Typology of pitfalls for causal analyses in physics The case of capillary forces

Laurence Viennot[1]

[1]Université Paris Cité, Matter and Complex Systems, UMR 7057, 75205 Paris Cedex 13

Abstract. The reasoning errors of experts in teaching or popularization situations are always surprising. However, they often correspond to "complexity-reducing" schemes, some of which have been categorized for a long time, especially concerning the role of causality. If we restrict ourselves to cases where the reasoning in question leads to the correct result and excludes technical errors, there is still a range of relatively unknown types of invalidity - even though they are in fact very well represented in teaching and popular science practices. The aim of this presentation is to highlight the importance of situating these types of reasoning in a list of problematic explanatory situations previously identified. After a quick reminder of the most well-known elements of such a list, such as "functional reduction" and "linear causal reasoning", the presentation will focus on cases where the demonstration used seems to surreptitiously remove from the explanatory landscape one of the relevant variables of the phenomenon to be explained. Various examples will show that such a case can be observed in very diverse fields of physics. A case concerning capillary forces - the "liquid bridge" - will introduce a discussion we can have as educators according to our more or less informed treatment of these explanatory situations. At stake: conceptual coherence.

1 Introduction

In the context of a general advocacy for the development of critical thinking skills in our students [1], a lot of research has been carried out, in particular to try to determine the respective shares of a general ability in this area and of more content-specific skills. Based on the standpoint that both entries are relevant, this article focuses on the critical analyses of explanatory texts in physics. A major point that emerges from the various studies in this area is that reasoning traps do not only affect novices in physics but also experts, including university teachers, textbook authors and authors of research or popular articles. In the perspective of improving the critical thinking of our students, it seems all the more reasonable to foster a better knowledge of the main trends and difficulties in this field within our academic communities. The goal of this paper is to contribute to this project regarding explanations in physics. This will be done in strong relationship with the results of studies about common reasoning in physics [2], that is, a set of ideas and lines of reasoning that seem “natural” to most of us, and are therefore both frequently undiscussed and commonly invoked. Four major trends in common reasoning in physics will first be recalled, with an emphasis on those that involve an inappropriate simplification of causal analysis (section 2). The core of the paper will then develop a lesser known and more subtle case, the 'cause-in-the-formula' syndrome (section 3). In this case, one or more relevant variables of a phenomenon somehow disappear from the analytical expression of the result and only the variables seen in this formula are considered as "causes" of the phenomenon. This will be illustrated in an electrostatic situation and with two capillary phenomena: the tensiometer and the liquid bridge. The implications for teaching will then be briefly discussed (section 4).

2 Some major trends in common reasoning

Two very well-known trends in common reasoning are psycho-cognitive phenomena that are widely observed across very diverse contexts: the confirmation biais and the confusion between correlation and cause. The first one consists in favouring a line of reasoning that leads to a conclusion considered (by the author of the line of reasoning in question) as correct. Its importance in discussions between students and/or teachers is major [3]. The confusion between correlation and causation, which consists in thinking that if two variables are statistically correlated, one of them (why this one?) is the "cause" of the other, is also a crucial determinant of erroneous arguments and fake news in the general public [4]. Two other trends, of importance nearly similar to the preceding ones, are more extensively developed below because they appear as precursors of the more subtle case which is the target of this article. Both are related to causality

2.1 Functional reduction

This trend consists in ignoring certain relevant variables of a phenomenon, often to keep only one in the proposed explanation. The case of the pierced bottle is emblematic in this respect. In many images from popular and more specialised sources (including da

Vinci [5], see Figure 1), it is suggested that the range of the jets is greater the deeper the hole is in the container.

