

1-1-2019

Problems of Proof for the Ban on Female Athletes with Endogenously High Testosterone Levels

Silver Lin

The University of Chicago Law School

Follow this and additional works at: <https://chicagounbound.uchicago.edu/cjil>



Part of the [Law Commons](#)

Recommended Citation

Lin, Silver (2019) "Problems of Proof for the Ban on Female Athletes with Endogenously High Testosterone Levels," *Chicago Journal of International Law*. Vol. 20: No. 1, Article 6.

Available at: <https://chicagounbound.uchicago.edu/cjil/vol20/iss1/6>

This Article is brought to you for free and open access by Chicago Unbound. It has been accepted for inclusion in Chicago Journal of International Law by an authorized editor of Chicago Unbound. For more information, please contact unbound@law.uchicago.edu.

Problems of Proof for the Ban on Female Athletes with Endogenously High Testosterone Levels

Cover Page Footnote

The University of Chicago Law School J.D. Candidate, 2020. I would like to thank the CJIL Board and staff for all of their guidance and support in drafting this Comment. I am especially grateful to Professor Tom Ginsburg for his feedback, ideas, and overall contributions to this piece.

Problems of Proof for the Ban on Female Athletes with Endogenously High Testosterone Levels

Silver Lin*

Abstract

At the time of this writing, a new International Association of Athletics Federations regulation preventing women with naturally high testosterone from competing in certain international athletics events has reignited the controversy over the male-female distinction in sports and its implications on individuals' right to compete. A recent case filed by runner Caster Semenya and Athletics South Africa challenging this regulation before the Court of Arbitration for Sport, an arbitral tribunal that adjudicates disputes in international sports, sought to have the regulation overturned as discriminatory against women with a genetic intersex condition. Drawing on established international arbitration law, international norms in arbitrations, and relevant precedent, this Comment explores the evidentiary issues before the Court of Arbitration for Sport in Semenya's challenge. In particular, this Comment argues that, given the high stakes of the case as well as the inequity in resources between the parties, the Court of Arbitration for Sport should have adopted unconventional rules with respect to the allocation of the burden of proof, the requisite standard of proof, and the evaluation of scientific evidence to ensure a fair hearing on the matter. The Comment ultimately concludes that the suggested changes are well within the discretion and ability of the Court of Arbitration for Sport to implement, slight challenges to the adoption of each proposed measure notwithstanding.

Table of Contents

I. Introduction.....	219
II. Semenya, the IAAF, and CAS.....	222
A. The Principal Parties	222

* The University of Chicago Law School J.D. Candidate, 2020. I would like to thank the CJIL Board and staff for all of their guidance and support in drafting this Comment. I am especially grateful to Professor Tom Ginsburg for his feedback, ideas, and overall contributions to this piece.

B. Relevant CAS Precedents	224
III. Current International Arbitration Laws and Norms	226
A. Absence of Binding Arbitral Laws	226
B. International Arbitration Norms on the Burden of Proof	228
C. International Arbitration Norms on the Standard of Proof.....	232
D. International Arbitration Norms on the Evaluation of Evidence	235
IV. CAS Proceedings Then and Now: An Argument for Changes to the Evidentiary Rules.....	237
A. CAS Should Have Shifted the Burden of Proof to Regulating Bodies.....	238
B. CAS Should Have Applied a Standard of Proof of Comfortable Satisfaction, Not Balance of Probabilities, to the Claims Brought by the IAAF	242
C. CAS Should Have Adopted the Daubert Standard for Considering Scientific Evidence.....	245
1. The <i>Chand</i> decision did not resolve how CAS panels should evaluate evidence.	245
2. The <i>Semenya</i> panel should have considered adopting the <i>Daubert</i> standard for evaluating scientific evidence.	246
V. Conclusion.....	250

I. INTRODUCTION

Since the late 1960s, international sports organizations have conducted gender testing to police the divide between men's and women's competitions.¹ Due to improvements in science and technology, the nature of such testing has evolved over the years.² As international sports organizations have developed a greater recognition of the lack of a binary gender classification, regulating bodies have likewise walked back from mandatory gender testing.³ Yet this reduction in gender testing cannot be seen as a complete victory for those opposed to such testing, for the mandatory tests have only been replaced by testing on a case-by-case basis in some sporting bodies, most notably in the International Olympic Committee (IOC) and in the International Association of Athletics Federations (IAAF).⁴ In particular, new understandings of hyperandrogenism, "a congenital disorder [that] leads to elevated levels of testosterone due to an androgen insensitivity,"⁵ has generally led to a greater acceptance of hyperandrogenic women,⁶ or women with biological features more commonly associated with masculinity.⁷ Given this new understanding of intersex conditions, the argument for eliminating gender testing based on problematic metrics like testosterone has gained much traction in the sporting world and has furthered the controversy surrounding the current system of case-by-case testing.⁸

Despite advances in biology that cast doubt upon the traditional gender divide, the IAAF announced on April 26, 2018 new regulations limiting participation of female athletes whose testosterone levels are outside the normal female range (as defined by the IAAF's medical and science experts) in certain track events, such as the 400-meter and 1600-meter races.⁹ Setting the threshold

¹ See Maayan Sudai, *The Testosterone Rule—Constructing Fairness in Professional Sport*, 4 J. L. & BIOSCIENCES 181, 182 (2017).

² Erin Elizabeth Berry, *Respect for the Fundamental Notion of Fairness of Competition: the IAAF, Hyperandrogenism, and Women Athletes*, 27 WIS. J. L. GEND. & SOC'Y 207, 212–13 (2012).

³ Dutee Chand v. Athletics Federation of India (AFI) & The International Association of Athletics Foundation (IAAF), CAS 2014/A/3759, ¶¶ 294–95 (July 24, 2015).

⁴ Berry, *supra* note 2, at 208.

⁵ Anna Boyd, *Back to the Binary: How the Olympics Struggle with Separation of Male and Female*, 14 DEPAUL J. SPORTS L. & CONTEMP. PROBS. 1, 13 (2018).

⁶ See generally Erin Buzuvis, *Hormone Check: Critique of Olympic Rules on Sex and Gender*, 31 WIS. J. L. GEND. & SOC'Y 29 (2016) (advocating hybrid approaches of uniform gender identity and hormone rules in categorizing competitors and suggesting openness to redefining gender norms in international sports).

⁷ See *id.* at 53 (defining hyperandrogenism).

⁸ See Berry, *supra* note 2, at 209.

⁹ See *IAAF Introduces New Eligibility Regulations for Female Classification*, IAAF (Apr. 26, 2018), <http://perma.cc/3CRF-29J4>.

for participation in competition at 5 nanomoles per liter (nmol/L) of testosterone, the regulation bans women with 46 XY differences in sexual development (DSDs) and thus endogenously, or naturally-occurring, high testosterone levels from competing in the named events unless they undergo treatment to lower their testosterone levels or choose to participate in men's competitions.¹⁰

The new regulations come after a similar rule that restricted competition participation for women with over 10 nmol/L of testosterone was temporarily invalidated by the Court of Arbitration for Sport (CAS) in July 2015 in *Dutee Chand v. Athletics Federation of India (AFI) & IAAF*.¹¹ CAS temporarily enjoined enforcement of the competition ban on the grounds that the IAAF, *inter alia*, did not provide enough evidence that an endogenous testosterone level differential resulted in unfair competitive outcomes.¹² Because the new regulations supplant the policies at issue in the 2015 case, that case is now rendered moot. With respect to the latest regulation on women with high testosterone, the IAAF claims to have strong evidence supporting its conclusion that high DSDs such as endogenously elevated testosterone levels lead to differences in athletic performance, arguably providing the necessary justification for the new regulation.¹³ Although the new regulations were scheduled to go into effect on November 1, 2018, South African runner Caster Semenya, who has been in the spotlight as a female athlete with naturally-elevated testosterone levels, filed a challenge to the regulation before CAS in June 2018.¹⁴ In light of Semenya's challenge, the IAAF agreed to suspend implementation of the regulation until CAS resolves the matter.¹⁵

From February 18–22, 2019, panelists the Honorable Dr. Annabelle Bennett, the Honorable Hugh L. Fraser and Dr. Hans Nater heard Semenya's challenge to the DSD regulations and the IAAF's response.¹⁶ On May 1, 2019, CAS released the decision in Semenya's challenge, dismissing the requests for arbitration because Semenya and Athletics South Africa (ASA) were unable to prove the invalidity of

¹⁰ See *id.*; see also CAS, *CAS Arbitration: Caster Semenya, Athletics South Africa (ASA) and International Association of Athletics Federations (IAAF): Decision*, CAS MEDIA RELEASE (May 1, 2018), <http://perma.cc/B6X5-LAAC> [hereinafter 2019 CAS Media Release].

¹¹ See *Chand*, *supra* note 3, at 160.

¹² *Chand*, *supra* note 3, at ¶¶ 547–48.

¹³ See IAAF, *supra* note 9; see also Audie Cornish & Katrina Karakazis, *IAAF Creates Rule to Ban Women with Naturally High Testosterone Levels from Competition*, NPR (Apr. 30, 2018), <http://perma.cc/E57L-Y3MH>.

¹⁴ See Dan Roe, *Caster Semenya's Pro Eligibility is Still in Question. Here's Why That's Total Bullshit*, RUNNER'S WORLD (Oct. 17, 2018), <http://perma.cc/KRW8-SJBM>; see also Nick Zaccardi, *Caster Semenya on New IAAF Rule: 'Discriminatory, Irrational, Unjustifiable'*, NBCSPORTS (June 18, 2018), <http://perma.cc/7D9N-6MXY>.

¹⁵ See Roe, *supra* note 14.

¹⁶ 2019 CAS Media Release, *supra* note 10.

the new DSD regulations.¹⁷ Referencing the as-yet unpublicized panel decision, the release simply stated that “[t]he Panel found that the DSD Regulations are discriminatory but the majority of the Panel found that . . . such discrimination is a necessary, reasonable and proportionate means of achieving” fair competition.¹⁸ Although the media release announced the panel’s ultimate conclusion and the panel’s own reservations about such an outcome, the statement failed to explain the process by which the panel arrived at its decision.¹⁹ As of June 3, 2019, however, the Federal Supreme Court of Switzerland temporarily suspended the application of the DSD Regulations to Semenya pending further hearing.²⁰

As a result of the controversy surrounding the IAAF’s new ban and its adverse effect on certain female athletes, Semenya’s challenge has been closely scrutinized and compared to the 2015 *Chand* case, in part due to the potential ramifications of the outcome of this challenge.²¹ One resounding allegation against the ban is that it discriminates unjustly against a small group of elite female athletes on the basis of poor scientific evidence.²² While scholars have commented on the discriminatory nature of the new regulation and debated the merits of the underlying science,²³ procedural issues have largely been overlooked. This Comment addresses the problems Semenya’s challenge faced in an arbitral hearing before CAS given the framework laid out in *Chand*. By exploring the generally accepted rules of evidence in international arbitration, this Comment argues that CAS procedures should have been updated in Semenya’s case—and should be updated in future cases—to ensure athletes receive a fair hearing on the merits of their claims. The Comment focuses on CAS proceedings not simply because of the possibility of a remand of Semenya’s case to CAS, but also because of CAS’s history of dealing with such procedural issues and its scope for improvement in future cases.

