Running head: OPPOSITION TO STEROID USE

What's wrong with using steroids?

Exploring whether and why people oppose the use of performance enhancing drugs

Justin F. Landy, Daniel K. Walco, and Daniel M. Bartels

University of Chicago Booth School of Business

DOI: 10.1037/pspa0000089

Author Note

Justin F. Landy, Center for Decision Research, University of Chicago Booth School of Business; Daniel K. Walco, Department of Behavioral Science, University of Chicago Booth School of Business; Daniel M. Bartels, Department of Marketing, University of Chicago Booth School of Business.

We are grateful to Halley Bayer, Johannes De Nova, Alex Kristal, Andrew Xue, and Austin Zheng for their assistance with this research, and to Sydney Scott and the members of the Chicagoland Morality Researchers for their insightful comments and feedback.

Correspondence concerning this article should be addressed to: Justin Landy, Center for Decision Research, 5807 S Woodlawn Avenue, Chicago, IL, 60637, USA. Email: justinlandy@chicagobooth.edu

Abstract

The use of performance enhancing drugs (PEDs) elicits widespread normative opposition, yet little research has investigated what underlies these judgments. We examine this question comprehensively, across 13 studies. We first test the hypothesis that opposition to PED use cannot be fully accounted for by considerations of fairness. We then test the influence of ten other potential drivers of opposition in an exploratory manner. We find that health risks for the user and rules and laws prohibiting PED use reliably affect normative judgments. Next, we test whether these patterns generalize to a related domain—the off-label use of cognitive-enhancement drugs. Finally, we sketch a framework for understanding these results, borrowing from Social Domain Theory (e.g., Turiel, 1983). We argue that PED use exemplifies a class of violations with properties of moral, conventional, and prudential offenses. This research sheds light on a widespread, but understudied, normative judgment and illustrates the utility of exploratory methods.

On September 24, 1988, Canadian sprinter Ben Johnson won the gold medal in the men's 100-meter sprint at the Summer Olympic Games in Seoul, with a then-world-record time of 9.79 seconds. This was the highlight of Johnson's career, and he was hailed as a national hero in Canada. Less than two days later, Johnson was stripped of his medal and disqualified from the games, after his blood and urine tested positive for the presence of stanozolol, an anabolic steroid. He lost his medal and his status as a hero to the people of Canada: Johnson's steroid use elicited widespread shock, outrage, and disappointment (Gordon, 2012).

Johnson's situation is not unique. For instance, since Major League Baseball instituted a new drug policy in 2005 that included random testing, 59 players have been suspended for using Performance Enhancing Drugs (PEDs; "Steroid Suspensions," 2017). Several prominent stars, including Barry Bonds and Roger Clemens, have been excluded from Major League Baseball's Hall of Fame due to their connection with PED use ("Barry Bonds, Roger Clemens Not Voted into Baseball Hall of Fame", 2017). Dozens of players in the NBA and NFL have faced similar penalties. Opposition to PED use has substantial financial consequences as well; in 2015 alone, the United States Anti-Doping Agency spent over $9 million dollars on tests for banned substances, across 85 different amateur and professional sports (U. S. Anti-Doping Agency, 2015). Opposition to PED use is widespread, and in this paper, we aim to understand what drives this commonly expressed normative judgment.

Prior research on opposition to PED use is fairly sparse, but the most common theoretical claim is that PED use is opposed because it violates standards of fairness. Several studies have examined the role of fairness in driving opposition to PED use. For example, people recognize that athletic competitions are zero-sum affairs; whatever one competitor gains by using PEDs is another competitor's loss. This helps to explain why opposition to PED use in the athletic

domain is more extreme than in non-zero-sum domains, like some academic settings (Dodge, Williams, Marzell, & Turrisi, 2012). Also, perceptions of both "competitive" fairness (whether a drug provides an unfair advantage) and "distributive" fairness (whether everyone has equal access to a drug) both reliably affect normative judgments of using cognitive-enhancement drugs (Scheske & Schnall, 2012; see also Fitz, Nadler, Manogaran, Chong, & Reiner, 2014).

Fairness has also been used in several studies as a measure of opposition, with the implicit assumption being that perceptions of unfairness are a primary driver of, or perhaps even synonymous with, opposition to PED use. For example, using a cognitive enhancement drug was seen as quite unfair when only some people could afford it (i.e., when it was distributed unfairly). This was especially true when the user did not have to work hard to earn the money to buy it (Fitz et al., 2014). Furthermore, the majority of a sample of German university students stated that the use of cognitive enhancement drugs among other students was unfair (Forlini, Schildmann, Roser, Beranek, & Vollmann, 2015). Lastly, people consider it less fair for others to use performance-enhancing products than for they themselves to do so (Williams & Steffel, 2014).

Ben Johnson might agree with the emphasis on fairness in this prior research. He defended his use of steroids by claiming that everyone he was competing against used them, and so he was not gaining any advantage by using them himself. In other words, he contended that his use of steroids was not unfair (Gordon, 2012). Johnson's coach, Charlie Francis, concurred. While testifying before the Canadian government, he said that he "couldn't find a single case of performance-enhancing drugs *not* being used" among world-class runners at the time (Janofsky, 1989, emphasis added). This argument did not stop the International Amateur Athletic

Federation (now known as the International Association of Athletics Federations) from banning Johnson from international competition for two years (Gordon, 2012).

Although we agree that concerns about fairness are an important contributor to opposition to PED use, we think that there are likely more considerations driving this opposition. Johnson himself, despite the fact that he did not consider using steroids to be unfair, noted, "I didn't tell my mother, because if I told my mother, I wouldn't have done it. She wouldn't let me and I would never disappoint her" (Gordon, 2012). He apparently recognized that, even if he was not gaining any unfair advantage over his competitors, his mother would still not approve of his using steroids. So, there may be more contributing to normative judgments of PED use than fairness concerns, but what these other contributors are is not well understood.

There is little current research on non-fairness-related drivers of opposition to PED use. Outside of the athletic domain, the judged "moral acceptability" (broadly construed) of drugs that enhance psychological traits predicts opinions on whether the drugs should be legal (Riis, Simmons, & Goodwin, 2008). Although these moral concerns might relate to fairness, other considerations may also be at play. Use of cognitive enhancement drugs is more strongly opposed if they can cause severe side effects, if they are illegal to possess without a prescription, and if they violate local (university) regulations (Sattler, Forlini, Racine, & Sauer, 2013). Using biomedical enhancements of various sorts was judged less wrong if they were taken for selfless ends, if they were "natural", and if they still required effort on the part of the user (Hester, Schein, Sirbu, & Gray, 2015). Lastly, opposition to the use of cognitive-enhancement drugs is affected by whether or not they pose any risk to the user's health, whether they are "natural", and how they are administered, with injected drugs being more strongly opposed than drugs taken as

pills or additives to drinks (Scheske & Schnall, 2012). Similar concerns might predict judgments of PED use more generally.

Overview of the Current Research

The current research aims to answer two related research questions:

RQ1: Do people still oppose the use of PEDs when there is no violation of fairness, i.e., even if there is no competitive advantage to be gained (Studies 1, 12, and 13)?

We hypothesize that the answer to RQ1 is yes. No research has directly tested this prediction; one study found that the use of cognitive enhancement drugs by students was still largely opposed even if everyone else was using them (Sattler et al., 2014). However, this study did not specify whether the academic context was structured like a competition, so it is not clear whether there is any violation of fairness in this scenario (see Dodge et al., 2012). However, anecdotal evidence like the story of Ben Johnson, and the similar – and similarly sordid – story of Lance Armstrong (Selinger, 2012) suggests that even when an athlete is just "keeping up" with the competition, PED use still prompts considerable opposition. Should we find support for this hypothesis, it would raise a subsequent question:

RQ2: What considerations, other than fairness, underlie opposition to PED use (Studies 2-13)?

Although a small amount of prior research has examined this question, it has been fragmented, with each project examining a handful of explanations, to the exclusion of others. We take a different approach, with the goal of examining this question more comprehensively. One tack researchers often take is to commit to a particular theoretical position and test predictions derived from this position using a variety of methods. In that vein, they often test the

effects of one psychological construct across a range of dependent variables, in different ways. We characterize this approach as "broad", in that it examines the implications of the construct of interest in various different situations. Our approach is "narrower", in the sense that—for most of our studies—we examine one outcome measure using variations on a single stimulus, and test the effects of a multitude of psychological constructs on this outcome.

What this approach lacks in "breadth", it makes up for in "depth", allowing us to study the domain of interest (opposition to PED use) comprehensively. Our "deep dive" method resembles John Platt's concept of "strong inference", in that it is capable of falsifying or supporting the most prominent explanation for the phenomenon being studied (i.e., opposition to PED use is due to fairness concerns), and it allows us to test multiple explanations to make rapid progress and arrive at a fuller understanding of the multiple causes of this phenomenon. We think that this approach is likely to be especially useful when applied to phenomena that have not yet been extensively studied, as we do here. In addition to shedding light on a widely expressed, but relatively understudied normative judgment, this research provides an example of how exploratory, "deep dive" studies can be used to illuminate social psychological phenomena that are not yet well understood.

Possible Explanations

Based on our reading of the psychological literature on normative judgments, participants' thought listings, and popular press articles about PED use in sports, we identified ten factors that might underlie opposition to PED use. We tested for effects of each of these factors on normative judgments in an exploratory manner to answer RQ2. The ten factors that we examined are:

Laws and league rules. Social Domain Theory (SDT) argues that acts can be "wrong" in qualitatively different ways (see, e.g., Nucci & Nucci, 1982; Smetana, Jambon, & Ball, 2014; Turiel, 1983, 2002). Violations of social conventions offend against shared norms of conduct, but are not considered to be wrong if they take place in normative contexts where there is no prohibition against them – e.g., where they are not against the rules, or where popular consensus has ruled them permissible. Classic examples include dress code violations, or addressing authority figures informally. Moral violations, on the other hand, violate standards of fairness or cause direct harm to someone other than the actor, and are considered wrong even in contexts where there are no rules prohibiting them (Turiel, 1983).[1] Steroid use is typically prohibited by law, and by the governing bodies of competitive sports, so perhaps it represents a particularly serious kind of conventional violation. If so, then we would expect that it would no longer be opposed in the absence of rules and/or laws prohibiting it (see Sattler et al., 2014).

Prudence. SDT also posits another class of behaviors: "prudential violations" are considered wrong, not for classically moral reasons like fairness or harm to others, but because of harm they cause to the actor (Nucci, Guerra, & Lee, 1991). Non-drug users generally consider the use of illicit drugs to belong to this class. Also, studies have found that people disapproved more of using cognitive-enhancement drugs when they posed a risk to the user (Sattler et al., 2014; Scheske & Schnall, 2012), and it seems likely that this applies to other PEDs as well. Research also indicates that perceptions of harm severity predict normative judgments of

[1] Consistent with SDT, we reserve the term "moral" to refer to issues of fairness and direct interpersonal harm. We use the term "opposition" to refer to the widespread judgment that PED use is wrong or blameworthy in some sense. We will use the term "normative", as distinct from "moral", to describe this judgment, because it indicates that a judgment is about what one should or should not do, but does not necessarily imply unfairness or harm. For example, a judgment that one *should not* wear pajamas to work (Huebner, Lee, & Hauser, 2010) would, in our terminology, be a "normative" judgment, but not a "moral" one, because the act in question violates social conventions, but does not directly cause harm or violate standards of fairness.

another, more extreme type of self-harm, suicide (Gray, 2014). Steroid use may similarly be considered a prudential violation, wrong by virtue of the harm it causes to the actor.

Punishment. Like Ben Johnson, athletes caught using PEDs are generally subject to some kind of punishment. People may therefore infer that one should not use PEDs, *because one could be punished.* This thinking represents "Stage 1: Punishment-Obedience Orientation" in Kohlberg's (1971) framework, where what is wrong is defined by what is punished. This type of thinking is primarily observed in children, but we thought it worthy of investigation nonetheless.

Naturalness. Naturalness is highly valued in Western cultures (Rozin, Fischler, & Shields-Argelès, 2012). Prior research has found that biomedical enhancements that are considered "unnatural" elicit more opposition than "natural" alternatives (Hester et al., 2015; Scheske & Schnall, 2012), consistent with research linking perceptions of naturalness to normative judgments in other domains (Rottman & Kelemen, 2012; Rozin et al., 2004; Sunstein, 2005). We expected to replicate this finding.

