

Marquette Sports Law Review

Volume 25
Issue 1 *Symposium: Sports Law: Marquette
University Law School Faculty Perspectives*

Article 12

2014

Of Umpires, Judges, and Metaphors: Adjudication in Aesthetic Sports and Its Implications for Law

Chad M. Oldfather

Follow this and additional works at: <https://scholarship.law.marquette.edu/sportslaw>



Part of the [Entertainment, Arts, and Sports Law Commons](#)

Repository Citation

Chad M. Oldfather, *Of Umpires, Judges, and Metaphors: Adjudication in Aesthetic Sports and Its Implications for Law*, 25 Marq. Sports L. Rev. 271 (2014)

Available at: <https://scholarship.law.marquette.edu/sportslaw/vol25/iss1/12>

This Symposium is brought to you for free and open access by the Journals at Marquette Law Scholarly Commons. For more information, please contact elana.olson@marquette.edu.

OF UMPIRES, JUDGES, AND METAPHORS: ADJUDICATION IN AESTHETIC SPORTS AND ITS IMPLICATIONS FOR LAW

CHAD M. OLDFATHER*

I. INTRODUCTION

While baseball's status as America's pastime may be in peril,¹ it undoubtedly remains the source of America's favorite analogy for judging. The notion of the "judge as umpire" reached its high point during Chief Justice Roberts' confirmation hearing.

Judges and justices are servants of the law, not the other way around. Judges are like umpires. Umpires don't make the rules; they apply them.

The role of an umpire and a judge is critical. They make sure everybody plays by the rules.

But it is a limited role. Nobody ever went to a ball game to see the umpire

. . . .

I have no agenda, but I do have a commitment. If I am confirmed, I will confront every case with an open mind. I will fully and fairly analyze the legal arguments that are presented. I will be open to the considered views of my colleagues on the

* Professor of Law, Marquette University Law School. Many thanks to Roy Englert, Ryan Scoville, Jordan Singer, and Howard Wasserman for their comments on an earlier draft.

¹ E.g., Jonathan Mahler, *Is the Game Over?*, N.Y. TIMES (Sept. 28, 2013), http://www.nytimes.com/2013/09/29/opinion/sunday/is-the-game-over.html?pagewanted=all&_r=1&; Kyle Smith, *America's Pastime*, N.Y. POST (Dec. 18, 2011), <http://nypost.com/2011/12/18/americas-pastime/>; Joey Spitz, *Is America's Pastime Dying a Slow Death?*, HUFFINGTON POST, http://www.huffingtonpost.com/joey-spitz/is-americas-pastime-dying_b_3686366.html (last updated Oct. 1, 2013).

bench. And I will decide every case based on the record, according to the rule of law, without fear or favor, to the best of my ability. And I will remember that it's my job to call balls and strikes and not to pitch or bat.²

The Chief Justice was, of course, hardly the first to draw this analogy,³ and he will certainly not be the last. The comparison is natural, for both roles require their occupant to “make the call,” or, more formally, to serve as the presumptively final adjudicator of the rights of competing parties.

Yet while the “judge as umpire” seems to have naturally captured the imagination of judges, commentators, and laypersons alike, it has also come in for its share of critique, with critics pointing out the various ways in which the comparison is inapt.⁴ Perhaps the most frequently mentioned distinction is that umpires, unlike judges, play no role in creating or refining the rules they must apply.⁵ Thus, commentators have proposed substitute analogies, including the “justice as commissioner”⁶ and “justice as color commentator.”⁷

I do not intend in this essay to add to these critiques. Instead, I will suggest that there is a better analogy—the somewhat less euphonious notion of the

² Transcript: Day One of the Roberts Hearings, WASH. POST (Sept. 13, 2005), available at <http://www.washingtonpost.com/wp-dyn/content/article/2005/09/13/AR2005091300693.html>.

