

Students' ability to analyse empirical data in practical work

Pols, C. F.J.

DOI

[10.1088/1742-6596/1287/1/012001](https://doi.org/10.1088/1742-6596/1287/1/012001)

Publication date

2019

Document Version

Final published version

Published in

Journal of Physics: Conference Series

Citation (APA)

Pols, C. F. J. (2019). Students' ability to analyse empirical data in practical work. *Journal of Physics: Conference Series*, 1287(1), Article 012001. <https://doi.org/10.1088/1742-6596/1287/1/012001>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

PAPER • OPEN ACCESS

Students' ability to analyse empirical data in practical work

To cite this article: C F J Pols 2019 *J. Phys.: Conf. Ser.* **1287** 012001

View the [article online](#) for updates and enhancements.



IOP | ebooks™

Bringing you innovative digital publishing with leading voices to create your essential collection of books in STEM research.

Start exploring the [collection](#) - download the first chapter of every title for free.

Students' ability to analyse empirical data in practical work

C F J Pols

University of Technology, Delft, The Netherlands

abstract.

Although students in upper secondary education are often asked to carry out practical work independently to a large extent, it is questionable whether they already have the knowledge to do so successfully. This study looks into the data analysis skills students aged 15 should possess according to various curriculum documents and investigates whether this is the case. To do so, the 51 students involved carried out two practicals using worksheets with questions in which they ought to apply each of the ten identified data analysis skills at least once. The analysis of their work shows that students master the skill of visualising data in graphs. In applying each of the other skills, students make many mistakes or do not know what to do or how to proceed. Students do not have a preconceived plan but work mostly ad hoc. They are able to draw valid but superficial conclusions which could have been drawn without an extensive 50 minute practical. Their lack in data analysis skills may result in not attaining the learning goals set in other practicals. Students thus should be taught how to analyse empirical data, the practicals used in this study may serve as a basis for this.

1. Introduction

In physics education, practicals are frequently carried out. During these practicals, students often have to establish a quantitative relation between variables. The idea is that students discover or confirm a physical law or get a feeling for the theory [1]. Regardless of the precise aim of the practicals, a proper understanding of data and the ability to analyse empirical data is essential. This is illustrated by the PACKS-model in which the different types of knowledge are related to the choices made during the practical [2], see Figure 1. This model shows that the understanding of criteria for evaluating the quality of empirical evidence (category D) influences each single step in designing and carrying out a practical. For example, more relevant knowledge of data and empirical can lead to repeating measurements a number of times. During the collection of data, students may become aware that the chosen instruments do not provide the desired accuracy. Using their knowledge about data and knowing that only accurate and reliable data can lead to unambiguously determine and interpret patterns in the dataset [3], students may decide to use different instruments or devise a new method.

However, students understanding of data is limited and students encounter various problems when analysing data. This holds true for young students who have limited ideas of what counts as valid and reliable data or do not know how to treat anomalous data [4-7] but also accounts for first year University physics students who have difficulties e.g. with handling situations with non-covariation [8, 9]. Still, analysing and interpreting the data is often set as homework and addressed by each student on his own [10]. But if students do not know how to process data to establish a quantitative relationship, the learning goals of the practicals may not be attained. Students' lack of data-analysis



may thus partly explain the limited learning outcomes of practical work often reported in literature [11, 12].

Although research has been carried out focussing on the understanding of empirical data with young students up to age 14 and students in their first year of University, no studies are conducted with students entering upper secondary education at an average age of 15. However, the understanding and ability to manipulate empirical data seems especially important to this group as students entering a science based program in upper secondary education are expected to work more and more independently. They thus have to have a firm knowledge about data-analysis in order to successfully carry out practicals and draw relevant conclusions. Students who have not opted physics as a subject (physics is not mandatory in upper secondary education in the Netherlands) should have attained a basic knowledge of data and data analyse as this is a prerequisite to become a scientific literate individual [13]. The basic knowledge of data and data analysis has to be attained in lower secondary education.