Section II describes the principal parties involved in this most recent challenge and CAS’s role. This Section also touches on CAS procedures in *Chand* and how similar procedural decisions may have affected Semenya’s challenge. Section III discusses the current laws governing the taking of evidence in international arbitration, and the norms that operate in the absence of codified evidentiary laws. Finally, Section IV advances a three-fold argument for a novel

¹⁷ *Id.*

¹⁸ *Id.*

¹⁹ *Id.*

²⁰ See Yomi Kazeem, *Caster Semenya can compete again without testosterone-reducing drugs—for now*, QUARTZ AFRICA (June 4, 2019), <http://perma.cc/SP5W-XC8M>.

²¹ See Roe, *supra* note 14.

²² See Jeré Longman, *Track’s New Gender Rules Could Exclude Some Female Athletes*, N.Y. TIMES (Apr. 25, 2018), <http://perma.cc/A769-XBD9>.

²³ See Roe, *supra* note 14.

CAS procedural approach to high-stakes cases such as those cases effectively preventing athletes from international competition. Grounded in examples from other international decisions and an evidentiary framework borrowed from United States (U.S.) jurisprudence, this Comment argues normatively for the following changes to the current arbitration procedures: (1) that the burden of proof for the necessity of the ban should be on the regulating entity; (2) that the standard for this burden of proof should be to the CAS panel's comfortable satisfaction; and (3) that the CAS panel should use a multifactor framework for assessing the quality of the scientific studies and other expert evidence presented by the parties.

II. SEMENYA, THE IAAF, AND CAS

A. The Principal Parties

Caster Semenya's story highlights the ongoing nature of the controversy surrounding the competition eligibility of athletes with atypical biology. Semenya burst onto the international track scene in 2008 with her first gold medal win in the 800 meter at the Commonwealth Youth Games.²⁴ She followed up her success at the junior level by winning the 800 meter and 1500 meter races at the 2009 African Junior Championships.²⁵ It was around the time of this initial success that word of Semenya being forced to undergo sex testing first leaked.²⁶ The rumored result of the sex testing was that Semenya had "the external genitalia of a female, but internal testes, instead of ovaries and a uterus,"²⁷ a result that was meant to be confidential and which the IAAF did not confirm.²⁸ Although Semenya was eventually allowed to retain her medals and cleared to continue competing, the controversy surrounding her sex did not end.²⁹ Despite, or perhaps because of, Semenya's tremendous success internationally, including her multiple gold medals at the 2016 Rio Olympics, Semenya's eligibility to race has continued to be a topic of conversation throughout her entire career.³⁰

²⁴ Shawn M. Crincoli, *You Can Only Race if You Can't Win? The Curious Cases of Oscar Pistorius & Caster Semenya*, 12 *TEX. REV. ENT. & SPORTS L.* 133, 154 (2011).

²⁵ *Id.*

²⁶ *Id.* at 137, n. 27.

²⁷ Anna Peterson, *But She Doesn't Run Like A Girl...: The Ethic of Fair Play and the Flexibility of the Binary Conception of Sex*, 19 *TUL. J. INT'L & COMP. L.* 315, 316 (2010).

²⁸ *Id.*

²⁹ *Id.*

³⁰ See Boyd, *supra* note 5, at 6; Crincoli, *supra* note 24, at 137–38. See also Peterson, *supra* note 27, at 316–17.

Partially responsible for this controversy is the IAAF, the governing body for international track competitions since 1912.³¹ One of the world's largest sporting organizations,³² the IAAF was created to act as the central authority on regulating athletics internationally and to serve as the official global recordkeeper.³³ In keeping with its original purpose, the IAAF enacts policies and regulations affecting athletes' participation in international track and field competitions. Due to the IAAF's role as "the primary regulator of athletics from the sub-national level all the way up to the Olympic level,"³⁴ athletes who wish to compete internationally arguably have no other recourse but to agree to be bound by its rules, including the DSD Regulations now in dispute.³⁵

When controversies between athletes and the IAAF inevitably arise, CAS serves as the forum for such dispute resolution.³⁶ The IAAF Competition Rules state that

[i]n all CAS appeals involving the IAAF, CAS and the CAS Panel shall be bound by the IAAF Constitution, Rules and Regulations. In the case of any conflict between CAS rules currently in force and the IAAF Constitution, Rules and Regulations, the IAAF Constitution, Rules and Regulations shall take precedence.³⁷

A private governing body, CAS was "designed specifically to adjudicate sports-related disputes and is essentially sport's ultimate umpire," often handing down the final word in such disputes.³⁸ CAS was intended to be "an institution capable of achieving the quick, efficient, inexpensive and binding resolution of sporting disputes."³⁹ In this capacity, CAS now "plays a key role within the global governance regime for athletics by holding the IAAF accountable in its regulatory activities."⁴⁰

Because CAS is a private tribunal, it creates and follows its own procedural rules⁴¹ rather than formally adopting all of the rules of a particular jurisdiction.

³¹ Michele Krech, *To Be a Woman in the World of Sport: Global Regulation of the Gender Binary in Elite Athletics*, 35 BERKELEY J. INT'L L. 262, 278 (2017).

³² *Id.* at 279.

³³ *Id.* at 278.

³⁴ *Id.*

³⁵ See Longman, *supra* note 22 ("Female athletes with elevated testosterone levels will essentially face a 'choice of no choice[.]'").

³⁶ Rachele Downie, *Improving the Performance of Sport's Ultimate Umpire: Reforming the Governance of the Court of Arbitration for Sport*, 12 MELB. J. INT'L L. 315, 317 (2011).

³⁷ Chand, *supra* note 3, at ¶ 438.

³⁸ Downie, *supra* note 36, at 316.

³⁹ *Id.* at 317.

⁴⁰ Krech, *supra* note 31, at 291.

⁴¹ Downie, *supra* note 36, at 318.

The fact that CAS sets its own procedural rules is significant because CAS panels often make their own evidentiary rules on a case-by-case basis and are not bound by other international laws governing evidentiary issues.⁴² Currently, CAS panels need only follow the CAS Code of Sports-related Arbitration (Code), and where the Code is silent on a relevant evidentiary issue, as it often is, the appointed CAS panel resolves the issue.⁴³ In reality, then, as “the supreme forum for international sport,”⁴⁴ a particular CAS panel can adopt evidentiary rules that in turn have wide-reaching effects on athletes by handing down decisions that substantively affect athletes’ rights to compete with little oversight or mechanism for review.

B. Relevant CAS Precedents

Since its inception, CAS has heard cases ranging from Fédération Internationale de Football Association (FIFA) disciplinary hearings⁴⁵ to doping violations challenges.⁴⁶ One of the most relevant precedents for the purposes of this Comment is the aforementioned case involving Dutee Chand, an Indian sprinter who underwent sex testing in 2014 after “several female athletes attending a training camp with Chand apparently expressed concern to the AFI President about her ‘masculine’ physique.”⁴⁷ After a round of testing, “the [Sports Authority of India] notified Chand that she would be excluded from the upcoming World Junior Championships and would not be eligible for selection to the Commonwealth Games because her ‘male hormone’ levels were too high.”⁴⁸ Instead of forcibly lowering her testosterone through treatment, Chand decided to challenge her participation ban before CAS, alleging that the Hyperandrogenism Regulations impermissibly discriminated against women with naturally high testosterone levels.⁴⁹

⁴² Chris Davies, *The ‘Comfortable Satisfaction’ Standard of Proof: Applied by the Court of Arbitration for Sport in Drug-Related Cases*, 14 U. NOTRE DAME AUSTL. L. REV. 1, 1 (2012) (CAS “is a non-judicial, international body which has made it clear that it does not have to follow the rules of evidence.”).

⁴³ Anthony Rigozzi & Brianna Quinn, *Evidentiary Issues Before CAS*, in INTERNATIONAL SPORTS LAW AND JURISPRUDENCE OF THE CAS – 4TH CONFERENCE CAS & SAV/FSA LAUSANNE 2012 1, 3 (Michele Bernasconi ed., 2014).

⁴⁴ Downie, *supra* note 36, at 317.

⁴⁵ Matthew J. Mitten, *The Court of Arbitration for Sport and Its Global Jurisprudence: International Legal Pluralism in a World Without National Boundaries*, 30 OHIO ST. J. ON DISP. RESOL. 1, 17 n. 84 (2014).

⁴⁶ *Id.* at 21.

⁴⁷ Krech, *supra* note 31, at 272.

⁴⁸ *Id.* at 273.

⁴⁹ *Id.*

The CAS panel reached its decision in Chand's case after three days of testimony from sixteen witnesses.⁵⁰ Throughout the hearing, both parties presented evidence in the form of expert testimony in support of their positions.⁵¹ Chand argued that the IAAF's Hyperandrogenism Regulations were unduly discriminatory on the basis of sex,⁵² while the IAAF took the position that such a regulation "was justified as a necessary, reasonable, and proportionate means of creating a level playing field for female athletes as whole."⁵³ The panel proceeded by first asking each party to lay out their requests regarding the burden and standard of proof before holding that the initial burden of proof would be on Chand, and the standard of proof would simply be on the balance of probabilities.⁵⁴ Then, each party presented evidence, including multiple experts' testimonies, regarding the acceptability of the science behind the Hyperandrogenism Regulations.⁵⁵ After ruling on the procedural elements of the arbitration, conducting the taking of evidence, and weighing the substantive evidence, the CAS panel determined that "the IAAF did not meet its burden of establishing that [the alleged] competitive advantage was of sufficient degree to warrant the exclusion of women with testosterone levels higher than 10 nmol/L,"⁵⁶ which experts for the IAAF in *Chand* asserted falls between .1 and 3.08 nmol/L for female athletes generally.⁵⁷ However, while the CAS panel did enjoin the IAAF from immediately implementing its regulations, it did so provisionally, allowing the case to be reheard if the IAAF presented more persuasive scientific evidence for the basis of the Regulations within two years of the decision.⁵⁸ Although CAS's holding in *Chand* has been mooted by the new regulations, this 2015 ruling matters not only because of its impact on Chand's case, but also because it sets persuasive precedent for other CAS cases, most recently, Semenya's latest challenge before the tribunal.

⁵⁰ *Id.*

⁵¹ *Id.* at 282–83.

⁵² *Id.* at 273.

⁵³ *Id.* at 274 (citing *Chand*, *supra* note 3, at ¶¶ 35(f), 230, 500).

⁵⁴ *Chand*, *supra* note 3, at ¶¶ 442–447. The balance of probabilities standard has been thought of as an event having at least a 51% chance of having occurred. *See Rigozzi & Quinn*, *supra* note 43, at 32.

⁵⁵ *Id.* at ¶¶ 451–453.

⁵⁶ Buzuvis, *supra* note 7, at 40.

⁵⁷ *Chand*, *supra* note 3, at ¶ 189.

⁵⁸ Buzuvis, *supra* note 7, at 40.

III. CURRENT INTERNATIONAL ARBITRATION LAWS AND NORMS

The outcome of *Chand* raises questions regarding the procedural decisions made by the CAS panel. Specifically, the apparent lack of justification for the panel's conclusions on the proper burden, standard, admissibility, and evaluation of proof call into question the panel's decision-making process. This section explores the relevant law that governs international arbitral tribunals generally and as applied to CAS, from codifications such as the International Bar Association's (IBA) Rules on the Taking of Evidence or the Swiss Rules of International Arbitration, to non-codified norms developed through precedential decisions in international adjudications generally. This section concludes that the current state of the law gives individual arbitral panels a high level of discretion in developing procedural rules for admitting and considering evidence.

