Supplementary Material for "An agent-based model of hierarchical information-sharing organizations in asynchronous environments"

[Authors here]

This document covers extra analysis mentioned in the original paper.

Using a group size of 10 versus smaller groups

Most larger organizations do not have more than 10 offices reporting to the head office. Some larger government agencies have more than 10 offices reporting to the top cabinet-level post (usually Secretary), but also include a position equivalent to a Deputy Secretary, reducing the management burden. Therefore, a group size of 10 should be sufficiently valid for even the top management of large organizations. A helpful side effect is that keeping the group size at 10 also reduces computational burden when running simulations. However, we would be concerned if group sizes in the range of 4 to 10 displayed different behavior, rather than performance that scaled slightly with size.

In the paper, we write "We are most interested in the relative differences of the groups, which is most easily seen with larger groups: we use groups of 10 for this analysis. Smaller groups have the same differences but with smaller spread (see Supplementary Material)."

We can show this by running simulations with three group sizes (4, 7, and 10 workers) plotting the histograms of these different group sizes for comparison. The spead is most clear in the no-hierarchy histograms. The spread in the hierarchy condition does increase slightly, but not as much as the no-hierarchy condition. In addition, the hierarchy condition distribution is sensitive to even vs odd numbers of workers due to the manager taking the mode of the worker responses. When the number of workers is even, the number of worker responses for the manager to consider is even. That means that the mode can be both 0 and 1 (say, five workers report 0, five report 1). In that case, neither is more legitimate, so the manager chooses one at random. When the number of workers is odd, this can never happen. This difference creates the small disparities between even and odd worker distributions. This effect, in addition to the interaction inherent in the hierarchy processes themselves, makes the hierarchy condition distributions less clean in their spread compared to the no-hierarchy condition distributions.

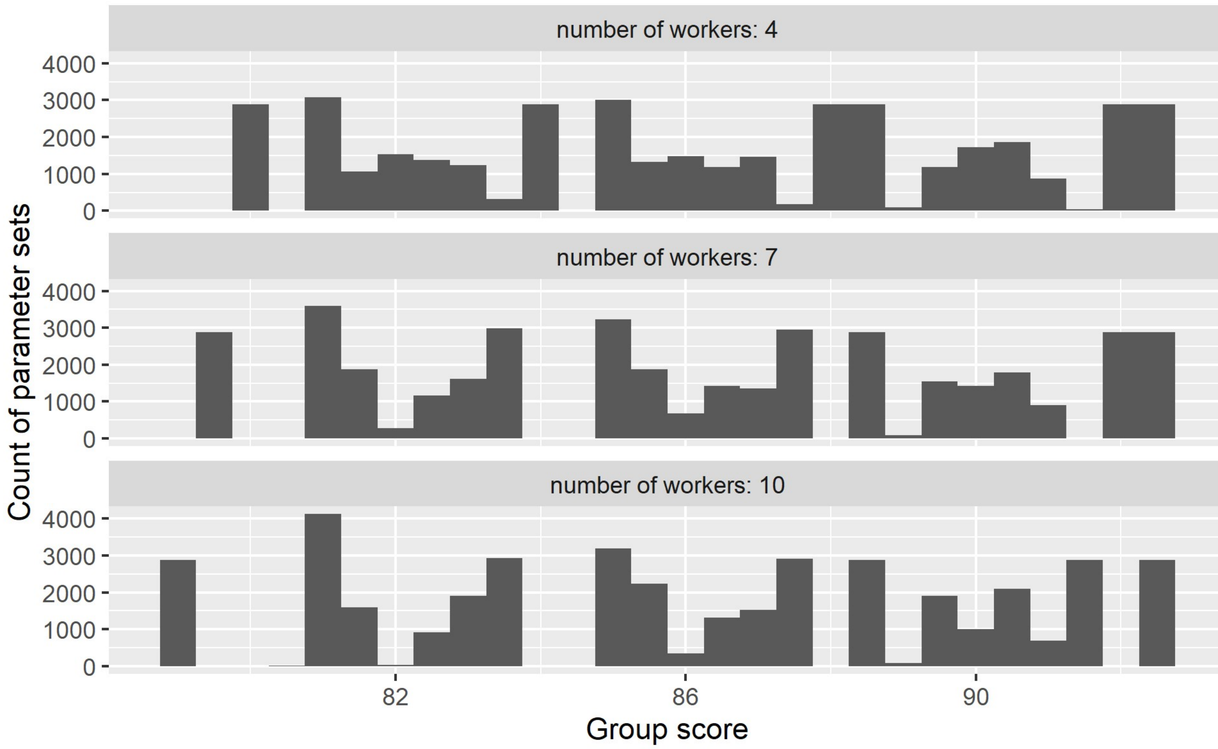

Figure 1. Histograms of the no-hierarchy condition with different numbers of workers. These show that increasing the number of workers increases the spread of scores slightly.

