Paving the cowpath in research within pure mathematics – A medium level model based on text driven variations.

Heuer, Karl; Sarikaya, Deniz

Published in:
Studies in History and Philosophy of Science, Part A

DOI:
10.1016/j.shpsa.2023.05.006

Publication date:
2023

Studies in History and Philosophy of Science 100 (2023) 39–46

Contents lists available at ScienceDirect

Studies in History and Philosophy of Science

journal homepage: www.elsevier.com/locate/shpsa

Paving the cowpath in research within pure mathematics: A medium level model based on text driven variations.

Karl Heuer[a,**,1], Deniz Sarikaya[a,b,c,1,*]

[a] Department of Applied Mathematics and Computer Science, Technical University of Denmark, Kongens Lyngby, Denmark
[b] Centre for Logic and Philosophy of Science of the Vrije Universiteit Brussel (VUB), Belgium
[c] Department of Science Education, University of Copenhagen, Copenhagen, Denmark

ARTICLE INFO

Keywords:
Text-driven variations
Problem-posing
Creativity
Philosophy of mathematical practice
Pursuit-worthiness
Narratives of mathematics

ABSTRACT

In this paper we show how simple text-driven variations of given statements in mathematics can lead to interesting new problems and push forward a whole theory around simple initial questions. We exemplify this in two cases. Case 1 deals with problem-posing activities suitable for pupils and case 2 is a rational reconstruction of the organisation of mathematical knowledge within problems of graph colorings. Mathematicians learn to systematically look for subsequent problems around a given problem.

We argue that this toy-model captures a nontrivial part of professional mathematical research within the pure fields and conjecture that it even grasps high level developments in mathematics. By doing this, we implicitly encourage a very simplistic view on criteria, so to speak a "cowpath" approach to progress in mathematics. The term "cowpath" is borrowed from architecture and software design, where it is commonly used. While we can contemplate which pathways are ideal, we may also just plant grass and see where people choose to walk. Those pathways are also self-enforcing, since we are less hesitant to walk on those rather than criss-cross the landscape.

When citing this paper, please use the full journal title *Studies in History and Philosophy of Science*.

1. Introduction

The central question this paper addresses is: *How do we generate mathematical questions within pure mathematics*? In principle, mathematicians could consider arbitrary axiomatic systems and ask arbitrary questions about them. To some extent there seems to be substantial freedom: The 2020 *Mathematics Subject Classification* of the *American Mathematical Society* distinguishes over 6000 categories. The division of labour and level of specialisation in mathematics is immense.[2] However, if a group of mathematicians were to propose an *entirely* arbitrary set of axioms, they would not receive much attention from other mathematicians for their endeavour. Yet, how do we choose which problems are worth solving in pure mathematics, and which questions are worth pursuing?

We argue that there are no strong criteria for pursuitworthiness or - at least - that nontrivial parts of mathematical practice can be explained without recurring to them. Instead, we defend a relativistic idea of concrete questions.[3] Our slogan is: *I prove what I am able to prove (and publish)*. Instead of focussing on strong and exceptional ideas, we focus on systematic search. Mathematicians do not stray far from their knowledge base and background. They extend knowledge locally. If I were an expert on finite Matroids, it would be more salient for me to start working on infinite Matroids rather than changing to, say, Kähler Manifolds. My past work is much more likely to be useful for topics, questions and problems close to it. Even if I were to prove a set of results, only some of said results would then enter into the established corpus of knowledge, which a nontrivial number of practitioners would internalize. Here, we want to draw an analogy between the above and a design philosophy originating in architecture, but currently probably most present in web design: *Paving*

* Corresponding author. Vrije Universiteit Brussel, Department for History, Archeology, Arts, Philosophy and Ethics (HARP), Centre for Logic and Philosophy of Science, Pleinlaan 2, room 5B425, B-1050, Brussels, Belgium.
** Corresponding author.
E-mail addresses: karheu@dtu.dk (K. Heuer), densar@dtu.dk, deniz.sarikaya@vub.be (D. Sarikaya).
1 Both authors share first authorship and contributed equally.
2 Cf. Kitcher (1990) for more general aspects of the division of labour.
3 Cf. in this volume Shaw (2022a) who "propose[s] a challenge to the possibility of even minimal criteria of pursuitworthiness".

https://doi.org/10.1016/j.shpsa.2023.05.006
Received 21 May 2021; Received in revised form 25 May 2023;
Available online 15 June 2023

the cowpath (or sometimes called *paving the desire-paths*). Imagine you build a new campus. You then plant grass covering all building-free areas of the complex. Campus users will cross these open areas in line with their goals and, over time, trample damage will result in visible pathways of dead grass along the most used routes. After a while, you pave those paths of dead grass. This second part of paving is the important aspect for new concepts and theories; Here, it is about the organization of knowledge.

We want to demystify mathematical research. More concretely, we aim to draw a picture in which one can learn to conduct mathematical research without needing strokes of geniality, and without needing to see intricate connections between distant fields of study. In our picture, doing mathematics is a local activity. A researcher extends their knowledge by looking for phenomena close to themes that they are currently working on to generate new problems. We show how a simple tool of playful textual variations, i.e. playing with concrete formulations of a given theorem or problem, already opens up a vast space of possible questions. We claim that mathematicians choose among newly suggested problems those they think they can prove. Only later, next generations of mathematicians choose some results that they find useful for their work. Those are – so to speak – "paved", i.e., codified in textbooks, included in lectures, etc. Only then, in hindsight, the value of studying those questions emerges. Not only does this determine success in an inner-scientific sense but also in terms of embedding novel results into existing mathematical structures.[4] Again, we import a well-known concept of software-design into the discussion of mathematical practice. In software design, we are told:

> [L]ook where the paths are already being formed by behavior and then formalize them, rather than creating some kind of idealized path structure that ignores history and tradition and human nature and geometry and ergonomics and common sense. (Crumlish & Malone, 2009).

