Better those than that

Anu Alahari, a freelancer working in Caen, France, asks:
‘Could you please take a few minutes of your time to give your opinion on a simple language issue. Which of the following phrases is grammatically most correct? And which one would you use in medical writing?

The time-weighted average leukocyte counts:
1. were higher in group X than group Y.
2. were higher in group X than in group Y.
3. were higher in group X than that in group Y.
4. were higher in group X than those in group Y.

Now, it is a tall order to decide which is grammatically ‘most correct’.

It is easy to exclude sentence [3] as grammatically incorrect, because that refers back to a plural noun, counts, in the subject of the sentence. This means that [4] is fine, because instead of that we have those referring back to the counts. But does those actually have to be there? Why not just use [2]? Or even just [1], because it is unlikely that any reader is going to misunderstand the import of this sentence, whether you use [1], [2] or [4]. Even though in is not repeated before group Y in [1], this seems acceptable because group X immediately precedes group Y in the same adverbial phrase, and the prepositional effect of in is easily transferred to group Y when reading. We do this a lot in short adverbial phrases, particularly when speaking.

Let’s see what happens if we change the word order (excluding [3]):

The time-weighted average leukocyte counts:
5. in group X were higher than group Y.
6. in group X were higher than in group Y.
7. in group X were higher than those in group Y.

Interposing the verb between group X and group Y in [5] make it unacceptable to leave out the in as in [1]. With [5], the reader has the feeling that something is missing. This is because the prepositional effect of in is not transferred in the reader’s mind because we have two separate adverbials split by the verb, and the reader has to contend with the verb before reading about group Y. So with this word order, we should choose [6] or [7].

There are prescriptivists out there who would insist that [4] and [7] are the ‘most correct’ way of expressing this type of information, because they both include those, which means that it is absolutely clear that we are talking about the time-weighted average leukocyte counts in both groups X and Y. [2] and [6] would be good enough for me because the maxim I always apply is ‘Can this sentence be misunderstood?’ and for me, [2] and [6] cannot be misunderstood.

Every situation has to be assessed individually by the author or editor, however, and there may well be instances where the additional ‘security’ of adding in those is useful for comprehension or reader comfort. And this certainly applies to the following examples:

The group X time-weighted average leukocyte counts:
8. were higher than group Y.
9. were higher than in group Y.
10. were higher than those in group Y.

[8] can be discounted for the same reasons as [5]. The two groups are too far apart in [9] for this to be a comfortable option, although it is just about all right. So we are left with [10] where those definitely contributes to reader comfort and comprehension. However, I expect that most of us would only opt for [9] or [10] because of word-count constraints as they are certainly the least elegant.

So, my rank order of choices is [2], [6], [4], [7], [10].

Alistair Reeves
a.reeves@ascribe.de