**Beauty, Truth and Understanding**

(in *The Aesthetics of Science: Beauty, Imagination and Understanding*)

Milena Ivanova

**1. Introduction**

Many scientific theories have been praised for their aesthetic qualities. Newtonian mechanics, Einstein’s theory of relativity are given as examples of a beautiful theory. The beauty of scientific theories is often used in the evaluation of their likelihood of being true or in the estimation of their expected empirical success. That is, often scientists place epistemic import on the aesthetic values of theories, deciding whether to commit to a theory in light of its aesthetic appeal, especially in situations when sufficient empirical data is not available to guide such a decision. The question then arises whether we can trust aesthetic considerations to be playing an epistemic role in science and informing our attitudes towards scientific theories.

 In this chapter I outline accounts that have defended the epistemic role for beauty and aesthetic values in science, claiming that there is a link between an aesthetically appealing theory and its likelihood to be true. After challenging the plausibility of these accounts, I turn to an alternative defence for the relevance and importance of aesthetic considerations in science. It is argued that science has many goals, truth and empirical success being the usual favourites, but it also aims at offering understanding of phenomena and such understanding can be achieved in the absence of truth. By focusing on the concept of understanding, I argue that aesthetic factors are intricately linked to our own cognitive make up and desire to understand the world around us, shaping our inferential patterns and guiding the construction and acceptance of scientific theories.

**2. Beauty in the practice of science**

Just like many different forms of artwork can be praised as beautiful or aesthetically appealing, in science many different objects are claimed to be beautiful or possess aesthetic properties. We find beauty in the phenomena, such as solar eclipses, Aurora Borealis, diffracted light rays, pictures of oscillating particles, scientific models like the structure of DNA molecules, mathematical proofs like Euclid’s Elements, experiments like Rutherford’s explorations of uranium radiation, and, of course, scientific theories. In this paper, my focus is primarily on the aesthetic properties of scientific theories. In what way have theories been considered beautiful? There is certainly a plurality of ways in which a theory can be beautiful. A theory can be simple, elegant, unified, obey symmetry principles, be consistent and coherent. Theories can exemplify these aesthetic qualities or virtues in multiple ways. For instance, Newtonian mechanics can be regarded as simple because it can describe the motion of bodies with three rather simple laws of motion and a law of gravitation featuring few parameters. The general theory of relativity is certainly more complex than Newton’s theory of gravity, the number of equations is larger and there are more parameters in them. But the theory offers a much more economical way of understanding the concepts involved in describing motion and forces by unifying gravity and inertial mass, the concepts of space and time into a spacetime continuum, so in this way it provides us with simpler and more unified understanding of gravitational phenomena. Mendeleev’s periodic table elegantly classifies all 117 elements by their common property of atomic number.

 Beauty is often seen as a motivator in scientific enquiry, but is also given significant epistemic weight by being regarded as an indicator of the theory’s truth. When it comes to its motivational role, scientists often claim that they study nature in order to find the beauty within it. During the heated debates in France at the turn of the 20th century regarding the aim of science and scientific progress, aiming to oppose science skepticism motivated by the argument from the ‘bankruptcy of science’, Henri Poincaré (2001) offers a defence of science appealing to its beauty. Poincaré argues that we cannot justify the idea that science is valuable in its own right, since we need to take practical constraints into account, but we cannot regard science as valuable only insofar as it delivers us products, such as technological advances. Science is valuable, he argues, because in constructing scientific theories and uncovering underlying relations in the phenomena, we experience an aesthetic response. As Poincaré states, "[t]he scientist does not study nature because it is useful to do so. He studies it because he takes pleasure in it, and he takes pleasure in it because it is beautiful" (Poincaré 2001, p. 368). Poincaré clarifies how he understands this concept of beauty: “I am not speaking, of course, of the beauty which strikes the senses, of the beauty of qualities and appearances. I am far from despising this, but it has nothing to do with science. What I mean is that more intimate beauty which comes from the harmonious order of its parts, and which pure intelligence can grasp” (ibid.). In similar fashion, the Nobel laureate Subrahmanyan Chandrasekhar also claims that “in the arts as in the sciences, the quest is for the very same elusive quality: beauty” (Chandrasekhar 1987, p.52).

 Many scientists have appealed to the concept of beauty when providing justification of their commitment to a theory. Often, the aesthetic features of a theory, such as simplicity and unity, are seen as convenient considerations in theory choice. As Ernst Mach (1984) argues, we should always aim to explain the phenomena in the most conceptually economical manner, because simple theories are easier to use. He argues that science should be seen an economical description of our observations, with simplicity being a guiding principle in the construction and evaluation of scientific hypotheses. In the *Science of Mechanics*, Mach argues that: “The aim of science is to be an ‘economy of thought’, to offer us a simple and concise classification of the observable appearances and enable the prediction of phenomena” (Mach 1984, p.577). in a similar manner, Poincaré argues that there is a relationship between utility and simplicity, in that a simpler theory is easier to use, making simplicity a regulative principle in the construction and choice of theories. According to him, “care for the beautiful leads us to the same selection [of theories] as care for the useful. Similarly economy of thought, that economy of effort which, according to Mach, is the constant tendency of science, is a source of beauty as well as a practical advantage” (Poincaré 2001, p.369).