We argue that textual proximity is key to understanding how mathematicians choose which topics to pursue and which problems to work on. To us, this boils down to the ease of converting the formulation of one problem into the formulation of a novel one. To give an example: Dropping an adjective from the antecedent normally yields to a generalization, such a change is already textually salient. Other generalizations might be further away from the syntactical form, like finding a structure theorem, which has our original theorem as an easy corollary. We conjecture there is a story to tell here in terms of iterations of our simple variations. Or to stress the metaphor: The big green field of possible results can be crossed in many ways, i.e. it is possible to vary questions easily to pose new ones. Only some of those random walks, will yield to paths that more people will trample upon, i.e. only some questions start a proper program, which many mathematicians follow and only in hindsight we judge which direction yield to places we want to be. But this history is written after creating the paths. This is a first model that will need more refinement and hopefully capture larger parts of mathematical practice more accurately in the long run.

In general, this work can be seen as part of the philosophy of mathematical practice research program, as we aim to render our picture of mathematics in accordance with actual practice. This tradition is becoming rather influential so that it might not be necessary to revisit the beginnings of the tradition and the influence of landmark books and articles such as Aspray and Kitcher (Eds. 1988), Buldt et al. (2008), Corfield (2003), Löwe & Müller (Eds. 2010), Van Kerkhove & Van Bendegem (Eds. 2007), Kitcher (1985), Maddy (1997, 2007) or Mancosu (Eds. 2008). For an overview, historical remarks and methodological comments see e.g. Löwe (2016), Kant et al. (2021), Perez-Escobar (2022) or Carter (2019).

We develop the idea of *syntactic variations* as a way of mechanizing the discovery process for interesting new questions and suitable definitions. We argue that this model.

(1) reflects important parts of mathematical research and
(2) is useful for mathematical education.

This article is structured as follows: First, we exemplify this method using an example of tilings of the plane (Section 2). We then rationally reconstruct more historical developments within the field of graph colorings (Section 3). Following this, we comment on more general inner-theoretical criteria (Section 4) and finally focus on possible critiques to our view (Section 5).

2. The idea of textual variations and mathematical creativity

In this section we present a simple heuristic to propose new problems and to organize mathematical knowledge within a research field. In fact, this heuristic is implemented in enrichment programmes for gifted high school students. Here we follow the idea of open problem fields.[5] In general, when we work on open problem fields, we start off with an initial question. This question should be carefully chosen and (in the context of enrichment for pupils) be easily accessible. Starting from this question, we change certain words of a given phrasing of the question while staying in the same mathematical context. This process may be iterated. Experience will teach a mathematician how to alter the words in order to get fruitful new questions. For instance, words related to numbers can often be altered, usually by other numbers, like when we change the dimension of a problem. Leaving out assumptions corresponds to generalisations. We tend to replace words by other conceptually related ones. If this yields conceptional mistakes or problems, we argue that this actually forces us to adapt the notions in play, yielding a more creative output. Apparently, mathematicians do take semantical considerations into account immediately, so the textual perspective is not everything there is.[6] However, it captures astonishingly many aspects of problem posing. In addition, this phrasing for the proposed method should make it more accessible for use in mathematics education. So, the issue of which question to pursuit seemingly vanishes; it is rather that some questions force themselves upon the practitioner.

The material presented in this section of the paper is developed within programs that (as many others in this sector) aim to mimic mathematical research and often include working mathematicians in their design and implementation stage. We can acquire more experience with the problem field in such classes, as we see them tackled by many groups of pupils and they are normally studied within sessions (and not years, or whole careers). So, we see this more as a lab experiment of mathematical research. This said, let's visit a problem field.

A problem field should be suitable for children within the classroom, assuming a heterogeneous level of motivation, engagement, and mathematical abilities. Here we want to boil down this whole problem field and stress the textual proximity that all involved questions have.

We initiate the whole problem field with the following question.

(P1) Which convex regular n-gons admit a tiling of the plane using only one type of tiles?

We recall that a *regular* n-gon is an n-sided polygon with all angles equal in measure and whose sides have all the same length. The n-gon

[4] The same holds for other sciences, as argued by Lichtenstein (2021). It is not merely about some success.

[5] For more on the idea of open problem fields consider Nolte (2002, 2012), Nolte and Pamperien (2017) or Kieβwetter (1985, 2009). In Heuer and Sarikaya (2019) we find material for a potential class.

[6] For literature on the bridging of textual positions and other notions of meaning see f.i. Fisseni et al., 2019 or Carl, Cramer, Fisseni, Sarikaya, & Schröder, 2021.

being *convex* means that for any two points within the area enclosed by the *n*-gon, the straight-line segment connecting these two points does also lie within the enclosed area. A *tiling* of the plane is a collection of tiles (subsets of the plane) which covers the plane without gaps or overlaps.

Some elementary angle calculations show that the only possibilities for the values of *n* yielding a desired tiling are 3, 4 and 6. In the style of Platonic solids, such tilings are called *Platonic tilings*. It might be too demanding to ask a student to find suitable follow-up questions and generalisations, but if we stress that a generalisation is often achieved by dropping antecedents, we get the following generalisation via a purely textual manipulation.

(P2) Which convex regular *n*-gons admit a tiling of the plane using only one type of tiles?

Or.

(P3) Which "**figures**" admit a tiling of the plane using only one type of tiles?

Dropping conditions like convexity or regularity as in (P2) results in more general *n*-gons that we allow for our tiling. Following this attempt to allow more general tilings, we used the word "figure" here without precise definition, but as a placeholder for a more (or even most) general shape our tiles can have. So, moving from (P2) to (P3) might even be seen as a process of allowing or generating many more questions of this type.

A slightly different procedure, but also yielding a more general question, would be to weaken the restriction on the number of allowed types of tiles, which can of course be iterated as well:

(P4) Which convex regular *n*-gons admit a tiling of the plane using only **more than one** types of tiles?

To give these considerations some perspective: The tilings which answer question (P4) are called Archimedean tilings. Combinations of these two presented textual variation attempts might be considered as well.