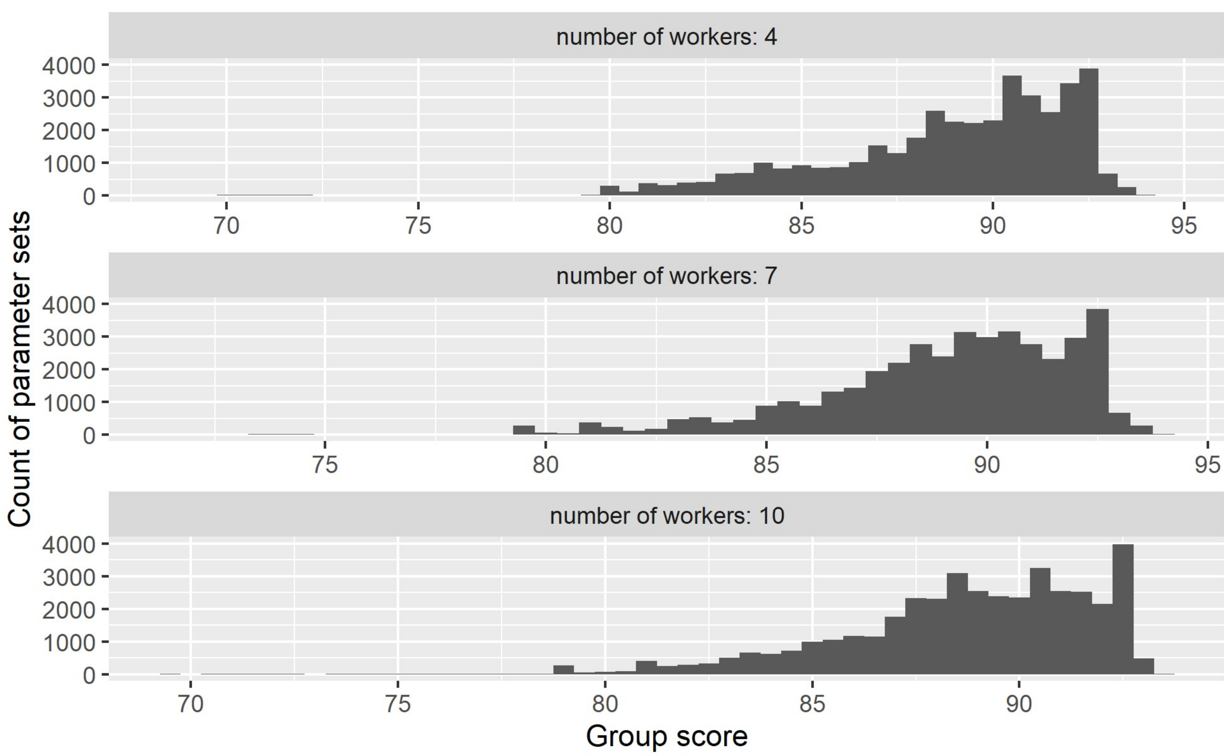

Figure 2. Histograms of the hierarchy condition with different numbers of workers. These show that increasing the number of workers increases the spread of scores slightly.

The variance in identical-parameter runs doesn't change the results

In the the paper we state, "We run 20 simulations for each unique parameter set, then calculate the mean score for those runs as the score of a typical run. The variance in score among the 20 runs per parameter set does not substantially change the results of our analysis (see Supplementary Material)."

This claim covers all the findings and figures in the original paper. Below, we recreate the figures while highlighting the data points that would impact the analysis.

The variance of simulations with identical parameter settings

Variance in the hierarchy conditions is has a long tail, as shown by plotting the variance in the sets of 20 runs in Figure 3.