Question addressed in this study is what students ought to know about data analysis and what we can expect of students entering upper secondary education at an age of 15 to do themselves. Can it be justified that they are allowed to independently draw conclusions once they have collected the data? These broad questions culminated in the following two research questions:

- 1 What might we expect of 15 year old students to do by themselves according to literature?
- 2 Which of these competencies do students indeed master?

2. Data analyse competencies in curriculum documents

Various curriculum document prescribe requirements of students ability to design scientific inquiries and carry out practical work, and thus describe what students ought to know about data analysis. Although the precise curriculum requirements will differ from country to country, we expect a general and common basis for the skills and knowledge that students must master at the age of 15. To establish that basis, the following curriculum documents have been used in which we determined the overlapping competencies:

- The Next Generation Science Standards [14, 15] which serves as guideline for curricula in many countries.
- The PISA framework [13] which describes the knowledge content to assess scientific literacy of students age 15 in Europe.
- The English science program key stage 3 [16] describing the subject specific competencies but also competencies which are related to working scientifically.

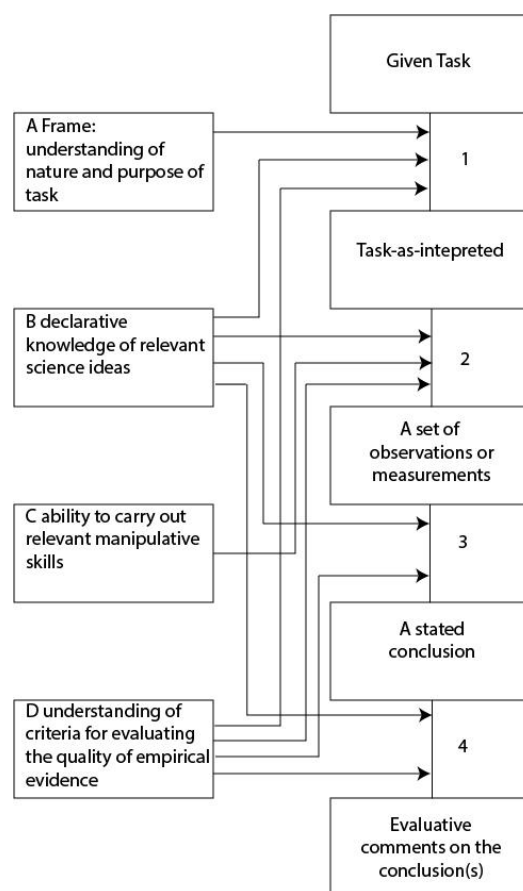


Figure 1: The PACKS-model [2] illustrates at which phases of a practical different types of knowledge influences the choices made.

- The Dutch lower secondary education curriculum [17] and an accompanied learning path to attain these requirements [18].
- The Dutch pre-University physics curriculum [19].

The overlapping ten competencies which could be found in these documents are shown in Table 1. These requirements are used in this study to determine to what extent students master each of competencies.

Table 1: The various curriculum documents seem to have a common basis in these 10 data analysis skills students aged 15 should possess.

1	Visualise data graphically
2	Use theoretical models to infer whether the origin constitutes a measurement
3	Draw a line of best fit through a data set and articulate that choice
4	Distinguish between a linear and non-linear relation
5	Qualitatively describe a dataset or trend
6	Describe qualitative similarities and differences between datasets
7	Draw a conclusion which is supported by the dataset
8	Indicate restrictions concerning the data-analysis and conclusions
9	Estimate the value of a point within the range of available data (interpolation)
10	Estimate the value of a point outside the range of available data (extrapolation)

3. Methods

In this study, two experiments were developed and used in two different classes in a regular secondary school in the Netherlands. In both practicals, the 51 students involved gathered data and answered various questions addressing the different analysis skills from Table 1. Students were familiar with the equipment used in these practicals and the practicals were easy to set up so that the attention could be focussed on the processing of the data rather than collecting these. During data collection, the teacher asked how they would approach the data analysis and what was necessary to determine a relation between the variables under investigation. The worksheets with students' answers were handed in by the students directly after the practical. These written answers and the audio recordings of the conversations with students were used as the main data sources.

