Scientific Online

EDUCATION PRACTICE AND INNOVATION
ISSN(Print): 2372-3092 ISSN(Online): 2372-3106
In Press

EDUCATION PRACTICE AND INNOVATION

Critical Reading of "Making Sense of Confusion" by Jason E. Dowd, Ives Araujo, and Eric Mazur[1]

Valentin Voroshilov*

Physics Department, Boston University, 590 Commonwealth Ave, Boston, Massachusetts, 02215, USA

***Corresponding author**: valbu@bu.edu

Abstract:
The scientific method developed to study physical phenomena presents a proven instrument for conducting research in any other field of science. Yet, vast amount of literature on physics education research does not represent examples of application of that scientific method, even if the researchers are physicists. In this paper the author offers a critical reading of one of recent papers published by Physical Review Letters. The goal of this work is to stimulate a conversation on how the scientific method developed to study physical phenomena can be applied to study phenomena in realm of education.

Keywords:

Physics represents a perfect example of how the scientific method should be applied to study, well, everything.

At first a scientist observes, collects facts, develops vocabulary, classifies objects and processes, tests some preliminary ideas, but in the end the scientist formulates postulates (a.c.a. axioms, or laws). It usually is impossible to test the postulates by direct measurements; the consequences derived/predicted from the laws, however, can and should be tested, and while the experiments agree with the predictions, we believe in the correctness of the postulates, we keep using the theory. Of course, every theory has limits, hence when experiments contradict the theory, a scientists starts thinking, is it something wrong with the experiments, or the limits of the theory finally have been reached?

The Newtonian Mechanics, the Maxwell's theory of electromagnetic phenomena, the Einstein's theory of Special Relativity, the Einstein's General Theory Relativity, the Euclidian Geometry are some of the bests and clearest examples of such approach.

Can the same approach be applied outside of physics, say, to study learning and teaching phenomena? The answer to this question depends on a personal view.

In 2002 Richard Hake wrote[2]: (begin the quote) "There has been a long-standing debate over whether education research can or should be "scientific" (*e.g.*, **pro**: Dewey 1929, 1966, Anderson *et al*. 1998, Bunge 2000, Redish 1999, Mayer 2000, 2001, Phillips and Burbules 2000, Phillips 2000; **con**: Lincoln and Guba 1985, Schön 1995, Eisner 1997, Lagemann 2000). In my opinion, substantive education research must be "scientific" in the sense indicated below. My biased prediction (Hake 2000*a*) is that, for physics education research, and possibly even education research generally: (a) the bloody "paradigm wars" (Gage 1989) of education research will have ceased *by the year 2009* (italic by Valentin Voroshilov), with, in Gage's words, a "productive rapprochement of the paradigms," (b) some will follow paths of pragmatism or Popper's "piecemeal social engineering" to this paradigm peace, as suggested by

Gage, but (c) most will enter onto this ”sunlit plain” from the path marked ”scientific method” as practiced by most research scientists” (end of quote).

Thirteen years later this prediction looks overly optimistic.

In many papers, even written by scientists who have been using the scientific method in their field, the authors do not seem applying the same way of reasoning when writing a paper on education. At the least, that indicates the fact that the authors do not believe that the same scientific method applied to study physics (chemistry, mathematics) *should* be applied to study education. At the most, that indicates the fact that the authors do not believe that the same scientific method applied to study physics (chemistry, mathematics) *can* be applied to study education.

For example, let us read the latest publication by Eric Mazur[1] and his colleagues.

The main statement I want to make after reading the article is that the methodology (which we call “a scientific method”) which had been developed and being used to study physical phenomena *can* and *should* be used for conducting research like the one described in the paper, but the paper does not show the use of that methodology.

Below I will try to support this statement by analyzing the study described in the paper. Clearly, my analysis of the study is based on certain assumptions I made during the reading.

The first assumption is that one of the goals of the study was to find a correlation between: (a) the fact that students are offered to answer questions designed to generate confusion, and assess how confused they are: and (b) learning outcomes. This assumption is based, in part, on the statement: “We ask the following question: To what extent are course performance, . . . related to confusion?”.