³ For an overview of the history of the “judge as umpire,” see Aaron S.J. Zelinsky, *The Justice as Commissioner: Benching the Judge-Umpire Analogy*, 119 YALE L.J. ONLINE 113, 114–18 (2010), available at <http://www.yalelawjournal.org/forum/the-justice-as-commissioner-benching-the-judge-umpire-analogy>. See also Chad M. Oldfather, *The Hidden Ball: A Substantive Critique of Baseball Metaphors in Judicial Opinions*, 27 CONN. L. REV. 17, 42–46 (1994).

⁴ See, e.g., Theodore A. McKee, *Judges as Umpires*, 35 HOFSTRA L. REV. 1709, 1709–11 (2007); Oldfather, *supra* note 3, at 35–39; Vaughn R. Walker, *Moving the Strike Zone: How Judges Sometimes Make Law*, 2012 U. ILL. L. REV. 1207, 1210–23 (2012); Kim McLane Wardlaw, *Umpires, Empathy, and Activism: Lessons from Judge Cardozo*, 85 NOTRE DAME L. REV. 1629, 1633–34 (2010); Zelinsky, *supra* note 3, at 118–24. But see Michael P. Allen, *A Limited Defense of (at Least Some of) the Umpire Analogy*, 32 SEATTLE U. L. REV. 525, 527–28 (2009). See generally Charles Fried, *Balls and Strikes*, 61 EMORY L.J. 641 (2012).

⁵ As Judge Posner wrote of Chief Justice Roberts' analogy,

Neither he nor any other [knowledgeable] person actually believed or believes that the rules that judges in our system apply, particularly appellate judges and most particularly the Justices of the U.S. Supreme Court, are given to them the way the rules of baseball are given to umpires. We must imagine that umpires, in [addition] to calling balls and strikes, made the rules of baseball and changed them at will.

RICHARD A. POSNER, *HOW JUDGES THINK* 78–79 (2008).

⁶ Zelinsky, *supra* note 3, at 118–24.

⁷ Adam Benforado, *Color Commentators of the Bench*, 38 FLA. ST. U. L. REV. 451, 454 (2011) (citations omitted).

“judge as judge.” It may be, in other words, that the best sporting analogy to judges in law is not the umpire or referee but rather their namesakes in sport—the judges who provide the authoritative scoring in aesthetic sports, such as gymnastics and figure skating.

Judges in aesthetic sports, like their counterparts in law, derive a substantial portion of their authority and legitimacy from the expertise they bring to the position. Both, as we will see, play a significant role in shaping the content of the governing standards and draw, to a considerable degree, on their “tacit knowledge” in shaping and applying the rules.⁸ At the same time, however, their reliance on this sort of expertise—which necessarily entails the application of criteria unavailable or at least opaque to the lay observer or even the less expert participant—creates an opening for skepticism. Observers will often be unable to determine for themselves whether the judges have accurately assessed the merits, and their efforts to do so will often lead them to assessments that differ from the judges’ assessments, likely because they have not accounted for all the factors that an expert judge considers. Because these observers cannot access all the factors driving the experts’ decisions and, thus, cannot fully understand them, they can easily conclude that the judges’ decisions are being driven by improper factors. Of course, judges in aesthetic sports have been accused, and occasionally found to be guilty, of corruption, as well as less pernicious forms of bias.⁹ The same is true of judges in law. Indeed, given the enormous cognitive demands associated with judging aesthetic sports, it is unsurprising that some extraneous factors influence officials’ determinations. That, too, can be said about judging in law.

Of course, metaphor and analogy have a tendency to obscure as well as to enlighten,¹⁰ so it is wise not to press the parallel too far. There are significant differences between the two contexts, as there will be for any sport-law comparison.¹¹ Still, the comparison is worth developing, for similarity between law and aesthetic sports extends beyond the characteristics of the judicial role to include features of the larger systems in which those two types of judges operate. Whatever the allure of analogizing judges to umpires and other officials in what philosophers of sport call “*purposive*” sports, the more realistic comparison may be to judges in “*aesthetic*” sports.

⁸ See *infra* notes 7–32 and accompanying text.

⁹ See *infra* Part III.