3.1 Experiment 1: The pendulum motion of Spiderman

In the first experiment, students were shown a movie scene in which Spiderman uses his web to sling from A to B to catch a villain. The research question posed in this practical is whether this motion would happen in real as is displayed in the movie. Using the movie scene, students estimated the length of the web to be 30 m and timed half the period to be roughly 3,5 second. Although students could not determine the period of a real 30 m long pendulum themselves, they could determine the relation between the length of the pendulum and its period using a 1,0 m long cord and subsequently extrapolate to a length of 30 m. As an aid, students first had to use their dataset to predict the period of a 5,0 m long pendulum. Subsequently, the exact period was measured in the stairwell by the teacher. Students were allowed to use this additional measurement to adjust their first prediction.

To determine the exact time for such a pendulum it is first necessary to visualise the data (1), determine whether it is likely that the origin should be part of the dataset (2), draw a line of best fit by

weighing whether it should be straight or curved (3 &4). Students then had to use some kind of extrapolation (10) to determine the period and subsequently conclude whether the motion of Spiderman shown in the movie scene is realistic (7&8).

3.2 Experiment 2: The almost uniform motion of a marble on a horizontal track

In the second experiment, students released a marble on a wooden track consisting out of two parts. On the first part of the track, the marble accelerated due to a small slope. Using the same starting height each time, the marble rolled with the same velocity along the second, horizontal part of the track, see Figure 2. Students were asked to precisely describe the motion of the marble on the second part of the track. For that, they measured the time it took for the marble to role various distances along the horizontal track. They were then asked to compare their outcomes with the outcomes of others.

To successfully describe and compare the motion of the marble, students had to visualise the data again (1), determine whether the origin should be part of the dataset (2), draw a line of best fit (3&4), and subsequently describe the trend and compare the findings with others (5&6). In the worksheets they were asked to predict two values, one within (9) and one outside the measured range (10). Finally, they were asked to draw a conclusion related to the marble's motion.

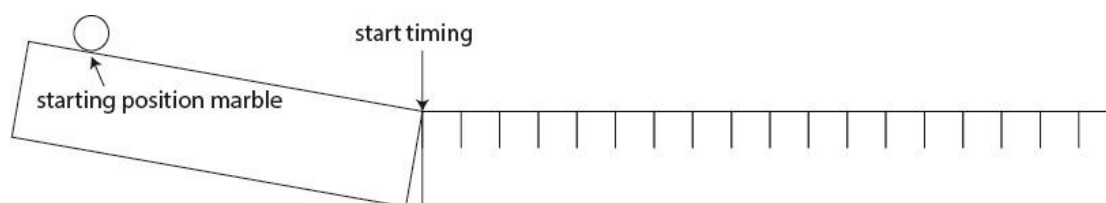


Figure 2: In the second practical students measured the time it took for the marble to run down a specific length of the track.

3.3 Data analysis

By investigating whether the students written answers met the pre-set requirements regarding each competency in each task and the completeness and correctness of their analysis, the extent to which each of the competencies are mastered by the students is determined.

4. Results

Before starting with describing the quality of applying each of the data analysis skills in the practicals (numbered below), it is worthy to mention that students hardly knew how they would process their data once collected. This appeared so from the recorded conversations the teacher held with the students during the data collection phase. During this phase, students had not thought about the processing of the data and how they would determine the relation between the two variables under investigation. Most students explained that they would draw a graph and then decide what to do. Some of them considered comparing the graph with the graphs given in the textbook. Students did not work using a clear preconceived plan or strategy but worked mostly ad hoc.

