sheet for each book. The results showed that the testing score of

experimental group was statistically higher than the control group ($F_{(1, 223)} = 6.64$, $p = .01$, $d =$

$.22$) in reading comprehension, indicating that middle school students' science literacy could be

improved by explicit reading instruction suggested in their study.

In addition, Kern (1989) had 53 college students, who enrolled in intermediate French,

participated in his study in order to find out whether the explicit reading instructions he provided

(i.e., word analysis, sentence analysis, discourse analysis, and reading for specific purposes)

would improve L2 students' reading comprehension. The results revealed a statistically

significant difference between the group received explicit reading instruction and the group with

no such instruction, and the finding in his study indicated that explicit instruction in reading

offered a positive effect on students' reading comprehension.

Other strategies that could also help teachers improve L2 learners' reading comprehension and language application of scientific language included vocabulary building, noun extension, sentence completion, paraphrasing, sentence stripping, developing awareness of signposts (Fang, 2007). Wellington and Osborne (2001) also suggested teachers' attention focus on explicit instruction in grammar learning of the specialized language of science, because science remained a foreign language to many L2 learners when they did not consciously understand the unique grammar of science.

Additionally, L2 learners' reading comprehension is hard to be optimal because they tend to focus more on the surface structure of the target language and the lack of ability to effectively implement high-level interpretive processes due to the deficient word recognition skills (Kern, 1989). Shen (2003) also suggested that explicit instruction in reading be a necessary tool to improve ESL and EFL reading comprehension. Grammar instruction strategies for the recognition and understanding of text meanings have also been considered by a large body of studies (Bernard, n.d.; Celce-Murcia, 1992; Lightbown & Pienemann, 1993; Zhang, 2012). Shen (2003) noted that L2 learners may need explicit grammatical knowledge to be able to understand intrasentence (meanings within a sentence) and intersentence (meanings between sentences) meanings of the target language. That is to say, good comprehenders in reading had more explicit knowledge of syntax (Gaux & Gombert, 1999).

Furthermore, in Taylor, Stevens, and Asher's (2006) meta-analysis study, they reported that reading strategy instruction for L2 learners was not effective with students in elementary level and in their first year of L2 learning, while it was effective with students who were older and who had greater L2 learning experience (Taylor, Stevens, & Asher, 2006; cited in Goldenberg, 2010). One of the main purposes of the present study was to provide pedagogical

implications to English instructors in order to help college L2 learners improve their reading comprehension in terms of the passive forms, thus the implications from this study have theoretical support according to Taylor et al.'s (2006) study.

When is Passive Voice Used

Which voice authors should use when writing reports or papers in the STEM field would not be an appropriate question to ask. Rather, asking when the active and passive voice should be used is more appropriate, because it is really not an "either/or" situation. Anderson (2015) indicated that it was not a black or white approach in writing scientific reports, as scientific writers neither must nor had to use one specific voice in that kind of reports. She quoted Watson and Crick's article (1953) and demonstrated that it would be beneficial for parallel structures and fluent sentences and paragraphs to use both active and passive voice flexibly. Moreover, by citing Williams and Colomb (2010) and Sutton (1992), she suggested writers avoid the use of the first person in the Method section and pointed out that scholarly writers seldom used the first person to describe experimental procedures they conducted. That is, they never wrote: "I added acid to the test tube" (Anderson, 2015, p.50).

Interestingly, the idea of avoiding the first person in scientific writing resonated with Ding's (2002) opinion of falsifiability of science – distance the researcher/author from the writing implying the experiment should be duplicated and repeatable by other scientists. Here, the special discourse function the passives were executing was to create distance between author and the text, thus explaining the reason in which most passive constructions were found in scientific discourse (Baratta, 2009). If passive voice was used, the emphasis was on the object but not the agent. Therefore, that was the reason why many writing guides of University Writing Center (UWC) pointed out that passive voice could be a better option if the writer did not care who performed the action or if the actor was irrelevant. In general, we use passive voice when

the doer of an action is not the most important thing of an event or irrelevant, as in *All the tickets for the Broadway show Hamilton were purchased.*

Dawson (1987) presented the following rules to guide authors when to use the passive voice in technical writing:

a) Use the active voice when the performer is mentioned.

b) Use the passive voice when the performer of the action is not known or need not be mentioned.

c) Do not use the passive voice to make vague statements in which the performer is not mentioned but should be. Instead, use the active voice with the performer as the subject. (p.171)

Dawson also provided tips of when to use passive voice, which basically concerned with the importance of the performer/doer (if the performer is of no concern, then use the passive voice) (Dawson, 1987). His guideline also reconciled with many other university writing guidelines demonstrating that the passive voice undoubtedly had its special discourse functions and should not be blindly avoided, especially not in the scientific field.

Another discourse function that passive voice execute is to reflect author's stance. Authors hide themselves as performers in order to make the readers focus on what is done instead of who conducted the action (Baratta, 2009). Passive stance can be regarded as a subtle way for writers to announce personal opinions and feelings within their written discourses. Passive voice is also a syntactical way of constructing an impersonal expression to avoid a subjective point of view. In order to understand students' passive stance in academic writing, Baratta (2009) analyzed essays from three undergraduate students at the University of Manchester. The findings led her to the conclusion that passive stance was a "rhetorical reality"

(p. 1406) as well as a subtle way to reveal writer's thoughts and feelings within their writings. One example from the student's essay in (1) revealed the student's resistant feeling about being under linguistic pressure, which did not simply express the student's linguistic challenges but also conveyed her views and created an alternate perception in a more direct manner.

(1)   When I was 16 my family moved to England. This had a great impact on my language as I was forced to communicate in a language that I was not very confident in, especially not with the terminology used in the different lessons at school and among the people my age (p. 1414).

Moreover, the study of Emeksiz (2015) presented a source-based pattern for the use of passives in academic writing reinforcing that passive constructions was useful for the construction of stance in discourse. By analyzing the distribution of the passive clauses from 10 research articles containing 1233 clauses in linguistics and educational sciences, Emeksiz summarized how authors reflected their stance through the functions of passive voice and their choice of using passives due to difference purposes according to the source of information. When the source was the writer themselves, they made use of the passive voice to 1) refer to a status of their study, 2) direct the readers to some part of the text, 3) make assertions, predictions, and suggestions. She also indicated that the passives were desired by the authors when they were "citing the contemporary work and reporting generic assumptions and shared knowledge" (Emeksiz, 2015, p. 21).

Furthermore, it is crucial to compose an organized and cohesive textual structure in academic writing, and passive voice satisfies this function. Passive voice can help with textual cohesion so that the information in the text flows smoothly and is easy for readers to follow. The structural flow in turn helps create cohesiveness of the text. Textual cohesion is achieved

45

through a theme-rheme structure proposed by Halliday (1970). The theme refers to the first element within the given information functioning as "the point of departure of a message" (Halliday & Matthiensen, 2004, p. 64), and the rheme is regarded as the remainder of the message in a clause where theme is developed (as cited in Wang, 2007). Theme typically includes acknowledged information, while rheme contains new or unfamiliar information. The following example shows how the theme-rheme structure allows information to flow from one sentence to another smoothly might be seen from the distinguished Watson and Crick paper (1953):

(2)   We wish to suggest a structure for the salt of deoxyribose nucleic acid (D.N.A.). This structure has novel features which are of considerable biological interest. A structure for nucleic acid has already been proposed by Pauling and Corey. (p. 737)

In the first sentence of (2), "structure for the salt of D.N.A." appears as new information, i.e., the rheme. In the remaining sentences, it becomes a familiar information, i.e., the theme, therefore, providing textual cohesiveness to what has been done before. Meanwhile, the passive structure in the third sentence also constructs parallel structure by repeating the theme of the second sentence, which contributes to textual cohesion as well (Anderson, 2015).

Besides scientific writings, research on the Business field has also found preference of the passives in particular scenarios. Business communication writers prefer to use passive rather than active voice when communication represents the company acting in its entirety instead of an individual senior management. In business, many people think passive voice can make certain contexts more authoritative and professional, and by avoiding the prominent subject "I", officer or management staff can shrewdly depersonalize one from the messy business of personal confrontation. Biondi (2002) provided a good example of a business scenario, where the use of

46

the passive voice was not only favorable but necessary. In a warning letter to Dagwood, who was an electronic engineer of the aerospace industry, regarding his recent discipline violation, the use of passive voice in example (3) made the message more direct and vigorous:

(3)   *TO: D. Bumstead*

FROM: Human Resources Department, Disciplinary Action Division

It has been noted that your lunch periods have exceeded the allotted half-hour on Monday, Tuesday, and Wednesday of this week. According to the employee handbook, a written warning must be given when an employee exceeds the time allowed for lunch breaks for three consecutive days. If an employee exceeds the time allowed for lunch breaks for five consecutive days, disciplinary action, up to and including termination, will be taken. (p.3)

On the contrary, the tone would sound completely different and even awkward if Dagwood receives a letter written in active voice:

(4)   *TO: D. Bumstead*

FROM: J. C. Dithers

Bumstead, I've seen you sneaking back into the office late from lunch for the last three days. Don't do it again or I'm going to fire you. (p.3)

On the whole, we should not neglect the special discourse functions of passive constructions but need to properly view them as an essential element to the wholeness of a composition. In addition, knowing the conventionalized rhetoric style can also be helpful for both native-speaking and nonnative-speaking writers. Therefore, the next section will illustrate the convention of passive usage in academic writing.

In What Context is Passive Voice Used in Academic Writing

Passive style was highly presented in the Abstract section of the articles of *Philosophical Transactions* in 1825, and has maintained a high proportion from then onward. Recent studies have revealed that passive constructions are commonly seen in Methods and Results sections. In the experimental procedure section described in Laemmli's article, not a single human agent appeared, and "*of all the 48 transitive verbs, 42 are in the passive form with grammatical subjects referring to things*" (as cited in Ding, 2002, pp. 144-145). Ahmad (2012) found that 70 per cent of passives were deployed in the Method section of Medical and Natural Science articles. Despite that, passive voice is not considered as a preferred way of writing according to many writing instructional book authors and teachers. However, different passive voice constructions actually take up a considerable percentage of academic and other discourses based on various empirical studies. Proportionally, passives account for 25 per cent and 15 per cent in of all finite verbs in academic and news prose respectively (Biber, 1999). Among different passive constructions, short passives are predominant in all syntactic positions, while long passives are found the most common in news and academic prose (Biber, 1999).

In particular disciplines, the percentages of using passive voice is relatively high. Quite a few studies have examined the use of passive voice in academic writing in the field of science and the results showed a higher frequency of passive voice in science writing. Alvin (2014) examined the proportion of passive voice used in science journals. In his study, 60 research articles comprising 250,947 words were selected from six well-known science journals published in 2013. The results showed that about 30 per cent of all clauses were passive clauses, and of all passive clauses 29 per cent were found in the Method section. Importantly, Alvin noted that it would be impossible to see the number of passive constructions drop and the use of the passive voice in the Methodology section would continue to be widespread. Similarly, Amdur et al.

48

(2010) evaluated the frequency of passive voice in medical journal articles. What was different from Alvin's study was that that study included opinion papers and review articles in addition to original research reports. They examined 90 articles from three mainstream medical journals and found that the frequency of passive voice usage was similar across all three medical journals. The distribution of the proportion of passive frequency revealed the same results as Alvin (2014) indicating the highest percentage of passive usage in the Method section (M = 44%). However, due to the percentage being statistically higher compared to the frequency of passive voice in nonmedical publication ($p < .0001$), Amdur et al. (2010) suggested setting an upper limit of 10 per cent of passive voice proportion in medical articles instead of keeping this conventional writing style because they argue that passive voice makes texts difficult to read.

In contrast, Millar et al. (2012) have a different position regarding the use of passive voice in medical journal articles. Their study described how the passive voice was used in medical journals by using a corpus that contains 297 primary research articles in the medical field. The results showed that the Methodology section had the highest proportion of passive voice frequency, which was about 68 per cent and the Abstract section has the lowest percentage of frequency, which was about 27 per cent. They claimed that the high frequency of passive voice represented an established convention of the field, and it was even preferable in certain cases.

Even though simple transformation of verbs from passive to active voice was possible, it did not necessarily result in better reporting (Millar et al., 2012). Moreover, in some experimental research articles, Riley (1991) also found consistently high percentage of passive frequency in the Methods and Results sections, whereas rare examples were seen in the Introduction and Discussion sections. Therefore, this is once again showing how over-simplistic

it is to advise writers to use active voice whenever possible. Moreover, the results of other research studies are also consistent with what Millar et al. (2012) have found that over half the sentences in the Method section are in the passive forms (Biber & Finegan, 1994; Swales, 1990).

Besides the field of science, researchers also studied other disciplines, such as the use of passive voice in American soap opera dialogue (Schwarz, 2015), which represents spoken language of the period. Schwarz (2015) examined the change of *be*-passive and *get*-passive in contemporary American English in ten American soap operas from the years 2001 to 2012. The results indicated that there was a decline of *be*-passive constructions. However, the declined percentage still reached 30 per cent raw frequency (compared to 35 per cent). The *get*-passive did not have a dramatic change. It went from 4.4 per cent to 4 per cent in terms of raw frequency. The author attributed the decline to the colloquialization as one possible reason. Meanwhile, it was still worth noticing that the raw frequency of passive voice was relatively high in spoken language if all past participles were included (Schwarz, 2015). In other literature, such as narratives describing domestic violence, Frazer and Miller (2009) concluded that there was an increased frequency of passive voice when perpetrators were male, which suggested that comparing to violent or negative acts in general, writers particularly preferred to use passive voice when describing male-on-female violence.

In conclusion, passive structures are indeed frequently and widely used in many kinds of literature, especially in scientific discourse. It is important for grammar book authors and educators to consider the essential role passive voice plays in academic writing. It is also crucial for writers to understand when and where should passive voice be appropriately used. In addition, L2 learners also need to learn and know how to apply the structure of passive voice accurately.

The Importance of Passive Voice for NSs and L2 Learners

Passives are considered as more complex than actives. Since passive constructions are more difficult to process in comprehension, they are acquired and mastered later in language learning compared to active structures (Diessel, 2004; LoMaglio & Robison, 1985; Turner & Rommetveit, 1967), passives are admitted being cognitively and syntactically difficult (e.g., MacDonald, 2013). Passive voice is not only a difficult grammatical feature for NNSs, but also for NSs as well (Biondi, 2001; Folse, 2016; LoMaglio & Robison, 1985; Mirdamadi & De Jong, 2015; Pullum, 2009; Pullum, 2014; Street & Dabrowska, 2006).

Jaeggli (1986) noted that producing passive structures required additional computations compared to active sentences. In addition, the passives were syntactically more complex than their active counterparts due to the increased components in the passive sentence (it can add as much as 50 per cent component to the sentence). This standpoint was reinforced by Ferreira's (1994) study in which participants were more likely to spend more time formulating the passive sentences than the active alternatives. For example, in Ferreira's first experiment, statistically significant difference of the average formulation time was found ($p < 0.01$) between active responses (3996 ms) and passive responses (4635 ms). Other experiments in her study also revealed similar result with the first one. Furthermore, studies that examined NSs' comprehension on actives versus passives revealed that comprehension of passives was slower than of actives (e.g. Kharkwal & Stromswold, 2014), and that NSs, who have lower educational levels sometimes show difficulty in comprehending passives (Ferreira, 2003; Street and Dąbrowska, 2010).

Why ELs Need to Particularly Learn Passive Voice

Hinkel (2004) analyzed the patterns and median frequency rates of native and nonnative speakers' uses of three English tenses, two aspects and passive constructions in academic

writing. She analyzed 746 essays in a total of 226,054 words and found significant difference between NSs and NNSs in terms of the usage of English tense and the passive voice. The results showed that advanced NNSs seemed to have difficulty in using English tenses, aspects and the passive voice in academic writing even after years of English learning and use.

White (2008) investigated and reviewed the effectiveness of different types of form-focused instruction provided within communicative context through three intervention studies focusing on possessive determiners (PDs) specifically "his" and "her." Research carried out in communicatively-oriented L2 classrooms showed that although elementary school-age learners could figure out a lot about the language they were learning on their own while engaged in meaning-focused activities, they still needed help from their teachers in acquiring those linguistic features that remained troublesome, even after many hours of classroom exposure. For some L1 speakers of Romance languages such as French, Spanish and Catalan, the English possessive determiners (PDs) "his" and "her" represented such a learning challenge, and it was not uncommon to hear proficient adult speakers of English use one form when the context requires the other. Therefore, grammatical features were not only important because they compose perfect prose but also crucial to be explicitly taught for L2 learners.

Given that NSs have difficulty in acquiring passive forms and that passive voice is one of the last structures mastered by young NSs (LoMaglio & Robison, 1985; Turener & Rommetveit, 1967), non-native ELs would find passive structures even harder to master. Meanwhile, we need to know that passive production of NNSs is also influenced by other factors, such as direct language transfer, proficiency level, input instruction, etc. The most difficult passive form for L2 learners to master is the agentless form, which unfortunately happens to appear frequently in textbooks and reports in science and engineering (LoMaglio & Robison, 1985). Granger (2013)

observed a large number of ESLs and EFLs underuse the passives. The passive structures were quite complex and, in most circumstances, the corresponding structures in L2 learners'' L1 were less frequent than the passive in English. There was a great number of errors in the passive uses, especially the lower level learners, who constantly overextended the passive from to active contexts (e.g. a part time job *can be improved* student's communication skill) (Granger, 2013). However, as L2 learners' proficiency level went up, the frequency of this error slightly went downward possibly because of a growing knowledge of verb morphology.

As L2 learners' English proficiency improves, they need to understand the writing convention in their own field. As discussed in the previous section, authors use more passives in the Methods section of scientific articles, however, they also occasionally use first person subject in those sections probably in response to guideline that use active voice whenever possible. Truly, there is no huge distinction between *We conducted the experiment in an aseptic condition* and *The experiment was conducted in an aseptic condition*, but in fact, if the article has only one author, then the first person "we" is still not able to identify whom the "we" represents. It is even more awkward to write *I conducted the experiment in an aseptic condition* because it might sound egotistical. Once again, scientists are more interested in how the experiment was being conducted than who performed the experiment, and by not emphasizing the performer, the experiment becomes a professional practice rather than personal choices of individual researchers (Ding, 2002). Therefore, passive voice is more appropriate in this case.

The reason why passives need to be purposely learned by L2 learners is probably because passive sentences take longer to formulate than the active ones (Ferreira, 1994). First, since L2 learners produce a large number of erroneous passive forms and structures, teachers should help them correct those errors. Hameed (2017) indicated that the most frequent error of passives

made by ESL learners was the incorrect use of past participle (accounts for 74 per cent of the total errors), while the least frequent error made by ESL learners was the incorrect use of irregular of have/had (accounts for 10 per cent of the total errors).

Thus, it is clear that grammar class should not only focus on teaching grammatical structures, but in the passive cases, vocabulary instruction of past participles is of importance in ESL grammar teaching. Passive voice is used in scientific discourse, so L2 learners need to know when and how to use the passive voice. Most L2 learners lack this guidance and instruction because the passive voice has not been recommended to be used in scholarly writing. Studies show that most language learners significantly underuse the passives. For instance, L2 learners from China, Japan, Korea, Indonesia, Vietnamese and Arabic were observed to employ consistent underuse of the passives (Hinkel, 2004). Moreover, many L2 learners were found to misuse the passives, which largely related to the overpassivization of verbs reported in the literature, e.g., *disappear; happen; occur* (Cowan, Choi, & Kim, 2003; Folse, 2016; Oshita, 2000).

Prystupa (2017) identified the most difficult grammatical features for Chinese ESL students to acquire were the past perfect tense in both active and passive voice and passive voice for all grammar facets from a self-created corpus of Chinese students' academic writing comprising 510 texts with 701,440 words in total. In addition, he pointed out that the ability to utilize such grammatical forms was related to language proficiency. The results showed that the frequency of error-free passive sentences produced by low English proficiency students was 53 per cent and the frequency of error-free passive sentences produced by high English proficiency students was even higher (63 per cent).

There were research studies focusing on passive unaccusative errors such as *What was happened.* Oshita (2000) suggested these unaccusative structure of passives may be caused by an overt maker of NP movement and an overgeneralization of the passive morphosyntax of English. Besides Chinese and Japanese L2 learners, many Korean EFL learners tended to omit either the be or -(e)d in their passive sentence, and they often found it challenging as to appropriately use adjectival participles (e.g., *confused/confusing*) (Kim, 2002; Yeo, 2002).