¹⁰ See Allen, *supra* note 4, at 527–28; Oldfather, *supra* note 4, at 24–30.

¹¹ For reflections on the sport-law comparison, see, e.g., Mitchell N. Berman, “Let ‘Em Play” *A Study in the Jurisprudence of Sport*, 99 GEO. L.J. 1325, 1351–1364 (2011); J.S. Russell, *Limitations of the Sport-Law Comparison*, 38 J. PHIL. SPORT 254, 261 (2011).

My goals in this essay are, first, to outline and refine the work done by philosophers of sport with respect to the role of officials in aesthetic sports and, in particular, to highlight the significance of tacit knowledge to that role. Part of the analysis will be conceptual, and part will involve bringing together two disparate lines of research: that done by philosophers of sport, and that done by psychologists who have examined both the nature of the expertise underlying sports officiating, and the various biases that can afflict such officiating. The two lines of work have implications for one another, and the work of the psychologists is particularly useful in informing the work of the philosophers.

The second goal is simply to explore some of the implications of this work for our conceptions of the legal system. Given the above-noted limitations of metaphor and analogy, the progress made under this second goal is necessarily tentative and more suggestive of further lines of inquiry than of any firm conclusions. Nonetheless, this sort of inquiry, which is a species of comparative analysis,¹² can help to provide new perspectives on, and otherwise illuminate entrenched features, of the system. As I will suggest, the comparison illuminates several avenues of potential cross-fertilization between the two contexts.

II. PERSPECTIVES FROM THE PHILOSOPHY OF SPORT

Philosophers of sport have divided the world of sport into two rough camps. I will work with the distinction as first articulated by David Best, which is that between *purposive* and *aesthetic* sports.¹³ In purposive sports, the goal of the sport is, to a considerable degree, independent of the means of obtaining it. A football team may score a touchdown through the flawless execution of an elegantly designed play, or it may do so via a broken play featuring an assortment of fumbles, missed tackles, and other miscues along the journey to the end zone. Thus, in a purposive sport, one may “win ugly,” since what matters is obtaining a higher score than one’s opponent rather than doing so in a graceful or otherwise pleasing way. In aesthetic sports, in contrast, to the extent that a goal exists,¹⁴ it is likely to be trivial—say, getting from the diving platform into the water—while the means of achieving the goal (or, if one concludes there is no goal, executing the performance) takes center stage. One can never win ugly in

¹² See Chad M. Oldfather & Matthew M. Fernholz, *Comparative Procedure on a Sunday Afternoon: Instant Replay in the NFL as a Process of Appellate Review*, 43 IND. L. REV. 45, 48 (2009).

¹³ See GRAHAM MCFEE, *SPORT, RULES AND VALUES: PHILOSOPHICAL INVESTIGATIONS INTO THE NATURE OF SPORT* 90–91 (2004). Cf. Howard Wasserman, *Defining Sport: Intrinsic and Instrumental (Not Utilitarian) Values*, PRAWFSBLAWG (last updated Aug. 4, 2012, 7:03 PM), <http://prawfsblawg.blogs.com/prawfsblawg/2012/08/defining-sport-intrinsic-and-utilitarian-values.html>.

¹⁴ Bernard Suits argued that there is no prelusory goal in a sport such as diving. See Bernard Suits, *The Trick of the Disappearing Goal*, 16 J. PHIL. SPORT 1, 1 (1989).

an aesthetic sport, at least when compared to one's direct competitors, for an ugly performance is, by definition, a bad performance.