We get an even bigger variety of questions by substituting words. For instance, by substituting "plane" with other objects, we get:

(P5) Which … admit a tiling of a given **rectangle** using … ?
(P6) Which … admit a tiling of a **unit disc** using … ?
(P7) Which … admit a tiling of the (3-dimensional) **Euclidean space** using … ?

The last question illustrates something special. One very narrow point of view regarding (P7) could be to say that a conceptual mistake had taken place while posing this question, since 2-dimensional tiles are not meant to fill or even partition the space in any sense. At this point, we omit a discussion making the case that this is in principle possible, even with 1-dimensional objects like space filling curves etc. We will revisit this problem in a moment.

This reflects a whole design-philosophy for problem sheets. This particular way of developing the problem field together with pupils has been done with children within regular classes grad 8–12 of the German school system with only self-selection as a hurdle to participate. A session lasted for around 3 h and the problems mentioned here were solved within two sessions. We believe this offers a partial answer to the second goal mentioned in the introduction, i.e. it shows that this model is useful in mathematics education.

To a certain degree, this concept comes into conflict with the way we normally teach mathematics. Original mathematical results from pupils are very rare. We would like to understand this *novelty* relative to the knowledge background of the person. This should be contrasted with the mere application of known procedures. This perspective has received some attention in the discussions in mathematics education, e.g. famously within Freudenthal's "Realistische Mathematische Erziehung" (Realistic Mathematical Education). One of his key-components was the idea of "guided rediscovery" (see Freudenthal (1991) or Treffers (1987), so that it feels like a genuine discovery for the pupil. It is a math-educational task to find problems which are indeed solvable by the pupils, by successfully choosing the level of difficulty which is challenging but not yet frustrating. This means in a way that harder problems become tangible without increasing the difficulty so rapidly that the material stops being engaging. This is indeed one key aspect of pursuit-worthiness, but unfortunately it is often only accessible in hindsight. This is the main cost in a "economy of research" discussed by Achinstein (1993) for pure mathematics. We only have limited time to do research and would like not to invest this time in dead ends. The amount of other costs, like cost of tools, is very limited in pure mathematics. Similarly, we would not like to demotivate the children and waste their limited resources of 'willingness to engage with the material'. We look at open problem fields as something where the discovery of new questions is part of the activity. This aligns much more with mathematical practice, where a large part of the community states new conjectures and problems quite often.

We reduced questions of problem posing to the questions of rephrasing given questions. Of course, we could try to make this textual proximity more precise. There is a variety of metrics for textual proximity, like the Levenshtein-Distance. However, it is more about the fact that this can be learned without any discourse about inner theoretical virtues of questions. The concrete questions that are actually asked are of course again dependent on the background knowledge of the practitioners. Have they made good experiences with one kind of variation? Have they learned a tool other mathematicians do not know. Was the given proof of the initial question of a type that can often extended to the new domain? It remains a empirical question how good these variations can be judged in advance but there is one important point to be made here. If the 'big names' decide which mathematics is valuable, maybe even indirectly by employing naturally people asking questions of their interest, then this creates a feedback loop. A loop where no bad intentions exist, but where marginalized mathematicians will have problems promoting their topics. It is also hard to compare the paved paths with the unpaved. By the very history of the field, we have explored the paved paths more than the others and alternative history is a genre where clear results are complicated anyways. But the theme of re-invention could be a possible case study to compare different paths, like the reconstruction of infinitesimals with modern tools, or toposes with properties, which where hold to be true, but where dismissed (f.i. concerning the continuity of functions).

(P7) Which … admit a tiling of the (3-dimensional) **Euclidean space** using … ?

Let us go back to our case study and revisit for a moment. We mentioned that this is problematic, but this also forces a conceptual innovation upon us. A constructive point of view would be to notice the limits of the given context and its definitions. In this case, we would need to consider the meaning of tiling and its dependencies regarding the other parts of the question, namely the plane, a rectangle, a unit disc or the space. We would deal with these contextual dependencies and realise the need for adaption of the involved definitions. While the generation of new questions might not seem a creative act per se because of its mechanical appearance as an application of our proposed text-driven variation method, the subtle definitions or concepts introduced in order to resolve tensions might give rise to new fruitful notions and fields to study. This mirrors the way actual research develops. Thus, in a quite unpredicted way, our variation process resulted in the creation of something new. It is an act of creativity. In this case one could start to ask which n-gons could form a polyhedra, i.e. a shape with flat polygonal faces, straight edges and sharp corners or vertices. We could also ask which polyhedra could be space filling as an analogous question to the two-dimensional case.

But how do we evaluate the value of such new concepts? Here there are some similarities to the other sciences. McMullin, (1976) or (Nickles 1989, 2006) develop issues on the heuristic value of theories.[7] A simple test case is whether a mathematical concept helps us to organize our knowledge and whether we can prove new things with it. The latter might sound more reasonable but the value of the first should not be underestimated: for instance, we find both aspects in the laudatory speech for the Fields Medal of Peter Scholze

> Moreover, although Scholze has made major additions to the elaborate theoretic foundations of arithmetic geometry, at the same time his ideas have dramatically simplified and clarified our field. This is a characteristic feature of his universal approach to and vision of mathematics. (Raport 2019)

Knowing the value of reorganization lets us also study how our model can help us organizing mathematical knowledge and give a story of developments of proper research in graph theory.

3. The case of organizing mathematical knowledge: A rational reconstruction of an important part of graph theory

As we cannot look into the future to see how long-term projects will develop from now on, we offer an organization of a big body of mathematical knowledge. We choose the field of graph colourings as the notions are rather easy to digest, but we deal with a proper research field. As the developments fit – as we will see in this section – our picture of syntactical variations, we can see how de facto the judgements on pursuit-worthiness of generations of mathematicians fitted the toy model we described.