A. Absence of Binding Arbitral Laws

Perhaps one of the most notable characteristics of CAS is its lack of guidance from substantive law. Currently, CAS is governed by Chapter 12 of the Swiss Private International Law Act (PILA), which includes rules for international arbitrations in Switzerland,⁵⁹ such as the rules for arbitrator selection, jurisdiction, arbitral choice of law, and the finality of awards.⁶⁰ However, the Swiss rules as codified provide little procedural guidance to CAS panels due to their brevity on evidentiary issues. For instance, the Swiss Rules of International Arbitration only include three clauses on evidence, none of which considers the quality of evidence, and one of which simply states: "The arbitral tribunal shall determine the admissibility, relevance, materiality, and weight of the evidence."⁶¹ Similarly, Chapter 12 of the PILA provides scant guidance—indeed, Article 184, Section 3: Taking of Evidence only states that "[t]he arbitral tribunal shall itself conduct the taking of evidence" and that "the arbitral tribunal may request the assistance of the state judge at the seat of the arbitral tribunal."⁶²

The lack of firm guidance on evidentiary procedure appears to pervade international arbitration in most, if not all, substantive areas,⁶³ and international

⁵⁹ Rigozzi & Quinn, *supra* note 43, at 2.

⁶⁰ *Chapter 12: International Arbitration*, Swiss Private International Law Act [hereinafter Private International Law Act] (Switz.).

⁶¹ Swiss Chambers' Arbitration Institution, *Swiss Rules of International Arbitration*, art. 24 (2012).

⁶² Private International Law Act, *supra* note 60, at art. 184 § 3.

⁶³ Abhinav Bhushan, *Standard and Burden of Proof in International Commercial Arbitration: Is There a Bright Line Rule?*, 25 AM. REV. INT'L ARB. 601, 602 (2014) ("Most major international arbitration institutions do not provide for specific evidentiary standards in their rules but do provide for

sports arbitrations are no exception. This deference to arbitral tribunal discretion on a case-by-case basis is exemplified by the IBA Rules on Evidence, which “are a relevant, if not the pre-eminent, body of evidentiary rules in modern arbitral practice, even if they are often considered ‘non-binding’ in many instances.”⁶⁴ As the primary standard-setter for evidentiary rules in international disputes, and given the lack of procedural guidance and the choice of law provisions in the Code,⁶⁵ the IBA Rules act as a default for international arbitrations.⁶⁶ Ideally, to provide greater guidance to international dispute resolution fora, the IBA Rules on Evidence would outline a set of rules or standards for international arbitral tribunals to consider when various evidentiary questions arise, such as how to allocate the burden of proof between parties, or “which party has to prove what, in order for its case to prevail,”⁶⁷ and whether to admit or bar certain types of evidence from hearings. However, the IBA Rules do not actually provide much guidance for arbitral tribunals on the “burdens or standards of proof and other issues such as presumptions or inferences in evidence.”⁶⁸ For instance, Article 9: Disclosure and Admissibility of Evidence of the IBA Rules on Evidence simply provides that

[t]he Arbitral Tribunal shall, at the request of a Party or on its own motion, exclude from evidence or production any Document, statement, oral testimony or inspection for . . . considerations of procedural economy, proportionality, fairness or equality of the Parties that the Arbitral Tribunal determines to be compelling.⁶⁹

Article 9.3 then goes on to emphasize the arbitral tribunal’s discretion in taking fairness and equality into account when making evidentiary rulings,⁷⁰ reaffirming the vague standard of fairness to the parties in international arbitration without providing more concrete guidance. As such, the IBA Rules on Evidence provide insufficient direction to CAS panels by failing to codify specific rules or even overarching standards on many facets of evidence consideration.

arbitrators’ discretion.”); *see also* Jennifer Vanderhart, *Gatekeeping: U.S. Litigation vs. Arbitration*, 12 EXPERT WITNESSES 3, 3 (2016).

⁶⁴ NATHAN D. O’MALLEY, RULES OF EVIDENCE IN INTERNATIONAL ARBITRATION: AN ANNOTATED GUIDE 10 (2012).

⁶⁵ Court of Arbitration for Sport, *Code of Sports-related Arbitration* ¶ R44.3 (2019).

⁶⁶ O’MALLEY, *supra* note 64, at 10.

⁶⁷ *Romp petrol Group N.V. v Romania*, ICSID Case No. ARB/06/3, ¶ 178 (May 6, 2013).

⁶⁸ Kabir A. N. Duggal, *Evidentiary Principles in Investor-State Arbitration*, 28 AM. REV. INT’L ARB. 3, 5 (2017).

⁶⁹ PETER ASHFORD, THE IBA RULES ON THE TAKING OF EVIDENCE IN INTERNATIONAL ARBITRATION: A GUIDE 141 (2013).

⁷⁰ *Id.* at 141–42.

Internationally adopted guidelines on the taking of evidence outside of the pure international arbitration context are likewise silent on how to allocate the burden of proof between parties and the standards of proof to be used in various circumstances.⁷¹ For instance, “that arbitrators have discretion to determine the evidentiary weight of evidence is **generally accepted** and expressly codified . . . in Art. 27(4) of the [United Nations Commission on International Trade Law (UNCITRAL)] Rules and Art. 24(2) of the Swiss Rules.”⁷² Further emphasizing arbitral tribunal discretion,

[t]he International Court of Justice has stated that ‘[t]he appraisal of the probative value of documents and evidence appertained to the discretionary power of the arbitrator is not open to question.’ This discretion provides arbitrators with the ability to alter the applicable procedure according to the requirements of the subject matter of the dispute.⁷³

In short, the lack of codified guidance combined with the presumptive deference to arbitral discretion can render international arbitration a highly unpredictable dispute resolution forum for parties. Absent more formal rules, norms developed over the course of different international arbitrations may supply panels with further guidance.

B. International Arbitration Norms on the Burden of Proof

While one may expect that the lack of a strong legal framework in the IBA Rules on Evidence would create many challenges for independent arbitral tribunals such as CAS, the loose framework does provide some benefits. The greatest advantage of having so much flexibility is that parties can then operate within this framework to argue for a different burden or standard of proof. In fact, the lack of a strong framework has led to the establishment of a number of norms regarding the burden of proof in international arbitration. Examples from various international dispute resolution fora, including CAS itself, highlight how values such as party consensus and fairness generated norms regarding burden of proof.

One such norm is that evidentiary issues are often decided by the parties themselves as long as norms of procedural fairness are adhered to, in keeping with the idea of party autonomy.⁷⁴ Because parties have so much say in determining

⁷¹ Bhushan, *supra* note 63, at 604 (“There are few institutional rules that address the issue of burden of proof.”)

⁷² TOBIAS ZUBERBÜHLER ET AL., IBA RULES OF EVIDENCE COMMENTARY ON THE IBA RULES ON THE TAKING OF EVIDENCE IN INTERNATIONAL ARBITRATION 167 (2012) (emphasis in original).

⁷³ Bhushan, *supra* note 63, at 602 (citing Case Concerning the Arbitral Award Made by the King of Spain on Dec. 23, 1906 (Hond. v. Nicar.), Judgment, 1960 I.C.J. Rep. 192, 215–216 (Nov. 18)).

⁷⁴ Bhushan, *supra* note 63, at 601. See also Francisco Blavi & Gonzalo Vial, *The Burden of Proof in International Commercial Arbitration: Are We Allowed to Adjust the Scales*, 39 HASTINGS INT’L & COMP.

how an arbitral panel will consider evidence, parties wishing to establish a particular evidentiary rule can argue “that a specific procedural agreement entered into by the parties should prevail over the general procedural agreement . . . Any such specific agreement should always prevail, unless it is not consistent with what one could call the ‘mandatory procedural rules’ contained in the Code,”⁷⁵ which would rarely be the case, given the Code’s silence on the burdens and standards of proof.

Another very important international arbitration norm that has developed is the generally accepted principle that “the burden of proving the existence of an alleged fact shall rest on the person who derives rights from that fact.”⁷⁶ This principle originates from the phrase *onus probandi actori incumbit*, which roughly translates to “the burden of proof lies on the plaintiff,” where plaintiff does not “mean the plaintiff from the procedural standpoint, but the real claimant in view of the issues involved.”⁷⁷ In other words, “each party has to prove the facts on which it relies to support its case,”⁷⁸ and that may not necessarily be the party bringing the case. There are a number of policy justifications for such an allocation of the burden of proof, such as the fact that it is often “difficult to distinguish between parties as claimant and respondent in international procedure,” and “[i]t is the duty of the parties to co-operate [sic] with international tribunals so as to establish the truth of a case.”⁷⁹

International tribunals tend to adopt rules that track the principle of having each party prove the facts upon which they rely. For instance, Article 19 of the American Arbitration Association (AAA) International Rules requires each party to prove the facts that party relies upon. Article 24(1) of the UNCITRAL requires the same.⁸⁰ Similarly, “tribunals in [International Chamber of Commerce (ICC)]

L. REV. 41, 54 (2016) (“[W]hat seems to be undisputed is that the rules governing the arbitration depend primarily on the parties’ agreement, and therefore, special deference shall be given to their decisions regarding different procedural and substantive issues, which includes the burden of proof.”).

⁷⁵ Rigozzi & Quinn, *supra* note 43, at 3.

⁷⁶ *Id.* at 15, quoting Schweizerisches Zivilgesetzbuch [ZGB], Code Civil [CC], Codice Civile [CC] [Civil Code] Dec. 10, 1907, SR 210, art. 8 (Switz.).

⁷⁷ O’Malley, *supra* note 64, at 197–98 (citing Asian Agricultural Products Ltd. v. The Republic of Sri Lanka, Final Award, ICSID Case No. ARB/87/3, 27 June 1990 in Y.B. COMMERCIAL ARB., vol. XVII 121 (Albert Jan van den Berg ed., 1992)).

⁷⁸ Blavi & Vial, *supra* note 74, at 44 (describing the general rule in international commercial arbitration, which can be generalized to other dispute resolution fora).

⁷⁹ Duggal, *supra* note 68, at 38.

⁸⁰ Bhushan, *supra* note 63, at 604.

arbitration cases also routinely hold that each party must prove the facts on which it relies in support of its claims and defenses[.]”⁸¹

International judgments, both arbitral and non-arbitral, have accordingly assigned the burden of proof to the claimant asserting a fact. In *U.S.—Measures Affecting Imports of Woven Wool Shirts and Blouses from India*,⁸² the World Trade Organization (WTO) Appellate Body stated “that the burden of proof rests upon the party, whether complaining or defending, who asserts the affirmative of a particular claim or defence.”⁸³ In that case, the Appellate Body upheld the Panel’s decision to allocate to India the burden of proof regarding its claims, and to require the U.S. to rebut the presumption if India met its initial burden.⁸⁴ In a case before the International Court of Justice (ICJ), the court had to decide whether the Temple of Preah Vihear was situated in Cambodia or Thailand, and therefore which sovereign had the right to occupy the area.⁸⁵ Part of the ICJ’s decision hinged on the burden of proof for each of the parties, on which the court noted, “[b]oth Cambodia and Thailand base their respective claims on a series of facts and contentions which are asserted or put forward by one Party or the other. The burden of proof in respect of these will of course lie on the Party asserting them or putting them forward.”⁸⁶ All of this is to say that, even in different contexts in international adjudication, tribunals have adhered to the maxim that each party bears the burden of proving the facts it asserts.