Bernard Suits drew the basic distinction between kinds of sports in a somewhat different way—those that are refereed, which fall within his definition of “games,” and those that are judged, which are not games but rather “performances.”¹⁵

The Olympics (as well as the Commonwealth Games, and so on) contain two distinct types of competitive event, what I have elsewhere called judged as opposed to refereed events. One is a performance and so requires judges. The other is not a performance but a rule-governed interplay of participants, and so requires not judges but law-enforcement officers, that is, referees. Performances require rehearsal, games require practice.¹⁶

Here it is useful to introduce another significant distinction, namely that between *constitutive* and *regulatory* rules. Constitutive rules are those that constitute the game, in the sense that the sport of baseball, for example, would not exist without a set of rules defining the nature of the activity.¹⁷ Regulatory rules, in contrast, place restrictions on the manner in which an activity may permissibly be undertaken. Thus, for example, the notion of a “pitch” exists only as a product of the constitutive rules of baseball, while the prohibition against doctoring the ball is the product of a regulatory rule. The distinction is, perhaps ultimately, more one of degree than of kind (one could argue that a pitcher who throws a doctored ball has not truly thrown a “pitch”), but then so, too, is the distinction between purposive and aesthetic sports, since most, if not all, sports share some characteristics of both.

In the case of what Suits calls a performance—roughly, an aesthetic sport—the assessment of any given instance of performance is to be undertaken not by reference to constitutive rules, but rather to rules of skill, which are in turn derived from a conception of what the ideal performance would look like.

¹⁵ See Bernard Suits, *Tricky Triad: Games, Play, and Sport*, 15 J. PHIL. SPORT 1, 2–3 (1988).

¹⁶ *Id.* at 2.

¹⁷ The distinction between constitutive and regulatory rules is elaborated and critiqued in MCFEE, *supra* note 13, at 35–36. As McFee characterizes it, the distinction is between “those rules which modify already existing practices (*regulative* rules) from those upon which the existence of a practice (in our case, a game or sport) is logically [dependent] . . . which do not so much determine what players do as part of the game, but rather create new forms of action.” *Id.* at 35.

In games, *rules* . . . are the crux of the matter. Just these rules generate just these skills. In performances, ideals are the crux of the [matter]. Just these ideals generate just these skills. That is why it is possible to speak of a perfect performance, at least in principle, without fear of contradiction¹⁸

As Suits suggested, the distinction between the two types of sport has implications for the nature of the role of officials in these sports. Officials in purposive sports—umpires and referees—are primarily concerned with ensuring that the contestants comply with the rules and that what gets counted for scoring purposes truly merits that distinction as a consequence of the proper application of the constitutive and regulatory rules. This undoubtedly calls for the frequent exercise of judgment: Was this pitch in the strike zone, did that defender interfere with an offensive player, and so forth. But the goal of the sport is defined by the constitutive rules and remains independent of the officials' conception of the goal. This results in a state of affairs in which, at least theoretically and, increasingly, actually, officials' rulings are subject to correction by some mechanical system.

John Russell builds on this point in considering whether umpires' calls amount to "performative utterances"—that is, whether a call creates the reality it describes (in the sense that whether a batter is out or not is purely a function of the call) and, thus, serves as "a linguistic act that brings a particular event or state of affairs into being."¹⁹ Russell concludes that they do not. If calls were simply performative utterances, Russell contends, it would not make sense to speak of an umpire's call as being correct or not. Instead, it would simply *be*, by virtue of the fact that the umpire decreed it to be so. A pitch would not be a strike or a ball, or a runner safe or out, until the umpire made the appropriate declaration. But that is obviously not correct, for an umpire in making such a call,

is also reporting on an antecedent state of affairs, and his calling you out or safe can be either right or wrong, true or false, according to how well his report reflects the events as they actually took place. Because it is, in part, a report or description of a state of affairs, we can intelligibly ask whether the runner was

¹⁸ Suits, *supra* note 15, at 6.

¹⁹ J.S. Russell, *The Concept of a Call in Baseball*, 24 J. PHIL. SPORT 21, 21 (1997). For a similar argument more directly applied to law, see generally Richard Lempert, *Error Behind the Plate and in the Law*, 59 S. CAL. L. REV. 407 (1986).

in fact out or safe according to that state of affairs.²⁰

The antecedent state of affairs Russell refers to exists because of conditions specified in the constitutive rules of the sport, which identify circumstances that trigger the applicable rule and that can be said to exist independently of the umpire's perception and characterization of them.

