How to justify a ban on doping?

Christof Breitsameter

ABSTRACT

Background This article deals with arguments that challenge the possibility of an ethical justification for a doping ban.

Hypothesis/purpose It shows that a justification for the doping ban is only possible if its implementation can be safeguarded.

Study design Systematic review.

Methods Based on the proposition of the game theory, this article examines the scope of the arguments of naturalness, health, equal opportunity and fairness used in scientific literature.

Results/conclusions Ceteris paribus, athletes will always prefer a situation that presents no health risk to a situation in which they face a threat to their health. They will therefore consent to a doping ban on the condition that it is ensured that all parties are bound to this rule, so that anyone complying with the rules will not be afraid of losing the competition as a result. For even if we condoned self-harm, it could still be argued plausibly that the individual should not suffer more disadvantages than absolutely necessary for the sake of gaining an advantage over others. Of course, it is possible to plead for a restricted approval of doping measures with acceptable risk. But even taking minor risks would not seem sensible under the condition that all participants without exception adhere to the same conditions when there is the option to renounce the (avoidable) risks. So as far as the use of performance-enhancing substances or methods is concerned, we can maintain that even if minor health risks are to be expected, a ban on doping can be justified from an ethical point of view.

In the discussion on the question of whether a ban on doping can be ethically justified, a number of terms are used in different constellations and classifications, especially terms like naturalness, health, fairness and equality of opportunity. This paper attempts to critically examine existing arguments by clarifying their normative status and argumentative relation. The aim is to provide a coherent argumentative base for the ethical justification of a ban on doping.

JUSTIFICATION OF THE DOPING BAN

To consider doping from an ethical point of view initially means asking for reasons why certain substances or methods that increase physical performance are banned, and to contrast these with reasons that speak for an approval of such substances or methods. This raises the issue of what actually is considered doping and what is not. There are two definitional strategies fundamentally conceivable. A broad definition—given the case that the ban on certain performance-enhancing substances and methods can be justified—comprises the permitted as well as prohibited forms of doping. A narrow definition is confined to those forms of doping that must be prohibited (this, of course, does not release from the obligation to justify the ban). For the following considerations, I will combine the two strategies in such a way that, as long as the justification for a ban is still in question, I have chosen the broad definition, but when discussing the questions of implementation once the questions of justification have been clarified, the narrow definition shall be applied.

First of all, we shall define doping as the use of substances or methods that serve athletic performance enhancement. Besides the ‘classic’ doping methods, what is known as gene doping promises new possibilities. This is the technology which is used to influence the genetic makeup and—through the expression of respective genetic information—to modify physiological properties with the aim of improving performance. This is aimed at influencing growth, structure, strength, endurance or the regeneration of skeletal muscle, or improving oxygen dispersal and energy supply. The modification of the expression of genetic information is principally achieved with or without the application of genetic elements. We shall differentiate between methods that apply genes or genetic elements to individual tissue cells (gene doping in the stricter sense), and methods that modify the transmission of genetic information (gene doping in the broader sense), both with the intent of influencing the specific control processes.