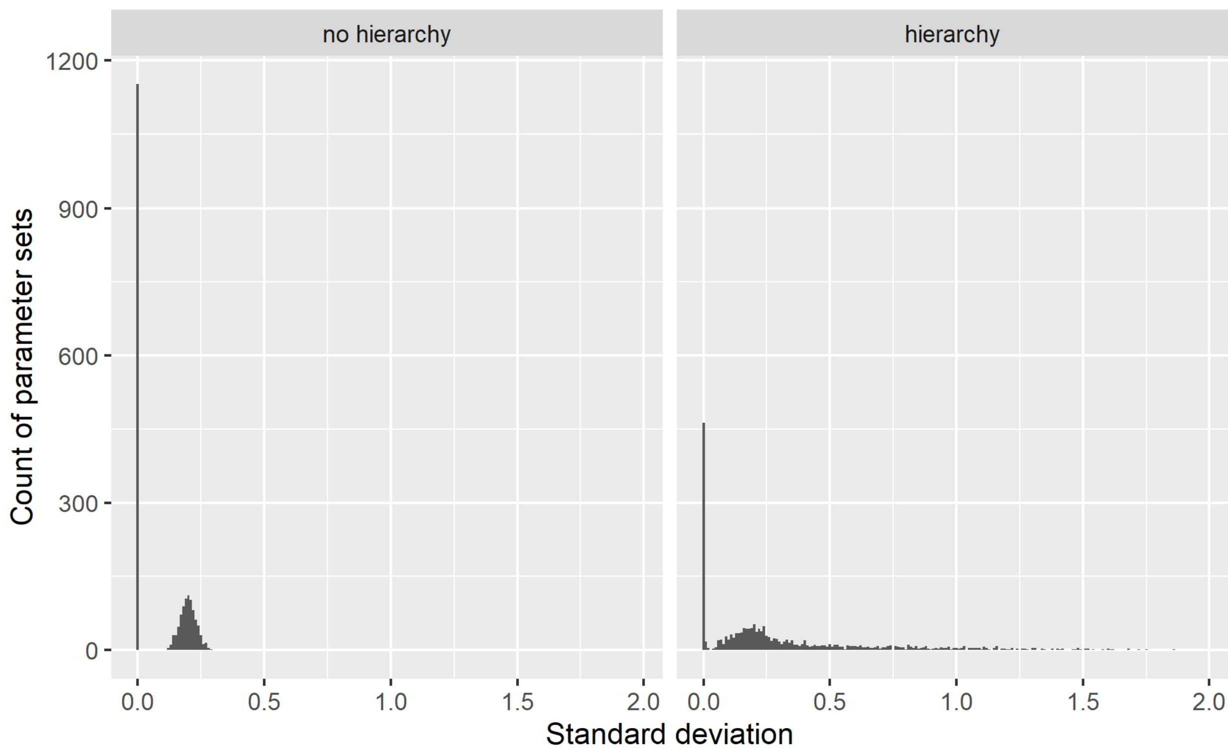

Figure 3. Histograms of the variance in scores between runs with identical parameter sets, plotted for the no-hierarchy and hierarchy conditions. The hierarchy condition has a distribution of variance with a long tail, which provides sufficient reason to investigate whether variance affects the results of the study. It does not, as we show with further analysis below.

Assessing the impact of sets of runs where hierarchy straddles 0 in performance

In the paper we write, "If those replication-sets with bounds that straddle zero are considered as zero instead of their mean, hierarchy does better in 75% of cases, worse in 4%, and makes no difference in 15% (see Supplementary Material)."

In other words, some parameter sets, when run 20 times, have variance in their performance. In the paper we use the mean of these runs. However, with the variance in run scores, it may be that sometimes hierarchy outperforms no-hierarchy, but sometimes does worse, depending on which particular runs one picks out of the two sets of 20 (20 for hierarchy, 20 for no-hierarchy, all other parameters identical). Thus, we need to see how many comparisons might be flawed because their score distributions straddle 0.

We do this by calculating the difference between the min hierarchy group scope and max no-hierarchy score, and also calculating the difference between the max hierarchy and min no-hierarchy. We don't know which will give us the higher or lower scores, so we just have to calculate them and assign the lowest score the lower bound and the highest score the higher bound. This gives us the most reasonable bounds on scores that we can use as "error bars" for the distribution of scores for a particular parameter set.

We can then identify the sets with score distributions that straddle zero, and see if accounting for these sets in any way changes the results of the study. We recreate the figures in the paper to highlight where these sets sit in the plots (see Figures 4 to 9).

sets straddling 0 included?	hier_better	hier_zero	hier_worse	ratio hier/no-hier	hier better %	hier zero %	hier worse %
FALSE	1686	135	111	15.19	0.87	0.07	0.06
TRUE	1450	281	68	21.32	0.75	0.15	0.04

Table 1. Removing sets whose score distributions straddle 0 does not change the overall results. The two rows, one that includes the sets that straddle 0, and one that removes them, have similar percentages for the number of sets where hierarchy performs better, versus sets where no-hierarchy is better.