In aesthetic sports, in contrast, the dynamic is different. The officials do, in a meaningful way, create the reality described by their judgments. The content of the ideal performance cannot be fully captured in words. Thus, it is not primarily a question of whether a given performance was conducted within the rules (though there will typically be some such rules that must be complied with), but, rather, with how well that performance matches up against a hypothesized ideal performance. The rules of the sport can only obliquely capture the content of the ideal. The criteria in figure skating, for example, reference various technical requirements but also incorporate concepts such as “flow and effortless glide,” “cleanness and sureness,” “balance,” “projection,” “carriage,” “style and individuality/personality,” “phrasing and form,” and so on.²¹ Similarly, “[t]he International Gymnastics Federation's (FIG) scoring system for men's and women's gymnastics . . . incorporates credit for the routine's content, difficulty and execution, as well as artistry for the women.”²² Judges of equitation classes at horse shows are to look for “a workmanlike appearance, seat and hands light and supple, conveying the impression of complete control should any emergency arise.”²³ In implementing the applicable standards in these classes, judges are cautioned, among other things, not to be overly influenced by the rider's body shape or attractiveness²⁴ and not to put “too much emphasis on

²⁰ Russell, *supra* note 19, at 22.

²¹ For an overview, with links to more detailed information, see US Figure Skating, *International Judging System (IJS)*, US FIGURE SKATING, http://www.usfsa.org/New_Judging.asp?id=289 (last visited Dec. 21, 2014).

²² USA Gymnastics, *Women's Scoring*, USA GYMNASTICS, <https://usagym.org/pages/women/pages/scoring.html> (last visited Dec. 21, 2014).

²³ U.S. EQUESTRIAN FED'N, INC., HUNTER SEAT EQUITATION MANUAL 4 (2012), available at <http://horse.public.iastate.edu/Extension%20Program/Judging/2012HuntSeatEqBooklet.pdf>.

²⁴ *Id.* at 5.

When evaluating a rider's skill a judge must make an effort not to be overly influenced by the body shape or [attractiveness] of the contestant. Some riders have a build that is more appealing on a horse. A lovely overall appearance cannot and should not be discounted but a judge should also emphasize technical ability.

Id.

any one detail” or “waste time on personal dislikes.”²⁵ In snowboarding, the touchstone is “overall impression,” which is to be shaped by judges’ evaluation of criteria such as amplitude, difficulty, execution, variety, progression, and risk-taking.²⁶

In applying these sorts of standards—which is to say, in assessing the extent to which any given performance measures up to the performative ideal—judges in aesthetic sports draw upon tacit knowledge. The idea is most pithily captured in Michael Polanyi’s phrase, “we can know more than we can tell.”²⁷ It is a matter of degree. While Polanyi insists that “strictly speaking nothing that we know can be said precisely,”²⁸ there are some things as to which our knowledge almost completely outruns our ability to articulate what we know. Polanyi uses the example of riding a bike, which is something most everyone knows how to do, while at the same time, being unable to say much of anything meaningful about how they do it.²⁹ So it is with what he calls “connoisseurship,”—a concept that captures the task of the aesthetic judge.

Although the expert diagnostician, taxonomist and cotton-classer can indicate their clues and formulate their maxims, they know many more things than they can tell, knowing them only in practice, as instrumental particulars, and not explicitly, as objects. The knowledge of such [particulars] is therefore ineffable, and the pondering of a judgment in terms of such particulars is an ineffable process of thought.³⁰

The sort of knowledge involved cannot be, at least not completely, passed along through instruction in rules or maxims. Some undoubtedly can—we are, after all, dealing here with a continuum of articulability—but the core of such knowledge, the part that separates the expert from the novice, can be transferred only through exposure to examples. “By watching the master and emulating his efforts in the presence of his example, the apprentice [unconsciously] picks up

²⁵ *Id.*

²⁶ See INT’L SKI FED’N, JUDGES’ MANUAL SNOWBOARD 10–12 (2013/2014), available at http://www.fis-ski.com/mm/Document/documentlibrary/Snowboard/04/21/07/FISSnow-boardJudgesbook1314_Final_English.pdf.

