Introducing argumentation in inquiry—a combination of five exemplary activities

Pols, Freek; Dekkers, Peter; de Vries, Marc

DOI
10.1088/1361-6552/ab2ae5

Publication date
2019

Document Version
Accepted author manuscript

Published in
Physics Education

Citation (APA)
Pols, F., Dekkers, P., & de Vries, M. (2019). Introducing argumentation in inquiry—a combination of five exemplary activities. *Physics Education*, *54*(5), [055014]. https://doi.org/10.1088/1361-6552/ab2ae5

Introducing argumentation in inquiry – a combination of five exemplary activities

Freek Pols[1], Peter Dekkers[1] and Marc de Vries[1]

[1] Science Education and Communication, University of Technology Delft, Delft, The Netherlands

E-mail: c.f.j.pols@tudelft.nl

Received xxxxxx
Accepted for publication xxxxxx
Published xxxxxx

Abstract

Successfully carrying out a secondary school physics inquiry requires a considerable amount of procedural and content knowledge. It further requires knowledge of how and why maintaining scientific standards produces the best available answer to the given research question. To this purpose, a series of five inquiry activities was developed and tested in a single case study with students aged 14. The test shows that students indeed come to use a more scientific approach to inquiry tasks and understand why they should do so. We believe that this series of activities can serve as a starting point for more complex physics inquiries.

Keywords: Scientific Inquiry, Practical work

1. The problem of teaching inquiry skills

Inexperienced students often use inadequate procedures in scientific inquiry of, e.g., the pendulum. They frequently choose only two values for the length instead of a wide range, measure only once at each length instead of repeating and calculating averages, and draw a straight line through the data-pattern that (to us) clearly looks curved [1]. Textbooks often 'help' students so they merely have to fill in a table as instructed, calculate averages and square roots, and plot a graph that is *meant* to be straight. This often precludes their exploration of further assumptions about the pendulum, and many remain mystified as to why the square root was taken. Worse, however, is that if these issues are not addressed at an early stage they will re-emerge years later and cause further problems. Yet, explaining *why* procedures should be followed rarely helps. While students tend to comply and do as they are told, they stop doing so when we stop telling them to. Could it be that they fail to see the point of doing so if all we ask is: how does the period of a pendulum depend on its length? Can we expand the students' aim from answering the research question to finding the *best possible* answer, and *demonstrating* that it is? We present a series of five activities designed for this purpose and our experiences in a class of 21 students aged 14.

2. Activity 1 – Investigating what they know in the Pirates' Pendulum

During the making of a pirate film Captain Jack Sparrow and his mates are spectacularly swinging between ships of war, explosions going off and razor-sharp weapons flashing everywhere. Students need no convincing that the stunt coordinator must have a thorough understanding of the swinging, since Jack should arrive at a given spot immediately *after* the explosion, not during.

Students explore the physics of a pendulum to provide the stunt coordinator with the required information. Students identify factors they think influence the swing time and investigate these in small groups. The teacher monitors, asking supporting questions with the final discussion in mind:

- Can you explain what you are doing there? Why? What are you trying to find out?
- How do you carry out your measurements? What instruments do you use, and why?

- What will you report to the stunt coordinator? Why should he trust your results?
- What could you do to make your results even more trustworthy?

Students' actions and conclusions are as usual, a report to the fictitious coordinator is the only new element. Rather than on the findings, however, the final discussion focuses on the question: if you were a stunt(wo)man, knowing what information the stunt coordinator received, *would you jump*? This shows students quite directly why typical conclusions such as '*if the rope is longer the swing takes longer*' are unsatisfactory. As one student put it: '*my conclusion is of no* [expletive] *use to him!*' (authors' translation). Teacher feedback on the lab report, in our experience, rarely has this effect. Students appreciate that actual filmmaking depends on similar research impacting, e.g., the safety of stuntmen. They conclude that the stunt coordinator needs a report that is *convincing* (optimally informative, trustworthy and useful) and that theirs is not.