Gowers makes the point, which he attributes to Atiyah (1974), that:

> So much mathematics is produced that it is not possible for all of it to be remembered. The processes of abstraction and generalization are therefore very important as a means of making sense of the huge mass of raw data (that is, proofs of individual theorems) and enabling at least some of it to be passed on. (Gowers, 2000, p. 68)

This might be overlooked as mathematicians tend to shorten the path to those definitions quite a lot. Thus, it is often unbelievably beautiful, astonishing, and indeed genius how a definition has precisely those features that are needed to develop a fruitful theory when we stumble upon it in a script during our studies.

For the following consideration let us start with two basic definitions.

Definition. A graph G is a pair (V, E) where V and E are sets whose elements are called the vertices and *edges*, respectively, of the graph. Each edge is a set consisting of precisely two different vertices.

Given an edge $e = \{u, v\}$, we say that e connects u and v or that u and v are adjacent. Usually, we think of edges as lines connecting their corresponding vertices.

Definition. A vertex-colouring of a graph G is a function which assigns each vertex of the graph a colour. However, we additionally demand that adjacent vertices receive different colours. If a vertex-colouring uses only k colours, for some finite number k, then we call the vertex-colouring also a k-vertex-colouring.

With these definitions we might be tempted to ask, given a graph G, what the minimum number k of colours is such that a k-vertex-colouring of the graph G exists.

(Q1) How few colours k suffice to colour a given graph with a k-vertex-colouring?

This might be an enjoyable puzzle, but to get to a mathematical theory it is often good to search for broader results. While playing with some graphs and vertex-colourings as examples, one might tend, just for the sake of clarity, to draw graphs on a sheet of paper without edges crossing each other. Such graphs are an important and quite special class of graphs, called planar graphs. This then leads to the more restricted question and our first variation.

(Q2) How few colours k suffice to colour a given planar graph with a k-vertex-colouring?

This is already a famous mathematical problem (in disguise). The history was the other way around: maps were considered, and countries coloured. Then graphs were used to model this by associating each country with its own vertex and adding edges between bordering countries. It was observed by Francis Guthrie in 1852, while trying to colour the map of counties of England, that four colours suffice. So, the following conjecture was born.

(C1) Every planar graph can be vertex-coloured with 4 colours.

This conjecture has been verified first in Appel and Haken (1977), Appel et al. (1977), although their proof was quite controversial. Later, shorter proofs appeared for example in Robertson et al. (1996, 1997). So now, the former conjecture is known as the Four-Colour Theorem and it turned out to be a very complicated one, which remained unsolved for over 100 years and for which there is still no proof known that avoids computer aid. So, when working on this problem to find an elementary proof, one might be tempted to relax the question and ask for a 5-vertex-colouring instead.

(C2) Every planar graph can be vertex-coloured with 5 colours.

This is indeed a more tangible problem for which short and elegant proofs are known, see for example Diestel (2017). The first proof reaches back to Kempe (1879), who claimed to have a proof of C1. Later, Heawood (1890) found an error in the work of Kempe, but the methods from that article enabled him to prove C2.

We could now start to vary question Q2 by trying to change a classical parameter, namely the dimension. So instead of drawing a graph on a sheet of paper, which then corresponds to the plane, we might consider higher dimensional spaces instead. However, it easily turns out that every graph can be drawn in the three-dimensional space without any two edges crossing each other. However, there is an infinite amount of further two-dimensional surfaces that differ from the plane. Classical examples might be the torus, the cross-cap or the Klein bottle. So, we arrive at the following variation, which widens the original question again after being restricted by question Q2.

(Q3) How few colours k suffice to colour a given graph that can be drawn in a two-dimensional surface without crossing edges, with a k-vertex-colouring?

We should note here that question Q3 has been solved completely. Heawood (1890) stated a certain function which only depends on the genus (or the Euler characteristic) of the given surface, and then conjectured that this function answers question Q3. The correctness of this function regarding Q3 has later been approved for all surfaces except for the Klein bottle in Ringel and Youngs (1968). In the case of the Klein bottle, only 6 colours are needed while the function by Heawood predicts 7, which had been proved in Franklin (1934).

Now let us introduce another way of varying our initial question. Instead of colouring vertices of a graph, we may colour the edges of that graph.

[7] For more recent case studies outside of mathematics cf. Šešelja and Weber (2012) for the theory of continental drifts.

However, since we did not allow arbitrary assignments of colours to vertices when considering vertex-colourings, we should restrict the allowed assignments of colours to edges as well. As we do not colour two vertices with the same colour if they are connected by an edge, it seems natural to disallow two edges coloured with the same colour, if they are connected via a vertex. It is precisely in this way that the notion of edge-colourings of graphs is defined. Analogously to vertex-colourings, we define, for some finite number k, a k-edge-colouring of a graph G as an edge-colouring of G that uses only k colours. So, we proceed to the following question.

(Q4) How few colours k suffice to colour a given graph with a k-edge-colouring?

So far, we varied the initial question while always considering vertex – or edge – colourings for certain graphs. However, we might be tempted to vary the question by considering objects slightly different from graphs. One example might be directed graphs. These structures are just like graphs, but now edges have a direction, pointing always from one vertex to the other. Differently from graphs, now we might have two different edges between two vertices, since they might be directed in one and the other direction. If we see the direction of an edge as a restriction where moving along that edge is allowed, we no longer connect two vertices by just one edge between them. In order to 'connect' two vertices, say u and v, we rather have to find a path from v to u and vice versa, both respecting the direction of the involved edges. This actually gives rise to a sequence of directed edges such that we can start at any of the two vertices, say v, move along those edges, visit u and end up at v again. Let us call such a sequence a directed closed walk containing u and v.

Following this adapted restriction for vertex-colourings of directed graphs, we note that we might have many edges whose end vertices have the same colour, given such a vertex-colouring of a directed graph, but never a monochromatic directed closed walk containing any two different vertices. Such vertex-colourings were introduced by Neumann-Lara (1982). These colourings are part of current mathematical research (e.g. Hochstättler, 2017) and many questions regarding them, even with respect to fundamental dependencies to other invariants of directed graphs, are still open. So, let us now post the variation of question Q1 regarding directed graphs.

(Q5) How few colours k suffice to colour a given directed graph with a k-vertex-colouring?

