Comparing online with paper pencil testing mode when solving stock flow problems using different problem solving strategies

Vivien Lungwitz

Technische Universität Chemnitz
W.-Raabe-Str. 43, 09120 Chemnitz, Germany,
+49-371-531 34143
vivien.roeder@psychologie.tu-chemnitz.de

Keywords: dynamic systems; stock–flow systems; online testing; paper pencil testing; general problem solving; problem solving strategy; gender effect

Introduction: Classic data collection requires participants to come into a laboratory and solve the tasks given on site in a paper pencil format. In past years, it became more and more popular to implement tests via internet. However, it seems not always clear whether results from both methods can be compared that easily. For some fields of research, comparative studies can be found. As of the author’s knowledge, no such comparison has been explicitly done for stock flow tasks yet. This study gave five stock flow tasks to two independent samples of participants either in a classic paper pencil setting at university or via internet. For each setting, about half of the participants were prompted to solve the tasks either with a conventional (CPS) or general (GPS) problem solving strategy. Slight improvements in subjects performance on stock flow tasks were recently observed when participants were primed to use analytical thinking right before being confronted with a stock flow task (Lakeh & Ghaffarzadegan, 2015, 2016) or when they used a GPS strategy instead of a CPS strategy (Röder, 2017). The idea to manipulate the problem solving strategy used by participants before working on stock flow tasks was derived from Youssef-Shalala, Ayres and Schubert, (2014).

This study tried shed light on the following questions: Does it make a difference for the average stock flow performance whether the testing is carried out in a paper pencil setting or via internet? The hypothesis was that it does not make a difference as long as the demographics of the sample are comparable. A second question was, whether prompting the use of a GPS strategy instead of a CPS strategy could boost stock flow performance. Based on previous studies (Röder, 2017; Youssef-shalala et al., 2014) a slight advantage for the GPS was expected. Additionally, the simultaneous use of drawing as well as question based stock flow tasks allowed a direct comparison between different task formats as dependent variables. Finally, the
impact of additional variables that have shown to influence stock flow results in the past, should be calculated (e.g. gender, grade in mathematics, field of study, motivation).

**Method:** Participants were acquired using mailing lists of different universities in Germany and advertising the studies in social media. The overall process was the same for everybody: After a general information about the study and the agreement to participate voluntarily and anonymous, participants (125 online, 46 paper pencil) read a short introduction on simple dynamic systems and then were given a total of five stock flow problems. Their understanding of how the stock develops was measured by drawing the development of the stock over time (four tasks) and by answering four questions (one task). Both types of tasks were used in past research. Tasks varied in difficulty and had been used in prior studies.

In the general problem solving condition \((n = 92)\), participants were not directly asked for a decision on the development of the stock but rather first prompted to think deeply about the given flow diagram and asked to gather all the information possible before asked about the stock.

**Results (Fig.1):** Overall, the online sample yielded somewhat better results, as did the GPS strategy for the four drawing tasks. Solution rates for the question based task were lower in general and did not produce a clear outcome towards setting or problem solving strategy. Gender had an impact in all groups, favoring men.

**Fig. 1.** The average performance in the four drawing tasks (left) and the question based task (right) separated by setting, problem solving strategy and participants’ gender. Error bars indicate the 95% confidence interval.

Other measures were collected to see whether the samples drawn differed in regard to the setting and whether those demographic data influenced the results. Multiple regression analysis were calculated for drawing and question based tasks. Gender is the best predictor in both models. For the drawing tasks the manipulation of setting and problem solving strategy also remain highly significant even after controlling for additional variables. Participants’ self-rated interest in mathematics also explains a huge deal. The last high school grade in mathematic misses significance on a 5% level but would remain in the model with a more modest alpha cut off. For the question based task the overall predictive value of the regression model is a lot lower than for the drawing tasks. This reflects in only two variables adding significant value to the model: gender and participants’ last high school math grade.
**Discussion:** Concerning the manipulated variables of the study it can be said that this study found an effect for setting (contrary to the prior hypothesis) and a moderate advantage for the general problem solving strategy (in line with the hypothesis). The self-selection bias in the online setting might play an important role in the different results. The rather small sample size in the paper pencil setting makes it difficult to draw far-reaching conclusions.

In addition, the quantitatively different outcomes for drawing and question based tasks raised the question, whether both tasks are comparable and measure the same. Three difficulties in interpreting the results between the different task types are that there was only one question based but four drawing tasks and the question based task was always presented last. Furthermore, the question based task could only be rated as wrong (0%) or correct (100%), whereas the drawing task got a rating between 0 and 100 percent.

Another major finding was a pronounced gender effect on performance in all conditions.

**References**


http://doi.org/10.1016/j.obhdp.2008.03.003


Veldhuis, G., & Korzilius, H. P. L. M. (2012). Seeing with the mind - The role of spatial ability in inferring dynamic behaviour from graphs and stock and flow