SCHOOL STUDENTS' SKILLS TO COMPARE DATASETS

Burkhard Priemer and Wiebke Musold
Humboldt-Universität zu Berlin, Institut für Physik
Newtonstr. 15, 12489 Berlin, Germany
priemer@physik.hu-berlin.de

In a study with 156 school students, we investigated the influence of the number of decimal places in measurements on students' decision to change or keep an initial hypothesis in physics. Participants were introduced to two experiments for the same quantity, provided with datasets (three groups with different numbers of decimal places), asked to compare the sets, and reconsider their hypothesis. Results show: When the number of decimal places increases, 1. the number of students that change from an incorrect hypothesis to the correct one decreases, and 2. the number of students that change from the correct hypothesis to an incorrect one increases. This shows that exact results may hinder learning due to students' lack of knowledge in core statistical quantities like the variance.

INTRODUCTION AND THEORETICAL BACKGROUND

Being able to judge the quality of evidence is a core competence school students should acquire (Chinn & Malhotra, 2002). Evidence is often drawn from empirical data that can be obtained for example from experiments in science. In order to make inferences with experimental data in justifications - which means to use data as evidence for a certain claim - data have to be analyzed and interpreted (see for example the different models of experimental work: Klahr & Dunbar, 1988; Osborne, 2014; Millar, Le Maréchal, & Tiberghien, 1999). The relevance of a justification depends on the quality of experimental data and the quality of quantitative experimental data depends – among other things – on measurement uncertainties. Hence, it is obvious that an estimation of measurement uncertainties is a necessary requirement in scientific argumentation that is based on quantitative experimental data.

Comparing datasets obtained from measurements is an example that shows the importance of estimating uncertainties in measurements, for example when investigating if the mass of a bob of a simple pendulum has an influence on the time of oscillation. However, research shows that students' competences in statistical methods are often very weak. This difficulty has been demonstrated for example in the field of evaluating measurements uncertainties (Buffler, Allie, Lubben & Campbell, 2001; Lubben & Millar, 1996; Priemer & Hellwig, 2016). Due to a lack of deeper background knowledge in statistics, students may rely on their limited conceptions (e.g., knowing how to calculate a mean) or use their intuition or heuristics they've heard of (e.g., "more measurements are better"; similar to p-prims suggested by di Sessa, 1983). Kanari and Millar (2004, p. 749) were able to show that students keep their beliefs - that are based on prior knowledge and everyday life conceptions - when data are not unambiguous. However, if the data reach a certain level of clarity, students are more willing to change their beliefs (Kanari & Millar, 2004, p. 762). Thus, it seems likely that data features like the variance influence students' conclusions. There is little research on what data characteristics students refer to when judging datasets and making inferences, and knowledge of such practices can help in understanding the misconceptions students hold, as well as in how to address them.

RESEARCH QUESTION

The study investigated the influence of the number of decimal places in the result of a physics experiment on school students' decision to keep or change their initial hypotheses about a physics context. More precisely: What is the influence of the number of decimal places in the result of a physics experiment on school students' choice a) to reject an incorrect initial hypothesis in favor of a correct hypothesis and b) to reject a correct initial hypothesis in favor of an incorrect hypothesis.

METHODS

Participants were $n = 156$ high school students attending the 8th, 9th, and 10th grade of an urban school in Germany (average age: 14 years). First, students were introduced to a physics experiment in a 5-minute video: the comparison of the motion of one object (a ball) when it a) falls freely from a certain height to the ground (setting 1) and b) when it rolls down a ramp and thus getting a non-zero initial horizontal velocity and then falls freely from the same height to the ground (setting 2, see figure 1). After demonstrating the two experimental settings successively - without making measurements - students were asked to predict (multiple choice question) and justify (open text question) the time it takes the ball to cover the same distance (from the same height to the ground) in the two different settings (with a zero and a non-zero initial horizontal velocity). After the students stated their hypothesis (which we denote as their initial hypothesis) they were randomly assigned to three different groups and were provided with quantitative results of the experiment (a dataset with six measurements for each of the two experimental settings) in three different forms: with two, three, or four decimal places. The students in each group were asked to analyze the given data and to review their initial hypothesis in the light of the quantitative results provided (multiple choice question).
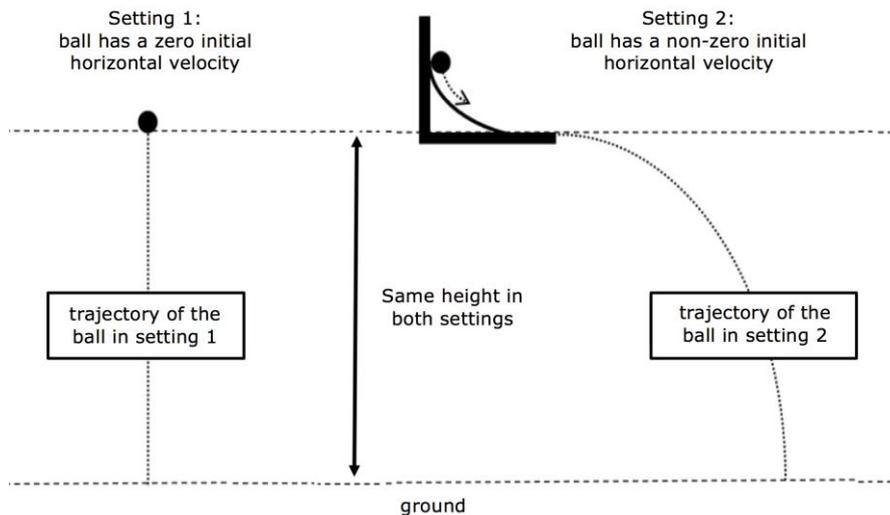