Second, a number of teaching pedagogies have demonstrated effective learning outcome on production of passives by ESL and EFL learners. Although grammatical voice is time consuming to learn and challenging to teach, it is still teachable and also has instructional value. For example, the findings in Johnson and Lyddon (2016) confirmed the challenge of teaching the various aspects of voice to L2 learners with low proficiency level, however, they also affirmed that significant gains were made by the learners. Bailey (2016) examined academic writing of advanced English language learners (ELLs) in terms of three corpus-identified grammatical features (passive voice, reduced relative clauses, and modal *would*). Advanced L2 learners in an Intensive English Language Program (IELP) were divided into experimental group, who received extended instruction on these three grammatical features, and control group, who received the standard instruction students typically receive at IELP. Independent t-tests did not reveal significant differences between two group essays on the frequency of passive voice. However, the results showed that the amount of passive uses between two groups created quantitative difference (the experimental group used more passives than the control group), but the difference was not enough to have a statistically significant difference.

Teaching implication should concentrate on approaches to improve L2 learners' use of passive voice. In Xiao, McEnery, and Qian (2006), they explored the distinction of passive

constructions between English and Mandarin and concluded that passive constructions in the two languages were contrasted in a structured way, indicating that passives instruction was necessary for Chinese L2 learners. Moreover, Kim and McDonough (2016) investigated 82 Korean undergraduate EFL learners' production of English passives using prime repetition. Students were divided into three groups (1. with passive primes and repetition, 2. without passive primes, 3. with active primes) and participated in three activities over four weeks. All three groups received prompts (e.g., *books/*write) to facilitate sentence production, whereas only the repetition group received the passive primes (e.g., *the counter was wiped by Bob/Bob wiped the counter*). During each activity, participants in group 1 repeated the passive primes prior to generating sentences using prompts, students in group 2 did not receive any passive primes thus no repetition prior to producing sentences from the given prompts, and group 3 received active primes instead of passive primes. To compare learners' production of English passives among the three treatment groups, a pre-test, an immediate post-test and a delayed post-test were conducted. The results of one-way ANOVA showed significant differences of passive production among the three priming conditions ($F_{(2, 47)} = 52.496$, $p < 0.001$), in which the mean passive production of group 1 (M = 27.21) was much higher than group 2 (M = 20.71) and group 3 (M = 0) and the mean active production of group 1 (M = 8.57) was lower than group 2 (M =9.53) and group 3 (M = 35.53). The researchers also found statistically significant main effect among three testing time and groups ($F_{(2, 94)} = 45.77$, $p < 0.001$), and statistically significant interaction effects between time and group ($F_{(4, 94)} = 14.51$, $p < 0.001$).

These results indicated the passive primes and repetition could increase learners' passive production (group 1 pre-test M = 0.29, immediate post-test M = 8.43; group 2 pre-test M = 0.12, immediate post-test M = 4.47; group 3 pretest M = 0.05, immediate post-test M = 0.11), and

learners were able to retain the ability of producing passive sentences. The findings suggested that communicative priming activities may facilitate the development of English passives. Therefore, in order to produce a passive structure, "a learner must learn that certain meaning relationships (patient having something done to it by an agent) map onto certain functional elements (subject and oblique objects), which map onto critical features of constituent sequences" (Kim & McDonough, 2016, p. 338). Nevertheless, it should be noted that even though participants' passive production increased due to the prime repetition, the researchers did not provide any information regarding accuracy of their passive production. Therefore, it is unclear that whether prime repetition can contribute to the correctness of constructing passives. In terms of teaching methods, Lee (2007) also promoted an approach of textual enhancement (underlining, color-coding, bold facing, capitalizing, italicizing or using different fonts), which improved Korean EFLs acquisition of passive forms.

For Chinese L2 learners, direct language transfer also influences their choices of the passive voice use since the passive voice is usually used in academic writing, and therefore Chinese ELs are often found to use the passive structure in academic writing. For instance, Prystupa's (2017) study reflected a high frequency of passive usage by Chinese L2 learners. However, English teachers should be aware that there are often some sentences that sound perfectly natural in Chinese turn out to be awkward when translate into English. For example, it is perfectly natural to produce a passive sentence in Chinese like (5) when the emphasis is on the object, while it is awkward and even unnatural to express the same information using English in (6).

(5)　　蛋糕　　被　我　吃　　了

　　　　dangao　bei　wo　chi　　le

　　　　cake　　　by　me　eaten　　past tense particle

(6)　　The cake was eaten by me.

The English translation (6) is not grammatically erroneous, and L2 learners who have learned the passive structures can produce the same kind of sentences. However, direct language transfer and insufficient input might cause the production of (6), so English teachers need to see L2 learners' need and intentionally teach them the nature utterances and the convention of voice choices in academic writing.

For Japanese L2 learners, passive constructions might be mistakenly produced because in Japanese, intransitive verbs can be made passive (which is called an adversative passive) due to many transitive and intransitive equivalents (Machida, 2002). However, in English, only transitive verbs can be converted to the passive voice, whereas the exceptions of transitive verbs, such as *resemble* and *cost,* cannot be used in the passive voice. Grammatical rules in both languages are complex enough to memorize, thus instructions on vocabulary is also necessary for Japanese L2 learners when teaching passive voice in class.

For Dutch L2 learners, studies that looked into the passive production found that passive constructions were orally formulated slower than actives, and that Dutch NSs who learned English demonstrated this disfluency when produce passive constructions (Ferreira, 1994; Segaert et al., 2011). Mirdamadi and De Jong (2015) adopted Ferreira (1994)'s three words verbal elicitation (e.g. "the mechanic", "the car", "to repair") approach and conducted a study to investigate the effect of producing active and passive sentences on Dutch and English utterance fluency. The researchers predicted that ESL learners would have more difficulty in producing a

passive than an active sentence, even for those who were able to use the passive construction correctly. However, the results did not support this hypothesis. In contrast, the researchers found that the effect of complexity on disfluency was stronger in the L1 compared to the L2. They explained the result by considering the fact that participants had much more response times in the active production in the L2 than in L1 speech, thus the discrepancy in length of response times between actives and passives became shorter in L2 speech than in the L1, which led to a less strong effect of syntactic complexity on hesitation occurrence. Although there was no statistically significant difference between actives and passives on disfluencies in the L2, both active and passive utterances took longer time to process and the syntactic complexity of passives did cause more hesitations both in the L1 and L2. Therefore, English teachers should pay more attention to passive voice instruction.

In order to give a better idea of the passive constructions in different languages, Table 4 provides information of passive constructions in seven other languages, making comparison to the constructions in English, so that it is clear to both teachers and learners what kinds of errors can be easily made and how to avoid those errors.

Table 4.

Passive Constructions in Other Languages (Adopted from Folse, 2016)

Language	Arabic	Chinese	French
Notes on Passive Voice	1. Active voice dominates. Passive voice seldom uses. Thus, ELs tend to avoid passive voice in their writing or speaking. 2. Active and passive forms look similar but sound different, so Arabic-speaking students might be easily confused by active and passive forms in English. 3. *By* +agent is expressed by the preposition *b-*. The agent is usually a thing not a person in passive constructions.	1. The passive form in Chinese does not follow the *be + past participle* formula. 2. The agent is marked by *bei*, which is consistent with the *by* + agent structure.	1. French has a similar passive construction as *be + past participle* in English. However, the past participle must agree with the subject in number and gender. 2. French uses reflexive verbs to express passive voice. 3. The use of the pronoun *on*, which translates as "one", is another common way to express passive voice in French. For example, *on parle français ici* means "French is spoken here", but literal translation is "one speaks French here".

Passive Constructions in Other Languages (Adopted from Folse, 2016)

Language	Japanese	Korean	Russian	Spanish
Notes on Passive Voice	1. There is no *to be* or a past participle in Japanese passive forms. 2. The subject in a passive construction is usually animate. 3. The *by* + agent is expressed by *ni,* which usually means "in".	1. There is no *to be* or a past participle in Korean passive forms. 2. The *by* + agent is often rendered in English by the word *from*.	1. There is no *to be* or a past participle in Russian passive forms. 2. Passives in progressive and perfect tenses could be affected by modifications to the verb because these are indicated by separate verbs in Russian.	1. There are two principles of passive forms: a. Spanish uses reflexive verbs to express passive voice. *Aqui se habla español* means "Spanish is spoken here", but the literal translation is "[it] speaks Spanish to itself here". b. Spanish has a similar passive construction as *be* + *past participle* in English, however, there are two verbs for *to be*, which may confuse Spanish students. Also, the past participle must agree with the subject in number and gender. 2. The subject of a passive voice sentence must be the direct object of the verb in active voice.

Corpus Linguistics

A corpus refers to a computerized systematic collection of authentic (naturally occurring) texts from both written and spoken language. Corpus linguistics is an approach of conducting linguistic analyses on authentic discourses from these computerized corpora. In the last two decades, corpus linguistics has become an important topic in language and education study. It has not only shed light on providing vocabulary information in terms of its overall occurrences, but also on specific disciplines and genres these words tend to occur and even how they act in relation to other words (e.g., collocations).

Because corpus linguistics analyzes real language samples as the words and structures are spoken or written by non-linguists, it tells us how a language actually operates, but not how someone "thinks" it should operate. Quantitative analyses are crucial for corpus linguistics studies. For instance, if we wanted to compare the language use for the word *hard* in the meaning of *not easy* versus *not easily broken* as well as the collocations co-occur with each of this adjective, corpus linguistics can tell us the relative frequency of the word *hard* used for each meaning and for the co-occur collocations. Thus, corpus linguistics focuses on the descriptive rather than the prescriptive. However, Hunston (2002) put forward limitations of corpus linguistics in the study of language:

    a) Corpus can only show how frequent or infrequent an item is, but it does not give information if something is possible or not. Therefore, it can answer whether something occurred and how frequent it was, but it cannot answer if it is acceptable in a language. NSs can answer that question probably with intuition, but English learners will face difficulties on what is acceptable or not.

b) Corpus can give evidence but not information.

c) Corpus can only show its contents, and it presents language out of its own context.

To summarize, Hunston claimed that corpus linguistics can present what is available or what is known about a language, but it cannot explain why a language works the way it does (Hunston, 2002). Nonetheless, Hunston's argument collapses itself due to an extensive use of qualitative method in corpus-based studies (Deroey & Taverniers, 2011; Hasko, 2012). Qualitative corpus analysis is an important methodology for conducting investigations on "real" language data to interpret in-depth linguistic phenomena, thereby great amount of effort has inevitably been devoted by means of qualitative corpus-based studies (Biber et al., 2004). While corpus data can be especially helpful to nonnative language teachers (Tsui, 2005) who may not have enough English proficiency to teach certain semantic and syntactic points, Biber and Conrad (2001) pointed out how NS teachers can also benefit from corpus findings.

By far, the *Longman Grammar of Spoken and Written English* (Biber et al., 1999), which consists of 40 million words of British and American English texts, is one of the most noted and influential corpus works. This phenomenal corpus contains various genres including conversation, fiction, newspapers, academic prose, nonconversational speech, and general prose. It also reports on lexicogrammar (the combination of syntactic structures with lexical items, see Liu & Nelson, 2016), such as the most frequently used past participles in passive voice by genre, and the most frequently used noun suffixes (*-tion, -ity, -ism, -ness*).

Generally, a good corpus comprises a minimum of 1-3 million words when looking for high frequency words for quantitative analysis (Brysbaert & New, 2009). Table 5 lists existing English corpora with at least a million words, in order to provide corpora information for researchers who are interested in quantitative corpus linguistics study.

Table 5.

Summary of Existing Corpora with At Least 1 Million Words

Corpus	Language Variety/ies	Text type(s)	Total word count
ACE (The Australian Corpus of English)	AusE	Written	1 million
ANC (American National Corpus)	AmE	Written + spoken	20 million
ARCHER (A Representative Corpus of Historical English Registers)	AmE & BrE	Written + spoken	1.7 million
BNC (British National Corpus)	BrE	Written + spoken	100 million
Brown	AmE	Written	1 million
The Collins Wordbanks Online English corpus	AmE & BrE	Written + spoken	56 million
COCA (The corpus of Contemporary American)	AmE	Written + spoken	520 million
COHA (The Corpus of Historical American English)	AmE	Written + spoken	400 million
Early Modern English Medical Texts	BrE	Written	2 million
FLOB (The Freiburg-LOB corpus of British English)	BrE	Written	1 million
Frown (The Freiburg-Brown corpus of American English)	AmE	Written	1 million

Summary of Existing Corpora with At Least 1 Million Words

Corpus	Language Variety/ies	Text type(s)	Total word count
ICAMET (Innsbruck Computer Archive of Machine-Readable English Texts)	BrE	Written	5.6 million
ICLE (International Corpus of Learner English)	LearnerE	Written	2.5 million
Kolhapur	IndE	Written	1 million
Lampeter Corpus of Early Modern English Tracts	BrE	Written	1.1 million
LOB (The Lancaster-Oslo/Bergen Corpus of British English)	BrE	Written	1 million
MICASE (The Michigan Corpus of Academic Spoken English)	AmE	Spoken	1.8 million
PCEEC (Parsed Corpus of Early English Correspondence Sampler)	BrE	Written	2.2 million
WSC (Wellington Corpus of Spoken New Zealand English)	NZE	Spoken	1 million
WWC (Wellington Corpus of Written New Zealand English)	NZE	Written	1 million

Corpus-based Evidence of Common Passive Constructions

Researchers can indicate which passive constructions are the most common ones using corpus-based analysis. For instance, corpus evidence has shown that *get*-passive is less common in both spoken language and formal language settings. In a corpus-based study using the Cambridge and Nottingham Corpus of Discourse in English (CANCODE), Carter and McCarthy (1999) found only 139 *get*-passives of the form *X get + past participle* (by Y) (e.g., if you get yourself locked out) out of a 1.5-million-word sample. Biber et al. (1999) also confirmed that rare instances of *get*-passives were found in spoken language. Likewise, *get*-passives were found to have far fewer instances than *be*-passive in two other British English corpora: The Freiburg-LOB Corpus of British English (FLOB) and the British National Corpus (BNC) (Xiao et al., 2006). Nevertheless, the results demonstrated in Mindt (2000) indicated a different situation in which *get*-passives are mostly used in spoken discourse.

With regard to the frequency of *get*-passives in written language, Collins (1996) claimed that *get*-passives were also rarely seen in formal written form. In that study, he examined a corpus of 5.25 million words with mixed spoken and written language and reported altogether only 291 *get*-passives. After examining the selected corpora in his study, Collins found that central *get*-passives were the most prevalent construction (28.8 per cent usage rate) and psychological *get*-passives (14.7 per cent usage rate) were the weakest representation among the five subclasses.

Moreover, Hundt (2001) investigated all forms of *get*-constructions, passive-like constructions and their frequencies from a self-developed corpus by collecting data from the ARCHER corpus, the Brown corpus, the LOB corpus, the Frown corpus and the FLOB corpus, which shed light on the patterns of how *get*-passives were used in American and British English throughout the decades. Results showed that the frequencies of *get*-construction were different

between the American and British English. In general, the *get + past participle* structure was found to function more as an adjective to the doer than to the action, and this structure appeared to co-occur more frequently with unanimated objects.

When comparing a native corpus (*BNC Baby*) with a learner corpus (*International Corpus Network of Asian Learners of English, ICNALE*) in terms of the frequency of the passive structure *be + past participle*, Granger (2013) found a noticeable variety of passive uses among the learner populations, however, all the L2 learner populations underuse the passives. The relative frequency of the passive forms in *BNC Baby* was 1436 in every 100,000 words, while the highest frequency of the passive forms in *ICNALE* was only 942 (L2 learner population from Philippine), which has 34 per cent less passives than the native corpus. The lowest frequency was only 574 (L2 learner population from Indonesia), which has 60 per cent less passives.

The corpus-based study Alvin (2014) conducted summarized the occurrence of the following passive forms, and the results indicated that the most commonly used forms were the basic and bare passives:

a)   Basic (be + Ven);

b)   Progressive (be + being + Ven);

c)   Perfective (have + been + Ven);

d)   Modal (modal + be + Ven);

e)   Modal perfective (modal + have + been + Ven);

f)   To-infinitive (to + be + Ven);

g)   Non-finite –ing (being + Ven);

h)   Bare (Ven)

Corpus-based Evidence of Lexical Associations of the Passive Forms

It is acknowledged that active and passive constructions have a strong association with lexical choices (Biber et al., 1999; Granger, 2013; Gries & Stefanowitsch, 2004). Some verbs exhibit strong association of passive usage, while others conventionally repel the use of passive constructions. Granger (2013) summarized the passive ratio of 12 verbs in the academic section. The passive ratios of those 12 verbs ranged from the highest 68.2 per cent (oblige) to the lowest 0.8 per cent (*want*), which offered an incredibly powerful piece of information in support of a strong association between passive constructions and verb uses. In addition, in-depth analyses of passive constructions and verb uses revealed that the same polysemous verbs also had different passive ratios. For example, the five meanings of the polysemous verb *meet* had respectively five different passive ratios, and the observation showed even more discrepancies: 0% ('come together'), 2% ('come across'), 21% ('standard'), 72% ('arrival') and 81% ('reaction') (Haas, 2008). Therefore, with the help of corpus data, English teachers can teach "with examples of prototypical uses of the passive and lists of verbs that often passivize, rather than the 'can't do' approach that is still dominant in many current grammars." (Granger, 2013, p. 13).

It is important to teach L2 learners the correct passive forms. Likewise, it is even more important to teach them some of the most commonly used vocabulary for these grammar forms and functions. These grammar items are easy to teach, and the functions are finite, but the vocabulary is infinite and requires effort to memorize (e.g., the differences of passive constructions between Japanese and English transitive and intransitive verbs). Folse (2015) provided lists of past participles frequently used in various constructions, summarized in Table 6.

Table 6.