Combining this shared burden of proof with the wide discretion given to arbitrators, an international arbitral panel “may, as influenced by the logical sequence of facts involved in a claim or as imposed by the substantive law or other circumstances of the case, allocate to one side or the other the risk of not producing the evidence in support of their case.”⁸⁷ Arbitral tribunals have thus been known to re-allocate the burden of proof to the responding party when a claimant has established *prima facie* the truth of its claims.⁸⁸

⁸¹ Andrea J. Menaker & Brody K. Greenwald, *Proving Corruption in International Arbitration Who Has the Burden and How Can it be Met?*, in ADDRESSING ISSUES OF CORRUPTION IN COMMERCIAL AND INVESTMENT ARBITRATION 77-102, 80 (Domitille Baizeau & Richard Kreindler eds., 2015).

⁸² Appellate Body Report, *U.S.—Measures Affecting Imports of Woven Wool Shirts and Blouses from India*, 14, WTO Doc. WT/DS33/AB/R (adopted Apr. 25, 1997) [hereinafter *Measures Affecting Imports of Woven Wool Shirts*].

⁸³ Duggal, *supra* note 68, at 39 (quoting *Measures Affecting Imports of Woven Wool Shirts*).

⁸⁴ See *Measures Affecting Imports of Woven Wool Shirts*, *supra* note 82, at 12–14.

⁸⁵ See Case Concerning the Temple of Preah Vihear (Cambodia v. Thailand) Merits, Judgment, 1962 I.C.J. Rep. 6, 15–16 (June 15) [hereinafter *Case Concerning the Temple of Preah Vihear*].

⁸⁶ *Id.* at 16.

⁸⁷ O’MALLEY, *supra* note 64, at 205.

⁸⁸ Angeliki K. Rousaki, *Corruption in International Commercial Arbitration*, 20 VINDOBONA J. OF INT’L COM. LAW & ARB. 29, 34–35 (2016) (“It has been asserted that in cases where *prima facie* evidence

Notably, the CAS panel in *Chand* adopted a variant of the shifting standard of proof:

During the course of the hearing, the Panel requested the parties to set out their position concerning the burden and standard of proof. The parties agreed that the Athlete bore the burden of proving that the Hyperandrogenism Regulations are invalid. The parties also agreed that, once a *prima facie* case of discrimination is established, the burden shifts to the party responsible for the discriminating measure to justify the discriminatory effect.⁸⁹

In sum, while the 2015 *Chand* panel appeared to shift the burden of proof by adopting the position that Chand need only establish, *prima facie*, that there was a discriminatory effect, it still found that the competing evidence “as to whether endogenous and exogenous testosterone have the same or different effects on the body did not enable the Panel to draw a conclusion one way or the other.”⁹⁰ But the panel did definitively say that Chand bore the onus of showing “that the Hyperandrogenism Regulations [were] unsupported by, or not based on, scientific data and that a difference does exist between the effects of endogenous and exogenous testosterone” and that she failed to meet the onus of the burden of proof.⁹¹ Such a determination appears to support the position that, although the panel facially stated that the burden of proof would shift to the IAAF to establish that the Regulations are justifiable as reasonable and proportionate after Chand made a *prima facie* showing of discrimination,⁹² in fact Chand bore the burden of proof of her claims. This case accordingly suggests that some instances of burden of proof shifts are not as complete as panel decisions would make them appear.

One salient example in international arbitration in which the burden of proof is actually shifted is in cases of alleged doping violations before CAS. In recent doping cases, CAS and other tribunals have shifted the burden of proof to the anti-doping organization, even when the athlete is the party filing suit, due to the punitive nature of such anti-doping findings. In fact, in a rare instance of codification of the burden of proof in international law, Article 3.1 of the 2009 World Anti-Doping Code (WADC) provides that “[t]he Anti-Doping Organization shall have the burden of establishing that an anti-doping rule violation has occurred.”⁹³ Courts hearing challenges to doping sanctions have applied this shifted burden of proof to the responding party, particularly when an

of corruption exists, the burden of proof should be shifted to the suspected party . . . in some cases arbitral tribunals seem to have engaged this practice.”).

⁸⁹ *Chand*, *supra* note 3, at ¶ 441.

⁹⁰ *Id.* at ¶ 488.

⁹¹ *Id.*

⁹² *Id.* at ¶ 443.

⁹³ Rigozzi & Quinn, *supra* note 43, at 16.

anti-doping allegation is based on circumstantial evidence.⁹⁴ For example, in *USADA v. Collins*,⁹⁵ the [North American Court of Arbitration for Sport Panel] required the U.S. Anti-Doping Agency (USADA) to prove beyond a reasonable doubt that the athlete used a prohibited substance or technique.⁹⁶ The takeaway from the example of doping is simply that, when the stakes are high and evidence relevant to the merits of the claims are in the hands of the respondent, not the initial claimant, shifting the burden of proof to the respondent party (which also happens to be the regulating body) can help achieve the norms of fairness and equality as outlined by the IBA.⁹⁷

C. International Arbitration Norms on the Standard of Proof

In addition to deciding how to allocate the burden of proof between parties, courts in international arbitration must also establish the standard of proof, which “defines how much evidence is needed to establish either an individual issue or the party’s case as a whole.”⁹⁸ The most relevant norm in this context consists of the standard of proof that is most often applied in substantively similar cases. As Professor Antonio Rigozzi and Brianna Quinn assert, in international sports law, the familiar balance of probabilities standard is typical for civil proceedings, while a standard of beyond a reasonable doubt is applied for criminal proceedings.⁹⁹ Additional standards include:

- a) ‘Comfortable satisfaction’ – a standard of proof that is stated to be lower than the criminal standard of beyond reasonable doubt, but higher than the civil standard of balance of probabilities;
- b) ‘Personal conviction’ – most akin to comfortable satisfaction and found in [Article] 97 of the FIFA Disciplinary Code; and

⁹⁴ Richard H. McLaren, *An Overview of Non-Analytical Positive & Circumstantial Evidence Cases in Sports*, 16 MARQ. SPORTS L. REV. 193, 194 (2006) (describing different standards of proof and circumstantial evidence used in various example cases).

⁹⁵ *USADA v. Collins*, AAA No. 30 190 00658 04 (Dec. 2004).

⁹⁶ McLaren, *supra* note 94, at 198 (citing *USADA v. Collins*).

⁹⁷ *See, for example*, in the international human rights context, *Gómez-Palomino v. Peru*, Merits, Reparations, and Costs, Judgment, Inter-Am. Ct. H.R. (ser. C) No. 136, ¶ 106 (Nov. 22, 2005) (pointing out “that forced disappearance . . . requires the State to comply with its international obligations in good faith and to provide all necessary information . . . Consequently, any attempt to shift the burden of proof to the victims or their next of kin is contrary to [the international obligations imposed upon the State]”).

⁹⁸ Menaker & Greenwald, *supra* note 81, at 78 (quoting *Rompotrol Group N.V. v Romania*).

⁹⁹ Rigozzi & Quinn, *supra* note 43, at 26.

c) ‘Preponderance of the evidence’ – akin to the balance of probabilities and featured in Section 3(G)(a) of the Uniform Tennis Anti-Corruption Program.¹⁰⁰

Given the civil law nature of many cases and the discretion afforded to panels to apply a particular standard of proof, the default standard of proof in most disputes in international arbitration is currently the balance of probabilities.¹⁰¹

However, international sports law has generally differed from international arbitration law in the sense that it is more willing to deviate from the balance of probabilities standard because “the comfortable satisfaction standard of proof works very well for sport, particularly as it is clearly defined and then applied in a consistent manner.”¹⁰² Practically speaking, a comfortable satisfaction standard requires the bearer of the burden of proof to present a stronger case than it would have had to if the standard was the balance of probabilities.¹⁰³ In fact, CAS has broken from international legal tradition by often requiring one standard of proof for the sporting organization, and a lower one for the athlete.¹⁰⁴

This international sports law norm has developed perhaps as a response to the overarching norm of requiring a heightened standard of proof under qualifying circumstances. “[F]or matters that have serious implications for a tribunal’s jurisdiction[] or are quasi-criminal in nature (e.g., allegations of fraud, corruption etc.) or if the arbitration rules specif[y] a heightened standard, tribunals have applied a heightened standard of proof.”¹⁰⁵ In particular, CAS panels have adopted the comfortable satisfaction standard, which is commonly used in the international sports arbitration context,¹⁰⁶ when confronted with “cases involving personal reputation and professional misconduct.”¹⁰⁷ For instance, in *N., J., Y., W. v. FINA*,¹⁰⁸ the second case in which a standard of comfortable satisfaction was applied, the panel noted that “a lower standard of proof than is required in a

¹⁰⁰ *Id.*

¹⁰¹ Duggal, *supra* note 68, at 42–43.

¹⁰² Davies, *supra* note 42, at 22.

¹⁰³ *See Rigozzi & Quinn, supra* note 43, at 26.

¹⁰⁴ *See id.* at 15–16.

¹⁰⁵ Duggal, *supra* note 68, at 42. *See also* Davies, *supra* note 42, at 4 (highlighting the *Briginsbaw* test, which established that “that the more serious the allegation and its consequences, the higher the level of proof required for a matter to be substantiated. The standard is not beyond reasonable doubt, but the more serious the allegation, the more persuasive the proof must be.”).

¹⁰⁶ Rigozzi & Quinn, *supra* note 43, at 27.

¹⁰⁷ Peter Charlish, *The Biological Passport: Closing the Net on Doping*, 22 MARQ. SPORTS L. REV. 61, 66 (2011).

¹⁰⁸ *Award, CAS Case No. 1998/A/208 22 December 1998*, in DIGEST OF CAS AWARDS II 1998-2000, 234 (Matthieu Reeb ed., 2002).

criminal case is appropriate in doping cases . . . because disciplinary cases are not of a criminal nature. Rather, according to the panel, disciplinary cases are of a private law of association nature.”¹⁰⁹

Again, athlete challenges to doping sanctions provide a robust example. In addition to allocating the burden of proof to the anti-doping agency, Article 3.1 of the WADC states that “[t]he standard of proof shall be whether the Anti-Doping Organization has established an anti-doping rule violation *to the comfortable satisfaction of the hearing body, bearing mind the seriousness of the allegation which is made.*”¹¹⁰ The subsequent adoption of the WADC by international sports federations such as the IAAF has normalized this heightened standard of proof in doping cases.¹¹¹ The application of the comfortable satisfaction standard has grown such that a 2014 CAS media release noted that the standard of comfortable satisfaction was a constant across decisions for over thirty athletes in anti-doping cases.¹¹² One justification for this widespread use of a heightened standard of proof in doping challenges is due to “the gravity of the allegations . . . [in which] [t]he only thing that can be said for certain is that a higher standard of proof will be required in these cases.”¹¹³ Interestingly, the international sporting community accepts this heightened standard of proof in cases where the challenged sanction is a two-year-long ineligibility, leading one to believe that CAS would have no problem applying the comfortable satisfaction standard where the ineligibility period is much longer, as in Semenya’s case.¹¹⁴

Doping challenges are not the only cases in which CAS has applied the standard of comfortable satisfaction. For example, “[w]ith respect to FIFA disciplinary proceedings, CAS has also found the standard of comfortable satisfaction to be applicable on the basis of the wording in Article 97 of the FIFA Disciplinary Code (FIFA DC),”¹¹⁵ particularly in quasi-criminal contexts. In the alleged bribery case of *Bin Hammam v. FIFA*,¹¹⁶ the CAS panel applied a standard of comfortable satisfaction when considering Mr. Bin Hammam’s appeal of his

¹⁰⁹ Michael Straubel, *Enhancing the Performance of the Doping Court: How the Court of Arbitration for Sport Can Do its Job Better*, 36 LOY. U. CHI. L.J. 1203, 1268 (2005).