²⁷ MICHAEL POLANYI, *THE TACIT DIMENSION* 4 (1966).

²⁸ MICHAEL POLANYI, *PERSONAL KNOWLEDGE: TOWARDS A POST-CRITICAL PHILOSOPHY* 87–88 (1962) (citation omitted).

²⁹ *Id.* at 49–50, 88. “To assert that I have knowledge which is ineffable is not to deny that I can speak of it, but only that I can speak of it adequately, the assertion itself being an appraisal of this inadequacy.” *Id.* at 91.

³⁰ *Id.* at 88.

the rules of the art, including those which are not explicitly known to the master himself.”³¹ Put differently, the content of tacit knowledge cannot “be [captured] in context-independent or purely linguistic terms. The articulation of the content requires practical demonstration.”³²

This is a point of significant difference between purposive and aesthetic sports. Because an ideal of the sort implemented in aesthetic sports cannot be fully specified, it cannot exist independent of the officials who are responsible for applying it. As Graham McFee has observed, “[T]he scoring must be achieved by judges who ‘look and see’ *that* such-and-such a move was executed, and *how*; and who know what that is worth in terms of the scoring in the sport.”³³ To a considerable degree, then, judges in aesthetic sport bring a state of affairs into being rather than simply reporting an antecedent state of affairs.

There is more: the nature of the determinations that such judges make, and the content of the knowledge they draw on in doing so are, to a significant degree, products of the manner and circumstances of the judges’ acculturation in the sport. The performative ideal in an aesthetic sport does not exist independently of the judges’ conception of it, and that conception is in turn tied to the ideal as understood in the sport more generally, and more specifically as understood by the preceding cohort of judges. This is a necessary byproduct of the typical manner by which judges in aesthetic sports are developed and promoted.³⁴ Beginner judges must generally have experience as participants in the sport for which they hope to serve as a judge. They must then serve a period of apprenticeship during which they work under the tutelage of experienced judges. Approval and advancement require, among other things, making evaluations that are sufficiently consistent with those of other, longer serving judges. The consequence is that the mechanisms by which officials are selected, trained,

³¹ *Id.* at 53.

³² NEIL GASCOIGNE & TIM THORNTON, TACIT KNOWLEDGE 191 (2013).

³³ Graham McFee, *Officiating in Aesthetic Sports*, 40 J. PHIL. SPORT 1, 3 (2013). McFee acknowledges that officials in purposive sports must perform a similar task. *Id.* at 5. But it is not outcome-determinative in the way that it is in aesthetic sport. *Id.*

³⁴ See generally Cheryl Litman & Thomas Stratmann, *Judging on Thin Ice: Affiliation Bias in Judging Figure Skating* (June 2013) (CESifo Area Conf. on Applied Microeconomics, Working Paper), available at http://www.cesifo-group.de/dms/ifodoc/docs/Akad_Conf/CFP_CONF/CFP_CONF_2014/Conf-am14-Gollier/Papers/am14-Stratmann_19107173_en.pdf. For outlines of the steps necessary to become a judge in figure skating, gymnastics, and equestrian sport, respectively, see U.S. Equestrian Federation, *Procedures for Recorded (r) Hunter, Hunter Seat Equitation, Jumper Judge Licenses, and Registered (R) Hunter Breeding Judge License*, U.S. EQUESTRIAN FEDERATION, <https://www.usef.org/documents/licensedOfficials/LicenseForms/LearnerPermit.pdf> (last visited Dec. 21, 2014); U.S. Figure Skating, *So You Want to be a Figure Skating Judge?*, U.S. FIGURE SKATING, <http://www.usfsa.org/About.asp?id=108> (last visited Dec. 21, 2014); USA Gymnastics, *Judging Women’s Artistic Gymnastics*, USA GYMNASTICS, <https://usagym.org/pages/women/pages/judging.html> (last visited Dec. 21, 2014).

and employed are an important part of what constitutes the sport, in the sense that if the officials were different, the sport would be different.

