

Article

The passive voice in scientific writing. The current norm in science journals

Leong Ping Alvin

ABSTRACT: *In contrast to past consensus, many authors now feel that the passive voice compromises the quality of scientific writing. However, studies involving scientific articles are rare. Using a corpus of 60 scientific research articles from six journals, this study examined the proportion of passives used, and the contexts and forms in which they occurred. The results revealed that about 30% of all clauses were passive clauses. The canonical form was most pervasive, followed by the bare passive; together, they constituted more than a quarter of all clauses analyzed. Passives were typically used in main clauses, followed by relative and adverbial clauses. Roughly 29% of all passives were located in the methodology section. Based on the results, the proportion of passives in scientific writing may stabilize at about 30%. It is unlikely to dramatically drop any further since the trend suggests that passives are still widely used in the methodology section.*

KEYWORDS: *Scholarly publishing, science writing*

1 Context and objectives

A cursory glance at the numerous writing guides available, both on the Internet and in printed form, reveals a clear shift in consensus regarding the use of the passive voice in scientific writing. While the prevalent use of the passive voice was characteristic of scientific discourse in the major part of the 20th century [1, 2], the situation today is rather different. Many writing guides favor the use of the active voice for reasons of clarity and conciseness [3]. In their writing guide for scientific English, Day and Sakaduski [4] state this preference even more directly — “the passive voice should be avoided”.

Yet, as Ding [1] has shown in his work involving corpora from the 18th century to the end of the 19th century, early scientific discourse actually favored the active voice. The move towards the passive voice in the 20th century arose as a result of the increasing demands for scientific discourse to be objective, in the sense that the writing should ideally represent the world “in terms of objects, things, and materials” [1] rather than humans. As a consequence, scientific writing took on a decidedly object- or thing-oriented character [5, 6]. With the preference for the active voice in present-day scientific writing, we appear to have come full circle.

Indications from an informal survey among postgraduates at a Singaporean university, however, suggest that not many may be aware of this current preference. In fact, 90 of

the 99 students surveyed felt that the passive voice should be the norm in academic writing. This is unlikely to be a localized problem since the respondents were international students. Many of them held fast to the belief that the passive voice, among other strategies, was necessary to make the writing objective; one of them even noted that it helped to make the writing look ‘professional’. Few saw it as a problem that the overuse of the passive voice could make the writing unclear and wordy. As publications are expected of postgraduates in their doctoral programs and future careers, it is therefore important for such students (and, indeed, new faculty and researchers) to understand the norms concerning the use of the passive voice. This will go some way to help them craft texts that are clear and concise.

This in turn raises two questions for investigation. The first concerns the stability of present-day norms. Seoane [7], for instance, points to a “pronounced decline” in the use of the passive voice; the data from two time periods (1961 vs. 1991–92) show a fall of between 10.3% and 24.0% in the relative proportion of passive clauses [8]. Despite the decrease, the proportion of passive clauses in the 1991–92 corpora remains somewhat high, accounting for about half of all clauses (41.9–56.7%). Has this proportion fallen even further in more recent examples of scientific writing? The second concerns the specific details of these norms. While many writing guides encourage the use of the active voice, they are silent about the extent and distribution of the passive voice in current scientific writing. Specifically, in what contexts are passives typically used, and in what form?

To address these research questions, this paper analyzed 60 recently published articles from six journals. The articles covered various topics and disciplines, giving a broad view of the general trend concerning the use of the grammatical voice in the sciences. As stylistic conventions are not cast in stone, it is imperative for students and researchers alike to be aware of the current expectations and so make adjustments as necessary. This is of particular importance in the area of publishing, since non-compliance may result in time wasted in the re-writing of the manuscript. The *Journal of Trauma and Dissociation*, for instance, has the following piece of advice for authors on its website [9]:

Use the active voice whenever possible: we will ask authors that rely heavily on use of the passive voice to re-write manuscripts in the active voice.

2 Related studies

Investigations into the use of the passive voice in various types of writing have included rape descriptions [10], nurses’ notes [11], and undergraduate essays [12], among others. These have largely centered on the rhetorical function of passives, which primarily involves a shift in focus (from actor to affected), thus revealing the author’s stance [12].

Detailed studies involving scientific writing, however, are rarer. Early published accounts on the need for a more direct style of writing took the form of personal viewpoints [13–15], but given that the passive voice was pervasively used in scientific papers at the time, even such viewpoints were far and few between (but see also series of exchanges in the journal *Nature* [16–20]).

A somewhat more focused attempt at arguing against the overuse of the passive voice in academic writing came about two decades later in an opinion article by Kirkman [21], who pointed out how easily an insistence on the use of the passive could make the writing “ponderously roundabout, abstract and heavy to read”. He noted several potential shortcomings of passive-voice writing, including uncertainty about the agent (1), stilted language (2), and change of meaning (3):

- (1) It has been found by experiment that . . .
- (2) It is strongly considered that . . . (*instead of ‘I firmly believe/consider that. . .’*)
- (3) I do not accept the idea \neq The idea cannot be accepted

Other authors [22] have also commented on similar shortcomings.