¹¹⁰ World Anti-Doping Code, art. 3.1 (2015) (emphasis added).

¹¹¹ *Id.*

¹¹² CAS, *The Court of Arbitration for Sport (CAS) Delivers Two Reasoned Awards in the Matter of 39 Russian Athletes VS/ The IOC*, CAS MEDIA RELEASE (Apr. 23, 2018), <http://perma.cc/XFM9-DHHP>.

¹¹³ McLaren, *supra* note 94, at 212.

¹¹⁴ *See id.*

¹¹⁵ Rigozzi & Quinn, *supra* note 43, at 29.

¹¹⁶ CAS 2011/A/2625 (July 19, 2012).

“lifetime ban from the [FIFA] for allegedly offering bribes to buy votes in a FIFA election” for FIFA President.¹¹⁷ The CAS panel in this case

overturned FIFA’s lifetime ban, even though it found ‘it to be more likely than not that Mr[.] Bin Hammam was the source of the money,’ [because of its conclusion that the evidence] does not permit the majority of the Panel to reach the standard of comfortable satisfaction in relation to the matters on which [Mr. Bin Hammam] was charged.¹¹⁸

The *Bin Hammam* case thus provides another example of when CAS has chosen to apply a heightened standard of proof: the standard of comfortable satisfaction was required because this was a case in which the sporting organization abridged an individual’s right to participate in a sporting activity.

D. International Arbitration Norms on the Evaluation of Evidence

Just as codified international law on arbitration proceedings does not supply much guidance on applications of the burden and standard of proof, it also fails to shed much light on the admissibility and evaluation of evidence before a tribunal,¹¹⁹ especially for the type of scientific evidence relevant to cases like Semenya’s. First, with respect to admissibility of evidence, “both Article 184(1) PILA and 9(1) of the IBA Rules support the ability of the arbitrators to decide whether or not a given piece of evidence is admissible, with Article 9(1) specifically providing that the ‘Arbitral Tribunal shall determine the admissibility . . . of evidence.’”¹²⁰ Even with an entire article devoted to the “Admissibility and Assessment of Evidence,” the IBA Rules on Evidence primarily focus on production issues, such as protecting privileged and confidential documents, but not necessarily the evaluation of evidence.¹²¹ Rather, Article 9(1) reflects the general tone of allowing the arbitral tribunal determine the evidentiary rules according to the panel’s best judgment in keeping with “considerations of procedural economy, proportionality, fairness or equality of the Parties.”¹²² Further, in international sports, “[t]here is no rule in CAS Code to define what may, or may not, be admitted in terms of evidence.”¹²³

¹¹⁷ Menaker & Greenwald, *supra* note 81, at 86.

¹¹⁸ *Id.*

¹¹⁹ Isabelle Van Damme, *The Assessment of Expert Evidence on International Arbitration*, 9 J. OF INT’L DISP. SETTLEMENT 401, 401 (2018).

¹²⁰ Rigozzi & Quinn, *supra* note 43, at 39 (quoting IBA, *IBA Rules on the Taking of Evidence in International Arbitration*, art. 9.1 (May 29, 2010)).

¹²¹ See ASHFORD, *supra* note 69, at 143.

¹²² *Id.* at 141.

¹²³ Rigozzi & Quinn, *supra* note 43, at 39.

International law outside of the sporting arbitration context likewise provides little guidance on the evaluation of evidence. Whereas the IBA Rules on Evidence at least dictate overarching principles arbitrators should adhere to when admitting evidence and hearing expert testimony, the Rules are largely silent on how an arbitral panel should evaluate that evidence.¹²⁴ Similarly, the WTO Appellate Body provides only cursory guiding principles, mandating that panels “must ‘verify that the scientific basis comes from a respected and qualified source,’” and that “‘the views must be considered to be legitimate science according to the standards of the relevant scientific community,’”¹²⁵ but does not detail specifics on how panels should weigh and consider such evidence. In fact, the dearth of guidance on evaluating evidence has led commentators to conclude that “[t]here is no steadfast rule determining what kind of evidence must be considered of higher or of better quality per se.”¹²⁶

Since arbitrators rely on their discretion to determine the rules in their proceedings, some commentators have argued that arbitrators should explain: the type of proof they are looking for and whether proof rendered by the parties is sufficient or not; their level of satisfaction from the proof provided by the parties and whether they will draw an adverse inference, if one party does not submit a certain type of proof, requested by the arbitrators; the methodology by which they will evaluate evidence.¹²⁷

The above framework is particularly informative for CAS, for “each CAS panel has the freedom to decide the evidentiary weight of any evidence on the record unless such freedom is limited in the relevant regulations.”¹²⁸ Because “[t]he CAS Panel also has the freedom to choose between contradictory elements of evidence in the decision-making process,”¹²⁹ it is important for the arbitrators to fully detail their evidentiary procedures in fairness to the parties. Thus, the fact that international arbitrators are accorded high discretion and deference in determining the rules of evidence in their respective proceedings need not be a detriment to fairness to the parties; rather, it can be an avenue for furthering justice by providing maximal flexibility for each panel to adapt to the needs of its current case.

¹²⁴ See generally ASHFORD, *supra* note 69; ZUBERBÜHLER ET AL., *supra* note 72.

¹²⁵ Appellate Body Reports, U.S. – Continued Suspension of Obligations in the EC – Hormones Dispute/Canada – Continued Suspension of Obligations in the EC – Hormones Dispute, ¶ 591, WTO Doc. WT/DS320/AB/R, WT/DS321/AB/R (adopted Nov. 12, 2008).

¹²⁶ O’MALLEY, *supra* note 64, at 197.

¹²⁷ Bhushan, *supra* note 63, at 610.

¹²⁸ Rigozzi & Quinn, *supra* note 43, at 52.

¹²⁹ *Id.*

The 2015 *Chand* case put the need for greater guidance on the admissibility and evaluation of evidence on stark display. In *Chand*, the CAS panel confronted what can only be described as a confusing array of scientific opinions. For instance, Professor Richard Holt, an expert testifying for Chand, explained that in his opinion,

the Hyperandrogenism Regulations [were] ‘scientifically unsound.’ Professor Holt identified two scientific flaws in the assumption that elevated endogenous testosterone in female athletes confers a competitive advantage. First, endogenous testosterone does not explain the difference between male and female athletic performance. Second, there is no convincing evidence that endogenous testosterone enhances athletic performance in female athletes, including those with hyperandrogenism.¹³⁰

Another expert for Chand, Dr. Sari van Anders, stated that “the existing scientific research does not establish that endogenous testosterone is the basis for successful athletic performance.”¹³¹ During the hearing, even the IAAF’s experts “agreed that, ‘we don’t have much evidence’ and there was no ‘definitive proof’ of the link [between testosterone levels and elevated athletic performance in women],” but opined that the “available science suggests that testosterone is ‘the most important factor’ that could explain the difference.”¹³² Further, the *Chand* panel had to grapple with direct versus indirect evidence, as another of the experts, Dr. Stéphane Bermon, “accepted that the evidence was indirect and said that it was not possible to obtain direct evidence”¹³³ on the effect of elevated testosterone levels on athletic performance. The barrage of conflicting scientific evidence caused the panel to ultimately defer on a final decision, asking the IAAF to present more concrete evidence on the necessity of the Hyperandrogenism Regulations. Without a clearer standard for admitting and evaluating scientific evidence, however, subsequent CAS panels hearing similar issues may confront the same problems as the *Chand* panel.

IV. CAS PROCEEDINGS THEN AND NOW: AN ARGUMENT FOR CHANGES TO THE EVIDENTIARY RULES

Having discussed the contours of international law governing international arbitral tribunals in general, this Comment now turns to Semenya’s challenge of the new IAAF regulation barring women with testosterone levels above 5 nmol/L from certain international competition track and field events.¹³⁴ While the

¹³⁰ *Chand*, *supra* note 3, at ¶ 136.

¹³¹ *Id.* at ¶ 163.

¹³² *Id.* at ¶ 198 (citations omitted).

¹³³ *Id.* at ¶ 214.

¹³⁴ *See* Roe, *supra* note 14.

arguments against the new regulation range from gender discrimination (there is no comparable ban for male athletes with atypical testosterone levels) to potential racial discrimination (many women with elevated levels of testosterone come “from developing nations who do not conform to Western standards of femininity”¹³⁵), this Comment focuses on the fairness of the CAS proceeding itself. As the sections below demonstrate, the CAS panel in the *Chand* case, which was the last panel to have considered a challenge to a similar ban, did not establish evidentiary rules that reflected the parties’ relative access to the requisite evidence. Further, based on the limited information in CAS’s media release on the outcome of Semenya’s case, the most recent official publication on the decision as this Comment goes to press, the proceedings in Semenya’s case did not substantially differ from those in *Chand*.¹³⁶ Because the media release did not expound further on the procedural aspects of the case,¹³⁷ this Comment assumes that the panel did not deviate from the proceedings laid out in *Chand*. In light of this reality, this Comment argues that the *Semenya* panel should have: (1) shifted the burden of proof onto the regulating body; (2) applied a heightened standard of proof of comfortable satisfaction, rather than balance of probabilities, due to the high-stakes nature of such cases; and (3) adopted a *Daubert*-like standard for evaluating scientific evidence, particularly scientific studies.¹³⁸

A. CAS Should Have Shifted the Burden of Proof to Regulating Bodies

As mentioned in Section III, the lack of binding authority on most international arbitral tribunals has led to the development of a number of norms governing the allocation of the burden of proof. Perhaps the most notable of these norms is that tribunals elect to shift the burden of proof to the party that asserts a fact under certain circumstances. Such a nonbinding legal framework affords international tribunals great latitude in adjusting to the circumstances of each case. For many cases in international arbitration, the claimant and the respondent appear to have high parity in resources and access to information (as in arbitration between sovereign nations, for example).¹³⁹ However, this parity of power is unlikely to exist in many cases before CAS, because CAS serves as the tribunal for

¹³⁵ See Longman, *supra* note 22.

¹³⁶ 2019 CAS Media Release, *supra* note 10.

¹³⁷ *Id.*

¹³⁸ See *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579 (1993). See Section IV.C below for a more detailed analysis.

