swiss physical society

Nr. 58
Juli 2019

SPG MITTEILUNGEN
COMMUNICATIONS DE LA SSP

# AUSZUG - EXTRAIT

## Progress in Physics (68)

### Physics Education Research - An Applied Science (part 1)

*Andreas Müller, Université de Genève, Faculté des Sciences / Section de Physique*
*and Institut Universitaire de Formation des Enseignants, 1211 Genève*

This article has been downloaded from:
https://www.sps.ch/fileadmin/articles-pdf/2019/Mitteilungen_Progress_68.pdf

© see https://www.sps.ch/bottom_menu/impressum/

# Progress in Physics (68)

## Physics Education Research - An Applied Science (part 1)

*Andreas Müller, Université de Genève, Faculté des Sciences / Section de Physique and Institut Universitaire de Formation des Enseignants, 1211 Genève*

### 1 Introduction: Research and Development in Physics Education

Physics education research (PER) is a broad term covering a whole range of research and development: The former means research e.g. about how learners think, difficulties for learning, how well a given approach succeeds, and by what factors success (or failure) might be influenced [1]. The latter means development of new and innovative teaching and learning ideas and materials (such as simplified approaches to difficult topics, new experiments, use of information and communication technology (ICT), etc.

In the following, I will illustrate the branches "research" and "development" by some recent outcomes [1]. Obviously, a connection between them in the form of research-based development is desirable, which I will discuss. I will also cover some other aspects such as the importance of PER for teacher education, or its service functions for physics departments and institutions in the second part of this article.

### 2 Research – Why not try a scientific approach to physics education?

Learning is a dynamical process, driven by relevant forces and interactions, and leading from an initial to a final state; in that sense, a current metaphor is that of a "learning trajectory" [2, 3]. As much as physics itself, physics education research tries to understand such trajectories, and design them to attain a given target. The same holds, if one tries to increase the physics interest of students. In the last decades, physics (science) education is increasingly inspired by approaches of science itself, as promoted e.g. an article by C. Wieman “Why not try a scientific approach to science education?” ([4], providing the title of the section).

As much as in physics, physics education research then needs measurement methods for the characteristics of the initial and final (intended) state, and an understanding of the intervening "dynamics", i.e. teaching interventions and other, maybe uncontrolled and unwanted, influences. In the following, some illustrative examples of PER for these aspects will be given, sometimes also referring to important results from other areas of STEM (Science, Technology, Engineering, Mathematics) education [2].

### 2.1 Measurement

> *Figuring out what to measure, and how well to measure it, is critical in all fields. […] Although the specifics for how to do this are different between physics and education, the basic methods are much the same.*
> ***C. Wieman [5]***

1 The present paper is based on a related text about teacher education (arXiv preprint: 1807.00974 [physics.ed-ph]) and on numerous fruitful discussions with many collegues (see acknowledgements).

2 Throughout the paper, it is understood that PER is strongly linked to other parts of STEM education research, i.e. many of the statements hold also for other areas of science learning, or are based on findings in these.

There is an increasing awareness for the need of reliable measurement in physics education, and the development of methods for diagnosis, measurement and assessment has received sustained attention over the past two decades [5 - 9]. In particular, teachers have to know, whether a given approach really increases understanding or motivation, and they have to be able to test this in their own classrooms.

As an example of particular importance, we discuss conceptual learning, considered as fundamental for the scientific literacy of the future citizen [10], as well as for the more advanced competences of a scientist, such as calculation, experimentation, and problem solving ([11], part II): there are no reliable empirical or theoretical results, nor convincing ideas or innovations without understanding of what you are talking about. Moreover, it is a striking strength of physics that it can explain a large range of phenomena by a small set of basic concepts [12, 13]. However, science education research has provided ample evidence for many conceptual obstacles, or "misconceptions" [1], such as the idea that a moving object "looses force" or that the seasons are caused by different distances to the sun (see [14] for some further illustrative examples). Such ideas are found to be widespread and hard to change, even among university students (see e.g. [15, 16, 17]).

Figure 1 presents three classical examples: The first shows that a majority of learners think even at the beginning of university that space between particles of a gas is not empty (there is vapor or oxygen: ≈ 40%, pollutant: ≈ 20%, correct answer "nothing": 40%); this is reminiscent of the "horror vacui" of Aristotelian and medieval thinking ([23], ch. 4), and in