I argue, that if one wants to study such a correlation, one can (and should) use the same methodology which had been developed and being used to study physical phenomena. In the latter approach, one has to compare two (at least) study cases: “Case 1” is when students do not have to answer questions designed to generate confusion and do not have to assess how confused they are; “Case 2” when students have to answer questions designed to generate confusion, and have to assess how confused they are (the “confusion” element becomes a part of a learning experience). The scientific method also demands that the “Cases” should not be different from each other by anything else but the “confusion” element, which means: student body in both “Cases” should be similar by the number of students, by the age, race, background distribution (for large classes it is reasonable to assume that these conditions are satisfied), students’ course work should be very similar (except the “confusion” element), faculty involvement should be similar, learning outcomes should be measured by the same measuring. If these conditions are not held the learning outcomes of student might be affected by many uncontrolled parameters and the examined correlation cannot be established.

While reading the paper, however, one cannot find any indication on how the introduction of the “confusion” element to some students influenced their learning outcomes, compared with students for whom the “confusion” element has not been a part of the learning process. It is not clear either the authors did not use the scientific method on purpose, or used it but the paper does not provide a clear description of doing that (the further analysis indicates that the former is more probable than the latter). The presence of this ambiguity in the description of the study makes the study scientifically deficient (I consider an ambiguity of a scientific study as a deficiency). Many similar studies experience a similar deficiency. It might have helped for a reader to navigate through a paper if at the beginning of the paper the authors would clearly state if they meant using the scientific method (the one developed to study physics), or they did not mean to use the scientific method on purpose.

Another deficiency of the paper (as well as many other similar publications) is the fact that the use of the scientific method would have eliminated the need for spending time and effort on collecting data which, when scrutinized, do not really support or contradict the hypothesis of a study. Instead, the conclusions of a study could have been derived from a set of well-established facts, a.c.a “postulates” or laws.

Below I provide several illustrations to the statement made above.

Setting the terminology aside, the introduction of the paper tells us that: (a) sometimes students get confused (and we know about that because students express their confusion in words or in actions); (b) students often have their own opinion on how good or how bad they can be when doing physics in general or when solving a specific problem; (c) helping students to reflect on their own thoughts, actions, and feelings may help them to perform better.

To this point we see a complete agreement with everyone’s teaching experience.

(a) Every teacher knows that students ask questions; what to do about it and how to manage each question (or how to initiate questioning from students who never talk) is a different conversation.

(b) The fact that different students may have different thoughts about themselves (in the variety of contexts) is also an everyday experience of every teacher (and again, we will not discuss in this paper what the best strategy is for a teacher teaching a class with students who have different self-perceptions).

(c) The correlation between “help” and “performance” can be derived from a more general principle (which is used use as one of the postulates of the Teachology: a practical science of teaching and learning), *i.e.* for most people (who do not have extraordinary deviations from average abilities) *learning outcomes are directly proportional to the volume and variety of learning experiences* (below, the “Postulate”).

For example, a teacher teaches a standard course (lectures, labs, discussions, homework). That leads to certain learning outcomes. If we accept the Postulate, we have to make a conclusion that, if the teacher will make students to do something else (reasonably related to the material) and do it on a regular bases over a long period of time, the teacher can expect learning outcomes to be better. In particular, making students (in addition to what they would have done before) watching movies, or reading additional texts, or discussing qualitative questions, or making them to reflect on what they read and how they felt will result in better learning outcomes.

One can compare any two teaching strategies by counting the amount of learning activities students will have to perform in each. If the material covered is similar by the topics and the volume, but the use of one strategy results in a visibly larger number of learning activities, that strategy will lead to better learning outcomes.

Ballet trainers, sport coaches, parents use this “rule” every day; people say: “practice makes perfect”, and that works every time as long as the practice provides a sufficient volume and variety of learning experiences.

A question like: “Will it affect learning outcomes if in addition to what students have done in the past they will be forced to do such and such?” does not always represent a research question. If “such and such” is related to the learning material, learning outcomes will be better. If learning outcomes did no improve, hence using “such and such” was the wrong choice, or “such and such” has not been used for a long enough time. The question a teacher should ask is “how can I make students to do “such and such” in addition to what they already do?” This question, however, is not a research question; this is a practical (*i.e.* social by its nature) question. Of course, the teacher assumes that the additional learning experience (“such and such”) will lead to better learning outcomes. But this assumption is an assertion (“I believe in the Postulate”) and not a scientific hypothesis, even if it looks like such (like, the assumption that “if I take this root I arrive home faster” is not a scientific hypothesis). I do not think that any possible question should be called a hypothesis, and any possible activity which leads to an answer should be called a research (my essay[2] provides a broader discussion of the difference between a scientific research and a social project).

