编校英文科技论文应注意正确“比较”

刘 园

浙江大学出版社期刊分社《信息与电子工程前沿》(英文)编辑部,310028,杭州

摘 要 英文科技论文以及中文科技论文英文摘要中如何正确地“比较”是一个常见的问题。一些编辑常因不能正确编校比较语句而导致出现问题。这些比较问题可以归为3类:比较对象不一致,比较短语表述不恰当,以及比较数据不正确。本文详细分析了最常见的2种比较数据差错,即把百分点当作百分比使用以及倍数比较错误。本文支持times more than不等同于times as many as的观点,并提供times + reduction/less表示更小更少的倍数关系的案例。通过分析11个典型案例,提供相应中文,给出具体修改意见,旨在给编辑同行提供比较语句的编校参考以及帮助学者提高英文写作水平。

关键词 英文编校;科技论文;比较

Correct“comparison” when editing English scientific papers//LIU Yuan

Abstract How to correctly compare sentences in English scientific papers and English abstracts of Chinese scientific papers is a common problem. Some editors find this challenging. Issues related to comparison can be classified into three categories: inconsistent comparison targets, inappropriate phrase structures, and incorrect comparison data. This paper provides a detailed analysis of two most common comparison data errors, i.e., using percent point as percentage and using incorrect comparison of multiple, supporting the view that “times more than” is not equal to “times as many as”, and provides a case study of times + reduction/less representing a smaller or less multiple relationship. This paper analyzes 11 typical casesin detail, provides corresponding Chinese meanings of these cases, and makes specific suggestions for revision. This paper provides some guidance for fellow journal editors in dealing with comparison sentences, and may help researchers improve their English writing.

Keywords English copyediting; scientific papers; comparison

Author's address Editorial Office of Frontiers of Information Technology & Electronic Engineering, Journal Publishing Center, Zhejiang University Press, 310028, Hangzhou, China

DOI:10.16811/j.cnki.1001-4314.2024.01.010

笔者在学术论文编校工作中发现,很多作者往往不能在英文科技论文以及中文科技论文的英文摘要中做到正确“比较”,经过英文外专润色过的文章依然存在不少这方面的错误。目前国内还没有编辑同人系统阐述比较语句中存在的这类问题。很多科技期刊编辑遇到此类问题时难以找到编校参考,以至于忽视这类问题,造成编校差错。本文案例均来自如《信息与电子工程前沿》(英文)等计算机、信息类的英文科技期刊,并提供所有案例的中文表述,旨在给编辑同行提供编校参考以及帮助论文作者提高英文写作水平。

1 比较中存在的问题

英文科技论文以及中文科技论文英文摘要中常见直观的比较词主要包括than、compared to/with、similar to、same as等,不那么直观的比较词主要包括exceed、equivalent to、like、superior to等。根据科技论文编校经验,本文将比较语句中存在的问题归为3类:1)比较词前后比较对象不一致:当句子结构较为复杂时,容易忽略比较词前后比较对象的一致性。2)比较短语表述不恰当:如用不恰当的副词修饰比较级,短语表述缺少介词,以及than和compared to/with混用等。3)比较的数据不正确:很多作者没有搞清楚百分比和百分点概念的区别,错误地把百分点作为百分比直接进行比较,正文数据和图表数据不吻合,以及比较的数据计算结果错误。另一类常见的数据比较错误是倍数的表达,很多作者错误地将times more than等同于times as many as。在中文论文中不用倍数表达更少更小的关系,但是在一些专业领域的英文论文中会使用更小更少的倍数关系表示性能的提升。针对这种情况,本文提供一个详细的案例分析,并给出其恰当的中文表达。

2 典型案例分析

2.1 比较对象不一致

比较对象不一致主要体现在比较词前后比较对象不统一。编辑和学者在处理科技论文比较语句时,尤其是当句子结构较为复杂时,往往容易省略比较词后的比较对象,直接把某一性能结果与方法进行比较。建议根据比较对象名词的性质,为保证句子简洁,在比较词后采用代词指代比较词前的比较对象;当比较对象为单数或者不可数名词,在比较词后用that指代比较词前的比较对象,当比较对象为可数名词复数形式,在比较词后用those指代比较词前的比较对象,确保比较对象的一致性。

示例1 The results indicate that the prediction accuracy of EEMDN-SABiGRU is superior to the comparable models. (结果表明,EEMDN-SABiGRU模型的预测精度优于同类模型。)

该句用的比较词是 superior to,作者直接将 EEMDN-SABiGRU 模型的预测精度与同类模型进行比较,导致比较词前后的比较对象不一致。在句中比较词 superior to 后需要用 those 指代同类模型的预测精度,建议改为:

The results indicate that the prediction accuracy of EEMDN-SABiGRU is superior to those of the comparable models.