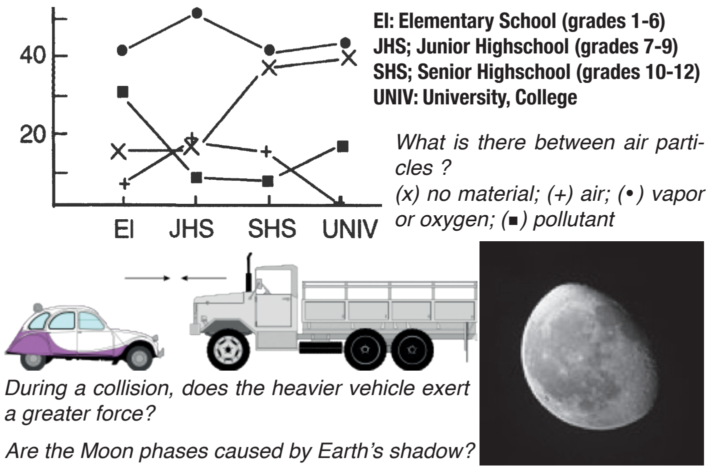



*Figure 1: Some examples of misconceptions in physics and astronomy*
*a) microscopic structure of matter [18]*
*b) Newton's third law; % correct: end of Gymnasium 24% [19]; begin / end of introductory physics course 34% / 62% [20] / [21]*
*c) origin of moon's phases, % correct: begin / end of introductory college astronomy (without addressing the misconception explicitly) 12% / 32% [22]*
*https://upload.wikimedia.org/wikipedia/commons/7/7d/2013-01-02_00-00-55-Waning-gibbous-moon.jpg*

fact this "parallelism" between individual and historical conceptual obstacles has been found also for other topics [24]. The second concerns a wide-spread difficulty with Newton's 3rd law, i.e. that in a collision the heavier vehicle (objects) also exerts more force. It is related to an unclear distinction of force and momentum (and maybe kinetic energy), again something that has also taken a long path in the history of physics [25]. Third, the misconception that the lunar phases are caused by Earth's shadow. The fact that it persists even when in contradiction with easy-to-make observations (a gibbous moon with its concave dark region, a crescent moon next to the sun) is an example of what is called the 'persistence' of misconceptions.

In order to measure the presence of a given set of misconceptions ("initial state"), any improvement of conceptual understanding ("final state"), and to compare the impact of various classroom interventions ("dynamics") the physics and science education research community has developed a series of concept tests, e.g. about the structure of matter, Newtonian mechanics, basic astronomy, and other fields [26]. These tests (also named "instruments" or "inventories") have been thoroughly discussed and improved (see e.g. [27] for mechanics), characterized according to standard psychometric characteristics [28], and analysed with more advanced methods [19]. Today comprehensive collections and reviews about concept tests are available for practitioners and researchers [26]. Another important element of physical reasoning known to be a major difficulty for learners are "multiple representations", i.e. the fact that understanding the link between scientific phenomena and their conceptual basis requires the learner to deal with multiple representations at different levels of abstraction. These levels comprise e.g. a verbal description of a geometrical optics experiment, a photograph of it, a schematic description through ray diagrams, and a formal description by the magnification equation. All these representations are necessary to achieve a proper understanding of an image formation process. As Kohl et al. put it "*good use of multiple representations is considered key to learning physics*", yet there is ample evidence that this "*good use*" is difficult even for learners up to university level [27, 29, 30 ch. 6, ch. 8]. Much as in the case of conceptual understanding, measurement of learning to use multiple representations is necessary, and various tests are available [31, 32]. Test for numerous other aspects of physics education exist, such as problem solving [33, 34 ch. 5] use of math [35] and other competencies [8, 36], and also for affective aspects (interest!) and attitudes [37].

### 2.2 Interventions

> *[R]eform in science education should be founded on "scientific teaching", in which teaching is approached with the same rigor as science at its best.*
> ***J. Handelsman et al. (2004), Science, 304, 521 [38]***

Of course, one does not only want to *measure* educational outcomes, but also to *improve* them. I present two examples of effective interventions in physics education. The first is about low physics interest at school, especially among girls and how to improve it, a problem teachers know very well and are facing in their daily teaching. One wide-spread approach to counter this is context based science education (CBSE), i.e. "using concepts and process skills in real-life contexts that are relevant to students from diverse backgrounds" [39]. Making science issues relevant to students and their everyday life can counter the wide-spread perception of physics as being dry, impersonal and irrelevant, and this is supposed to have positive effects on motivation and learning [40, ch. 19.4.3]. However, look at Figure 3: it shows a striking "scissor"-form, with perceived relevance of physics increasing over the years, yet interest showing a marked decrease. Thus, just taking account of physics contexts in the sense of "making it relevant" [41] is *not* sufficient to maintain (or generate) pupil's interest [3]. Moreover, this decrease shows a strong gender bias: At the end of 5th grade, physics lessons are perceived as interesting or very interesting by about 40% of the girls and 60% of the boys, at the end of grade 10, it is about 20% and still 60%, respectively. Girls thus most likely will avoid physics in favour of biology, for their educational and career choices. We observe the consequences in the matura tracks of the gymnasium, and our physics classes at university: a gender ratio ($N_{♀} / N_{♂}$) of 0.2 – 0.3 [43], and the decisive step for this does already happen at the end of secondary level one [4].