Interfering with the body. Similar to the above point, people might also take issue with artificially tampering with the natural functioning of the human body, whether via a "natural" or "unnatural" substance. In the language of moral psychology, such tampering would be a violation of bodily purity (e.g., Chakroff et al., 2015; Graham, Haidt, & Nosek, 2009; Horberg, Oveis, Keltner, & Cohen, 2009) or the ethic of divinity (Rozin, Lowery Imada, & Haidt, 1999). We therefore aimed to test whether improving performance via illicit methods that do not interfere with the body's natural functioning are considered less wrong than doing so by using PEDs.

Effort. Studies have found that PED use is opposed less when the user still has to exert significant effort to achieve the desired results (Hester et al., 2015). Anabolic steroids mimic testosterone, speeding up protein synthesis in muscles (Brodsky, Balagopal, & Nair, 1996), which may be why athletes report being able to train harder and recover more quickly while using them (Fahey, 1998). Steroid users may actually be able to train *more* than non-users, and intense resistance training seems to be necessary for any improvement in performance while using steroids (Fahey, 1998). In other words, steroid users must actually still exert substantial effort to achieve the desired increase in performance, a point that we suspected most people would not be aware of. Opposition to steroid use may be reduced if one is made aware of this fact.

Competitive context. At least part of why PED use by athletes is opposed seems to be because athletic competitions are a zero-sum context (Dodge et al., 2012). People might over-apply a heuristic judgment that PED use is wrong in zero-sum competitions, even in circumstances where all participants are using PEDs, and there is no actual advantage to be gained. If so, this would imply that PED use should be considered less wrong in non-competitive contexts.

Athletes as role models. In the popular press, one oft-raised issue is that professional athletes are role models for younger athletes, and young fans more generally. These young fans, the argument goes, may be more likely to engage in bad behaviors if athletes model steroid use as acceptable (Bruinius, 2005; Jenkins, 2005). On this explanation, PED use by athletes should be opposed because of the negative influence it might have on the next generation. This requires a belief that the use of PEDs is wrong for some other, logically prior reason; for instance, because it violates a rule and/or carries some risk. If it did not, the athletes would not be

modeling rule-breaking or risky behavior, and there would be no negative influence on young fans. This explanation for opposition to PED use might therefore be considered a derivative or special case of other types of explanations above. We considered it worthy of investigation on its own because it is so often discussed in the popular media.

Fairness to past competitors. If Ben Johnson was right, and all of his contemporary competitors were using PEDs, one still could argue that his steroid use was unfair to past competitors. Current athletes are often compared to athletes from past generations, particularly when records are at stake. If those former competitors set records without the benefit of PEDs, then current PED use could be viewed as unfair to those past competitors. This is often mentioned as a reason to oppose PED use in the popular media (e.g., Barra, 2007; Bruinius, 2005; Jenkins, 2005).

Terminology. Lastly, perhaps people are averse to terms like "drug" and "steroid" that have negative connotations. If an aversion to these terms underlies opposition to PED use, then this opposition should be reduced if the performance enhancers are described using more evaluatively neutral terms like "supplements". Framing effects like these are often observed in political polling. For example, in 2013, 46% of Americans reported opposing "Obamacare", whereas only 37% reported opposing the Affordable Care Act, but both terms refer to the same legislation (CNN Political Unit, 2013).

Preview of Results and Theoretical Contribution

The methods in this research are exploratory, but converge on the conclusion that opposition to PED use is attributable to classically moral (fairness), social conventional (laws and league rules), and prudential (harm to the self) concerns, as articulated in SDT, with the

conventional concerns exerting the largest influence on judgments. Our results also suggest that steroid use is an example of a heretofore unstudied class of violations that does not clearly exemplify only one of these three domains of behavior, but exhibits properties of all three. In this paper, we aim to illuminate how people make normative judgments of PED use and to explicate the properties of this novel class of violations. We sketch an account of people's normative judgments of PED use (and possibly of other cases in this class of violations) more fully in the General Discussion.

Study 1

Method

Participants. Participants in all 13 studies reported in this paper were recruited through Amazon Mechanical Turk, and were located in the United States. Participants could not take part in more than one study, with one exception. Study 13 was run more than four months after Studies 1-12, and used quite different stimuli, so we allowed participants from Studies 1-12 to take part in Study 13. Because we did not have an *a priori* estimate of the effect sizes of interest, we aimed to recruit at least 50 participants per condition in Study 1, following the recommendations of Simmons, Nelson, and Simonsohn (2013). After exclusions for not completing the full study, we were left with a final sample of $N = 150$ (89 male, $M_{Age} = 33.61$ years, $SD_{Age} = 10.06$).

Procedure. Participants were randomly assigned to one of three conditions: separate evaluation – competitive advantage, separate evaluation – no advantage, and joint evaluation. In the latter condition, participants viewed and responded to the scenarios from both separate evaluation conditions on the same page, with the order of presentation counterbalanced. Joint

evaluation produces more reflective, careful responses (Bazerman & Messick, 1998; Hsee, Loewenstein, Blunt, & Bazerman, 1999), and several normative judgments differ across these modes of evaluation (e.g., Paharia, Kassam, Greene, & Bazerman, 2009; see Bartels, Bauman, Cushman, Pizarro, & McGraw, 2016, for a review), so we included both modes of evaluation in this study.

Participants read a scenario, or a pair of scenarios, that described "Joe", a professional baseball player who had never before used PEDs, but decided to start using anabolic steroids. In the competitive advantage condition, participants read that no other player in Joe's league was using steroids. So, Joe was gaining a large competitive advantage. In the no advantage condition, participants read that every other player was using steroids, and therefore Joe was not gaining any advantage by also using them. Note that in the no advantage condition, there is no violation of fairness, because all of Joe's competitors are also using anabolic steroids, so Joe is just keeping up with them. Participants rated how wrong it was for Joe to use steroids on a 1-9 scale, ranging from "Not at all wrong" to "Extremely wrong", with the scale midpoint labeled "Moderately wrong".

On the following page, participants were presented with an unexpected second task, where the scenario(s) were presented again, and they were asked to explain their responses in one to two sentences. Participants in the joint evaluation condition were asked to justify both of their responses. After responding to this open-ended question(s), participants completed a brief demographics questionnaire (age, sex, political affiliation, and education), and were debriefed, thanked, and paid. No unreported measures were collected in any study in this paper.

Results

Wrongness ratings. As can be seen in Figure 1, wrongness judgments were similar in the Separate Evaluation and Joint Evaluation conditions. Using steroids was seen as significantly more wrong when no one else was ($M_{SEPARATE}$ = 7.64, $SD_{SEPARATE}$ = 1.72; M_{JOINT} = 7.92, SD_{JOINT} = 1.65) than when everyone else was ($M_{SEPARATE}$ = 5.88, $SD_{SEPARATE}$ = 2.41; M_{JOINT} = 5.79, SD_{JOINT} = 2.46), in both separate evaluation, $t(86.22) = 4.22$, $p < .001$, $d = .85$,[2] and joint evaluation, $t(48) = 6.15$, $p < .001$, repeated-measures $d = .89$.[3] More importantly, and supporting our prediction above, participants still opposed "Joe's" use of steroids, even when he did not gain any competitive advantage. In fact, mean ratings in the no advantage condition were significantly greater than the scale midpoint (labeled "Moderately wrong") in both separate, $t(48) = 2.55$, $p = .014$, $d = .36$, and joint evaluation, $t(47) = 2.23$, $p = .030$, $d = .32$.

[2] Throughout this paper and the Supplemental Materials, *t*-tests with fractional degrees of freedom indicate that Levene's test was significant, so equality of variances was not assumed.

[3] Morris & DeShon (2002) present an accessible explanation of how the repeated-measures *d* should be interpreted, and how it relates the more familiar independent-groups *d*.

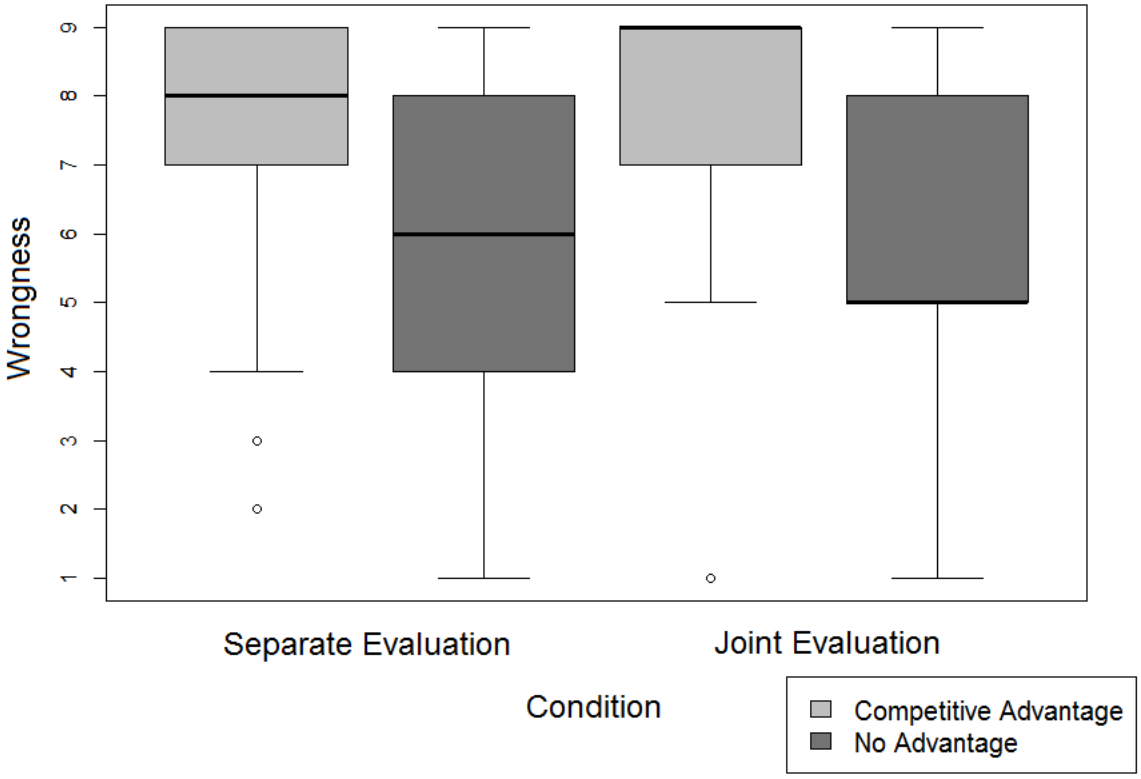

Figure 1. Wrongness ratings in Study 1.

Justifications. Initially, two research assistants, blind to condition and to the purpose of the study, coded participants' open-ended justifications for their responses. They assigned each response to one of 12 categories adduced from a preliminary reading of the responses by the first author, or to an "other" category, or coded them as being uninterpretable. The full coding scheme is presented in the Supplemental Materials. Agreement between the coders was only moderate; their classifications agreed for 125 out of 198 total justifications (63.1%). Therefore, we had a third research assistant (also blind to condition and purpose) code the justifications as well. We intended to retain any classification that at least two of the three coders agreed on, and treat disagreements among all three as uninterpretable. Twenty-one responses were deemed uninterpretable based on this criterion, leaving 177 responses with a usable classification. We

analyze those classifications here; separate analyses of each of the three coders are presented in the Supplemental Materials.

The continued opposition to steroid use in the no advantage condition is not attributable to participants misunderstanding the scenario. In the competitive advantage condition, 50 out of 89 usable responses (56.2%) appealed to fairness concerns, with the remaining 39 appealing to other reasons. By contrast, in the no advantage condition, only 18 out of 88 usable responses (20.4%) appealed to fairness and only eight of these 18 rated steroid use as at least "moderately wrong". The remaining 70 responses in the no advantage condition cited other reasons, $\chi^2(1) = 23.87$, $p < .001$, Cramér's $V = .37$. So, participants in the no advantage condition generally recognized that there was no violation of fairness, yet continued to oppose steroid use to a substantial degree.

What did participants in the no advantage condition appeal to in support of their normative judgments? Among participants who rated Joe's steroid use as at least "moderately wrong", the most common non-fairness-related responses argued that consensus is insufficient to make steroid use permissible (23 out of 58 participants) or cited a rule or law prohibiting steroid use (17 participants). No other classification was assigned to more than ten responses.

Discussion

In Study 1, participants still opposed steroid use, even when there was no competitive advantage to be gained, suggesting that fairness considerations cannot fully explain disapproval of PEDs, as we hypothesized. Participants frequently cited a rule or law as a reason for their judgments. This justification implies that steroid use would *not* be wrong in the absence of such a rule, a property typically ascribed to violations of social convention in SDT (see, e.g., Turiel,

1983). This lends some preliminary support to the idea that PED use is viewed similarly to canonical violations of social convention: wrong in normative contexts where it is prohibited, but less wrong otherwise. We test this possibility, and the other nine articulated above, in Studies 2-11.