Of course, we can try to adapt this question even further by incorporating the above variations for graphs now for directed graphs. This then leads to analogous or new concepts, for example in the case of edge-colourings, since, as for vertex-colourings of directed graphs, we have to define what being connected should mean for two directed edges. Since we only wanted to illustrate the variation process itself, we do not discuss the analogue variation for edge-colourings for directed graphs here.

Another way to vary the initial question Q1 when we focus on changing graphs to a different structure, might be moving on to *hypergraphs*. These structures also consist of a set of vertices, but an edge might consist of more than two vertices. In case all edges contain the same number n of vertices, which is the case for usual graphs where edges contain always two vertices, we call the hypergraph n-uniform. As for directed graphs, we now have to adapt the definition of vertex-colourings. One natural adaptation seems to view two vertices as connected via an edge if they are contained in a common edge. This then gives rise to an adapted version of vertex-colourings by imposing the condition that no two vertices that lie in a common edge are allowed to receive the same colour. So, we arrive at the following question.

(Q6) How few colours k suffice to colour a given n-uniform hypergraph with a k-vertex-colouring?

Here, many follow-up problems arise. One possibility would be to drop the condition of uniformity of the hypergraph. However, finding suitable analogues for planarity and edge-colourings paves the way for new definitions and concepts extending the given theory.

4. Which paths will be paved?

So far, we offered a basic model of how to find new research questions. We did this by showing how parts of mathematical research can (partly) be rationally reconstructed by varying text elements in the formulations of the problems. This can also be used to teach mathematical research beyond what is normally done in schools. This was just an exploration; it remains to be shown that a bigger part of mathematical practice can be reconstructed in this way. It also remains open to define hard criteria for distinguishing the quality of different variations. But we will try to do some first steps here. It might well be that many of the chosen paths after performing text-driven variations lead to nothing too interesting, even after several iterations of the process. Nevertheless, the fruitful ones remain and while the way of their discovery might not be part of scientific knowledge, their use is incorporated in teaching of future generations of mathematicians. With a clean picture of how to use fruitful concepts and experience built up on the use or the concepts, the ideal of a well-constructed notion with a lot of knowledge about its usefulness might rather emerge to people that do not know about the underlying genesis. This is also the reason why we decided to use the cowpath metaphor. As argued by Kant and Sarikaya (2021) a better narrative of mathematics is both describing mathematical practice more adequately and of societal value. Wagner (2022) argued about the importance of stressing the historical developments for similar reasons (to tackle the under and overevaluation of mathematics). We believe that the stressing of the arbitrariness and historicity is helpful to motivate people to engage in mathematics.

For the individual research mathematician doing day-to-day mathematics and publishing good but not the most important papers, the main question is whether a result can be published. In Section 2 we argued that it should be challenging but not too hard to solve the formulated problem. We note two things: First that this is of course a question that depends on the background of each individual. If I know the right tools one question is easier to me than to somebody who does not know those tools. Second, this shows a trade-off caused by the particular reward structure of mathematics (cf. Morris, 2021). More challenging questions should give more credit so it is a question of tactics whether one could invest a lot of time that might be fruitless in terms of credit. This opens up the possibility to use agent based modelling to study the change of trends in mathematics.[8] This does not make every question of the same interestingness, one could for instance study which question is likely to be judged as sufficiently new and complex to be accepted by reviewers (if a proof is found). This does not mean that we should strive for an abstract ideal theory of peer reviewing but of a study of the actual practice. Inglis and Mejia-Ramos (2009) and Inglis et al. (2013) show how different standards are evoked for the criteria of correctness of mathematical proofs but other criteria that go into the reviewing process are much fuzzier, like questions of interestingness and novelty and Löwe (2022) explores what empirical data of the reviewing process could say about the position that is epistemically of an exceptional position. Another mechanism that comes into play is simply prestige surrounding a question. Prestige helps to draw attention to a result. The prestige is a function of many factors: How long is the question already open and how many smart mathematicians worked on the question? For the latter a proxy could be how famous the people who asked this question were. Here prizes play a role as well. A famous case is the Hungarian mathematician Paul Erdős. He is well known for being very productive: he wrote 1500 articles and had many collaborators, and besides his many results, he is famous for the many conjectures he stated. For many of those there are also awards attached to their solution. This begs the question of causality

[8] Cf. Šešelja, 2021 for such an approach for the sciences in general.

but such a track record would make it plausible that this work has a lot of impact even if we would come to believe that the mathematics is rather uninteresting (which we do not!). How big his impact was might be read out as an analysis of papers published within the most prestigious journal of mathematics: the Annals of Mathematics.[9] Erdős' work was crucial for many branches of combinatorics, but his work on combinatorics was never featured in the Annals. We might think this shows how he established the field as we currently easily find articles from combinatorics in the most prestigious journals (while combinatorics might still be underrepresented, see again (Pak, 2012)).

Another point is how accessible the question is (not to speak about its solution). This is most famously the case in number theory and there may be a point about why this field is often regarded as the queen of mathematics.[10] A famous example is Fermat's theorem, which actually follows from a heuristic of variation we mentioned, changing the power within the Pythagorean theorem. Most pupils can understand the question, so all mathematicians can at least understand that here is something to solve. In this case the history of the question is important as well, as it is quite curious: Fermat famously claimed to have found a proof but did not write it down due to lacking space in the margin of the book he found the problem in.[11]

A final aspect is the one we mentioned shortly while discussed the Scholze's results. This does not need to come hand in hand with new mathematics as in the Scholze case. The organization is an important feature and many communities profit a lot from important text books, like the community of Homotopy Type Theory and their magnum opus which might be outdated nowadays, cf. (The Univalent Foundations Program 2013). Or as Freudenthal puts it:

> A systematic textbook is a thing of beauty, a joy for its author, who knows the secret of its architecture and who has the right to be proud of it. (Freudenthal, 1968, p. 7)

Finally, a short note. While we drew a picture of pure mathematics which only cared about the opinions of pure mathematicians, there are of course factors from outside the community that can help some fields/ results. When a result is important to physicists they will quote it, they will learn about it, and this might be a factor contributing to the codification of the theorem.