1 In both practicals, most students produced graphs that satisfy the scientific conventions [20], see **¡Error! No se encuentra el origen de la referencia.** for an example. Sometimes students interchanged the horizontal and vertical axes, where the scientific conventions demands that the independent variable is placed on the horizontal axes.

2 Regarding the second competency, half the students chose to take the origin into account using the pattern shown in the graph. The other half decided whether to take the origin into account based upon

theoretical grounds. Theoretical grounds could especially be used in the second practical where it is evident that the ball has not travelled any distance when $\Delta t = 0$ s. The origin should thus be part of the dataset. Still, many of the students did not use this line of reasoning.

3&4 Following from both practicals in which students drew a line of best fit, students seem to be able to draw a line of best fit when they have access to accurate data and more data points. **¡Error! No se encuentra el origen de la referencia.** illustrates this. Students correctly distinguished between a linear and a non-linear relation for the first data set up to 1,0 m. However, they connected their first data set with the later received measurement for a pendulum length of 5,0 m using a straight line. Only three out of 25 groups expressed their doubt about connecting the first data set with the later received measurement using a straight line. They asked the teacher for help. However, he did not help them by giving the right answer but asked them how they could find out. They suggested that an additional measurement would help them out and subsequently measured the period for a pendulum with a length of 2,5 m and decided the line of best fit should be curved. Next to curved and straight lines some students connected the measurements with each other (connecting the dots) or drew a smoothly curved line through the dataset. This might indicate that some students do not know what is meant by a line of best fit.

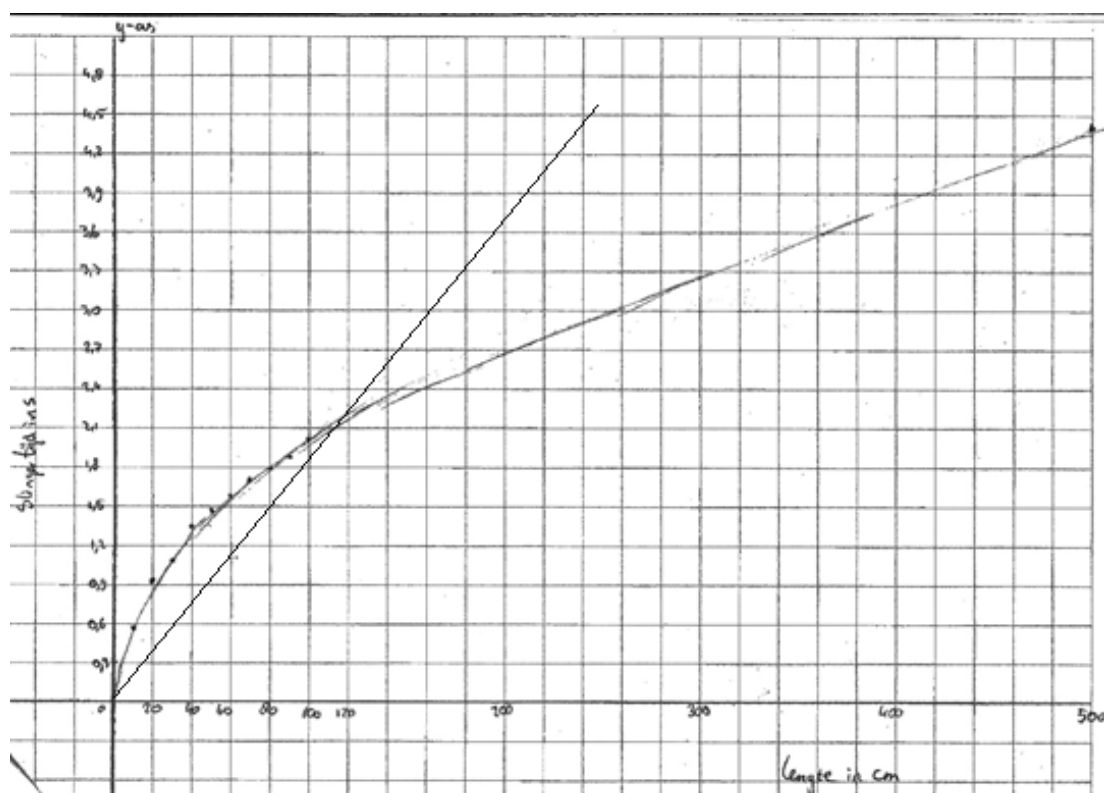


Figure 3: Students were able to represent their data in a graph that satisfies the scientific conventions but experienced many difficulties when drawing a correct line of best fit or predict values using extrapolation.