The first figure looks at whether the two types of environmental asynchrony have different effects on performance (Figure 2 in the paper).

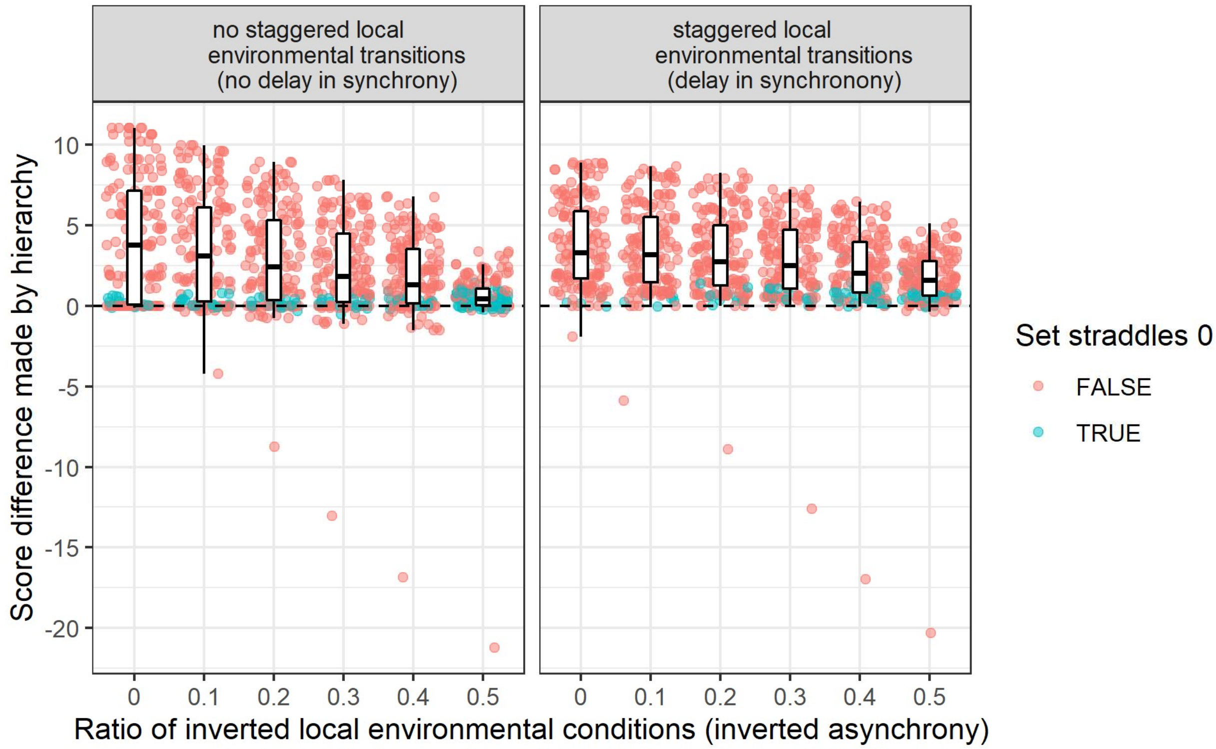

Figure 4. Group performance by the extent of environmental asynchrony. The y-axis shows the score difference by hierarchy, with positive values indicating hierarchical groups perform better, negative values that they perform worse. The x-axis shows increasingly disordered environments using the ratio of local conditions that are the inverse of the base environmental condition. Figure 2 in the original paper, recreated to take account for run sets that straddle 0.

The next figure re-creates figure 3 in the paper, investigating the the importance of autonomy. As before, hierarchical groups with agents with this autonomy do better in nearly all conditions than non-hierarchical groups, including in highly asynchronous environments (those where half the lanes in the environment are inverted).