 In addition to its heuristic role, beauty is often taken to stand in a special epistemic relationship to truth. Many scientists hold that a beautiful theory is more likely to be true, so if faced with a choice between two theories, the simplest or more beautiful theory is to be epistemically privileged. Paul Dirac is often stated to have adopted such an epistemic role for beauty in science, claiming that “one has a great confidence in [a] theory arising from its great beauty, quite independently of its detailed successes” (Dirac 1980, p.40). Dirac takes beauty to be an indicator of the theory’s truth, such that we can be confident in the truth of a beautiful theory independently of whether the theory is supported by the empirical data. He claims that “[o]ne has an overpowering belief that [the theory’s] foundations must be correct quite independently of its agreement with observation” (ibid.). Dirac claims that one had good ground to believe in the truth of the general theory of relativity even prior to Arthur Eddington’s expeditions, which were regarded as its first empirical confirmation, on the grounds of the theory’s beauty. This conviction was shared by many physicists at the time, including Eddington himself as well as Werner Heisenberg, who argued that “[i]f nature leads us to mathematical forms of great simplicity and beauty we cannot help thinking that they are “true”, that they reveal a genuine feature of nature” (Heisenberg 1971, p. 68).

 Chandrasekhar similarly argues that it is reasonable to believe that “we have evidence that a theory developed by a scientist, with an exceptionally well-developed aesthetic sensibility, can turn out to be true even if, at the time of its formulation, it appeared not to be so” (Chandrasekhar 1987, p.64). In his 1979 lecture “Beauty and the Quest for Beauty in Science” presented at the Fermi National Accelerator Laboratory, Chandrasekhar argues that scientists aim to uncover the beauty in nature and this is their primary motivation in science. He articulates an account of beauty, drawing from Francis Bacon’s thoughts that beauty implies some ‘strangeness in the proportion’, understood as unexpected wonder or surprise, and Heisenberg’s definition of beauty as ‘conformity of the parts to one another and to the whole’[[1]](#endnote-1) (ibid., p.70). Drawing upon this definition, Chandrasekhar goes on to explain in detail why Einstein’s theory of general relativity is beautiful, achieving the unification of our fundamental concepts of space and time, and the concepts of matter and motion with “unerring sense for mathematical elegance and simplicity” (ibid., 71). He argues that the discovery of beauty in nature is the most significant of achievements, that it is an “incredible fact that a discovery motivated by a search after the beautiful in mathematics should find its exact replica in Nature” (ibid., p.54).[[2]](#endnote-2)

The employment of beauty is particularly interesting in the context of contemporary particle physics, with the research being driven to discover ‘susy’ (super symmetric particles) and the symmetry principle becoming something of an imperative in the community. The Nobel laureate Marry Gell-Mann, whose use of symmetry principles led to the advancement of the standard model and the discovery of the previously unknown particle now called omega minus, has also argued for the deep connection between beauty and truth (Hossenfelder 2018, p. 37). But while many remain enthusiastic about the fertility of the symmetry principle and building more colliders to discover susy particles, others are on the fence as to whether this principle should be trusted so much in contemporary physics and whether beauty might be a systematic bias in contemporary physics, leading physicists to pursue research programmes that are not fruitful (Hossenfelder 2018).

**3. Can beauty tell us about the truth of our theories?**

The previous section showed that scientists do use aesthetic considerations in their decision-making and in deciding whether to trust a theory. This fact is particularly prevalent in contemporary high-energy physics, with some arguing that beauty claims have misled the physics community into fruitless projects. So should we trust beauty to be an indicator of truth? What grounds could we have to place trust in non-empirical factors such as aesthetic values of a theory? Traditionally the link between beauty and truth has been developed in both the rationalist and empiricist tradition. On the one hand, we can argue that beautiful theories correctly capture facts about the world and their aesthetic qualities latch on to beauty in the world. On the other hand, one can place confidence in a beautiful theory simply because, inductively, beautiful theories seem to have had a good track record of empirical success. In this section I outline both defences and focus primarily on the empiricist approach to defending the link between beauty and empirical adequacy or truth, as developed in the work of James McAllister (1989, 1996).

While rationalist defences of beauty are not particularly popular in the current philosophical literature, they have been employed since antiquity in attempt to show that there is an intrinsic relationship between beauty and truth, and are implicitly assumed by many scientists. In the *Philosophiæ Naturalis Principia Mathematica* Isaac Newton proposed several methodological rules in science, one of which is to avoid superfluous causes in describing nature. The justification for this principle is that nature is itself simple. Such arguments are taken to be circular: our confidence in the claim that nature is beautiful, for instance for its simplicity, is justified upon reflection on our current best theories, which in their turn are meant to be true in light of their reflection of the beauty or simplicity of the world. It is also problematic to assume the beauty of nature a priori, since reflecting upon some of our theories can also lead us to make the contrary conclusion. For instance, if we take the standard model, often claimed to be an ugly and inelegant theory, to be reflecting the nature of reality, we would perhaps have to conclude that the world is not simple or unified as we would like it to be, and perhaps there is more complexity and irregularity in the world. I now turn to discussing empiricist of the role of beauty in science.