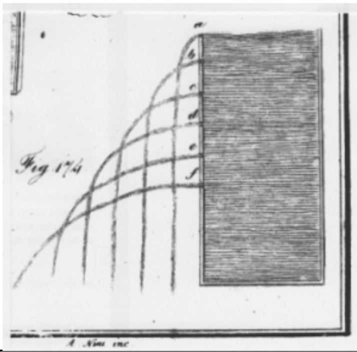
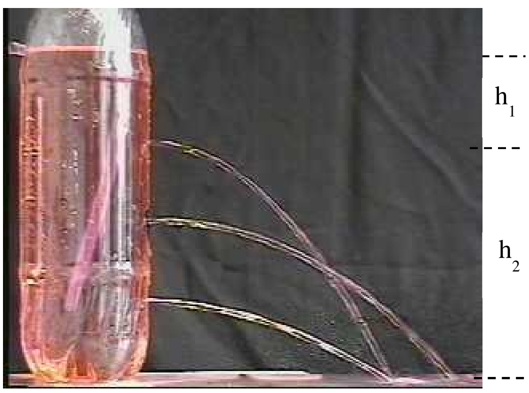


Fig. 1. Leonardo da Vinci suggested, as have many authors since, that for a pierced bottle, the closer the hole is to the bottom of the bottle, the greater the range of the emerging jet.

In fact, it is for the middle hole that the range is greatest. In images similar to those in Box 1a, the conclusion seems to be based solely on the exit velocity, which is in the order suggested. In fact, another variable must be taken into account: the duration of the fall from the bottle. An elementary calculation using two variables - exit velocity and duration of fall - (see Figure 2) is sufficient to justify what is observed.

This example may seem anecdotal but, although denounced as early as Torricelli, its remarkable permanence in academic sources attests to the strength of a tendency to 'reduce' causal analysis. Innumerable examples could have been cited here [6, 7, 8], if only all those comments of the type "fewer molecules, therefore less pressure" which ignore the role of temperature in the pressure of gases.



$V^2_{hor} = 2gh_1$ $\quad$ $t^2_{fal} = (2/g)h_2$

range of the jet = $2\sqrt{h_1 h_2}$

Fig. 2. Pierced bottle: the greatest range is for the middle hole.

2.2 Linear causal reasoning

The expression *linear causal reasoning* designates a linear chain of implications, each link mentioning a single phenomenon (ϕ), relating to the evolution of a single quantity: $\phi_1 \rightarrow \phi_2 \rightarrow \phi_3 \rightarrow ...\phi_n$. This brings us back to a form of reduction in the number of variables, or "functional reduction", just mentioned. A more specific characteristic is the status of the arrow which is used to symbolize the relationship between two successive phenomena in the explanatory sequence. An arrow between the predicates could represent a logical implication, that is, "therefore". Or it could indicate the occurrence of a subsequent event: "next". In fact, using the "then" connector offers a comfortable way of not choosing between a logical implication and a chronological sequence. Note that these features are in term-to-term opposition with what characterizes a quasi-static analysis of a system, where *several* variables are seen as changing *simultaneously* under the *permanent constraint* of relationships (Table 1).

Table 1. A term-to-term contrast between quasi-static physics and linear causal reasoning

Quasi-static physics	Linear causal reasoning
Several variables	Simple phenomena (one variable at a time)
change simultaneously	are taken into account simultaneously
constrained by permanent relationship	hence temporarily

We find an example of the implications of linear causal reasoning in an explanation of the siphon by Marie Curie [9]: "The water in the long arm of the siphon flows out. A vacuum is created and the atmospheric pressure makes the water rise in the short arm, which is immersed in the container". There is a characteristic structure in this argument: water ... flows out → a vacuum is created → atmospheric pressure causes the water to rise, i.e. simple and localised events are cited sequentially in a linear causal chain. The flaw is that the atmospheric pressure in the air at both ends of the tube is the same, and so the atmosphere cannot explain any global movement of water. Apart from the vaporisation of water in a vacuum, a siphon could function in a vacuum (but not without gravity). This flaw goes mostly unnoticed because of the sequential character of the explanation: by the end of the story, the beginning has been forgotten. This example is also an opportunity to show that a relevant analysis of a reasoning can guide the construction of a pedagogical scenario that focuses on a critical point of the incriminated reasoning. For example, hiding a part of the system, for example the side of the tank, and asking for a prediction about what will happen when the other end of the tube is unclogged, makes an appropriate answer hazardous. This is because the question of whether the water will "flow" or be sucked into the tank depends on the *relative* level of water at the two ends of the tube; this point is undecidable if one side of the siphon is hidden. To answer the question, *both* ends of the system must be considered *at the same time*.