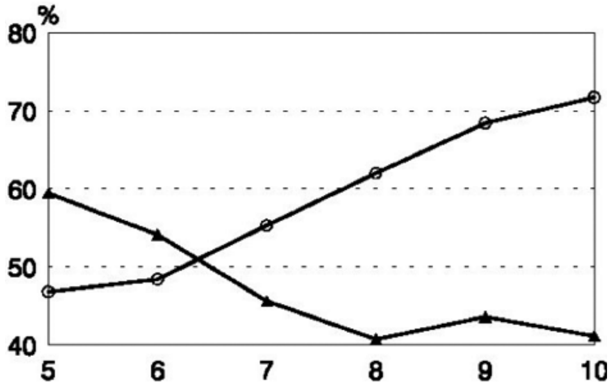



*Figure 2: Development of interest (▲) for and of perceived relevance (○) of physics from begin to end of secondary level I in Germany (x-axis: grade; y-axis: % of maximum value, [42]).*

What can be done? Figure 2 offers a revealing finding in that respect: When asking about the same physics topic related to different contexts, it turns out that girls interest for a biomedical context is much larger than that for a technical context; in fact it *increases* even against the general tendency during adolescence stated above (Figure 3; for boys the interest level for both contexts is similar and rather high (70%)). Note that context (kind of applications and activities) was shown to have a much stronger influence on physics interest than the specific content (subject matter): 80% of the variance of physics interest across the items of this study can be attributed to context, and only 20% by content. The Infobox "Context matters" presents concrete research-based examples of motivating contexts, leading to considerable effect sizes [5].

3 The interpretation is that in the period from the end of childhood (5th grade) to the young adults (10th grade), youths increasingly get aware of the large relevance of physics (science) for the world they live in, but that does not entail personal relevance and interest for themselves. This latter missing, one sees the a decline of interest, well-known from developmental psychology for this age group from almost all other school subjects as well.

4 Secondary level one is the school attendance period after primary till the end and of compulsory school, secondary level two the period after that (in the US system, this would be junior and senior high school, respectively).

5 The basic definition of an effect size (ES) is (MT – MC) / SD, where MT


## Context matters

Context (of physics teaching) matters, and considerable beneficial effects have indeed been found e.g. in an extensive project about the research-based development and evaluation of teaching sequences using biomedical contexts for secondary level II: effect sizes for interest development before and after the sequence are *ES* = +0.45 with context vs. *ES* = –0.52 without context; [47, 48]. By way of example, the Fig. on the left shows an excerpt form the unit on forces on the backbone.

Other empirically validated forms of contexts interesting for young people have been provided by research [49]. Interestingly, among all areas of science, astronomy topics are among the most interesting ones for young people, much more interesting than many conventional school topics [50]. Another example for secondary level I physics is learning with problems based on newspaper articles and the real-life contexts provided by them (Newspaper Story Problems). When comparing the treatment group with a control group (learning with conventional tasks, but otherwise the same content, lesson plan, and the same teacher) there is a significant improvement of motivation (*ES* = 0.85, up to 1.3 for various topics), and a rather sustainable one (considerable improvement still after almost 4 months) [51]. An example for elementary kinematics is shown below, tasks are on average velocities and their comparison. For learning effects by the same approach see below.

**Transatlantik-Weltrekord**

(*si/apa*) Der Amerikaner Steve Fossett und seine neunköpfige Mannschaft haben am Mittwoch einen Transatlantik-Weltrekord (von West nach Ost) für Segelboote aufgestellt. Mit einem 38-Meter-Katamaran legten sie die 5417 Kilometer zwischen New York und der Südwestküste Englands in 4 Tagen, 17 Stunden und 28 Minuten zurück. Der Millionär Fossett unterbot den Rekord des Franzosen Serge Madec aus dem Jahr 1990 (6 Tage, 13 Stunden und 3 Minuten) um mehr als 43 Stunden.

*Neue Züricher Zeitung, 11.10.2001*

In sum, context is decisive for the development of physics interest, often more than content, but it has to be empirically tested *which* contexts are really interesting for young people; being "relevant" for adults (educators, researchers) is not sufficient (see also 2.4).

The second example is about physics learning in lecture type courses. Lecturers of introductory university courses often face the problem of having hundreds of students in their

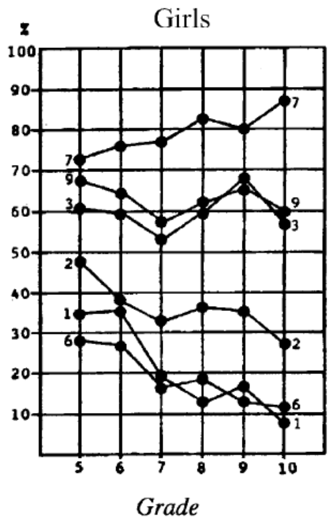



*Figure 3: Percentage of girls with "great" and "very great" interest in selected contexts for the topic of mechanics (motion, force, pressure; [44]). Curve 7: artificial heart as blood pump; curve 2: pumping petrol from great depths (see [44] for the other contexts)*

auditory, in particular in large enrolment universities or in physics minor classes. The classical solution for this is the frontal lecture, but there are doubts about the active intellectual engagement by students in this setting (look at the number of students using their smartphones in class, and this is only what you *see* of them not being involved). Are their more effective ways of conveying the basics of physics we consider as indispensable, and also e.g. to biology or medical students? Look at Figure 4, which shows the effects of a research-based teaching intervention in a introductory university course on electricity and magnetism. The teaching intervention introduces the following elements to enhance active intellectual participation of students [52]:

- preclass reading assignments + quizzes
- in-class clicker questions with student-student discussion
- targeted in-class instructor feedback (based on clicker data).