Some scientific realists have argued that virtues of theories, such as simplicity, symmetry and scope, are evidential, making a theory possessing such aesthetic properties more plausible. Richard Boyd (1984) for instance argues that theories possessing theoretical virtues that have been instantiated by previously successful theories will be ranked as more plausible. Wesley Salmon (1990) goes further in showing how these virtues can bear upon the confirmation of a theory by showing that the past record of success will ultimately determine the prior probability of a theory. Stathis Psillos (1999) endorses Boyd and Salmon’s account, explaining that

[t]he past record of mature scientific theories can be seen as the background knowledge in the light of which the plausibility of emergent scientific theories can be judged and estimated […] if theories which have not been subjected to ad hoc adjustments have tended to be better supported by the evidence than theories with ad hoc features, then this consideration should be used in assessing the prior probability of other theories, in order to rank higher theories with no ad hoc features. Naturally, finding out which theoretical virtues have been associated with well-confirmed theories can be the outcome only of substantive empirical-historical research” (Psillos 1999, p. 172).

Independent of the scientific realism debate, James McAllister has also offered a track record account of aesthetic values. His classic work *Beauty and Revolutions in Science* (1996) offers a sophisticated account of aesthetic values in science, showing how one could justify confidence in a theory based on its beauty even if the theory has not received sufficient empirical support. The reliability of aesthetic values is supported by appealing to the established aesthetic canon, which is formed upon reflection on the aesthetic properties instantiated by the most empirically successful theories of the past. The argument relies on a mechanism known in the psychology literature as the ‘exposure effect’. Studies performed on subjects have found that an agent’s aesthetic preference towards an object tends to increase with repeated exposure to the object (Cutting 2003)[[3]](#endnote-3). Similarly, scientists learn from their experience, by habituation and exposure, what aesthetic values have been associated with successful theories in forming an established aesthetic canon, which then guides the development and evaluation of new theories. For instance, if we reflect upon our current and past successful theories, we can place high confidence in the value of unity, since Maxwell, Newton and Einstein’s theories have all achieved unification of phenomena. We can then project an expectation that when a theory comes along and has unifying power, we can expect it to be successful since it exemplifies properties conforming to the established aesthetic canon. The aesthetic canon allows us to inductively project that when a new theory exhibits the properties associated with the canon, we have good inductive ground to predict the theory will be empirically adequate or true even if at the time the theory lacks empirical support.

The established aesthetic canon aids inductive inferences, justifying trust in a theory that might at the time lack empirical support. By making an aesthetic induction, based on the aesthetic properties of successful theories of the past, one can infer that a theory that possesses the aesthetic properties associated with the established aesthetic canon is highly likely to enjoy long empirical success and be true. This account is both projectivist and fallibilist, it does not take aesthetic properties to have timeless and objective validity, allowing that new empirically successful theories can completely revise the accepted aesthetic canon, making aesthetic values relative to a particular framework and revisable by empirical advancement. One of McAllister’s main aims is to defend the rationality of science in light of these changes in the aesthetic canon and offer an explanation of why decision making, based on aesthetic values, is rational even in light of a dynamic notion of beauty. If our aesthetic canons are revised in light of empirical findings and the improvement of our scientific theories, confidence in the aesthetic canon can at least be weakly justified inductively. McAllister takes aesthetic discontinuities to occur in theory transitions, and claims that what drives the revision of an aesthetic canon is ultimately the empirical success of the superseding theory. Theory change, and in fact scientific revolutions, are regarded as changes in the aesthetic canon, seeing theory transitions as battles between progressive scientists, who adopt a theory based on its empirical success, and conservative scientists, who delay the acceptance of a new theory because it does not conform to the established aesthetic canon. Such revolutions can result in a particular aesthetic property becoming irrelevant (such is the case of visualisation since the acceptance of quantum mechanics) or with new properties becoming part of the established canon (an example here is symmetries, which because highly desirable since the success of general relativity and quantum mechanics).

McAllister’s account takes scientists’ aesthetic talk literally rather than adopting a reductivist interpretation of aesthetic claims. For McAllister, we should not reinterpret the aesthetic language used by scientists to be reducible to emotive praise of the empirical success of a theory. Rather, aesthetic talk should be taken as discourse about aesthetic features of the theories. I agree with this approach and think we can motivate it on two grounds. First, in addition to the fact that scientists use aesthetic language when they praise theories, there is also evidence they undergo aesthetic experience when presented with a beautiful equation or proof[[4]](#endnote-4). Second, aesthetic considerations feature in the acceptance of theories that have not been confirmed and cannot technically be considered empirically adequate. This is particularly pressing in theoretical physics, in cases such as string theory, where there is not available evidence to support the theory and yet scientists want to place their epistemic commitment to the theory in light of its aesthetic appeal[[5]](#endnote-5). These aesthetic judgements, however, in this case cannot be claimed to reduce to empirical adequacy.