¹³⁹ See, for example, *Waguih Elie George Siag & Clorinda Vecchi v. the Arab Republic of Egypt*, ICSID Case No. ARB/05/15 (2009); Alan T. Blackwell, *Acid Rain: Corrosive Problem in Canadian-American Relations*, 47 SASK. L. REV. 1, 35 (1982).

managing disputes between athletes and the international federations governing sports. Therefore, in order to adhere to the overarching principles of fairness to the parties as well as procedural economy, CAS should exercise its discretion in employing a particular burden of proof allocation for cases that involve limitations on individuals' right to participate in international sporting events.

First, CAS should consider procedural fairness when assigning the burden of proof between parties. The IBA Rules on Evidence repeatedly reaffirm the idea that the taking of evidence should be conducted with "an efficient, economical and fair process"¹⁴⁰ in mind. While this wording can be construed to speak primarily to the time and money spent on the arbitrations, it also reminds arbitral tribunals that, fundamentally, all parties deserve a fair hearing, for "[i]t is axiomatic that a party who is not afforded a fair opportunity to present its evidence will not have been afforded due process."¹⁴¹ Further, it has been noted that, "[t]o provide procedural fairness and substantive justice, a private legal system for resolving Olympic and international sports disputes must have, at a minimum . . . a full and fair opportunity for all parties to be heard."¹⁴²

Given this norm of procedural fairness to the parties, Semenya should have submitted, and the panel should have considered and accepted, an argument that the IAAF bears the burden of definitively demonstrating that women with testosterone levels above 5 nmol/L have an unacceptable advantage over other women in athletic competition because the IAAF is the party moving to infringe on Semenya's human right to compete.¹⁴³ Whether it is proving that women with such testosterone levels do possess a notable, unacceptable advantage, or that the regulation as written is reasonable and proportionate, the IAAF should carry the burden of proof. It would have been unfair for Semenya to bear the onus of proving the IAAF's premise false, when in reality the IAAF was the party that had gone on the offensive and deprived Semenya of her eligibility to compete without any wrongdoing on her part.¹⁴⁴ One can even favorably compare this to the principle that in criminal cases, the prosecutor bears the burden of proving the accused guilty, because the prosecutor is asserting a fact which, if true, would deprive the defendant of his or her human right to liberty.¹⁴⁵

¹⁴⁰ IBA, *IBA Rules on the Taking of Evidence in International Arbitration*, art. 2.1 (May 29, 2010) [hereinafter IBA Rules].

¹⁴¹ O'MALLEY, *supra* note 64, at 5.

¹⁴² Mitten, *supra* note 45, at 20.

¹⁴³ Peterson, *supra* note 27, at 329.

¹⁴⁴ In fact, upon announcing the new regulation, the IAAF itself noted that the regulations were not "intended as any kind of judgment . . . on any athlete." See IAAF, *supra* note 9.

¹⁴⁵ See Menaker & Greenwald, *supra* note 81, at 85.

Another argument in favor of shifting the burden of proof to preserve procedural fairness is related to the clear information asymmetry between Semenya and the IAAF. When it announced the new DSD Regulations, the IAAF asserted that its regulation was grounded in “a decade and more of research.”¹⁴⁶ The IAAF’s apparent possession of a wealth of research actually cuts in favor of shifting the burden of proof, as it proves that the IAAF had the resources to defend its assertion that allowing women with naturally high testosterone levels to compete alongside women who have “normal” ranges of testosterone results in unfair competition.¹⁴⁷ Further, the fact that scientists are often commissioned by sports regulating bodies increases the risk that the scientific studies conducted will be skewed in favor of the regulating organizations’ viewpoint.¹⁴⁸ Here, since the IAAF commissioned the study, those conducting the study likely had a conflict of interest resulting in a non-neutral study. Add on the fact that the hearing before CAS had been expedited in this case, effectively preventing Semenya from commissioning robust studies of her own to refute the IAAF’s presentation of evidence, even assuming she had the resources to do so, and the inequity between the parties becomes even more striking. In light of this egregious imbalance in access to evidence regarding the effect of endogenous testosterone levels on fair competition, combined with the fact that the IAAF is the party moving to abridge the rights of a class of individuals, the case for a shifted burden of proof was urgently compelling.

One challenge to this approach is that CAS has not applied such a shift in the most relevant, though non-binding, precedent. In its decision in *Chand*, that CAS panel noted that Chand initially “suggested that IAAF bore the burden of establishing the scientific basis of the Hyperandrogenism Regulations to the comfortable satisfaction of the Panel . . . [but] appeared to resile [sic] from that position during the hearing by expressly accepting that Athlete bears the burden of proof on the issue of scientific basis.”¹⁴⁹ While it is unclear why Chand changed her position, the fact remains that ultimately, despite saying that the burden of proof shifted “to the IAAF to establish that the Hyperandrogenism Regulations are necessary, reasonable and proportionate for the purposes of establishing a level playing field for female athletes,”¹⁵⁰ the panel still implied that Chand needed

¹⁴⁶ See IAAF, *supra* note 9.

¹⁴⁷ See Roe, *supra* note 14.

¹⁴⁸ See Maureen A. Weston, *Doping Control, Mandatory Arbitration, and Process Dangers for Accused Athletes in International Sports*, 10 PEPP. DISP. RESOL. L.J. 5, 36-37 (2009) (discussing how the “Code of Silence” among WADA laboratories create a burden on the athlete challenging the laboratories’ findings).

¹⁴⁹ *Chand*, *supra* note 3, at ¶ 442.

¹⁵⁰ *Id.* at ¶ 450.

to prove that the available science did not support Hyperandrogenism Regulations. Essentially, the CAS panel held that because the parties presented competing scientific evidence and the IAAF demonstrated that the Hyperandrogenism Regulations were based on data available to it, Chand still needed to show that the IAAF did not have a proper basis for its Regulation.¹⁵¹ Ultimately, the CAS panel's ruling did not clearly assign the burden of proving the necessity of the Hyperandrogenism Regulations to the IAAF. Semenya could have argued that, by deferring implementation of the 2014 Hyperandrogenism Regulations, the CAS panel acknowledged the IAAF's evidentiary burden when it sought to exclude individuals from competition based on a genetically predetermined condition.

Mitigating the challenge presented by the *Chand* panel's equivocation on the burden of proof, however, is the fact that a shifted burden of proof is not unprecedented, even in the limited realm of cases before CAS: the challenges to doping violations serve as a comparison point for Semenya. For just as CAS sought fit to shift the burden of proof onto the regulating body in doping cases, so too should it have done the same in Semenya's case, where the right to participate in international sporting events is similarly, if not more heavily, implicated. The argument would have been that the regulating body in this case, the IAAF, should bear the burden of proof, since it initiated the limitation on participation in international sports. This is akin to the scenario in which the anti-doping organization that is limiting participation in sports through sanctions on the accused athlete must prove the accuracy of its allegations.¹⁵² In short, persuasive precedent exists for both positions on allocating the burden of proof between the two parties. The CAS panel assigned to Semenya's case should accordingly have considered the factors of procedural economy and overarching fairness as described above in deciding the appropriate allocation of the burden of proof.

¹⁵¹ *Id.* at ¶ 488.

¹⁵² Meredith Lambert, *The Competing Justices of Clean Sport: Strengthening the Integrity of International Athletics While Affording a Fair Process for the Individual Athlete under the World Anti-Doping Program*, 23 TEMP. INT'L & COMP. L.J. 409, 422 (2009); see also Weston, *supra* note 148, at 34 ("In the 2003 Code, an athlete successfully shifted the burden of proof 'back' to the antidoping agency . . . The anti-doping agency then had to prove to the 'comfortable satisfaction of the hearing body' that the departure had not caused the adverse analytical finding."); Rigozzi & Quinn, *supra* note 43, at 53.

B. CAS Should Have Applied a Standard of Proof of Comfortable Satisfaction, Not Balance of Probabilities, to the Claims Brought by the IAAF

In the context of selecting a standard of proof, international norms also provide arbitral tribunals with great flexibility.¹⁵³ Panels hearing cases on international sports law often apply either the balance of probabilities or the comfortable satisfaction standard of proof.¹⁵⁴ Regardless of the final allocation of the burden of proof, the standard of proof required in cases that implicate an athlete's right to participate in sports or that otherwise involve high stakes consequences in the sporting world should be higher than merely "more likely than not[.]" In the present case, the CAS panel should only have accepted the IAAF's claim that allowing women with endogenously high testosterone to compete would interfere with fair and meaningful competition if it was comfortably satisfied that the evidence supported this proposition.

The argument for adopting such a standard of proof once again relies on the fact that this case involves important substantive rights for an entire class of athletes. According to the Olympic Charter, National Olympic Committees must work to "ensure that no athlete 'has been excluded for racial, religious or political reasons or by reason of other forms of discrimination.'"¹⁵⁵ This directive reflects the universal value that "[t]he practice of sport is a human right,"¹⁵⁶ and there is no intelligible reason why other international sporting organizations should not be similarly enjoined from discriminatorily excluding an athlete. Recognizing that "the 'comfortable satisfaction' standard must be applied with respect to both the seriousness of the offence and the type of the offence,"¹⁵⁷ CAS panels have discretionarily applied this standard in numerous cases. Specifically, the comfortable satisfaction standard

is most commonly applied in disciplinary proceedings, due to the often serious repercussions of being found guilty of a relevant offence [because t]he use of this standard offers somewhat of a safeguard to the accused, requiring the satisfaction of the offence to a higher standard than that typically used in civil proceedings.¹⁵⁸

¹⁵³ See Bhushan, *supra* note 63, at 605 ("Even though it is largely accepted as a general rule that the burden of proof lies on the party seeking to establish its particular claim or defense, the standard of proof . . . in an international arbitration has not evolved into a general rule.")

¹⁵⁴ Rigozzi & Quinn, *supra* note 43, at 27.

¹⁵⁵ Mitten, *supra* note 45, at 7.

¹⁵⁶ Peterson, *supra* note 27, at 329.

¹⁵⁷ Rigozzi & Quinn, *supra* note 43, at 28.

¹⁵⁸ *Id.* at 27.

Once again, challenges to doping violations sanctions prove instructive. “The comfortable satisfaction standard is well-established in doping cases, and features in various articles in the WADC.”¹⁵⁹ For instance, the CAS panels in *Oliveira v. USADA*¹⁶⁰ and *Querimaj v. IWF*,¹⁶¹ cases in which the accused athletes had ingested forbidden substances thought to enhance athletic performance, applied the standard of comfortable satisfaction and held that the athlete needed only prove that “ingestion of [the specified substance] was not intended to enhance” athletic performance.¹⁶² Likewise, the CAS panel in *Veerpalu v. International Ski Federation*¹⁶³ “applied the standard of comfortable satisfaction to invalidate a regulation for a human growth hormone testing procedure for doping cases.”¹⁶⁴ Although not every doping challenge before CAS warrants an application of the comfortable satisfaction standard,¹⁶⁵ such a standard is commonly accepted in doping violations challenges because the stakes are high for the challenging athlete. Semenya’s case exemplified the scenarios in which a comfortable satisfaction standard should apply: were the new DSD Regulations to take effect, they would deprive an entire class of athletes from certain international sporting events indefinitely, which constitutes a deprivation of an individual’s right to participate in sports.¹⁶⁶

The most formidable challenge to raising this argument in Semenya’s case was the fact that the closest precedent, *Chand*, did not apply the standard of comfortable satisfaction. In fact, the panel in *Chand* explicitly considered the comfortable satisfaction standard and yet adopted the balance of probabilities standard, referencing the *Pistorius v. IAAF*¹⁶⁷ panel. The *Pistorius* panel “noted that the standard of proof is clearly not the ‘beyond reasonable doubt’ standard applicable in criminal cases in most jurisdictions.”¹⁶⁸ The *Chand* panel went on to reject the comfortable satisfaction standard, although it did mitigate its statements by conceding “that the requisite standard to justify discrimination of a fundamental right, including the right to compete that is recognised in the

¹⁵⁹ *Id.*

¹⁶⁰ *Oliveira v. USADA*, CAS 2010/A/2107 (Dec. 6, 2010).