5. Is this view too simple?

At a first glance all the above seems overly reductionistic. Of course, there are long projects like the *Langlands program* and some devote their life to solve a problem that they encountered when they were children, like in the case of *Fermat's last theorem.* This is more than the (close to) random selection of possible questions by our proposed method. One can be very critical of whether such an approach could understand big developments: Why are mathematicians so interested in Scholze's work, why do so many consider *number theory* as a central branch, and why would everybody love to see a proof of the Riemann hypothesis? However, this article is less about these deliberations and more on the issue of day-to-day business. Mathematicians publish many articles, most of them not earth shaking. Many results are forgotten. Where to draw a line remains an empirical question. Do most mathematician live in the day-to-day business of publishing articles? Did most have at least one proper program carefully contemplating the greater impact of their (possible) results? We stay agnostic on that but wanted to stress the locality of mathematical research. Liljedahl and Sriraman (2006) found two very fitting self-assessments from famous mathematicians: first the person who proved Fermat's last theorem, and second the Hungarian mathematician as well as mathematics educator Pólya, who described mathematical research as follows:

> One stumbles around bumping into the furniture, and gradually, you learn where each piece of furniture is, and finally, after six months or so, you find the light switch. You turn it on, and suddenly, it's all illuminated. Wiles (1997).

> It is like going into an unfamiliar hotel room late at night without knowing even where to switch on the light. You stumble around in the dark room, perceive confused black masses, feel one or the other piece of furniture as you are groping for the switch. Then, having found it, you turn on the light and everything becomes clear. (Pólya, 1962, p. 54)

It is unknown to us whether Wiles was influenced by Pólya when he chose a very similar metaphor. Similarly, Lakatos, in his *Proof and Refutations*, stresses how local reparation plays a role in proving theorems. Famously, he also included aspects of the change of the notions in play. Those changes are often very local in character, like monster bearing, where we disallow concrete counterexamples to a given theorem. This work is very much in the spirit of Pólya and Lakatos. We believe to extend the scope of their work and give a clear heuristic that can and should be implemented in mathematics education. We furthermore try to stress the historicity of value judgements concerning mathematical results and how the history forms which problems are well connected to the rest of mathematics. Those pathways are also self-enforcing, since we are less hesitant to walk on those rather than criss-cross the landscape.

We believe that our philosophical literature is distorted towards an overrepresentation of those earth-shaking results. Apparently, when a long-standing open problem is solved, we may find media coverage even at the level of mainstream news. Examples for this are the solution of Poincaré's conjecture (see e.g. (Overby 2006)), or the (currently considered to be unsuccessful) attempts to prove the ABC-conjecture (see e.g. Chang (2012)). Indeed, some of those proofs come with new techniques as well, opening up many new possibilities to prove further open problems, e.g. elliptic curves have been used to prove Fermat's last theorem. However, this is not the only type of task that mathematicians concern themselves with. It is important to find new problems to work on, state conjectures, introduce new concepts, give definitions or shift the interests of the community by organizing conferences, writing books which summarise the knowledge of the field and make it accessible to a new generation of mathematicians, and much more. All these activities are creative as well. Some of those can be considered as forming the cowpath and others of those are already paving existing paths. The creative part comes due to the shaping of the problem fields and dictating what questions to pursue next. A literature review "realised that no articles addressing both problem posing and creativity have been published before 1994" (Joklitschke, Baumanns and Rott 2019, p. 64). Here we argued how important it is to propose new questions in pure mathematics and that our model is useful to understand this.[12]

The first aspect is that our model captured much of that research practice that is not high-profile. However, we believe that it might capture the high-profile cases as well. Conducting mathematical research or problem solving is a non-trivial task. Many people encounter difficulties during their primary and high school times or in their university studies. Furthermore, people tend to express these problems in several occasions. This seems to be evidence for the claim that mathematics is perceived as a non-trivial task. Thus, we do not want to repeat this narrative and in contrast stress that real mathematical creativity can be (partly) systematised and learned by adapting our toy-model. This does not make mathematical research trivial and in no way possible to automatise (at

[9] Pak (2012).

[10] Anglin (2012).

[11] See for instance Vandiver (1946).

[12] While it has recently been argued how problematic the line between pure and applied mathematics is (cf. Perez-Escobar and Sarikaya 2021 it remains that purer branches do not live that directly under real world constraints, they do not need expensive tools and the self-image devalues applicability often even within other scientific disciplines. For a problematization of the pure/applied distinction in the sciences in general see (Shaw, 2022b).

least not with our current computer science techniques; of course it would be great if this view could actually be deployed by an expert system at some point in the future) but should help to attack the paralysing effect that an overblown picture of mathematical practice tends to have.

Acknowledgments

The first author has been supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (ERC consolidator grant DISTRUCT, agreement No. 648527). The second author is thankful for the financial and ideal support of the Studienstiftung des deutschen Volkes and the Claussen-Simon-Stiftung as well as the Research Foundation Flanders (FWO) [grant number FWOAL950], the paper was finalized during a visit of the second author in Denmark (University of Copenhagen & Technical University of Denmark) funded by a DAAD-Kurzzeitstipendium with the project "Theoretical virtues of conjectures and open questions in mathematical practice". We furthermore would like to thank Louis Bellmann, Sophie Nagler, Armin Rezaia Nia, José Antonio Perez-Escobar, and Mira Sarikaya for useful comments. The views stated here are not necessarily the views of the supporting organizations and individuals mentioned in this acknowledgement.

References

Achinstein, P. (1993). How to defend a theory without testing it. *Midwest Studies In Philosophy, 18*, 90–120. https://doi.org/10.1111/j.1475-4975.1993.tb00259.x. Philosophy Documentation Center.

Anglin, W. S. (2012). *The queen of mathematics: An introduction to number theory, 8*. Springer.