The above account offers a plausible explanation for the role of aesthetic values in science and is well motivated by the historical development of science. An important aspect of this account is that it aspired to address the question concerning the status of aesthetic judgments, whether aesthetic values are objective and fixed across different times and places, or subjective and relative to a particular individual, society or temporal framework. It has been widely recognised in philosophy of science that scientists can differ substantially with respect to how they prioritise theoretical virtues, with some valuing elegance over other virtues, while others preferring the unifying power of a theory, for instance. In such cases theory choice is contingent on the virtues most valued by the particular scientists or community, so there can be disagreement as to which theory should be preferred on the grounds of them exemplifying differently prioritised virtues. In addition, each such theory virtue can receive different interpretations in different contexts, making an objective weighting comparison between the aesthetic values of two competing theories inconclusive.[[6]](#endnote-6) Defending a dynamic conception of beauty can accommodate ~~for~~ the fact that aesthetic values seem to change during theoretical transitions. For instance, the demand for visualisation was slowly abandoned with the increasing success of quantum mechanics. Quantum mechanics has introduced a highly difficult phenomenon to visualise – the superposition of particle states. We cannot visualise a particle in a superposition, this phenomenon is counter intuitive to us, but with the theory being highly successful we can revise our aesthetic canon and no longer regard visualisability as part of this canon, allowing aesthetic appreciation of a theory that is not visualisable. Furthermore, during the Copernican revolution and Kepler’s developments of the heliocentric model of the solar system, the use of perfect circles to describe the movement of the heavenly spheres was slowly replaced by ellipses, as the elliptical models became more predictively successful. This transition was not quick, since the community was reluctant to deviate from the established aesthetic canon, and endorse what was considered an imperfect geometrical shape to describe planetary motion, but such criticism slowly became irrelevant as the elliptical model gained more and more empirical support. McAllister argues that the aesthetic canon guiding the research community is ultimately formed by reflecting on the values possessed by highly successful research projects of past and contemporary science.

Leaving the aesthetic canon open to change and tying it to the empirical performance of theories seems at first sight uncontroversial. However, one can worry about the descriptive merit of this account, in particular whether it adequately describes what happens in scientific practice. The aesthetic induction, while plausible in the noted cases, seems to not be able to explain certain cases in science where some values persist and the community resiliently follows them. Such ‘persistent values’ are appreciated and desired in the community even if contemporary theories have failed to exemplify them. In particular, by following the aesthetic induction we would expect certain values to be abandoned or weakened, as they are not properties of contemporary successful theories. However, such a phenomenon has not always been observed in the scientific community. If aesthetic values are revisable and driven by the empirical success of the theories that exemplify them, we would expect to observe more revisions to our aesthetic canons. However, values such as simplicity and unity hold a centrality in science. Simplicity has been an ideal even before the development of successful theories and continues to be an ideal followed by scientists despite the prominence of successful theories that are not regarded elegant or simple. If the aesthetic canon is driven by the values exemplified by contemporary successful theories, we would expect the appreciation for elegance and simplicity to at least weaken and perhaps be replaces by other values. As Montano (2014) argues, certain aesthetic values have the status of ‘historical constants’, meaning that their appreciation in the community appears to be stable throughout theory change. This puts into question how predictively successful McAllister’s account is and whether exposure to successful theories is a sufficient condition for the formation of aesthetic canons.

A common idea shared by the epistemic accounts of aesthetic values is that aesthetic values are formed by habituation. McAllister (1996), Montano (2014) and Kuipters (2002) each regard the exposure effect as the mechanism responsible for the formation of our aesthetic preferences in science. It is, however, unclear whether exposure is sufficient for the formation of such preferences in science. While exposed to successful complex theories (theories, for example, with great number of equations or free parameters), most physicists do not report increased appreciation of complexity in their theories, on the contrary, they negatively compare complex explanations of phenomena to simple ones. In the case of the standard model, for instance, the community openly expresses the idea that the theory is aesthetically unappealing, despite its enormous empirical success. As Kaku and Thompson argue, “there has been no experimental deviation from the Standard Model. Thus, it is perhaps the most successful theory ever proposed in the history of science. However, most physicists find the Standard Model unappealing because it is exceptionally ugly and asymmetrical. […] The reason why the Standard Model is so ugly is that it is obtained by gluing, by brute force, the current theories of the electromagnetic force, the weak force, and the strong force into one theory” (Kaku and Thompson 1997, p.75). This dissatisfaction with the standard model is due to its failure to satisfy persistent aesthetic values, rather than what the aesthetic induction would predict, which would be that these aesthetic ideals would be revised in light of the success of the standard model.

Similarly in mathematics, as argued in Montano (2014), proofs by cases have not gained the appreciation of mathematicians, despite their successful employment. Montano focuses on case-by-case proofs to challenge McAllister’s prediction that as these proofs obtain a strong record of success, we can expect them to gain aesthetic appeal in the mathematical community. Such appreciation has not been observed. What such computer-assisted proofs do is check case by case which violates the parsimony criterion. In the case of Appel and Haken’s proof of the four colour theorem, the computer assists in checking hundreds of cases. The fact that the proofs violate the desideratum of simplicity, which Montano claims is associated with understanding a proof, is the reason why aesthetic appeal has not been observed despite their successful employment, challenging the prediction made by the aesthetic induction. The conclusion I draw from these two cases is that while empirical success is an important element in the formation of trust in the performance of a theory with specific aesthetic merit, it is not sufficient to explain why certain values seem persistently appreciated while others do not gain aesthetic appreciation despite being exemplified by successful theories.