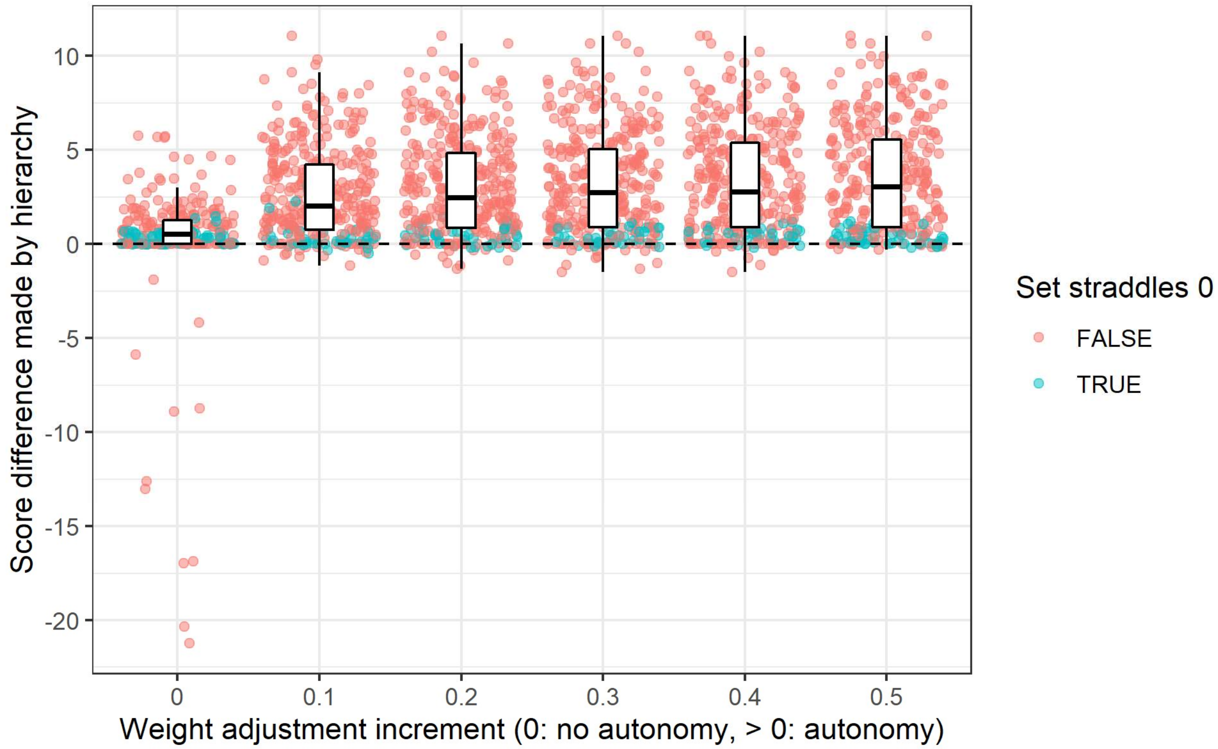

Figure 5. Distributions of group performance by their weight adjustment increment. The y-axis shows the score difference by hierarchy, with positive values indicating hierarchical groups perform better, negative values that they perform worse. The x-axis indicates the increment by which an agent adjusts the weight they put on manager input, where the weight can be 0 to 1. Agents adjust their input based on whether they or their manager got the local environmental condition right or wrong. Figure 3 in the original paper, recreated to take account for run sets that straddle 0.