The Most Commonly Used Past Participles in Canonical Passives, Get-passives and as

Adjectives (Adopted from Folse, 2015, pp. 135, 147, and 144)

15 Past Participles Frequently Used in Passive Voice	12 Past Participles Frequently Used in Get-Passive	20 Past Participles Frequently Used as Adjectives
1. be used	*1. get rid (of)*	*1. unidentified*
2. be made	*2. get married*	*2. concerned*
3. be done	*3. get started*	*3. involved*
4. be found	*4. get caught*	*4. supposed*
5. be called	*5. get paid*	*5. interested*
6. be taken	*6. get involved*	*6. united*
7. be expected	*7. get done*	*7. married*
8. be seen	*8. get dressed*	*8. used*
9. be considered	*9. get hit*	*9. increased*
10. be born	*10. get elected*	*10. surprised*
11. be given	*11. get fired*	*11. limited*
12. be based	*12. get arrested*	*12. tired*
13. be killed		*13. so-called*
14. be told		*14. armed*
15. be designed		*15. broken*
		16. lost
		17. advanced
		18. complicated
		19. unknown
		20. scared

Yannuar et al. (2014) analyzed the frequency of active and passive voice constructions of selected Biber's (2004) stance verbs from a corpus named C-SMILE (Corpus of State University of Malang Indonesian Learners' English), which comprised of 1,587,059 words from a collection of 124 undergraduate theses from the English Department of State University of Malang. The results revealed that the most frequently used passive constructions of stance verb was *expect,* which occurred 555 out of 645 times (frequency rate 86 per cent) in the form of present and past tense. Figure 1 shows the most frequently used passive stance verbs from the corpus in Yannuar et al. (2014). Overall, the active voice constructions occurred 29.6 per cent more than the passive voice constructions, and the frequent occurrence of passive voice construction was 35.2 per cent, which was higher than the average frequency rate of passive constructions in academic writing from the literature.
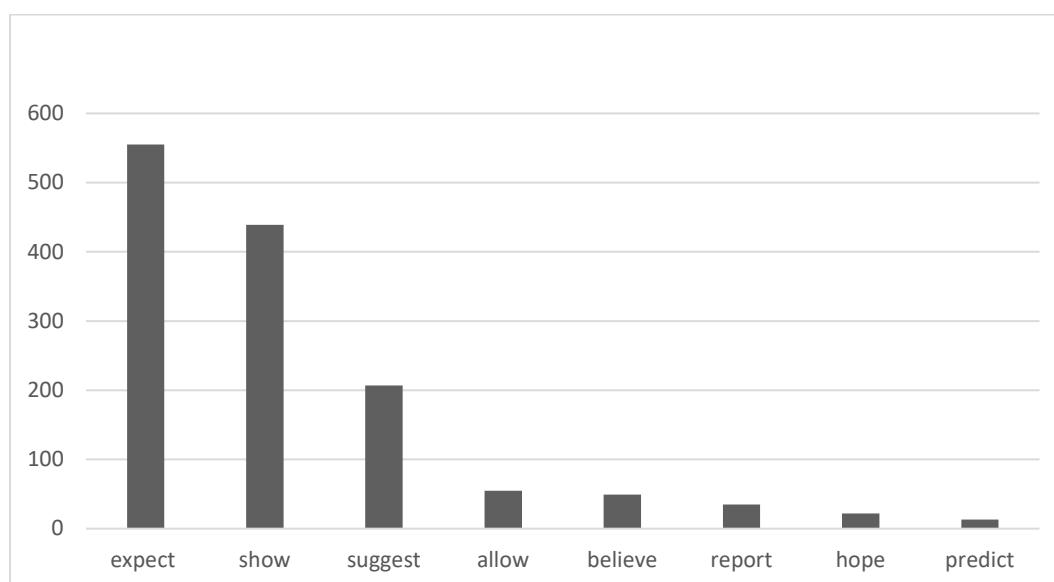
Figure 1. Frequency of passive voice constructions of stance verbs from C-SMILE

Millar et al. (2012) assessed how passive voice was used and the effect of journal guidelines on the use of passive voice. The researchers compiled a corpus of 297 primary research articles from five top medical journals and compared the New York Times corpus

(NYT), a corpus of General English, along with selected medical articles from the Randomised

Controlled Trials (RCT) corpus in 2005 in order to investigate the stylistic writing discourse

differences between both genres. They first identified all the verbs in the corpus and then

analyzed the occurrence of the passive voice in both the NYT corpus and RCT corpus. Finally,

they found 20 common verbs associated with passive voice constructions (See Figure 2). They

noted that the usage of the passive voice was unavoidable, and was sometimes necessary to focus

the reader's attention on important points. Even though the medical journals guidelines

influenced the usage of the passive voice in some extent, the guidelines needed to be clearer as

when to use the passive voice or active voice (Millar et al., 2012).
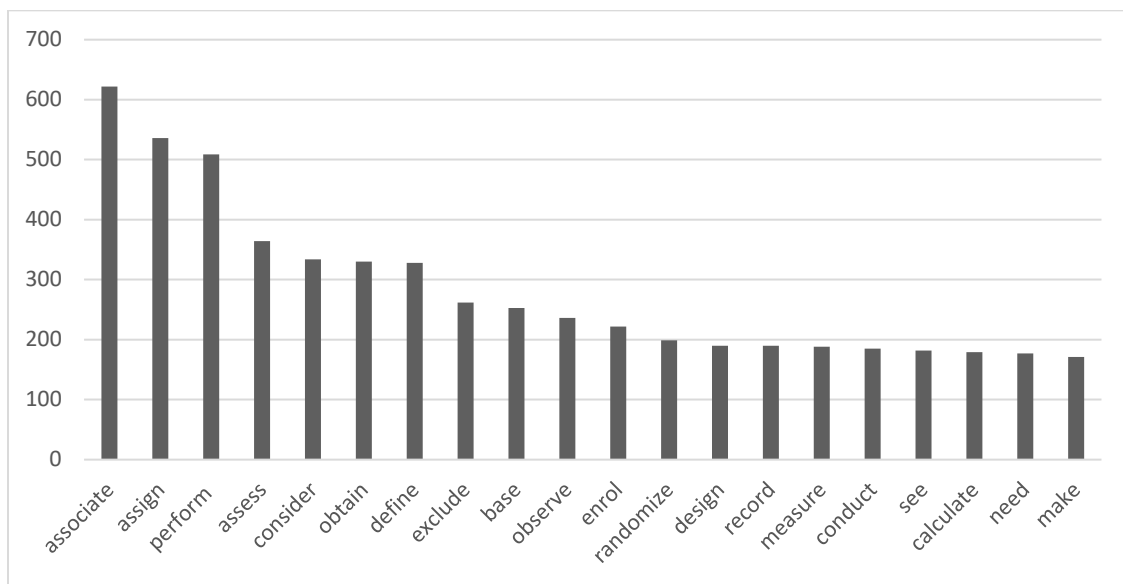
Figure 2. Verbs strongly associated with passive constructions

Although plenty of previous research has indicated high proportions of passive structures

in scientific writing, the frequency of passive usage in STEM content in academic discourse is

still unknown. Therefore, the present study aims to contribute complementary corpus-based data

to this discussion by investigating STEM textbooks.

Conclusion

The chapter of literature review has pointed out some important aspects of grammar learning and the controversy upon whether passive voice should be used or avoided. It appears that the main adverse claim against passive voice concerned with the clearness of statements, which accuses the meaning of passive voice to be indirect and/or less forceful than an active alternative (Albert 2000; Day 1989; Hall 1994; Matthews et al. 1996; Sheen 1982; Strunk & White 2000; Taylor 2005). However, arguments that passive voice will necessarily result in less powerful writing are also open to investigation. As presented earlier, passive forms are commonly found in the Methods section of research articles, and according to Ding (2002), the reason why passive voice is frequently found in such section is because authors are inclined to claim two social values in science, namely falsifiability of science and cooperation among scientists.

Grammar is indispensable in SLA and is one of the crucial components to construct native-like language structure in any context. The fact that the rubrics of all English proficiency tests for writing, such as TOEFL and IELTS, underlines grammatical errors are important assessment criteria of an effective writing. Grammar serves as a beautifier to language and without grammar, verbal expression cannot be perfect. Among many grammatical features in English grammar, passive voice seems to be the most disputable one between scholars and educators. Additionally, STEM is a field that is very necessary for educational subjects in order to prepare STEM professionals that meet the needs for current labor market. Some studies have looked at journal articles, some investigated students' writings, whereas very few has examined textbooks. In fact, most studies looked at journal articles, however, college students are not trying to study journal articles but read and study academic textbooks. Therefore, the present

study focuses on investigating the current norm of passive voice in STEM writing in

undergraduate textbooks.

CHAPTER 3: RESEARCH METHODOLOGY

Given that there has been argument about whether the passive voice should be avoided in academic writing, especially in scientific areas, the importance of this grammar point need to be clarified for English instructors as well as grammar book authors. For English instructors who prepare L2 learners to study in a college classroom in the United States, knowing how important the passive voice is and how often it is used in undergraduate textbooks can inform curricula of English reading and writing. This could adjust language practices to increase reading comprehension and make writing skills close to authentic in English for academic purposes programs. As reviewed in Chapter Two, studies have revealed that the passive voice is indeed used frequently in the scientific discourse. However, since no study has investigated the passive usage in undergraduate STEM textbooks, the present study attempted to fill this research gap.

This chapter delineated the quantitative method for exploring passive usage in terms of frequency, constructions and dispersion in undergraduate STEM textbooks. Areas of methodology listed below was discussed:

1. A general orientation of the research design.

2. A general description of the selected textbooks and sampling procedures for investigation.

3. A detailed description of data collection procedures, which included the variables it was designed to measure, inter-rater reliability, and the steps that the researcher followed in using it.

4. The instrument used to identify the most frequently occurred verbs that are strongly associated with passive constructions.

5. Data analysis procedures and statistical analysis procedures implemented to obtain results.

Orientation to Research Design

Based on the above research goals, the present study undertook a quantitative, descriptive and inferential statistical analyses. Twenty undergraduate STEM textbooks were analyzed to gather total passive usage in the selected textbooks so that the frequency of passive usage in terms of passive constructions and dispersion would be calculated. Passive constructions and verb dispersion were ranked and organized according to their frequencies. To investigate the most frequently used verbs in passive forms, an online corpus software, Compleat Lexical Tutor (Cobb, 2017) was used. For the overall quantitative analysis, a non-experimental descriptive design was mainly utilized to answer the three main questions of the frequency, constructions and dispersion in the selected textbooks.

For the second research question, a MANOVA test was used for determining whether the frequency of passive constructions differed among disciplines. Figure 3 provided a visual outline of the research design, which was further elucidated in the following sections.
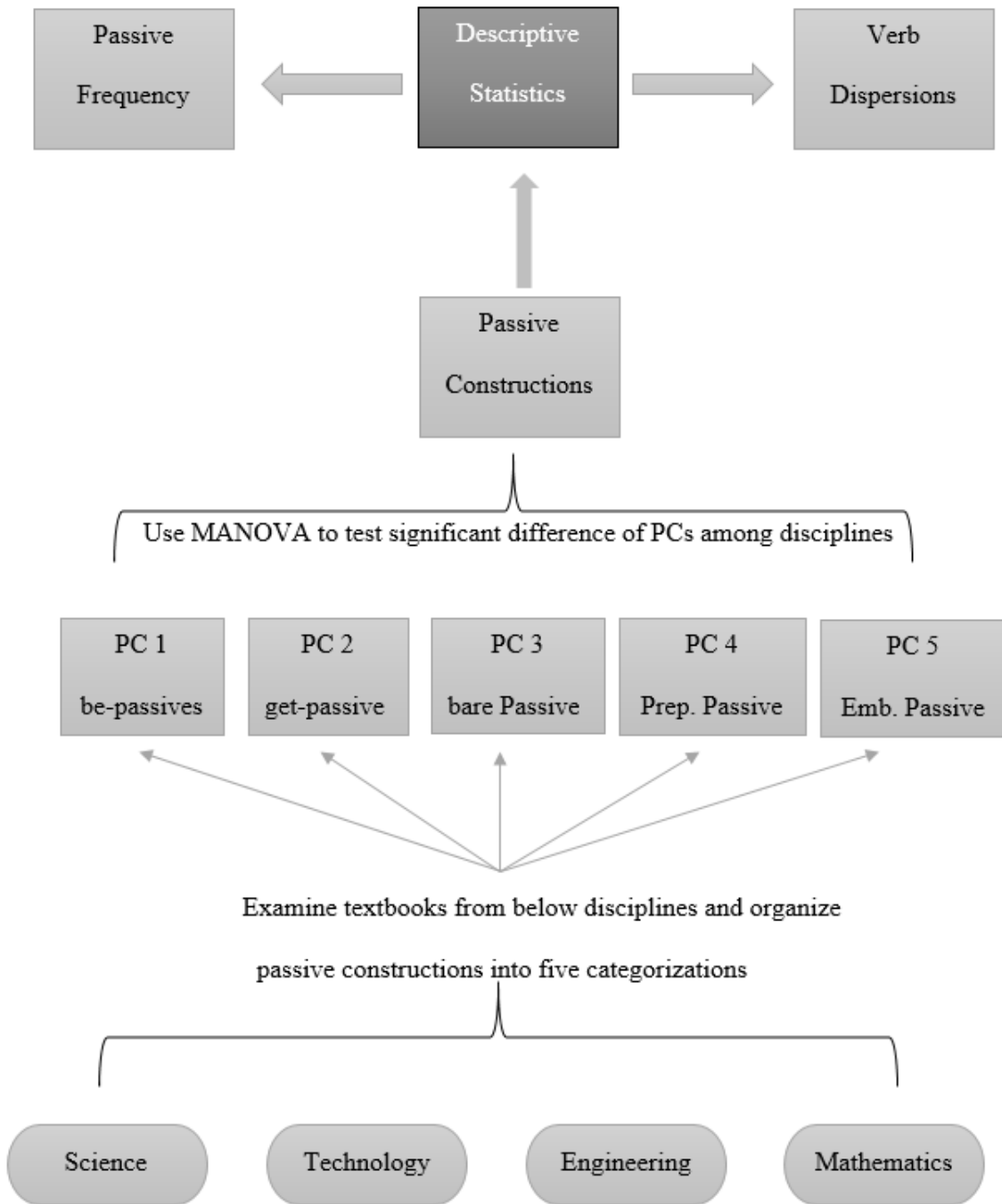
Figure 3. Orientation to research design

Sampling Procedures

This study followed a convenience sampling method using undergraduate STEM courses at a large university in southeastern United States as detailed in the selected textbooks section above. The convenience sampling method was selected where samples of the target population suited the research objectives (Fraenkel & Wallen, 2009). After the courses were determined, a random and systematic sampling (Kalton, 2017) was used to select every other ten pages from the examined textbooks until each book reached around 50,000 words. In this way, relatively more chapters were included into analysis to prevent topic bias. As a result, twenty textbooks were selected from the undergraduate catalog with five books for each discipline and approximately 50,000 words in analysis for each book to yield a corpus that consisted of at least 1 million words when looking for high frequency words for quantitative analysis (Brysbaert & New, 2009). Procedures on how this study was going to reach these figures were presented as follows.

Sample Size Determinations

There are around 400 textbooks used in all STEM majors at the targeted university. To determine an appropriate sample size for a MANOVA test, Cohen's (1992) statistical power analysis was consulted. Therefore, G*Power was conducted to determine a sufficient sample size at a .05 level of significance, a power of .80 so that a medium effect size of .40 could be achieved (Faul, Erdfelder, Buchner, & Lang, 2013). As a result, a desired sample size of at least 18 textbooks was needed to conduct statistical analyses. In order to collect adequate running words for the corpus, and since eighteen cannot be divided by four with no remainder, twenty textbooks were randomly selected for the study. In addition, having five books for each discipline also prevents subject bias of the study, given that the language expression used in a book is limited to specific author(s) and editor(s).

<center>**Recruitment of Subjects**</center>

Selected Textbooks

There were no human participants involved in the present study. Data was collected from twenty undergraduate STEM textbooks with five books for each area. Ideally, the subject bias would reduce if multiple majors were chosen, thus data included five courses from five different majors of one area. As for those areas that comprised less than five majors, five courses from the existing majors were selected to diminish subject bias. Consequently, there were five books for each discipline. The university *Find Course Materials* website was used to search for the required textbooks of the selected courses. If the number of the required textbook for one course exceeds one, the one that has adequate length that satisfies the sample size of the current study would be chosen. All textbooks used for this study were electronic books. For the books that did not have e-version available, previous edition was used. The differences between latest and previous editions were checked to ensure the contents were consistent. Only if there were no distinct differences between two editions could the previous one be used in this study.

Raters

The researcher independently analyzed all selected textbooks; however, to ensure the reliability of the examined data, a second rater, who is a professor in TESOL, was invited to analyze a sample text randomly selected from one of the examined textbooks prior to formal data collection.

<center>**Instrumentation**</center>

Compleat Lexical Tutor

To examine the most frequently used verbs in passive structures, an online computer software, Compleat Lexical Tutor (CLT) (Cobb, 2017), was employed. The CLT is a

<center>78</center>

comprehensive website that was primarily designed for learning vocabulary, and then it was

expanded to include a wide range of resources for learning and teaching vocabulary and

grammar. It is freely available online and allows researchers to either upload files that do not

exceed maximum size limit or copy and paste texts into the program for analysis

(http://www.lextutor.ca/freq/eng/). The majority of the programs on CLT have been noted to be

validated and have been used in empirical peer-reviewed papers (Cobb, 2004). The CLT is

mainly divided into three sections: Learners, Researchers, and Teachers. In the Researchers

section, there are several notable applications that can be utilized to examine word frequency

from texts, such as *VocabProfile*, *TextLex Compare*, *Frequency* and *Concordance tools*.

VocabProfile has three subcategories, namely *VP-Classic*, *VP-kids*, and *VP-Compleat*. *TextLex*

Compare allows researchers to compare word frequency in each text as well as shared word

frequency appeared in both texts. Additionally, both *Frequency* and *Concordance tools* can

process with links to English and French programs. For the objective of the present study, the

researcher used *Frequency* to obtain frequency counts of passive verbs. The Web Frequency

Indexer v 1.3 generates a word frequency list after uploading a text file into the program, thus it

is useful to obtain results for the third research question. With a column of word frequency list

in hand, the researcher can easily locate where the number of frequency drop dramatically so that

a word list of verbs that are strongly associated with passive forms in the STEM fields can be

created.

Data Collection Procedures

The purpose of the present study was to examine how frequently the passive voice is used

in undergraduate STEM textbooks. Required courses of the undergraduate programs that were

STEM related at a metropolitan university in Florida were chosen due to their generalizability across the target undergraduate programs at the university.

Given that there was no human being involved in the study, the study was IRB exempt. Thus, data collection procedures occurred in the following steps:

1. STEM related undergraduate majors offered at the university, organized in Table 7. The researcher used Random Integer Generator (https://www.random.org/integers/) to determine five majors for Science and Engineering because both fields consisted of more than five majors.

2. The undergraduate catalogs were screened in order to select courses for the study. Only core requirements were selected into a course pool. Random Choice Generator (https://www.textfixer.com/tools/random-choice.php) was used to select courses. The same major that comprised different tracks were categorized as one in the pool.

3. Therefore, this process indicated twenty courses in total, namely Biology, Statistics, Physics, Chemistry, Social Sciences, Microbial Metabolism, Molecular Biology I, Operating System Concepts, Security in Computing, Computer Architecture, Computer Science, Construction Engineering, Environmental Engineering, Industrial Engineering, Aerospace Engineering, Advanced Calculus I, Matrix and Linear Algebra, Linear Algebra, Abstract Algebra, and Logic Proof in Mathematics. The selected majors and textbooks were listed in Table 8.

4. After the courses are determined, the corresponding textbooks were found via the course materials website (http://ucf.bncollege.com/webapp/wcs/stores/servlet/TBWizardView?catalogId=10001&langId=-1&storeId=16552) or from the courses' syllabi.

5. Looking for textbooks online and/or at campus library and bookstore. Priority of textbook version was first electronic and then printed book, because e-books facilitate data entry process and total word count of a book would be more precise. Due to the selection of the courses were random, the corresponding textbooks were therefore randomly selected.

6. The present study followed the procedure of random and systematic sampling and examined texts of every other ten pages of the textbook until reached 50,000 words in each book. Altogether, there were approximately 1 million words included in the corpus.

7. To determine the total word count of the analyzed texts, the researcher randomly selected a page from each book. On the selected page, picked ten consecutive lines that have typical lengths of the book. Counted the words on those ten lines and then divided by ten to obtain estimate number of words per line. Then counted the number of lines on the page and multiplied by the number of words per line to get estimate words per page. Finally, multiplied this number by the total number of pages in the book. Discounted a little to account for blank or partially full pages. Hyphenated words, such as high-impact, were counted as one word. According to Conner and Upton (2004), a specialized corpus study, which contains 20,000 words, is considered a sufficient size. Therefore, the size of the present study was considered as sufficient.

8. Prior to formal data collection, a pilot study was conducted in order to increase reliability of data collection. Randomly choosing a 1,000-word text from one selected book, the researcher and the second rater analyzed the text and circled all the passive forms. Then the two analyses were compared to see their consistency. After

the raters reached a consensus by using the following coding scheme (shown in Table 9.), the second rater stepped out and the researcher resumed data collection.

9. All the main verbs and passive voice were individually examined. After all textbooks were analyzed, the percentage of passive usage was then calculated by dividing the total number of passive verbs by the total number of main verbs. Transitive and intransitive verbs were coded differently as main verbs, because there were two sets of denominators, one excluded intransitive verbs and one included both transitive and intransitive verbs in the denominator. Each passive construction was recorded and organized in an excel sheet so that the numbers of passive constructions appeared in the texts could be obtained.

10. After all textbooks were analyzed, all the passive sentences were recorded into an Excel sheet, and they were organized in terms of their constructions. Then the number of passive constructions in each discipline would be obtained. Next, data was entered into the computer statistical analysis program, called Statistical Package for the Social Sciences (SPSS) 24, to conduct a MANOVA test to find out if statistically significant difference of passive constructions among the four disciplines could be found.

11. Similarly, using the organized record from the second research question, the data was submitted to CLT Web Frequency Indexer v 1.3 in order to determine what verbs tended to be strongly associated with the passive forms. Before inserting data into SPSS 24 for the MANOVA test, the frequency of each passive construction used was first counted and recorded in an Excel sheet (e.g., be-passive, appeared 138 times).

Table 7.

STEM Programs

Science	Technology	Engineering	Mathematics
1. Actuarial Science	1. Biotechnology	1. Aerospace Engineering	1. Mathematics
2. Biology	2. Information Technology	2. Civil Engineering	
3. Chemistry		3. Computer Engineering –	
3. Chemistry – Biochemistry Track		Comprehensive Track	
4. Forensic Science – Biochemistry Track		3. Computer Engineering - Digital Circuits Track	
4. Forensic Science – Chemistry Track		4. Computer Science	
5. Mathematics (Science Track)		5. Construction Engineering	
6. Physics		6. Electrical Engineering – Communications and Signal Processing Track	
7. Psychology		6. Electrical Engineering – Comprehensive Track	
8. Social Sciences		6. Electrical Engineering – Power and Renewable Energy Track	
9. Sport and Exercise Science		6. Electrical Engineering – RF and Microwaves Track	
10. Statistics		7. Environmental Engineering	
		8. Industrial Engineering	
		9. Mechanical Engineering	
		10. Photonic Science and Engineering	

Table 8.