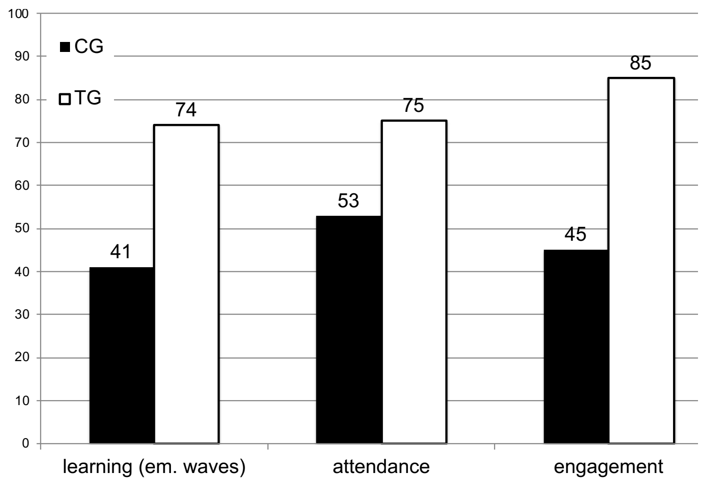



*Figure 4: Comparison of effects of two teaching approaches on a learning test, attendance and active participation in a introductory university course on electricity and magnetism (y-axis: % of maximum value): CG (control group): traditional lecture style; TG (treatment group): "active learning" intervention, see text [52].*

The results of a controlled comparison show a clear superiority of this approach over traditional lecturing for a learning test, attendance and active participation. The latter was assessed by an observation protocol for classroom behaviour; details of this and other elements of methodology and controls can be found in [52]. The effect on learning is 2.5 SDs, a very large value which the authors attribute to an instructional design which "maximized productive engagement". Moreover, this course method was well appreciated by students (e.g. "I would have learned more if the whole physics course would have been taught in that style": 77% agreement, 7% disagreement). The effectiveness of this

and MC are the means (of some variable of interest) for the treatment and control group, respectively, and SD is either the pooled standard deviation or that of the control group [45]. In simple terms, d thus measures the impact of an intervention in units of standard deviations of the sample under consideration. Usual effect-size levels (as established from comparison of a great many of studies in different areas) are small ($0.2 < d < 0.5$), medium ($0.5 \le d < 0.8$) or large ($0.8 \le d$) [45]. Many modifications and refinements of the concept of "effect size" have been developed and are used in the literature, see e.g. [46] for an overview.

kind of educational approach, called (inter-)active learning or interactive engagement according to various authors is backed by solid evidence over more than a decade, in particular by a large scale meta-analysis involving 225 studies and several ten thousands of students (see box "Active learning, interactive engagement").

There are many other examples of effective, evidence-based approaches for improving learning, such as selection of appropriate contexts for science learning. They may not only have positive effects on physics interest, as stated before, but also on learning. One example are the above-mentioned "Newspaper Story Problems" (effect sizes *ES* = 0.9 - 1.3 for the topics of elementary kinematics or of energy, [51]), or approaches integrating science, technology, and society [62]. Table 1 gives an overview of meta-analytic results for learning by various educational approaches.

| **Educational approach** | ***ES*** | **ref.** |
|---|---|---|
| feedback [a)] | 0.72 | [58] |
| cooperative vs. individualistic science learning<br>– school (primary and secondary)<br>– university (undergraduates) | <br>0.95<br>0.51 | <br>[59]<br>[60] |
| enhanced questioning strategies [b)] | 0.74 | [59] |
| active learning approaches (for lectures)<br>– across all sciences<br>– physics | <br>0.47<br>0.72 | [53] |
| homework<br>– with feedback<br>– without feedback | <br>0.83<br>0.28 | [61] |

*Table 1: Meta-analytic results for learning (of science, if not stated otherwise) by some educational approches*

*a) across all disciplines and a range of feedback methods*

*b) e.g., increasing wait time, adding pauses at key student-response points, including more high-cognitive-level questions, etc.*