示例 2　For the test images, the average PSNR was 61.50 dB after hiding 10 000 bits, which exceeds the recently reported pairwise PEE, pairwise IPVO, LPVO, dual pairwise PEE, and CNN PEO methods by 3.19, 1.92, 1.22, 1.39, and 0.73 dB, respectively. (对于测试图像,隐藏 10 000 bit 后的平均峰值信噪比为 61.50 dB,比最近报道的成对 PEE、成对 IPVO、LPVO、双成对 PEE 和 CNN PEO 方法分别高 3.19、1.92、1.22、1.39 和 0.73 dB。)

与示例 1 简单句相比,示例 2 用了“, which”引导的非限定性定语从句,句子结构较为复杂。示例 2 将平均峰值信噪比与所提 5 种方法直接比较,导致比较词 exceed 前后比较对象不一致。建议用 those 指代 5 种方法的平均峰值信噪比,将原句改为:

For the test images, the average PSNR was 61.50 dB after hiding 10 000 bits, which exceeds those of the recently reported pairwise PEE, pairwise IPVO, LPVO, dual pairwise PEE, and CNN PEO methods by 3.19, 1.92, 1.22, 1.39, and 0.73 dB, respectively.

示例 3　Although methods using a discretizing mechanism like Chu et al. (2018a) converge more easily during training, their output angles are limited to the preset classification set. [尽管使用 Chu 等人(2018a)所使用的离散机制的方法在训练过程中更容易收敛,但它们的输出角度仅限于预设的分类集。]

该句比较词 like 后直接跟参考文献,比较对象显然不一致。建议在比较词 like 后用 that 指代机制 mechanism 把表述补充完整,以增加句子可读性:

Although methods using a discretizing mechanism like that used in Chu et al. (2018a) converge more easily during training, their output angles are limited to the preset classification set.

2.2　短语表述不恰当

短语表述不恰当一般表现为用不恰当的副词修饰比较级,短语缺少介词,代词指代不明确,以及 than 和 compared to/with 混用等形式。

在科技论文中,经常用副词修饰形容词比较级的比较程度。一般常见的可用来修饰比较级的副词有 a little、a bit、slightly(表示“稍微”“一点”),以及 much、far、a lot、a great deal、rather 等(表示“……得多”)。具体用什么副词修饰比较级取决于比较词后面是可数名词还是不可数名词。

示例 4　Morrison et al. (2018) proposed a pixel-level grasp detection network with many fewer network parameters and less computation, making it fully meet the real-time requirement of robot grasping. (Morrison 等人(2018)提出一种像素级抓取检测网络,具有更少网络参数以及更少计算量,完全满足机器人抓取的实时性要求。)

Many 含义为“许多”,当比较级 more 后连接复数名词时,可用 many 修饰,表述为 many more,含义为“更多”。但 many 不可修饰比较级 fewer,many fewer 为错误表述。可用 a lot/much fewer 等表述“更少”。该句建议改为:

Morrison et al. (2018) proposed a pixel-level grasp detection network with much fewer network parameters and less computation, making it fully meet the real-time requirement of robot grasping.

示例 5　As for search time, the BO algorithm needs at least 20 min to search CNNs, GPPSO takes slightly more than BO, about 40 min, and PSO takes the most time at an average of 100 min on CIFAR-10 and 280 min on CIFAR-100. (在搜索时间上,BO 算法搜索 CNNs 至少需要 20 min,GPPSO 算法的搜索时间稍微地多于 BO 算法,约为 40 min,而 PSO 算法的搜索时间最多,在 CIFAR-10 数据集上平均为 100 min,在 CIFAR-100 数据集上平均为 280 min。)该句中的数据如图 1 所示。

单从表述上看,副词 slightly 可修饰比较级 more than,表示“稍微”“略”的含义。然而通过图 1 中的表 3 可知,BO 所需最少时间是 17 min,GPPSO 所需平均时间(45 + 39) min/2 = 42 min,42 是 17 的 2.4 倍,GPPSO 所花时间远多于 BO 所花时间,用 slightly(稍微地)修饰明显不妥,建议用 much/rather/a great deal more than 表示“……多得多”。基于 CIFAR-10 和 CIFAR-100 数据集,PSO 所花平均时间分别为(50 + 150)/2 min = 100 min 和(192 + 212) min/2 = 202 min,而非句中 100 min 和 280 min。该句建议改为:

As for the search time, the BO algorithm needs at least 17 min to search CNNs, GPPSO takes much more than BO, 42 min on average, and PSO takes the most time at an average of 100 min on CIFAR-10 and 202 min on CIFAR-100.