Selected Majors and Textbooks

Science	Technology	Engineering	Mathematics
1. Biology – *Essential Cell Biology*	1. Microbial Metabolism – *Physiology and Biochemistry of Prokaryotes*	1. Computer Science – Programming Erlang	1. Advanced Calculus I - *Advanced Calculus*
2. Statistics – *Applied Regression and Other Multivariable Methods*	2. Molecular Biology I – *Molecular Biology of the Gene*	2. Construction Engineering – *Estimating in Building Construction*	2. Matrix and Linear Algebra – *Linear Algebra with Applications*
3. Physics – *Principle of Electronic Materials and Devices*	3. Operating System Concepts – *Operating Systems*	3. Environmental Engineering – *Principle of Geotechnical Engineering*	3. Linear Algebra – *Linear Algebra*
4. Chemistry – *Fundamentals of Biochemistry*	4. Security in Computing – *Computer Security*	4. Industrial Engineering – *DeGarmo's Materials and Processes in Manufacturing*	4. Abstract Algebra – *The First Course in Abstract Algebra*
5. Social Sciences – *The Essentials of Political Analysis*	5. Computer Architecture Concepts – *Computer Organization and Design*	5. Aerospace Engineering – *Vector Mechanics for Engineering: Statics with Connect Access Card*	5. Logic Proof in Mathematics - *Mathematical Proofs*

Table 9.

Passive Coding Scheme

Code	Passive construction	Example
a.	*Be*-passive	*The house was built by his father.*
a1.	*Be*-passive non-finite-ing construction	*Our travel plan is being made.*
a2.	*Be*-passive infinite	*The main campus is going to be expanded.*
a3.	*Be*-passive perfect tense construction	*The Nobel Peace Prize ceremony has been held at three different places.*
a4.	*Be*-passive modal	*How the prizes should be awarded was discussed by the academy.*
b.	*Get*-passive	*He got fired yesterday.*
c.	Bare passive	*"28 injured in accidental detonation"* (*Los Angeles Times,* 5 July 2013)
d.	Phrasal verb passive	*The meeting was called off by the board.*
e.	Embedded passive	*The government had the case investigated by the police.*

Criteria of identifying main verbs and passive constructions

The overall principle of identifying main verbs in the selected texts in this study was to identify finite verbs, which has an expressed subject and can function as the root of an independent clause. Transitive verbs that include active and passive verbs, intransitive verbs, and command verbs were considered as main verbs. Phrasal verbs were considered transitive verbs in this study. The criteria of identifying passive constructions should be agreed and carefully followed.

However, not all the passive constructions founded within the main verbs would be counted. For example, concealed passives were not included in the passive repertoire due to their irregular passive structure (lack of passive designator) and uncommonly frequency of occurrences. Furthermore, due to the controversial debate of adjectival and stative passives regarding their syntactical functions and morphological features, they were not categorized as one independent variable but were embedded into the constructions listed as variables in Table 9. Additionally, fixed phrases, such as *to be born*, *it is called* and so on were also not counted as passive in this study. The reason for recruiting the regular passive constructions was not only that they were the most commonly used in expression (Alvin, 2014), but they also provided a more conservative result to the study.

Research Questions

The present study aimed to investigate the following research questions that based on the review of literature:

1. What is the percentage of passive voice usage in undergraduate STEM textbooks?

2. Is there any statistically significant difference of seven passive constructions between the four STEM disciplines? If so, what are the differences?

86

3. What are the most frequent verbs tending to be associated with passive form in those textbooks?

Hypothesis

The null and directional hypothesis for the second research question is as follows.

- H_0: There is no statistically significant difference of the common passive constructions between the four STEM disciplines.

- H_1: There is statistically significant difference of the common passive constructions between the four STEM disciplines.

The estimation that scientific discourse would demonstrate significantly higher amount of passive constructions was based on previous research studies that investigated the use of passive voice in scientific journal articles and the deficiency of related studies in the area of Technology, Engineering and Mathematics (Alvin, 2014; Atkinson, 1996; Bazerman, 1988; Ding, 2002; Harmon, 1992; Hundt et al. 2016; Rodman, 1994).

Data Analysis Procedures

In order to answer the research questions investigating the frequency of passive usage, passive constructions and verb dispersion of undergraduate STEM textbooks, the study mainly used descriptive statistics to present the frequency of passive usage, the most commonly used passive constructions and the most frequent verbs that were strongly associated with passive forms in the examined undergraduate STEM textbooks. Although descriptive research is the most basic quantitative research, it is a type that concerns making conscientious descriptions of educational phenomena, thus it is equally important in educational research (Gall, Gall, & Borg, 2007).

For the second research question exploring the difference of the use of passive constructions among the four disciplines, a multivariate analysis of variance (MANOVA) test was conducted. Tabachinik and Fiddell (2007) noted that MANOVA test was an optimal statistical technique when examining more than one dependent variables simultaneously. In the present study, there were four dependent variables involving in the analysis, which were the four examined disciplines.

Conclusion

This chapter detailed the research design, selected texts for the study, sampling procedures, operational research questions, and the data collection and analysis procedures. Due to the population of the current study excludes human being, the ethical considerations were not provided within the study. Chapter Four presented the results from the data collection and analysis procedures discussed above to answer the research questions. Following Chapter Four, the final chapter, Chapter Five discussed research and pedagogy implications along with future directions for research.

CHAPTER 4: RESULTS

This chapter presented the results of the current study that investigated the usage of passive voice usage in undergraduate STEM textbooks. The chapter first reviewed the research questions and then presented the data screening and assumption checks that were run prior to data analysis. Finally, this chapter presented the research findings from descriptive statistics and a MANOVA test, which provided answers to each research question.

Research Questions

The present study aimed to explore the overall passive voice usage in college textbooks in terms of the percentage of passive voice usage, constructions, and dispersion. The following three research questions guided the current study:

1. What is the percentage of passive voice usage in current undergraduate STEM textbooks?

2. Is there a statistically significant difference of seven passive constructions between the four STEM disciplines? If so, what are the differences?

3. What are the most frequent verbs tending to be associated with passive form in those textbooks?

Sampling Procedures

The data collection process lasted for two months from October 2017 to November 2017. The texts that analyzed in the corpus in the current study were collected from four areas of study, and each area comprised of five majors. Therefore, there were twenty selected textbooks examined in this study (Table 10).

To add to what we know about STEM instruction, one of the main goals of this current study was to examine a broad spectrum of STEM textbooks. Therefore, all four STEM areas, namely Science, Technology, Engineering, and Mathematics, were included. While several earlier studies focused on the area of Science (Alvin, 2014; Ding, 2002; Harmon, 1992), they neglected the other three. The inclusion of Technology, Engineering, and Mathematics helped add to the body of literature by identifying (a) the percentage of passive voice usage in all four areas of STEM, and (b) any significant differences in the amount of passive voice usage between the four areas of study, and (c) the most commonly used past participles within these areas.

In order to have a corpus that consisting of at least 1 million running words (Brysbaert & New, 2009), data was collected from every other ten pages of the textbooks until each book reached the minimum 50,000-word threshold. However, if the running words of one textbook was over the 50,000 threshold, data collection would still continue until the exact ten pages were included. As seen in Table 10, numbers of running words varied from textbook to textbook.

Table 10.

Numbers of Running Words in Textbooks

Discipline	Textbook	Running Words
Science	*Essential Cell Biology*	51,164
	Learning SAS by Example	50,339
	Principle of Electronic Materials & Devices	50,080
	Fundamentals of Biochemistry	50,011
	Getting from College to Career	50,100
	Physiology + Biochemistry of Prokaryotes	50,141
	Molecular Biology of the Gene	50,172
Technology	*Operating Systems*	50,237
	Computer Security	50,167
	Computer Organization and Design	50,091
	Programming Erlang	50,050
Engineering	*Estimating in Building Construction*	50,331
	Principles of Geotechnical Engineering	50,291
	DeGarmo's Materials and Processes in Manufacturing	50,245
	Vectors Mechanics for Engineers	50,199
	Advanced Calculus	50,107
	Linear Algebra with Applications	50,174
Mathematics	*Linear Algebra*	50,394
	First Course in Probability	50,273
	Mathematical Proofs	50,237

Description of the Sample

Descriptive statistics and the common measures of central tendency revealed that the mean length of the textbooks was 50240.15 running words ($SD = 240.47$; range, 50,011-51,164; Table 11). The mean length of textbooks in Science was 50,338.8 running words ($SD = 477.60$;

range, 50,011-51,164). In Technology, the mean text length was 50,161.6 running words (*SD* = 52.99; range, 50,091-50,237). As for Engineering, the mean length was 50,223.20 running words (*SD* = 108.72; range, 50,050-50,331). Last but not least, the mean length of textbooks in Mathematics was 50237 running words (*SD* = 108.18; range, 50,107-50,394; Table 11).

Table 11.

Running Words Length

	N	Range	Minimum	Maximum	*M*	*SD*
Total Corpus	20	1153	50,011	51,164	50,240.15	240.47
Science	5	1153	50,011	51,164	50,338.80	477.60
Technology	5	146	50,091	50,237	50,161.60	52.99
Engineering	5	281	50,050	50,331	50,223.20	108.72
Mathematics	5	287	50,107	50,394	50,237.00	108.18

Testing Statistical Assumptions

Prior to data analysis, data screening was needed to test normality, linearity, homogeneity of variance, and homoscedasticity for the variables in the study. There were originally nine dependent variables in total, including the common passive constructions, namely *be*-passive, *be*-passive non-finite-ing, *be*-passive infinite, *be*-passive perfect tense, *be*-passive modal, *get*-passive, bare passive, phrasal verb passive, and embedded passive. All the dependent variables were continuous level data.

First, the dependent variables were screened for outliers by checking standardized values (Z scores) and deleting the Z scores that were greater than 2. The initial data screening identified 89 cases as outliers out of a total of 298 cases among the dependent variables. Thus, there were 209 cases that proceeded to the next phase of data screening test.

The skewness and kurtosis values and Kolmogorov-Smirnov (K-S) tests were conducted to evaluate the normality of the nine dependent variables. When all data were included, the values were highly skewed, and the K-S test indicated significant deviation from normal distribution (Table 12). Data re-analysis excluding the outliers indicated acceptable skewness and kurtosis values; however, it also displayed non-normality distributions (Table 13). Nevertheless, according to Fidell (2013), interpretation of the K-S test should be prudent and can be inaccurate when data sets exceed 100 cases. Therefore, histograms and Q-Q plots were all taken into consideration to visually check for normality. The results revealed proximate normal distributions with the exception of *get*-passive and embedded passive. Both variables were positively skewed and had high kurtosis, indicating these two passive constructions rarely appeared in the examined texts. There was no data transformation to achieve normalized distribution because the benefit of normalizing transformation is low, and it was not recommended by the scholars (Glass, 1972). In addition, it would be hard to interpret the results of the transformed data, thus no data transformation was conducted in this study.

Table 12.

Test of Normality for Passive Constructions Sub-scale Occurrences

	All Data Points			Multivariate Outliers Excluded		
	Skewness	*Kurtosis*	*K-S Test*	*Skewness*	*Kurtosis*	*K-S Test*
Science						
Be	1.343	3.143	$p < .05$	0.714	1.242	$p > .05$
non-finite-ing	2.063	5.011	$p < .05$	1.456	2.596	$p < .05$
infinite	2.243	6.725	$p < .05$	0.902	0.406	$p < .05$
perfect tense	1.208	0.531	$p < .05$	1.286	0.949	$p < .05$
modal	2.232	7.59	$p < .05$	1.618	2.912	$p < .05$
get-passive	2.607	8.946	$p < .05$	2.561	4.769	$p < .05$
bare passive	2.337	8.93	$p < .05$	0.99	1.193	$p > .05$
phrasal verb passive	1.519	1.351	$p < .05$	1.11	0.017	$p < .05$
embedded passive	2.528	4.515	$p < .05$	3.595	11.433	$p < .05$
Technology						
Be	1.174	2.078	$p < .05$	0.443	0.839	$p > .05$
non-finite-ing	2.189	6.292	$p < .05$	1.255	0.868	$p < .05$
infinite	3.495	17.535	$p < .05$	0.945	-0.026	$p < .05$
perfect tense	2.507	9.937	$p < .05$	0.675	-0.06	$p < .05$
modal	1.989	5.645	$p < .05$	1.29	2.818	$p > .05$
get-passive	4.86	22.2	$p < .05$	5.094	24.847	$p < .05$
bare passive	0.551	0.084	$p > .05$	0.165	-0.882	$p > .05$
phrasal verb passive	1.722	2.713	$p < .05$	1.334	0.861	$p < .05$
embedded passive	6.08	35.892	$p < .05$			$p < .05$
Engineering						
Be	0.848	0.491	$p < .05$	0.418	-0.056	$p > .05$
non-finite-ing	1.477	1.632	$p < .05$	1.608	1.993	$p < .05$
infinite	2.896	9.515	$p < .05$	1.703	3.852	$p < .05$
perfect tense	2.294	6.921	$p < .05$	1.068	1.232	$p < .05$
modal	1.909	5.029	$p < .05$	0.691	-0.358	$p < .05$

Test of Normality for Passive Constructions Sub-scale Occurrences

	Skewness	Kurtosis	K-S Test	Skewness	Kurtosis	K-S Test
get-passive	5.257	31.651	$p < .05$	2.446	4.144	$p < .05$
bare passive	0.915	1.528	$p < .05$	0.622	-0.919	$p < .05$
phrasal verb passive	2.008	3.572	$p < .05$	1.882	2.74	$p < .05$
embedded passive	7.08	52.192	$p < .05$			$p < .05$
Mathematics						
Be	1.618	4.682	$p < .05$	1.651	4.294	$p < .05$
non-finite-ing	2.584	9.188	$p < .05$	1.241	0.055	$p < .05$
infinite	2.818	11.745	$p < .05$	1.335	1.588	$p < .05$
perfect tense	1.817	2.52	$p < .05$	1.646	1.673	$p < .05$
modal	1.078	1.557	$p < .05$	1.103	1.506	$p < .05$
get-passive	5.089	25.495	$p < .05$	7.616	58	$p < .05$
bare passive	1.117	1.386	$p < .05$	1.179	1.949	$p < .05$
phrasal verb passive	6.933	51.418	$p < .05$	2.674	5.332	$p < .05$
embedded passive	5.784	32.381	$p < .05$			$p < .05$

Table 13.

Test of Normality for Passive Constructions Composite Occurrences

	All Data Points			Multivariate Outliers Excluded		
	Skewness	*Kurtosis*	*K-S Test*	*Skewness*	*Kurtosis*	*K-S Test*
Be	1.247	2.058	$p < .05$	0.837	0.898	$p < .05$
non-finite-ing	2.116	5.379	$p < .05$	1.537	2.061	$p < .05$
infinite	3.888	20.174	$p < .05$	1.357	1.745	$p < .05$
perfect tense	2.303	8.358	$p < .05$	1.141	0.754	$p < .05$
modal	3.296	15.823	$p < .05$	1.274	1.647	$p < .05$
get-passive	4.641	27.118	$p < .05$	3.489	10.273	$p < .05$
bare passive	1.499	4.628	$p < .05$	0.769	-0.291	$p < .05$
phrasal verb passive	3.331	17.993	$p < .05$	1.797	2.374	$p < .05$
embedded passive	4.905	25.708	$p < .05$	8.225	66.285	$p < .05$

Next, a linearity test was performed by looking at the correlations between the dependent variables (Table 14). Although most of the variables were strongly related, the correlations did not entirely reach the .07 threshold for linearity (Tabachnick & Fidell, 2013). However, *be*-passive and bare passive were highly correlated at .74 ($p < .01$). In addition, the correlation matrix indicated that most of variables were related at the significant level ($p < .05$) with the exception of *get*-passive and embedded passive. Consequently, *get*-passive and embedded passive were removed from the MANOVA test in research question two in order to meet the linearity assumption.

Table 14.

Correlations between Passive Constructions Occurrences

	be	ing	infinite	perfect	modal	get	bare	prep.	embedded
be	--	.274**	.318**	.473**	.351**	-.100	.740**	.288**	-.098
ing	.274**	--	.354**	.191**	.261**	-.037	.244**	.162*	-.085
infinite	.318**	.354**	--	.280**	.324**	-.025	.243**	.175*	-.057
perfect	.473**	.191**	.280**	--	.318**	.112	.460**	.200**	-.073
modal	.351**	.261**	.324**	.318**	--	-.110	.399**	.054	-.096
get	-.100	-.037	-.025	.112	-.110	--	-.074	.074	.129
bare	.740**	.244**	.243**	.460**	.399**	-.074	--	.243**	-.081
prep.	.288**	.162*	.175*	.200**	.054	.074	.243**	--	.074
embedded	-.098	-.085	-.057	-.073	-.096	.129	-.081	.074	--

*Note. *$p < .05$ level (2-tailed), **$p < .01$ level (2-tailed)

Levene's Test of Equality of Error Variances was consulted in order to determine whether the homoscedasticity assumption was met. Results indicated that only *be*-passive non-finite-ing ($F_{(3, 205)} = 2.3$, $p > .05$) and *be*-passive perfect tense ($F_{(3, 205)} = 1.85$, $p > .05$) were not significant for Levene's Test, the rest dependent variables were significant for this test, which

indicated non-homoscedasticity (Table 15). However, ANOVA test was robust to the violation of this assumption, therefore, interpretation of the results was continued.

Lastly, the homogeneity of variances and covariance test was run. The Box's Test of Equality of Covariance Matrices indicated the assumption of homogeneity of variances was violated $F_{(84, 89390)} = 3.20$, $p < .05$. However, MANOVA test was robust to the violation of this assumption, thus the analysis for research question two could be proceeded to interpret the results.

Table 15.

Levene's Test of Equality of Error Variances for Dependent Variables

Dependent Variable	F	df1	df2	p
Be	3.497	3	205	.017
non-finite-ing	2.306	3	205	.078
infinite	4.708	3	205	.003
perfect tense	1.845	3	205	.140
modal	7.924	3	205	< .001
bare passive	11.543	3	205	< .001
phrasal verb passive	20.120	3	205	< .001

Research Question 1

The first research question aimed to investigate the percentage of passive voice usage in the twenty examined textbooks. Overall, 32.9% of all verbs in this STEM corpus of 1,004,803 words were in passive voice. If only transitive verbs are considered, 37.6% of verbs are in passive voice.

The percentage of passive voice usage in each discipline was also determined. When excluding the intransitive verbs, the results indicated the highest percentage of passive usage was in Engineering (49.8%), followed by Technology (41.6%), Science (28.7%), and the lowest

97

percentage was in Mathematics, which only contained 26.9% of the passive voice usage. The highest percentage of passive usage was also found in Engineering when intransitive verbs were included (43.9%), Technology had the second highest percentage (36.8%), and the last two disciplines had slightly different percentage, yet Mathematics (23.7%) still remained last when intransitive verbs were added to the denominator (Table 16). As stated in Chapter One and Two, given that intransitive verbs cannot be made passive in English (Folse, 2016), and the percentage of passive voice usage might be influenced by a bigger number of total main verb count in the denominator (Ding, 2002), therefore the present study considered two situations separately.

Table 16.

Percentage of Passive Usage

	Transitive Verbs	Intransitive Verbs	Passive Voice	Percentage
Science	11602	2858	4673	28.7% (T) 24.4 (T+I)
Technology	9964	2222	7088	41.6% (T) 36.8% (T+I)
Engineering	8363	2262	8300	49.8% (T) 43.9% (T+I)
Mathematics	8556	1516	3142	26.9% (T) 23.7% (T+I)
Overall	38485	8858	23203	37.6% (T) 32.9% (T+I)

Research Question 2

In order to determine whether the common passive constructions were different among each discipline, a multivariate analysis of variance (MANOVA) was conducted. The independent variable was area of study (Science = 1, Technology = 2, Engineering = 3,

Mathematics = 4), and the seven common passive constructions consisted of the dependent

variables. Results showed significant differences of the common passive constructions in

different areas of study ($F_{(21, 603)} = 5.91$, $p < .05$, $\eta^2 = .17$). Pillai's Trace was used to interpret

effect size due to the violation of homogeneity of variance (Tabachnick & Fidell, 2013).