3 Complex causality: A more subtle case

3.1 Electric field: A reduced causal analysis and associated misunderstanding

Based on Gauss theorem, the expression of the electric field at a point M near a conducting body at equilibrium is $\mathbf{E} = \sigma/\varepsilon_0\, \mathbf{n}$ where σ is the surface density of charge on the conductor, ε_0 the permittivity of empty space and $\mathbf{n}$ is a unit vector normal to the surface pointing outwards near M. The formula is very simple but may create a misunderstanding. Indeed, in answer to the question : "what are the sources of the field?", many students at undergraduate level would answer " the charges on the body" [10], thus neglecting the contribution of external charges, i.e., in passing, the principle of superposition of fields.

A causal analysis of this situation may explain this fact. In accepted physics, an external charge creates some charges by induction on the conducting body which themselves contribute to the local density of charge σ which permits to find the field near the body. But the direct contribution of an external charge to the field is still there at equilibrium as a component of this field. Fig. 2- 3 sums up this double causal path between an external charge and the field near a conducting body at equilibrium, one indirect, by induction, and one direct. The principle of superposition is satisfied.

The misunderstandings observed in university students ("the field created by the charges on the conductor is perpendicular to the surface") can be explained by the fact that the formula somehow "swallows" the direct contribution of the external charges. To sum up this example, the value of the electric field near a conducting body in equilibrium incorporates the coulombic action of external charges and that of surface charges of the conducting body while only the latter contribution appears in the formula $\mathbf{E} = \sigma/\varepsilon_0\ \mathbf{n}$ This creates a difficulty when it comes to taking all the sources of the field into account. The invalid causal analyses often observed are well in line with the trends detailed in the two preceding sections : functional reduction and linear causal reasoning. The misunderstandings observed in university students ("the field created by the charges on the conductor is perpendicular to the surface") can be explained by the fact that the formula somehow "swallows" the direct contribution of the external charges.

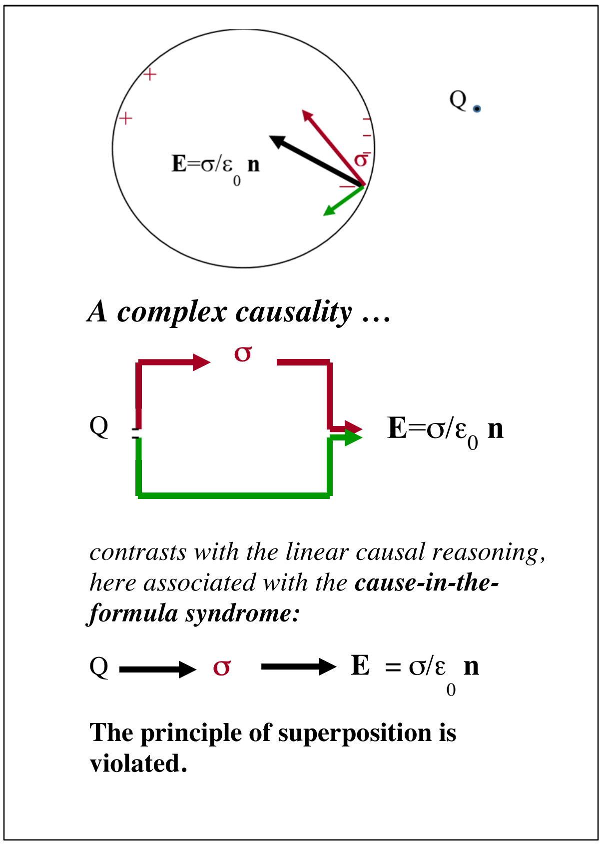


Fig. 3. The electric field near a conducting body at equilibrium is created by *all* the charges in the universe, not only by the surface charges on the body, σ.