Appel, K., & Haken, W. (1977). Every planar map is four colorable. Part I. Discharging. *Illinois J. Math, 21*, 429–490.

Appel, K., Haken, W., & Koch, J. (1977). Every planar map is four colorable. Part II. *Reducibility, Illinois J. Math., 21*, 491–567.

Aspray, W., & Kitcher, P. (Eds.). (1988). *History and philosophy of modern mathematics. Minnesota studies in the philosophy of science: XI* (pp. 3–57). Minneapolis: University of Minnesota Press.

Atiyah, M. F. (1974). How research is carried out. *Bull. I.M.A., 10*, 232–234.

Buldt, B., Löwe, B., & Müller, T. (2008). Towards a new epistemology of mathematics. *Erkenntnis, 68*(3), 309–329.

Carl, M., Cramer, M., Fisseni, B., Sarikaya, D., & Schröder, B. (2021). How to frame understanding in mathematics: A case study using extremal proofs. *Axiomathes, 31*, 649–676.

Carter, J. (2019). Philosophy of mathematical practice — motivations, themes and prospects. *Philosophia Mathematica, 27*(1), 1–32.

Chang, K. (2012). A possible breakthrough in explaining a mathematical riddle. *The New York Times*. Online accessible via: https://www.nytimes.com/2012/09/18/science/possible-breakthrough-in-maths-abc-conjecture.html.

Corfield, D. (2003). *Towards a philosophy of real mathematics*. Cambridge: Cambridge University Press.

Crumlish, C., & Malone, E. (2009). *Designing social interfaces*. O'Reilly.

Diestel, R. (2017). *Graph theory*. Berlin, Heidelberg: Springer Berlin Heidelberg Imprint Springer.

Fisseni, B., Sarikaya, D., Schmitt, M., & Schröder, B. (2019). How to frame a mathematician. In S. Centrone, D. Kant, & D. Sarikaya (Eds.), *Synthese library (studies in epistemology, logic, methodology, and philosophy of science): 407. Reflections on the foundations of mathematics*. Cham: Springer.

Franklin, P. (1934). A six color problem. *J. Math. Phys., 13*, 363–379, 1934.

Freudenthal, H. (1968). Why to teach mathematics so as to be useful? *Educational Studies in Mathematics, 1*, 3–8.

Freudenthal, H. (1991). *Revisiting mathematics education. China lectures*. Dordrecht: Kluwer Academic Publishers.

Gowers, W. T. (2000). The two cultures of mathematics. In V. Arnold, M. Atiyah, P. Lax, & B. Mazur (Eds.), *2000. Mathematics: Frontiers and perspectives* (pp. 65–78).

Heawood, P. J. (1890). Map-colour theorems. *The Quarterly Journal of Mathematics, 24*, 332–338.

Heuer, K., & Sarikaya, D. (2019). Group theory via symmetries for enrichment classes for gifted youth. In M. Nolte (Ed.), *Including the highly Gifted and creative students – current Ideas and future directions Proceedings of the 11th international Conference on mathematical Creativity and giftedness (MCG 11)* (pp. 257–263).

Hochstättler, W. (2017). A flow theory for the dichromatic number. *European J. Combin., 66*, 160–167.

Inglis, M., & Mejia-Ramos, J. P. (2009). The effect of authority on the persuasiveness of mathematical arguments. *Cognition and Instruction, 27*(Issue 1), 25–50 (Informa UK Limited).

Inglis, M., Mejia-Ramos, J. P., Weber, K., & Alcock, L. (2013). On mathematicians' different standards when evaluating elementary proofs. *Topics in Cognitive Science, 5*(Issue 2), 270–282 (Wiley).

Joklitschke, J., Baumanns, L., & Rott, B. (2019). The intersection of problem posing and creativity: A review. In M. Nolte (Ed.), *Including the highly Gifted and creative students – current Ideas and future directions Proceedings of the 11th international Conference on mathematical Creativity and giftedness: 59–67*.

Kant, D., Pérez-Escobar, J. A., & Sarikaya, D. (2021). Three roles of empirical information in philosophy: Intuitions on mathematics do not come for free. *KRITERION – Journal of Philosophy, 35*(3), 247–278.

Kant, D., & Sarikaya, D. (2021). Mathematizing as a virtuous practice: Different narratives and their consequences for mathematics education and society. *Synthese, 199*, 3405–3429. https://doi.org/10.1007/s11229-020-02939-y

Kempe, B. (1879). On the geographical problem of the four colors. *Amer. J. Math., 2*, 193–200.

Kießwetter, K. (1985). Die Förderung von mathematisch besonders begabten und interessierten Schülern - ein bislang vernachlässigtes sonderpädagogisches Problem. *Mathematisch-naturwissenschaftlicher Unterricht, 38*(5), 300–306.

Kießwetter, K. (2009). Was sollte und was kann Hochbegabtenförderung im Bereich Mathematik leisten? In S. Schiemann (Ed.), *Talentförderung mathematik* (pp. 43–69).

Kitcher, P. (1985). *The nature of mathematical knowledge*. Oxford: Oxford University Press.

Kitcher, P. (1990). The division of cognitive labor. *Journal of Philosophy, 87*(1), 5–22.

Lichtenstein, E. I. (2021). Mis)Understanding scientific disagreement: Success versus pursuit-worthiness in theory choice. *Studies In History and Philosophy of Science Part A, 85*, 166–175 (Elsevier BV).

Liljedahl, P., & Sriraman, B. (2006). Musings on mathematical creativity. *For the Learning of Mathematics, 26*(1), 17–19.

Löwe, B. (2016). Philosophy or not? The study of cultures and practices of mathematics. In S. Ju, B. Löwe, T. Müller, & Y. Xie (Eds.), *Cultures of mathematics and logic, selected papers from the conference in Guangzhou, China, 9-12 November 2012* (pp. 23–42). Basel: Birkhäuser.

Löwe, B. (2022). Measuring the agreement of mathematical peer reviewers. *Axiomathes, 32*(3), 1205–1219.