Therefore, areas of study accounts for 17% of the variance in common passive constructions

(Table 17).

Table 17.

Multivariate Analysis of Variance between of Common Passive Constructions by Areas of Study

Effect		Value	*F*	Hypothesis *df*	Error *df*	*p*	η^2
	Pillai's Trace	51	5.91	21	603	< .001	.17
Areas of	Wilks' Lambda	56	6.16	21	571.97	< .001	.18
Study	Hotelling's Trace	68	6.37	21	593	< .001	.18
	Roy's Largest Root	44	12.55	7	201	< .001	.30

When separately considered the occurrences of the dependent variables within the

MANOVA test, all six dependent variables were also significantly different by areas of study

with the exception of be-passive non-finite-ing (Table 18). Areas of study accounted for 17.3%

of the variance in *be*-passive construction ($F_{(24, 600)} = 14.28$, $p < .05$, $\eta^2 = .173$), 3.6% of the

variance in *be*-passive non-finite-ing ($F_{(24, 600)} = 2.56$, $p > .05$, $\eta^2 = .036$), 6.8% in *be*-passive

infinite variance ($F_{(24, 600)} = 5.01$, $p < .05$, $\eta^2 = .068$), 5% of the variance in *be*-passive perfect

tense ($F_{(24, 600)} = 3.57$, $p < .05$, $\eta^2 = .05$), 12.9% of the variance in *be*-passive modal ($F_{(24, 600)} =$

10.14, $p < .05$, $\eta^2 = .129$), 22.5% of the variance in bare passive ($F_{(24, 600)} = 19.83$, $p < .05$, η^2

$= .225$), and 9.8% of the variance in phrasal verb passive ($F_{(24, 600)} = 7.44$, $p < .05$, $\eta^2 = .098$).

Table 18.

Tests of Between-Subjects Effects for Common Passive Constructions by Areas of Study

Source	Dependent Variable	Type III SS	df	MS	F	p	η^2
	Be	10931.74	3	3643.91	14.28	< .001	.173
	non-finite-ing	7.55	3	2.52	2.56	.056	.036
	infinite	38.63	3	12.88	5.01	.002	.068
Areas of Study	perfect tense	44.94	3	14.98	3.57	.015	.050
	modal	1861.61	3	620.54	10.14	< .001	.129
	bare passive	5760.96	3	1920.32	19.83	< .001	.225
	phrasal verb passive	14.55	3	4.85	7.44	< .001	.098

All in all, it was showed that *be*-passive appeared frequently greater among all the other dependent variables in Science ($M = 31.36$, $SD = 15.71$), Technology ($M = 43.22$, $SD = 19.35$), Engineering ($M = 34.24$, $SD = 16.27$), and Mathematics ($M = 23.72$, $SD = 11.86$). Bare passive (Science; $M = 17.13$, $SD = 10.28$; Technology; $M = 23.75$, $SD = 10.49$; Engineering; $M = 18.29$, $SD = 12.10$; Mathematics; $M = 9.6$, $SD = 5.85$) *and be*-passive modal (Science; $M = 7.11$, $SD = 5.78$; Technology; $M = 10.86$, $SD = 8.27$; Engineering; $M = 15.35$, $SD = 10.05$; Mathematics; $M = 8.95$, $SD = 6.39$) displayed relatively frequent appearances among the rest of the dependent variables. It should be noticed that the least frequent passive construction was the non-finite-ing passive (Science; $M = .69$, $SD = .90$; Technology; $M = .98$, $SD = 1.19$; Engineering; $M = .71$, $SD = 1.08$; Mathematics; $M = .47$, $SD = .73$) and phrasal verb passive (Science; $M = .82$, $SD = 1.05$; Technology; $M = .62$, $SD = .89$; Engineering; $M = .47$, $SD = .86$; Mathematics; $M = .10$, $SD = .31$), which was averagely used less than one time in the four disciplines (Table 19). In addition, Figure 4 provided a clearer visual displaying the means of passive usage for each passive construction.

Table 19.

Means for Areas of Study

					95% CI	
Dependent Variable	Areas of Study	*M*	*SD*	*SE*	LL	UL
Be	Science	31.36	15.71	2.38	26.66	36.05
	Technology	43.22	19.35	2.15	38.97	47.47
	Engineering	34.24	16.27	2.24	29.83	38.65
	Mathematics	23.72	11.86	2.10	19.59	27.86
non-finite-ing	Science	.69	.90	.15	.40	.98
	Technology	.98	1.19	.13	.72	1.25
	Engineering	.71	1.08	.14	.43	.98
	Mathematics	.47	.73	.13	.21	.72
infinite	Science	1.58	1.42	.24	1.11	2.05
	Technology	2.04	2.02	.22	1.61	2.46
	Engineering	1.29	1.67	.22	.85	1.74
	Mathematics	.90	1.18	.21	.48	1.31
perfect tense	Science	2.20	2.44	.31	1.60	2.80
	Technology	2.38	2.09	.28	1.84	2.93
	Engineering	1.98	1.79	.29	1.42	2.55
	Mathematics	1.21	1.89	.27	.68	1.74
modal	Science	7.11	5.78	1.17	4.81	9.41
	Technology	10.86	8.27	1.06	8.78	12.93
	Engineering	15.35	10.05	1.10	13.19	17.51
	Mathematics	8.95	6.39	1.03	6.92	10.97
bare passive	Science	17.13	10.28	1.47	14.24	20.03
	Technology	23.75	10.49	1.33	21.13	26.36
	Engineering	18.30	12.10	1.38	15.58	21.01
	Mathematics	9.60	5.85	1.29	7.06	2.15
phrasal verb	Science	.82	1.05	.12	.59	1.06
passive	Technology	.62	.89	.11	.40	.83
	Engineering	.47	.86	.11	.25	.69
	Mathematics	.10	.31	.11	-.11	.31
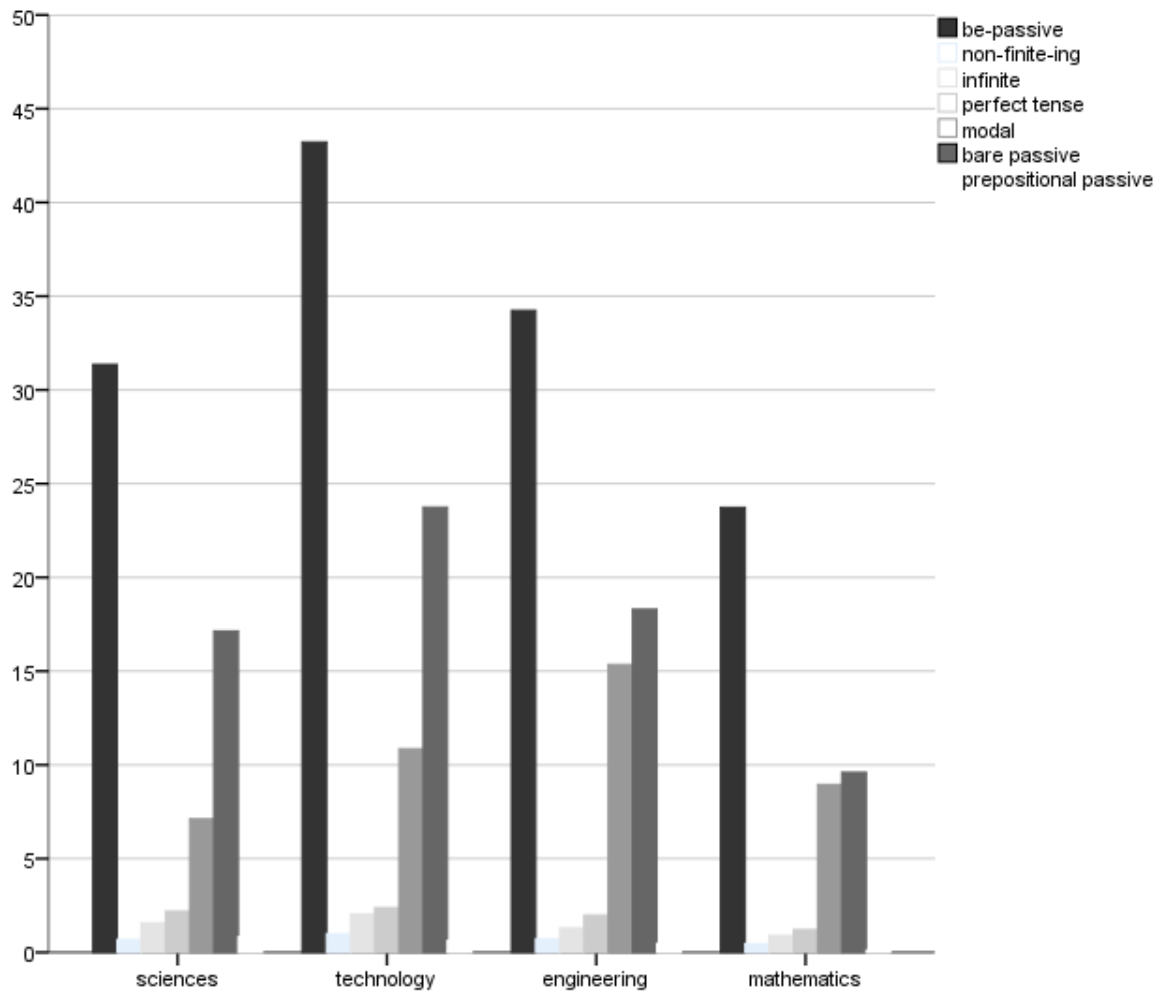
Figure 4. Means for areas of study

Research Question 3

The third research question continued to look further into the specific verbs that were used most frequently in passive voice in the examined texts. The goal was to identify the most commonly used vocabulary for these passive forms, and it was crucial to teach students these past participles that appear with this grammar point (Folse, 2016).

The results showed that the most commonly used past participle in the study was *be used* ($n = 1727$), followed by *be called* ($n = 1238$). The occurrences of the remained past participles

ranged from 926 to 201. The list only included 20 past participles due to the dramatically

dropped occurrences of the rest past participles (Table 20).

Table 20.

20 Most Frequent Past Participles Used in Passive Voice

Past Participle	Occurrence
be used	1727
be called	1238
be given	926
be required	730
be shown	711
be defined	513
be known	509
be obtained	387
be made	379
be found	309
be written	268
be based	267
be determined	263
be expressed	238
be described	237
be applied	234
be said	233
be discussed	220
be needed	217
be considered	201

Furthermore, looking at the most commonly used past participles in each discipline, the

results indicated *be called* ($n = 288$) was used the most frequently in Science, *be used* occurred

the most in both Technology ($n = 491$) and Engineering ($n = 830$), and *be given* ($n = 294$) was

commonly used in Mathematics. These three past participles also took up the first three places

on the above list.

Additional Findings

To review statistically significant difference of the passive constructions among the four disciplines, the post hoc test of Scheffe was run due to unequal sample size in each cell. The results showed that the usage of *be*-passive had statistically significant difference between Science and Technology ($M = -11.86$, $SE = 3.21$, $p < .05$), Technology and Engineering ($M = 8.98$, $SE = 3.11$, $p < .05$), Technology and Mathematics ($M = 19.49$, $SE = 3.01$, $p < .05$), and Engineering and Mathematics ($M = 10.51$, $SE = 3.07$, $p < .05$). The results also indicated that the usage of *be*-passive infinite ($M = 1.14$, $SE = .30$, $p < .05$) and *be*-passive perfect tense ($M = 1.17$, $SE = .39$, $p < .05$) had statistically significant difference between Technology and Mathematics. The usage of *be*-passive modal had significantly difference between Science and Engineering ($M = -8.24$, $SE = 1.60$, $p < .05$), Technology and Engineering ($M = -4.50$, $SE = 1.52$, $p < .05$), and Engineering and Mathematics ($M = 6.40$, $SE = 1.50$, $p < .05$). Moreover, the usage of bare passive had significantly difference between Science and Technology ($M = -6.61$, $SE = 1.98$, $p < .05$), Science and Mathematics ($M = 7.53$, $SE = 1.96$, $p < .05$), Technology and Engineering ($M = 5.45$, $SE = 1.91$, $p < .05$), Technology and Mathematics ($M = 14.14$, $SE = 1.85$, $p < .05$), and Engineering and Mathematics ($M = 8.69$, $SE = 1.89$, $p < .05$). Finally, there were statistically significant difference of the usage of phrasal verb passive between Science and Mathematics ($M = .72$, $SE = .16$, $p < .05$), and Technology and Mathematics ($M = .51$, $SE = .15$, $p < .05$; Table 21).

Table 21.

Multiple Comparisons for Passive Constructions by Areas of Study

Dependent Variable	(I) Disciplines	(J) Disciplines	*M*	*SE*	*p*
Be	Science	Technology	-11.86*	3.211	.004
		Engineering	-2.88	3.267	.855
		Mathematics	7.63	3.174	.126
	Technology	Science	11.86*	3.211	.004
		Engineering	8.98*	3.106	.042
		Mathematics	19.49*	3.007	< .001
	Engineering	Science	2.88	3.267	.855
		Technology	-8.98*	3.106	.042
		Mathematics	10.51*	3.067	.010
	Mathematics	Science	-7.63	3.174	.126
		Technology	-19.49*	3.007	< .001
		Engineering	-10.51*	3.067	.010
non-finite-ing	Science	Technology	-.29	.199	.541
		Engineering	-.02	.203	1.000
		Mathematics	.22	.197	.733
	Technology	Science	.29	.199	.541
		Engineering	.28	.193	.563
		Mathematics	.52	.187	.057
	Engineering	Science	.02	.203	1.000
		Technology	-.28	.193	.563
		Mathematics	.24	.190	.661
	Mathematics	Science	-.22	.197	.733
		Technology	-.52	.187	.057
		Engineering	-.24	.190	.661

Multiple Comparisons for Passive Constructions by Areas of Study

Infinite	Science	Technology	-.46	.322	.568
		Engineering	.28	.328	.862
		Mathematics	.68	.318	.209
	Technology	Science	.46	.322	.568
		Engineering	.74	.312	.132
		Mathematics	1.14*	.302	.003
	Engineering	Science	-.28	.328	.862
		Technology	-.74	.312	.132
		Mathematics	.40	.308	.644
	Mathematics	Science	-.68	.318	.209
		Technology	-1.14*	.302	.003
		Engineering	-.40	.308	.644
perfect tense	Science	Technology	-.18	.412	.978
		Engineering	.22	.419	.965
		Mathematics	.99	.407	.117
	Technology	Science	.18	.412	.978
		Engineering	.40	.398	.797
		Mathematics	1.17*	.386	.028
	Engineering	Science	-.22	.419	.965
		Technology	-.40	.398	.797
		Mathematics	.77	.393	.279
	Mathematics	Science	-.99	.407	.117
		Technology	-1.17*	.386	.028
		Engineering	-.77	.393	.279

Multiple Comparisons for Passive Constructions by Areas of Study

modal	Science	Technology	-3.74	1.572	.133
		Engineering	-8.24*	1.600	< .001
		Mathematics	-1.84	1.554	.706
	Technology	Science	3.74	1.572	.133
		Engineering	-4.50*	1.521	.035
		Mathematics	1.91	1.472	.643
	Engineering	Science	8.24*	1.600	< .001
		Technology	4.50*	1.521	.035
		Mathematics	6.40*	1.502	.001
	Mathematics	Science	1.84	1.554	.706
		Technology	-1.91	1.472	.643
		Engineering	-6.40*	1.502	.001
bare passive	Science	Technology	-6.61*	1.978	.012
		Engineering	-1.16	2.013	.954
		Mathematics	7.53*	1.955	.002
	Technology	Science	6.61*	1.978	.012
		Engineering	5.45*	1.913	.046
		Mathematics	14.14*	1.852	< .001
	Engineering	Science	1.16	2.013	.954
		Technology	-5.45*	1.913	.046
		Mathematics	8.69*	1.889	< .001
	Mathematics	Science	-7.53*	1.955	.002
		Technology	-14.14*	1.852	< .001
		Engineering	-8.69*	1.889	< .001

Multiple Comparisons for Passive Constructions by Areas of Study

		Technology	.20	.162	.664
	Science	Engineering	.35	.165	.213
		Mathematics	.72*	.160	< .001
		Science	-.20	.162	.664
	Technology	Engineering	.15	.157	.829
		Mathematics	.51*	.152	.011
phrasal verb passive		Science	-.35	.165	.213
	Engineering	Technology	-.15	.157	.829
		Mathematics	.37	.155	.136
		Science	-.72*	.160	< .001
	Mathematics	Technology	-.51*	.152	.011
		Engineering	-.37	.155	.136

Note. *p < .05 level (2-tailed)

Conclusion

All in all, Chapter Four provided the answers to the proposed research questions and presented the results of the statistical analyses that investigated passive voice usage in college STEM textbooks in terms of the percentage of total passive usage, the difference of the common passive constructions among disciplines, and the frequent occurred past participles in passive forms. Altogether, there were over one million running words that were examined out of 20 textbooks in the four areas of study. Descriptive statistics displayed an overall 37.6% of passive usage in the examined texts. The percentage was reduced (subtracting 4.7% form the initial 37.6%) when intransitive verbs were added to the denominator. Nevertheless, the overall percentage remained greater than 30%, which reconciled with the findings in Alvin (2014). A MANOVA test indicated that there was statistically significant difference of the common passive constructions among the four disciplines. The usage of passive forms tended to have greater

differences between Mathematics and the other three disciplines. Moreover, descriptive statistics in research question three identified 20 most commonly used past participles in the four areas, where *be used* was inclined to occur most frequently, which was also consistent with Folse's list of 15 most-common past participles in passive voice (Folse, 2016).

Chapter Five interpreted and summarized the findings from the descriptive statistics and statistical analysis. The chapter also discussed the significance of the findings as well as the limitations of the study. Finally, it concluded with pedagogical implications for teaching passive voice to ELs and recommendations for future research.

CHAPTER 5: DISCUSSION

This chapter examined the results of the current study investigating the passive voice usage in college STEM textbooks. This chapter first reviewed the purpose of the study, and then discussed the findings of the study. Next, the significance and the limitations of the study were considered. Finally, the chapter ended with pedagogical implications and recommendations for future research.

Purpose of the Study

Dissimilar to first language acquisition, SLA goes through a systematic language acquisition process, and grammar learning has been taken into consideration due to this crucial linguistic characteristic. Of all the grammatical features in English, passive voice has been viewed as one of the most important yet controversial features by scholars, practitioners and L2 learners. It is important for grammar learning because it restructures a regular active sentence and grammatically complicates it; thus, L2 learners need targeted training on this specific grammar point so that they can comprehend knowledge in their readings (Ferreira, 1994; Folse, 2016; LoMaglio & Robison, 1985; Prystupa, 2017). It is, at the same time, controversial because there have been two sides of arguments surrounding this grammar point fighting on whether passive voice should be used in academic writing (positive: Alvin, 2014; Biondi, 2002; Ding, 2002; Millar et al., 2012; negative: Beason & Lester, 2003; Matthews et al., 1996; Sigel, 2009). Many research studies have provided empirical evidence supporting that passive voice is favored in scientific discourses (Alvin, 2014; Atkinson, 1996; Bazerman, 1988; Ding, 2002; Rodman, 1994). However, literature discussed so far has limited its spectrum exploring the passive voice usage in only a few areas of study, where a research gap was opened to the question that whether

the usage of passive voice is equally prevalent in other areas of study (e.g., Technology, Engineering, and Mathematics), as well as other types of writing. In addition, an exceedingly optimistic job prospect for STEM graduates has made these majors extremely desirable especially for international college students in recent years. Thus, there is a need for teaching English to increasing numbers of international students majoring in these fields.

Undergraduate L2 learners who study in any of the STEM majors, need to comprehend their textbooks in order to acquire knowledge and keep good academic performance. Given that passive voice is also confirmed to be conventionalized in these areas of study, it is thus important for students to understand how it is constructed and how to interpret its meanings. According to noticing hypothesis (Schmidt, 1990), it would be a risk to neglect passive voice in English instruction, for students are probably not going to notice the passive structures in their readings.

Similarly, for L2 learners who study at the postgraduate level, it is important to be familiar with the writing convention in their fields if they want to publish their research in academic journals. Although many writing guides advocate avoiding passive voice in the academic publications, it will be advisable to look at what has really been the practice in academic writing. Moreover, it is also important for English instructors and grammar book authors to be familiar with the current writing convention regarding passive voice usage in STEM areas, so that they know what and how to teach in EFL or ESL classes, as well as what and how to illustrate this specific grammar point in grammar books.

