Science, Engineering and Bibliometrics: What is the confusion?

Emil Alégroth
Chalmers University of Technology
Department of Computer Science and Engineering
SE-412 96 Göteborg, Sweden
Emil.Alegroth@Chalmers.se

***Abstract*—Science, Engineering, Software Engineering, Software Engineering Science? In this report I give my views, supported by references from the scientific community, on the difference and meaning of these terms.**

***Keywords*-Sience; Engineering; Bibliometrics**

I. Introduction

What is science? Depending on who you ask you will get a different answer. Nevertheless, the question is important, and perhaps not asked enough within our academic community. In this paper, I give my views, not answers, to this question and extend the discussion through deeper probing questions regarding this subject within the area of Software Engineering. In addition, this report takes an initial, and brief, look into the importance of citation and bibliometrics and the role these play within the research community.

II. What is science?

Science is the process of how we improve our understanding of how nature works, including everything from the marco universe, e.g. how galaxies rotate, to the micro universe, e.g. how quarks interact to form atoms. Hence, science drives our knowledge and understanding of nature. However, science does not always drive this process of understanding through fact, but rather through best guesses that are verified through experimentation or observation in nature. These guesses are referred to as hypotheses, which, after rigorous testing, can become scientific fact. A scientific fact is a hypothesis that has been so repeatedly tested, and so much reliable evidence exists to support it, that it would be perverse or irrational to deny it [1]. If many of these scientific facts are grouped together we can form a theory, which in itself should be subject to rigorous testing. However, as we know, testing can never prove the absence of fault, only strengthen our convictions of the possibility of absence of fault. Hence, a theory, just as the scientific facts from where it originates, is still not a fact, but rather a concept that can withstand rigorous testing and probing without being proven false. However, this does not mean that a theory is inherently true, many theories have been shown to be wrong and have therefore been changed or corrected based on new scientific facts. Hence, a theory could be considered the model of a fact, since a fact can be proven right or wrong through a new discovery, whilst a theory is generally only strengthened or weakened by it.

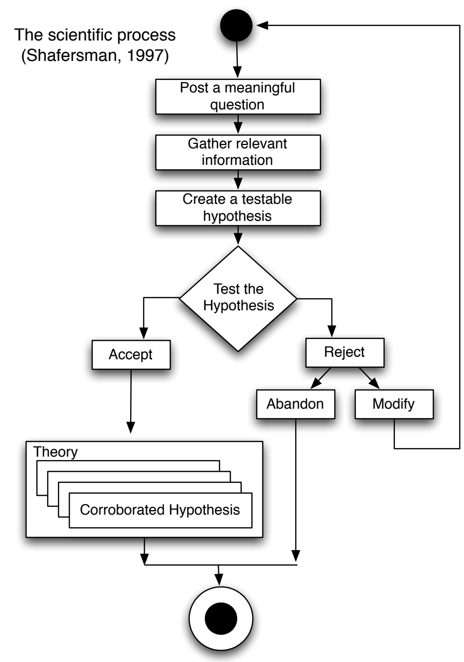

Fig. 1. The scientific process as described by Shafersman [1].

In order for science to reach high quality and rigor it should be conducted following a process, i.e. the scientific process as presented by Shafersman [1], visualized in Figure 1. In addition, for the process to be successful, the researcher has to adopt logical- and critical-thinking, i.e. thinking, free from emotion, that challenges new information that is not based on relevant and corroborated evidence.

A more simplistic way of viewing science is to refer to science as the concept of state-of-art knowledge and understanding of our reality, i.e. our current best guesses. In contrast to state-of-art, we have state-of-practice, which is generally built upon state-of-art but has been refined into practical solutions that are of use to someone or something, e.g. theoretical physics lead to the creation of nuclear power. The development of state-of-practice can in itself help strengthen or weaken state-of-art, forming a tight dependency between them where one improves and/or challenges the other, and vice versa.

III. How is science different from engineering?

Whilst science aims to advance our knowledge and understanding, engineering aims to apply it. Hence, in engineering you take the theoretical concepts, which may be both state-of-practice and/or state-of-art, and realize them into an applied solution for use in the real-world [2]. The core difference then, between science and engineering, is that scientists observe and do experiments in nature to understand it, whilst engineers change, decompose and reassemble nature to reach a desired result, i.e. a product [3]. Scientists in different fields do realization as well, e.g. proof of concept solutions or products, but engineers focus on the realization. In addition, realization through engineering, i.e. products, are aimed to be of use and serve a practical purpose, whilst scientific realization mostly serves the purpose of corroborating or contradicting a hypothesis or theory.