PhiMSAMP. Philosophy of mathematics: Sociological aspects and mathematical practice. In Löwe, B., & Müller, T. (Eds.), *Texts in philosophy, 11*, (2010). London: College Publications.

Maddy, P. (1997). *Naturalism in mathematics*. Oxford: Oxford University Press.

Maddy, P. (2007). *Second philosophy. A naturalistic method*. Oxford: Oxford University Press.

Mancosu, P. (Ed.). (2008). *The philosophy of mathematical practice*. Oxford: Oxford University Press.

McMullin, E. (1976). The fertility of theory and the unit for appraisal in science. In R. S. Cohen, P. K. Feyerabend, & M. Wartofsky (Eds.), *Essays in Memory of Imre Lakatos. Reidel* (pp. 395–432).

Morris, R. L. (2021). Intellectual generosity and the reward structure of mathematics. *Synthese, 199*, 345–367.

Neumann-Lara, V. (1982). The dichromatic number of a digraph. *Journal of Combinatorial Theory - Series B, 33*, 265–270.

Nickles, Thomas (1989). *Heuristic appraisal: A proposal. Social Epistemology 3, 3*, 175–188.

Nickles, T. (2006). Heuristic Appraisal: Context of Discovery or Justification? In J. Schickore, & F. Steinle (Eds.), *Revisiting Discovery and Justification* (pp. 159–182). Springer.

Nolte, M. (2002). Förderansätze für mathematisch besonders begabte Grundschulkinder. In H. L.f. Pädagogik (Ed.), *Grundlagen - Förderkonzepte und Praxisbeispiele - Unterstützungsangebote: 10. Besondere Begabungen - eine Herausforderung für Lehrerinnen und Lehrer*.

Nolte, M. (2012). Mathematically gifted young children - questions about the development of mathematical giftedness. In H. Stöger, A. Aljughaiman, & B. Harder (Eds.), *Talent Development and excellence* (pp. 155–176).

Nolte, M., & Pamperien, K. (2017). Challenging problems in a regular classroom setting and in a special foster programme. *ZDM, 49*(1), 121–136.

Overby, D. (2006). Elusive proof, elusive prover: A new mathematical mystery. *The New York Times*. Online accessible via https://www.nytimes.com/2006/08/15/science/15math.html.

Pak, I. (2012). *How do you solve a problem like the Annals?* Online accessible via: https://igorpak.wordpress.com/tag/the-annals-of-mathematics/.

Pérez-Escobar, J. A. (2022). Showing mathematical flies the way out of foundational bottles: The later Wittgenstein as a forerunner of Lakatos and the philosophy of mathematical practice. *KRITERION – Journal of Philosophy, 36*(2), 157–178.

Pérez-Escobar, J. A., & Sarikaya, D. (2021). Purifying applied mathematics and applying pure mathematics: How a late Wittgensteinian perspective sheds light onto the dichotomy. *European Journal for Philosophy of Science, 12*(Issue 1). https://doi.org/10.1007/s13194-021-00435-9. Springer Science and Business Media LLC.

Pólya, G. (1962). *Mathematical discovery: On understanding, learning, and teaching problem solving, 1965*. New York: John Wiley.

Rapoport, M. (2019). *The work of Peter Scholze. In proceedings of the international congress of mathematicians (ICM 2018). International congress of mathematicians 2018*. WORLD SCIENTIFIC. https://doi.org/10.1142/9789813272880_0004

Ringel, G., & Youngs, J. W. T. (1968). Solution of the Heawood map-coloring problem. *Proceedings of the National Academy of Sciences of the United States of America, 60*, 438–445.

Robertson, N., Sanders, D. P., Seymour, P. D., & Thomas, R. (1996). A new proof of the four colour theorem. *Electronic Research Announcements of the American Mathematical Society, 2*, 17–25.

Robertson, N., Sanders, D. P., Seymour, P. D., & Thomas, R. (1997). The four colour theorem. *Journal of Combinatorial Theory - Series B, 70*, 2–44.

Šešelja, D. (2021). Exploring scientific inquiry via agent-based modelling. *Perspectives on Science, 29*(4), 537–557.

Šešelja, D., & Weber, E. (2012). Rationality and irrationality in the history of continental drift: Was the hypothesis of continental drift worthy of pursuit? *Studies In History and Philosophy of Science Part A, 43*(Issue 1), 147–159. https://doi.org/10.1016/j.shpsa.2011.11.005. Elsevier BV.

Shaw, J. (2022a). On the very idea of pursuitworthiness. *Studies in History and Philosophy of Science, 91*, 103–112. https://doi.org/10.1016/j.shpsa.2021.11.016. Elsevier BV.

Shaw, J. (2022b). Revisiting the basic/applied science distinction: The significance of urgent science for science funding policy. In *Journal for general philosophy of science*. Springer Science and Business Media LLC. https://doi.org/10.1007/s10838-021-09575-1.

The Univalent Foundations Program. (2013). *Homotopy type theory: Univalent foundations of mathematics*. Institute for Advanced Study.

Treffers, A. (1987). *Three dimensions: A model of goal and theory description in mathematics instruction - the Wiskobas project*. Dordrecht: Kluwer Academic Publishers.

Van Kerkhove, B., & Van Bendegem, J. P. (Eds.). (2007). *Perspectives on mathematical practices. Bringing together philosophy of mathematics, sociology of mathematics, and mathematics education. Logic, epistemology, and the unity of science: 5*. Berlin: Springer.

Vandiver, H. S. (1946). Fermat's last theorem: Its history and the nature of the known results concerning it. *The American Mathematical Monthly, 53*(10), 555–578.

Wagner, R. (2022). Mathematical consensus: A research program. *Axiomathes, 32*(3), 1185–1204. https://doi.org/10.1007/s10516-022-09634-2

Wiles, A. (1997). *The proof*. NOVA, 1997, updated 2000. Aired on PBS on October 28, 1997, online available via: http://www.pbs.org/wgbh/nova/proof/wiles.html. (Accessed 5 January 2003).