Several participants argued that consensus is insufficient to nullify the wrongness of steroid use. This argument resembles a property ascribed to paradigmatic moral violations in SDT, that their wrongness is independent of culture or popular opinion (see, e.g., Turiel, 1983, Royzman, Leeman, & Baron, 2009; Royzman, Landy & Goodwin, 2014). This suggests that SDT may provide a useful theoretical starting point for understanding opposition to steroid use. We return to this point in the General Discussion, after a series of exploratory studies testing for effects of the ten factors articulated above.

Studies 2-11

Having found support for the hypothesis that people will still oppose steroid use even when it does not provide an unfair advantage, we now turn to answering RQ2: in the absence of fairness violations, what does drive opposition to PED use? We separately tested for effects of each of the ten factors enumerated above in Studies 2-11. Before collecting any data, we determined that all manipulations that produced significant effects ($p < .05$) on wrongness ratings would be combined in a later study (Study 12) to test for interactive effects. Manipulations that produced non-significant effects ($p > .05$) would be excluded from this later study. Establishing this criterion *a priori* constrained our "researcher degrees of freedom" (Simmons, Nelson, & Simonsohn, 2011) for the tests that follow.

Method

Participants. After exclusions for incomplete responses, a total of 2,620 participants took part in these ten studies (1432 male, M_{Age} = 34.55 years, SD_{Age} = 11.01), with approximately 125 participants in each of 21 total between-subjects conditions. This sample size was determined before collecting data. According to power analyses conducted using the G*Power software package (Faul, Erdfelder, Lang, & Buchner, 2007), this sample size allows us to detect a medium-small effect (d = .35) with reasonable power (.79), and a medium effect (d = .50) with high power (.98).

Procedures and manipulation checks. Each study manipulated one of the ten factors above, in a between-subjects design. In all studies, participants read a variant of a scenario describing Joe, a professional athlete who begins using anabolic steroids. All of the scenarios were adapted from the same base scenario, presented below. In this scenario, fairness concerns were removed; as in the no advantage condition in Study 1, everyone in Joe's competitive circuit was currently using steroids:

> Joe is a professional, competitive weightlifter. Joe has never used performance-enhancing substances like steroids, but he is considering starting to use steroids to improve his performance. Joe is fully aware that he is the *only* person in his competitive weight-lifting circuit who is not currently using steroids. *All of the people Joe competes against* are currently using steroids. Knowing this, Joe decides to start taking anabolic steroids to improve his performance.

In what follows, the condition in each study that is intended to most closely match people's usual beliefs about steroids will be referred to as the "control" condition. The condition meant to remove a reason for opposition and reduce wrongness judgments will be called the "experimental condition" (with one exception in Study 11, see below). Participants rated how

wrong it was for Joe to use steroids, using the 1-9 scale from Study 1. After responding to the scenario, participants completed the same demographics questionnaire as in Study 1 and were debriefed, thanked, and paid.

Developing manipulations of each of the ten concerns above required substantial pre-testing. For each manipulation, we recruited approximately 10 participants per cell, and tested the effect of our draft manipulation on a face-valid manipulation check (all described in more detail below). Before running the full study, we verified that the manipulation produced at least a clear directional effect on the manipulation check. The number of pre-tests to obtain validated manipulations ranged from one to seven. Full details are presented in the Supplemental Materials. We describe each manipulation below, and full materials from each study and pre-test are presented in the Supplemental Materials.

Study 2: Laws and league rules. Study 2 modified the base scenario above by adding clauses explicitly indicating that steroid use "violates the law and the rules of [Joe's] circuit" (in the control condition) or "does not violate the law or the rules of [Joe's] circuit" (in the experimental condition). Following the wrongness question, participants responded to a manipulation check "did Joe break any rule or law by taking anabolic steroids?" on a separate page. Participants could respond "Yes", "No", or "Not sure". The manipulation was successful: 110 participants out of 127 (86.6%) in the control condition correctly indicated that Joe had broken a rule, and 113 out of 126 (89.6%) in the experimental condition correctly indicated that he had not, $\chi^2(2) = 166.2$, $p < .001$, $V = .81$.[4]

[4] Some of our manipulation checks used categorical responses and others used continuous responses, so we did not exclude on the basis of "failed" manipulation checks (because there is no obvious definition of "failure" for continuous measures). Manipulation checks were used only to confirm that our manipulations were successful, on average, across our samples.

Study 3: Prudence. The control condition in Study 3 consisted of the base scenario above, with an additional sentence at the end noting that "the steroid he decides to start using, like most steroids, has potential side effects, and poses some threat to Joe's health." The experimental condition replaced this with a sentence explaining that "the steroid he decides to start using is specially formulated to avoid the typical side effects of steroids, and poses no threat to Joe's health, in the short term or the long term". The experimental condition therefore removed from consideration any risk to Joe. After indicating how wrong it was for Joe to use steroids, participants responded to a manipulation check, "How dangerous to Joe's health is the steroid that he took?", on a separate page, on a 1-9 scale. The manipulation was successful: participants in the control condition indicated that the steroid was significantly more dangerous (M = 6.59, SD = 1.75) than did participants in the experimental condition (M = 2.65, SD = 2.25), $t(232.12) = 15.44$, $p < .001$, $d = 1.96$.

Study 4: Punishment. In Study 4, the control condition added a sentence to the end of the baseline scenario stating that "the steroid [Joe] decides to start using, like most steroids, poses some risk of detection, and therefore there is a chance that Joe will be caught and punished." The experimental condition replaced this with a sentence stating that "the steroid he decides to start using is specially formulated to avoid detection, and therefore there is no chance that Joe will be caught and punished." On the following page, participants responded to a manipulation check, "How likely do you think it is that Joe will be caught using steroids and punished for it?" on a 1-9 scale. Participants in the control condition indicated that it was more likely that Joe would be punished (M = 5.36, SD = 1.84) than did participants in the experimental condition (M = 3.78, SD = 2.22), $t(231.06) = 6.08$, $p < .001$, $d = .78$.

Study 5: Naturalness. The primary component of the lay understanding of "naturalness" has to do with how something comes into existence; human intervention makes something unnatural (Rozin, 2005; Rozin et al., 2012). However, we felt that an anabolic steroid that existed without any human intervention would seem implausible. Therefore, in Study 5, we manipulated a secondary element of lay conceptions of naturalness: "origin in nature" (Rozin et al., 2012). The stimuli in this study indicated that the steroids that Joe started using were "derived from chemicals developed in a laboratory" (in the control condition) or "derived from chemicals that naturally occur in the human body" (in the experimental condition). We reasoned that the latter would be seen as having its origin closer to nature than the former, despite both being developed via human intervention. On the following page, participants responded to a manipulation check, "How unnatural is it for Joe to use the anabolic steroids that he took?", on a 1-9 scale. Participants in the control condition indicated that Joe's steroid use was more unnatural (M = 6.62, SD = 1.97) than did participants in the control condition (M = 5.70, SD = 2.36), $t(242.40) = 3.36$, $p = .001$, $d = .42$, supporting this reasoning and demonstrating the effectiveness of the manipulation.

Study 6: Interfering with the body. The control condition in Study 6 was the unaltered base scenario presented above. The experimental condition presented a version of the scenario where Joe opted to start using a banned pair of gloves that improved his grip, enhancing his performance as much as steroids would. This manipulation was loosely inspired by "mechanical doping" in the sport of cycling (Associated Press, 2016), "sticky gloves" in football (Reimer, 2015), and other such technological innovations that improve athletic performance. The idea was to depict a method of improving performance that was effective, and highly unnatural, yet did not in any way tamper with the usual functioning of the body. On the following page,

participants responded to a manipulation check, “did Joe do anything that changed the normal functioning of his body?” Participants in the control condition indicated that Joe had changed his body’s normal functioning more (M = 7.27, SD = 2.49) than did participants in the experimental condition (M = 3.43, SD = 2.92), $t(246.13) = 11.25$, $p < .001$, $d = 1.41$.

Study 7: Effort. As noted above, steroid users must actually exert significant effort to gain muscle mass while taking steroids. A large part of how steroids aid the user is by allowing them to recover more quickly from a workout, so that they can train more frequently. We suspected that people may not know this, and that understanding that steroids require substantial effort to produce any effect, might reduce opposition to steroid use. The control condition in Study 7 was the unaltered base scenario. The experimental condition added one sentence explaining (in greatly simplified terms) how steroids affect the body: “The steroids work by allowing Joe’s body to recover more quickly from workouts, so that he can lift more often, allowing him to get stronger, faster.”

On the following page, participants responded to an open-ended manipulation check: “Please briefly explain how the steroids that Joe took improve his performance. What did they do? How do they work? If you are not sure, please indicate this.” Responses were coded by the second author, blind to condition, and by a research assistant, blind to condition and to the purpose of the study. Responses were coded as either explicitly indicating no understanding of how steroids affect the body, correctly indicating that they allow for faster recovery and/or more workouts, or indicating a different, incorrect mechanism. The coders agreed on their classification of 227 out of 245 responses (92.6%). For disagreements, we retained the classification from the research assistant, though disagreements were infrequent, so this does not substantively affect the results.

Of 120 participants in the control condition, 36 (30.0%) indicated that they did not know how steroids work, 4 (3.3%) approximated the correct mechanism, and 81 (67.5%) articulated an incorrect mechanism. Of 125 participants in the experimental condition, only 4 (3.2%) indicated that they did not know how steroids work, 92 (73.6%) approximated the correct mechanism, and 27 (21.6%) articulated an incorrect mechanism, $\chi^2(2) = 128.1$, $p < .001$, $V = .72$. So, we successfully conveyed to participants in the experimental condition that effort is still required while taking steroids.

Study 8: Competitive context. The control condition in Study 8 was the unaltered base scenario above. In the experimental condition, Joe was described as a "recreational, non-competitive" weightlifter instead of a "professional, competitive" weightlifter, and as the only person in his gym not using steroids, rather than the only person in his competitive circuit. In the experimental condition, not only did Joe not gain an advantage over anyone by using steroids, he was not even in a context where such an advantage was logically possible. On the following page, participants responded to a manipulation check, "did Joe take steroids to gain a competitive edge over someone else"? on a 1-9 scale. We considered this to be a conservative check on the effectiveness of this manipulation, because Joe did not *actually* gain an advantage over anyone in either condition. Despite this, participants in the control condition more strongly endorsed the idea that Joe took steroids to gain a competitive advantage ($M = 5.68$, $SD = 2.99$) than did participants in the experimental condition ($M = 3.93$, $SD = 2.88$), $t(250) = 4.73$, $p < .001$, $d = .60$.

Study 9: Athletes as role models. In Study 9, we manipulated whether or not Joe was a well-known athlete, "considered to be a role model for many young athletes" (in the control condition) or "not considered to be a role model for any young athletes" (in the experimental condition). Following the wrongness question, participants responded to a manipulation check,

"how likely is Joe's decision to use anabolic steroids to negatively influence other people's behavior?" on a 1-9 scale. Participants in the control condition considered this to be more likely (M = 5.59, SD = 2.28) than did participants in the experimental condition (M = 4.51, SD = 2.47), $t(252) = 3.64$, $p < .001$, $d = .45$.

Study 10: Fairness to past competitors. In Study 10, we specified that Joe competes in an event that originated 15 years ago. The control condition was otherwise identical to the base scenario, providing no information about past competitors. The experimental condition, however, specified that "every other competitor in the 15-year history of Joe's event has used steroids", removing any backward-facing concerns about fairness to past competitors. The manipulation check in this study asked "is Joe's use of steroids unfair to any record-holders, past or present?", on a 1-9 scale. Participants in the control condition considered Joe's steroid use to be more unfair (M = 6.05, SD = 2.54) than participants in the experimental condition (M = 4.53, SD = 2.94), $t(239.62) = 4.33$, $p < .001$, $d = .55$.

Study 11: Terminology. Unlike Studies 2-10, Study 11 consisted of three conditions, a control condition modeled closely on the base scenario above, and two experimental conditions. In one experimental condition, the term "steroids" was replaced with "performance-enhancing drugs" or "drugs" throughout the scenario. In the other experimental condition, the term "steroids" was replaced with "performance-enhancing supplements" or "supplements". After responding to the wrongness question, participants were asked, "based on what you know from the scenario on the previous page, what kind of performance-enhancing substance did Joe decide to start using? (check all that apply)". This was followed by five options: "an anabolic steroid", "a drug", "a supplement", "other" and "not sure". The correct answer was the most common answer in all three conditions (steroids condition: 97.58% of participants; drugs condition:

37.69%, supplements condition: 49.59%). Because participants could select more than one response, we analyzed responses using a 3 (condition) x 5 (response) mixed-model ANOVA, with repeated measures on the latter factor. There was a significant interaction, $F(8, 1488) = 62.67$, $p < .001$, $\eta_p^2 = .25$, which, in combination with the pattern of responses mentioned above, indicates that our manipulation was successful.[5]

Results

Descriptive and inferential statistics for the wrongness ratings from Studies 2-11 are presented in Table 1 and the distributions of responses are graphed in Figure S1 in the Supplemental Materials. Across the ten studies, only the manipulations in Study 2 (laws and league rules) and Study 3 (prudence) significantly affected wrongness ratings. In agreement with prior research (Hester et al., 2015; Scheske & Schnall, 2012), the manipulation of naturalness in Study 5 approached significance, though this effect size was considerably smaller than the effects observed in Studies 2 and 3. We also observed a marginally significant effect of effort in Study 7, though in the direction opposite our predictions. We return to this in the General Discussion.