5&6 In this study it was difficult to see whether students could give a clear description of graphs as their description was very concise. However, essential features were often missing in their description of the data and graphs. An illustrative example of a description of a graph given by students during the second practical: *The graph runs upwards, at the beginning an increasing rise and then a part*

more or less constant. Distance travelled on the horizontal axis and time on the vertical axis. Without seeing the graph, it is impossible to explain what is meant by *a part more or less constant*. Do the students mean a steady horizontal line or a steady increase? This example also illustrates that students describe the graph and not what it represents. In the given description students could easily refer to a (non) uniform motion of the marble.

Comparing data sets is not done thoroughly by students. Their explanation is even less descriptive and sometimes even cryptic: *Similarities, the line is rising. Differences, the steepness and speed of the line differs in the graph.* Although the speed of the line may seem a translation error here, it is not. We cannot give any other translation here because the students truly refer to the speed of the line.

7&8 Students were expected to draw conclusions related to the practicals, for instance that the almost straight pattern in the graph indicates that the marble was rolling with a near constant speed of 0,14 m/s down the horizontal track. However, students often drew only superficial conclusions which were foremost qualitative of nature without any links to the data obtained. In some cases, the conclusion could be drawn without such a practical. One group merely concluded that *the longer the ball travels, the further it comes*. Such a conclusion serves no practical or scientific purpose, is therefore unimportant and is not reliable as no link between data and claim has been made (although it is questionable whether this is necessary for such a superficial conclusion).

Of all conclusions only one group was able to draw a conclusion that consist of most elements that make up a good conclusions. This group concluded about the motion of Spiderman: *Quicker than would happen in reality, assuming that the length is 30 m and that we did not make any mistakes in our measurements and calculations.* Although this conclusion is very brief, they mention their certainty of their claim using an assumption. In general the conclusions drawn in the first practical were better than those in the second practical although they were still very limited in reasoning.

9&10 Most students were able to predict a value outside the measured range in the second practical by doubling one of the known values which gave them a neat estimate. However, none of the groups mentioned any restriction to this prediction or looked at more detail to the somewhat curved pattern indicating that friction plays a role. As friction plays a role, the velocity decreases and one can not double time to predict values. Or at least, one should state their assumption.

Predicting the period of a 30 m long pendulum was more difficult to students. Some students indicated in a previous question related to the observed pattern that the graph was curved. Still, these same students used a linear relation to predict the value. This is illustrated in Figure 3, where the straight line is drawn by us to demonstrate what is mathematically done by the students. Other groups drew a straight line through their data set, excluded the origin but subsequently used a direct proportional relationship to calculate the value.

Although predicting and subsequently measuring the period of a 5,0 m long pendulum increased the accuracy of the prediction for the period of a 30 m long pendulum, only one group was able to precisely predict the right value but could not explain why their reasoning was valid. They reasoned that 0,1 m and 1,0 m has the same ratio as 0,5 and 5,0 m and the period probably would have the same ratio. Following this line of reasoning they could determine the period of a 30 m long pendulum.

5. Conclusions and implications

Although the various curriculum documents prescribe what competencies students at an age of 15 should possess, this small-scale study shows that these students are hardly able to meet any of those requirements. This account for students who opted for physics as a major subject in upper secondary education and are thus likely to be good at physics, while those requirements also apply to students who opt for a non-science-based program. We can only imagine that their score on this test would

even be worse. This thus infers that it is very unlikely that the requirements to become a scientifically literate person are met by these students.