Table 3 Comparisons with the basic algorithms of GPPSO on CIFAR-10 and CIFAR-100 datasets

Method	CIFAR-10					CIFAR-100				
	Test accuracy (%)	Training accuracy (%)	Number of parameters ($\times 10^6$)	Search time (min)	Training time (min)	Test accuracy (%)	Training accuracy (%)	Number of parameters ($\times 10^6$)	Search time (min)	Training time (min)
Manual (ResNet20)	92.25	98.71	0.38		116	68.96	90.78	0.40		111
BO	92.91	99.92	4.70	20	274	66.47	90.62	4.70	20	279
BO_ac	92.26	99.96	5.26	17	367	65.64	75.60	3.57	23	326
PSO	84.17	86.09	4.65	50	258	54.25	28.20	4.70	192	265
PSO_ac	74.57	74.75	6.17	150	361	51.77	54.77	4.74	212	296
GPPSO	91.85	99.95	4.77	45	270	65.91	94.00	4.74	33	330
GPPSO_ac	95.26	99.96	5.26	39	261	76.36	97.65	4.44	39	304

图1 表3的复制品

示例6 From Table 7, the time complexity of the proposed EEMDN-SABiGRU model is the same level as other comparable models.(从表7可以看出,本文提出的EEMDN-SABiGRU模型的时间复杂度与其他类似模型处于同一水平。)

"处于同一水平"应表述为"at the same level","is the same level"的含义为"是同一水平",短语表述中in/at等介词处于重要地位,不可缺少。且句中比较对象不一致,应用those指代所比较方法的时间复杂度。该句建议改为:

From Table 7, the time complexity of the proposed EEMDN-SABiGRU model is at the same level as those of the other comparable models.

示例7 The simulations indicated only how the size of the quantum from 1×10^3 to 1×10^7 affected PPBP. When the quantum size was too small, the number of CPU or GPU requests processed in batches became insufficient. It was equivalent to FRFCFS when the quantum size was infinitely small.(模拟只显示了从1×10^3~1×10^7的量子尺寸如何影响PPBP。当量子尺寸过小时,批量处理的CPU或GPU请求数量不足。当量子尺寸为无限小时,它与FRFCFS方法等效。)

句子"It was equivalent to FRFCFS"前一句主语是量子尺寸,显然句子中"It"指代的不应该是量子尺寸,而应该是一种方法,才能确保比较对象的一致性。咨询作者得知,此处It指代是前文相隔较远的PPBP方法。遇到代词离指代对象较远时,建议不用代词,宜直接给出比较对象,增加句子可读性。该句建议改为:

The simulations indicated only how the size of the quantum from 1×10^3 to 1×10^7 affected PPBP. When the quantum size was too small, the number of CPU or GPU requests processed in batches became insufficient. The PPBP was equivalent to the FRFCFS when the quantum size was infinitely small.

示例8 The network congestion starts later than with other algorithms, and the bandwidth utilization is higher than with other schemes.(网络拥塞开始时间比其他算法晚,带宽利用率比其他算法高。)

句中"than with"表述错误,猜想是作者将than和compared to/with混用,且句中第2处比较的对象不一致,应用that指代所比较的带宽利用率。提供2种修改方案:

The network congestion starts later than the other algorithms, and the bandwidth utilization is higher than that of the other schemes.

或The network congestion starts later compared with the other algorithms, and the bandwidth utilization is higher compared with that of the other schemes.

2.3 数据比较差错

2.3.1 百分比和百分点错误使用

这类数据差错主要体现在2方面:1)百分比和百分点是数据比较中经常出现的2个概念,很多编辑和作者没有完全搞清楚百分比和百分点的区别,错误地把百分点作为百分比直接进行数据比较。2)正文数据和图表数据不吻合以及比较的数据计算结果错误。这2种错误经常一起出现,现通过以下2个案例说明此类差错。

示例9 As shown in Table 2, for classification accuracy, GPPSO ranks the fifth among the 10 algorithms compared on the CIFAR-10 dataset, with up to 11.09% higher accuracy than PSO and 1.52% lower accuracy than the first algorithm (CNN-GA). As for CIFAR-100, GPPSO achieves good performance that ranks fifth among the eight algorithms, 22.21% better than the last and only 3.11% lower than the first.(如表2所示,在CIFAR-10数据集上,GPPSO的分类准确率在10种算法中排名第5,比PSO的准确率高出11.09%,比第一种算法CNN-GA的准确率低1.52%。在CIFAR-100数据集

