A note on the scenario task complexity progression across freeze points

The scenario tasks that the teams were working on in the current experiment were intended to start by posing a complex, dynamic and unstructured problem to the teams, that would gradually be resolved and become clearer as the teams worked with the task. Based on this, we expected the degree of shared mental models (SMM) and situation awareness (SA) for team members to increase across scenario time. This would manifest as a linear increase in the SMM and SA measures across the four measurement times (the four freeze points in the scenario). This yielded the following hypotheses: *H3a: Individual SMM will increase across scenario time* and *H3b: Individual SA accuracy will increase across scenario time*.

However, during our data collection it became clear that there was a discrepancy between the planned and the actual execution of the scenario exercises. In line with their traditional approach, our industry partners adjusted the complexity of the task and the intensity of the workload to pose challenges to their teams in order to improve the training effect. Explicitly, in most or all of the scenario exercises in the current data collection, an unexpected change was introduced to the task about halfway through the scenario. This practice became clear to the researchers after the data collection was underway, and was confirmed to us by our industry partners towards the end of the data collection.

Given this new information, the test of hypotheses H3a and H3b are unlikely to find a linear increase in SMM and SA over scenario time, as the task complexity is designed to increase between the second and third freeze point. Rather, the actual task complexity can be said to progress according to an ABAB pattern (or a complexity that may progress roughly as 10 - 5 - 8 - 3, in arbitrary units). Thus rather than a linear relationship as described in the hypotheses, we may expect the relationship to be cubic in nature.

In addition to testing the pre-registered analysis of linear relationships for H3a and H3b, we will thus also test for cubic relationships. To some extent, this issue may also have an impact on hypotheses H4a and H4b, but as the implications there are less clear there, we will refrain from changing these hypotheses.

Although the pre-registration noted that H3 and H4 could be tested with ANOVAs, it may be more appropriate to use simple regressions to test for linear or polynomial changes in the dependent variables between freeze time points 1, 2, 3 and 4.