Insofar as recent empirical studies are concerned, the one by Amdur, Kirwan and Morris [23] on medical papers parallels the focus of this present work. Their analysis of 90 papers (opinion, review, and original research papers) in three high-impact medical journals revealed a median frequency of passive-voice sentences of 20–26% of all sentences used. The range in percentages, however, was large; some articles had frequencies of 46–47%, i.e., almost one in two sentences was written in the passive voice. Apart from the typical reason of maintaining objectivity and detachment, the authors also suggested the pressure to publish as a further possible reason for the pervasive use of the passive voice [23]:

Several veteran editors attribute the problem to a change in the role academic publications play in modern society and the workload of journal editors: “It seems that many medical articles are written to be published and cited, but not to be read.”

More recently, Millar, Budgell and Fuller [24] investigated the use of the passive voice in 297 research articles, also in the field of medicine. Unlike the study by Amdur, Kirwan and Morris [23], the authors computed the mean (rather than the median) frequency of passives per sentence. They reported a figure of 43% for the research articles. In view of the median-mean difference, a comparison of the results from both studies is difficult. The work of Millar, Budgell and Fuller [24] also relied on an automatic process to retrieve the passive constructions by using the following string:

[any form of BE] [negation]^{optional} [up to 3 adverbs]^{optional} [past participle]

There are, however, two problems with this approach. The first is that the string restricted the analysis to only declaratives, since interrogatives have a slightly different grammatical structure. The second is that the string excluded truncated passives. Such passives take the form of either *bare* passives [25], which are headed by only past-participle verbs, and passives where the repeated ‘be’ auxiliary in a conjoined clause is omitted to avoid repetition:

- (4) . . . soil and gas samples *collected* again from the same site in 2007. (*bare passive*)

- (5) Statistical analysis was performed using Statistica version 10, and *deemed* statistically significant at $p < 0.05$. (*missing 'be' auxiliary in the conjoined clause*)

In consequence, the findings of Millar, Budgell and Fuller [24] may be under-reported.

In a slightly different vein, Seoane [8] compared the use of passives in scientific articles published in two time periods, 1961 and 1991–92. The articles, in both British English and American English, covered topics in the natural sciences and medicine. The comparison revealed a marked fall in the use of passives, with the fall being more pronounced in American English articles (24%) than in British English articles (10.3%). She argued that the decline reflected the prevailing pressure “to convey scientific knowledge as efficiently and clearly as possible” [8]. This suggests a clear change in the way scientific discourse is perceived by those who partake in it, resulting in a shift in emphasis towards making the writing clear and concise. It should be noted, though, that the relative proportion of passive clauses in the 1991–92 corpora was 41.9–56.7%, i.e., about one in two clauses was a passive clause. This is not a low figure.

Seoane’s [8] data, though, offer some empirical support against Gross, Harmon and Reidy [2], who are of the view that the passive voice in scientific writing in the last century reached “a point of evolutionary stability”. Although Gross, Harmon and Reidy [2] relied on various corpora spanning 1901 to 1995, they used a rather curious measure — the occurrence of passives per 100 words — to arrive at that conclusion. This does not appear to be an appropriate measure since the passive voice is a form of the verb phrase; the number of words in a passage reflects merely its length. If all the verb phrases under consideration are passives, the measure will be higher in the case of a short passage than a long one; such a difference, however, makes little sense when one considers the fact that 100% of the verb phrases used in both passages are passives.

This, nevertheless, raises an interesting question. While Seoane’s [8] data point to a decline in the use of the passive voice towards the end of the 20th century, they are based on articles published more than 20 years ago. It is unclear if this decline has continued in more recent scientific articles or if it has since stabilized. This study hopes to offer some answers in this regard.

I should hasten to add that while the available studies and commentaries do point out the rhetorical effects of the passive voice (e.g., depersonalization), such statements often appear as either brief remarks [2] or personal viewpoints [22]. It is far less clear whether scientists and researchers themselves are aware of these effects and whether they make careful decisions about the use of the grammatical voice in their writing. I am doubtful that many are able to recognize the different forms of the passive or even the contexts in which the passive, rather than the active, may be more appropriate. Unfortunately, I have not been able to locate any empirical evidence in support of this, but we do sometimes find indications in sentiments expressed in established sources. In 1996, in reaction against a piece written by the outgoing editor of *Nature*, Leather [19] wrote a letter to the same journal and argued passionately for the use of the passive voice in scientific writing. He saw the rhetorical effects of the passive voice as a virtue, and ended his letter with a rather strong statement — “The active voice encourages carelessness, partisanship and,

Table 1. Summary details of corpus.