[5] Because the dependent variables in this analysis were binary, ANOVA is not strictly appropriate. We followed up on this analysis with a repeated-measures binary logistic regression. The expected interactions, between Condition = Drugs and Check = Drug, and between Condition = Supplements and Check = Supplement, were observed, *B*s 6.47 and 7.96, respectively, *p*s < .001. This indicates that participants were more likely to indicate that Joe had taken a drug in the "Drug" condition than the "Steroids" condition, and more likely to indicate that he had taken a supplement in the "Supplement" condition than the "Steroids" condition.

Table 1. Descriptive and inferential statistics for wrongness judgments in Studies 2-11.

		Condition Means and SDs		
Study	**Manipulated Factor**	**Control**	**Experimental**	**Inferential Statistics**
2	Laws and League Rules	6.18 (2.60)	3.23 (2.61)	$t(251) = 9.01$, $p < .001$, $d = 1.13$
3	Prudence	6.66 (2.23)	5.02 (2.24)	$t(247) = 5.77$, $p < .001$, $d = 0.73$
4	Punishment	6.37 (2.24)	6.44 (2.22)	$t(244) = -0.24$, $p = 0.809$, $d = -0.03$
5	Naturalness	5.76 (2.44)	5.16 (2.53)	$t(247) = 1.92$, $p = 0.056$, $d = 0.24$
6	Interfering with the Body	5.56 (2.38)	5.57 (2.19)	$t(250) = -0.05$, $p = 0.962$, $d = -.00$
7	Effort	5.41 (2.41)	5.93 (2.37)	$t(243) = -1.70$, $p = 0.090$, $d = -0.22$
8	Competitive Context	5.82 (2.41)	5.91 (2.75)	$t(250) = -0.27$, $p = 0.784$, $d = -0.03$
9	Athletes as Role Models	5.95 (2.26)	5.77 (2.40)	$t(252) = 0.59$, $p = 0.556$, $d = 0.08$
10	Fairness to Past Competitors	5.88 (2.42)	5.45 (2.40)	$t(243) = 1.38$, $p = 0.169$, $d = 0.18$
11	Terminology	"Steroids": 5.41 (2.40)	"Drugs": 5.55 (2.12) "Supplements": 4.93 (2.22)	$F(2, 372) = 2.66$, $p = 0.071$, $\eta_p^2 = 0.014$

Discussion

Across Studies 2-11, we found that only the presence or absence of rules and laws prohibiting steroid use and the risk posed to the user reliably attenuated opposition to steroid use. Following our original research plan, we next crossed these manipulations with a manipulation of fairness modeled after Study 1, to test for interactive effects.

Study 12

Method

Participants. The sample size for this study was determined based on an *a priori* power analysis using the MorePower software package (Campbell & Thompson, 2012), which can compute statistical power in complex factorial designs like the 2 x 2 x 2 x 2 design of this study. We aimed to recruit at least 50 participants per cell, for a total of 800 participants across 16 cells. This sample size provides good statistical power (.98) to detect an effect size of $\eta^2 = .02$, conventionally considered to be a "small" effect (Cohen, 1988), and adequate power (.80) to detect a smaller effect of $\eta^2 = .01$.

We recruited an initial wave of $N = 102$ participants to examine failure rates on the attention check included in the study (described below). No analyses were conducted on these participants' data, other than tallying failed attention checks, until after the full sample was recruited. Five of these participants (4.9%) failed the attention check. We therefore aimed to over-recruit by approximately 5 percent, for a total of at least 840 participants, to obtain a usable sample of at least 800. Our final sample size was $N = 848$. Twenty-four participants failed the attention check (2.8%), leaving a final, usable sample of $N = 824$ (396 male, $M_{Age} = 35.43$ years, $SD_{Age} = 11.81$).

Procedure. This study used a 2 (fairness: advantage vs. no advantage) x 2 (prudence: risk vs. no risk) x 2 (legality: illegal vs. legal) x 2 (league rules: banned vs. permitted) between-subjects design. We manipulated legality and league rules separately, as a pre-study (Study S1 in the Supplemental Materials) suggested that neither of these regulations was likely to be fully responsible for the results of Study 2. Stimuli were again based on the scenario used in Studies 2-11: the fairness manipulation was modeled after Study 1, the prudence manipulation, Study 3, and the legality and league rules manipulations, Studies 2 and S1. Because it was always our intention to eventually run a large study manipulating whichever factors produced reliable effects, this scenario was designed to accommodate all of these manipulations with no major revisions. Full materials are presented in the Supplemental Materials.

After consenting to participate, participants were randomly assigned to one of 16 conditions, and presented with the appropriate version of the scenario. Participants then indicated how wrong it was for Joe to use steroids, on the same 1-9 scale used in prior studies. On the next page, participants responded to four manipulation checks, "did Joe gain an unfair advantage over anyone" (1-9 scale), "did Joe break any law by taking anabolic steroids" (yes/no/not sure), "did Joe break any rules of his competitive circuit by taking anabolic steroids" (yes/no/not sure), and "how dangerous to Joe's health is the steroid that he took" (1-9 scale). Participants then completed the demographics questionnaire from the prior studies. On the final page of the survey, participants were presented with a surprise attention check, where they had to select Joe's profession from among five choices, with the correct answer being "Joe is a professional athlete". After responding to the attention check, participants were debriefed, thanked, and paid.

Results

Manipulation checks. All four manipulations were successful. Participants much more strongly believed that Joe had "gained an unfair advantage" when no one else was using steroids (i.e., in the advantage condition, $M = 7.32$, $SD = 2.18$) than when everyone else was (i.e., in the no advantage condition, $M = 3.55$, $SD = 2.66$), $t(764.98) = 22.19$, $p < .001$, $d = 1.56$). Also, participants indicated that steroid use was more "dangerous to Joe's health" when the steroids were not specially formulated to be without health risks (i.e., in the risk condition, $M = 6.77$, $SD = 1.62$) than when they were (i.e., in the no risk condition, $M = 3.04$, $SD = 2.50$), $t(653.74) = 25.10$, $p < .001$, $d = 1.79$. A majority of participants (84.9%) correctly said that Joe had broken a law in the illegal condition, and a majority (88.4%) correctly said that he had not, in the legal condition, $\chi^2(2) = 615.74$, $p < .001$, $V = .86$. Lastly, a majority of participants (95.5%) correctly said that Joe had broken a rule of his competitive circuit when steroids were banned, and a majority (86.5%) correctly said that he had not, when they were permitted, $\chi^2(2) = 538.10$, $p < .001$, $V = .81$. As would be expected, effects of the manipulations on other manipulation checks were considerably smaller, and are summarized in the Supplemental Materials.

Wrongness ratings. Condition means and standard deviations are presented in Table 2. As in our previous studies, all four manipulations had significant main effects on wrongness ratings. First, participants more strongly opposed Joe's use of steroids when it gave him a competitive advantage ($M = 6.68$, $SD = 2.46$) than when he was just keeping up with his competitors ($M = 5.76$, $SD = 2.56$), $F(1, 808) = 35.37$, $p < .001$, $\eta_p^2 = .042$. They also considered using steroids that presented some risk ($M = 6.61$, $SD = 2.40$) to be more wrong than using steroids that were entirely safe ($M = 5.83$, $SD = 2.64$), $F(1, 808) = 31.38$, $p < .001$, $\eta_p^2 = .037$. Also, using steroids that were illegal ($M = 6.97$, $SD = 2.17$) was considered more wrong than using steroids that were legal ($M = 5.46$, $SD = 2.68$), $F(1, 808) = 112.31$, $p < .001$, $\eta_p^2 = .122$.

Lastly, using steroids that were banned by the sport (M = 7.24, SD = 2.17) was considered more wrong than using steroids that were permitted (M = 5.17, SD = 2.70), $F(1, 808) = 204.12$, $p < .001$, $\eta_p^2 = .204$. League rules produced a medium-large effect (Cohen, 1988), and laws produced a medium effect, whereas fairness and prudence produced relatively small, though highly significant, effects.

Table 2. Descriptive statistics for wrongness judgments in Study 12 (1-9 scale) and Study 13 (0-100 scale).

				Study 12		**Study 13**	
League/University Rules	**Legality**	**Prudence**	**Fairness**	***M***	***SD***	***M***	***SD***
Banned	Illegal	Risk	Advantage	8.23	1.10	75.04	24.50
			No Advantage	7.30	1.98	75.69	25.78
		No Risk	Advantage	7.90	1.49	73.82	23.71
			No Advantage	6.92	2.18	69.02	24.02
	Legal	Risk	Advantage	7.89	1.62	73.32	21.80
			No Advantage	6.31	1.65	67.89	22.45
		No Risk	Advantage	6.84	2.03	67.81	23.61
			No Advantage	6.00	2.11	58.50	23.53
Permitted	Illegal	Risk	Advantage	6.74	2.24	73.93	23.17
			No Advantage	6.89	2.14	69.63	22.36
		No Risk	Advantage	6.02	2.40	68.89	24.50
			No Advantage	5.24	2.31	64.47	26.97
	Legal	Risk	Advantage	4.86	2.55	66.03	23.34
			No Advantage	4.10	2.52	63.87	25.04
		No Risk	Advantage	4.04	2.65	50.41	31.11
			No Advantage	2.81	2.09	45.84	29.81

Two significant two-way interactions also emerged in the this analysis, between league rules and legality, $F(1, 808) = 24.43$, $p < .001$, $\eta_p^2 = .029$, and league rules and prudence, $F(1, 808) = 4.24$, $p = .040$, $\eta_p^2 = .005$. The former interaction is plotted in Figure 2. Simple effects tests revealed that the presence or absence of a law prohibiting steroid use reliably affected wrongness judgments whether the governing body of the sport banned steroid use, $F(1, 808) = 16.46$, $p < .001$, $\eta_p^2 = .020$, or permitted it, $F(1, 808) = 117.40$, $p < .001$, $\eta_p^2 = .127$. However, the effect was over six times as large in the latter case; participants gave substantially more weight to the (il)legality of steroid use when rules governing the sport did not prohibit it.

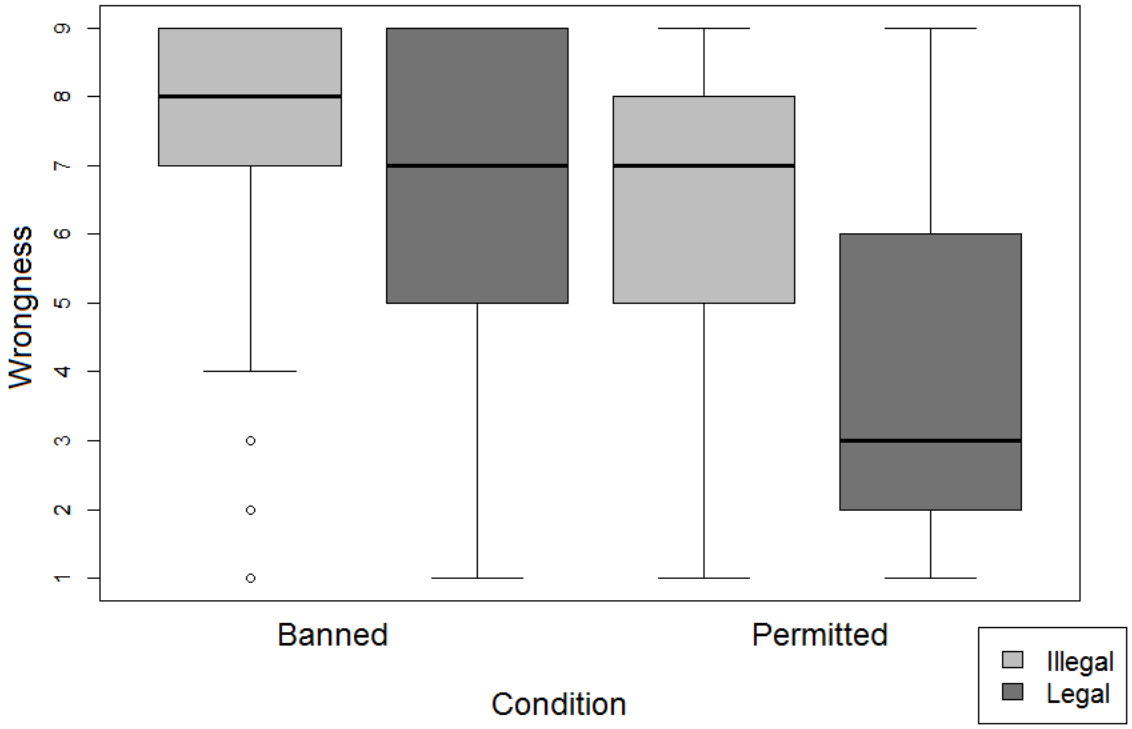

Figure 2. Effects of legality and league rules on wrongness ratings, Study 12.