Consequently, the main purpose of this study was to explore the overall usage rate of passive voice and how the common passive constructions were used among the investigated disciplines. The research designs aimed to achieve three empirical goals. The first and the third objectives were to find out the percentage of passive voice usage and the dispersion of past

111

participles used in passive forms in the examined textbooks. The second goal compared the difference of commonly used passive forms among the investigated areas of study.

To accomplish these objectives, the study selected 20 textbooks that comprised over 1 million running words to examine. The past participles in passive voice and the main verbs were labeled for the general descriptive analyses. The common passive constructions were categorized for the analysis of research question two. A MANOVA test was conducted to compare the difference of the common passive forms among targeted disciplines. One instrument called Compleat Lexical Tutor (Cobb, 2017) was utilized to identify frequently used past participles in passive voice in research question three.

Summary of the Findings

A sample corpus of over 1 million running words, altogether 289 cases was collected for two months during the Fall 2017 semester. Numbers of cases were reduced to 209 due to the elimination of outliers. Finally, the cases in Science ($n = 45$), Technology ($n = 55$), Engineering ($n = 51$), and Mathematics ($n = 58$) were included in the data analysis in the second research question. Descriptive statistics revealed that the mean text length of the total corpus was 50240.15 running words ($SD = 240.47$; range, 50011-51164), which exceeded the one-million-word threshold. The transitive and intransitive verbs, past participles in passive voice, and passive constructions were marked on the texts and transformed into numerical and categorical data on a spreadsheet. After the examination of the texts was done, the answer to the first research question was computed in Excel, the data were inserted in the Statistical Package for the Social Sciences (SPSS) 24.0 for data analysis for the second research question, and the labeled texts were submitted to the Compleat Lexical Tutor (Cobb, 2017) to obtain the answer for the last research question.

Prior to the data analysis, statistical assumptions were tested. Although there were violations on the homogeneity and the homoscedasticity tests, MANOVA test is robust to the violation of these assumptions. Two passive constructions (*get*-passive and embedded passive) were removed from the MANOVA test because they were noncorrelated with other dependent variables. Given that no major violations were found, data analysis proceeded to interpret the second research question – was there a statistically significant difference of the common passive constructions between the four STEM disciplines? If yes, what were those differences?

Results

Research question one. The first research question aimed to examine the usage rate of passive voice in the targeted fields. The result revealed that the overall usage rate of passive voice across the four areas of study was 37.6%. Even though when intransitive verbs were included in the denominator, the usage rate was still over 30% (32.9%). It is worth noting that whether intransitive verbs were added to the denominator would influence the percentage of passive voice usage in a great extent, however, given that intransitive verbs cannot be conjugated to passive forms and that students are not going to learn how to conjugate those verbs to passive voice, we should focus on the percentage of when transitive verbs were included.

The finding is consistent with previous studies investigating passive voice usage in academic writing in scientific discourses. Although the passive usage in scientific writing has dropped greatly from 75% (Bazerman, 1988) to 53% (Harmon, 1992) within a hundred years, and from 53% to around 30% (Alvin, 2014) after another two decades, it is still a dominant writing convention found in the methodology section. Alvin's (2014) study of examining the proportion of passive usage in science journals indicated that passive clauses comprised 30.02% of the total number of clauses examined in 60 science articles. In addition, when looked at individual discipline, the results showed that the highest passive voice usage rate was in

Engineering, 49.8% or 43.9% depending on whether intransitive verbs were included, which was close to what Harmon (1992) found in his study examining 50 most-cited scientific literature.

Although the passive voice usage in Science and Mathematics were less than 30% (Table 16), based on the current study's results and the previous studies (Alvin, 2014; Harmon, 1992), the percentage of passive voice usage exhibited a consistent radio in existing college STEM textbooks.

Research question two. The second research question focused on the mean differences of the common passive constructions among the four disciplines. It was hypothesized that there would be no difference of the common passive constructions among the examined disciplines. A multivariate analysis of variance (MANOVA) test indicated that the examined passive forms occurred differently across disciplines ($F_{(21, 603)} = 5.91$, $p < .05$, $\eta^2 = .17$). Results of between-subjects effects revealed that only *be*-passive non-finite-ing was rarely used among all four disciplines ($F_{(24, 600)} = 2.56$, $p > .05$, $\eta^2 = .036$). The most commonly used passive construction was *be*-passive, and it was favored differently between all disciplines except between Science and Engineering ($M = -2.88$, $SE = 3.27$, $p > .05$). The second and the third frequently used passive forms were bare passive and *be*-passive modal. Results of between-subjects effects showed that they were used significantly different among the disciplines (*be*-passive modal, $F_{(24, 600)} = 10.14$, $p < .05$, $\eta^2 = .129$; bare passive, $F_{(24, 600)} = 19.83$, $p < .05$, $\eta^2 = .225$), however, a post-hoc analysis exhibited the differences actually fell between Science and Engineering ($M = -8.24$, $SE = 1.60$, $p < .05$), Technology and Engineering ($M = -4.50$, $SE = 1.52$, $p < .05$), and Engineering and Mathematics ($M = 6.40$, $SE = 1.50$, $p < .05$) in *be*-passive modal, and between Science and Technology ($M = -6.61$, $SE = 1.98$, $p < .05$), Science and Mathematics ($M = 7.53$, $SE = 1.96$, $p < .05$), Technology and Engineering ($M = 5.45$, $SE = 1.91$, $p < .05$), Technology and

Mathematics ($M = 14.14$, $SE = 1.85$, $p < .05$), and Engineering and Mathematics ($M = 8.69$, $SE = 1.89$, $p < .05$) in bare passive.

Moreover, the Estimated Margin Means plot indicated that almost all passive constructions, except for the phrasal verb passive, appeared mostly in Technology. Almost all passive constructions appeared least in Mathematics with the exception of *be*-passive modal. Phrasal verb passive was more preferred in Engineering than in Technology, and *be*-passive modal was more disfavored in Science than in Mathematics (Figure 5 to Figure 11).
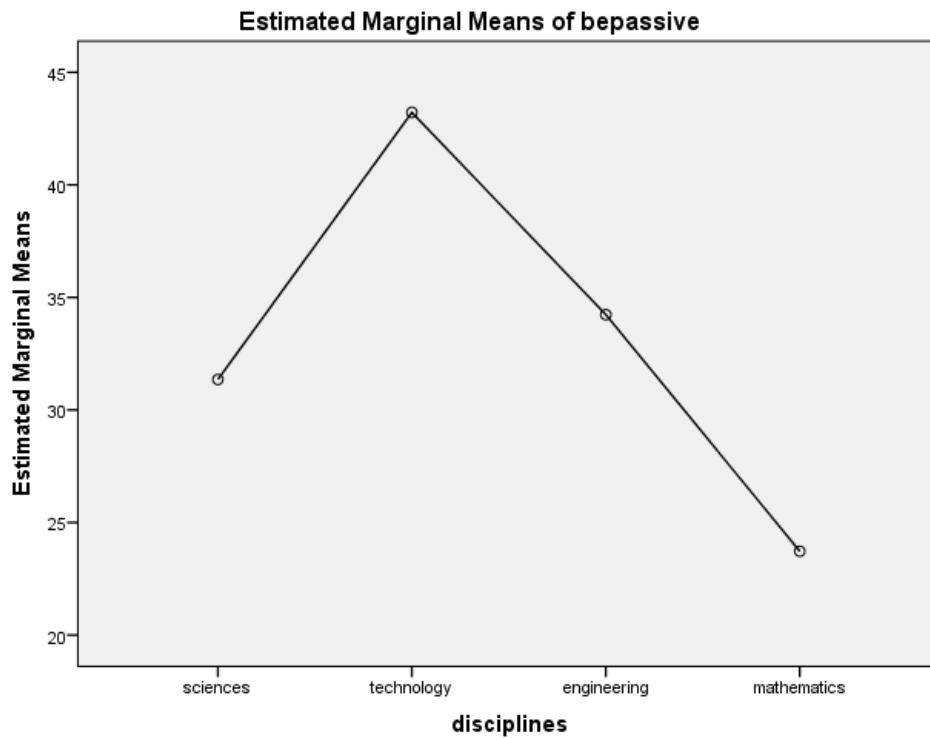
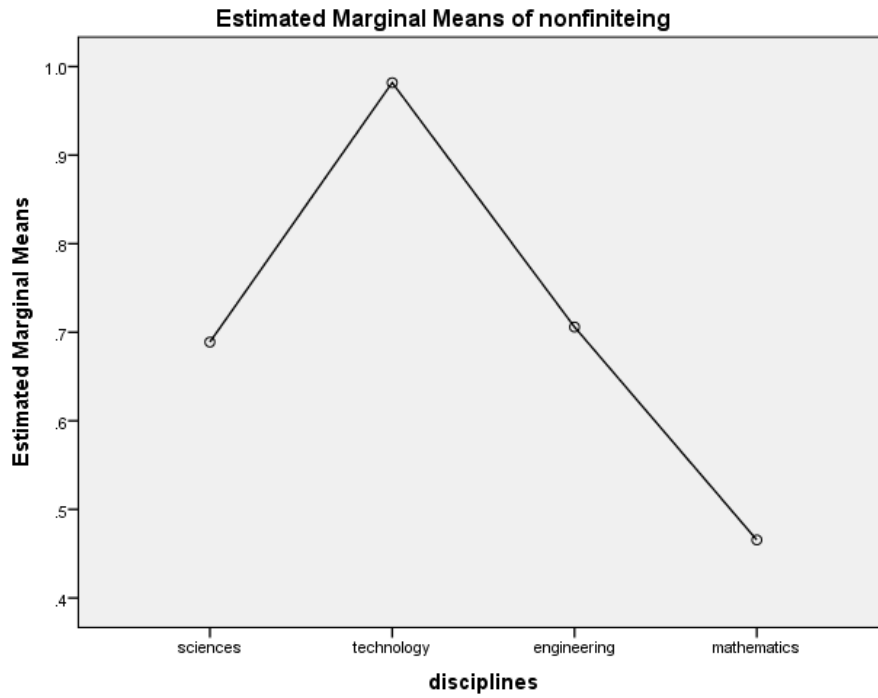*Figure 5. Passive constructions by disciplines (be-passive)*

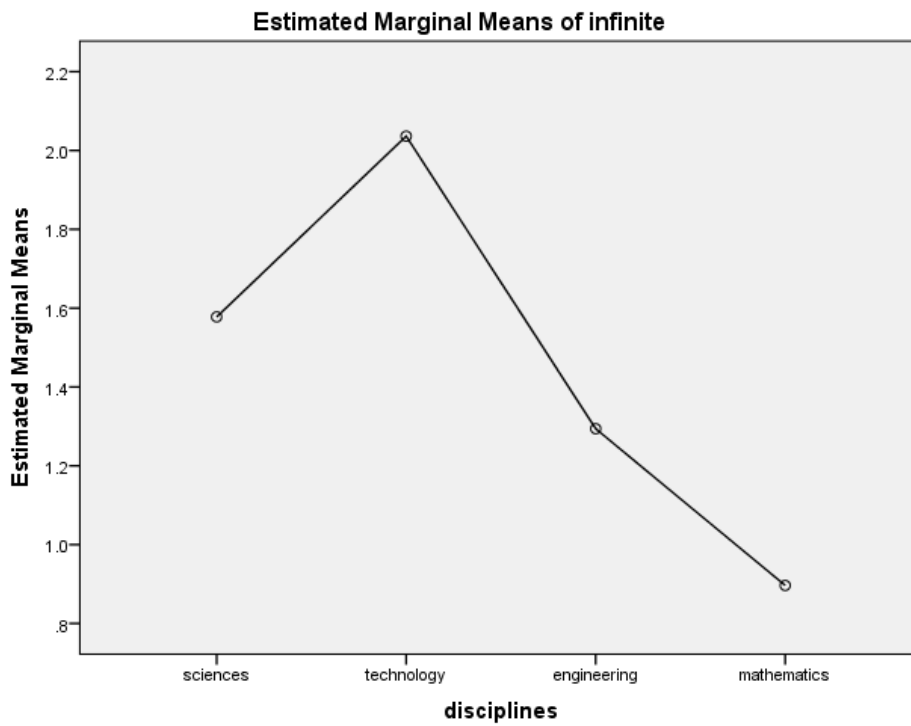*Figure 6. Passive constructions by disciplines (non-finite-ing)*

Figure 7. Passive constructions by disciplines (infinite)
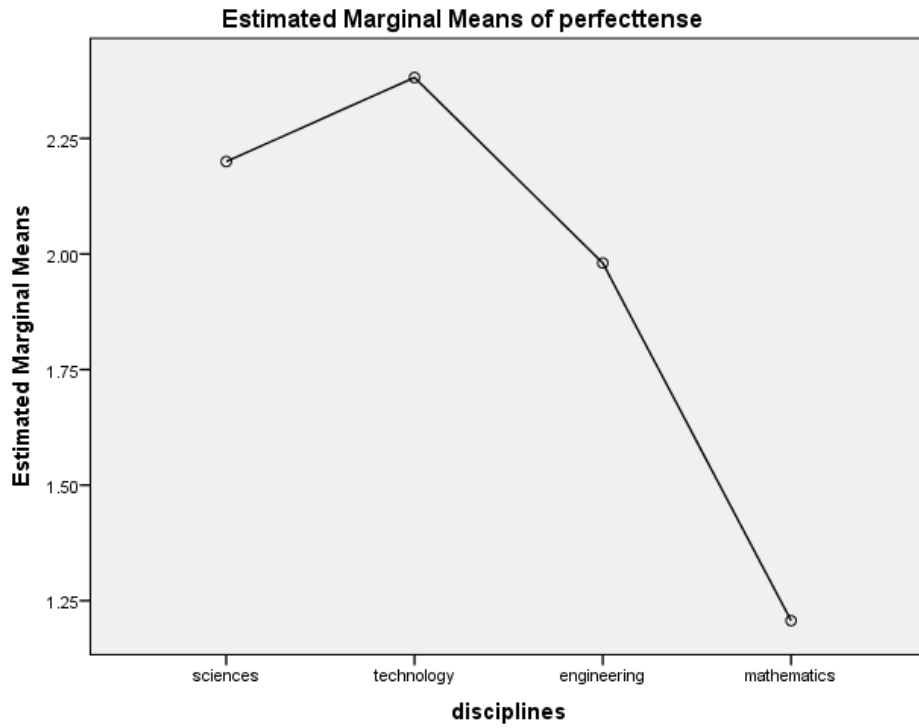
Figure 8. Passive constructions by disciplines (perfect tense)
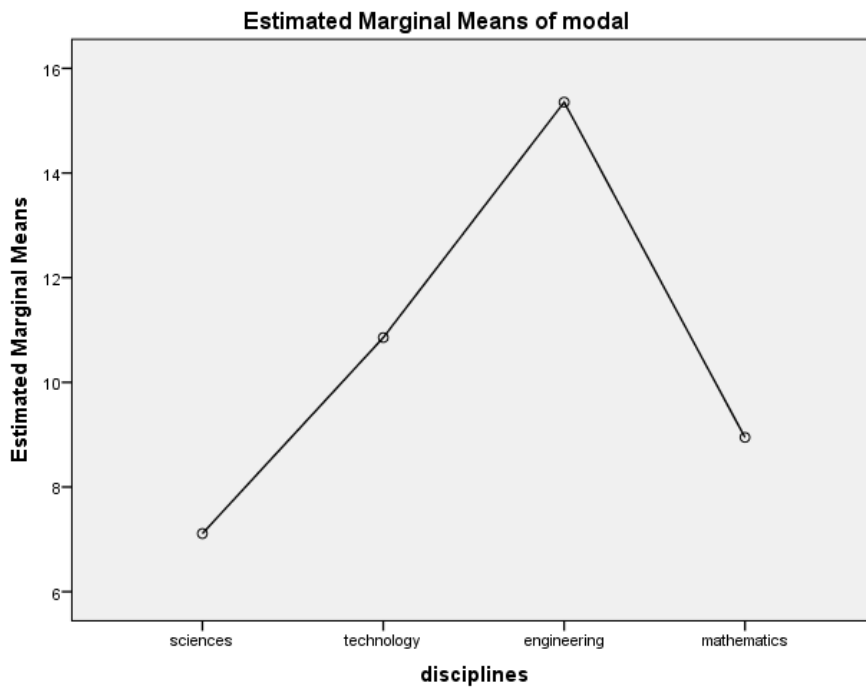

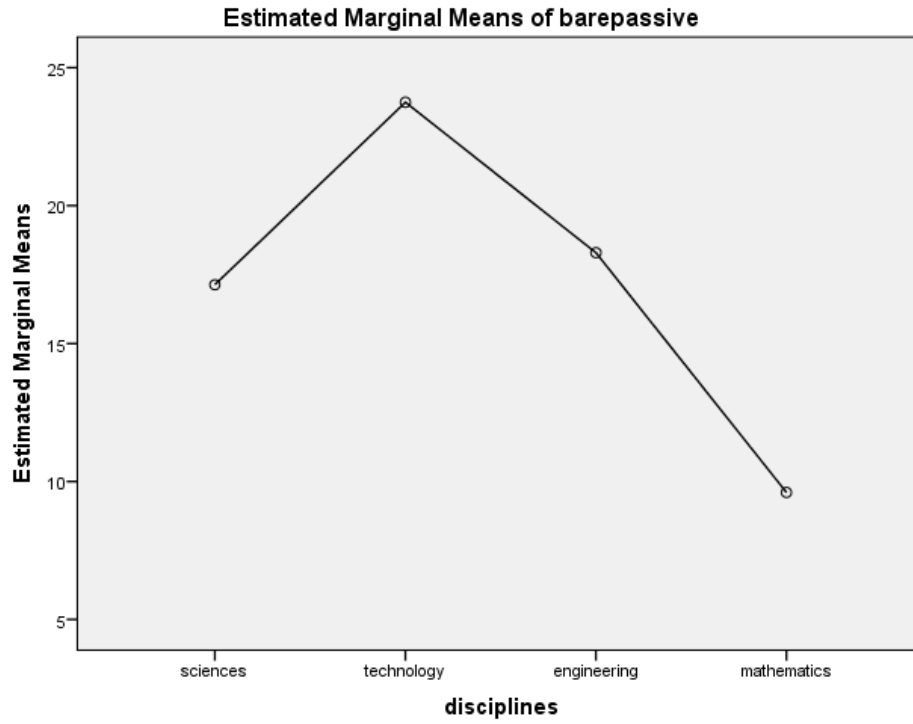*Figure 9. Passive constructions by disciplines (modal)*

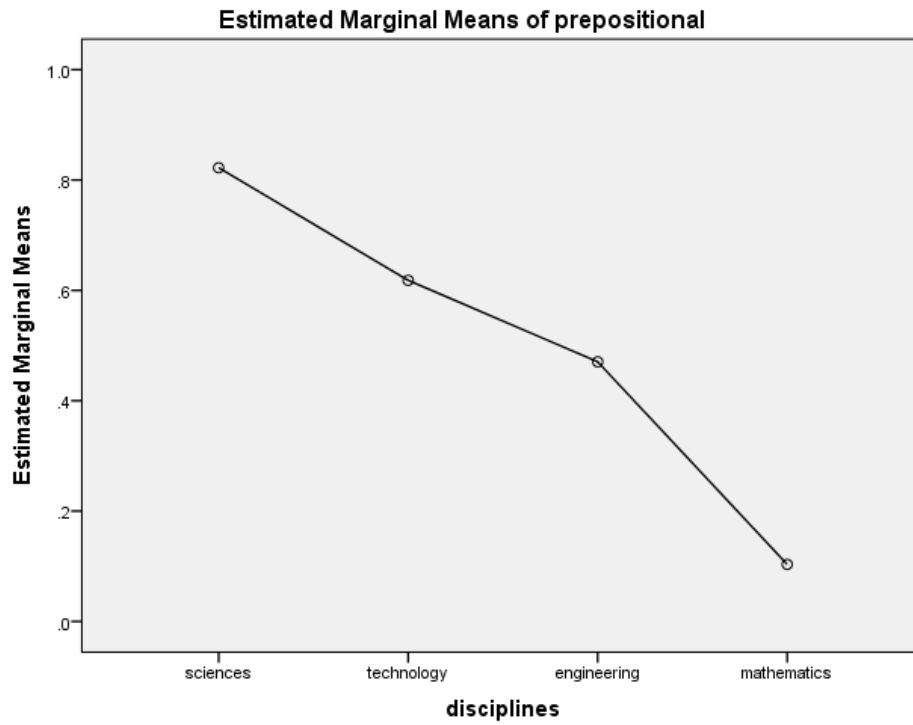*Figure 10. Passive constructions by disciplines (bare passive)*

Figure 11. Passive constructions by disciplines (phrasal verb passive)

Research question three. The last main research question was interested in finding out the most commonly used past participles in passive voice. The results revealed that *be used* occurred most frequently (*n* = 1727) across the four disciplines, which was similar with Folse's *15 Past Participle in Passive Voice* list (Folse, 2016). Other vocabulary concluded in the present study that also fell on Folse's list were *be called* (*n* = 1238), *be given* (*n* = 926), *be made* (*n* = 379), *be found* (*n* = 309), *be based* (*n* = 267), and *be considered* (*n* = 201). It should be noted that his list was based on the data from the Corpus of Contemporary American English (COCA), which contains genres among spoken, fiction, magazines, newspapers, and academic texts, thus, the list that is provided in the current study targeted a more specific genre (academic discourse) and fields of study.