	Journals		Articles	Clauses	
A	Nature	Nat	10	3,241	
	Science	Sci	10	3,311	
	Journal of Trauma and Dissociation	Trm	10	3,700	10,252
B	ACS Combinatorial Science	ACS	10	1,710	
	Global Change Biology	GCB	10	4,478	
	ChemPhysChem	CPC	10	3,907	10,095
	<i>Total</i>		60		20,347

as used by many of its adherents, does no favours to the English language or science”. This provoked a series of responses, many of them negative. The rejoinder that perhaps best captured the discrepancy in Leather’s writing came from Perlman [20] :

I consider Leather’s letter an outrageous display of scientific hypocrisy. He makes dogmatic pronouncements on a subject he knows nothing about . . . Leather should practise what he preaches: of the 18 transitive sentences in his letter, only four were in the passive.

This, naturally, points to the need for scientists and researchers to be made aware of the norms governing the use of the grammatical voice in current scientific writing. One cannot take it as a given that writers have this knowledge. Based on the findings of this study, a grounds-up approach is recommended in section 5 to, hopefully, rectify this issue.

3 Methods

3.1 Corpus and preparation

Sixty original research articles were selected from six well-known science journals (see Table 1). The articles were all published in 2013, and the 10 most-recent articles of the year were selected from each journal. At the time of data collection, these articles were the most recent additions to the journals. The articles, totaling 250,947 words, covered several disciplines and topics, ranging from chemistry to archaeology. Three of the journals (Group A journals) had specific statements on their websites encouraging authors to use the active voice whenever possible. The other three journals (Group B journals) made no such mention, although one of them, *ACS Combinatorial Science*, had a link to the American Chemical Society style guide [26] exemplifying different contexts in which the active and passive voices were deemed appropriate.

The research articles were first divided into clauses, defined as “units containing their own verbs” [27]. These included not merely clauses which could be easily segmented into separate grammatical units, but also embedded clauses that were part of larger clauses, as illustrated in (6–7), respectively (the number on the left of each clause is the reference number assigned to that clause in the article concerned):

- (6) 18 They can transfer charge and energy,
 19 and their absorption and photoluminescence wavelength can be tuned
 comparatively easily.
- (7) 42 No substantial batch effects in the data <...> were detected.
 43 < that might confound analyses >

Since each clause has, by definition, one and only one verb phrase, dividing up the articles into clauses this way allowed the number of passives in the corpus to be expressed as a proportion of the total number of verb phrases used.

Only the clauses in the main text of the articles were analyzed. Peripheral segments, such as titles, headings, captions, footnotes, and appendices were excluded.

3.2 *Passive forms*

The basic form of the passive is very well known, and its structure is well documented in grammar books [27, 28]. Most descriptions of the passive leave the impression that the ‘be’ auxiliary is always present in the verb phrase [27]. However, I agree with Puckica [25] that such a view of the passive is limiting. As he notes:

The ‘BE + past participle’ construction is understandably the main focus of presentations of the passive voice in Present-Day English (PDE). Yet, the label ‘passive voice’ cannot be reduced to that single construction. A passive construction may involve a verb other than BE combined with a past participle, as in the case of the GET-passive, but it may also not involve any kind of ‘helping’ verb.

Hence, in addition to the usual range of passives, the analysis also considered the presence of bare passives in the corpus. The full range of passives identified in the corpus is exemplified below in (8–15).

Basic (be + Ven)

- (8) The slab of PDMS with the molecular film *was dried* ...

Progressive (be + being + Ven)

- (9) The dissociative subtype of PTSD *is being* officially *adopted* by the upcoming DSM–5.

Perfective (have + been + Ven)

- (10) However, these predictions *have* not yet *been tested*.

Modal (modal + be + Ven)

- (11) These modulators *can be fabricated* simply, monolithically and at low cost.

Modal perfective (modal + have + been + Ven)

- (12) ... transient and delayed tree responses that *would have been missed* ...

To-infinitive (to + be + Ven)

- (13) ... for adjacent channels *to be placed* tens of micrometres apart ...

Non-finite -ing (being + Ven)

- (14) ... long, straight and stout, *being caudally directed* with no significant dorsal projection.

Bare (Ven)

- (15) ... with the lowest prices *accepted* by sellers in actually concluded trades.

Other passive forms commonly described in grammar books — such as the *get*-passive (e.g., ‘he got arrested’) and *have*-pseudo-passives (e.g., ‘they had him arrested’) — were not found in the corpus.

The clauses in which the passives occurred were classified as either main or subordinate clauses. For subordinate clauses, further distinctions were made as follows:

Relative clause

- (16) ... which is increasingly debated as a trigger for forest decline ...

Adverbial clause

- (17) When the coupling combination was changed to DCC/HOBt and DIPEA in DMF, ...

Reported clause

- (18) Subjects were informed + *that one choice situation would be randomly selected*...

Nominalized clause

- (19) A limitation of the present study is + *that the data were collected from a single source*...

3.3 Statistical analysis

IBM SPSS Statistics version 21 was used for all statistical analyses: the independent-sample *t*-test and analysis of variance (ANOVA). The Tukey honest significant difference post-hoc test was used for significant ANOVA results. The significance level for all tests was $\alpha = 0.05$.