The league rules x prudence interaction is plotted in Figure 3. Similar to the interaction described above, simple effects tests found that whether or not using steroids posed a risk reliably affected wrongness judgments, whether steroid use was banned by the sport, $F(1, 808) =$

6.46, $p = .011$, $\eta_p^2 = .008$, or permitted, $F(1, 808) = 28.54$, $p < .001$, $\eta_p^2 = .034$. But, the effect was over four times larger in the latter case. In other words, the effects of laws and prudence on opposition to steroid use both depended on whether it was permitted by the sport's governing authority, whereas fairness affected judgments independent of this consideration. No other interactions were statistically significant, Fs(1, 808) < 3.18, ps > .07, η_p^2s < .005.

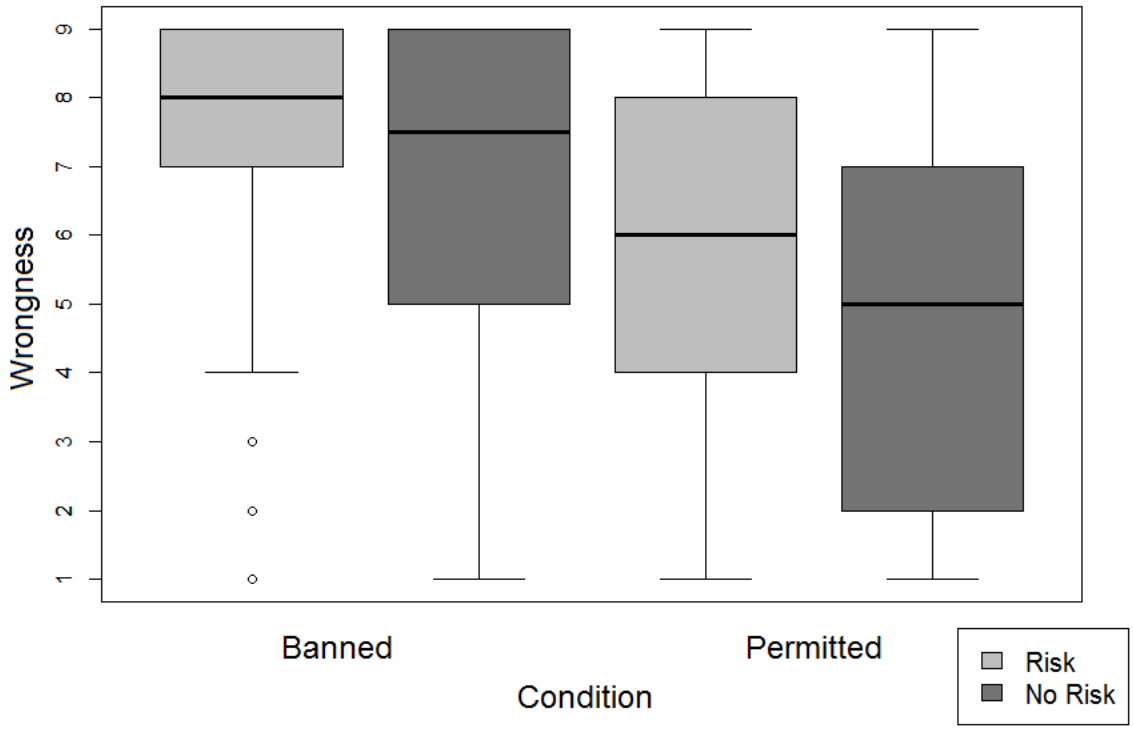

Figure 3. Effects of league rules and prudence on wrongness ratings, Study 12.

Perhaps the most meaningful finding in this study is that the four manipulations, in combination, nearly eliminated opposition to PED use. As we would expect, the advantage/risk/illegal/banned condition, which we think most closely matches people's usual beliefs about steroid use, elicited the most opposition of any condition in this study ($M = 8.23$, $SD = 1.10$). The no advantage/no risk/legal/permitted condition removes every consideration that we have found to reliably affect normative judgments and elicited the least opposition ($M =$

2.81, *SD* = 2.09), $t(58.57) = 15.57$, $p < .001$, $d = 3.43$. We take the large difference in wrongness ratings between these conditions to indicate that we have successfully identified much of what underlies opposition to steroid use.

Discussion

Like Studies 1, 2, and 3, Study 12 found that fairness, prudence, laws, and league rules all affect normative judgments of steroid use. Study 12 also revealed two intriguing interactions. The effects of the legality and prudence of steroid use were both substantially larger when steroid use was permitted under the sport's rules.

Overall, this study indicates that the most important driver of opposition to steroid use is the presence or absence of legitimate rules prohibiting it. League rules and laws produced considerably larger effects on normative judgments than fairness and prudence, and the effects of laws and prudence depended on league rules. This suggests that PED use is conceptualized as being somewhat like a particularly serious violation of social convention, which would not be (as) wrong if it were not prohibited. We elaborate on this point in the General Discussion.

Study 13

Studies 1-12 all examined normative judgments of a particular type of performance enhancement (anabolic steroid use), in a particular context (professional sports), by a particular target (a male athlete named Joe). These studies are a sort of "experimental case study" in that they examine a single, paradigmatic case of PED use in great depth.[6] But, a case study is more useful if the insights that it generates are found to be applicable to other cases as well. We therefore felt that it was important to test whether the results of these studies generalize to a

[6] We thank an anonymous reviewer for suggesting this very useful way of conceptualizing these studies.

different kind of PED (cognitive-enhancement drugs), a different context (academics), and a different target (a female college student named Anne). Study 13 was a conceptual replication of Study 12 designed to examine whether the effects of fairness, prudence, laws, and rules (in this case, a university's rules) would generalize to these new circumstances. The materials, procedure, and analysis plan for Study 13 were preregistered on the Open Science Framework (https://osf.io/8sfvv/).

Method

Participants. As noted above, Study 13 was conducted four months after Studies 1-12 concluded. Therefore, we did not exclude participants who had participated in a prior study. Participants were recruited in the same manner as the prior studies, and after exclusions for incomplete responses, we obtained a final sample of $N = 1{,}044$.[7] Thirteen participants failed the attention check and 12 more did not answer the normative judgment question, leaving a final, usable sample of $N = 1{,}019$.

Procedure. The design of this study was a 2 (fairness: advantage vs. no advantage) x 2 (prudence: risk vs. no risk) x 2 (legality: illegal vs. legal) x 2 (university rules: banned vs. permitted) between-subjects design, modeled after Study 12. The stimuli were variations of a new scenario describing a college student, "Anne", who bought prescription stimulants from another student and started taking them to improve her performance in a highly competitive class. The stimulants were described as being similar to Adderall or Ritalin. Full materials are

[7] Our initial intention (detailed in our pre-registration) was to recruit a sample of roughly the same size as Study 12. However, on the day that we ran this study, there were widespread technical issues on Amazon Mechanical Turk, due to problems with Amazon Web Services (O'Brien, 2017). This resulted in many participants completing the survey, but being unable to submit the work for payment; approximately 200 participants submitted responses, but did not count toward the target sample of $N = 840$ on Mechanical Turk. We retained all of the data that were collected, resulting in a larger sample size than we initially planned.

presented in the Supplemental Materials, with descriptions of several pretests that we conducted while developing this scenario. Briefly, these pre-tests found that participants performed well on a series of comprehension checks, so we can be confident that the scenario was comprehensible to members of the population that we sampled in the final study.

After consenting to participate, participants were randomly assigned to one of 16 conditions, and presented with the appropriate version of the scenario. Participants then rated how right or wrong it was for Anne to use stimulants to improve her academic performance on a bipolar sliding scale. The scale ranged from "the right thing to do" (coded as 0) to "the wrong thing to do" (coded as 100), with the midpoint labeled "neither right nor wrong". Numeric values were not presented to participants. The prior studies' unipolar scale, ranging from "not at all wrong" to "extremely wrong", may have communicated to participants that they were supposed to consider PED use to be wrong to some degree. The new bipolar scale allowed participants to provide positive, negative, or neutral normative judgments. After responding to the scenario, participants then answered manipulation checks based on those in Study 12 on a separate page, then completed the demographics questionnaire from the prior studies. On the final page of the survey, participants were presented with a surprise attention check, where they had to select Anne's profession from among five choices, with the correct answer being "Anne is a college student". After responding to the attention check, participants were debriefed, thanked, and paid.[8]

Results

[8] Participants who informed us that they could not submit the study for payment due to the issues with Amazon Web Services were later paid separately.

Manipulation checks. All four manipulations were successful. Participants indicated that Anne had gained more of an "unfair advantage" when she was the only person in her class taking stimulants (i.e., in the advantage condition, $M = 6.85$, $SD = 2.07$) than when all of her classmates were taking them as well (i.e., in the no advantage condition, $M = 3.38$, $SD = 2.55$), $t(976.86) = 23.89$, $p < .001$, $d = 1.50$). Also, participants indicated that using stimulants was more "dangerous to Anne's health" when they had potential side effects (i.e., in the risk condition, $M = 5.47$, $SD = 1.63$) than when they did not (i.e., in the no risk condition, $M = 2.42$, $SD = 2.02$), $t(981.79) = 26.52$, $p < .001$, $d = 1.66$. A majority of participants (83.0%) correctly said that Anne had broken a law in the illegal condition, and a majority (92.9%) correctly said that she had not, in the legal condition, $\chi^2(2) = 540.97$, $p < .001$, $V = .73$. Lastly, a majority of participants (83.8%) correctly said that Anne had had violated the rules of her university in the banned condition, and a majority (94.4%) correctly said that she had not in the permitted condition, $\chi^2(2) = 571.87$, $p < .001$, $V = .75$. Effects of the manipulations on other manipulation checks were considerably smaller, and are summarized in the Supplemental Materials.

Normative Judgments. Condition means and standard deviations are presented in Table 2. All four manipulations had significant main effects on participants' normative judgments, conceptually replicating Study 12 with a different target of judgment, and a different kind of performance enhancement. First, participants more strongly opposed Anne's use of stimulants when it gave her a competitive advantage ($M = 68.91$, $SD = 25.34$) than when all of her classmates were also using stimulants and she was just keeping up ($M = 64.36$, $SD = 26.16$), $F(1, 1003) = 7.55$, $p = .006$, $\eta_p^2 = .007$. They also considered using stimulants that presented some risk to Anne's health ($M = 70.68$, $SD = 23.72$) to be more wrong than using stimulants with no such risk ($M = 62.66$, $SD = 27.21$), $F(1, 1003) = 28.43$, $p < .001$, $\eta_p^2 = .028$. Also, using

stimulants that were illegal to take without a prescription (M = 71.04, SD = 24.62) was considered more wrong than using stimulants that were legal (M = 62.23, SD = 26.31), $F(1, 1003) = 37.79$, $p < .001$, $\eta_p^2 = .036$. Lastly, using stimulants was considered more wrong when it was against the university's rules (M = 69.76, SD = 24.10) than when it was permitted (M = 63.58, SD = 27.12), $F(1, 1003) = 21.54$, $p < .001$, $\eta_p^2 = .021$. The effect sizes in this analysis are noticeably smaller than their analogs from Study 12; however, all four manipulated variables still produced reliable effects on normative judgments of using prescription stimulants to improve academic performance.

We observed a significant interaction between university rules and legality, $F(1, 1003) = 3.91$, $p = .048$, $\eta_p^2 = .048$ (see Figure 4). Similar to Study 12, simple effects tests found that the legality of using prescription stimulants significantly affected normative judgments whether it violated university rules, $F(1, 1003) = 8.60$, $p = .003$, $\eta_p^2 = .009$, or not, $F(1, 1003) = 33.35$, $p < .001$, $\eta_p^2 = .032$. Though, the effect was over three times larger in the latter case.