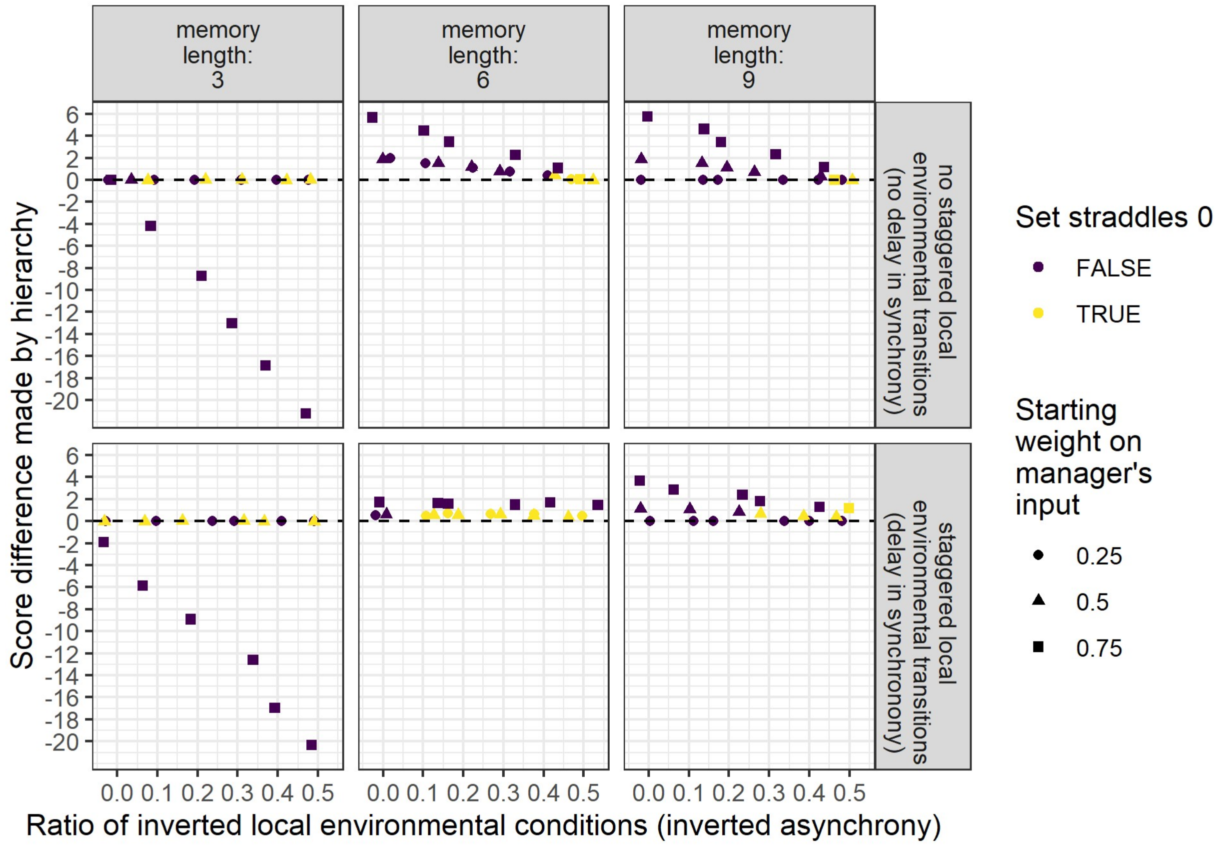

Figure 6. How starting weight relates to group performance in agents *without* autonomy. The x-axis indicates the starting weight an agent places on manager input, where the weight can be 0 to 1. For example, an agent putting a weight of 0.75 on their manager's input would place 0.25 weight on their own memory of their local environmental conditions. In this no-autonomy condition, agents are never allowed to change their weights. Figure 4 in the original paper, recreated to take account for run sets that straddle 0.

