EDUCATION
RESEARCH SERVICE

ORIGINAL ARTICLE

The passive voice and comprehensibility of biomedical texts: *An experimental study with 2 cohorts of chiropractic students*

Neil Millar, PhD and Brian S. Budgell, DC, PhD

Objective: Authors in the health sciences are encouraged to write in the active voice in the belief that this enhances comprehensibility. Hence, the purpose of this study was to compare objectively measured and subjectively perceived comprehensibility of texts in which one voice or the other was highly prevalent.
Methods: Objectively rated comprehensibility was obtained by presenting 161 2nd-year chiropractic students with questions pertaining to 2 methods sections of biomedical articles, each presented in its original form with high prevalence of the passive voice, and in a manipulated form with all main verbs in the active voice. The difficulties and sensitivities of questions were compared for the 2 forms of each text. Comprehensibility was obtained by asking students to rate the comprehensibility of authentic sentences from biomedical manuscripts and matched manipulated form in which the voice of the main verb had been changed. Differences in comprehensibility between the 2 texts were assessed with a dependent t test.
Results: There were no significant differences in the difficulties or sensitivities of questions pertaining to the 2 original texts written in the passive voice versus the active voice ($p > .35$ for all comparisons). Students rated sentences written in the passive voice as marginally more comprehensible than sentences written in the active voice ($p = .003$ per 2-tailed paired t test).
Conclusion: The texts written in the active voice were not more comprehensible than texts written in the passive voice. The results of this study do not support editorial guidelines that favor active voice over passive voice.

Key Indexing Terms: Literacy; Reading; Linguistics; Education; Publishing; Chiropractic Education

J Chiropr Educ 2019;33(1):16–20 DOI 10.7899/JCE-17-22

INTRODUCTION

Writings in the health sciences tend to be complex. This is natural, given what are often the relatively esoteric themes and the argumentative (in a scientific sense) nature of writings in the health sciences; for example, establishing cause-effect relationships from samples of populations. Efforts to render such writing more comprehensible are therefore laudable, especially as English becomes the lingua franca of health sciences education and research worldwide. As increasing numbers of health professionals undergo training in English, it becomes apparent that the language of their chosen discipline may present challenges.[1,2] This is likely as true for chiropractic students as for students in other areas of the health sciences.[3] Furthermore, even health sciences students who are so-called native English speakers may be challenged by the vocabulary and writing conventions in their chosen discipline.[4]

Thus, the more understandable learning resources are, the more efficient the learning process will be.[5] The overabundance of articles on how to write readable medical prose is testimony to this conventional wisdom (eg, see Alexandrov,[6] Bredan and van Roy,[7] and Steen[8]). However, commonly proposed strategies for comprehensible writing may not be grounded in good evidence. For example, it is not uncommon to see writers encouraged to favor the active voice over the passive voice—"the doctor treated the patient" versus "the patient was treated by the doctor"—perhaps based on the misconception that the passive voice results in longer and less-comprehensible constructions. Even the *American Medical Association Manual of Style* continues to encourage authors to "use active voice whenever possible,"[9] while the *British Medical Journal* still advises authors to "use active voice but avoid 'we did' or 'we found.'"[10]

On the other hand, it has recently been argued that there are instances in medical writing in which the passive voice is preferable to the active voice, and that sweeping

Table 1 - Characteristics of 2 Convenience (Nonrandomized) Cohorts in Study of Comprehensibility of Passive Versus Active Voice Texts

	Cohort 1 ($n = 79$)	Cohort 2 ($n = 82$)
Sex	39 males, 40 females	38 males, 44 females
Mean age, y (range)	25 (22–44)	24 (21–30)
Native English speakers	54 (68%)	59 (72%)
Completed high school in English	73 (92%)	79 (96%)
Obtained undergraduate degree	44 (54%)	42 (51%)
Nonscience/health degrees	2	2

editorial guidelines may be overly simplistic and so do not serve their intended purpose.[11] To date, however, no experimental studies have challenged the assumption of enhanced comprehensibility for the active voice. In this age of evidence-based care, it is somewhat ironic that the teaching of health sciences literacy remains largely opinion based. Hence, the purpose of this study was to compare the comprehensibility of authentic samples of biomedical text expressed in the active and passive voices.