Figure 1. The two experimental settings.

RESULTS

Choice of the hypothesis before and after the experiment

Before the experiment was conducted, 32 % of the participants stated a correct initial hypothesis: The time it takes the ball to cover the distance is identical in both experimental settings. After students were provided with the datasets, the number of correct choices increased to 61 % (see Table 1). The two distributions of the percentages of students' choices of the hypotheses (before and after the experiment) differ significantly from each other ($X^2 = 46.69$, $df = 2$, $p < .001$).

Table 1. Percentage of students choosing the different hypotheses before and after the experiment.

The time it takes the ball to cover the distance…	Initial hypothesis (before the experiment); Percentage of students	Hypothesis after the students were provided with quantitative experimental results; Percentage of students
…is higher when the ball has a zero initial horizontal velocity.	9	26
…is higher when the ball has a non-zero initial horizontal velocity.	59	13
…is the same in both experimental settings.	32	61

Around 60 % of the students changed their initial hypothesis in face of the data provided: 35 % changed from one of the incorrect hypotheses to the correct one, 7 % changed from the correct hypothesis to an incorrect hypothesis, and 19 % changed from one incorrect hypothesis to the other incorrect hypothesis. From the 40 % of the students who kept their initial hypothesis, 26 % kept the correct initial hypothesis and 14 % kept one of the incorrect initial hypotheses.

The influence of the number of decimal places on changing or keeping the initial hypothesis
To answer the research question, we analyzed students' decisions to keep or change their initial hypothesis depending on their assignment to the three groups in which different numbers of decimal places in the datasets were provided (see Figure 2). A qualitative view shows that increasing the numbers of decimal places in the datasets results in a decrease of the percentage of students a) keeping a correct hypothesis and b) changing from one incorrect hypotheses to the correct hypothesis. Further, increasing the numbers of decimal places in the datasets results in an increase of the percentage of students changing from the correct hypothesis to one incorrect hypotheses.
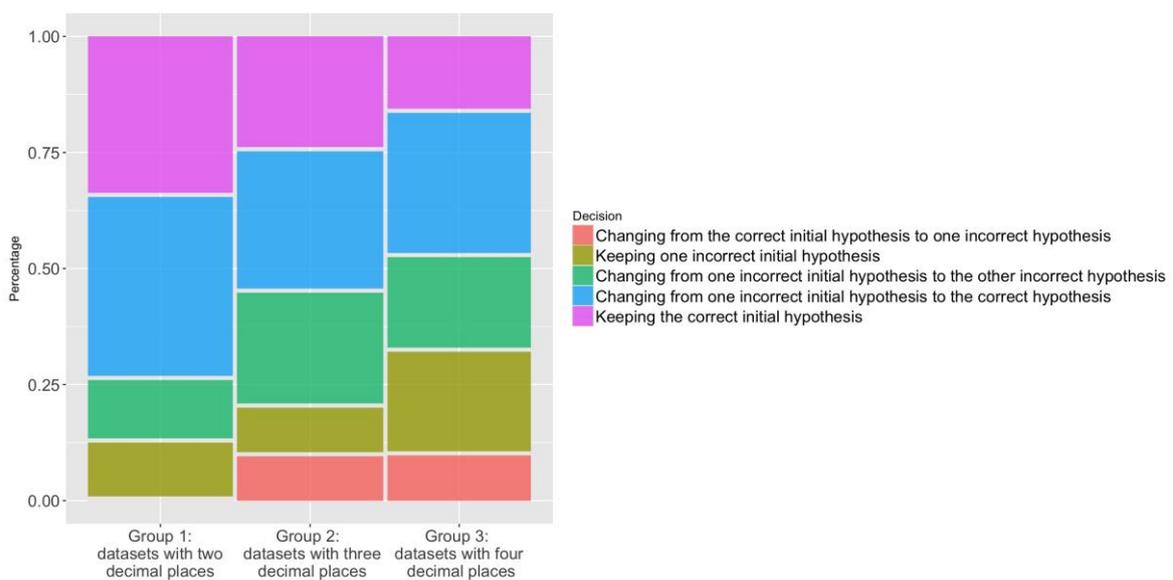

Figure 2. Percentage of students who are keeping or changing their initial hypothesis in the three different groups with differing numbers of decimal places in the provided datasets.