Yet how can the ban on doping be justified? One of the most frequently discussed models of justification is the naturalness argument. In this case, doping is defined as an unnatural increase in athletic performance. You should not use performance-enhancing substances (or methods, as shall always be implicit in the following) that either the body cannot produce or fail to produce in sufficient quantities. In case of a descriptive intention, it is possible to strictly differentiate between substances that are produced by the body itself (this also applies to gene doping, where effective elements are not applied directly, but indirectly to the body, which in turn activates individual body cells to produce the relevant performance-enhancing substances). It is doubtful, however, whether this can be turned into a prescriptive statement. This is because, on the one hand, foreign substances with performance-enhancing effect can also be used to soothe pain or to cure injuries, and consequently there is no reason to generally discredit such agents; on the other hand, it is because substances...
produced naturally in the body can be used for doping purposes and then the question arises, what distinguishes the natural from the unnatural use of the respective substances. The first part of this reasoning can be countered by arguing that the medically indicated use of foreign substances, which also have a performance-enhancing effect, may lead to a distortion of competition (this is particularly relevant for the distinction between gene doping and gene therapy—e.g. when considering methods for muscle or bone regeneration—which is difficult once such methods are established). Although his argument is certainly apt, in individual cases, for justifying the exclusion from a competition for reasons of fairness, not, however, for generally founding a ban on doping on the criterion of naturalness. This would necessitate a fundamental ban on the medically non-indicated use of foreign substances, but then the reasons for this ban would be unclear. Moreover, according to the criterion of naturalness—and this concerns the second part of the reasoning—increasing performance through the reproduction or enhancement of substances produced naturally in the body would also have to be dismissed, which, of course, gives rise to the question of how the legitimate supply of substances that are not produced in sufficient quantities by the body should be distinguished from the illegitimate supply. For reasons of fairness, it may seem worth considering—theoretically, at least—imposing a ban on individual methods (such as altitude training) as competitive distortion (quite apart from the practicability of this). It is then again the argument of fairness and not that of naturalness which allows the imposition of a general ban on respective methods.

The second important justification model is the argument of the athlete’s health. The use of performance-enhancing substances is only safe if the athlete’s health does not suffer any short-term or long-term damage. This is the point at which it is important to discuss the dimensions of what may be classified as harmful. The intake of anabolic steroids, for example, may lead to an increased risk of arteriosclerosis: the heart muscle mass increases, the myocardial perfusion deteriorates and the liver is damaged, leading to liver failure. Erythropoietin (EPO) primarily leads to ‘thick blood’ and the danger of thromboses, with the additional risk of heart attack. In gene doping, the use of, for example, viral transport vectors (gene shuttles) may have health effects that go beyond the risks encountered in tradition doping (e.g. when building the cardiac muscle leads to undesired side effects). Besides the place of action, problems may also arise when controlling the amount and the point in time for the production of the desired effect. Immune reactions, leukaemia-like conditions or even death are examples of the risks associated with the use of gene doping.

In traditional doping as in gene doping, it is often argued that there are always transitional areas between what can be considered harmful and what not. However, there are also cases where there is a clear-cut distinction that can be drawn with good judgement (we will return to this issue later on). Moreover, there is the question of whether there are substances that do not directly harm the athlete’s health, but which may lead to addiction and should be banned for this reason. Medical research must find answers to all these questions on the basis of empirical findings. Yet these questions do not fundamentally speak against the legitimacy of founding the doping ban on the criterion of harmfulness. It is often argued against the argument of health that certain disciplines in sports are in themselves harmful, yet continue to be practised. However, exactly this can—at least from an ethical point of view—provide the reason for banning unhealthy sports types or practices. The extended argument that sports as a whole bears risks can be rejected with the statement that—with reference to a certain set of rules—you can expect risks to be imminent in the respective game, but to avoid these you would have to give up the game itself. It would be sensible to accept the risks imminent in a certain sport, which are (only) in this sense unavoidable, and simultaneously take precautions that all additional, and thus avoidable, risks should be eliminated. Therefore, the first conclusion could be that harming the athlete’s health (in his or her own interest) should be avoided by banning certain substances or methods. Of course, we take (far higher) health risks in everyday life to achieve effects unachievable without these risks (regardless of a ban on doping, the risks associated with a specific type of sport remain). However, it is rational to avoid unnecessary risks. Given equal circumstances, athletes would reasonably be in favour of a setting without doping as opposed to one with doping, because there is no alteration with regard to the competitive output. The participants therefore collectively gain an advantage by abstaining from doping. It is necessary to demonstrate that the only way of preventing an individual disadvantage is by ensuring that the ban on doping can be implemented for all participants in equal measure.