Furthermore, not all engineering is based in science, for instance the steam engine had existed for two centuries before the scientific field of thermodynamics was developed, i.e. providing an example where engineering pawed the way for a new scientific field [3]. In addition, to be a scientist, and an expert in your field, you need to be up-to-date on the latest research within your field. In contrast, as an engineer you need to have a broad understanding of all the factors you need to take into consideration in order to build a good product, but not necessarily the very latest research results [2].

These descriptions paint a picture of a clear distinction between scientists and engineers, but is the distinction really that clear cut? Given, most science aims to push our understanding of the world forward, but many scientists have designed and constructed products, and engineers have also made important scientific discoveries. Furthermore, if one takes a look at the name of our own research field, i.e. Software Engineering, one cannot help but to be confused. So we must ask ourselves, does the distinction between a scientist and an engineer lie in the taxonomy and definitions that we use, or is it our contributions that matter? I tend to agree with the latter.

IV. How is science different from a process improvement project in industry?

In science, academics strive to find the best general solution to solve a problem or problems within a given domain. Hence, state-of-art tends to be quite comprehensive but still modifiable in certain regards to make the solution more widely applicable. In a process improvement project, state-of-art is generally molded to fit the organization of the target company in its own domain. Consequently, in a process improvement project, the general state-of-art solution is molded to a more specific solution that solves a specific problem or set of problems at the given company. However, one important part of science is to do verification and validation of theoretical concepts, which is done by applicability studies, either in experiments or in a real-world context. Hence, a subset of process improvement projects, can be considered science, given that these projects are governed by academic practices that put more rigor into success, or failure, measurement. A question does however remain if results from these applied projects actually corroborate the hypotheses set by the original work since the conditions, compared to the original work, are different. The ideal-world, where all the prerequisites of a scientific theory are fulfilled, does not exist. Therefore, we must accept the chaos and unpredictability of the real-world and accept that there are factors that we cannot control. However, does this mean that our scientific process for applied science is flawed and that the best we can strive to achieve are approximations of the truth? As stated by Shafersman, “science and naturalism reject the concept of ultimate or absolute truth in favor of a concept of proximate reliable truth that is far more successful and intellectually satisfying than the alternative, the philosophy of supernaturalism” [1]. As also stated by assistant professor Matthias Tichy, when asked about this concept, ”We live in an analog world, therefore we will never find absolute answers, those we only find in the digital world, instead we have answers within intervals that we can deem more or less acceptable”.

V. What kind of science is Software Engineering?

Software engineering is an applied science, meaning that the solutions we develop are aimed at practitioners to solve, or mitigate, real-world problems, issues and challenges. Hence, in contrast with more theoretical research, Software Engineering research has a larger emphasis on the business perspective with the goal of serving a market need or want, e.g. lowering cost and raising quality of software being developed by industrial practitioners. However, the term Software Engineering, with emphasis on engineering, is the cause of much dispute and confusion within both the public and scientific community. The confusion lies in if an engineer can also be a scientist since engineers develop applicable products, or entities, whilst scientists work with theoretical concepts that need not have applicable use. However, does this mean that all science is mutually exclusive from engineering? To answer this question I think we have to redefine our views of what a scientific contribution can be, and ask ourselves what differentiates a scientific contribution from a product? Looking at the definition of a product, ”A good, idea, method, information, object, or service that is the end result of a process and serves as a need or want satisfier” [4], we see that there is nothing prohibiting us from calling a scientific contribution a product. Thereby we have bluntly answered our original question, what kind of science is Software Engineering?, and can state that Software Engineering is a science focused on the development of scientific contributions that serve, or will serve, a need or want within the field, i.e. product engineering. This reasoning leads us to the hypothesis that Software Engineering science actually equals engineering. I welcome all scientific fact that contests this hypothesis, even though support seems easier to identify [5].

VI. Why is publication important in science?

In order for any scientific field to advance, it is important to build upon the work of others as not to (re-)invent the same thing over and over. Publication is therefore important since it serves as a means of providing the scientific community with successes, and failures, to build upon, and drive our understanding forward. However, in Software Engineering there has yet to arise a process in which we, "Stand on the shoulders of giants" as so adequately stated by Isaac Newton [6]. Instead we are fostered into a traditional and narrow-minded view in which scientists should have deep knowledge rather than broad knowledge. Broad knowledge is required in engineering, in order to consider all factors that make up a good product, but is Software Engineering not an engineering discipline? The hypothesis, stated in Section V, tells us that it is.