### 2.3 Influences

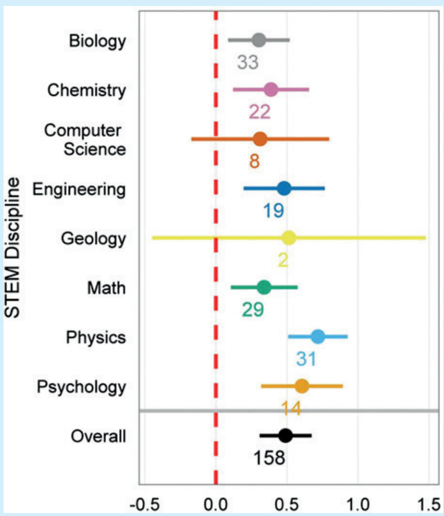
To understand (and design) educational processes one has to analyse them on a more fine-graded level than just comparing two approaches. Often a given educational measure will not work in the same way for different groups of learners (e.g. for boys and girls, see example regarding contexts above) or for different settings, and one has to study which influences are at play, from the side of the learner, teacher, setting, etc. A first informative example is about homework, a topic strongly debated among parents, teachers, and researchers: overall effects are between small and medium size (*ES* from 0.36 to 0.65), but a much clearer picture is obtained when taking account of feedback (by teachers comments or other methods) as influence, which leads to a large contrast (with: *ES* = 0.83, without: *ES* = 0.28; [61]). Of course, this does not close the discussion about homework, as feedback on homework is not always possible due to lack of time, or because several other influencing factors might be important in a given context. But this result is a useful element of this discussion: homework is *not* ineffective in general, as it is sometimes claimed, and feedback is decisive for its effectiveness.

Another example of an important influence, this time on the information processing level, is presented in the Infobox "Working Memory and Science Learning".

In general, as a recent review [71] puts it "*student success is influenced by such things as: demographics (age, gender, race, ethnicity, native language); intelligence and working memory; background knowledge and misconceptions; motivation, self-regulation, ability to pay attention, and persistence; self-concept and goal orientation, […]*" etc. It is e.g. a common situation at school that teachers face quite heterogeneous groups, and a recurrent problem is then that some innovation might work better for the those who score higher on some desirable attributes (interest, knowledge) already at the beginning than for those who score lower ("Mathew-effect", [72]). For physics learning, PER (together with general background form educational science) has provided many useful results about such influences, important for teaching practice, and in particular teacher education. The present article is limited to an example of an important influ-

**Active learning, interactive engagement**

The effectiveness of this kind of educational approach is backed by a large scale meta-analysis across several disciplines and many universities (see Figure and [53]) comparing active learning courses vs. traditional lectures. It yields an overall increase of learning by 0.47 SDs (for course exams, concept tests, and other assessment) and 0.88 SDs for concept tests in particular, as well as a decrease of failure rate from 34% to 22% (225 studies, ≈ 46000 and 29000 students for learning assessment and failure rate, respectively [53]).
For physics, a still larger effect was found for the overall effect on learning (0.72 SDs), a finding corroborated by another large study on introductory physics courses (600 classes, ≈ 45 000 students [54]). Similar approaches have been largely studied also at high school level, e.g. the predict-observe-explain sequence [55, 56, 57], which works even for demonstration experiments (which still are widespread and necessary for cost or security reasons), with the same main message: "Active learning increases student performance in science, engineering, and mathematics" [53], and it does so with considerable effect sizes, and is possible even in many teaching settings, including "frontal" ones [3].




*Comparison of active learning courses vs. traditional lectures: Meta-analytic effect sizes (in SDs5) for learning (examination scores, concept inventories, or other measures) by discipline [53]. Error bars are 95% confidence intervals, below each data point the number of included studies is given.*

### Working Memory and STEM Learning

Several studies have yielded substantive correlations between working memory and STEM achievement, across age groups and across science disciplines (Table 2). All values found are larger (up to a factor 2) or at least comparable to the correlation between motivation and achievement [1], considered as very important e.g. by many science outreach initiatives by physics departments. Note that this is about working (or short term) memory, not long term memory. The interpretation, in terms of a metaphor, is that it is the CPU which limits interrelatedness and complexity which can be treated at a given time, not the hard disk drive capacity, and that these factors are decisive for the quality of knowledge and understanding, in particular for complex topics like physics. This might appear obvious, but it has very important implications for reducing "cognitive load" (working memory demand) for teaching and learning (limitations of space do not allow to present this in detail here, but see e.g. [63]).

| STEM discipline (age/age group) | *r* (correlation coefficient) | comments | ref. |
|---|---|---|---|
| physical sciences | | | |
| – chem. (uni. freshmen) | 0.28 – 0.75 | 9 independent studies | [64] |
| – phys. (sec. I) | 0.30 | | [65] |
| other STEM disciplines | | | |
| math (14 yrs) | 0.54 | | [66] |
| science* (14 yrs) | 0.50 | | |
| math (11 yrs) | 0.58 | large sample study, n ≈ 5000 – 10000 depending on test and age group | [67] |
| math (16 yrs) | 0.63 | | |
| science* (11 yrs) | 0.46 | | |
| science* (16 yrs) | 0.60 | | |
| bio. (sec. I) | 0.62 | | [68] |
| comparison value | | | |
| motivation-achievement | 0.3 – 0.4 | n > 500 000(!) | [69] [70] |