Millar, Lubben [2] regard inquiry as the implementation of 'procedural and conceptual knowledge in science' (PACKS). Their PACKS model builds on so-called Concepts of Evidence (CoE), '*certain ideas which underpin the collection, analysis and interpretation of data [that] have to be understood before we can handle scientific evidence effectively*' [3]. The concept at hand is called 'practicality of consequences'. While concepts of reliability and validity are still abstract and remote, our students *can* consider the costs of implementing their findings, as a step towards developing these targeted concepts. For this, activity 1 uses six design principles:

1. Students carry out their own inquiry. This provides a baseline on students' PACKS.
2. In the first activity they make the usual mistakes so that it can become a constructive 'bad example' – an episode that reinforces how *not* to address an issue [4].
3. Students experience the context as realistic and demanding of high quality answers.
4. Students take the roles of 'producers' *and* 'consumers' of knowledge. The context is suggestive of evaluation criteria such as useful, trustworthy, informative as characteristic of a *cogent* result.
5. Only basic knowledge and skills are needed, if the inquiry fails it does so in terms of the students' own criteria. They find out for themselves what is needed to do inquiry properly.
6. The activity is 'closed' in that all ought to draw the same conclusions concerning the purposes of inquiry and how to approach them. These conclusions are explicitly formulated as '*rules for doing proper investigations*' by the students in their own words at the end of each activity.

3. Activity 2 – Observation vs. inference with Tricky Tracks

Once students feel a *need* for cogent conclusions, developing a method for constructing and evaluating these is in order. We adapted 'Tricky Tracks' [5] for this purpose. Young students may regard an observation and its interpretation as one 'fact'. If the possibility of multiple interpretations of a single data set is non-existent, contesting its interpretation makes no sense and inferences need no justification. The claim 'is' the data. Our version of 'Tricky tracks' addresses this by asking students, in turn, to state what they observe in figure 1, but without repeating any previous statement. Soon, observations (e.g., '*the shapes are of two different sizes*') are mixed with inferences (e.g., '*the shapes are footprints*'). As all statements are displayed the teacher asks:

- Do you agree with all observations made so far? Why, or why not?
- Can we be sure that birds made these tracks? That they were present at the same time? What makes you think they fought/played/one flew away?
- If you could visit this place what would you do, or pay special attention to? Why?
- What would be a better term than 'observations' for statements we cannot agree upon?

Figure 1. Tricky tracks adopted from Lederman and Abd-El-Khalick [5] to teach the difference between observations and inferences.

Similar situations where a dataset has various acceptable interpretations are explained as common in science. But lacking a unique correct interpretation, we can still seek out and report the best ones available [6] and draw some tentative conclusions from our data, provided we specify *how we arrive at them* and *how certain we are*:

> *If* <u>*this is a pattern in loose dirt*</u> *it is* <u>likely</u> *that* ***it was recently produced by animals,*** **because this is what footprints look like. Since it consists** *of* <u>*two shapes that differ in size*</u>*, it is* <u>*likely*</u> *that* ***two animals produced it****. If both animals were present at the same time, we can conclude* **from the usual shape of feet** *that they* <u>must</u> ***have come together in the middle****.* <u>*There one set of tracks ends*</u>*. We can* <u>*firmly*</u> *conclude that* **this animal** ***did not leave the scene walking unless footprints were erased****. We may* <u>*speculate*</u>*:* ***is it still present, did it fly away, was it eaten or did it climb on the back of the other animal****?*

Generalising this account, a simplified version of Toulmin's [7] 'model of argumentation' (figure 2) provides a method for constructing a cogent conclusion; construct a ***claim*** (e.g., the answer to the research question), moderated by <u>qualifiers</u> and supported by **inferences** (i.e. warrants and backings) based on the <u>*data*</u>. These aspects of arguments have been highlighted similarly, with underlining, italics and bold, in the preceding section.

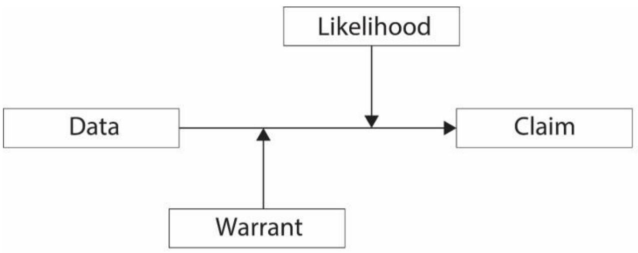

Figure 2. A reduced and simplified version of Toulmin's argumentation model is introduced to help structure conclusions.

Students practice the approach by analysing a short online article of the (highly respected) National Dutch Broadcasting Foundation that claims that conclusive evidence has finally been found of the existence of the *Abominable Snowman* or *Yeti*. Students identify the different aspects of the (exceedingly flimsy) argument and evaluate whether they find it convincing.