Additionally, in contrast to engineered products that are judged by society, publications are judged and evaluated by academic peers before publication that helps ensure a rigor and validity of the research. Hence, the peer review process ensures that the level of a scientific contribution and the standards of what is to be considered "good research" are upheld. At least, that is how it is ideally supposed to work. How well this process works in practice, in the context of political decision making, financing, etc, is a subject for a whole other discussion.

VII. Why is scientific/research community central for scientific processes?

As mentioned, the scientific community has the important purpose of reviewing and judging the rigor of scientific work to ensure not only the quality of the work being done but also that the work is driving our understanding forward. Without a scientific/research community, information distribution among researchers would come to a halt together with our scientific advances. Hence, the scientific/research communities responsibilities are two-fold. First, to judge and review the scientific work. Second, to drive the respective scientific fields forward by promoting good research and spreading this research within the community.

VIII. What are the two most common bibliometric measures out there and how are they calculated?

Bibliometrics are based on analysis of scientific literature, e.g. citations, i.e. how many times a work has been referenced by peers within a scientific community. There are several ways of creating bibliometrics from citations, but the two most common ones are based on the collected citations of a group, e.g. a research group, institute, etc, or individual authors. There a several indices for measuring the citations of individual authors, but one of the most commonly used is the H-index, which is calculated as the h value for x number of publications, from an author, with h number of citations, where x is higher than h.

Bibliometrics are also used to rank the quality of journals, i.e. the impact factor of the journal, which is calculated as the average number of citations of the publications of said journal. Hence, a journal with a higher impact factor is considered to be more important than a journal with a lower impact factor. The impact factor is important for instance when evaluating work that has not yet been cited, i.e. publications within prestigious journals are considered to be of higher quality than publications in less prestigious journals. In addition, the impact factor of a publication is also, in general, a better indicator of scientific rigor than the publication's number of citations, since a publication can be poor in terms of scientific rigor but still well cited due to good positioning within the scientific field.

Bibliometrics can, and are, however misused, e.g. the popular ISI impact factor that can be used to assess the quality of a research group, institute, etc., by studying the journals they have published in. However, this is not what the impact factor was intended for and its use for this purpose is therefore unadvisable [7]. Because of misuse and the bias of bibliometrics, e.g. that more senior researchers gain from it, several alternative indices have been proposed, e.g. the future H-index, the HYPE(R) index, etc. These indices have in common that they not only take the researchers previous work into account, but also what he/she is expected to achieve and the impact the researcher has had on the scientific community. Hence, leveling the playing field for senior and junior researchers.

IX. Summary

What is science? How does the scientific community affect us as researchers and how we perform, and convey, our research? In this report I have given my views, supported by references from the scientific community, on these questions. However, the discussion is far from over, and even though some questions have been answered, new ones have also been formed. In addition, based on the findings from the academic community, it is clear that there is not a consensus as to what the answers to the posted questions are, especially regarding what type of science Software Engineering is. Therefore, I post to you, can we accept multi-faceted answers to these questions, i.e. our best approximations of the truth. Or, do we need to strive to find a common consensus? That, ladies and gentlemen, is the real question!

References

[1] S. Schafersman, "An introduction to science: Scientific thinking and the scientific method," *online whitepaper, January*, 1994.

[2] D. Parnas, "Software engineering programmes are not computer science programmes," *Annals of Software Engineering*, vol. 6, no. 1, pp. 19–37, 1998.

[3] H. Petroski. (2010, Dec.) Engineering is not science: And confusing the two keeps us from solving the problems of the world. [Online]. Available: http://spectrum.ieee.org/at-work/tech-careers/engineering-is-not-science

[4] B. Dictionary. (2012, Oct.) Product. [Online]. Available: http://www.businessdictionary.com/definition/product.html

[5] L. Briand, "Embracing the engineering side of software engineering," *Software, IEEE*, vol. 29, no. 4, p. 96, july-aug. 2012.

[6] B. Selic, "What will it take? a view on adoption of model-based methods in practice," *Software and Systems Modeling*, pp. 1–14, 2012, 10.1007/s10270-012-0261-0. [Online]. Available: http://dx.doi.org/10.1007/s10270-012-0261-0

[7] T. K. I. B. P. Group, "Bibliometrics - publication analysis as a tool for science mapping and research assessment," *online*, 2008.