*Table 2: Correlation between working memory capacity and science achievement for different STEM disciplines and age groups (sec. I: secondary level one, uni.: university), to be compared to the correlation between motivation and science achievement, as comparison value. There is only one result for physics with a somewhat lower value, for which I do not have an explantion*

**Note that "science" is taught in many countries not differentiated by disciplines.*

1 This holds also for the only study regarding physics, yielding a somewhat lower value than most other studies.

ence factor on the learner level [6] (working memory capacity), the method level (homework with or without feedback), and the teacher level (see 3.2). More complete discussions can be found e.g. in [1] and [73].

6 These influences are sometimes called "aptitude-treatment" or "attribute treatment" interactions in edcuational research.

### 2.4 Illusions (and ideology)

Let us come back to the necessity of measurement for taking educational decisions. Regarding science interest, we have already seen that it can be illusionary to believe that mere relevance is sufficient (Figure 3). A related result is about the PISA tasks, whose philosophy to ensure "relevance to students' interests and lives" is essential to PISA's understanding of scientific literacy ([10], p. 27). However, results show that for the available PISA items, students' perceived interest is at best medium, contrary to the basic assumption of PISA, and that it is also strongly overestimated by teachers [74]. Regarding learning, a first example is homework, already discussed in 2.3, where one sees that simplistic, ideological convictions ("homework is necessary"; "homework is punishment of families" [7]) do not lead anywhere.

Another example highly relevant for science education is inquiry-based learning (IBL), which finds strong support on the political level [75]. However, there is little empirical support for the effectiveness of IBL (*ES* = 0.31, [58]), and looking more closely, a meta-analysis on IBL in science by Furtak et al. [76] has shown that teacher guidance is a decisive influence factor: effect sizes are more than twice as large with guidance than without (*ES* = 0.65 and *ES* = 0.25, respectively). Unguided inquiry, as IBL is often understood, is thus not an effective approach for science learning. As in the case of homework, a differentiated, quantitatively based stance has to replace a simplistic, and sometimes ideological one.

**Acknowledgements:** This paper has benefited from numerous fruitful discussions with many collegues, in particular Hans Peter Beck (CERN), Bernhard Braunecker (Rebstein), Alice Gasparini (Geneva), Jean-Sébastien Graulich (Geneva), Hanns-Ludwig Harbey (Sierksdorf), Martin Pohl (Hamburg), Laura Weiss (Geneva).

*Part 2 of this article will cover practical tasks of PER and services it can provide to the community (such as for teacher education). The contribution contains many empirical data and sources not necessarily well-known to physicists. In order to provide a proper documentation for these sources, the reference list is longer than usual in the "Progress in Physics" series.*

### References

[1] Docktor, J. L., & Mestre, J. P., Physical Review – Physics Education Research, **10**(2), 020119 (2014).
[2] Lehrer, R., and Schauble, L. (2006). Scientific thinking and science literacy. In K. A. Renninger, W. Damon, I.E. Sigel, and R.M. Lerner (Eds.), Handbook of child psychology (vol IV): Child psychology in practice. Hoboken, NJ: Wiley
[3] Meltzer, D. E., & Thornton, R. K., American journal of physics, **80**(6), 478-496 (2012).
[4] Wieman, C.,Change: The Magazine of Higher Learning, **39**(5), 9-15 (2007).
[5] Wieman, C. E., Educational Researcher, **43**(1), 12-14 (2014).
[6] Wieman, C., & Perkins, K., Physics today, **58**(11), 36 (2005).
[7] National Research Council (2001). Knowing what students know: The science and design of educational assessment (Committee on the Foundations of Assessment ; Pelligrino, J., Chudowsky, N., and Glaser, R., editors). Washington, DC: National Academy Press.
[8] Adams, W. K., & Wieman, C. E., International Journal of Science Education, **33**(9), 1289-1312 (2011)
[9] Liu, X. (2012). Developing Measurement Instruments for Science Education Research. Second International Handbook of Science Education.