Finally, my main objection to the above track-record account is methodological. The above arguments have relied on the historical record of successful theories to justify confidence in the truth of contemporary theories that exemplify certain aesthetic properties. I do not take such arguments from the history of science to be sufficient to defend an ‘optimistic’ over a ‘pessimistic’ argument. If we take the current state of the scientific realism debate as an example, we can see that arguments both in defense of and against realism have been made by inductive inferences from the history of science. Anti-realists, following Larry Laudan’s (1984), argue against realism on the grounds that lots of past successful theories have been abandoned, undermining the claim that success can be indicator of truth. And scientific realists take it that no better explanation of the predictive track record of theories can be given unless we take them to be true, at least approximately. Such inductive arguments are, however, inconclusive and cannot pull towards a realist reading of science.

The problem is that there are many aesthetically pleasing theories that have turned out to be false and many aesthetically displeasing theories that fit with the highest requirements of the realist to count as true, at least ‘approximately’. On the one hand, we have the beautiful theories that failed. Copernicus’ heliocentric system was abandoned in favor of Kepler’s despite the initial negative reaction to the idea of abandoning a system based on perfect cycles to epicycles, considered at the time aesthetically inferior. And when it comes to theoretical projects in high-level physics, unificationist projects have been abandoned due to difficulties despite being recognized for their aesthetic appeal. The Kaluza-Klein theory, for instance, provided an elegant way to unify gravity and the other gauge fields but was abandoned despite it being highly regarded due to its elegance and simplicity. Hossenfelder’s (2018) *Lost in Math: How Beauty Leads Physics Astray* is full of illuminating examples from theoretical physics in which theories of great aesthetic merit are abandoned due to their lack of predictive success. Hossenfelder argues that we should be wary of relying on beauty to guide our belief in theories, and that beauty can often be a systematic bias that leads away from developing successful theories. And when it comes to placing confidence in the track record of ‘ugly’ theories, our current best theories in physics – the standard model and quantum mechanics – provide examples of great empirical and predictive success but are rarely used as examples of beautiful theories, quite the contrary.

The problem with drawing inductive inferences from the history of science is that we can find plenty of examples to base an inductive argument on, some in favour of the link between beauty and truth and some against it. Arguments relying on the track record of science have been well utilized in the scientific realism debate, but have not produced a conclusive choice between realism and anti-realism. Similarly in this context, it seems like inferences based on the history of science do not have the power to pull us conclusively in either direction when it comes to the question of whether there is a link between aesthetic values and truth and a further argument is needed to strengthen the realist or anti-realist position.

So far I have examined how one could support the epistemic significance of aesthetic values by inferring the likelihood of a theory from the track record of aesthetic values exemplified by previously successful theories. The objections developed above challenge the idea that there is a link between beauty and truth, but this does not necessitate accepting the idea that aesthetic values are merely heuristic. In the next section I develop an alternative account, arguing for the epistemic significance of aesthetic values. I shall argue that by reconsidering the aims of science and taking our primary aim to be understanding rather than truth, we can recognise the regulative role played by aesthetic values in achieving these epistemic aims.

**4. How beauty guides our understanding**

Science delivers epistemic goods, but these goods are rarely characterised as literally true theories. Often our understanding of phenomena is achieved by constructing models that deviate significantly from the truth, and the history of science teaches us that even our very best theories can turn out to be false. The fact that truth is not present in these ‘layers’ of scientific theorising should not be taken to undermine the epistemic success of science. This is because science has many aims, truth and empirical adequacy are certainly the usual suspects, but science also advances our understanding of the world around us. Such understanding can be achieved with theories that deviate from the truth, but this fact does not undermine their importance for the advancement of our epistemic aims. Before I defend the epistemic role of aesthetic values in science, I would like to further motivate the account of understanding I defend.

In the recent literature in philosophy of science, much focus has been on explaining how we use models and idealisations in science. Much of the practice of making a model in science involves distortion, abstraction and idealisation, which often equates with making knowingly false assumptions. When we construct economic models, we often make assumptions such as the perfect rationality of agents. When we study the behaviour of gases, we often use laws like the ideal gas law, which represents the relationship between pressure, temperature and volume in a gas, while it falsely assumes that gases don't exhibit intermolecular attraction or that gas molecules are spherical. The use of such devices in science has led some philosophers claim that the aim of science is not truth but rather understanding of worldly patterns, and this account can better explain scientific practice (most instructively developed in the work of Catherine Elgin (2018) and Angela Potochnic (2016)).

 Apart from concerns about idealisation in science, claiming that the aim of science is truth has been ridden with some well-known problems, including the problem from theory change. This problem was pressing at the beginning of the twentieth century when systematic philosophical work emerged addressing the aim of science and the limits of scientific knowledge. Reflecting on the constant revisions of theories and the apparent discontinuities occurring during such transitions, philosophers like Pierre Duhem and Henri Poincaré explicitly distanced themselves from the idea that science aims at discovering truths about that world, holding instead that scientific theories aim at prediction and lead to the deeper understanding of the relations between the phenomena[[7]](#endnote-7). Duhem and Poincaré worried as to whether we have good grounds to trust that our theories can reveal the true nature of reality, since many scientific theories once believed to be true descriptions of reality were abandoned and replaced by new ones. This argument, known at the time as the argument from the ‘bankruptcy of science’, was employed to question our optimism that scientific theories can reveal the underlying nature of reality. This problem motivated the position that the aim of science is not to offer us true theories, but to provide understanding of relations in the world.