METHODS

This study was approved by the research ethics board of Canadian Memorial Chiropractic College. As part of an annual assessment of incoming students, 161 2nd-year chiropractic students were recruited into the study by an in-class announcement on September 10, 2012. Once seated in the testing room and in sequence according to their seating, the students alternatively received 1 of 2 forms of the assessment. Hence, the allocation to cohorts was not randomized. Nonetheless, post hoc analysis of the cohorts showed that they were quite similar (Table 1).

The assessments used 2 strategies to test the effect of the passive voice on the comprehensibility of biomedical texts.

1. Objectively Rated Comprehensibility

In the first instance, each of the 2 cohorts of students was presented with 1 sample of text in its original published form, with frequent use of the passive voice, and a sample of text in which the passive voice was replaced throughout the text with the active voice (text available from corresponding author). The text samples were extracted from the methods sections of 2 original research papers referenced within the undergraduate curriculum of the students.[12,13] The comprehensibility of the 2 examples of biomedical text was tested in the original published form and in the rewritten form. Thus, in more detail, 1 cohort received the original version of text 1 and the rewritten version of text 2. The 2nd cohort received the rewritten version of text 1 and the original version of text 2.

The readability statistics for the texts were obtained from the spelling and grammar checking function of Microsoft Word 2007 (Microsoft Corp, Redmond, WA), and as shown in Table 2, they were essentially identical for the original and rewritten versions of the texts. Additionally, the profile of the vocabulary was obtained by processing the text samples using the online software Web VP Classic v.4.0.[14] Thus, the words within each version of each text sample were classified as belonging to (1) the general service list, the approximately 2000 most commonly occurring word families in the English language; (2) the academic word list, the approximately 570 word families, other than general service list words, commonly found in academic writings; and (3) all other words not belonging to the general service list or academic word list, so-called "off-list" words. Off-list words are more likely to be technical words specific to the theme of the text. As shown in Table 2, the vocabulary profiles were essentially unaltered by rewriting the sample texts in the active voice.

The students were asked to answer 10 true-or-false questions dealing with each of the 2 texts. Applying item response theory, the difficulties and sensitivities of each of the 10 questions were calculated for each of the 2 forms (active voice versus passive voice) of the text. The difficulty of each question was defined as the mathematical inverse of the prevalence of correct answers. For example, if half of the students answered a question correctly, that question would have a difficulty of 2.0 (the inverse of 1/2). The sensitivity of each question was defined as the prevalence of correct answers among the top-scoring quartile of the cohort divided by the prevalence of correct answers among the bottom-scoring quartile of the cohort.

Comparisons were also made between the raw and processed (see below) scores of the 2 cohorts for each of the 2 versions (original published form and rewritten form) of each of the 2 texts, essentially asking whether cohort scores were on average different depending on whether the text was presented in the active or passive voice. Raw scores were calculated simply as the percentage of correct answers. Processed scores were calculated by assigning each correct answer a weight equivalent to its calculated difficulty and then dividing the student's score by the total weight of all questions.

2. Subjectively Rated Comprehensibility

Twenty sentences were extracted from the Biomedical and Health Linguistics randomized controlled trials corpus[11] and presented in their original form (active or passive voice) to 1 cohort and then in the alternate (rewritten) form (text available from corresponding author) to a second cohort; for example:

Active voice version presented to 1 cohort: Question 21. "We also observed this minimal level of correlation in the subgroup of patients who subsequently had recurrent coronary events ($r = 0.1$, $P = 0.004$)." *New England Journal of Medicine*

Passive voice version presented to the other cohort: Question 21. "This minimal level of correlation was also observed in the subgroup of patients who subsequently had recurrent coronary events ($r = 0.1$, $P = 0.004$)." *New England Journal of Medicine*

Each version of the test contained the same number of sentences in each voice. Hence, within each cohort (ie,

Table 2 - Readability Statistics and Vocabulary Profiles for Texts in Original (Passive Voice) Versions and Rewritten (Active Voice) Versions