Significance of the Study

Language and education-related research serve primarily to improve effectiveness in teaching and learning so that learner's gap is addressed by scholars and practitioners. Therefore, the present study is significant for adding empirical evidence to raising recognition of an aspect of language knowledge, to providing suggestions to enhance teaching pedagogy, and to guiding student's learning.

First, there is the controversial argument that passive voice in academic writing is ambiguous to both L2 learners and English instructors. It is very easy to overlook this grammar point if learners and instructors think it is not important or they are not sure if it is worth teaching or learning. However, according to passive voice studies, no matter how many times writing guides ask authors to avoid passive voice, its usage rate remains stable in academic scientific writing. There have been many arguments presented about whether passive voice should be avoided in scientific writing. To the best of the researcher's knowledge, pervious passive voice

studies have not touched upon the areas in STEM separately or in combination. The main purpose of this investigation was to explore the passive voice usage and to fill in the research gap in that area. This study reconciled with other research that advocate treating passive voice objectively. The results revealed a consistent passive usage rate of over 30% across the four examined disciplines; Engineering had an even higher percentage for almost 50% of passive voice usage, which demonstrated that passive voice was indeed commonly used in these fields.

Second, Swales (1990) also noted that passive voice consisted of a large portion across academic writings. Since passive voice usage cannot be generally applied to all academic areas, L2 learners should be taught specifically and differently how passive voice was to be used in which circumstances. For example, comparing to other disciplines, Engineering textbooks showed more preference in passive voice, and *be*-passive modal seemed to be used more frequently than the rest of the common passive forms in this discipline. English instructors then can modify their lesson plans to meet students' need appropriately so that students could read and write competently to achieve various rhetorical goals in their study. Findings of Hinkel (2004) pointed out the importance of explicit instructions on how to use passive voice even for advanced L2 learners. Thus, by emphasizing passive voice in Intensive English Program (IEP) and in English for Academic Purposes (EAP) programs, both teachers and students would know how to teach and what to learn in order to succeed academically.

Furthermore, it is imperative for grammar book authors to organize book content based on corpus-based research so that a carefully designed and directed teaching material of passive voice instruction and exercises can be included in their textbooks. Consequently, the results of this study revealed current grammar trend of passive voice usage in college STEM textbooks, thus providing necessary guidance to fulfill the pedagogical objective stated above.

Limitations of the Study

There are several limitations of the methodology in the present study that needed to be addressed. The first limitation concerned the generalizability of this study. Due to the fact that this study used convenience and systematic sampling, it was subject to systemic bias, and thus resulted in limitations of the generalizability of results (Fraenkel & Wallen, 2009). However, this potential bias was addressed in the study's methodology to ensure that the sample of this study met the research goals and answered the research questions. Moreover, the findings were limited in a sense that they were not to be generalized in other types of STEM discourses. Data was collected from college STEM textbooks at undergraduate level at a metropolitan university in a southeastern state of the United States only, and only a sample of twenty textbooks used at this university were examined. Therefore, it is uncertain how these findings would apply to academic articles or reports in the same fields. Furthermore, it was also unknown how these findings would apply to the same types of discourses in other countries or at a different academic level. Thus, applications of the teaching pedagogy of this study should be with caution when the language or learning environment, or the academic level was different.

Second, the results revealed that the overall percentage of passive usage agreed with what other studies have reported, which was approximately 30%. The results also indicated that passive voice is not only a sociocultural selection or writing pattern in Science discourses (Ding, 2002; Seoane, 2013), it is also an "ineluctable" writing pattern in Technology and Engineering writings. However, a solid conclusion cannot be made to affirm that the proportion of passive voice usage was high compared with active voice usage, since there were no criteria or standards to determine a high usage rate of passive voice.

Third, the current study did not include any confounding variables to the statistical model; different results may have been obtained had text length, genre or literary style been

controlled as confounding variables in the data analysis. In addition, there were a few assumptions that have been violated for the MANOVA test. Although the sample size of this study was large and MANOVA test was robust to the violation of these assumptions, different results may have been produced had all the assumptions been met.

Pedagogical Implications of the Findings

The findings of this study reiterated and highlighted the importance of teaching and learning passive voice in English learning. The results indicated that passive voice and passive constructions were commonly and variously used in broad fields. Given that relatively little explanation can be found to guide the L2 writers about when or how a particular voice can be used, in particular in academic writing passive voice should not be overlooked in grammar teaching and should be paid attention by English instructors and grammar book authors. Writing is as important as reading in academic study, and passive voice is an important grammatical feature that influences the accuracy of academic writing. Biber (1993) noted that passive voice was much more common in scientific writing than in spoken conversation or fictional writing.

One way to gauge the status of current perceptions regarding the use of passive voice in academic writing is to peruse universities' writing guides from their writing centers. It appears that all the writing guides suggested using active voice or switching from passive voice to active voice to strengthen and clarify writing. Yet, very few noted that passive voice was used in certain academic contexts and may in fact be preferred in certain cases. For instance, the UWC at Texas A&M University indicated that even though many writing guides suggested active voice over passive voice, there were two circumstances that passive voice was the better option in two cases: (a) when the performer of the action could be ignored or is irrelevant and (b) when the object needed to be emphasized) ("Active & Passive Voice," n.d.). The Writing Center at

University of North Carolina-Chapel Hill agreed that passive voice could obscure writing, yet they still recognized situations when passive voice was preferable ("Passive Voice," n.d.). The Purdue University Online Writing Lab suggested to generally avoid passive voice because it could cause writing monotony and create awkward sentences ("Active and passive voice," n.d.). Moreover, the UWC of the University of Nevada, Reno also suggested avoiding passive voice whenever possible, while they pointed out that passive voice was often used in scientific prose ("Passive voice: What it is and how to reduce usage," 2013). Similarly, the UWC of the University of Central Florida agreed on avoiding passive voice but acknowledged the tendency of passive usage in scientific writing (Pitts, 2010). Although it is obvious that every writing guide addresses the suitability of passive voice in scientific discourse, no mention is made for any other disciplines. The findings from this study suggest that passive voice should also be emphasized in Technology and Engineering.

As discussed in Chapter 2, the *by*-passive can make sentence difficult because when we transform active sentence to *by*-passive, a new word is added to the sentence. For example, *the dog chased the boy* vs. *the boy was chased by the dog*. It is harder than the *be*-passive because in the *be*-passive (e.g., *the data was collected in the lab)*, only the subject and the object are switching positions. However, in this simple *be*-passive, the agent is missing so that it is not clear that what the object is being acted upon, or in another word, who is the doer. Therefore, many scholars have been criticizing passive voice as being unclear. Nevertheless, this weakness can be easily fixed by adding the preposition *by* after the past participle (see examples in Table 22).

Table 22.

Comparing the Difference among Active voice and Passive Sentence with and without

Preposition 'by'

Active Voice VERB	Passive Voice be + Past Participle	Passive Voice be + Past Participle + by + indirect object
The dog chased the boy.	The boy *was chased*.	The boy was chased *by the dog*.
Shakespeare *wrote* Romeo and Juliet.	Romeo and Juliet *was written*.	Romeo and Juliet was written *by Shakespeare*.
Police shot a gunman.	A gunman *was shot*.	A gunman was shot *by police*.
Michelangelo created the famous sculpture *David* between 1501 and 1504.	The famous sculpture *David was created* between 1501 and 1504.	The famous sculpture *David* was created *by Michelangelo* between 1501 and 1504.
Researchers are collecting data in the lab.	The data *are being collected* in the lab.	The data are being collected *by researchers* in the lab.

However, it may sound awkward to name the doer in some circumstances. We should

not or could not name the doer when the information is common sense or when we do not care

who performed the action or the doer is just not known. For instance, it is common sense that

every president is elected *by the people*, thus adding the phrase *by the people* to *President*

Washington was reelected by the people for a second term in 1792 does not provide any new nor

important information, thereby can be omitted (Folse, 2016). In other words, omitting the agent in passives is indeed conventional in such circumstance. In some other cases, the performer in a passive sentence is being neglected, especially in scientific research, because the success of an experiment is not limiting by the person who did the experiment, yet it is repeatable by other researchers following the same procedure (Ding, 2002). In general, instructional details of using specific voice is not properly addressed in certain degree. This may become the result of the traditional separation of teaching grammar and teaching writing (Hinkel, 2004). The teaching of grammar covers the use of tenses and voices without relating those to the academic writing, while, in writing instruction, the use of voice varies depending on the purpose of the author stance (Baratta, 2009).

L2 learners should learn how to form passive voice of verb tenses and sentence structures, and know when to use it so that they are more precise and efficient in expressing their intended message in their writings. However, indications from Alvin's (2014) study have brought up the idea that it was possible international students were never given instructions on how to appropriately use the grammatical voice in their writing classes. Passive voice is not a land mine in academic writing, and there is nothing wrong with using passive voice. Moreover, there are some cases when passive voice is actually desired over active voice, and it is important for English teachers to clarify them to their students. The findings of this study revealed that the percentage of passive voice usage was almost 50% in Engineering, that is to say, one in two verbs in Engineering texts were in passive voice. Therefore, English teachers should pay attention to teaching passive voice in different writing conventions so that students could write more scientifically and professionally in their fields.

In addition, the findings of this study indicated that the common passive constructions appeared differently among different disciplines, suggesting that English instructors should modify their teaching objectives and emphasize appropriate learning goals for students majoring in different STEM fields. For example, L2 learners are easily confused by bare passive, yet it was used the most frequently in Technology. Therefore, one of the teaching and learning objectives for L2 learners majoring in Technology could be to comprehend bare passive in reading and to be able to apply in productive language skills. Moreover, the results also revealed that Science tends to use phrasal verb passive more frequently than the other three disciplines, thus it should also be highlighted to students in this field. It should also be noteworthy that in order to comprehend phrasal verb passive, L2 learners need to first understand the phrasal verbs used in this passive form. For instance, if students do not understand the meaning of *put off,* it will be unlikely that they can interpret the sentence "*The meeting was put off*". As a result, English instructors should also pay attention to phrasal verbs instruction.

Since vocabulary is recognized as an essential component in language proficiency, student acquisition of vocabulary is seen as a key for academic success (Nagy & Townsend, 2012). In consideration of students' need to know much vocabulary to achieve academic success, perhaps the biggest contribution of corpus-based linguistics studies is to create word lists to improve language teaching. Powerful corpus software and computer programs have produced various vocabulary lists that serve language learners with different needs. Lessard-Clouston (2012) and Folse & Youngblood (2017) concluded extensive information about various ESL vocabulary lists, for example, the Academic Word List (AWL) (Coxhead, 2000) of academic vocabulary, the New General Service List of high frequency words for ESL learners

(Browne, Culligan, & Phillips, 2013), the Business Word List (Konstantakis, 2007) of Business vocabulary, and the Basic Engineering List (Ward, 2009) of vocabulary in Engineering.

Teachers, especially those in IEP and EAP programs want their students to master the appropriate vocabulary necessary to function in future academic courses. Similar to the AWL, which specifically addresses the academic vocabulary found in university textbooks, the findings of this study identified the most commonly used past participles in passive voice found in college STEM textbooks. Therefore, English teachers could take advantage of this list when teaching vocabulary in their classed.

Recommendations for Future Research

The findings of the current study are one of the first steps to supporting the necessary role of passive voice in STEM writing. In this study, confounding variables, such as text length and literary styles, were not considered in the statistical model. Future research studies could add covariates into data analyses to further provide comprehensive interpretation to the results.

When considering literary styles, the texts that have lower passive proportions are those written in narratives. For instance, the textbook *From College to Career* was mostly written in narratives because the author was telling a lot of personal experience and giving constructional advice to job-hunting students. Many imperative sentences, such as *Show recruiters that you have a strong vocabulary by varying the words and phrases*, *List internships, volunteer work, and unpaid summer jobs*, and *Be honest about what each situation entailed*, were used in the book. However, in imperative sentences, verbs cannot be in passive forms, thus the percentage of passive voice in such literary style displays lower usage rate than others written to illustrate, for example, procedures of operation, such as *Many of these structures were constructed on silt and soft clay layers* and *Single and double quotes can be used interchangeably*. The present

127

study touched upon how often passive voice was used quantitatively, yet without going into it qualitatively as in what contexts were passive voice typically used, and in what form.

Therefore, to validate and strengthen the conclusion drawn in the present study, researchers could replicate this study using more STEM textbooks at both undergraduate and postgraduate level, creating a larger corpus, adding covariates to the model, or conducting a mixed method study. Moreover, quantitative analyses could be done to examine the difference between passive voice usage in various disciplines, such as Arts and Humanities, Business Administration, Hospitality Management, etc., but not limiting to the areas related to STEM. These replications could provide suggestions on how to properly apply passive voice in reading and writing for students in different majors at different academic levels.

Conclusion

This study investigated the passive voice usage in college textbooks specifically in the areas of Science, Technology, Engineering, and Mathematics. Quantitative analyses showed that the overall percentage of passive voice usage was consistent with previous studies, which indicated that passive voice was not used sparingly as writing guides had suggested. In addition, common passive constructions were used differently among the examined disciplines, and the past participles intended to be used in passive voice also occurred differently across the disciplines. These findings imply that teachers and students should pay more attention to passive voice teaching and learning, and instructions of writing guides should be carefully provided according to different discipline-specific writing conventions and genres.

APPENDIX:
IRB APPROVAL LETTER

NOT HUMAN RESEARCH DETERMINATION

From:	**UCF Institutional Review Board #1** **FWA00000351, IRB00001138**
To:	**Huiyuan Luo**
Date:	**February 07, 2018**

Dear Researcher:

On 02/07/2018, the IRB determined that the following proposed activity is not human research as defined by DHHS regulations at 45 CFR 46 or FDA regulations at 21 CFR 50/56:

Type of Review:	Not Human Research Determination
Project Title:	PASSIVE VOICE USAGE IN UNDERGRADUATE STEM TEXTBOOKS
Investigator:	Huiyuan Luo
IRB ID:	SBE-18-13769
Funding Agency:	
Grant Title:	
Research ID:	N/A

University of Central Florida IRB review and approval is not required. This determination applies only to the activities described in the IRB submission and does not apply should any changes be made. If changes are to be made and there are questions about whether these activities are research involving human subjects, please contact the IRB office to discuss the proposed changes.

This letter is signed by:

Signature applied by Gillian Morien on 02/07/2018 02:12:36 PM EST

Designated Reviewer

REFERENCES

Active & passive voice. (n.d.). In *Texas A&M University Writing Center online*. Retrieved from

http://writingcenter.tamu.edu/Students/Writing-Speaking-Guides/Alphabetical-List-of-

Guides/Drafting/Active-Passive-Voice

Active and passive voice. (n.d.). In *Purdue OWL online*. Retrieved from

https://owl.english.purdue.edu/owl/owlprint/539/

Ahmad, J. (2012). Stylistic features of scientific English: A study of scientific research articles.

English Language and Literature Studies, 2(1), 47-55.

Alvin, L. P. (2014). The passive voice in scientific writing. The current norm in science journals.

Journal of Science Communication, 1, 1-16.

Amdur, R. J., Kirwan, J., & Morris, C. G. (2010). Use of the Passive voice in medical journal

articles. *American Medical Writers Association Journal, 25*(3), 98-104.

Anderson, M. (2015). Passive, active or both? Which voice to use when writing science reports.

Teaching Science, 61(1), 50-51.

Atkinson, D. (1996). The "philosophical transactions of the royal society of London." 1675-

1975: A sociohistorical Discourse Analysis. *Language in Society, 25*(3), 333-371.

Bailey, A. (2016). *The effect of extended instruction on passive voice, reduced relative clauses,*

and modal would in the academic writing of advanced English language learners.

Retrieved from Dissertation and Theses.

Baratta, A. (2009). Revealing stance through passive voice. *Journal of Pragmatics, 41*, 1406-

142.

Bazerman, C. (1988). *Shaping written knowledge: The genre and activity of the experimental*

article in science. Madison, WI: University of Wisconsin Press.

Bazerman, C. (1997). The life of genre, the life in the classroom. In W. Bishop & H. Ostrom

    (eds.), *Genre and writing: Issues, arguments, and alternatives* (pp. 19-26). Portsmouth,

    NJ: Boynton/Cook.

Beason, L., & Lester, M. (2003). *A commonsense guide to grammar and usage*. Boston:

    Bedford/ St. Martin.

Bernard, M. (n.d.). *The importance of grammar in reading comprehension*. Retrieved from

    https://ortongillinghamonlinetutor.com/the-importance-of-grammar-in-reading-

    comprehension/

Berndt, S. R., Mitchum, C. C., Burton, M. W., & Haendiges, A. N. (2004). Comprehension of

    reversible sentences in aphasia: The effects of verb meaning. *Cognitive Neuropsychology,*

    21, 229-44.

Biber, D. (1993). Using register-diversified corpora for general language studies. *Computational*

    Linguistics, 19, 219–241.

Biber, D. (2004). Historical patterns for the grammatical marking of stance. *Journal of Historical*

    Pragmatics, 5(1), 107-136.

Biber, D., & Finegan, E. (1994). Intra-textual variation within medical research articles. In N.

    Oostdijk and P. de Haan (Ed.), *Corpus-Based Research into Language: In Honour of Jan*

    Aarts (pp. 201–222). Amsterdam: Rodopi.

Biber, D., Concord, S., & Leech, G. (2003). *Longman student grammar of spoken and written*

    English. Harlow, Essex: Pearson Education Ltd.

Biber, D., Concord, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language*

    structure and us. Cambridge: Cambridge University Press.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman Grammar of Spoken and Written English*. London: Longman.

Biondi, C. (2002). Active versus passive sentence construction: Which is the best choice? *American Communication Journal, 5*(2), 1-4.

Bock, J., & Warren, R. K. (1985). Conceptual accessibility and syntactic structure in sentence formulation. *Cognition, 21*, 47–67. doi:10.1016/0010-0277(85)90023-X

Browne, C., Culligan, B., & Phillips, J. (2013). *The New General Service List*. Retrieved from http://www.newgeneralservicelist.org/

Brysbaert, M. & New, B. (2009). Moving beyond Kucera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, 41*(4), 977-990.

Carter, R., & McCarthy, M. (1999). The English get-passive in spoken discourse: Description and implications for an interpersonal grammar. *English Language and Linguistics*, *3*, 41–58.

Carter, R., & McCarthy, M. (2006). *Cambridge grammar of English: A comprehensive guide: Spoken and written English grammar and usage*. Cambridge: Cambridge University Press.

Celce-Murcia, M. (1992). Formal grammar instruction: An educator comments. *TESOL Quarterly*, *26*, 406-409.

Celce-Murcia, M., & Larsen-Freeman, D. (2015). *The grammar book: An ESL/EFL teacher's course (2nd edition.).* Boston, MA: Cengage Learning.

Chomsky, N. (1969). *Aspects of the theory of syntax*. Cambrideg, MA: MIT Press.

Chomsky, N. (1980). *Rules and representations*. Oxford: Basil Blackwell.

Christianson, K., & Ferreira, F. (2005). Conceptual accessibility and sentence production in a

      free word order language (Odawa). *Cognition, 98*, 105–135.doi:

      10.1016/j.cognition.2004.10.006

Close, K. (2016, January 21). STEM majors will earn highest starting salaries this year. *Money*.