¹⁶¹ *Querimaj v. IWF*, CAS 2012/A/2822 (Sept. 12, 2012).

¹⁶² *Oliveira*, *supra* note 160, at ¶ 9.17.

¹⁶³ *Veerpalu v. International Ski Federation*, CAS 2011/A/2566 (Mar. 25, 2013).

¹⁶⁴ *Chand*, *supra* note 3, at ¶ 446.

¹⁶⁵ *See Mitten*, *supra* note 45, at 34–35.

¹⁶⁶ *Peterson*, *supra* note 27, at 329.

¹⁶⁷ *Pistorius v. IAAF*, CAS 2008/A/1480 (May 16, 2008).

¹⁶⁸ *Chand*, *supra* note 3, at ¶ 446.

Hyperandrogenism Regulations themselves, requires the IAAF to overcome positively the onus to establish that justification.”¹⁶⁹

In response to this unfavorable precedent, Semenya should have relied on another accepted norm, that “CAS panels are not bound by findings of other arbitrators in earlier cases,”¹⁷⁰ particularly when a unique set of facts counsels against being bound by inapplicable rulings in previous cases. Indeed, one major difference between this most recent case and *Chand* is the fact that this time, the new DSD Regulations seem to be targeted, and thus more egregiously discriminatory, than the ones at issue in *Chand*. Specifically, the DSD Regulations appear to be directed at Semenya because the restricted events named in the regulations range include “400m, hurdles races, 800m, 1500m, one mile races and combined events over the same distances,”¹⁷¹ yet the studies purported to justify the regulations suggest “that female athletes with higher levels of testosterone had an advantage in five events[:] the 400m, 400m hurdles, 800m, hammer throw, and pole vault.”¹⁷² The discrepancy between the events listed in the study as being most susceptible to unfair competition and the events actually restricted by the DSD Regulations thus seems “like a very arbitrary and selective way in which to apply regulations, and [seems] targeted [toward] Semenya.”¹⁷³ Indeed, CAS’s executive summary of Semenya’s case noted “the paucity of evidence to justify the inclusion of” the 1500m and one mile events in the DSD Regulations restrictions,¹⁷⁴ bolstering the claim that the regulations were not grounded in scientifically proven outcomes, but rather were a direct response to Semenya’s personal successes. In light of this selective application of the ban, the panel should have weighed the higher risk of targeted discrimination in this case in favor of applying the higher standard of proof of comfortable satisfaction, especially since allegations of discriminatory intent generally warrant greater scrutiny.¹⁷⁵

Finally, in addition to the importance of the rights at stake for athletes who wish to compete internationally, the fact that the comfortable satisfaction standard has strong persuasive precedent in other cases before CAS, which the *Chand* panel

¹⁶⁹ *Id.* at ¶ 447.

¹⁷⁰ Dirk-Reiner Martens, *The Role of the Arbitrator in CAS Proceedings*, 2014-2 CAS BULLETIN 31, ¶ 6.2 (2014); *but see* Mitten, *supra* note 45, at 43 (“CAS panels are extremely reluctant to depart from precedent.”).

¹⁷¹ *See* IAAF, *supra* note 9.

¹⁷² *See* Andy Bull, *Caster Semenya and the IAAF: If the Science is Wrong, the Ruling is Wrong*, THE GUARDIAN (May 1, 2018), <http://perma.cc/5XTS-9J7V>.

¹⁷³ *See* Simon Franklin et al., *New Eligibility Regulations for Female Athletes with “Differences of Sex Differentiation”*, BJSM BLOG (Apr. 26, 2018), <http://perma.cc/NYL3-8Z3P>.

¹⁷⁴ CAS, *Semenya, ASA IAAF: Executive Summary*, CAS MEDIA RELEASE (May 1, 2018), <http://perma.cc/PX5S-MFJX> [hereinafter Executive Summary].

¹⁷⁵ *See generally* Peterson, *supra* note 27, for more on the antidiscrimination principle endorsed in the Olympic Charter.

itself acknowledged, weighed in favor of adopting the comfortable satisfaction standard over the balance of probabilities standard. First, the fact that such a standard is common in international sports arbitration signifies that CAS arbitrators are familiar with applying such a standard and would be able to do so effectively. Any argument that requiring a balance of probabilities standard of proof would lead to greater procedural economy would therefore not be persuasive. More importantly, adopting the comfortable satisfaction standard of proof in Semenya's challenge would have sent a strong message to the international sporting community that encroachments on athletes' rights to compete will be scrutinized carefully. By adopting this commonly accepted standard, the CAS panel in Semenya's case would have reaped the benefit of institutional knowledge by adhering to both persuasive, if non-binding, CAS precedent while also protecting the individual right to compete in athletics by attributing due weight to the claims against the ban.

C. CAS Should Have Adopted the Daubert Standard for Considering Scientific Evidence

Another evidentiary issue CAS panels must consider is what evidence to admit and how to evaluate the admitted evidence. In this context, again, the lack of international guidance allows a panel to decide the most appropriate standards on a case-by-case basis.¹⁷⁶ The nature of the evidence expected to be presented before the panel therefore becomes highly relevant.¹⁷⁷ And in Semenya's case, the controversy surrounding the IAAF's evidence in support of its DSD Regulations suggests that evidence was as contradictory evidence as in *Chand*, which resulted in a deferred implementation of the old Hyperandrogenism Regulations until more conclusive evidence came to light. The lack of guidance for both current and subsequent panels hearing such contradictory evidence rendered the issue ripe in Semenya's case. This Comment suggests the *Daubert* framework of policing and evaluating scientific evidence as a desirable standard for CAS arbitrations.

1. The *Chand* decision did not resolve how CAS panels should evaluate evidence.

The *Chand* panel left many questions unanswered, perhaps the most pressing of which was the quality of proof required to justify a ban as necessary, reasonable, and proportionate.¹⁷⁸ In that case, the panel heard undisputed evidence that there

¹⁷⁶ See Davies, *supra* note 42, at 23 (“CAS has stated that it is not bound by the rules of evidence.”).

¹⁷⁷ See Van Damme, *supra* note 119, at 402.

¹⁷⁸ Boyd, *supra* note 5, at 22–23 (“The Panel remained silent on what competitive advantage would make the Regulation valid, but the Panel reiterated that it required substantial difference.”).

is a “[10–12 percent] difference between elite male and elite female athletic performance.”¹⁷⁹ The panel held that the IAAF did not meet its burden of justifying the then-contemporary ban on women with over 10 nmol/L of testosterone from international athletic competition in part because the IAAF did not demonstrate a similar differential in athletic performance between women with testosterone levels above that cutoff as opposed to below the cutoff, or within the “normal” range for women.¹⁸⁰ However, the CAS panel did not provide more specific guidance on what would qualify as sufficient evidence to justify the IAAF’s Hyperandrogenism Regulations, for the CAS panel only suggested “that a slight advantage such as [one percent] ‘may not justify a separation of athletes within the female category, given other relevant variables that legitimately affect athletic performance.’”¹⁸¹

This lack of formal and precedential guidance only punted the problem down the road in future cases. In Semenya’s challenge, the inability to define the quality of evidence necessary to justify the Regulations created a high level of uncertainty for both Semenya and the IAAF, particularly since the scientific studies presented by the IAAF were already being challenged in the scientific community before the CAS proceedings began. For instance,

[c]ritics [of the new DSD Regulations] say IAAF researchers Stéphane Bermon and Pierre-Yves Garnier (who was sanctioned for interfering with an IAAF ethics investigation last year) essentially re-ran the statistical analyses on the same data [from *Chand*] until they produced a different conclusion. But three independent researchers—Roger Pielke of the University of Colorado-Boulder, Ross Tucker of the University of Cape Town, and Erik Boye of Oslo University Hospital—couldn’t reproduce Bermon’s and Garnier’s analyses with publicly available data.¹⁸²

If the IAAF merely presented a re-run of the studies used to justify the Hyperandrogenism Regulations at issue in *Chand*’s case, and those studies were clearly insufficient to justify that ban,¹⁸³ then the case for adopting a clear framework for evaluating the disputed evidence in future cases is even stronger.

2. The *Semenya* panel should have considered adopting the *Daubert* standard for evaluating scientific evidence.

In the absence of such guidance from international arbitration law, other frameworks could have informed the CAS panel’s decision on how to evaluate scientific evidence. One such framework was the *Daubert* standard from U.S. law,

¹⁷⁹ *Chand*, *supra* note 3, at ¶ 526.

¹⁸⁰ Buzuvis, *supra* note 7, at 40.

¹⁸¹ *Id.* (quoting *Chand*, *supra* note 3, at ¶ 527).

¹⁸² *See* Roe, *supra* note 14.

¹⁸³ Boyd, *supra* note 5, at 19–21.

which replaced the relevance and general acceptance standards previously used in American courts to evaluate evidence.¹⁸⁴ The modern *Daubert* framework arose out of three cases that “clarified and expanded the scope of the [American] Federal Rules of Evidence as they apply to expert witnesses.”¹⁸⁵ These three cases, often known as the *Daubert* trilogy, stand for the proposition that scientific evidence and expert testimony should be scrutinized for both the general methodology applied in the study as well as the soundness of the expert’s conclusions.¹⁸⁶

First, in *Daubert v. Merrell Dow Pharmaceuticals, Inc.*,¹⁸⁷ the U.S. Supreme Court overruled a precedential holding that made general acceptance the test for admitting scientific evidence.¹⁸⁸ Instead, the Court held that trial judges were to be gatekeepers of evidence and were “responsible for ensuring that an expert’s testimony was based on a reliable foundation and relevant.”¹⁸⁹ In addition to overturning the previous standard articulated in *Frye v. U.S.*,¹⁹⁰ “the Court [in *Daubert*] adopted a new framework for evaluating the reliability of scientific evidence based on four considerations: falsifiability, peer review, error rates, and ‘acceptability’ in the relevant scientific community.”¹⁹¹ In essence, the court endorsed a multi-factor standard emphasizing “the principles and methodology of the scientific proposition and not [] the proffered conclusions.”¹⁹²

¹⁸⁴ LLOYD DIXON & BRIAN GILL, CHANGES IN THE STANDARDS FOR ADMITTING EXPERT EVIDENCE IN FEDERAL CIVIL CASES SINCE THE *DAUBERT* DECISION 1 (2001).

¹⁸⁵ Vanderhart, *supra* note 63, at 4.

¹⁸⁶ *Id.* at 5.

¹⁸⁷ 509 U.S. 579 (1993).

¹⁸⁸ Vanderhart, *supra* note 63, at 4–5.

¹⁸⁹ *Id.* at 4–5.

¹⁹⁰ 293 F. 1013 (D.C. Cir. 1923).