A research question could have been stated in the following form: “Will learning outcomes improve if we keep the amount of learning activities and the total time of learning practically the same as in the previous course, but rearrange some activities or replace them with different ones?" Unfortunately, as it has been mentioned above, in many papers, including the one under the discussion, there is no available information, which would allow readers to see the specific procedural (technical) differences between the new and the previous learning processes.

Reading the article, however, indicates that in the study described in the paper students - in addition to their regular learning process - had been doing something else: "students were assigned 22 ... and 21 reading exercises", "in each assignment, the confusion question was posed before the two content-related questions, followed by a final opportunity to revise the response...". The statement that "at least two – and sometimes as many as three or four – researches and instructors reviewed and discussed each content-related question ..." also shows that during this particular teaching process students have been treated differently then students not participating in the study (the content-related questions were developed using a higher level of involvement of developers).

Based on what I read, I made a second assumption, namely, that courses taught during the study described in the paper were different from the courses taught before the study by the use of the "confusion" element. Based on this assumption and on the Postulate stated in part (c) I made a conclusion, that the results of the study should be obvious (*i.e.* should support the Postulate, or the design of the study should be reexamined). If we accept the Postulate, we should expect that the additional practice will be "positively related to a final grade". In a sense, this study supports the effectiveness of the Postulate (like a working clock supports the effectiveness of the Newton's laws).

Next I would like to address briefly one specific statement from the introduction.

"One cannot express confusion without engaging in metacognition, which involves knowledge and cognition about cognitive phenomena".

The purpose of this statement is to begin a discussion about metacognition. It is naturally to expect, however, that a student who can explain reasons for his or her actions "will be more strategic and effective in the educational setting" than a student who can just act without being able to explain why did he or she act the way he or she did. This conclusion is a straightforward consequence from the Postulate formulated above in part (c). An ability to explain the reasons for his or her actions does not come with a birth; it requires a specific type of practice and, of course, a designated time. Hence two students – one who can and another one who cannot explain the reasons for his or her actions – must differ by the volume and variety of learning experiences. No surprise that every research "consistently suggests that enhanced metacognition is positively related to learning outcomes". In the end students' results had been positive, which agrees with the following quote from the paper: "Specifically within physics, researches observe that adding metacognitive tasks to reading-comprehension exercises results in higher post-test scores when compared to a group of subjects who do not complete the metacognitive tasks". This is an example of a statement which often sounds like: "We divided students into two groups, in one group students were instructed to learn "that", in another group students were not instructed to learn "that", the result is, students in the first group learned "that", and students in the second group did not ".

The statement itself, however, is wrong; one can and very often expresses confusion without engaging in metacognition; expressing confusion in many cases is just an emotional reaction to inability to understand something which a person feels like to be expected from him or her to be understood. Every human being might experience many different states, like hunger, tiredness, angriness, confusion. Saying "I am confused" is no different from saying "I am tired", "I am angry", *etc*. It does not require any metacognition. Although, one could redefine "metacognition" by including in it any statement people make about themselves, however it would water down the sole meaning of this term and would make it useless.

The discussion regarding the effect confusion might have on students' outcomes leads, basically to a conclusion that sometimes confusions is good and sometimes is not.

Every experienced teacher, of course, will agree with this conclusion. However, the mere fact of expressing confusion should not lead to a large change in learning outcomes because it does not involve any additional mental work. The outcome depends on what work has been done to reduce that confusion. An interesting research question is what type of work (step by step guiding, giving away an answer, initiating peer-to-peer conversation, *etc*.) and under what circumstances would be the most efficient way to decrease or “eliminate” that particular confusion.

The technical realization of the study has been described very clearly and can be used by any instructor who would like to use for his or her purpose qualitative indicators of confusion and confidence.

References