Update to the Pre-Analysis Plan for Harnessing AI to Spark Curiosity: Experimental Evidence from Middle School Education" (AEARCTR-0017865)

We will start collecting the endline data for the project "Harnessing AI to Spark Curiosity: Experimental Evidence from Middle School Education" (AEARCTR-0017865) from June 2nd through late June.

This document is a short update to the original pre-analysis plan (PAP) concerning the Primary Outcome measuring student's willingness to pay for D- and I-curiosity tasks in the behavioral measure of curiosity (only measured at the endline):

Robustness checks for knowing the answer.
We separate between D-tasks (mysteries) for which a student selected "I am not interested", and in the follow-up questions selected "I definitely know the answer" and "I think I know the answer". While the first version of the PAP proposed to code the WTP in case either of the two options for knowing the answer was selected as missing, we realized that students answering "I think I know the answer" may still have different motivations to reveal the answer from individuals who are convinced that they know the answer with certainty. For this reason, we run two sets or robustness checks: 1) as in the original PAP, 2) coding WTP as missing only for D-curiosity mysteries for which "I am not interested" and "I definitely know the answer" was selected.

Maximum willingness to pay.
As a robustness check (secondary outcome), we also consider the maximum willingness to pay in a given class (D- or I-curiosity) instead of the average (our primary outcome) for both the D- and I-curiosity tasks combined (1 variable), as well as for D- and I-curiosity tasks separately (2 variables). If student interests are narrow or if they know answers for some D-curiosity mysteries, we may be interested in measuring the maximum willingness to pay.

Regression analysis.
We enrich the regression estimation proposed in the original PAP by including two other variables that our behavioral measure collects:

1. time it took to complete the trial task (in ms) and
2. a likert scale variable for self-reported task difficulty that we included in a post-task survey.

The aim is to increases precision by capturing the likely heterogeneity in how the real effort task (the price for obtaining information) is perceived by the students.

We also introduced a gendered example in the instructions to the behavioral task (Vincent and Bianca; see the full script of the behavioral tool) and we randomly manipulate whether a higher willingness to pay example is associated with either of the two individuals.

- We thus also include a control for whether the higher willingness to pay example was introduced by Bianca (female) (indicator variable).

The aim is to control for possible gender-congruent role-model effects or similar.

Translations.

- The original script was first written in English. All Czech materials were translated and edited by two of the authors who are native speakers and several other team members. All materials were then translated from the Czech version to Slovak and edited by one of the authors who is a Slovak native.

Pilot of the behavioral tool.

- The behaviral tool was piloted at one elementary (8th grade, three classes) and one high school (3rd grade of a gymnazium, one class) school with 14-15 year old students. Instructions were adjusted to reflect the comprehension and attentional issues faced by the students. The key design remains equivalent to the description in the original PAP.

Supplementary materials.

For transparency, we also attach in separate documents:

- The entire Qualtrics code of the final version of the WTP task
- The set of mysteries offered in the WTP task
- A calibration exercise using two major AI models that assesses the mysteries along several dimensions.