Table 2 Comparisons with the manually designed CNNs, non-OA-based methods, and OA-based methods on CIFAR-10 and CIFAR-100 datasets

Method	Peer competitor	Accuracy (%)		Number of parameters ($\times 10^6$)	Search time (GPU days)
		CIFAR-10	CIFAR-100		
Manually designed CNN	VGGNet	93.34	71.95	20.04	
	ResNet	93.57	78.84	1.7	
	DenseNet	94.76	75.58	0.8	
	Maxout	90.70	61.40		
	Network in network	91.19	64.32		
	Highway network	92.40	67.66		
	ALL-CNN	92.75	66.29	1.3	
	FractalNet	94.76	77.51	22.9	
Non-OA-based method	BO	92.91	66.47	4.70	0.02
	AK	88.56		1.7	
	NAS	93.99		0.8	22 400
	MetaQNN	93.08	72.86		100
	EAS	95.77		23.4	10
	Block-QNN-S	95.62	79.35	6.1	90
OA-based method	PSO	84.17	54.25	4.65/4.70	0.05/0.19
	Large-scale evolution	94.60	77.00	5.4/40.4	2750
	Hierarchical evolution	96.37			300
	CGP-CNN	94.02		1.68	27
	CNN-GA	96.78	79.47	2.9/4.1	35/40
	AE-CNN	95.70	79.15		27/36
	AE-CNN+E2EPP	94.70	77.98		8.5
	SHEDA-CNN	96.36	78.84	10.88/18.64	0.58/0.97
	Genetic CNN	92.90	70.97		817
Our method	GPPSO	95.26	76.36	5.26/4.44	0.04

The results before and after "/" are based on the CIFAR-10 and CIFAR-100 datasets, respectively

图 2　表 2 的复制品

中,GPPSO 取得了较好性能,在 8 种算法中排名第 5,比上一种算法提高 22.21%,比第一种算法仅低 3.11%。)

示例 9 主要存在 3 个错误:1)比较对象不一致,应用代词 that 指代分类准确率;2)that 引导限定性定语从句错误,宜用“,which”引导非限定性定语从句,因与本文主题无关,此处不展开具体分析;3)正文中错误地把百分点作为百分比直接进行比较,这类错误在科技论文中非常常见。从图 2 可知,95.26% −84.17% =11.09%,表达的含义是 GPPSO 分类准确率比 PSO 分类准确率高 11.09 百分点,在句中不可直接表述为 11.09%;(95.26% −84.17%)/84.17% =13.18%,因此 GPPSO 分类准确率比 PSO 分类准确率高 13.18%。其他比较数据处理类似。故示例 9 建议改为:

As shown in Table 2, for classification accuracy, GPPSO ranks the fifth among the 10 algorithms based on the CIFAR-10 dataset, with up to 13.18% higher accuracy than that of PSO and 1.57% lower accuracy than that of the first algorithm (CNN-GA). As for CIFAR-100, GPPSO achieves good performance, which is the sixth among the eight algorithms, 40.76% better than that of the last and only 3.91% lower than that of the first.

2.3.2　倍数的比较

此处特别关注数据比较中经常出现的倍数的比较问题,讨论 times more than 和 times as many as 的争议问题。

示例 10　According to our simulation results on the constructed CPU-GPU heterogeneous system model, the average number of GPU memory requests per unit time is 50 times more than the number of CPU memory requests.

表述 A is 50 times more than B 主要存在 2 种理解:A 比 B 多了 50 倍(A 是 B 的 51 倍)或 A 是 B 的 50 倍。《牛津高阶商务英汉双解词典》(第 10 版)中有表述:three times longer than sth 比某物长 2 倍,并列于 three times as long as sth 某物的 3 倍长[1];其认为 times more than 等同于 times as many as。但本文认为 more than 本身是比较级,应从加减法的角度比较,在比较级前加倍数应表示“比……大/小几倍”,three times more than 应是“多了 3 倍”。《实用英语语法教程》中的表述:“he deserved all these prizes because he spent three times more energy than I did on study。他在学习上花费的精力比我多 3 倍,所以他理应获得这些奖项。”[2]以及有道词典(科斯林)中的表述:“Its profits are rising four times faster than the average company。其利润正以快于一般公司 4 倍的速度增长。”可以很好支撑本文的观点。

针对关于倍数比较的表述,尤其是当数值较小时,编辑需要保持高度敏感,及时与作者沟通了解句子本意,并积极求证。此处经咨询作者得知,该句本意为:根据本文构建的 CPU-GPU 异构系统模型上的仿真结果可知,单位时间内 GPU 的平均内存请求数是 CPU

内存请求数的50倍。沟通后作者将该句改为：

According to our simulation results on the constructed CPU-GPU heterogeneous system model, the average number of GPU memory requests per unit time is 50 times that of CPU memory requests.