The three groups differ significantly in the distribution of the percentage of participants falling into the five categories of changing or keeping the initial hypothesis ($X^2 = 28.13$, $df = 8$, $p < .05$). The same effect between the groups was found when only the two categories a) "changing from one incorrect initial hypothesis to the correct hypothesis" and b) "changing from the correct initial hypothesis to one incorrect hypothesis" are considered ($X^2 = 11.67$, $df = 2$, $p < .05$).

Limitations
As it is always the case, our study has limitations that restrict the generality of the results. We used only one experiment in physics, made students to choose between three different hypotheses only, operationalized the exactness of the measurement only with three different levels of numbers of decimal places, and collected our data only in one school in Germany with a pretty narrow target group (with respect to sample size and age of the participants).

DISCUSSION
The results of our study indicate that the number of decimal places in quantitative experimental data may hinder students' learning of physics content. Better measurement equipment - that leads to more exactness by means of more decimal places in the measurement results and hence stronger evidence - can lead some students to reject a correct hypothesis. How can this be explained? While most of the students were familiar with calculating the mean values of their two

datasets (we know that from students' protocols in which they analyzed both experimental settings), they knew no method to judge if the calculated difference of the two means is relevant. Hence, they made their decision whether to change or keep their initial hypothesis solely by comparing two numbers. If the numbers differ, students assume that there is a main effect neglecting statistical variance. This is in line with another study that we've conducted (Priemer & Hellwig, 2016) in which students measured temperatures inside foam cubes of different sizes and used only the readings of the thermometer with all its digits to investigate if there are differences between the temperatures. Further, some students seemed to be irritated and puzzled by the statistical variance in the given measurement data, a result that was found as well in Pfeiler, Priemer, and Upmeier zu Belzen (2015). In their study, a student said that he distrusts and discards his own data because the numbers in the dataset vary. Thus, our study provides more evidence that students lack basic statistical understanding (here the concept of variance) which in turn leads them to wrong scientific inferences.

Given this somewhat alarming result - empirical data from school experiments may hinder learning - it becomes clear that students need to acquire competences in judging the quality of data. That means for example, that students must be able to estimate the uncertainty when making measurements (Priemer & Hellwig, 2016). Measurement uncertainties may origin from the devices used and from statistical variance. Calculating the variance or standard deviation (which quantifies one component of the uncertainty of a measurement), relating the variance or standard deviation to the mean (which generates an uncertainty interval around the mean), and comparing two measurement results (each written as the mean value with an uncertainty interval) with respect to intersections of the uncertainty intervals are steps to teach data analysis to high school students. Without such a fundamental understanding of the character of measurements and basic statistics students will not be able to make valid inferences on the basis of data in any empirical field.

REFERENCES

Buffler, A., Allie, S., Lubben, F. & Campbell, B. (2001). The development of first year physics students' ideas about measurement in terms of point and set paradigms. *International Journal of Science Education, 23*(11), 1137-1156.

Chinn, C. A. & Malhotra, B. A. (2002), Epistemologically authentic inquiry in schools: A theoretical framework for evaluating inquiry tasks. *Science Education, 86*, 175-218.

diSessa, A. A. (1983). Phenomenology and the evolution of intuition. In D. Gentner and A. Stevens (Eds.), *Mental models* (pp. 15–33). Mahwah, NJ: Lawrence Erlbaum.

Kanari, Z., & Millar, R. (2004). Reasoning from data: How students collect and interpret data in science investigations. *Journal of Research in Science Teaching, 41*(7), 748-769.

Klahr, D., & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive Science, 12*(1), 1-48.

Lubben, F. & Millar, R. (1996). Children's ideas about the reliability of experimental data. *International Journal of Science Education, 18*(8), 955-968.

Millar, R., Le Maréchal, J., & Tiberghien, A. (1999). 'Mapping' the domain: Varieties of practical work. In J. Leach & A. Paulsen (Hrsg.), *Practical work in science education. Recent research studies* (pp. 33–59). Frederiksberg: Roskilde University Press; Kluwer Academic Publishers.

Osborne, J. (2014). Scientific practices and inquiry in the science classroom. In N. G. Lederman & S. K. Abell (Eds.), *Handbook of Research on Science Education*. Routledge.

Pfeiler, S., Priemer, B. & Upmeier zu Belzen, A. (2015). Der kritische Umgang mit Primär- und Sekundärdaten im naturwissenschaftlichen Unterricht. In S. Bernholt (Hrsg.), *Heterogenität und Diversität - Vielfalt der Voraussetzungen im naturwissenschaftlichen Unterricht* (Bd. 35, pp. 672-674). Kiel: IPN Kiel.

Priemer, B. & Hellwig, J. (2016). Learning about measurement uncertainties in secondary education: A model of the subject matter. *International Journal of Science and Mathematics Education*. doi:10.1007/s10763-016-9768-0.