To sum up this example, the value of the electric field near a conducting body in equilibrium incorporates the coulombic action of external charges and that of surface charges of the conducting body while only the latter contribution appears in the formula $\mathbf{E}=\sigma/\varepsilon_0\ \mathbf{n}$.

This example of causal reassignment is not isolated, as the following section shows.

3.2 Forces exerted by a liquid on a solid: a common causal analysis and associated misunderstandings

An example of explanation in the capillary domain poses the question dealt with in this section: the liquid bridge. Two spherical solid grains are mutually attracted when some liquid is in contact with both of them like in Figure 3. The explanation proposed in a recent article [11] is reproduced in Figure 4, mentions the role of a Laplace pressure linked to the curvature κ of the liquid/air interface. It also mentions a second cause for the attraction between the two grains, i.e. the "pulling force" linked to the surface tension γ_{LG} which intervenes on the triple line solid/liquid/gas. The expression of the total attractive force on a grain given by the authors ($F_{cap} = \gamma_{LV}\ 2\pi r \cos\theta$, r radius of the sphere, θ angle of contact) is correct. A question of mere consistency is then raised: why do the authors take the curvature of the liquid/gas interface into account while they (seemingly) don't consider the curvature of the solid/liquid interface? is correct. The main elements of their analysis are recalled in Figure 4.

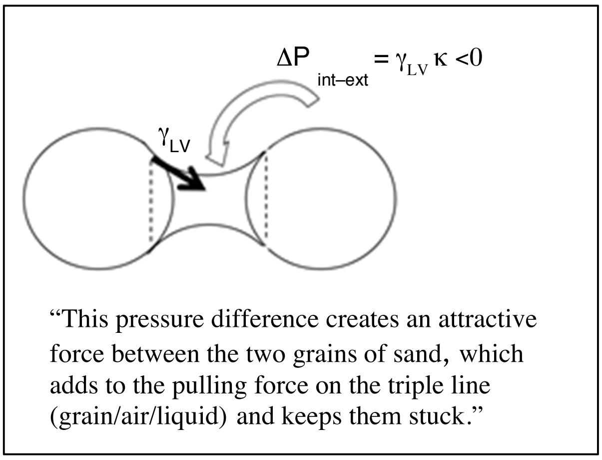


Fig. 4. An explanation of the attraction between two grains in hydrophilic contact with a liquid [11]

To resolve this apparent paradox, it is useful to remember the basics of capillary forces as proposed by Marchand et al. [12]. The main elements of their analysis are recalled in Figure 5. A first striking point in their model is that a glass molecule on the triple line is attracted by all the molecules of the liquid in the vicinity. At a corner, the solid is pulled towards the interior of the liquid. This contradicts a view of surface tension as a local force which would necessarily "pull" objects in contact with a liquid tangential to the liquid/gas interval. It is also worth noting that a free body diagram can be set up for the liquid wedge, with *no force pulling on it tangential from the outside*: forces *internal* to the drop are essential to understand what happens to the liquid wedge. In particular, If the liquid moves parallel to the stand toward the exterior of the drop, it is because it is pushed parallel to this plane by the liquid inside the drop. Note that the value of the component of the solid/liquid attraction, $\gamma_{LV}(1+\cos\theta)$, is far from obvious, but it is very clearly established for instance in [12].

In this light, the paradox raised above about the liquid bridge - a curvature seems to have been neglected - is added to a second one: why is "the pulling force on the triple line" represented tangentially at the liquid-gas interface? We may have the same question with a tensiometer (Figure 6a).