IMPLEMENTING THE DOPING BAN

If you argue from the point of view of the individual interest of the athlete not to harm his or her health, the problem of implementation subsequently emerges. Although we are able to justify why doping (now only used in its narrow definition) should be banned—it should not damage the athlete’s health—we must, however, consider ways of implementing this ban. For it is clear that the athlete, who follows the logic of this argument and renounces the use of harmful, performance-enhancing substances, will be outperformed by the one who does not, because we can expect him or her to lose the competition. This means that we are faced with the ethically relevant issue that individual self-commitment, for which there are substantial reasons, can be exploited. In more concrete terms, we are dealing with conflicting values, and the competing values are ‘health’ on the one hand and ‘success’ on the other. The question now is how to solve this conflict.

Let us consider the situation in which an athlete contemplates whether to resort to doping or not, following the maxim of the famous American coach Vince Lombardi, to whom the following sentence is attributed: ‘Winning isn’t everything, it’s the only thing!’ This means that for an athlete the value of success may be of higher value than his or her own health. If athletes are seeking success (something we can at least assume) and do not know which values the other players prefer, this results in the following logic that is visualised in a matrix (figure 1). From the point of view of A, there is the following consideration: if B has a non-doping strategy, A will achieve the best result if he resorts to doping; for if A also adopts a non-doping strategy, he can only achieve his second best result. Let us now consider the case in which B adopts a doping strategy. This is where A will achieve his third best result, if he also adopts the doping strategy; the worst strategy for him would be to adopt a non-doping strategy while B resorts to doping. In this case, doping is also the better strategy. The same result applies to B. The outcome of this constellation is that for both contestants doping is the dominant strategy (a setting commonly referred to in ethics and game theory as prisoner’s dilemma). Indeed, the use of game theory is not an entirely new approach. Gunnar Breivik focuses mainly on success and fairness as conflicting goals. This way, rules of fairness preventing detection can be established, but this
approach does not rule out all parties involved collectively agreeing on the doping strategy. In my opinion, this strategy can only be rejected on rational grounds if health is added as a goal. Another weakness of Breivik’s rationale is his suggestion to overcome defection through moral appeals and fair motives. On the one hand, there is empirical evidence suggesting the ineffectiveness of such appeals causing fair motives to deteriorate. This is why a moralisation may run the risk of tolerating the unilateral exploitation of moral motives. In this regard, dilemmas cannot be resolved without sanctionable rules (which Breivik acknowledges in addition to the moral appeal, although he does not establish a systematic connection between motives and rules).

Basically, there are three conceivable strategies. A first possibility is for athletes to change their order of preferences (eg, on grounds of a moral appeal, as suggested by Breivik) and promote health to their top consideration as far as values are concerned. In reality, this is rather unlikely to happen. However, we would merely be dealing with a contingent phenomenon; otherwise there would be no doping problem. A second possibility would be for the contenders to come to a mutual agreement to do without doping. However, they will not be able to rely on this simply because of the existence of such an agreement, as the incentive to defect and to exploit the moral advance of the others remains. A third possibility is to establish binding rules and regulations, which means making doping an offence and threatening any action against this norm with sanctions. This is the only way to reliably effect change in the individual’s judgement. Therefore, in addition to the value of health, which is the first restrictive condition when using non-natural performance-enhancing substances, we require a second restrictive condition, which is a legal perspective to guarantee that all athletes will renounce the intake of harmful agents. Against this backdrop it becomes evident that even the proposal to liberalise doping, on the proviso that athletes keep a (non-public) ‘drug diary’ in which they must record all employed substances and methods (there will be a threat of sanctions, such as publishing the records), is inappropriate, as this will not help to overcome the above-mentioned dilemma. (1) In no way will this eliminate, as is the claim, the incentive to create new substances and methods because—just as with the strict ban on doping—the use of previously unknown agents can mean victory and the athlete concerned may even accept the risk of being detected and punished. (2) If all athletes would follow this rule and lay open all substances and methods (because we assume that any kind of doping can be successfully detected), so that indeed there would be no incentive to create new substances as those could immediately be copied, thus preventing a competitive advantage, there would be the decisive disadvantage that this rule would establish a system of collective, self-inflicted damage as far as the athlete’s health is concerned. This is why the strict ban on doping is the better rule.