While students clearly came to distinguish observation from inference implementing this distinction and constructing cogent arguments was no simple matter, requiring further practice throughout the sequence.

4. Activity 3 – Establishing a relationship in advising the International Swimming League

Inquiry into relationships between variables is especially relevant in school science. In activity 3, students learn that relationships become more convincing if based on (1) more data collected from (2) a larger population, provided that (3) they are obtained through one and the same, appropriate procedure, in which (4) (human) error is avoided. Combining data sets (5) generally enhances trustworthiness, but (6) conclusions apply only to the researched population. The notion that (7) a conclusion is most convincing if it is optimally trustworthy, useful and informative is reinforced.

Reflecting on a 'newspaper article', students consider whether swimmers with relatively long arms have an unfair advantage, warranting the introduction of length classes in swimming. To start investigating the matter and advise the fictitious *International Swimming League* (ISL), students explore the relationship between human body length and arms' width. They measure each other in pairs, then share the data on the interactive whiteboard. A scatter graph gradually appears. They discuss:

- Were the first two data point enough to state a conclusion? Why, or why not?
- How reliable are our data, did everyone measure in the same way?
- What is the relation, if any, between arms' width and body length?
- How certain are we that this relationship really exists? How can we obtain more certainty?
- Is this relationship valid always and everywhere? How can we find out?
- If an additional data set is available should we combine them? What information do we need to decide?

An additional set of over 100 measurements (figure 3) is introduced. The class discusses how it affects the established relationship and previous answers. Next, in the role of ISL Chairperson, students discuss which of the following conclusions, appearing consecutively, is most satisfactory, and why:

1. Taller people have longer arms.
2. There is a relationship between body length and width.
3. Body length and width are directly proportional
4. For people of between 1,50 and 1,90 m in length, conclusion 3 is true.
5. Conclusion 4 is often true, but for one in three people this rule does not apply.

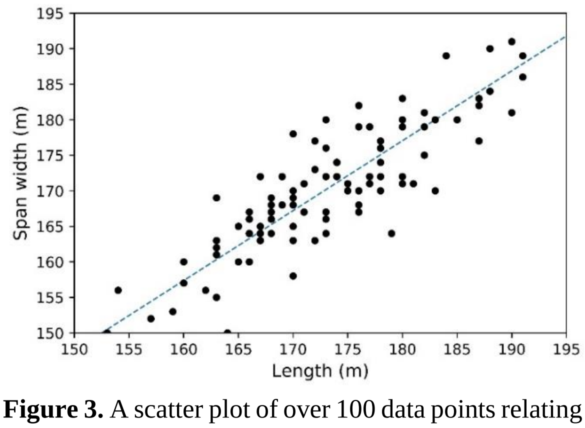

Figure 3. A scatter plot of over 100 data points relating human arms' width to body length.

Returning to their researcher roles, students then write a conclusion that is even better than these to the ISL, including also their personal view.

Students responded well, e.g., spontaneously discussing the fit with and meaning of the data pattern as data were still coming in. They identified limitations of the study and proposed appropriate expansions to take into account, e.g., a wider age range and other demographic characteristics.

5. Activity 4 – Data variability and the Fitch Barrier

For students without experience in inquiry repeating a measurement may seem pointless – if you measured correctly, why should it be different? This activity addresses understandings (1)-(7) again, but focuses on repeating measurements and verifying reproducibility. Students learn that variability in the measurements (8) is a natural, unavoidable characteristic that (9) if accounted for makes the conclusion more credible. Students also learn (10) how to deal with outlying data and discuss (11) how many repeats of a measurement suffice.

After his friend's terrible racing accident in 1955, John Fitch invented the Fitch Barrier [8] consisting of barrels filled with sand. A car crashing into these will decelerate, providing some protection for both the driver and spectators along the road. However if the car slows down too quickly the driver gets hurt – too slowly and the spectators remain unprotected. How many barrels are needed to decelerate the car just right?