Text	1. Cavanaugh et al[12]		2. Dishman et al[13]	
Statistic\voice	Passive voice	Active voice	Passive voice	Active voice
Words (tokens)	357	357	304	305
Characters	1857	1830	1591	1582
Paragraphs	1	1	1	1
Sentences	18	18	14	14
Sentences/paragraph	18.0	18.0	14.0	14.0
Words/sentence	19.8	19.8	21.7	21.7
Characters/word	5.0	5.0	5.1	5.0
Passive sentences	72%	0%	50%	0%
Flesch reading ease	36.6	36.6	39.0	39.3
Flesch-Kincaid grade level	13.0	13.0	13.1	13.1
General service list words	63%	63%	67%	67%
Academic word list words	11%	11%	7%	7%
Off-list words	26%	26%	26%	26%
Types (unique words)	177	176	146	147
Type-token ratio	0.48	0.48	0.47	0.47

version of the test), students were presented with 10 original and 10 rewritten sentences, and among these sentences, 10 were in the active voice and 10 were in the passive voice. The sentences came from the "big five" medical journals: *Journal of the American Medical Association*, *New England Journal of Medicine*, *Annals of Internal Medicine*, *British Medical Journal*, and the *Lancet*. Students were asked to subjectively rate the comprehensibility of each sentence using the following scale: 1, completely clear; 2, quite clear; 3, moderately clear; 4, quite unclear; 5, completely unclear.

Seventy-eight complete copies of test 1 and 82 complete copies of test 2 were returned. The numerical ratings for the 2 versions of each sentence were averaged, and the averages for the active-voice and passive-voice versions of each sentence were compared using a paired, 2-tailed *t* test.

RESULTS

1. Objectively Rated Comprehensibility

The readability statistics and vocabulary profiles for the 2 texts are presented in Table 2 and suggest that the readabilities of the 2 versions of each text (supplementary files) were not substantially different.

Table 3 shows, for text 1, the calculated difficulties and sensitivities of the 10 true-or-false questions when the text was presented in the passive versus the active voice. There was no statistically significant difference, per paired, 2-tailed *t* test, between the difficulties and sensitivities ($p =$ .69 and .35, respectively) for the true-or-false questions when tested against text 1 in the passive versus active voice.

Table 4 shows, for text 2, the calculated difficulties and sensitivities of the 10 true-or-false questions when the text was presented in the passive versus the active voice. There was no statistically significant difference, per paired, 2-tailed *t* test, between the difficulties and sensitivities ($p =$ .48 and .81, respectively) for the true-or-false questions when tested against text 2 in the passive versus active voice.

2. Subjectively Rated Comprehensibility

As shown in Table 5 and on the basis of subjectively rated comprehensibility, the versions of sentences presented in the active voice were, on average, somewhat less comprehensible (mean 2.21, SD 0.48) than the versions presented in the passive voice (mean 2.07, SD 0.47; $p = .003$ per 2 tailed, paired *t* test).

DISCUSSION

This study challenged the hypothesis that texts written in the active voice are more comprehensible than texts written in the passive voice. The study was conducted using students who were just commencing their 2nd year of study at a Canadian chiropractic college. Hence, all students already possessed at least an undergraduate university degree and had completed 1 year (approximately 1000 classroom hours) of additional study focusing on

Table 3 - Difficulty and Sensitivity for Each of 10 Questions Pertaining to Text 1, Which Was Presented in the Passive (Original) and Active (Manipulated) Voice

	Difficulty		Sensitivity	
Item	Original	Manipulated	Original	Manipulated
1	1.25	1.19	1.73	1.67
2	1.00	1.06	1.00	1.00
3	1.14	1.11	1.46	1.33
4	1.36	1.52	2.00	1.70
5	1.16	1.11	1.73	1.33
6	1.30	1.24	1.64	1.19
7	1.11	1.09	1.36	1.33
8	1.01	1.05	1.06	1.11
9	1.36	1.17	1.42	1.33
10	1.16	1.19	1.46	2.00
Mean (SD)	**1.19 (0.13)**	**1.17 (0.14)**	**1.49 (0.31)**	**1.40 (0.30)**

Table 4 - Difficulty and Sensitivity for Each of 10 Questions Pertaining to Text 2, Which Was Presented in the Passive (Original) and Active (Manipulated) Voice