**EQUAL OPPORTUNITY AND FAIRNESS**

The necessity to establish rules that collectively apply to competitive sports is handled by many authors under the terms of equal opportunity and fairness. Both terms mark constituents of competition and can therefore be called inherent characteristics of competitive sports: they issue from the logic of the competition itself. Whereas in the following I will consider the term equal opportunity as being the general conditions of admission, as they foremost guarantee the uncertainty of the outcome of the game (eg, the distinction between men and women or of weight classes in the individual disciplines), I will use the term fairness to denote the adherence to the underlying rules of the game or discipline (eg, offside or foul). The conditions of admission shall also include equipment or training methods, the admission or possible exclusion of which would follow pragmatic instead of ethical considerations (for even here a foreseen imbalance would prejudice the outcome of the competition). Under this consideration, doping would be a breach of sporting fairness that would destroy the fundamental safeguard of equal opportunity. Ensuring fairness and equal opportunity in this case is linked to the observance of rules.

Let us take a closer look at both terms. We could use the term equal opportunities to illuminate the following argument that would be in support of the liberalisation of doping: doping helps compensate for natural imbalances and therefore it is essential to ensure equality in the competition. There are two possible reactions to this. (1) Either all athletes react uniformly to the given method. Then it would be more sensible not to do anything at all in case of immanent health risks, as this would only prolong the existing imbalances, as shall be shown in more detail later on. (2) Or the athletes react differently to the given method. In this case, it would be possible to establish equality, given the technical possibilities. Even if we countered the argument that in this case the natural lottery of different talents is only substituted by the natural lottery of various reactions to one and the same measure with the argument, that knowledge founded on evidence about the respective effects could lead to a fair central distribution, one must consider that not all differences in performance can be levelled out through medical intervention. And this again raises the question of whether it is worth taking the risks, even if they are so minute and improbable, if the desired effect can only be predicted to a certain extent.

So let us turn to the term fairness: linking the rule of fairness to the guiding principles of health, the breach of which should be prevented in the interest of the individual athlete, results in the following: ceteris paribus, athletes will always prefer a situation that presents no health risk to a situation in which they face a threat to their health. They will therefore consent to a doping ban on the condition that it is ensured that all parties are bound to this rule, so that anyone complying with the rules will not be

---

**Figure 1** Doping as a prisoners’ dilemma.
afraid of losing the competition as a result. For even if we condoned self-harm, it could still be argued plausibly that the individual should not suffer more disadvantages than absolutely necessary for the sake of gaining an advantage over others. Of course, it is possible to plead for a restricted approval of doping measures with acceptable risk. But even taking minor risks would not seem sensible under the condition that all participants without exception adhere to the same conditions when there is the option to renounce the (avoidable) risks (even if, as already mentioned, there will always be a debate about what is healthy and what is individually unhealthy). So as far as the use of performance-enhancing substances or methods is concerned, we can maintain that even if minor health risks are to be expected, a ban on doping can be justified from an ethical point of view.

So we can only reliably solve the conflict of values—health versus success—by introducing an enforceable rule, and we can assume that all parties involved can agree to the introduction of this rule. The existence of dilemma structures, as depicted in the matrix, or—to put it differently—existing stimuli for defection must therefore be removed through the introduction of enforceable rules. There are two conclusions to be drawn here. (1) The obligatory force of the ban on doping is linked to control and sanction mechanisms that must be established by means of appropriate statutory regulations. This is where detection methods play a decisive role. Without doubt, the extent of sanctions and the rigour of their application are discretionary. However, to participate in a sport and accept its rules is a voluntary action. That is precisely what makes this situation different from that of any other citizen, who indeed needs to be protected from inconsiderate and excessive access by the state. In principle, unnecessary self-inflicted damage can be prevented by collective self-commitments, for example, the requirement to wear a seat belt. The connection of a useful action with an action causing only (collective) disadvantages can be dissolved in this manner. Tests are not undesirable interferences in the privacy of non-doping athletes, who refuse to dope because the doping competitors will be exposed through tests and are also not undesirable interferences in the privacy of doping competitors, because they, if given the choice, would prefer a world without doping over a world with doping. The potential of game theory lies precisely in its ability to illustrate that athletes do not control the outcome of a game individually but only collectively.