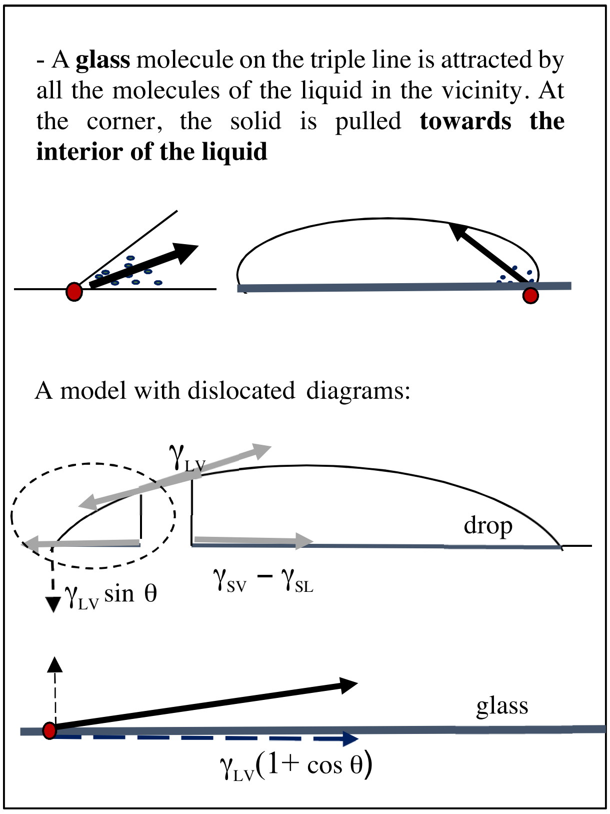


Fig.5. Elements of a model for the forces at play between a drop and its glass support, after [].

Here we need to consider what happens in the case of immersed solids. It is not only the liquid surface that acts on such solids, but the curvature of their immersed part creates a force towards the exterior of this liquid, like a net that pulls on this object from the outside. This mechanism is explained in Figure 7.

The expression of the total force due to this mechanism is extremely simple. For a tensiometer, for instance, it reads $\mathbf{F_c} = \gamma_{LV}\cos\theta\,\mathbf{u_z}$, θ being the contact angle, γ_{LV} the gas/liquid surface tension and $\mathbf{u_z}$ a unit vertical vector (Figure 6b). To sum up, the value of the force liquid–on-solid incorporates the local attraction liquid-solid and the repulsion due to the convexity of the solid. Only local attraction is suggested by the (correct) formula $F_{cap}= \gamma_{LG}\, 2\pi r \cos\theta$. This *wrongly* suggests that local attraction by the liquid acting locally at the corner, tangent to the interface, is the only cause of the attraction between the grains. The other component of the grain-liquid interaction is not localised at the corner but distributed over the whole liquid /solid interface.

Therefore, the structure of this situation has much to do with the preceding example, where the value of the electric field near a conducting body *incorporates* the direct effect of an external charge Q and the field created by the surface charges developed by the influence of Q on the conducting body, while only this indirect effect is evoked in the formula. In both cases we may observe a causal reassignment, that is, what is seen in the formula seems to designate the causes at play in the phenomenon.

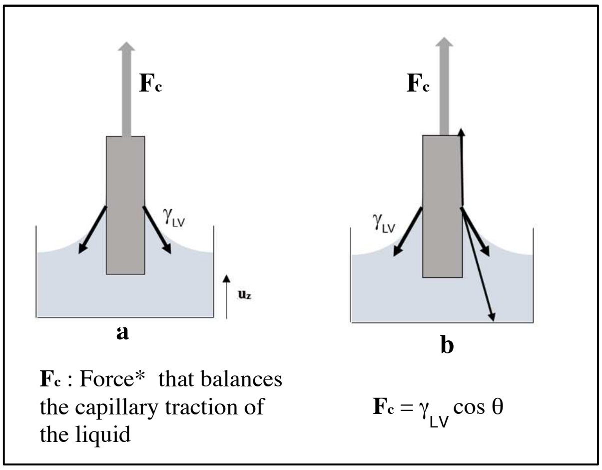


Fig. 6. A very common diagram for the tensiometer (a) and a more detailed account of the forces* at play (b): thin arrow downward: vertical component* of the local attraction water/glass $-\gamma_{LV}\,(1+\cos\theta)\,\mathbf{u_z}$; thin arrow upward: resultant* of the distributed forces due to the convexity of the immersed body, $\gamma_{LV}\,\mathbf{u_z}$. Both approaches lead to the same correct formula $\mathbf{F_c} = \gamma_{LV}\cos\theta\,\mathbf{u_z}$. *All "forces" are by unit length of triple line.