¹⁹¹ Edward K. Cheng & Albert H. Yoon, *Does Frye or Daubert Matter—A Study of Scientific Admissibility Standards*, 91 VA. L. REV. 471, 477 (2005). See also DIXON & GILL, *supra* note 184, at 3 (stating that *Daubert* held that,

[f]or a case involving scientific evidence . . . the Supreme Court [of the U.S.] provided a list of factors that judges might consider when determining whether a theory or methodology is scientifically valid:

- whether it can be (and has been) tested
- whether it has been subjected to peer review and publication
- the known or potential rate of error
- the existence and maintenance of standards controlling the technique’s operation
- whether it is generally accepted in the scientific community.)

¹⁹² DONALD E. SHELTON, FORENSIC SCIENCE IN COURT: CHALLENGES IN THE TWENTY-FIRST CENTURY 18 (2011).

The second case of the trilogy, *General Electric Co. v. Joiner*,¹⁹³ expanded the trial judge's role as the gatekeeper of admissible scientific evidence by applying the Daubert factors to a scientific expert's conclusions as well as methodology. *Joiner* indicated that the trial judge "has the discretion to totally reject and disallow an expert's opinion, even if based on an accepted methodology, if the judge finds that the expert's conclusion is not reliably based on that methodology."¹⁹⁴ And finally, the third case in the *Daubert* trilogy, *Kumho Tire Co., Ltd. v. Carmichael*,¹⁹⁵ "extended the standards of *Daubert* to include not only expert testimony about matters that are scientific but also expert testimony based on technical and other specialized knowledge."¹⁹⁶

While at first glance the *Daubert* framework appears to be geared toward the American legal system, it can still prove instructive in international arbitration. In CAS arbitrations, the panel must act like the American trial judge in that the panel must decide what evidence to admit.¹⁹⁷ However, the two systems differ in that the decision-maker in the American legal system is often a panel of non-expert jurors, not the trial judge, whereas in CAS arbitrations, the arbitral tribunal evaluates the evidence it admitted.¹⁹⁸ The unification of the acts of determining admissibility and considering evidence on the merits into one entity in international arbitrations renders the *Daubert* framework unsuitable for structuring rules of admissibility in the arbitral hearing, in part because "[n]owhere in international arbitration . . . is there a system like that which exists in U.S. domestic courts to challenge the admissibility of evidence presented by an expert."¹⁹⁹ However, because the *Daubert* framework was intended to only allow reliable, quality evidence to be heard, it still provides ample guidance for international arbitrators in the *consideration* and *evaluation* of evidence.

In Semenya's case, for instance, the IAAF should have been required to prove its assertion that endogenously elevated testosterone in females produces an unfair athletic advantage.²⁰⁰ To do this, the IAAF should have produced studies that isolate endogenous testosterone (as opposed to exogenous testosterone or

¹⁹³ 522 U.S. 136 (1997).

¹⁹⁴ SHELTON, *supra* note 192, at 18.

¹⁹⁵ 526 U.S. 137 (1999).

¹⁹⁶ Vanderhart, *supra* note 63, at 5.

¹⁹⁷ Rigozzi & Quinn, *supra* note 43, at 51 ("Finally, once the burden of proof, the relevant standard of proof, and any questions as to the admissibility of relevant evidence have been determined, the final task to be undertaken by a Panel will be the evaluation of the evidence so admitted.").

¹⁹⁸ *Id.*

¹⁹⁹ Vanderhart, *supra* note 63, at 4. *See also* Joan E. Donoghue, *Expert Scientific Evidence in a Broader Context*, 9 J. OF INT'L DISP. SETTLEMENT 379, 386 (2018).

²⁰⁰ *See Chand*, *supra* note 3, at ¶ 454.

other factors) as the sole driver of the difference in athletic performance in hyperandrogenic women. But because there are so many inputs that result in particular levels of athletic performance,²⁰¹ and attempting to isolate testosterone as the driver of such performance “glosses over the fact that innumerable other natural and environmental factors contribute to each athlete’s relative advantages and disadvantages—from height and lung capacity to coaching and training facilities,”²⁰²—any study that purports to identify endogenous testosterone as the sole cause of such performance levels must be highly scrutinized. To evaluate a study claiming to isolate testosterone, the CAS panel could have applied some of the *Daubert* factors by asking whether the study’s conclusion had been tested, subjected to peer review, established with an acceptably low error rate, and accepted by the scientific community.²⁰³ Similarly, the CAS panel could have used these factors to test whether other scientific conclusions presented in Semenya’s case were sufficiently well-established and reliable to justify upholding the high-stakes DSD Regulations.

As with the arguments for a shift in the burden of proof and adoption of a comfortable satisfaction standard of proof proposed above, the argument for implementing the *Daubert* framework in CAS proceedings is not without its challenges. In addition to the aforementioned limitation that “arbitral proceedings [are] more akin to a bench trial than a jury trial,”²⁰⁴ the fact remains that other international arbitral tribunals have not appeared to adopt a similar framework for admitting and evaluating evidence. However, since “the discretionary power of the arbitrator [to appraise evidence] is not open to question,”²⁰⁵ and “the tribunal is able to accord its own weight to any evidence presented,”²⁰⁶ a CAS panel adopting a clear test for evaluating evidence that it has the discretion to consider is hardly objectionable.²⁰⁷ The fact that no comparable framework is currently used in international law is not an argument *against* applying and clarifying such a standard now, but rather an argument *for* doing so.

It must also be noted that even if the CAS panel failed to adopt a *Daubert*-like approach to evaluating evidence, the scientific studies put forth by the IAAF

²⁰¹ See Longman, *supra* note 22.

²⁰² Krech, *supra* note 31, at 265. See also Peterson, *supra* note 27, at 335.

²⁰³ See SHELTON, *supra* note 192, at 17–18.

²⁰⁴ Vanderhart, *supra* note 63, at 5.

²⁰⁵ Arbitral Award Made by the King of Spain on Dec. 23, 1906 (Honduras v. Nicaragua), Judgment, 1960 I.C.J. Rep. 215-216 (Nov. 18, 1960).

²⁰⁶ Vanderhart, *supra* note 63, at 6.

²⁰⁷ See Brendan Plant, *Expert Evidence and the Challenge of Procedural Reform in International Dispute Settlement*, 9 J. OF INT’L DISP. SETTLEMENT 464, 465 (2018) (“One of the central features of international courts and tribunals is their immense flexibility.”).

did not even meet the lower general acceptance standard the *Daubert* trilogy replaced, nor the standards from other international dispute resolution fora.²⁰⁸ The overturned *Frye* standard simply held that “the thing from which the deduction is made must be sufficiently established to have gained general acceptance in the particular field in which it belongs.”²⁰⁹ Meanwhile, the WTO Appellate Body, for example, requires scientific evidence to “have the necessary scientific and methodological rigour to be considered reputable science.”²¹⁰ The evidence put forth by the IAAF met neither of these standards, for the high level of controversy surrounding the scientific evidence and expert testimonies first in *Chand*, and then in Semenya’s case, clearly prevented those “scientific findings [from being] ‘sufficiently established to have gained general acceptance in the particular field in which [they] belong[.]’”²¹¹ At a minimum, then, the CAS panel in Semenya’s case should have rectified the laxity of the panel in *Chand* by at least vetting the scientific evidence presented against a standard of generally accepted methodology. Indeed, the IAAF studies’ failure to meet the lower bar supplied by *Frye* and the WTO Appellate Body rules exemplified the need to implement a clarified approach that allows arbitral tribunals to scrutinize evidence carefully, especially when important substantive rights are involved. Whether by crafting clear standards for evaluating scientific evidence akin to the *Daubert* standard or taking on the more active role of appointing a neutral expert,²¹² the CAS panel should have leveraged its procedural flexibility to ensure that the evidence used to substantiate either party’s claims did so reliably.²¹³

V. CONCLUSION

For over 50 years, the international sporting community has struggled with notions of fairness when it comes to classifying athletic events on a male-female binary. Today, the question remains as unsettled as the first day on which sex-testing was introduced. In announcing its most recent DSD Regulations, the IAAF renewed the controversy over the eligibility of intersex individuals to

²⁰⁸ See generally Roger Pielke Jr. et al., *Scientific Integrity and the IAAF Testosterone Regulations*, INT’L SPORTS L. J. (2019), <http://perma.cc/H8TW-HK7J> (critiquing the study the IAAF cites as justification for the new DSD Regulations, found at: Stéphane Bermon & Pierre-Yves Garnier, *Serum Androgen Levels and Their Relation to Performance in Track and Field: Mass Spectrometry Results from 2127 Observations in Male and Female Elite Athletes*, 51 BRITISH J. SPORTS MED. 1309 (2017)).

²⁰⁹ *Frye*, *supra* note 190, at 1014.

²¹⁰ Appellate Body Reports, U.S. – Continued Suspension of Obligations in the EC – Hormones Dispute/Canada – Continued Suspension of Obligations in the EC – Hormones Dispute, ¶ 591, WTO Doc. WT/DS320/AB/R, WT/DS321/AB/R (adopted Nov. 14, 2008).

²¹¹ Cheng & Yoon, *supra* note 191, at 476.

²¹² See generally Donoghue, *supra* note 199, at 380 for more on appointed experts.

²¹³ For a longer discussion of evidentiary procedural reforms in international dispute settlement, see Plant, *supra* note 207, at 467–71.

participate as females in international sporting events, this time with respect to one particular athlete: Caster Semenya. The new regulations triggered a case that mirrored the 2015 *Chand* case, which also challenged the Hyperandrogenism Regulations then in effect for impermissible discrimination.

While *Chand* amounted to a temporary victory for hyperandrogenic women in that it stayed the implementation of the Hyperandrogenism Regulations at the time, the proceedings left many questions unresolved. Substantively, the *Chand* panel did not reach a conclusion as to whether hyperandrogenic individuals truly receive a significant advantage in athletic performance due to their elevated testosterone levels. Procedurally, the *Chand* panel also failed to answer important evidentiary questions related in international sports arbitration. Although the *Chand* panel assigned the burden of proof to the IAAF after Chand made a *prima facie* showing that the Hyperandrogenism Regulations were discriminatory, the CAS panel still seemed to hold Chand responsible for proving the claims in her case, rather than having the IAAF bear the burden. Additionally, the CAS panel rejected the comfortable satisfaction standard of proof in favor of the balance of probabilities standard of proof following nothing more than a cursory consideration of both standards. And finally, although the panel entreated the IAAF to bring more evidence demonstrating that endogenously high testosterone levels result in unfair competition, it did not specify the strength of the evidence needed to uphold the Hyperandrogenism Regulations.

Given the indeterminacy surrounding the taking of evidence in CAS arbitrations, in cases in which athletes' fundamental right to compete are involved, arbitral panels should: shift the burden of proof onto the party abridging those rights; employ a standard of proof such that the panel can only accept such an abridgement if the panel is comfortably satisfied that the abridgement is justified; and set out a clear framework, based on the *Daubert* standard in U.S. law, for evaluating the scientific evidence presented. Although challenges to implementing each of these evidentiary standards abound, the lack of binding evidentiary rules and natural flexibility of international arbitration law provide the latitude needed to adopt the proposed measures. And the interests of the impacted athletes in their individual right to participate in sporting events demand it.