4 Results and discussion

4.1 Comparisons among journals

Taking the corpus as a whole, passive clauses constituted 30.02% of the total number of clauses analyzed. This number is substantially lower than the 41.9–56.7% range reported by Seoane [8]. As might be expected, Group A articles contained fewer passive clauses

than Group B articles — 26.02% vs. 34.09% ($p < 0.0002$). A comparison among the individual journals, though, revealed mixed results. Figure 1 illustrates the mean proportions of passive clauses in the six journals.

A one-way ANOVA test revealed that at least one of the means was different from the rest ($p = 0.0002$). The post-hoc test showed significant differences in four comparisons: Nat vs. ACS ($p = 0.0005$), Sci vs. ACS ($p = 0.0372$), Trm vs. ACS ($p = 0.0094$), and Nat vs. CPC ($p = 0.0048$). The observed difference between Trm and CPC was only weakly insignificant ($p = 0.0785$).

On the face of it, the numbers appear to indicate equality in the means of four of the journals, with the mean of GCB (a Group B journal) being less than half a percentage point above the corpus mean. The higher means in the case of ACS and CPC could in part be due to the absence of any clear indication from the respective journals about the use of the passive voice. In the case of GCB, the articles had a wide range of proportions, from 20.52% to 45.19%, with the proportions of six of the articles below the corpus mean. This variation could also have been the result of a similar lack of indication about the use of the grammatical voice.

The same explanation, unfortunately, cannot be applied to another journal, Sci, which also displayed a similarly wide variation in proportions. The variation in Sci ranged from 14.36% to 41.31%, with the proportions of five of the articles exceeding the corpus mean. Unlike GCB, however, Sci is a Group A journal; the other two Group A journals (Nat and Trm) had no such wide variation. This makes the anomaly rather curious, given the following advice on Sci’s website [29]:

Use active voice when suitable, particularly when necessary for correct syntax (e.g., “To address this possibility, we constructed a λ Zap library . . .,” not “To address this possibility, a λ Zap library was constructed . . .”).

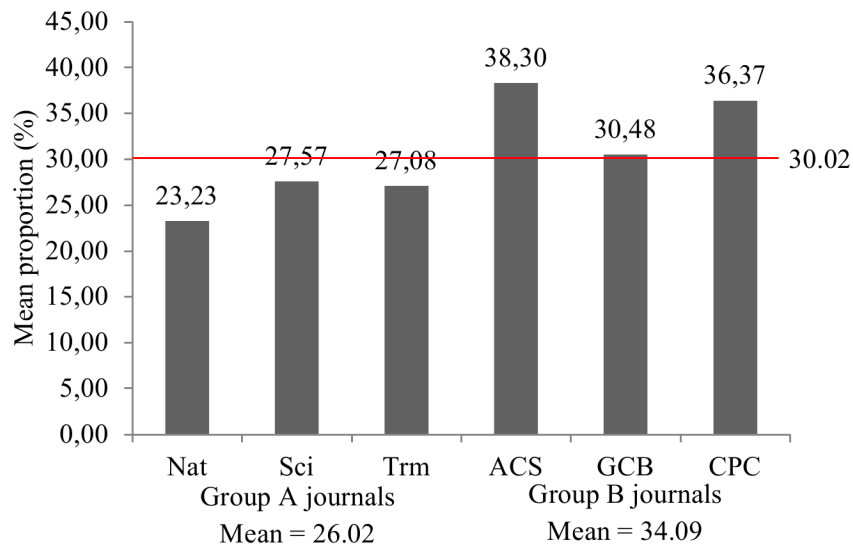


Figure 1. Mean proportions of passive clauses. The red horizontal line represents the corpus mean.

The findings, then, suggest a layered effect. At the macro level, as seen in the broad comparison between Group A and Group B articles, the results indicate that the presence of explicit statements on the use of grammatical voice does have an impact on the writing. At the micro, subtler level, 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. To this end, I agree with Millar, Budgell and Fuller [24] that the anomaly could have been due to the vague nature of the advice given to authors. It is unclear, for instance, what “when suitable” means. Further, apart from “correct syntax”, are there other contexts, grammatical or otherwise, that may be suitable? We shall return to this issue in section 4.3.

4.2 *Passive forms*

We turn next to the passive forms used in the corpus (Table 2). The two dominant forms were basic and bare passives: the former occurred in 16.34% of all clauses analyzed, and the latter, 9.71%. There were more basic passives in Group B journals than Group A journals (19.72% vs. 12.98%, $p < 0.0001$); differences in the use of bare passives across all the journals were statistically insignificant. Observed differences in the use of the other passive forms were also statistically insignificant. Progressive and modal perfective passives were rare; they occurred only five times each in the entire corpus.