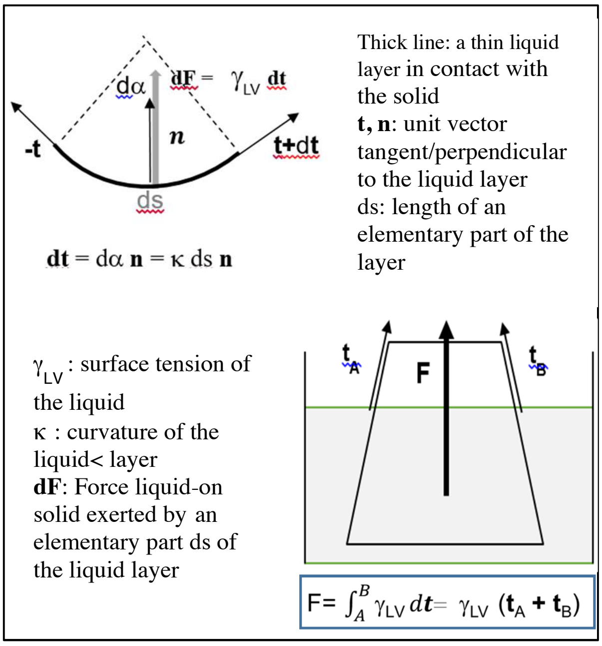


Fig.7 Elements to find the liquid-on-solid force for a partly submerged body

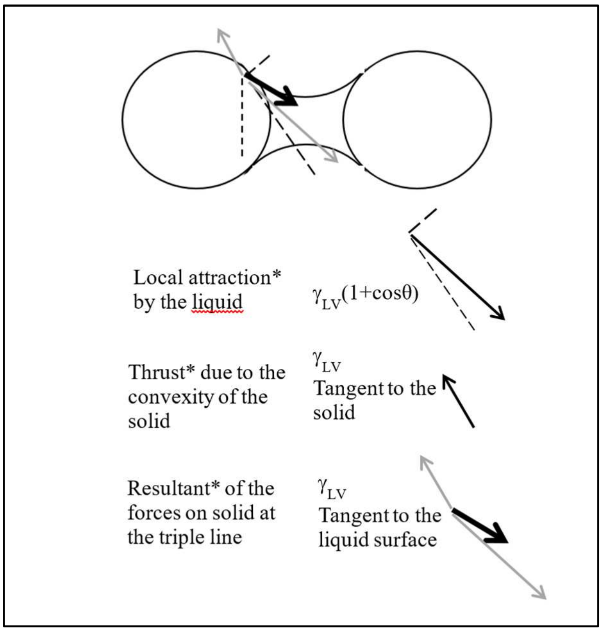


Fig. 8. Two types of forces act on a partly submerged body. The resultant* is tangent to the air/liquid interface. *All "forces" are by unit length of triple line.

Other versions of this kind of misleading interpretation of a formula are commonly observed in textbooks about capillarity forces. Thus, many explanations of capillary forces bear some similarity to Figure 9 which shows the edge of a drop on a hydrophilic support. Some forces are somehow invented to re-interpret a formula, here the Young relationship between coefficients of interfacial tensions. A liquid corner is identified and shown as being pulled at each corner by a single force (per unit length of the separation line): one horizontal force, say to the left, towards the gas, γ_{SG}, another parallel to the liquid-gas interface at the top right, γ_{LG}, and and another at the bottom right, γ_{SL}.

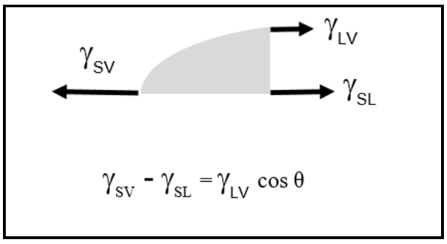


Fig. 9. A prototype of a causal reassignment (see for instance [13]). A correct analysis is given in Figure 5.