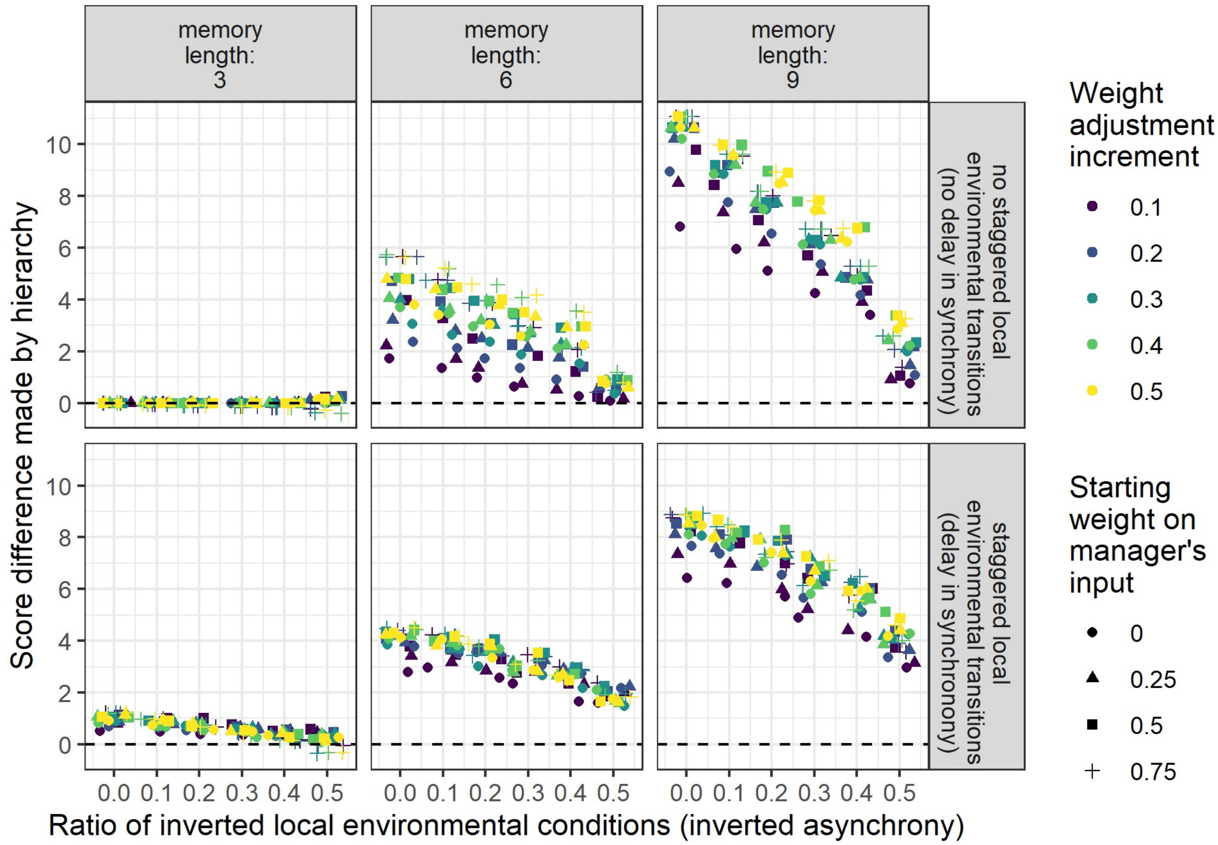

Figure 7. How starting weight relates to group performance in agents *with* autonomy. The x-axis indicates the starting weight an agent places on manager input, where the weight can be 0 to 1. For example, an agent placing a weight of 0.75 on their manager's input would place 0.25 weight on their own memory of their local environmental conditions. In this autonomy condition, agents are allowed to change their weights, once each time unit, by their set weight adjustment increment. Agents adjust their input based on whether they or their manager got the local environmental condition right or wrong. Figure 5 in the original paper.