在中文科技论文中一般不用“降低了……倍”的表述，但是在英文科技论文中，尤其是一些专业领域的英文论文表述中，会遇到用降低的倍数表示性能提升，降低的倍数越大，性能提升越大。本文提出用 times + reduction/less 表示更小更少的倍数关系，并提出用分数的形式表达其对应中文含义。

示例 11 Compared with original NOVA and DWARM (the state-of-the-art wearleveling-aware allocator of persistent memory file system), experimental results show that the proposed WMAlloc can reduce 4.11 ×, 1.81 × maximum write and achieve 1.02 ×, 1.64 × performance improvement with four workloads on average, respectively. Furthermore, the proposed WMAlloc-BMT outperforms WMAlloc by 1.08 × and achieves 1.17 × maximum writes reduction with four workloads on average.

经了解，作者想表达的含义是：实验结果表明，在4种工作负载下，所提方法 WMAlloc 的最大写次数（即系统在所有存储单元上施加的写操作的最大值，值越小代表预期寿命越高）是 NOVA 最大写次数的 1/4.11，是 DWARM 最大写次数的 1/1.81（此处中文不能表述为与 NOVA 和 DWARM 相比，WMAlloc 的最大写次数降低了 4.11 和 1.81 倍；意味着 WMAlloc 的平均使用寿命分别提高至 4.11 和 1.81 倍）。与 NOVA 和 DWARM 相比，所提方法 WMAlloc 的性能分别提高了 1.02 和 1.64 倍。此外，在4个工作负载下，WMAlloc-BMT 的性能是 WMAlloc 性能的 1.08 倍，WMAlloc-BMT 的最大写次数是 WMAlloc 最大写次数的 1/1.17。

参考在 times 后用 less 的方法表示“更小更少”的倍数关系[3]，本文建议在 times 后用 reduction 表示降低的倍数关系（为遵从作者领域内的写法，论文中用“×”表示倍数），该句修改为：

Experimental results show that, compared with the original NOVA and dynamic wear-aware range management (DWARM), which is the state-of-the-art wear-leveling-aware allocator of PM file systems, WMAlloc can, respectively, achieve 4.11 × and 1.81 × maximum write number reduction and 1.02 × and 1.64 × performance with four workloads on average. Furthermore, WMAlloc-BMT outperforms WMAlloc with 1.08 × performance and achieves 1.17 × maximum write number reduction with four workloads on average.

3 结束语

本文归纳了一些常见的直观比较词以及其他不那么直观的比较词。针对科技论文比较语句常见的3类错误，提出具体的修改意见。1）编辑在处理结构复杂的句子时需要注意保持比较词前后比较对象的一致性，为保持句子简洁，一般在比较词后采用代词指代比较词前的比较内容，该代词不可省略，具体可根据比较对象名词单/复数形式选择合适的代词。2）编辑对一些固定搭配的短语以及典型的可修饰比较级的副词要了解清楚，大胆质疑小心求证，同时要核对正文数据，确保副词使用得当。3）正确区分百分比和百分点的概念。在英文科技论文以及中文科技论文英文摘要中一般较少使用百分点概念，宜用百分比表述所提方法性能的提升。本文详细阐述了倍数的表达方式，支持 times more than 不等同于 times as many as 的观点，尤其是当数值较小时，编辑应及时与作者沟通交流，保证论文数据的正确性。通过案例分析用 times + reduction/less 表达更小更低的倍数关系，并提出用分数形式表达其对应的中文含义。本文通过详细分析11个典型案例，提供每个案例的中文含义以及具体修改意见，希望能在编辑同行遇到相似问题时提供编校参考。

感谢《信息与电子工程前沿》（英文）编辑部翟自洋老师提供的示例11，以及他对本文写作的指导。

4 参考文献

[1] 商务印书馆. 牛津高阶商务英汉双解词典[M]. 10版. 北京：商务印书馆，2023

[2] 张干周. 实用英语语法教程[M]. 杭州：浙江大学出版社，2010

[3] 陈用仪. 英语常用词疑难用法手册[M]. 2版. 杭州：浙江大学出版社，2020

（2023-10-18收稿；2023-12-15修回）