7 http://www.lemonde.fr/vous/article/2012/01/07/devoirs-scolaires-le-chatiment-familial_1627076_3238.html

Fraser, B. J., Tobin, K., McRobbie, C. J. (Eds.). Dortrecht: Springer
[10] OECD (2006). Assessing Scientific, Reading and Mathematical Literacy. A Framework for PISA 2006. Paris: Organisation for economic cooperation and development
[11] Fraser B. J., Tobin K. and McRobbie C. (Eds.) (2012). Second International Handbook of Science Education. Dordrecht: Springer
[12] Hund, F. (1969). Grundbegriffe der Physik, Mannheim: BI
[13] Kuhn, W. (2016). Ideengeschichte der Physik: Eine Analyse der Entwicklung der Physik im historischen Kontext. Heidelberg: Springer
[14] McDermott L. C. and Redish E. F. , Am. J. Phys. **67**, 755 (1999).
[15] Freeman S., et al., Proc Natl. Acad Sci USA **111**:8410–8415 (2014).
[16] Shtulman, A., & Valcarcel, J., Cognition, **124**(2), 209-215 (2012).
[17] Thacker, B., Dulli, H., Pattillo, D., & West, K., Physical Review - Physics Education Research, **10**(2), 020104 (2014).
[18] Novick, S., Nussbaum, J., Science Education **85**, 187 (1981)
[19] Planinic, M., L. Ivanjek, & Susac A., Physical Review - Physics Education Research, **6**, 010103 (2010)
[20] Wang, J., & Bao, L., American Journal of Physics, **78**(10), 1064-1070 (2010).
[21] Traxler, A., Henderson, R., Stewart, J., Stewart, G., Papak, A., & Lindell, R., Physical Review Physics Education Research, **14**(1), 010103 (2018).
[22] LoPresto, M. C., & Murrell, S. R., Journal of College Science Teaching, **40**(5), 14-22 (2011).
[23] Grant, E. (1981). Much ado about nothing: Theories of space and vacuum from the middle ages to the scientific revolution. Cambridge University Press.
[24] Krebs, R. E. (1999). Scientific development and misconceptions through the ages: A reference guide. Westport, Conn: Greenwood Press
[25] Jammer, M. (1957/1999). Concepts of force. New York: Dover
[26] Madsen, A., McKagan, S. B., & Sayre, E. C., American Journal of Physics, **85**(4), 245-264 (2017).
[27] Lichtenberger, A., Wagner, C., Hofer, S. I., Stern, E., & Vaterlaus, A., Physical Review Physics Education Research, **13**(1), 010115 (2017).
[28] Ding, L., & Beichner, R., Physical Review - Physics Education Research, **5**(2), 020103 (2009).
[29] P. B. Kohl, D. Rosengrant, and N. D. Finkelstein, Physical Review - Physics Education Research **3**, 1, 010108-1–010108-10 (2007).
[30] Treagust, D., Duit, R., Fischer, H. E. (2017). Multiple Representations in Physics Education. Dordrecht, Netherlands: Springer.
[31] P. Nieminen, A. Savinainen, and J. Viiri, Physical Review - Physics Education Research **6**, 2 (2010).
[32] Klein, P., Müller, A., & Kuhn, J., Physical Review - Physics Education Research, **13**(1), 010132 (2017).
[33] Docktor, J., & Heller, K. (2009, November). Assessment of student problem solving processes. In AIP Conference Proceedings (Vol. 1179) College Park: AIP, pp. 133-136.
[34] NRC (National Research Council, 2012). Discipline-Based Education Research: Understanding and Improving Learning in Undergraduate Science and Engineering, Washington, DC: National Academies Press.
[35] Caballero, M. D., Doughty, L., Turnbull, A. M., Pepper, R. E., & Pollock, S. J., Physical Review Physics Education Research, **13**(1), 010118 (2017).
[36] Etkina, E., Van Heuvelen, A., White-Brahmia, S., Brookes, D. T., Gentile, M., Murthy, S., ... & Warren, A., Physical Review - Physics Education Research, **2**(2), 020103 (2006).
[37] Khine, M.S. (2015) Attitude Measurements in Science Education: Classic and Contemporary Approaches. Charlotte, NC: Information Age Publishing.
[38] Handelsman, J., Ebert-May, D., Beichner, R., & Bruns, P. Science, **304**(5670), 521 (2004).
[39] Glynn, S., Koballa, T. R., (2005). The contextual teaching and learning approach. In R. E. Yager (Ed.), Exemplary Science. Best practices in professional Development. Arlington (VA): National Science Teacher Association Press, pp. 75-84.
[40] Kircher, E.; Girwidz, R.; Häußler, P.: Physikdidaktik - Theorie und Praxis, Berlin, Heidelberg, New York: Springer.
[41] Nentwig, P., & Waddington, D. (2005). Making it relevant: Context based learning of science. Münster: Waxmann.
[42] Muckenfuß, H., Naturwissenschaften im Unterricht, Physik, **7**/31, 20-25 (1996).
[43] Office fédéral de la statistique/OFS (2019), Etudiants des hautes écoles universitaires. Berne: OFS
[44] Hoffmann, L., Learning and instruction, **12**(4), 447-465 (2002)
[45] Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Erlbaum.
[46] Fritz, C. O., Morris, P. E., & Richler, J. J., Journal of experimental psychology: General, **141**(1), 2 (2012).