Based upon these outcomes, it can not be justified that students carry out experiments in which they analyse empirical data and draw conclusions out of these without the teacher's help. This also means that, if we want students to carry out practical work independently at this age, they first have to be trained in data analysis. The compiled list of competencies can be an aid for teachers in terms of what has to be taught. The practicals in this study are used as an assessment tool but might be used for training purposes as well. Although not studied here, we see many opportunities to use these practicals to teach about pattern recognition and drawing conclusions. During the analysis of the data, the teacher can pose various questions as 'how sure are you of your findings?' or 'are other explanations, or relations between the variables possible?'. The teacher should provide continuous feedback and increase his/her influence on the outcomes. In small groups, students might as well compare their conclusions and argue which of the conclusions is best. In this way they learn what constitutes a good conclusion. The practicals might be carried out as activities on their own but can be embedded in various stages of the existing curriculum as well.

6. References

1. Tamir, P., *Practical work in school science: an analysis of current practice*. Practical science, 1991: p. 13-20.
2. Millar, R., et al., *Investigating in the school science laboratory: conceptual and procedural knowledge and their influence on performance*. Research Papers in Education, 1994. **9**(2): p. 207-248.
3. Gott, R. and S. Duggan, *Practical work: its role in the understanding of evidence in science*. International Journal of Science Education, 1996. **18**(7): p. 791-806.
4. Lubben, F. and R. Millar, *Children's ideas about the reliability of experimental data*. International Journal of Science Education, 1996. **18**(8): p. 955-968.
5. Kanari, Z. and R. Millar, *Reasoning from data: How students collect and interpret data in science investigations*. Journal of Research in Science Teaching, 2004. **41**(7): p. 748-769.
6. Wellington, J., *Practical work in school science: which way now?* 2002: Routledge.
7. Kok, K., et al., *Students' conclusions from measurement data: The more decimal places, the better?* Phys. Rev. Phys. Educ. Res., 2019. **15**(1).
8. Lubben, F., et al., *Point and set reasoning in practical science measurement by entering university freshmen*. Science Education, 2001. **85**(4): p. 311-327.
9. Allie, S., et al., *First-year physics students' perceptions of the quality of experimental measurements*. International Journal of Science Education, 1998. **20**(4): p. 447-459.
10. Barton, R., *'IT' in practical work*, in *Practical work in school science: Which way now*, J. Wellington, Editor. 1998.
11. Hodson, D., *A critical look at practical work in school science*. School Science Review, 1990. **70**(256): p. 33-40.
12. Hofstein, A. and V.N. Lunetta, *The laboratory in science education: Foundations for the twenty-first century*. Science education, 2004. **88**(1): p. 28-54.
13. OECD, *PISA 2015: DRAFT SCIENCE FRAMEWORK*. 2013.
14. NRC, *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. 2012: National Academies Press.
15. NRC, *Next generation science standards: For states, by states*. 2013.
16. Department for education, *Science programmes of study: key stage 3. National curriculum in England*. 2013.
17. Ottevanger, W., et al., *Kennisbasis natuurwetenschappen en technologie voor de onderbouw vo: Een richtinggevend leerplankader*. 2014: SLO (nationaal expertisecentrum leerplanontwikkeling).

18. Spek, W. and M. Rodenboog, *Natuurwetenschappelijke vaardigheden onderbouw havo-vwo*. 2011: SLO, nationaal expertisecentrum leerplanontwikkeling.
19. Netherlands Institute for Curriculum Development, Retrieved from <http://international.slo.nl>, 2016.
20. Lachmayer, S., C. Nerdel, and H. Prechtel, *Modellierung kognitiver Fähigkeiten beim Umgang mit Diagrammen im naturwissenschaftlichen Unterricht (Modelling of cognitive abilities regarding the handling of graphs in science education)*. Zeitschrift für Didaktik der Naturwissenschaften, 2007. **13**: p. 161-180.