	Difficulty		Sensitivity	
Item	Original	Manipulated	Original	Manipulated
1	1.03	1.00	1.11	1.00
2	1.06	1.03	1.12	1.06
3	1.11	1.08	1.33	1.36
4	1.01	1.07	1.05	1.19
5	1.00	1.04	1.00	1.12
6	1.08	1.05	1.18	1.27
7	1.64	1.58	3.17	3.20
8	1.39	1.80	2.38	4.25
9	3.15	3.16	3.33	2.20
10	1.86	1.84	3.60	3.20
Mean (SD)	**1.43 (0.67)**	**1.47 (0.68)**	**1.93 (1.07)**	**1.99 (1.17)**

human biology but also involving a yearlong course in critical appraisal of research literature. The texts presented in this experiment were therefore representative of the reading materials which the students would be expected to access in their current year of study.

Objective measures of readability and vocabulary profiles were obtained for 2 samples of text which, in their original forms, made extensive use of the passive voice, as is conventional in biomedical texts.[11] When these texts were rewritten to eliminate the use of the passive voice, the changes in readability measures were miniscule or none. This contradicts the argument that the use of the passive voice produces more-complex or less-comprehensible texts. Furthermore, when students were asked questions pertaining to the original (passive voice) versus rewritten (active voice) versions of the 2 texts, there were no statistical differences in the mean difficulties or sensitivities of the questions. The individual comprehension questions satisfactorily discriminated between different levels of reading ability among the students. Specifically, 36 of 40 questions had sensitivities of >1.00, while 4 of 40 questions had sensitivities just equal to 1.00, and so overall the questions provide a valid measure of the students' reading comprehension. This indicates that the 2 versions of the same text were equally comprehensible to the students.

When students were asked to subjectively rank the comprehensibility of sentences written in the active versus passive voice, they indicated that overall sentences presented in the passive voice were somewhat (and statistically significantly) more comprehensible. When compared with the objective measures of comprehensibility, this seems to reflect a bias on the students' part in favor of the passive voice, and, given that students are familiar with medical texts that make frequent use of the passive voice, this general preference seems perhaps not surprising. On the other hand, it is not clear if students prefer to encounter particular verbs in 1 voice versus the other. Each of the main verbs in the sentences listed in Table 4 was presented 4 times—twice in the active voice and twice in the passive voice. The verb *obtain* was the only verb for

Table 5 - Subjectively Rated Comprehensibility for Sentences Presented in the Active and Passive Voice

Sentence*	Active Voice	Passive Voice
21 NEJM	2.5	2.6
22 Lancet	2.0	1.9
23 Lancet	2.9	2.8
24 NEJM	2.0	2.0
25 Lancet	1.4	1.5
26 JAMA	1.9	1.9
27 NEJM	2.1	2.0
28 Ann Int Med	3.0	2.6
29 JAMA	2.3	1.8
30 BMJ	2.2	1.8
31 JAMA	2.1	1.9
32 NEJM	1.6	1.6
33 BMJ	1.8	1.6
34 Ann Int Med	3.0	2.9
35 BMJ	1.8	1.4
36 Ann Int Med	2.4	2.5
37 Lancet	3.0	2.8
38 Ann Int Med	1.7	1.7
39 BMJ	2.4	2.3
40 JAMA	2.0	1.7
Mean (SD)	2.21 (0.48)	2.07 (0.47)

Abbreviations: NEJM, *New England Journal of Medicine*; JAMA, *Journal of the American Medical Association*; Ann Int Med, *Annals of Internal Medicine*; BMJ, *British Medical Journal*. *Number refers to the sentence number of the article retrieved from the journal listed.

which the passive voice was twice rated as much more comprehensible than the active voice. The verb *obtain* was previously shown to have a strong affinity for the passive voice in reports of randomized controlled trials.[11] On the other hand, other verbs which had previously shown a strong affinity for the passive voice (eg, associate, consider) were not consistently ranked as more comprehensible when presented in the passive voice in the present study. Hence, there is not a strong argument that students are conditioned by medical writing to prefer certain verbs in the passive voice.

A limitation of this study is that subjects were not randomized to the 2 test cohorts. Further, there may be important differences in the literacies of students at 1 institution versus another, or, indeed, 1 country versus another, limiting the generalizability of this work.