Distancing ourselves from the traditional ideal in epistemology, that sees science as obtaining truths about the world, is also motivated by the desire to make sense of the status of first principles in scientific theories. These principles are the fundamental building blocks of our theories, but are often considered to be neither true or false. Many of the founding principles in science, the very ‘first’ principles, on which scientific theories are built, often receive their justification not on the grounds of experience or a priori reasoning, but by convention. The rationality principle in economics, the laws of motion in Newtonian mechanics, the light postulate in the special theory of relativity theory, and the equivalence principle in the general theory of relativity are often taken to be neither empirically true nor a priori; they lack truth value in isolation. This account of principles makes it difficult to take theories are true descriptions of reality, since the very basic principles on which these theories are built to lack truth-value. There has been a lot of attention on how central principles of science are justified and what the epistemic implications are if we claim science is built on principles that are established on conventional grounds. However the status of these principles is settled, it is difficult to claim that our best scientific theories are approximately true, since the building blocks of these theories have no truth value. This problem further motivates us to reconsider the plurality of aims in science and focus our attention on epistemic goods other than truth, such as understanding.

The idea that we can have understanding without truth may appear deeply problematic at first sight. According to traditional epistemic accounts of understanding, truth is a necessary condition for understanding, and a kind of propositional knowledge. This account, however, is too restrictive to accommodate for the success of science. Nonfactive accounts of understanding, developed recently by Elgin (2009, 2018) and deRegt (2004), treat understanding not as a kind of knowledge but rather as a skill or ability to grasp how the facts fit together. On the face of it, there is a substantial difference between knowing a set of propositions or laws, and being able to apply these laws in different situations. This ability to ‘see’ how certain propositions relate to each other and apply them to different contexts is considered an essential element of understanding. Also, understanding does not always require the belief in true propositions. One’s ability to use Newtonian mechanics to calculate the gravitational pull of the Moon requires understanding of how to employ Newton’s laws together with a set of initial conditions in order to derive an answer. If truth were a presupposition for having understanding of gravitational effects, one would lack understanding in this case, given that the theory in play is strictly speaking false. As Henk de Regt (2015) argues, if truth is a necessary condition for understanding, it would follow that past scientists lacked understanding of phenomena for which they had advanced empirically successful (but from our perspective false) theories. Separating the two concepts and taking understanding not to require true theories, but rather an ability to manipulate and use a theory in a certain theoretical domain, avoids this problem and leads to a more plausible thesis of the aims in science and scientific progress.

 In the previous section I argued that aesthetic values can not serve as predictors or indicators of a theory’s truthlikeness. But that is not to say their use in science is not well motivated. They do in fact play a regulative role in achieving the aim of science, namely understanding the phenomena. Here I draw upon some of Poincaré’s thoughts on beauty in science recently reconstructed in Ivanova (2017). According to this account, the beauty in our theories -their elegance or unity- does not reflect some elegance or harmony in the world. Rather, they are properties of our theories only because we decided to construct them in such a way, and we decide to construct them this way because it is most convenient for us to operate with theories that satisfy our aesthetic and intellectual requirements. Poincaré explicitly argues against the idea that aesthetic values can be truth indicative. Questions like ‘is nature itself beautiful’ are not within the empirical method to address. On the contrary, aesthetic values such as simplicity and unity should not be taken to be properties of nature itself, but conditions of our making, regulating ideas in scientific inquiry. When we construct theories that conform to the principles of unity and simplicity, we are following ideals reflective of our cognitive makeup: “[i]n formulating a general, simple, and formal law, based on a comparatively small number of not altogether consistent experiments, we have only obeyed a necessity from which the human mind cannot free itself” (Poincaré 2001, p. 100). He further claims that this ‘harmony’ we experience when we reflect on our theories “is at once a satisfaction of our aesthetic requirements, and an assistance to the mind which it supports and guides” (ibid., p. 396-397). Thus, following these aesthetic values allows us to construct theories fit for achieving the goal of science, to offer us understanding of relations between the phenomena.

Taking aesthetic values in science as conditions of our cognitive makeup reflecting our intellectual interests and capacities explains why aesthetic values persist even when the best theories do not seem to quite fit our aesthetic requirements. Physicists search for a grand unifying theory because even though we have successful theories of the different forces, they are convinced our understanding of the world would be advanced by a unifying theory. Theorists in chemistry search for the best organisation of the elements even though they know all the elements, their atomic properties and structural patterns. What guides their enquiry is the idea that a more elegant representation of the elements would provide more understanding[[8]](#endnote-8). A theoretical description fitting with our aesthetic ideals seems better disposed to attain our epistemic aims.

This tendency to prefer simplicity and scope in our explanations seems deeply ingrained in human thinking and has been systematically studied in psychology and neuroscience. For instance, studies in cognitive psychology focused on an agent’s preferential inferences supports the idea that agents strive to develop and consistently support simpler explanations of broader scope. Tania Lombrozo explains that “when children and adults generate and evaluate explanations, they recruit explanatory virtues, such as simplicity and breadth, as evaluative constraints on reasoning. As a result, they are more likely to generate and favor broad and simple hypotheses, and to discover broad and simple patterns” (Lombrozo 2016, p.749). Our desire for simplicity and unity affects both our preferences for what theories we construct and what patterns in nature we focus to study and identify.

