Many thanks to the anonymous reviewers of our manuscript for their thoughtful comments and suggestions which were important to improve the quality of the paper. We revised the paper accordingly.

The most important changes are:

- Present dimensions in the Introduction with an improved Figure 1.
- Show how examples of the terminology are used in typical evaluation settings (Section 5).
- Revised writing to improve clarity.
- Corrected spelling and grammar errors.

We took each comment very seriously and put significant effort into this paper revision to address most reviewers’ suggestions. Additionally, we would like to comment specifically to the reviewers’ suggestions below.

Meta Reviewer 3

“I do agree with R2 that the paper is dense and hard to follow in places. It would benefit from the model figure being presented earlier and maybe presenting an additional table with the components (perspective, level of composition, ...).”

“For the camera-ready version, please try to improve the writing and add some clarity to the model description. E.g. by adding a more thorough overview list, table, or figure, introducing more subsections.”

We revised the complete text in order to improve clarity. We took up the advice and present a list of abstraction, composition, and perspective in the Introduction and move Figure 1 to the Introduction as well. We refrained from adding a table of the dimensions in order to avoid duplicated content. We changed the order of subsections in Section 2 and added paragraph headers in Section 3. Furthermore, the complete text has been revised carefully and many sentences have been rephrased.
“It would also be interesting to speculate on effects of changing terminology on visualization evaluation in general (e.g. how do you recommend to formulate the current “task description” of a regular user study?).”

“One aspect of the paper that could be more strongly argues is how a change in terminology (and the subsequent distinction between problems and actions) would impact evaluations of various types.”

We adopted this advice and the new Section 5 shows how our terminology applies to the concept of ‘task’ in typical evaluation settings.

“Is it that qualitative field evaluations are more concerned with problems that actions? Whereas controlled experiments are more concerned with actions?”

As we explain in Section 5, we would classify the stimuli of a typical controlled experiment in visualization as problem – even if the problem description is anticipates the answering action such as “click on the item representing the most critical condition”.

“Have the authors analyzed any recent user studies to see how people have mixed (or not) the two concepts?”

A thorough analysis would be necessary to avoid bias from few examples and that this would be too much effort in the given timeframe as well as go beyond the scope of this position paper. We have added it to the future work section.

“Also a few sentences on what the “community consensus glossary” is for the temporal database community would aid in understanding how the same thing could be useful for visualization.”

We describe more details about both the existing glossary for the temporal databases and the envisioned glossary for visualization.

“The acknowledgment text needs to be filled.”

We added the acknowledgment text.

**Assigned Reviewer 1**

“One small issue is that this topic is relevant to evaluation (the focus of BELIV) but probably not quite as central as some of the other submissions, I am guessing. The topic of this paper is relevant, but isn’t about evaluation methodologies exactly.”

We think that the paper is relevant to the audience of BELIV and are reconfirmed by the positive reviews. Evaluation is an essential aspect of visualization throughout design and development and either problems or actions or both play a relevant role in these evaluations. We discuss in the new Section 5 how our terminology applies to the concept of ‘task’ in typical evaluation settings.
“Some small improvements for the paper’s text:
Section 1: "we speak out the warning to use" Reword this, as it is awkward.
Section 2.4: Substitute "interconnectedness" for "intertwindingness".
Section 3: "could be broken down to include" (flip words)
Section 5: "we critically analyze" (not analyzes)"

We incorporated all these improvements.

**Assigned Reviewer 2**

“But I’m afraid the presented model doesn’t really address many of the issues,...”

In general, we agree that there are further issues to be addressed in future research. With this paper we hope to raise attention to the topic and provide a reference for other researchers.

“. . . the paper is so dense and hard to follow that I can’t see many people slogging their way through it.”

We completely revised the paper, changing the order of sections and reconsidering each sentence so that the writing is now clearer.

“The introduction is too long and doesn’t provide an overview of what is going to be presented. Rather than keep making the case for how important the issue is, it would have been way more helpful to briefly sketch the model that will be presented later.”

We adopted the advice, shortened the motivation and present a list of abstraction, composition, and perspective already on page 1.

“The Survey (which should be titled ”Literature Survey”, since survey implies asking people) provides a lot of data, but is also very dense.”

We changed the section heading as suggested and worked on the text to improve clarity.

“The model is very abstract, and doesn’t seem to provide much help in disambiguating tasks and problems. In fact, I’m more confused now about the distinction than I was before. I don’t see any demonstration of this model being superior to any of the other uses of task, etc. Why not tie this model back to some of the literature surveyed earlier? Or find some other way to actually use the model.”

We revised the structure and content of the paper to make the paper less confusing and thus more helpful. In particular, we restructured Section 3 to emphasize the examples and the distinction between each combination of dimensions. Furthermore, the new Section 5 demonstrates the terminology in practical evaluation settings.

“The best part of the paper is the Discussion section, even if the first sentence is really unfortunate.”
We rephrased the sentence.

Again we thank the anonymous reviewers for their thoughtful comments and suggestions that helped us to enhance the quality of our submission.

Best regards,
Alexander Rind, Wolfgang Aigner, Markus Wagner, Silvia Miksch, Tim Lammarsch