The distribution of the passive forms is interesting as it is consistent across all the journals. We do not see, for instance, a reversal of the basic-bare ordering in the journals, or the prevalence of any other passive form in a particular journal. This uniformity is perhaps reflective of the need for scientific writing to be as precise as possible. This may help to explain the preference for the basic passive over other passive forms in the corpus since the addition of other auxiliary verbs is either impossible (20) or inappropriate in context (21):

- (20) Such a fabricated multichannel device *is shown* in Figure 2d.
 ~ (?) Such a fabricated multichannel device *may be shown* in Figure 2d.
- (21) When the phase-matching condition *is met*, ... (*expressing a timeless condition*)
 ~ (?) When the phase-matching condition *has been met*, ...

The pervasiveness of the bare passive in the corpus also deserves mention, given how it is often bypassed in studies [7, 24]. The data in Table 2 show that almost one in ten clauses in the corpus contained a bare passive. As the bare passive is already in its reduced form, modifying it to a non-passive version is very likely to inflate the word count. Further, some of the bare-passive clauses frequently found in scientific writing are fossilized:

- (22) *As observed* in ...
- (23) ... also *known* as ...
- (24) ... as *compared* to ...

Table 2. Distribution of passive forms (percentages in terms of the clauses in each journal). *Note:* Prog = progressive; Perf = perfective; Mod = modal; Mod pf = modal perfective; To-inf = to-infinitive; NF *-ing* = non-finite *-ing* (see 8–15); %Cls = percentages expressed in terms of all clauses in the corpus.

Forms	Nat	Sci	Trm	ACS	GCB	CPC	%Cls
Basic	10.36	14.14	14.27	25.26	17.71	19.64	16.34
Prog	0.03	0.03	0.08				0.02
Perf	1.30	1.18	1.44	2.51	0.89	1.79	1.41
Mod	1.95	1.86	1.84	2.34	1.14	3.35	2.04
Mod pf	0.03	0.06			0.05		0.02
To-inf	0.46	0.33	0.67	0.17	0.17	0.33	0.37
NF <i>-ing</i>	0.12	0.03	0.24	0.18	0.09	0.05	0.11
Bare	8.98	9.94	8.54	7.84	10.43	11.21	9.71
<i>Total</i>	23.23	27.57	27.08	38.30	30.48	36.37	30.02

While a change in (22–23) is possible (e.g., ‘As we observe(d) in’, ‘which we also know as’), this is somewhat difficult to accomplish in (24); the alternative — ‘as we compare(d) X with Y’ — loses the elegance of the original version. Apart from unnecessary wordiness, attempting to convert some of the bare-passive clauses into non-passive versions may also at times result in some awkwardness:

- (25) ... a strict monolayer of carbon atoms tightly *packed* into a hexagonal honeycomb lattice, and fullerene, ...
 ~ ... a strict monolayer of carbon atoms that/which AGENT tightly packed into a hexagonal honeycomb lattice, and fullerene, ...

In the case of (25), what would the writer indicate as the agent?

Indeed, as many authors have pointed out, it is not the mere use of the passive that is at issue, but its *overuse*. We cannot completely avoid using the passive voice, and in some instances, using the active voice “may detract from the clarity and/or go against the established phraseological conventions” [24]. Day and Gastel [30] put it even more aptly (*emphasis mine*):

In any type of writing, the active voice is usually more precise and less wordy than is the passive voice. (*This is not always true; if it were, we would have an Eleventh Commandment: “The passive voice should never be used.”*)

4.3 Distribution of passives

The distribution of passives within main and subordinate clauses was fairly even (Table 3). Among the journals, only the observed difference in ACS was significant ($p = 0.0001$), but this is unlikely to be a characteristic of scientific writing as a whole. The difference in proportions at the corpus level (15.48% vs. 14.54%) was statistically insignificant.

Table 3. Distribution of passives in main and subordinate clauses (percentages in terms of the clauses in each journal). *Note:* Sub = subordinate clauses; %Cls = percentages in terms of all clauses in the corpus.

Clauses	Nat	Sci	Trm	ACS	GCB	CPC	%Cls
Main	9.35	11.93	14.00	25.49	16.12	19.84	15.48
Sub	13.88	15.64	13.08	12.81	14.36	16.53	14.54
<i>Total</i>	23.23	27.57	27.08	38.30	30.48	36.37	30.02

Table 4. Distribution of passives in subordinate clauses (percentages in terms of the subordinate clauses in each journal). *Note:* Rel cl = relative clauses; Adv = adverbial clauses; Rep = reported clauses; Nom = nominalized clauses; %DepCls = percentages in terms of all subordinate clauses in the corpus.