This is not to say that if systematically in error, we would continue to prefer the aesthetically pleasing theory. A pretty idea that is useless would not get us far. But there is certain resistance in accepting complexity over simplicity, narrowed and more limited scope to breath. This fact is also nicely highlighted by Lombrozo’s studies, arguing that agents continue to prefer simpler explanations even after they learn about the likelihood of the simple and more complex explanation and it takes a rather significant decrease of likelihood before they consider the plausibility of the complex explanation. She argues that: “In one set of studies, participants selected between explanations for an individual's symptoms that appealed to either a single common cause (simple) or to two independent causes (complex), and they were additionally provided with information about the base rate of each disease. Under these conditions, explanation choices were sensitive to both simplicity and probability. When two candidate explanations were equally likely, a majority of participants selected the simpler explanation. It was not until the complex explanation was 10-fold more likely than the simpler alternative that a majority of participants selected it as the most satisfying explanation for the symptoms. It thus appears that, when the probability of an explanation is uncertain, the relative simplicity of an explanation has a significant effect on its perceived quality” (ibid., p. 751). This phenomenon can be seen in science: when aesthetically pleasing explanations systematically fail, we move on and work with theories that do not fit these requirements, but that is not to say that the search for simplicity and elegance ends.

This point can be further illustrated by considering complexity theory, where simple explanations are likely to be useless and perhaps aesthetic appreciation for complexity can be developed. Here I again draw lessons from Poincaré, whose work on the three-body problem led him to consider how complexity theory fits with the idea of simple and unified science. In the case of simplicity, for instance, he claims that apparent simplicity can conceal deep complexity in the phenomena so it cannot be considered a reliable indicator of truth: “[a] century ago it was frankly confessed and proclaimed that nature loves simplicity; but nature has proved the contrary since then on more than one occasion. We no longer confess this tendency, and we only keep of it what is indispensable, so that science may not become impossible” (Poincaré 2001, p. 100). Poincaré argues that even though simplicity might not be a property of nature, a question we cannot answer, it nevertheless has to be a property of our theories and we should always try to generalise in the simplest possible ways: *“*those who do not believe that natural laws must be simple are still often obliged to act as if they did believe it. They cannot entirely dispense with the necessity without making all generalisations, and therefore all science, impossible. It is clear that every law can be generalised in a number of ways, and it is a question of choice. The choice can only be guided by considerations of simplicity” (Poincaré 2001, p.113). I thus take it that our experience of beauty, when engaging with a theory, stems from the ability of the theory to present a complex phenomenon in a simple way and give us insight into further phenomena and their relations which give us understanding of the world.

This brings me to my conclusion. That scientists use aesthetic values in the conception and evaluation of theories is a fact. That they credit theories on non-empirical grounds is also a matter of fact. The more remote from empirical data a discipline is, the more its proponents appeal to beauty as a source of justification. These systematic preferences are quite clear in every level of scientific theorising, but it is misleading to claim that such aesthetic judgments reflect anything about the world in itself, they should be seen as reflections of our own intellectual capacities and regulative ideals guiding our epistemic aims. The quest for simplicity and unification is well grounded in our desire to understand the world in which we operate and to construct theories we can easily employ and operate. But the fact we often succeed in constructing such highly successful theories with aesthetic appeal is not a reason to infer that this beauty is a reflection of a beautiful orderly nature, as nature might be very far from this ideal.

**5. Conclusion**

One way to think about aesthetic values in science has been to dismiss them as subjective and irrelevant factors in the context of justification. Another has been to show not only that such values have some heuristic role to play in the context of discovery and justification, but that they can be trusted to indicate the likelihood of a theory’s empirical adequacy or even truth. I have discussed the insights drawn from such empirical accounts as well as some difficulties. In this chapter I have explored a different way to address the question of whether aesthetic values can play an epistemic role in science by arguing that such values can be seen as requirements we impose upon theories for our intellectual purposes. The prominence of aesthetic values in scientists decision-making can be seen as part of our way of thinking about the world that are deeply dependent upon our capacities as agents, independently of the question whether such values can in any way guarantee that the theories we regard as beautiful are also likely to be successful.

**Acknowledgments**

I am very thankful to Steven French, Alice Murphy and Matt Farr for their insightful comments on the earlier draft of this paper. This work was presented at the universities of Bern, Cambridge, Durham, Edinburgh, Exeter and Leeds. I am grateful to the audiences for their questions and suggestions.

**Bibliography**

Chandrasekher, S. (1987), *Truth and Beauty: Aesthetics and Motivation in Science*. The University of Chicago Press

Cutting, J. (2003), ‘Gustave Caillebotte, French Impressionism, and Mere Exposure’, *Psychonomic Bulletin & Review*, 10: 319–43.

Dawid, R. (2013), *String Theory and the Scientific Method*. Cambridge University Press.

De Regt, H. (2015), ‘Scientific understanding: truth or dare?’ *Synthese* 192: 3781-3797.

Dirac, P.A.M. (1980), ‘The excellence of Einstein's theory of gravitation’, in M. Goldsmith, A. Mackay, & J. Woudhuysen (Eds.). *Einstein: The First Hundred Years*. Oxford: Pergamon Press.

Duhem, P. (1954[1906]), *The Aim and Structure of Physical Theory*. Princeton: Princeton University Press.

Elgin, C. (2018), *True Enough*. MIT Press.