The well-known formula , $\gamma_{SV} - \gamma_{SL} = \gamma_{LV}\cos\theta$, is deduced from the balance between these forces. But the left end of the drop is said to be pulled outwards. The formula is correct, but the analysis is not, because no lateral force can be exerted on a water molecule by a flat, horizontal glass surface (Figure 5).

4 Recapitulation and final remarks

After a brief reminder of some major trends in common reasoning related to causality, their importance in some erroneous reasoning used even by experts was highlighted. This article focuses on a complex case where both localised and distributed 'causes' are involved in the explanation of a given phenomenon, while only one type of cause is somehow 'visible' in the formula that accounts for this phenomenon. The label "cause in the formula" syndrome was used to designate the reduction of the causal analysis to this visible part, which is compatible with a correct use of the formula, but goes with some misunderstandings concerning the nature and the localization of the causes. In the case of the electric field near a conducting body, admitting that only the charges on the body create the field comes down to deny the principle of superposition. In the case of a partly submerged body, it wrongly suggests a purely local action of the liquid on the solid, at the triple line solid-liquid-gas.

It could be argued that this is not so serious, as much can be done with correct formulas, without going into the details of the causes. "Why bother since you can calculate the correct result", as it is commonly heard. A possible answer is that what is at stake is not only such and such "detail", it can be the very principle of scientific reasoning, that is, *consistency*, as when a curvature is taken into account and (seemingly) not another one (Figure 3), or when the principle of superposition is denied.

In all cases, a clear vision of the explanatory choices made for oneself or for the students is recurrently put forward by the student teachers at the end of preparation sessions to the critical analysis []. It is also often said to facilitate the discussions between colleagues. Even more broadly, thinking in physics is similar to playing a sport at a high level. It is therefore preferable to be aware of the tools at our disposal, the most successful techniques and the possible pitfalls of such a practice.

References

1. European Commission, Science Education for responsible citizenship Report EUR 26893, EN chair H. Hazelkorn (European Commission, Brussels, 2015)
2. L. Viennot, Reasoning in physics The part of common sense (Springer, Dordrecht, 2001)
3. J.B. Henderson, A. Macpherson, J. Osborne, A. Wild, *Beyond construction: Five arguments for the role and value of critique in learning science*. Int. J. Sc. Ed., **37**, 10, 1668-1697 (2015)
4. N. Gauvrit, *Statistiques méfiez-vous!* (Ellipse, Paris, 2007)
5. L. da Vinci, *Del moto e misura dell' acqua di Leonardo da Vinci. A spese di Francesco Cardinali*, Bologna (1828). Digitizedd copy of Harvard College Library, Google books.
6. L. Viennot, N. Décamp, *Activation of a critical attitude in prospective teachers: from research investigations to guidelines for teacher education* Phys. Rev. Phys. Educ. Res. **14**, 010133 (2018)
7. L. Viennot, N. Décamp, L'apprentissage de la critique Développer l'analyse critique en physique (EDP Sciences, Les Ulis, 2019);
8. L. Viennot, N. Décamp, Developing Critical Thinking in Physics The Apprenticeship of Critique, Contributions From Science Education Research, vol. 7 (Springer, 2020)
9. I. Chavannes. Leçons de Marie Curie aux enfants de nos amis, (EDP Sciences, Les Ulis, 1907-2003)
10. S. Rainson, L. Viennot, *Charges et champs électriques: difficultés et éléments de stratégies pédagogiques en Mathématiques Spéciales Technologiques,* Didaskalia **12**, 31-59 (1998)
11. A. Sauret, G. Saingier, P. Jop, *Erosion et accrétion de matériaux granulaires humid*es, Reflets de la physique **64**, 17-22. (2020)
12. A Marchand, J.H. Weijs, J H Snoeijer, B. Andreotti, *Why is surface tension parallel to the interface*, Am. J. Phys., 999 -1008 (2011)
13. P.G. de Gennes, F. Brochard-Wyart, D. Quéré Gouttes, bulles, perles et ondes (Belin, Paris, 2005)