[47] Colicchia, G. (2002): Physikunterricht im Kontext von Medizin und Biologie: Entwicklung und Erprobung von Unterrichtseinheiten zur Steigerung des Interesses und für den fachübergreifenden Physikunterricht. (Doctoral dissertation). München: LMU München, Fakultät für Physik
[48] Wiesner, H. & Colicchia, G. (2010). Improving students interest – medical and biological contexts in physisics education. In C. P. Constantinou (Hrsg.), Physics Curriculum Design, Development and Validation (GIREP Conference 2008) Nicosia: Learning in Science Group, University of Cyprus.
[49] Bennett, J., Lubben, F., Hogarth, S., Science Education, **91**(3), 347-370 (2007).
[50] ROSE – Relevance in Science Education (2007-2018), *http://roseproject.no/*; see in particular: Sjøberg, S., & Schreiner, C. (2010). The ROSE project: An overview and key findings. Oslo: University of Oslo and Sjøberg, S, Schreiner, C. (2007). Reaching the minds and hearts of young people: What do we know about their interests, attitudes, values and priorities? What about the interest for space science? Bern: International Space Science Institute, Bern, *http://roseproject.no./network/countries/norway/eng/nor-sjoberg-issi-2007.pdf*
[51] Kuhn, J., Müller, A., Perspectives in Science: Progress in Science Education **2**, 5-21 (2014).
[52] Deslauriers L., Schelew E., Wieman C., Science **332**:862-4 (2011)
[53] Freeman S., et al., PNAS **111**:8410–8415 (2014).
[54] Von Korff, J., Archibeque, B., Gomez, K. A., Heckendorf, T., McKagan, S. B., Sayre, E. C., ... & Sorell, L., American Journal of Physics, **84**(12), 969-974 (2016).
[55] Yin, Y., Science scope, **35**(8), 48-53 (2012).
[56] Yin, Y., Tomita, M. K., & Shavelson, R. J., International Journal of Science Education, **36**(4), 531-552 (2014).
[57] Liew, C. W., & Treagust, D. F., Australian Science Teachers' Journal, **41**(1), 68-71 (1995).
[58] Hattie, A. C. (2009). Visible Learning. A synthesis of over 800 meta-analyses relating to achievement. London, New York: Routledge
[59] Schroeder , C., Scott, T., Tolson, H., Huang , T., & Lee, Y., Journal of Research in Science Teaching **44**, 1436-1460 (2007).
[60] Springer, L., Stanne M., and Donovan S., Review of Educational Research, Vol. **69**, No. 1, 1999, pp. 21–52.
[61] Marzano, R., Gaddy, B., Dean, C. (2000). "What Works In Classroom Instruction". Alexandria/VA: Association for Supervision and Curriculum Development
[62] Yager, R. E., & Weld, J. D., International Journal of Science Education, **21**(2), 169 – 194 (1999).
[63] Sweller, J., Ayres, P., & Kalyuga, S. (2011). Cognitive load theory: Explorations in the learning sciences, instructional systems and performance technologies. New York, NY: Springer.
[64] Tsaparlis, G., Research in Science & Technological Education, **23**(2), 125-148 (2005).
[65] Chen, W. C., & Whitehead, R., Research in Science & Technological Education, **27**(2), 151-160 (2009).
[66] Gathercole, S. E., Pickering, S. J., Knight, C., & Stegmann, Z. Applied Cognitive Psychology **18**(1), 1-16 (2004).
[67] Donati, G., Meaburn, E. L., & Dumontheil, I., Learning and Individual Differences, **69**, 84-93 (2019).
[68] Chu, Y. C. (2008). Learning difficulties in genetics and the development of related attitudes in Taiwanese junior high schools. (Doctoral dissertation). Glasgow: University of Glasgow, Centre for Science Education
[69] Wild, E., Hofer, M. & Pekrun, R. (2001). Psychologie des Lerners. In A. Krapp & B. Weidenmann (Hrsg.), Pädagogische Psychologie – Ein Lehrbuch (4. Aufl., S. 207-270). Weinheim: Beltz Psychologie Verlag.
[70] Uguroglu, M. E., & Walberg, H. J., American Educational Research Journal 16(4), 375-389 (1979).
[71] Berry, A., Friedrichsen, P., and Loughran, J., (Eds.). (2015). Re-examining Pedagogical Content Knowledge in Science Education. New York: Routledge.
[72] Walberg, H. J., & Tsai, S., American Educational Research Journal **20**, 359-373 (1983).
[73] Koran, M. L. and Koran J., Journal of Research in Science Teaching **21**(8):793-808, (1984)
[74] Weiss, L., & Müller, A., Zeitschrift für Didaktik der Naturwissenschaften, **21**(1), 87-97 (2015).
[75] Rocard M. et al. (2007) Science Education Now. A Renewed Pedagogy for the Future. Brussels: European Commission
[76] Furtak, E. M, Seidel, T., Iverson, H., Briggs, D., Review of Educational Research, **82**(3), 300-329 (2012). doi: 10.3102/0034654312457206

**For editorial reasons this article is split in two parts. Part 2 will be published in the *SPG Mitteilungen* Nr. 59.**