Elgin, C. (2009), ‘Is Understanding Factive?’, in D. Prichard, A. Millar and A. Haddock (eds.) *Epistemic Value*. Oxford University Press

Heisenberg, W. (1971), *Physics and Beyond: Encounters and Conversations*. Translated by Arnold J. Pomerans. George Allen and Unwin.

Herfeld, C. and Ivanova, M. (forthcoming), ‘Principles in Science: Reflecting on the Justification and Epistemic Status of Principles Across the Sciences’, Special Issue on First Principles in Science. *Synthese*

Hossenfelder, S. (2018), *Lost in Math: How Beauty Lead Physicists Astray*. Basic Books

Ivanova, M. (2010), ‘Pierre Duhem’s good sense as a guide to theory choice’, *Studies in the History and Philosophy of Science*, 41: 58–64.

Ivanova, M., Paternotte C. (2013), ‘Theory Choice, Good Sense and Social Consensus’, *Erkenntnis,* 78: 1109-1132

Ivanova, M. (2014), ‘Is There a Place for Epistemic Virtues in Theory Choice?’, in A. Fairweather (ed.), *Virtue Epistemology Naturalized*. Synthese Library, Vol. 366, pp. 207-226

Ivanova, M. (2015), ‘Conventionalism About What? Where Duhem and Poincaré Part Ways’, *Studies in the History and Philosophy of Science*, 54: 80-89.

Ivanova, M. (2017), ‘Poincaré’s aesthetics of science’, *Synthese* 194: 2581-2594

Kuhn, T. (1977), ‘Objectivity, Value Judgment, and Theory Choice’ in *The Essential Tension* (pp. 320-353). The University of Chicago Press

Kaku, M. & Thompson, J. T. (1997), *Beyond Einstein: The cosmic quest for the theory of the universe*. Oxford University Press

Kuipers, T. (2002), ‘Beauty, A Road to the Truth’, *Synthese* 131: 291-328

Laudan L. (1981), ‘A Confutation of Convergent Realism’, *Philosophy of Science*, 48: 19*–*49

Mach, E. (1984), *The Analysis of Sensations and the Relation of the Physical to the Psychical*. Trans. by C. M. Williams. La Salle: Open Court.

Meskin A., Phelan, Moore and Kieran, M. (2013), ‘Mere Exposure to Bad Art’, *British Journal of Aesthetics*, 53: 139–164.

McAllister, J. (1989), ‘Truth and Beauty in Scientific Reasoning’, *Synthese* 78: 25-51

McAllister, J. (1996), *Beauty and Revolution in Science*. Ithaca, NY: Cornell University Press

Montano, U. (2014), *Explaining Beauty in Mathematics: An Aesthetic Theory of Mathematics*. Synthese Library Vol. 370

Poincaré, H. (2001), *The Value of Science: Essential Writings of Henri Poincaré*. (ed. Stephen Gould). New York: Modern Library.

Potochnik, A. (2017), *Idealization and the Aims of Science*. University of Chicago Press

Stump, D. (2007), ‘Pierre Duhem’s virtue epistemology’, *Studies in History and Philosophy of Science*, 38: 149-159

Zeki, S., Romaya J.P., Benincasa, D.M.T., Atiyah M.F. (2014), ‘The Experience of Mathamtical beauty and its Neural Correlates’, *Frontiers in Human Neuroscience*, 8: 1-12.

1. That we can find unexpected order in apparently unrelated phenomena can lead to surprise, which in this account related to the experience of beauty, offering an additional dimension to this concept and interestingly relating it to discovery and creativity. [↑](#endnote-ref-1)
2. While the notion of beauty in science in this quotation reflects the mathematical beauty of the physical theory, in the rest for the paper I adopt a much more pluralistic notion of beauty, which extends to theories that are not mathematicised. [↑](#endnote-ref-2)
3. This point has been called into question by recent work on the exposure effect in art conducted by Meskin, Phelan, Moore and Kieran (2013). They argue that exposure to ‘bad’ art does not correlate with increase in subjects’ aesthetic appreciation, suggesting that something over and above exposure must be responsible for subject’s aesthetic responses to art pieces. This results are important for our understanding of beauty in science, giving us grounds to challenge the assumption that exposure and habituation are responsible or sufficient for scientists’ appreciation of the aesthetic values of scientific theories. [↑](#endnote-ref-3)
4. Zeki et al. (2014) show that the same areas of the brain are active when a mathematician is exposed to equations they have previously described as beautiful as when they are appreciating pieces of art and music. [↑](#endnote-ref-4)
5. Of course not all defense of the plausibility of string theory is based on its aesthetic appeal. Richard Dawid has offered a different justification for the theory, explicitly arguing against the employment of beauty as a criterion due to its subjectivity (see Dawid (2013)). [↑](#endnote-ref-5)
6. Since Kuhn (1977), it is commonly accepted that scientists can legitimately disagree in theory choice if they favour different theoretical values (theory virtues such as simplicity, harmony, elegance, etc.). See Duhem (1954) for a discussion of theory choice and the involvement of judgment and good sense in weighting theory virtues, as well as Stump (2007) and Ivanova (2010, 2014). [↑](#endnote-ref-6)
7. See Ivanova (2013, 2015, 2017). [↑](#endnote-ref-7)
8. I am grateful to Hasok Chang and Klaus Ruthenberg for directing me to this debate. [↑](#endnote-ref-8)