(2) Apart from this, creating incentive structures is also conceivable, for example, rewarding performance maintained over years or promoting self-commitment to health checks, etc. Moreover, it may make sense to eliminate existing incentives that are considered unfavourable, for example, by splitting disciplines requiring extreme endurance into several individual stages. The model proposed here therefore identifies the insufficiency of the theoretical approaches that treat doping as being primarily an issue of individual ethics. Instead of individual self-commitment, the consensus of excluding harmful performance-enhancing methods through collective self-commitment comes to the fore. Individual self-commitment would be faced with the problem of anyone following this maxim would be a disadvantage in the sporting event. A discourse on the correct rules must therefore precede a discourse on the correct attitude. Only against a backdrop of generally obliging and enforceable rule based on consensus—social ethics—a moral attitude can be established: that of treating one’s own body with care and moderation, through the introduction of a rule acceptable to all parties concerned that will not present a disadvantage to the individual or the collective.

At this point, we need to revert to the argument of naturalness. It is argued that, if one viewed both the usage of substances foreign to the body and natural substances in the athletes’ body for performance enhancement as unnatural, the prescriptive inference of a ban of these technologies could be justified. In fact, this paper claims in descriptive terms (agreeing with Pawlenka) that it is plausible to consider the increased use of the body’s own substances as a means of performance enhancement as unnatural, respectively artificial in contrast to the natural production of these substances, which is the prerequisite for the prescriptive consequence illustrated above. If there is no descriptive difference, a corresponding prescriptive conclusion cannot be established. Hence, in prescriptive terms, the question is why the increased use of the body’s own substances should be subjected to a ban. In principle, it appears conceivable to permit all athletes to make use of these methods. The prohibition of this method could be established on the rationale of individual defection, that is, to prevent unfair distortions of competition, not because the method is unnatural. A prohibition could also be established with regard to collective defection, if the method is health damaging. One could also argue that the increased supply with endogenous substances is not, in descriptive terms, unnatural. Even in this case, the prescriptive consequences would not be definitive, that is, whether or not this method is to be considered doping and whether it should be banned or not. The argument that potentially health-damaging technologies should be forbidden by collective rationality if the same goals can be achieved without these methods also applies in this case. Therefore, the prescriptive application of the category of naturalness, respectively the distinction between natural and unnatural, is rejected in this context.

CRITICISING SOME ARGUMENTS CONCERNING ETHICS IN SPORT

It also becomes clear that doping cannot be considered an irrational action. The explanation for this is simple, and basically we have already seen it. If all other participants would abstain from doping, it would be the best solution for the individual to turn to doping. Doping, in this respect, would be individually rational. Of course, you may argue that the athlete must accept that his or her health will be harmed. Balancing health and success, however, can only be performed individually, which means that one cannot generally say that health is more important than success for a human being. Let us assume that an athlete could win international fame and prize money, a situation that will leave him financially secure for the rest of his life, and against this would be a reasonable risk of damaging his health, with which he would be able to live or which could even be cured. It could be rational to accept this risk in order to exploit the opportunity to win (the fact that this assumption is realistic is shown by ref. 18).

There is no external logic in this case, which could intervene in the internal rationale of the athlete and out the Lombardi logic. We can only intervene in the internal rationale of a person if we change the rules within which he or she acts such that the person who has to obey these rules is able to agree to the limitation of his freedom of action. This means that the athletes (participants) are faced with the problem that the individual and the collective view may diverge. Even if athletes agree that a world without doping and its negative consequences is better than a world with doping and the respective damages, if this view thus could be collectively rational, it could factually be the best solution for the individual to let all others renounce doping and then win the competition uncontended. Doping can
only not be considered as the result of a rational decision if a ban has been imposed and its implementation has been ensured. Then it will indeed be collectively and individually rational to renounce doping.