CONCLUSION

Collectively, these results indicate that for the test subjects and texts used in this study, use of 1 voice versus the other did not affect objectively rated comprehensibility, and that there was only a marginally greater subjectively rated comprehensibility for the passive voice. These results are congruent with a previous paper which proposed, on theoretical grounds, that the passive voice is appropriate in medical writings. Furthermore, these results argue against editorial guidelines which advocate for deliberate selection of the active voice in medical writings. Given that articles in leading journals and accessed by

health sciences students are already difficult to read,[4,15] editors of publications in the health sciences may wish to take a more evidence-based approach to developing recommendations on language use.

FUNDING AND CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest. No external funds were received in support of this work.

About the Authors

Neil Millar is an associate professor, faculty of engineering, information and systems, at the University of Tsukuba (1-1-1 Tennodai, Tsukuba, Ibaraki 305-8573, Japan; millar.neil.gm@u.tsukuba.ac.jp). Brian Budgell is a professor at the Canadian Memorial Chiropractic College (6100 Leslie Street, Toronto, Ontario, Canada M2H 3J1; bbudgell@cmcc.ca). Address correspondence to Brian Budgell, Canadian Memorial Chiropractic College, 6100 Leslie Street, Toronto, Ontario, Canada M2H 3J1; bbudgell@cmcc.ca. This article was received September 10, 2017, revised February 1, 2018, and accepted March 2, 2018.

Author Contributions

Concept development: BB, NM. Design: BB, NM. Supervision: BB, NM. Data collection/processing: BB, NM. Analysis/interpretation: BB, NM. Literature search: BB, NM. Writing: BB, NM. Critical review: BB, NM.

REFERENCES

1. Heming TA, Nandagopal S. Comparative difficulties with non-scientific general vocabulary and scientific/medical terminology in English as a second language (ESL) medical students. *Sultan Qaboos Univ Med J*. 2012;12(4):485–492.
2. Mann C, Canny BJ, Reser DH, Rajan R. Poorer verbal working memory for a second language selectively impacts academic achievement in university medical students. *Peer J*. 2013;1:e22. doi:10.7717/peerj.22.
3. Budgell B, Millar N, Fuller K. Assessing the literacy of modern chiropractic students. Proceedings of the 2016 World Federation of Chiropractic/Association of Chiropractic Colleges Education Conference; October 19–22, 2016; Montreal, Canada. *J Chiropr Educ*. 2016;30:e167.
4. Fuller S, Horlen C, Cisneros R, Merz T. Pharmacy students' reading ability and the readability of required reading materials. *Am J Pharm Educ*. 2007;71(6):111.
5. Olson MA. English-as-a-second language (ESL) nursing student success: a critical review of the literature. *J Cultural Diversity*. 2012;19(1):26–32.
6. Alexandrov AV. How to write a research paper. *Cerebrovas Dis*. 2004;18:135–138.
7. Bredan AS, van Roy F. Writing readable prose. *EMBO Reports*. 2006;7(9):846–849.
8. Steen RG. Writing for publication in a medical journal. *Indian J Endocrinol Metab*. 2012;16(6):899–903.
9. Christiansen S. Grammar. In: Iverson C, Christiansen S, Flanagin A, et al, eds. *American Medical Association Manual of Style*. 10th ed. New York, NY: Oxford University Press; 2009. doi:10.1093/jama/9780195176339.001.0001.
10. British Medical Journal. Article types at the BMJ. http://www.bmj.com/about-bmj/resources-authors/article-types. Accessed August 31, 2017.
11. Millar N, Budgell B, Fuller K. "Use the active voice whenever possible": the impact of style guidelines in medical journals. *Appl Linguistics*. 2013;34(4):393–414.
12. Cavanaugh JM, Lu Y, Chen C, Kallakuri S. Pain generation in lumbar and cervical facet joints. *J Bone Joint Surg 2006*. 2006;88-A(suppl 2):63–67.
13. Dishman JD, Dougherty PE, Burke JR. Evaluation of the effect of postural perturbation on motoneuronal activity following various methods of lumbar spinal manipulation. *Spine J*. 2005;5:650–659.
14. Cobb T. Web VP Classic v.4. http://www.lextutor.ca/vp/eng/. Accessed September 15, 2015.
15. Weeks WB, Wallace AE. Readability of British and American medical prose at the start of the 21st century. *BMJ*. 2002;325:1451–1452.