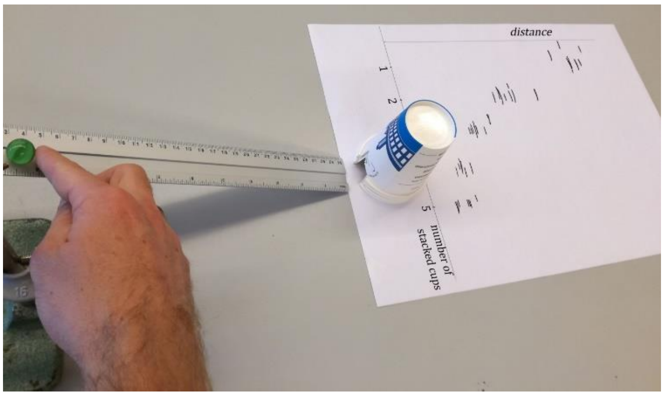

Figure 4. In this activity adopted from Farmer [9], a marble rolls into the stack of cups which then slide across the paper. A line is drawn at that spot. Repeating the procedure reveals variability in the measurements.

Students investigate this in an activity adapted from Farmer [9]. A marble rolling down an incline into stacked cups (with a hole in the side) will slide some distance and stop. The cups model the barrels, the marble represents a car. Possible extensions of the inquiry are easy to envisage. Students determine the relationship between the number of cups and their sliding distance. During the experiment the teacher monitors and asks questions like:

- Do you get the same results in repeated measurements? Why do the outcomes keep on changing? Does that mean you are not measuring properly? Can you reduce the spread?
- For a given number of cups, your results are roughly but not precisely the same. What is a trustworthy way to report this? How do you report the (real) value of the sliding distance?

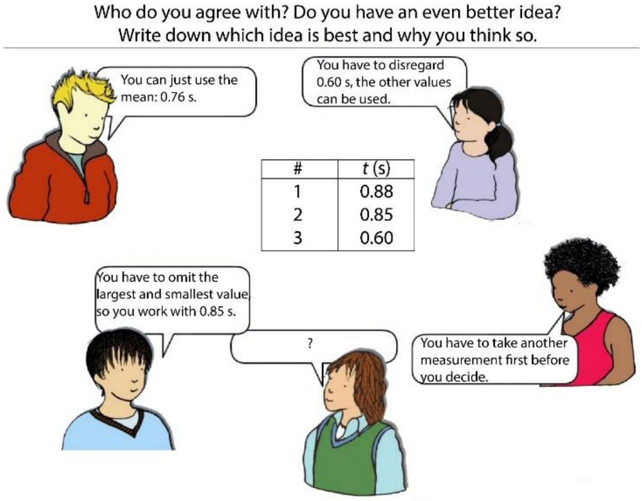

Figure 5. An adapted concept cartoon [10] is used to discuss the difficult concept of dealing with outliers.

Students became aware that no matter how well they tried to repeat the measurement, the stopping distances always varies, even though both it and its absolute variability become smaller with more cups (figure 4). Concept cartoons (figure 5) were discussed to decide how to deal with outlying data, and how many repeats of a measurement are needed to ensure a reliable value. Students also evaluated their inquiry by discussing issues of reporting reproducibility and data collection: they compared the results among the teams and reported similarities and differences. They discussed how trustworthiness increases if all data are reported as well as the method of collection. The teacher assisted in establishing the nonlinear relation by suggesting to students to look at changes when doubling the independent variable instead of using equal increments, something students are unlikely to figure out unaided.

Students noted without help that reporting the average measured values would not suffice here. Although the averages contribute to establishing the relation between mass and stopping distance, they argued that guaranteeing the safety of the people requires that the extreme measurements are also reported.

Activity 4 uses an additional design principle:

7. Developing Toulmin's model and understandings (1)-(11) is an explicit aim of learning in Activities 2-4. Understandings (1)-(11) invoke a range of the CoE

specified in Gott, Duggan [11] including, e.g., concepts such as 'datum', 'measurement', 'variability', '(in)dependent'. 'variable', 'repeat', 'reproduce', 'average', 'fair test', and 'line of best fit'.

6. Activity 5 – Practicing what was learned and NASA's Escape Pod

In Activity 5 students practice what was learned and consolidate their learning by reflection in helping NASA design a new escape pod for astronauts. The computer model designing the pod requires very accurate input, especially on factors influencing the frictional force. The pod is modelled as a paper cone (*e.g.*: Mooldijk and Savelsbergh [12], measurements involve its falling (figure 6). Potential factors are identified and allocated to research teams for further study: distance fallen, mass, diameter, top angle of the cone. During the investigation the teacher asks supporting questions about the cogency of students approach and the use of CoE's such as 'fair testing'.