Clauses	Nat	Sci	Trm	ACS	GCB	CPC	%DepCls
Rel cl	62.44	68.92	61.98	61.64	63.45	63.93	63.99
Adv	27.33	20.85	28.93	29.68	30.48	28.64	27.60
Rep	7.78	6.76	6.82	7.76	4.67	6.19	6.42
Nom	2.45	3.47	2.27	0.92	1.40	1.24	1.99
<i>Total</i>	100.00	100.00	100.00	100.00	100.00	100.00	100.00

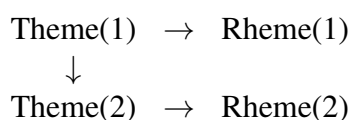
The majority of the subordinate passive clauses were relative clauses, constituting almost two-thirds of all subordinate passive clauses (Table 4). This is followed by adverbial clauses. Together, relative and adverbial clauses occurred in more than 90% of the subordinate passive clauses ($p = 0.00002$).

Taken together, the results in Tables 3–4 suggest that authors are as likely to use passives in main clauses as in subordinate clauses. The function of the passive voice — that it draws attention to the affected phenomenon rather than the agent — is well known and documented in numerous grammar books [27]. What the data show is that this is likely to be realized not merely in primary clauses in which “the situation expressed ... is foregrounded” [31], but also in secondary clauses which perform a supporting role. While a number of the passives in relative and adverbial clauses take the form of fossilized expressions (e.g., ‘as shown in’; see section 4.2), the others provide descriptive and contextual information.

- (26) These experiments however usually measure longer distances by small rate increases which *are* usually *interpreted* in a semiquantitative way yielding relative distance information.
- (27) In the United Kingdom, the obvious mechanism through which that goal could be achieved is reform of the European Union’s (EU’s) Common Agricultural Policy (CAP).
- (28) In fact, the equatorial-axial conformational energy difference in scopine (MP2: 13 kJmol⁻¹) is larger than in tropinone (~2 kJmol⁻¹) although this value *would be amplified* by a weak C–H···O interaction in equatorial scopine.

The passives in (26–28) draw attention to the affected phenomenon. While it is fairly straightforward to amend the passive relative clauses in (26–27) to the active voice — i.e., ‘which researchers usually interpret’, ‘through which the government could achieve’ — this is not true of passive subordinate clauses in general. At times, the theme-rheme structure of the sentence may be affected, which may in turn result in a coherence problem. According to Halliday and Matthiessen [32], the theme of a clause is the point of departure of the message in the clause; it is the first element that functions as a participant, adverbial, or process. The remainder of the clause is known as the rheme. In the case of (28), the two clauses have the following theme-rheme structure:

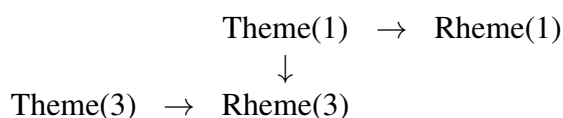
- (28) a. [In fact, the equatorial-axial conformational energy difference in scopine (MP2: 13 kJmol⁻¹)]^{Theme(1)} [is larger than in tropinone (~2 kJmol⁻¹)]^{Rheme(1)}
 [although this value]^{Theme(2)} [would be amplified by a weak C–H···O interaction in equatorial scopine.]^{Rheme(2)}



As (28a) shows, the second clause elaborates on the theme of the first clause (‘this value’ = ‘energy difference’). The Prague School linguists refer to such a pattern as constant thematic progression [33], which provides multiple developments on a single point of departure.

Rewriting the subordinate clause in (28) in the active voice would change the thematic pattern dramatically, as in (28b). In the amended version, the theme of the first clause is linked to the rheme of the second, resulting in what Enkvist [34] refers to as ‘rheme regression’. This leaves the theme of the second clause, its point of departure, unattached, thus creating a small coherence problem within the sentence.

- (28) b. ...[although a weak C–H···O interaction in equatorial scopine]^{Theme(3)}
 [would amplify this value.]^{Rheme(3)}



The coherence problem may seem minor in the context of a single sentence, but problems are likely to surface with modifications of all similar sentences at the text level. This returns us to a point alluded to earlier in section 4.2 that enforcing the use of the active voice may not always lead to better writing. There are situations where the passive voice may in fact be more appropriate.

There is at least one other area in which the use of the passive is commonplace and perhaps justifiable. Many researchers have noted the prevalent use of the passive voice in

Table 5. Distribution of passives in the methodology sections of 48 articles (%). Percentages are expressed in terms of the clauses in each methodology section.

Nat	Sci	Trm	ACS	GCB	CPC
100.00	49.02	36.56	76.32	40.16	58.33
100.00	69.70	32.14	77.78	34.29	77.27
100.00	57.32	53.23	91.30	42.70	42.31
20.69	50.00	39.13	47.89	41.18	54.17
69.23	23.23	25.17	76.92	64.96	27.45
71.43		42.06	68.00	53.04	66.67
		45.56	80.00	38.37	72.34
		33.96		56.64	61.11
		34.78		40.34	50.00
		53.03		43.43	63.64

the methodology section of articles [23, 24]. The passive is preferred since descriptions in this section lay emphasis on the procedures, instruments, and materials used rather than the agents (e.g., ‘LC-MS *were performed* on an Agilent 2100 system’).