The attempt to challenge the argument that doping is an infringement on equal opportunities with the suggestion that there are never equal opportunities in sports if you consider training practice and the infrastructure in which rival contenders are acting continues to be unacceptable. Of course, one must admit that there may be asynchronicities in the development of training methods or equipment, etc. All participants, however, should in principle be able to make up leeway. If it were—as already implied—clear from the start that these inequalities would continue to exist and it would be impossible for some to close the gap, a competition (which cannot be a veritable competition) could not even develop because nobody would be interested in it. If, however, equal opportunities would in principal be enabled through rules, then one could also approve doping for all parties, thus all above-mentioned criticism would lead nowhere on this issue.

This is why we, as we have seen, also consider the value of health on which we must found the ban on doping. To prevent any asymmetries from forming, that is, that only a few will abide by the ban and then surely lose, it is necessary to include the rule of fairness, or better: to establish fairness as an enforceable rule. Because doping presents the dominant strategy, provided that there is no binding rule, we can conclude that mere appeals without institutional provisions will have no effect, and provided that there is no binding rule, we can conclude that mere appeals without institutional provisions will have no effect, and in the long term will lead to the erosion of moral attitudes.21 This does not intend to discredit moral appeals in general.22 Moral appeals by institutions, providing information on possible health risks or on the consequences of breaching the rules of the competition, must be supported by enforceable rules ensuring that those adhering to the rules of the competition are not at a disadvantage. If the law cannot prohibit something efficiently, action cannot be taken against it, for example, in the case of potentially health-damaging substances, which cannot be detected in an athlete’s body. This becomes visible in the moral bindness as well: people can only be demanded by law to refrain from an action considered immoral, if the enforcement of such a rule can be guaranteed for all parties involved. Someone cannot be morally demanded to (not) do something, if it cannot be enforced legally. From this insight, the conclusion can be drawn that deficits in the legal enforceability need to be removed if possible. If it is not possible to remove these deficits, moral demands should not be directed at individuals.

PATERNALISM

When banning doping, can this be considered a paternalistic intervention, and can such an intervention be ethically justified? It is justified—and this is where relevant literature reveals great consensus—to intervene through action or omission where a person is harming himself or herself as a result of unknown or unintentional consequences. This also applies if there is no evidence for deliberate decisions or actions, where we can thus only talk of behaviour. For such cases, it is plausible that the cause of this behaviour or this decision or action should not be considered to be part of the respective person. Instead, we should deal with the underlying mental and external factors as a cause. This, for example, applies to wrong convictions regarding the relationship between cause and effect.23 Events that fall under this description can be treated as a natural phenomenon and can be captured with the harm-to-others principle (such as if a person had to be protected from the harmful effects of a natural phenomenon). Therefore, they do not present a case of paternalistic action. From an ethical point of view, it is therefore not legitimate and mandatory to inform an athlete about the underlying health risk factors.24 To put it differently: athletes must always be comprehensively informed about the risks their actions involve to take autonomous decisions for or against those actions. The case is different if a person harms himself or herself through known or intended consequences of doping. With regard to the competitive sport which is our concern here, one must observe that such a decision or action also has an effect on all other athletes in the competition, which, in turn, refers to (breached) principle of fairness. By means of a mutual self-commitment or collective voluntary agreement, the parties involved can submit to a rule, which they can universally approve of. Self-legislation is autonomy and needs to be distinguished from paternalism. Clearly, this only applies if the situation is designed to be resistant to exploitation. This way, individual and collective disadvantages are prevented. It is rational to grant an institution the ability to intervene in our freedom of action in a paternalistic manner, if the same goals can thus be achieved precisely without unnecessary risks. The freedom of the individual is not limited by paternalism in as far as we may assume that athletes engage in their sport by their own free will and always have the opportunity to give up their career.

Competing interests None declared.

Provenance and peer review Not commissioned; externally peer reviewed.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/

REFERENCES

Doping in sport