      Retrieved from http://time.com/money/4189471/stem-graduates-highest-starting-salaries/

Cobb, T. (2004). The compleat lexical tutor. *TESL-EJ, 8* (3). Retrieved from http://tesl-

      ej.org/ej31/m2.html

Cobb, T. (2017). *Web frequency indexer v 1.3* (computer program). Accessed 11 July 2017 at

      http://www.lextutor.ca/freq/eng/

Coetzee, J. (1980). The rhetoric of the passive in English. *Linguistics, 18*, 199-221.

Cohen, J. (1992). Power primer. *Psychological Bulletin, 112*, 155-159.

Collins, P. C. (1996). Get-passives in English. *World Englishes, 15*, 43–56. doi: 10.1111/j.1467-

      971X.1996.tb00091.x

Cowan, R., Choi, H. E., & Kim, D. H. (2003). Four questions for error diagnosis and correction

      in CALL. *CALICO Journal, 20*(3), 451-463.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34*(2), 213–238.

Day, R. A. (1989). *How to write and publish a scientific paper*. Cambridge University Press.

Day. R. A., & Sakaduski. N. (2011). *Scientific English: A guide for scientists and other

      professionals* (3rd edition). Santa Barbara, CA: Greenwood.

Deroey, K., & Taverniers, M. (2011). A corpus-based study of lecture functions. *Moderna språk,

      105*, 1-22. doi: 10.1002/9781405198431.wbeal0974

Diessel, H. (2004). *The acquisition of complex sentences*. Cambridge: Cambridge University

      Press.

Ding, D. D. (2002). The passive voice and social values in science. *Journal of Technical Writing and Communication, 32*(2), 137-154.

Emeksiz, Z. E. (2015). Stance taking and passive voice in Turkish academic discourse. *Journal of Language and Linguistics Studies, 11*(2), 13-22.

Fang, Z. (2005). Science literacy: A systemic functional linguistics perspective. *Science Education, 89*, 335-347.

Fang, Z. (2007). The language demands of science reading in middle school. *International Journal of Science Education, 28*(5), 491-520. doi: 10.1080/09500690500339092

Fang, Z., & Coatoam, S. (2013). Disciplinary literacy: What you want to know about it. *Journal of Adolescent & Adult Literacy, 56*(8), 627-632. doi: 10.1002/JAAL.190

Fang, Z., & Schleppegrell, M. J. (2010). Disciplinary literacies across content areas: Supporting secondary reading through functional language analysis. *Journal of Adolescent & Adult Literacy, 53*(7), 587-597. doi:10.1598/JAAL.53.7.6

Fang, Z., & Wei, Y. (2010). Improving middle school students' science literacy through reading infusion. *The Journal of Educational Research, 103*(4), 262-273. doi: 10.1080/00220670903383051

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2013). G*Power Version 3.1.7 [computer software]. Uiversität Kiel, Germany. Retrieved from http://www.psycho.uni-duesseldorf.de/abteilungen/aap/gpower3/download-and-register

Ferreira, F. (1994). Choice of passive voice is affected by verb type and animacy. *Journal of Memory and Language, 33,* 715–36.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology 47*, 164–203.

Folse, S. K. (2004a). Myths about teaching and learning second language vocabulary: What recent research says. *TESL Reporter, 37*(2), 1-13.

Folse, S. K. (2004b). *Vocabulary myths: Applying second language research to classroom teaching.* Ann Arbor, MI: University of Michigan Press.

Folse, S. K. (2015). *Clear grammar 3: Keys to grammar for English language learners* (2nd ed.). Ann Arbor, MI: University of Michigan Press.

Folse, S. K. (2016). *Keys to teaching grammar to English language learners* (2nd ed.). Ann Arbor, MI: The University of Michigan Press.

Folse, S. K., & Youngblood, A. (2017.) Survey of corpus-based vocabulary lists for TESOL classes. *MexTESOL Journal, 41*(1), 1-15.

Fraenkel, J.R. & Wallen, N.E. (2009). *How to design and evaluate research in education.* (7th ed.). New York: McGraw-Hill.

Frazer, A. K. & Miller, M. D. (2009). Double standards in sentence structure: Passive voice in narratives describing domestic violence. *Journal of Language and Social Psychology, 28*(1), 62-71.

Gall, M. D., Gall, J., & Borg, W. R. (2007). *Educational research: An introduction* (8th edition). Boston, MA: Allyn & Bacon.

Gaux, C., & Gombert, J. E. (1999). Implicit and explicit syntactic knowledge and reading in preadolescents. *Journal of Developmental Psychology, 17*, 169-188.

Gennari, S.P., Mirkoví́c, J., & MacDonald, M. C. (2012). Animacy and competition in relative clause production: across-linguistic investigation. *Cognitive Psychology, 65*, 141–176. doi: 10.1016/j.cogpsych.2012.03.002

Glass, G. V. (1972). Consequences of failure to meet assumptions underlying the fixed effects

analyses of variance and covariance. *Review of Educational Research, 42*(3), 237-288.

Goldenberg, C. (2010). Reading instruction for English language learners. In M. Kamil, P. D.

Pearson, E. Moje, & P. Afflerbach (Eds.). *Handbook of Reading Research, Vol. IV*.

Newark, DE: International Reading Association.

Grabe, W. & R. B. Kaplan (1996). *Theory and practice of writing*. London: Longman.

Granger, S. (2013). The passive in learner English corpus insights and implications for

pedagogical grammar, in Ishikawa (1ˢᵗ ed.). *Learner corpus studies in Asia and the world.*

School of languages and communication (pp. 5-15). Kobe University: Kobe.

Graves, M. E, Juel, C., & Graves, B. B. (2001). *Teaching reading in the 21st century* (2nd ed.).

Boston, MA: Allyn & Bacon.

Haas, F. (2008). Symmetric verbs and constraints on passivization: An English-German

comparison. *Zeitschrift für Anglistik und Amerikanistic, 56*(3), 269-282.

Halliday, M. A. K. (1970). Theme and information in the English clause, in Kress, G. (ed.)

(1976). *Halliday: system and function in language* (pp. 174-188). London: OUP.

Halliday, M. A. K., & Matthiessen, C. M. I. M. (2004). *An introduction to functional grammar*

(3rd edition). London: Arnold.

Hameed, A. N. (2017). *The use of the passive voice in the writing of Arab ESL learners*

(Master's thesis). Retrieved from

http://studentsrepo.um.edu.my/7207/2/FINAL_THESIS.pdf

Harmon, J. E. (1992). An analysis of fifty citation superstars from the scientific literature.

Journal of Technical Writing and Communication, 22, 17-37.

Hasko, V. (2012). Qualitative corpus analysis. In C. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics.* Malden, MA: Wiley-Blackwell.

Hsiao, Y., Gao, Y., & MacDonald, M. C. (2014). Agent-patient similarity affects sentence structure in language production: Evidence from subject omissions in Mandarin. *Frontiers in Psychology, 5.* doi: 10.3389/fpsyg.2014.01015. Retrieved from https://www.frontiersin.org/articles/10.3389/fpsyg.2014.01015/full

Hundt, M. (2001). What corpora tell us about the grammaticalisation of voice in 'get'-constructions. *Studies in Language, 25*(1), 49-87.

Hundt, M., Schneider, G., & Seoane, E. (2016). The use of the be-passive in academic Englishes: Local versus global usage in an international language. *Corpora, 11*(1), 29-61.

Jaeggli, O. A. (1986). Passive. *Linguistic Inquiry, 17*(4), 587-622.

Johns, A. M. (1991). Faculty assessment of ESL student literacy skills: Implications for writing assessment. In L. Hamp-Lyons (Ed.), *Assessing second language writing in academic contexts* (pp. 167-180). Norwood, NJ: Ablex.

Johns, A. M. (2008). Genre awareness for the novice academic student: An ongoing quest. *Language Teacher, 41*(2), 237-252. doi:10.1017/S0261444807004892

Johnson, N. H., & Lyddon, P. A. (2016). Teaching grammatical voice to computer science majors: The case of less proficient English learners. *English for Specific Purposes, 41*, 1-11.

Kalton, G. (2017). Systematic Sampling. *Wiley StatsRef: Statistics Reference Online*, 1–6.

Kern, R. G. (1989). Second language reading strategy instruction: Its effects on comprehension and word inference ability. *The Modern Language Journal, 73*(2), 135-149. Retrieved from http://www.jstor.org/stable/326569

Kharkwal, G., & Stromswold, K. (2014). Good-enough language processing: Evidence from

    sentence-video matching. *Journal of Psycholinguistic Research, 43*(1), 27-43. doi:

    10.1007/s10936-013-9239-5

Kim, B. (2002). The role of focus on form in the EFL classroom setting: Looking into recast and

    CR tasks. *English Teaching, 57*, 267–296.

Kim, Y., & McDonough, K. (2016). Prime repetition and Korean EFL learners' comprehension

    and production of passives. *International Review of Applied Linguistics in Language*

    Teaching, 54(4), 319-346. doi: 10.1515/iral-2015-0028

King, S. (2000). *On writing: A memoir of the craft.* New York City, NY: Simon and Schuster.

Konstantakis, N. (2007). Creating a business word for teaching Business English. *Elia: Estudios*

    de lingüística inglesa aplicada, 7, 79-102.

Krisch, J., & Houdek, F. (2015). The myth of bad passive voice and weak words: An empirical

    investigation in the automotive industry. Proceedings from RE '15: *23rd International*

    Requirements Engineering Conference. Ottawa, Canada.

Lee, (2007). Effects of textual enhancement and topic familiarity on Korean EFL students'

    reading comprehension and learning of passive form. *Language Learning, 57*(1), 87-118.

Lessard-Clouston, M. (2012). Word lists for vocabulary learning and teaching. *The CATESOL*

    Journal, 24(1), 287-304.

Levelt, W. J., Schriefers, H., Vorberg, D., Meyer, A. S., Pechmann, T., & Havinga, J. (1991).

    The time course of lexical access in speech production: A study of picture naming.

    Psychological Review, 98(1), 122–142. doi:10.1037/0033-295x.98.1.122.

Lightbown, P. P., & Pienemann, M. (1993). Comments on Stephen D. Krashen's "Teaching

    issues: Formal grammar instruction". *TESOL Quarterly*, *27*, 717-722.

Liu, D., & Nelson, R. (2016). *Teaching language as a system*. In G. Hall (Eds.), The Routledge handbook of English language teaching (pp. 413-427). London and New York: Routledge.

Loebell, K. H., & Morey, R. (1992). From conceptual roles to structural relations: Bridging the syntactic cleft. *Psychological Review, 99*, 150–171.doi: 10.1037/0033-295X.99.1.150

LoMaglio L. J., & Robinson V. J. (1985). The impact of passive voice on reading comprehension. *IEEE Transactions on professional communication, 28*(4).

MacDonald, M. C. (2013). How language production shapes language form and comprehension. *Frontiers in Psychology, 4*, 226. doi:10.3389/fpsyg.2013.00226

Mahon, B., Costa, A., Peterson, R., Vargas, K., & Caramazza, A. (2007). Lexical selection is not by competition: A reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(3), 503–535.

Matthews, J. R., Bowen, J. M., & Matthews, R. W. (1996). *Successful Scientific Writing*. Cambridge University Press.

McNeil, D. (1966). The genesis of language: A psycholinguistic approach. In F. Smith and G. Miller (Ed.), *Developmental Psycholinguistics* (pp. 15-84). Cambrideg: MIT Press.

Millar, N., Budgell, B., & Fuller, K. (2012). Use the active voice whenever possible: The impact of style guidelines in medical journals. *Applied Linguistics, 34*(4), 393-414.

Mindt, D. (2000). *An Empirical Grammar of the English Verb System*. Berlin: Cornelsen Verlag GmbH & Co.

Mirdamadi, F. S., & De Jong, N. H. (2015). The effect of syntactic complexity on fluency: Comparing actives and passives in L1 and L2 speech. *Second Language Research, 31*(1), 105-116.

Moreb, B. (2016). *The frequency of the passive voice in freshman academic books* (Unpublished master's thesis). University of Central Florida, Orlando.

Nagy, W., & Townsend, D. (2012). Words as tools: Learning academic vocabulary as language acquisition. *Reading Research Quarterly, 47*(1), 91-108.

Nation, K., & Snowling, M. J. (2000). Factors influencing syntactic awareness skills in normal readers and poor comprehenders. *Applied Psycholinguistics, 21*, 229-241.

Oshita, H. (2000). What is happened may not be what appears to be happening: a corpus study of 'passive' unaccusatives in L2 English. *Second Language Research, 16*(4), 293-324.

Passive voice: What it is and how to reduce usage. (2013). In *University Writing Center of Nevada, Reno online*. Retrieved from https://www.unr.edu/Documents/liberal-arts/writing-center/Editing-Passive-Voice.pdf

Passive voice. (n.d.). In *The Writing Center at UNC-Chapel Hill online*. Retrieved from http://writingcenter.unc.edu/handouts/passive-voice/

Pitts, J. (2010). Zombies and Vampires or the passive and active voice. In *The University of Central Florida Writing Center online*. Retrieved from http://uwc.cah.ucf.edu/wp-content/uploads/sites/9/2015/04/Passive_Voice_Zombies-Vampires.pdf

Prystupa, V. (2017). An investigation of Chinese students' grammar developmental sequence: A corpus study of academic writing. *Theoretical and Applied Linguistics, 3*(1), 71-75. Doi: 10.18414/2313-8912-2017-3-1-71-75

Pullum, G. K. (2009). 50 years of stupid grammar advice. *The Chronicle of Higher Education, 55*(32), B15.

Pullum, G. K. (2014). Fear and loathing of the English passive. *Language & Communication, 37*, 60-74. doi: 10.1016/j.langcom.2013.08.009

Riley, K. (1991). Passive voice and rhetorical role in scientific writing. *Journal of Technical Writing and Communication, 24,* 239-257.

Rodman, L. (1994). The active voice in scientific articles: Frequency and discourse functions. *Journal of Technical Writing and Communication, 24*, 309-331.

Rosinski, R. R., Golinkoff, R. M., & Kukish, K. S. (1975). Automatic semantic processing in a picture-word interference task. *Child Development. 46*, 247–253. doi: 10.2307/1128859

Safire, W. (1992, January 19). On language; tense encounter. *The New York Times.* Retrieved from http://www.nytimes.com/1992/01/19/magazine/on-language-tense-encounter.html

Schoonmaker, A. (2015). *Increasing metalinguistics awareness as a necessary precursor for preservice teachers* (Doctoral dissertation). Retrieved from http://stars.library.ucf.edu/cgi/viewcontent.cgi?article=2469&context=etd

Schwarz, S. (2015). Passive voice in American soap opera dialogue. *Studia Neophilologica, 87*(2), 152-170. doi: 10.1080/00393274.2015.1049831

Segaert, K., Menenti, L., Weber, K., & Hagoort, P. (2011). A paradox of syntactic priming: Why response tendencies show priming for passives and response latencies show priming for actives. *PLOS ONE, 6*(10), 1-14.

Seoane, E. (2013). On the conventionalization and loss of pragmatic function of the passive in late modern English scientific discourse. *Journal of Historical Pragmatics, 14*(1), 70-99. DOI: 10.1075/jhp.14.1.03seo

Sheen, A. P. (1982). *Breathing Life into Medical Writing: A Handbook*. Mosby.

Shen, H. (2003). The role of explicit instruction in ESL/EFL reading. *Foreign Language Annals, 36*(3), 424-433.

Sigel, T. (2009). How passive voice weakens your scholarly argument. *Journal of Management Development, 28*(5), 478-480. doi: 10.1108/02621710910955994

Stone, J. A. M. (2014). A study on scientific writing: Active versus passive voice. *American Acupuncturist, 66,* 20.

Street, J. A. & Dąbrowska, E. (2010). More individual differences in language attainment: How much do adult native speakers of English know about passives and quantifiers? *Lingua, 120*, 2080–94.

Street, J. A., & Dąbrowska, E. (2006). Individual differences in language attainment: Comprehension of passive sentences by native and non-native English speakers. *Language Sciences, 28*(6), 604-615. doi: 10.1016/j.langsci.2005.11.014

Strunk, W. & White, E. B. (2000). *The Elements of Style.* Harlow: Longman.

Sutton, C. R. (1992). *Words, Science and Learning.* Buckingham: Open University Press.

Swales, J. M. (1990). *Genre Analysis: English in Academic and Research Settings*. Cambridge University Press.

Tabachnick, B., & Fidell, L. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Pearson Education.

Tabachnick, B., & Fidell, L. (2013). *Using multivariate statistics* (6th ed.). Boston, MA: Pearson Education.

Tanaka, M. N., Branigan, H. P., McLean, J. F., & Pickering, M. J. (2011). Conceptual influences on word order and voice in sentence. *Journal of Memory and Language, 65*, 318-330. doi: 10.1016/j.jml.2011.04.009

Taylor, A., Stevens, J., & Asher, J. (2006). The effects of Explicit Reading Strategy Training on L2 reading comprehension. In J. Norris & L. Ortega (Eds.). *Synthesizing research on language learning and teaching* (pp. 213-244). Philadelphia, PA: John Benjamins.

Taylor, R. B. (2005). *The Clinician's Guide to Medical Writing*. Berlin: Springer.

The National Science Board. (2014). *What percentage of freshmen intend to major in an S&E field when they start college*? Retrieved from https://nsf.gov/nsb/sei/edTool/data/college-09.html

Thompson, D., Ling, S. P., Myachykov, A., Ferreira, F. & Scheepers, C. (2013). Patient-related constraints on get- and be-passive uses in English: Evidence from paraphrasing. *Frontiers in Psychology, 4*(848), 1-24. doi: 10.3389/fpsyg.2013.00848

Toyota, J. (2009). Fossilisation of passive in English: Analysis of passive verbs. *English Studies, 90*, 476-497.

Tseng, J. (2007). English Prepositional Passive Constructions. In Stefan Muller (Ed.): Proceedings of the 14th International Conference on Head-Driven Phrase Structure Grammar, *Stanford Department of Linguistics and CSLI's LinGO Lab* (pp. 271–286). Stanford, CA: CSLI Publications.

Turner, E. A., & Rommetveit, R. (1967). Experimental manipulation of the production of active and passive voice in children. *Language & Speech, 10*(3), 169-180.

Wang, L. (2007). Theme and rheme in the thematic organization of text: Implications for teaching academic writing. *Asian EFL Journal, 9*(1), 164-176.

Wasow, T. (1977). Transformations and the lexicon, in P. Culicover, T. Wasow, and A.

  Akmajian (Eds.). *Formal Syntax*. New York, NY: Academic Press.

Watson, J. D., & Crick, F. H. C. (1953). Molecular structure of nucleic acids: A structure for

  deoxyribose nucleic acid. *Nature, 171*, 737-738.

Wellington, J., & Osborne, J. (2001). *Language and literacy in science education*. Philadelphia,

  PA: Open University Press.

Wilkinson, A. W. (1992). Jargon and the passive voice: Prescriptions and Proscriptions for

  Scientific Writing. *Journal of Technical Writing and Communication, 22,* 319-325.

Williams, J. M., & Colomb, G. G. (2010). *Style lessons in clarity and grace.* United States of

  America: Longman.

Xiao, R., McEnery, T., & Qian, Y. (2006). Passive constructions in English and Chinese: A

  corpus-based contrastive study. *Languages in Contrast, 6*(1), 109–149. doi:

  10.1075/lic.6.1.05xia

Xue, Y., & Larson, R. (2015). STEM crisis or STEM surplus? Yes and yes. *Monthly Labor*

  Review, U.S. Bureau of Labor Statistics. Retrieved from

  https://doi.org/10.21916/mlr.2015.14.

Yannuar, N., Shitadevi, I. A., Basthomi, Y., & Widiati, U. (2014). Active and passive voice

  constructions by Indonesia student writers. *Theory and Practice in Language Studies,*

  4(7), 1400-1408. doi: 10.4304/tpls.4.7.1400-1408

Yeo, K. (2002). The effects of dictogloss: A technique of "focus on form". *English Teaching, 57,*

  149-167.

Yu, X., Luo, H., & Folse, S. K. (2016). *The use of passive voice in college introductory course*

  textbooks. Manuscript in preparation.

Zhang, D. (2012). Vocabulary and Grammar Knowledge in Second Language Reading

    Comprehension: A Structural Equation Modeling Study. *The Modern Language Journal,*

    96(4), 558-575.

Zygouris-Coe, V. (2009). Teaching reading comprehension skills. *Middle Matters, 17*(4), 1-2.