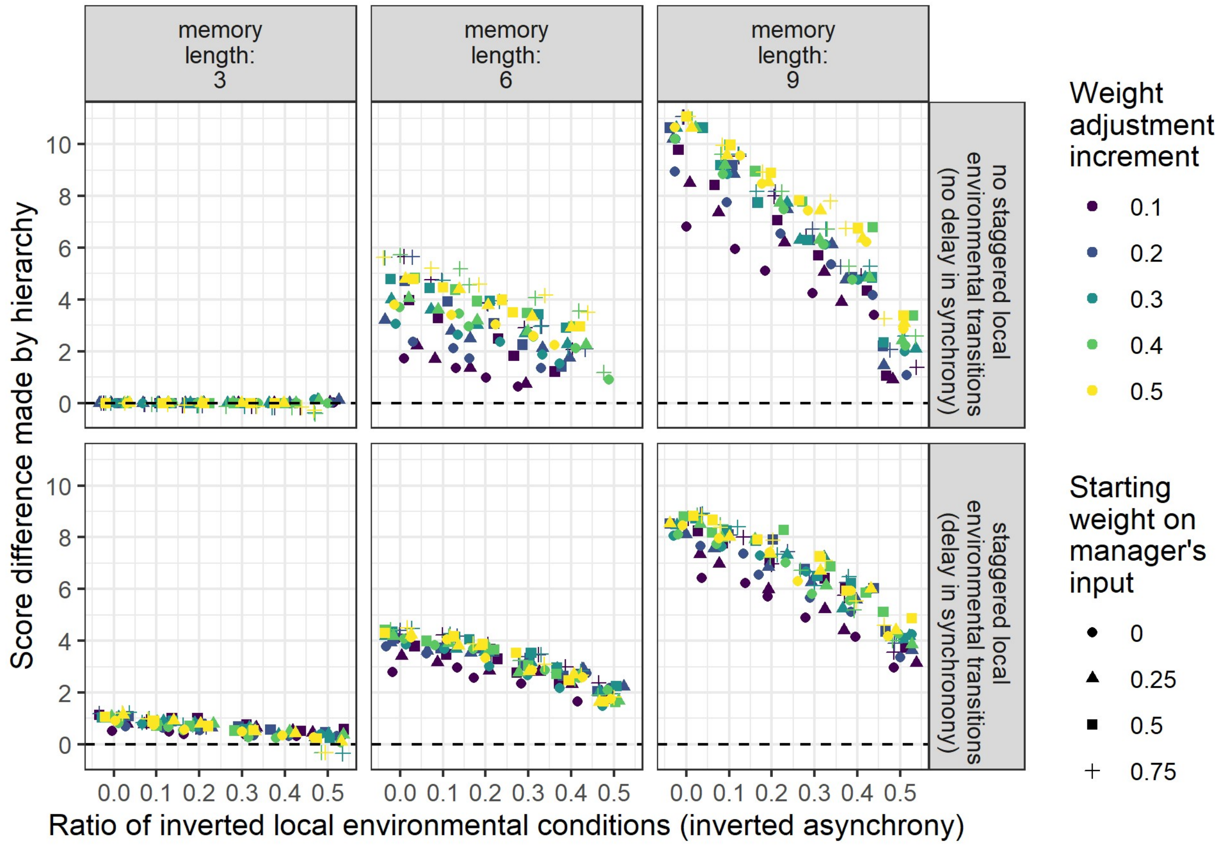

Figure 8. Figure 7 (Figure 5 in the original paper) with run sets removed that straddle 0 in their score distributions. Figure 5 in the original paper, recreated to take account for run sets that straddle 0. Note that random differences in the jitter of the two plots creates slight differences in the horizontal position of the points between the two plots.

With the whole set plotted, we can see the sets that straddle 0 are very close to zero, not spread out or otherwise biasing the data in a substantive way (see Figure 9).

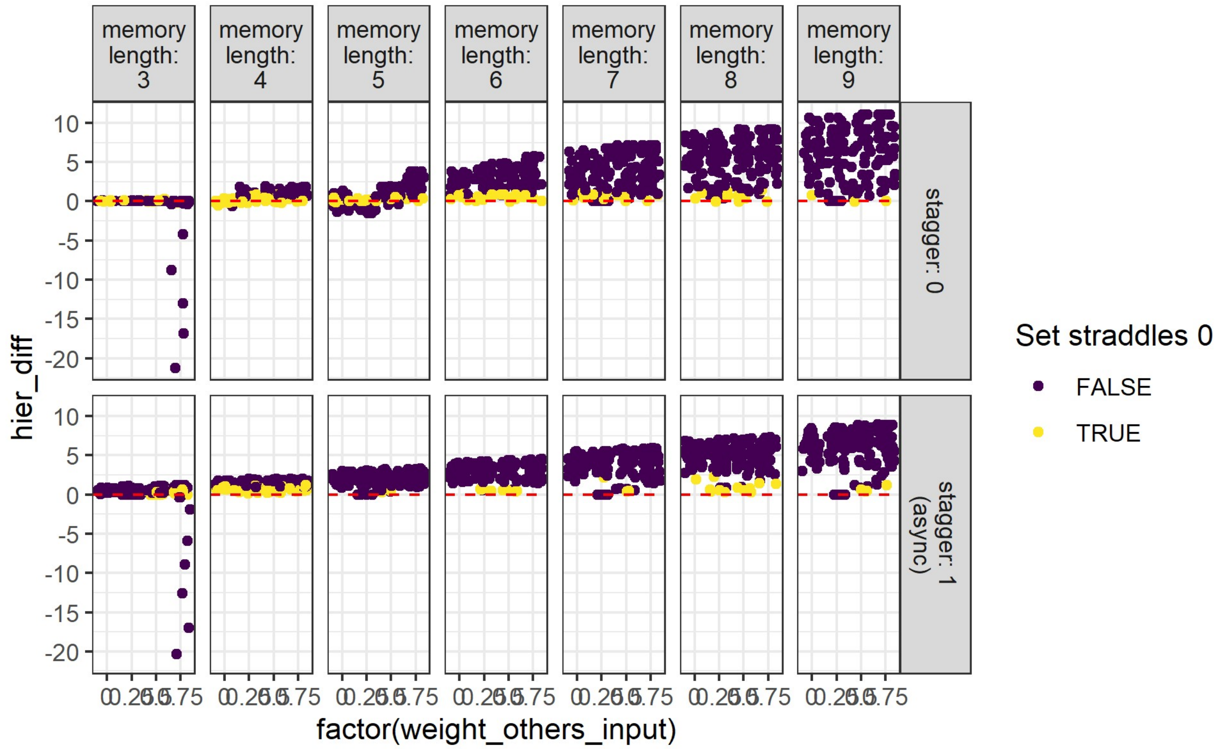

Figure 9. The sets with score distributions straddling 0, plotted among the other set data points. All sets are plotted by the mean of the scores.