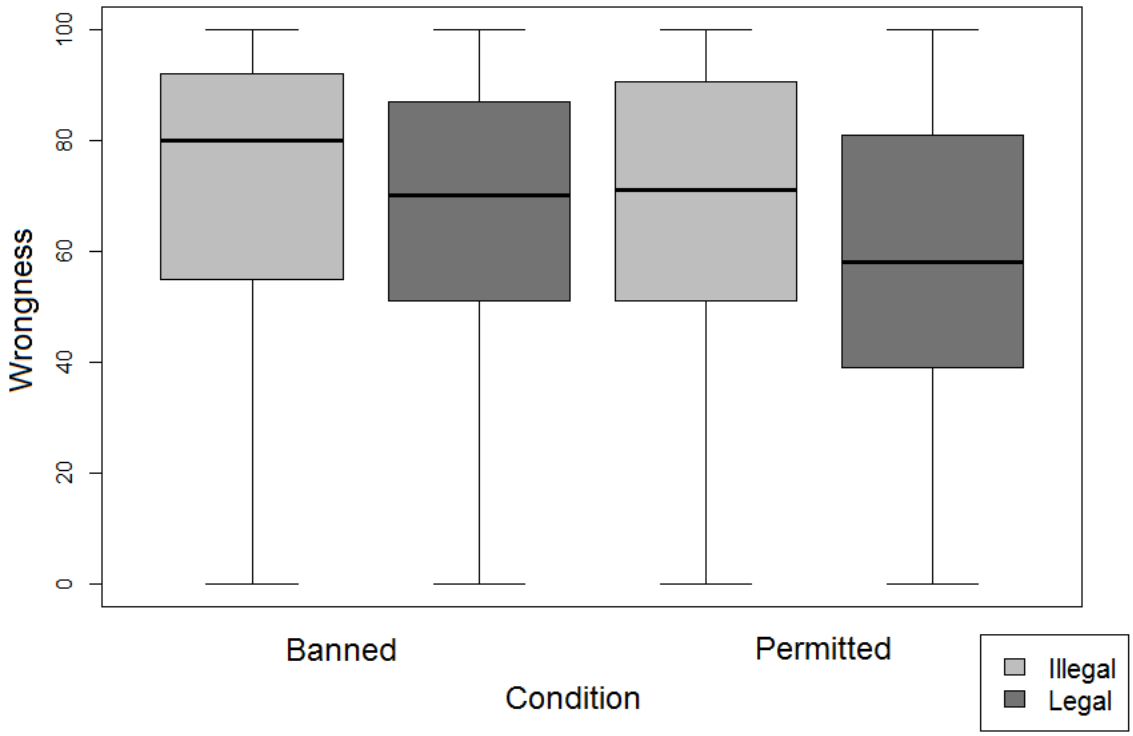

Figure 4. Effect of university rules and legality on normative judgments, Study 13.

The interaction between league/university rules and prudence observed in Study 12 was only marginally significant in this study, $F(1, 1003) = 2.84$, $p = .092$, $\eta_p^2 = .003$. Simple effects tests found a pattern similar to Study 12: whether taking stimulants was safe or not affected normative judgments whether it violated university rules, $F(1, 1003) = 6.58$, $p = .010$, $\eta_p^2 = .007$, or not, $F(1, 1003) = 24.88$, $p < .001$, $\eta_p^2 = .024$. However, the effect was over three times larger in the latter case. This interaction is illustrated in Figure 5.

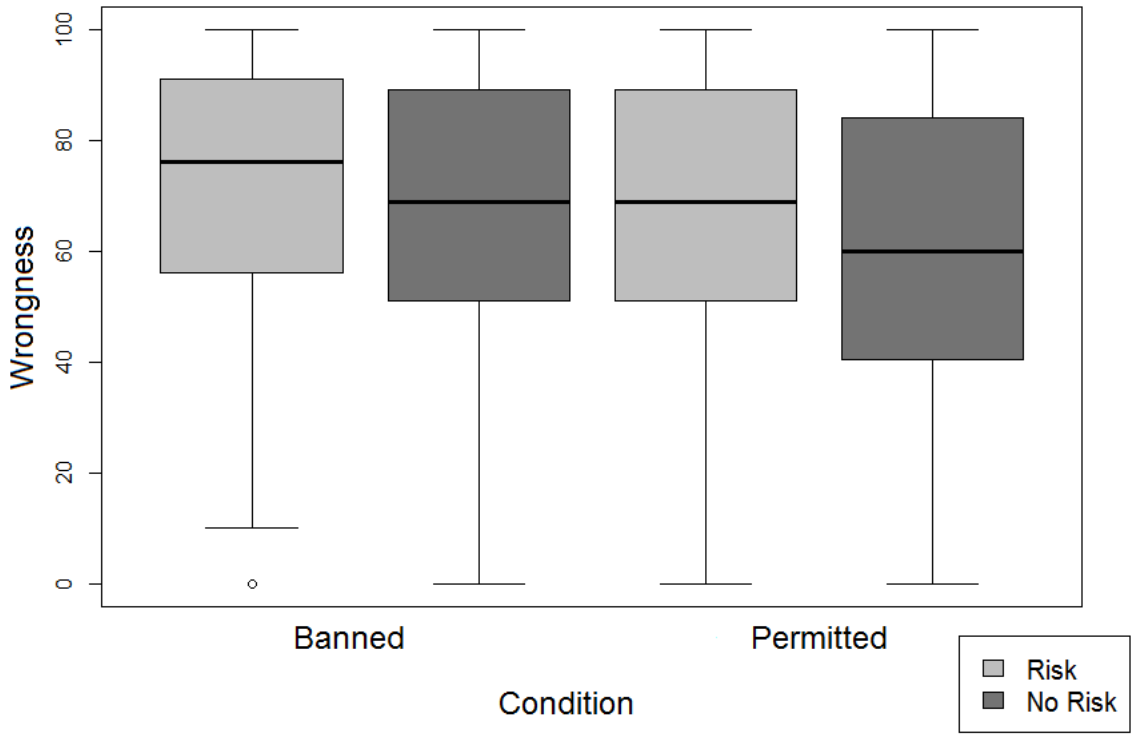

Figure 5. Effects of university rules and prudence on normative judgments, Study 13.

One more interaction emerged as significant in this analysis, between legality and prudence, $F(1, 1003) = 5.95$, $p = .015$, $\eta_p^2 = .006$.[9] This interaction is illustrated in Figure 6. Simple effects tests found that the risk, or lack thereof, associated with using stimulants significantly affected normative judgments whether doing so was illegal, $F(1, 1003) = 4.16$, $p = .042$, $\eta_p^2 = .004$, or legal, $F(1, 1003) = 30.36$, $p < .001$, $\eta_p^2 = 029$. Though, the effect was over seven times larger in the latter case. No other interaction approached significance, Fs$(1, 1003) < 1.74$, ps $>.18$, η_p^2s $< .003$.

[9] The analogous interaction in Study 12 did not approach significance, $F(1, 808) = 0.10$, $p = .749$, $\eta_p^2 < .001$.

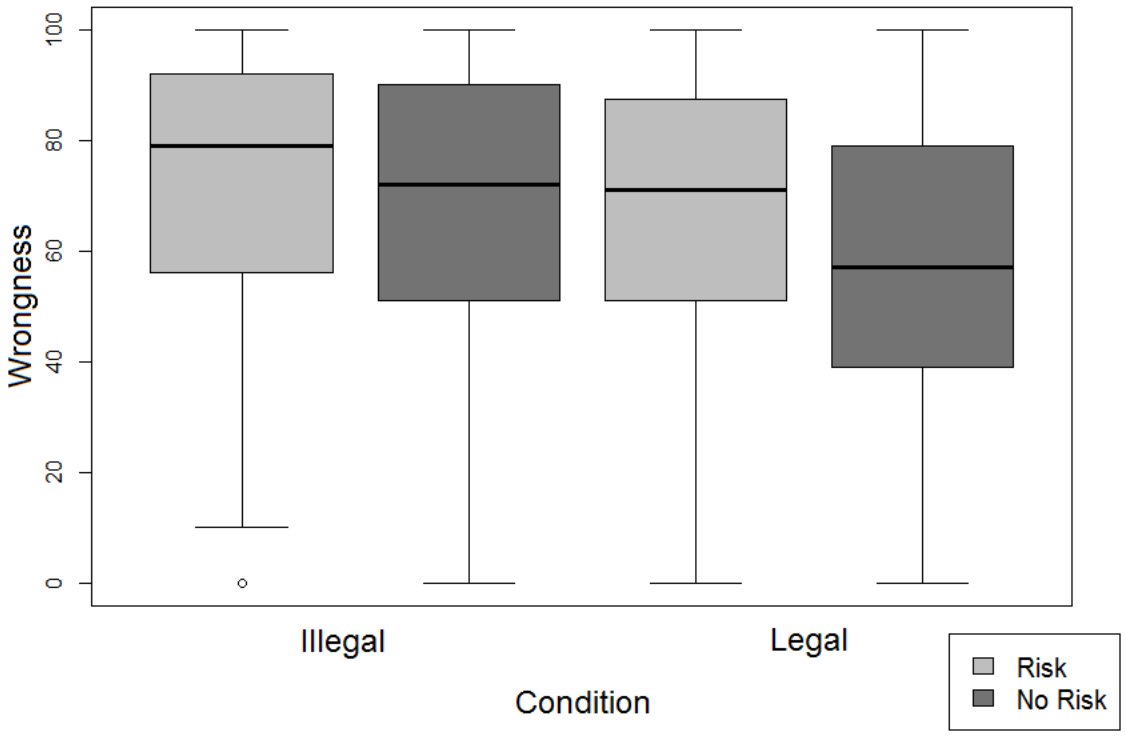

Figure 6. Effects of legality and prudence on normative judgments, Study 13.

As in Study 12, we compared the advantage/risk/illegal/banned condition with the no advantage/no risk/legal/permitted condition, which removes every consideration that we have found to reliably affect normative judgments. The former prompted substantial opposition (M = 75.04, SD = 24.50),[10] and the latter elicited less opposition than any other condition in this study, with mean ratings hovering around the scale midpoint, labeled "neither right nor wrong" (M = 45.84, SD = 29.81), $t(109) = 5.63$, $p < .001$, $d = 1.07$. The effect size from this comparison is notably smaller than its analog in Study 12. This seems to be because normative judgments in the advantage/risk/illegal/banned condition in Study 12 were nearly at ceiling, whereas in Study 13, they were about halfway between "neither right nor wrong" and "the wrong thing to do".

[10] Unlike in Study 12, the advantage/risk/illegal/banned condition did not elicit the most severe normative judgments in this study; the no advantage/risk/illegal/banned condition prompted slightly more opposition (M = 75.69, SD = 25.78), though this difference did not approach statistical significance, $t(112) = 0.14$, $p = .889$, $d = .03$.

This is consistent with some research finding that PED use is more strongly opposed in the athletic than the academic domain (Dodge et al., 2012; Sabini & Monterosso, 2005). The key observation here is that, as in Study 12, when Anne gained no advantage, did not endanger her health, and did not break any laws or rules, opposition to her use of PEDs was eliminated, with her actions being rated as "neither right nor wrong".

Discussion

Study 13 largely replicated the results of Study 12, using a new stimulus describing the use of prescription stimulants to enhance cognitive performance (rather than anabolic steroids to enhance physical performance), by a new target of judgment, in an academic context (rather than an athletic context). Once again, manipulations of laws and university rules produced larger effects than the manipulation of fairness, though in this study, the prudence manipulation produced a comparably-sized effect. We also found that the effect of prudence depended on the presence or absence of both laws and university rules. These results converge with those of the prior studies to suggest that the presence of legitimate rules and laws prohibiting PED use is the primary – though not the only – driver of judgments opposing it.

General Discussion

The present research used an exploratory approach to investigate why people oppose the use of performance-enhancing drugs, particularly anabolic steroids. Most of the past research on this topic has focused on concerns relating to fairness, though some recent work has explored predictors of normative judgments other than fairness. We presented the first direct evidence that people oppose PED use to a substantial degree even when it does not violate fairness norms (Studies 1, 12, and 13). We then explored the effects of ten potentially relevant factors, drawn

from the moral psychology literature, participants' thought listings, and popular press writings on steroid use, on people's normative judgments (Studies 2-13).

We found support for the most common theoretical claim in this area of research, that fairness concerns (at least partially) drive opposition to PED use, but the effects of fairness were relatively small. Two other factors also emerged as strong predictors of opposition, alongside fairness: the presence (or absence) of laws and/or rules instituted by the governing body of a sport prohibiting the use of PEDs, and the risk (or lack thereof) posed to the user, with the former typically exerting a larger influence on normative judgments. In the language of Social Domain Theory (SDT), steroid use can be viewed as primarily a violation of social convention (i.e., legitimate, but changeable rules), but also, simultaneously, as a violation of morality (i.e., fairness) and prudence (i.e., avoiding harm to the self). We return to this point shortly.