In this present study, a systematic comparison of the methodology sections of the articles was not possible as 12 of the articles did not have such a section. Notwithstanding this, 29.02% of all passive clauses were located in the methodology sections of the remaining articles. Table 5 summarizes the extent of the use of passives in these sections.

Although the variation in proportions appears wide (23.23–100.00%), the majority of the articles displayed a prevalent use of the passive voice. A total of 27 articles had proportions in excess of 50%, meaning more than half of the clauses in the respective methodology section were passive clauses. In three of these articles, the entire methodology section was written in the passive. A further ten articles had proportions in the 40% range.

5 Conclusions

Many books and guides on scientific writing currently advice authors to use the active voice whenever possible. The main argument is that the active voice helps to make the writing clear and concise. This represents a change from the situation in the major part of the 20th century, when the passive voice was characteristic of scientific writing in general. This paper investigated the occurrence of passives in 60 research articles from six well-known science journals to better understand the new norm in light of the current preference concerning the use of the active voice. This paper sought answers to two research questions:

- (a) How does the proportion of passive clauses identified in the present study compare with the results from other similar studies?
- (b) What is the distribution of passives in modern scientific articles, and in what forms do they typically occur?

At the macro level, 30.02% of all clauses in the corpus were passive clauses. Journals that encouraged authors to use the active voice had a lower proportion of passives than journals that did not (26.02% vs. 34.09%). The corpus mean of 30.02% is much lower than the 41.9–56.7% range reported by Seoane [8] or the 43% figure reported by Millar, Budgell and Fuller [24]. Although the work by Millar, Budgell and Fuller [24] is very recent, it analyzed only articles from medical journals. This present study, by contrast, used articles covering a range of scientific topics. As the journals included in this study (e.g., *Science*, *Nature*) are well known and widely read by serious researchers, the present norm is likely to influence the choice of grammatical voice in research articles in general over time. Based on the findings of this study, I expect the proportion of passive clauses in research articles to hover around 30% in the foreseeable future. Any marked decrease is unlikely since passives are still preferred in the methodology section, and bare passives cannot be easily changed.

At the micro level, the findings revealed that basic and bare passives were the most commonly used passive forms; together, they constituted more than a quarter of all clauses in the corpus. There was a roughly even distribution of passives in main and subordinate clauses. The majority of passive subordinate clauses were relative clauses (63.99%) and adverbial clauses (27.60%). The methodology sections of the articles had the greatest concentration of passive clauses; they contained 29.02% of all passive clauses.

I have also provided examples of bare passives and subordinate passive clauses to illustrate that the use of the active voice may not always be appropriate. This is a crucial point that may not have been expressed clearly (and strongly) enough in current journal style guides. In the absence of any clear directions, it can be confusing for novice writers who may have other rhetorical demands of writing to juggle with. The current emphasis on clear and concise writing requires the *appropriate* use of the grammatical voice, which may at times favor the passive voice. It is the overuse of the passive voice that we should avoid, not its complete avoidance.

What, then, can be done to heighten awareness regarding the appropriate use of the grammatical voice? The most direct course of action is for journals to include a style manual on their websites, indicating clearly not just the contexts in which the passive voice is entirely permissible (e.g., the methodology section) but why it may be more appropriate than the active voice (e.g., it lays focus on the affected). A more enduring course of action, though, is a grounds-up approach through education. It stands to reason that students who are guided early on in their university studies about the importance of clarity and directness in scientific writing will choose the appropriate grammatical voice (among other strategies) to achieve this. Writing courses offered in universities worldwide, whether at the undergraduate or postgraduate level, thus have an important role to play. It is easy to neglect the smaller things; in fact, indications from the informal survey mentioned at the start of this paper suggest that issues on the appropriate use of the grammatical voice were never addressed in the international students' previous writing courses. More certainly needs to be done in this respect.