Figure 6. Students drop paper cones with different frontal areas and measure the falling time using both a stopwatch and their mobile phone camera.

Since only adequate work is to be included in the final report, the teams evaluate each other's contributions to judge whether inclusion is warranted. Each team uses its own checklist of evaluation criteria drawn from the '*rules for doing proper investigations*' written up in the preceding activities.

Thus was explored whether the students apply the appropriate CoE adequately in their own work and can recognize this in the work of others.

The quality of students' work varied from two groups designing an Arduino-based electronic timing device eliminating response time, to a group who forgot to measure the cone's diameter in exploring the influence of its frontal area on falling time. Despite the variety in approaches, the vast majority of data collection procedures improved considerably to previous student practice. Students accounted for their choices in terms of the reliability of the data, showing understanding of the relation between research procedure and quality. The ability to analyse data and draw the most informative conclusion remained limited.

This final activity uses the following design principle:

8. Activity 5 is designed for students to consolidate previous learning as they engage in inquiry. They summarize and apply insights on how to do inquiry properly, and reflect on how they developed these insights. Students and teacher learn whether the intended understandings have been fully developed or require further clarification.

7. Conclusion

We wanted students to see why their usual conclusions in inquiry are unsatisfactory by scientific standards. Since students do not yet have these standards, they were asked to consider if their conclusion was good enough if their personal safety depended on it. They realised that it was not, as it was not optimally informative, trustworthy and useful.

Student then learned that a conclusion is in fact one interpretation of the research data while many tend to be possible. A conclusion in inquiry therefore should be an argument, consisting of a claim, the data, and the statements that link the two, providing support for the given claim.

Inquiry in science is directed at finding the best possible claim given the circumstances, i.e. the most informative, trustworthy and useful conclusion. In order to convince themselves and each other that a conclusion is the best available, scientists use a range of understandings. Eleven of these, all about optimizing the quality of the data and their interpretation, were developed in activities 3 and 4. Throughout the sequence, students drew up 'rules for doing proper inquiry'. In the final inquiry they used these to design and report an investigation of factors influencing air resistance and to evaluate the reports of others.

As expected, students did not develop straight away a high proficiency in applying Toulmin's model of argumentation or in applying the eleven understandings of evidence. They did, however, come to apply more appropriate data collection procedures, choose a wide range of many values for the independent variable, repeat measurements and calculate averages, take into account and report data variability, deliberately try to reduce or eliminate error, and consider various interpretations of any data set. Importantly, they clearly understood why they should do all of this. This, in our view, provides a useful starting point for more challenging kinds of scientific inquiry.

All teaching materials can be obtained via the first author. A more detailed manuscript of the theoretical background and design research on the learning outcomes is in preparation.

Acknowledgements

This work is part of a research programme for teachers financed by the Netherlands Organisation for Scientific Research (NWO).

References

[1] Gott, R. and S. Duggan, *Investigative Work in the Science Curriculum. Developing Science and Technology Education.* 1995.

[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] Kapur, M., *Productive failure*. Cognition and instruction, 2008. 26(3): p. 379-424.

[5] Lederman, N.G. and F. Abd-El-Khalick, *Avoiding de-natured science: Activities that promote understandings of the nature of science*, in *The nature of science in science education*. 1998, Springer. p. 83-126.

[6] Lipton, P., *Inference to the best explanation*. 2003: Routledge.

[7] Toulmin, S.E., *The uses of argument*. 2003: Cambridge university press.

[8] Fitch, J.C., *Energy absorbing deceleration barriers*. 1971, Google Patents.

[9] Farmer, S., *Real Graphs from Real Data: Experiencing the Concepts of Measurement and Uncertainty*. School Science Review, 2012. 346: p. 81-84.

[10] Naylor, S., Keogh, B., *Science Concept Cartoons*, 2008, Millgate House Education

[11] Gott, R., et al., *Research into understanding scientific evidence*. Website: http://www.dur.ac.uk/rosalyn.roberts/Evidence/cofev.htm, 2003.

[12] Mooldijk, A. and E. Savelsbergh, *An example of the integration of modeling into the curriculum: a falling cone*. Proceedings of the GIREP: Physics Teacher Education Beyond, 2000: p. 625-628.