Although there is some prior research on normative judgments of PED use, this is the first work to directly demonstrate that opposition persists in the clear absence of fairness violations, and to comprehensively test so many possible drivers of these judgments. Our results shed new light on an area of normative judgment that has received widespread attention in the popular media, but is relatively understudied in moral psychology.

Connection to Prior Research, and Remaining Questions

In agreement with most research on normative judgments of PED use, we found a reliable effect of fairness on opposition. Beyond this, we observed a sizable effect of legitimate laws and rules on these judgments (Sattler et al., 2015). We also replicated the finding that prudential concerns predict opposition to PED use (Sattler et al., 2015; Scheske & Schnall, 2012).

However, our results also differ from other findings in this area in several ways. First, we did not find that the effort required to achieve results predicts opposition to PED use (cf. Hester et al., 2015). If anything, this effect was marginally significant in the direction opposite our predictions, with participants rating steroid use as *more* wrong after learning that considerable effort is required to produce the desired results. This may be due to participants in the experimental condition having more information, leading them to be more confident in producing negative normative judgments of steroid use (see Oskamp, 1965).

Also, we observed only a marginal effect of naturalness on normative judgments, and the effect size was fairly small (cf. Scheske & Schnall, 2012). Of course, even in the fair/no risk/legal/permitted condition in Study 12, however, wrongness ratings were not at floor.[11] This suggests that we have identified most of what accounts for opposition to PED use, but perhaps not everything. We suspect that perceptions of naturalness do affect these normative judgments, though the effect is relatively small and is only likely to be reliably detected in well-powered designs. We adhered in this research to our original plan to only include in Studies 12 and 13 factors that produced statistically significant effects, but we view exploring interactive effects between naturalness and the factors that we have identified as an important future task.

Lastly, none of our studies manipulated how steroids were administered, which was found to affect judgments of cognitive enhancement drugs (Scheske & Schnall, 2012; see also Hall, Sowden, & Ellsworth, 2017). Whereas this finding is informative, it does not seem central to the current investigation. We are studying why people are opposed to the use of PEDs in general, rather than what *kinds* of PEDs people are more opposed to than others. We did,

[11] Wrongness ratings in this condition were still significantly greater than the scale minimum of 1.0, $M = 2.18$, $SD = 2.09$, $t(42) = 5.71$, $p < .001$, $d = .87$, likely due to the high statistical power of this study.

however, consider that PED use might be seen as disgusting. Theoretically, injecting oneself with steroids—as opposed to taking a pill—is a "body envelope violation" that should engender disgust (Rozin, Haidt, & McCauley, 2000). Some theories argue that the experience of disgust should then produce more severe normative judgments (see, e.g., Haidt, 2001; Schnall, Haidt, Clore, & Jordan, 2008; Wheatley & Haidt, 2005; though see Case, Oaten, & Stevenson, 2012; Johnson et al., 2016; Landy & Goodwin, 2015). Several pretests found that disgust did not respond to manipulations of how steroids were administered, so we ultimately abandoned this idea.[12] However, considering that our four focal manipulations essentially eliminated opposition to PED use in Studies 12 and 13, any effect of administration method would likely be fairly insubstantial. Aggregating across four pretests (total $N = 79$), we found this to be the case, $d = .15$, 95% CI [-.29, .59].

This work represents the most extensive investigation to date of the underlying causes of opposition to PED use. However, as some of our results are not obviously consistent with prior research, there remains considerable work to be done to fully explicate this domain of lay normative judgment.

Limitations

[12] We pretested three different manipulations of disgust where Joe injected himself with steroids, or took them in the form of a pill, and two different face-valid manipulation checks. Across three pre-tests, participants reported quite low levels of feeling "grossed out" (a lay term that captures the theoretical meaning of disgust, see Nabi, 2002), with a maximum mean across the three studies of 4.0 on a 1-9 scale. Also, in one of these pretests, participants reported feeling more grossed out when Joe took steroids as a pill than when he took them as an injection. A fourth pretest used an alternative measure of disgust, where participants viewed a canonical disgust facial expression (Ekman & Friesen, 1975) and indicated how well it matched their feelings about Joe's steroid use (see Royzman, Atanasov, Landy, Parks, & Gepty, 2014 and Rozin et al., 1999 for similar measures). In this pretest, endorsement of the disgust face was again quite low, and was higher in the pill condition (M = 4.50) than in the needle condition (M = 3.90). Aggregating the data from these four pretests reveals only a small effect of the manipulation on reported disgust, $d = .10$, 95% CI: [-.34, .54], and low levels of reported disgust, $M_{Needle} = 3.35$, $M_{Pill} = 3.10$. In our "Joe" scenario, at least, how steroids are administered did not appreciably affect feelings of disgust. So, we did not pursue this further.

Although the exploratory, "deep dive" approach that we have taken here allows researchers to make rapid progress in understanding a psychological phenomenon, it also has some clear limitations. Most significantly, the majority of our studies relied on minor variations on the same stimulus, so we do not know how generalizable our results are. The fact that we largely replicated our findings in another domain of performance-enhancement in Study 13 provides some assurance that our results are not due to some quirk of our "Joe" scenario. But, this is far from evidence for widespread generalizability. Future research in this area could expand this investigation to a broader sample of contexts (Wells & Windschitl, 1999), varying the context of enhancement (e.g., weightlifting, basketball, cycling, university classes, competitive trivia), target of judgment, method of enhancement (e.g., steroids, human growth hormone, blood doping, stimulants), and so forth. We would expect that, across this range of stimuli, reliable effects of fairness, rules and laws, and prudence would emerge, consistent with what we found in the present "deep dive". Such an investigation might also provide insight into the conditions that moderate the effect of other considerations, such as naturalness.

This leads to another potential limitation of the current research: our reliance on self-report measures and fictitious scenarios. Because we are studying judgments, we do not see any viable dependent variable besides participants' self-reports, but this research is still subject to the usual critiques of self-report methodologies. For one, participants may have altered their responses to present themselves in a better light. However, we do not see any reason why this would interact with our manipulations. Though there may be a main effect due to self-presentational concerns in our studies, we think it is unlikely that it undermines the results we have reported. Also, participants may lack introspective access to the judgments they would make in real cases of PED use, leading to a lack of ecological validity among our hypothetical

scenarios. This problem seems potentially soluble, in that there is no shortage of real news reports about athletes using PEDs, which could be used as stimuli in future research. We opted to use fictional scenarios because it allowed for very tight experimental control, and for manipulation of many different aspects of the situation, but future research could use more naturalistic stimuli.

A Tentative Framework for Interpreting Our Results

Despite these important drawbacks, our exploratory approach allowed us to paint a more comprehensive picture of what underlies opposition to PED use than we could have by committing in advance to a particular theoretical framework and testing predictions derived from it. We now turn to sketching a framework for understanding these judgments that might also be applicable to some other normative judgments that are quite different from those explored here.

PED use appears to be a transgression that exhibits properties of what are referred to in Social Domain Theory as moral, conventional, and prudential violations. PED use seems to most closely resemble a violation of social convention, because a powerful determinant of opposition is the presence or absence of rules and laws prohibiting it. This is not to suggest that if a sport were to expressly permit PED use, that such permission would eliminate opposition to it. Most PEDs would still be illegal, except when prescribed to treat medical conditions, and concerns about risks to users (i.e., prudential concerns) would remain. In other words, rules and laws are a primary driver of opposition to PED use, but not the only one. Opposition appears to be a function of classically moral, social conventional, and prudential concerns, where these factors produce unique and interactive effects.

"Domain-crossing" violations of this sort have received relatively little discussion in SDT research, with some exceptions. For example, illicit drug use is considered by some to be a moral violation, by others a prudential violation, and by still others as being permissible, and at the discretion of the user (Nucci, Guerra, & Lee, 1991). Also, some sexual acts are classified as "nonprototypical" violations that are inconsistently ascribed properties of moral violations and personal prerogatives (Turiel, Hildebrandt, & Wainryb, 1991). As far as we are aware, however, no investigations have found that a single action that can be considered wrong for classically moral (fairness), conventional (rules) and prudential (risk) reasons, simultaneously. PED use seems not to fit neatly into any of these three categories, but rather to exhibit properties typically associated with all of them. It may not be unique in this way, but may be one exemplar of a more general class of violations that blurs the lines between the moral, conventional, and prudential domains, and has not been previously articulated.

We speculate that this class of violations is characterized by a belief that a set of rules has intrinsic value or normative force, and actions that violate the rules offend against this, and may harm the actor as well. In contexts like this, agreed-upon rules established by legitimate authorities *create* the potential for violations of fairness (i.e., the moral component of the violation)—if there were no rules prohibiting steroid use, then everyone could use steroids, and there would be no fairness violation. Yet, these rules also retain their normative import even when everyone "fairly" violates them in the same way (i.e., the social conventional component). And, if violating the rules is dangerous to the actor (i.e., the prudential component), this is even worse. This confluence of moral, conventional, and prudential components describes competitive athletic contexts, but it may also occur in other contexts.

Some of these contexts are clearly similar to steroid use by athletes, such as off-label use of prescription stimulants by students, as in Study 13. Others are more far-flung. One example might be predatory pricing by firms, which is unfair to competitors, illegal, and – at least in the short term – harms the firm engaging in it. Other examples can be seen in international relations, such as limits on pollution or development of nuclear weapons. If one country violates an international agreement by polluting or building up their arsenal, this gives them an unfair advantage over other countries (economically in the former case, strategically in the latter). Yet, if every country "fairly" violated international law to the same extent, we suspect that people would still be opposed to these violations. This opposition could be rooted, at least in part, in the idea that these violations harm the countries (i.e., their citizens) directly in the case of pollution, and indirectly in the case of arms proliferation, by diverting resources away from the people. It could also be because international law is likely seen as having legitimate normative force in and of itself. Analogous arguments can be made relating to federal, state, and local laws as well.

These disparate examples share a structural commonality that we think relates to people's normative judgments of this class of violations. They all resemble prisoner's dilemmas where an imperfectly enforceable rule mandating cooperation has been instituted by a legitimate authority. To illustrate, consider the use of steroids in a two-person athletic competition. If Player 1 uses steroids (i.e., defects) and Player 2 does not (i.e., cooperates) Player 1 incurs a small cost (the risk of side effects, i.e., the prudential component of this violation). But, he is better off overall, in that he (unfairly) gains an advantage over his competitor. If both players use steroids (i.e., defect), they both incur a cost, and neither benefits, because their relative outcomes in the competition remain fixed (assuming the steroids are equally effective for both of them, etc.).

This state of affairs is, strictly speaking, “fair”, but prudential concerns disfavor it over a state of affairs where neither competitor uses steroids.

Now suppose that an authority (the governing body of the sport, the government, etc.) recognizes that both players have reason to defect, but they are better off if they both cooperate. This authority might create a rule mandating cooperation (i.e., banning steroid use). If the authority is viewed as legitimate, this rule may take on its own normative force in the eyes of third-party observers, such that, even if both players defect, “fairly” violating the rule, their actions will be opposed if they are discovered. It should be clear enough how this logic also applies to predatory pricing, pollution treaties, nuclear arsenals, and probably other domains as well. Based on this reasoning, we predict that this structure defines the class of violations that we have identified in this research. Defection in prisoner’s dilemma-type situations with imperfectly enforceable mandates for cooperation should be considered wrong for canonically moral, conventional, and prudential reasons, simultaneously. Testing this prediction in other domains is beyond the scope of this paper, but this serves as an illustration of how an exploratory, “deep dive” method that examines one narrowly-defined phenomenon can motivate new theorizing and predictions for future research.

Conclusion

In investigating what underlies opposition to the use of performance-enhancing drugs (PEDs), we began with the hypothesis that opposition to PED use cannot be fully explained by concerns about fairness. After finding support for this hypothesis, we explored ten other potential drivers of opposition individually, and found reliable effects of laws and league rules prohibiting steroid use, and of risk to the user. We found in a follow-up study that the effects of laws and risk depend on the presence or absence of league rules, but the effect of fairness is

largely independent, and that the effects of laws and rules are considerably larger than the effects of fairness and risk. We then found similar patterns in a different domain, the use of cognitive-enhancement drugs to improve academic performance. Lastly, we sketched a framework for interpreting our results that situates our research in the literature on Social Domain Theory, posits a heretofore unstudied class of violations, and makes predictions for future research.