References

- [1] D. Ding (1998), "Rationality reborn: Historical roots of the passive voice in scientific discourse", in J.T. Battalio ed., *Essays in the Study of Scientific Discourse: Methods, Practice, and Pedagogy*, Ablex, Stamford, CT, pp. 117–135.
- [2] A.G. Gross, J.E. Harmon and M. Reidy (2002), *Communicating Science: The Scientific Article from the 17th Century to the Present*, Oxford University Press, Oxford, U.K.
- [3] J.D. Humphrey and J.W. Holmes (2009), *Style and Ethics of Communication in Science and Engineering*, Morgan & Claypool, San Rafael, CA, U.S.A.
- [4] R.A. Day and N. Sakaduski (2011), *Scientific English: A Guide for Scientists and Other Professionals*, 3 edition, Greenwood, Santa Barbara, CA, U.S.A.
- [5] J.L. Kinneavy (1980), *A Theory of Discourse*, W.W. Norton & Company, New York, U.S.A.
- [6] A.M. Wilkinson (1992), "Jargon and the passive voice: Prescriptions and proscriptions for scientific writing", *J. Tech. Writ. Commun.* **22**: 319–325.
- [7] E. Seoane (2013), "On the conventionalisation and loss of pragmatic function of the passive in Late Modern English scientific discourse", *J. Hist. Pragmat.* **14**(1): 70–99.
- [8] E. Seoane (2006), "Changing styles: On the recent evolution of scientific British and American English", in C. Dalton-Puffer et al. eds., *Syntax, Style and Grammatical Norms: English from 1500–2000*, Peter Lang, Bern, Switzerland.
- [9] *Instructions for authors, Journal of Trauma and Dissociation*, http://www.tandfonline.com/action/authorSubmission?journalCode=wjtd20&page=instructions#.UqA_KuL85ac, retrieved 5 December 2013.
- [10] G. Bohner (2001), "Writing about rape: Use of the passive voice and other distancing text features as an expression of perceived responsibility of the victim", *Brit. J. Soc. Psychol.* **40**: 515–529.
- [11] K.L. Porter (2005), *The Frequency and Function of Passive Voice Use in Nurses' Notes*, Thesis, Department of English, University of North Carolina Wilmington, U.S.A.
- [12] A.M. Baratta (2009), "Revealing stance through passive voice", *J. Pragmatics* **41**(7): 1406–1421.
- [13] P.W. Bridgman (1950), *Reflections of a Physicist*, Philosophical Library, Inc, New York, U.S.A.
- [14] R.V. Ormes (1957), "Passive voice", *Science*, **125**(3247): 529.
- [15] J.T. Robinson (1957), "Passive voice and personal pronouns", *Science*, **125**(3258): 1160.
- [16] T.I. Baskin (1996), "Hazards of the passive voice" (correspondence), *Nature*, **381**(6585): 730.
- [17] N.W. Goodman (1996), "Hazards of the passive voice" (correspondence), *Nature*, **381**(6585): 730.
- [18] C. Jolly (1996), "Hazards of the passive voice" (correspondence), *Nature*, **381**(6585): 730.
- [19] S.R. Leather (1996), "The case for the passive voice", *Nature* **381**(6582): 467.
- [20] A.M. Periman (1996), "Passive voice", *Nature*, **382**(6587): 108.
- [21] J. Kirkman (1975), "That pernicious passive voice", *Phys. Technol.* **6**: 197–200.
- [22] T. Sigel (2009), "How passive voice weakens your scholarly argument", *Journal of Management Development* **28**(5): 478–480.
- [23] R.J. Amdur, J. Kirwan and C.G. Morris (2010), "Use of the passive voice in medical journal articles", *American Medical Writers Association Journal* **25**(3): 98–110.
- [24] N. Millar, B. Budgell and K. Fuller (2013), "'Use the active voice whenever possible': The impact of style guidelines in medical journals", *Applied Linguistics* **34**(4): 393–414.
- [25] J. Puckica (2009), "Passive constructions in present-day English", *Groninger Arbeiten zur Germanistischen Linguistik* **49**: 215–235.

- [26] A.M. Coghill and L.R. Garson (2006), “Writing Style and Word Usage”, in A.M. Coghill and L.R. Garson eds., *The ACS Style Guide: Effective Communication of Scientific Information*, Oxford University Press, New York, U.S.A., pp. 41–58.
- [27] R. Carter and M. McCarthy (2006), *Cambridge Grammar of English*, Cambridge University Press, Cambridge, U.K.
- [28] S. Greenbaum et al. (1985), *A Comprehensive Grammar of the English Language*, Longman, Harlow, U.K.
- [29] *Some notes on Science style*, <http://www.sciencemag.org/site/feature/contribinfo/prep/res/style.xhtml>, retrieved 13 December 2013.
- [30] R.A. Day and B. Gastel (2012), *How to Write and Publish a Scientific Paper*, 7 edition, Cambridge University Press, Cambridge, U.K.
- [31] M.J. Endley (2010), *Linguistic Perspectives on English Grammar: A Guide for EFL Teachers*, Information Age Publishing, Charlotte, NC, U.S.A.
- [32] M.A.K. Halliday and C.M.I.M. Matthiessen (2004), *An Introduction to Functional Grammar*, 3 edition, Arnold, London, U.K.
- [33] F. Daneš (1974), “Functional sentence perspective and the organization of the text”, in F. Daneš ed., *Papers on Functional Sentence Perspective*, Mouton, The Hague, Netherlands, pp. 106–128.
- [34] N.E. Enkvist (1973), “Theme dynamics and style: An experiment”, *Studia Anglica Posnaniensia* 5: 127–135.

Author

Leong Ping Alvin is a senior lecturer at the Language and Communication Centre, Nanyang Technological University, Singapore. He coordinates the postgraduate writing course for international doctoral students. His research interests are in systemic functional grammar and literacy studies. E-mail: alvin.leong@ntu.edu.sg

HOW TO CITE: Leong P.A., *The passive voice in scientific writing. The current norm in science journals*, *JCOM* 01(2014)A03.