We consider this research to be a useful step in demystifying a widely expressed but relatively understudied normative judgment, and an illustration of how exploratory methods can be useful for generating theoretical insights. Rather than test one or two focal hypotheses, we tested a multitude of different contributing and competing explanations to understand the phenomenon being studied. We think approaches like the one used here can promote rapid progress in understanding psychological phenomena, especially when applied to understudied, "wide-open" domains like the one under scrutiny here.

References

Associated Press. (2016, June 11). Cycling officials combat “mechanical doping”. *The New York Times.* Retrieved from http://www.nytimes.com

Barra, A. (2007, March 27). An asterisk is very real, even when it’s not. *The New York Times.* Retrieved from http://www.nytimes.com

Barry Bonds, Roger Clemens not voted into Baseball Hall of Fame. (2017, January 18). *Sports Illustrated.* Retrieved from http://www.si.com/mlb/2017/01/18/baseball-hall-of-fame-barry-bonds-roger-clemens

Bartels, D. M., Bauman, C. W., Cushman, F. A., Pizarro, D. A., & McGraw, A. P. (2016), Moral Judgment and Decision Making. In G. Keren & G. Wu (Eds.) *The Wiley Blackwell handbook of judgment and decision making* (pp. 478-515). Chichester, UK: Wiley.

Bazerman, M. H., & Messick, D. M. (1998). On the power of a clear definition of rationality. *Business Ethics Quarterly, 8*, 477-480.

Brodsky, I. G., Balagopal, P., & Nair, K. S. (1996). Effects of testosterone replacement on muscle mass and muscle protein synthesis in hypogonadal men – A clinical research center study. *Journal of Clinical Endocrinology and Metabolism, 81,* 3469-3475.

Bruinius, H. (2005, March 24). Will steroids alter baseball records, too? *The Christian Science Monitor.* Retrieved from http://www.csmonitor.com

Campbell, J. I. D. & Thompson, V. A. (2012). MorePower 6.0 for ANOVA with relational confidence intervals and Bayesian analysis. *Behavior Research Methods, 44,* 1255-1265.

Case, T. I., Oaten, M. J., & Stevenson, R. J. (2012). Disgust and moral judgment. In R. Langdon & C. Mackenzie (Eds.), *Emotions, Imagination, and Moral Reasoning* (pp. 195-218). New York, NY: Psychology Press.

Chakroff, A., Dungan, J., Koster-Hale, J., Brown, A., Saxe, R., & Young, L. (2015). When minds matter for moral judgment: Intent information is neutrally encoded for harmful but not impure acts. *Social Cognitive and Affective Neuroscience, 11*, 476-484.

CNN Political Unit. (2013, September 27). Poll: "Obamacare" vs. "Affordable Care Act". *CNN*. Retrieved from http://cnn.com

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.)*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Dodge, T., Williams, K. J., Marzell, M., & Turrisi, R. (2012). Judging cheaters: Is substance misuse viewed similarly in the athletic and academic domains?

Ekman, P., & Friesen, W. V. (1975). *Unmasking the face: A guide to recognizing emotions from facial cues*. Englewood Cliffs, NJ: Prentice-Hall.

Fahey, T.D. (1998). Anabolic-androgenic steroids: Mechanism of action and effects on performance. In T. D. Fahey (Ed.), *Encyclopedia of sports medicine and science.* Retrieved from http://www.sportsci.org/encyc/anabster/anabster.html

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191.

Fitz, N. S., Nadler, R., Manogaran, P., Chong, E. W. J., & Reiner, P. B. (2014). Public attitudes toward cognitive enhancement. *Neuroethics, 7,* 173-188.

Forlini, C., Schildmann, J., Roser, P., Beranek, R., & Vollman, J. (2015). Knowledge, experiences and views of German university students toward neuroenhancement: An empirical-ethical analysis. *Neuroethics, 8,* 83-92.

Gordon, D. (Director). (2012). *9.79** [Motion picture on DVD]. United States: ESPN.

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology, 96,* 1029-1046.

Gray, K. (2014). Harm concerns predict moral judgments of suicide: Comment on Rottman, Keleman, and Young (2014). *Cognition, 133,* 329-331.

Haidt, J. (2012). *The righteous mind.* New York, NY: Vintage Books.

Hall, M. P., Sowden, W. J., & Ellsworth, P. C. (2017). When cheating is fair: Performance-enhancing behaviors and the moral handicapping effect. *Manuscript in preparation.*

Horberg, E. J., Oveis, C., Keltner, D., & Cohen, A. B. (2009). Disgust and the moralization of purity. *Journal of Personality and Social Psychology, 97,* 963-976.

Hester, N., Schein, C., Sirbu, M., & Gray, K. (2015). Selfish, unnatural, and lazy: When biomedical enhancements are wrong. *Unpublished manuscript.*

Hsee, C. K., Loewenstein, G. F., Blount, S., & Bazerman, M. H. (1999). Preference reversals between joint and separate evaluations of options: A review and theoretical analysis. *Psychological bulletin*, *125*(5), 576.

Huebner, B., Lee, J. J., & Hauser, M. D. (2010). The moral-conventional distinction in mature moral competence. *Journal of Cognition and Culture, 10,* 1-26.

Janofsky, M. (1989, March 2). Coach tells a Canadian inquiry Johnson used steroids since ’81. *The New York Times.* Retrieved from http://www.nytimes.com

Jenkins, C. (2005, March 15). Players admit steroids changed baseball. *USA Today.* Retrieved from http://usatoday.com

Johnson, D. J., Wortman, J., Cheung, F., Hein, M., Lucas, R. E., Donnellan, M. B., Ebersole, C. R., & Narr, R. K. (2016). The effects of disgust on moral judgments: Testing moderators. *Social Psychological and Personality Science, 7,* 640-647.

Kohlberg, L. (1971). From is to ought: How to commit the naturalistic fallacy and get away with it in the study of moral development. In T. Mischel (Ed.), *Cognitive development and epistemology* (pp. 151-235). New York: Academic Press.

Landy, J. F. & Goodwin, G. P. (2015). Does incidental disgust amplify moral judgment? A meta-analytic review of experimental evidence. *Perspectives on Psychological Science, 10,* 518-536.

Morris, S. B. & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-group designs. *Psychological Methods, 7,* 105-125.

Nabi, R. L. (2002). The theoretical versus the lay meaning of disgust: Implications for emotion research. *Cognition and Emotion, 16,* 695-703.

Nucci, L. (1981). Conceptions of personal issues: A domain distinct from moral or societal concepts. *Child Development, 52,* 114-121.

Nucci, L., Guerra, N., & Lee, J. (1991). Adolescent judgments of the personal, prudential, and normative aspects of drug usage. *Developmental Psychology, 27,* 841-848.

Nucci, L. P. & Nucci, M. S. (1982). Children's responses to moral and social conventional transgressions in free-play settings. *Child Development, 53,* 1337–1342.

O'Brien, S. A. (2017, February 28). Here's why the Internet isn't working today. *CNN.* Retrieved from http://cnn.com

Oskamp, S. (1965). Overconfidence in case study judgments. *Journal of Consulting Psychology, 29,* 261-265.

Paharia, N., Kassam, K. S., Greene, J. D., & Bazerman, M. H. (2009). Dirty work, clean hands: The moral psychology of indirect agency. *Organizational Behavior and Human Decision Processes, 109*(2), 134-141.

Platt, J. R. (1964). Strong inference. *Science, 146,* 347-353.

Reimer, A. (2015, August 30). NFL could soon look into regulating sticky gloves. *SBNation.* Retrieved from http://www.sbnation.com

Riis, J., Simmons, J. P., & Goodwin, G. P. (2008). Preferences for enhancement pharmaceuticals: The reluctance to enhance fundamental traits. *Journal of Consumer Research, 35,* 495-508.

Rottman, J. & Kelemen, D. (2012). Aliens behaving badly: Children's acquisition of novel purity-based morals. *Cognition, 124,* 356-360.

Rottman, J., Kelemen, D., & Young, L. (2015). Hindering harm and preserving purity: How can moral psychology save the planet?. *Philosophy Compass, 10,* 134-144.

Royzman, E., Atanasov, P., Landy, J. F., Parks, A., & Gepty, A. (2014). CAD or MAD? Anger (not disgust) as the predominant response to pathogen-free violations of the divinity code. *Emotion, 14,* 892-907.

Royzman, E. B., Kim, K., & Leeman, R. F. (2015). The curious tale of Julie and Mark: Unraveling the moral dumbfounding effect. *Judgment and Decision Making, 10,* 296-313.

Royzman, E. B., Landy, J. F., & Goodwin, G. P. (2014). Are good reasoners more incest-friendly? Trait cognitive reflection predicts selective moralization in a sample of American adults. *Judgment and Decision Making, 9,* 176-190.

Royzman, E. B., Leeman, R., & Baron, J. (2009). Unsentimental ethics: Towards a content-specific account of the moral-conventional distinction. *Cognition, 112,* 159-174.

Rozin, P. (2005). The meaning of "natural": Process more important than content. *Psychological Science, 16,* 652-658.

Rozin, P., Fischler, C., & Shields-Argelès, C. (2012). European and American perspectives on the meaning of natural. *Appetite, 59,* 448-455.

Rozin, P., Haidt, J., & McCauley, C. R. (2000). Disgust. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (2nd ed., pp. 637–653). New York, NY: Guilford Press.

Rozin, P., Lowery, L., Imada, S., & Haidt, J. (1999). The CAD Triad Hypothesis: A mapping between three moral emotions (contempt, anger, disgust) and three moral codes

(community, autonomy, divinity). *Journal of Personality and Social Psychology, 76,* 574-586.

Rozin, P., Spranca, M., Krieger, Z., Neuhaus, R., Surillo, D., Swerdlin, A., & Wood, K. (2004). Preference for natural: Instrumental and ideational/moral motivations and the contrast between foods and medicines. *Appetite, 43,* 147-154.

Sabini, J. & Monterosso, J. (2005). Judgments of the fairness of using performance enhancing drugs. *Ethics & Behavior, 15,* 81-94.

Sattler, S., Forlini, C., Racine, E., & Sauer, C. (2013). Impact of contextual factors and substance characteristics on perspectives toward cognitive enhancement. *PLoS One, 8*(8), e71542.

Scheske, C. & Schnall, S. (2012). The ethics of "smart drugs": Moral judgments about healthy people's use of cognitive-enhancing drugs. *Basic and Applied Social Psychology, 34,* 508-515.

Schnall, S., Haidt, J., Clore, G. L, & Jordan, A. H. (2008). Disgust as embodied moral judgment. *Personality and Social Psychology Bulletin, 34,* 1096-1109.

Selinger, E. (2012, August 28). "But everybody's doing it!": Lance Armstrong and the philosophy of making bad decisions. *The Atlantic.* Retrieved from http://www.theatlantic.com

Simmons, J. P., Nelson, J. D., Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science, 22,* 1359-1366.

Simmons, J., Nelson, L. D., & Simonsohn, U. (2013). Life after p-hacking. In S. Botti & Labroo, A. (Eds.), *Advances in consumer research, volume 41* (p. 775).

Smetana, J. G., Jambon, M., & Ball, C. (2014). The social domain approach to children's moral and social judgments. In M. Killen & J. G. Smetana (Eds.), *Handbook of moral development* (pp. 23–45). New York: NY: Psychology Press.

Steroid Suspensions. (2017). *Baseball Almanac.* Retrieved from http://www.baseball-almanac.com/legendary/steroids_baseball.shtml

Sunstein, C. R., (2005). Moral heuristics. *Behavioral and Brain Sciences, 28,* 531-573.

Tisak, M. & Turiel, E. (1984). Children's conceptions of moral and prudential rules. *Child Development, 55,* 1030-1039.

Turiel, E. (1983). *The development of social knowledge: Morality and convention.* Cambridge, UK: Cambridge University Press.

Turiel, E. (2002). *The culture of morality.* Cambridge, UK: Cambridge University Press.

Turiel, E., Hildebrandt, C., & Wainryb, C. (1991). Judging social issues: Difficulties, inconsistencies, and consistencies. *Monographs of the Society for Research in Child Development, 56,* i-116.

U. S. Anti-Doping Agency (2015). *2015 USADA annual report.* Retrieved from http://www.usada.org/wp-content/uploads/2015_annual_report.pdf

Wells, G. L. & Windschitl, P. D. (1999). Stimulus sampling and social psychological experimentation. *Personality and Social Psychology Bulletin, 25,* 1115-1125.

Wheatley, T. & Haidt, J. (2005). Hypnotic disgust makes moral judgments more severe. *Psychological Science, 16,* 780-784.

Williams, E. F. & Steffel, M. (2014). Double standards in the use of enhancing products by self and others. *Journal of Consumer Research, 41,* 506-525.