There are several significant implications of this dynamic. Because aesthetic value cannot be specified as completely or precisely as the goals in a purposive sport, there is great potential for the development of different schools of thought concerning what constitutes the proper or preferred way of skating, vaulting, or riding a horse through a course of jumps. One must learn the content of the ideal and do so by inferring it from concrete examples of performances coming close to the ideal³⁵ because the inability to fully specify the content of the ideal makes it impossible to deduce it from some broader principle. This learning takes place during prospective officials' time as competitors and, more importantly, during their period of apprenticeship, as new officials internalize the preferences of senior officials. As with any such process, the mechanism by which the knowledge of the content of the idealized form of performance is passed down amongst the body of officials leaves open the possibility—more likely, the inevitable reality—that the content of the ideal will shift over time. The passing down of knowledge of the ideal amounts to a diffuse, collective game of “telephone” in which the message shifts over time.

Here, then, is another point of potentially significant contrast between purposive and aesthetic sports. Changes in what “counts” in purposive sports can come about only by way of modifications to the constitutive rules. Football officials have no ability to award points for especially admirable plays that did not quite result in a touchdown or even to shift the nature of how they call a game to reflect an evolving preference for a running game over a passing game. This is not to deny that the latter sorts of preferences exist and evolve within the sport, but they do so as a product of strategic choices by participants rather than through conceptions of officiating.³⁶ In aesthetic sports, in contrast, there is

³⁵ McFee, *supra* note 33, at 13. McFee continues:

Moreover . . . within the sport, the rules prioritise *this* over *that* – they award more marks for it. And judges *learn* this appreciation. Hence *this* is more valuable in pairs (more *aesthetic*, in the context of pairs) than *that*; and the reverse for ice-dance. Of course, such priorities can be reversed (say, by rule changes, perhaps reflecting changes in discrimination).

Id. at 14.

³⁶ It is of course true that governing authorities sometimes instruct officials to change their emphasis in how they call certain infractions—such as what should constitute pass interference, whether the strike zone in practice should reflect the strike zone as written, and so forth—but these are more akin to changes in the regulatory rules rather than the evolution of an aesthetic ideal.

plenty of room for a common law-like³⁷ evolution of the ideal over time. Because this evolution is a product of what the officials value, there is considerably less room for participants to drive the evolution of the point through their strategic choices. Put differently, while participants in aesthetic sports are free to make strategic decisions, they do so on the understanding that those decisions will pay off only if they play to the officials' conception of the ideal.

As with much in the philosophy of sport, these cleavages are not absolute. A purposive sport, like baseball, can accommodate similar sorts of evolution, though on a much more limited scale. For example, during the era when the American League (AL) and the National League (NL) had separate sets of umpires, it was generally understood that there was an AL strike zone and an NL strike zone,³⁸ neither of which conformed to the strike zone as described in the rules.³⁹ While there is thus a sense in which the rules of baseball were informally amended by the officials, Russell's point remains: The strike zone as it was implemented in practice could legitimately be said to be deviant when measured against the strike zone as defined by the rules.

One could make the case that the same holds with respect to aesthetic sports, in the sense that there is, in theory, an objectively ascertainable, correct conception of the idealized performance against which any official's scoring could be measured for correctness. On this view, phrases like "flow and effortless glide" and "artistry" specify a definite, discoverable ideal and standard for measurement. To be sure, uncovering the content of that ideal would require something of a Herculean (in both its informal and Dworkinian senses) effort involving the distillation of the views of the entire universe of judges in at least two respects—as to the nature of the performative ideal and as to how this particular performance ought to be scored given that ideal. What is more, performing that task would require answering a host of ancillary questions, such as whether the distillation is to be of the views of all qualified judges or only some subset; whether it is to be of their views on their best day versus an average day; and so on. It may ultimately not even be theoretically possible to perform such an inquiry due to the large number of contestable issues and sub-issues involved.

³⁷ Though different in significant respects, as discussed below. See *infra* notes 38–39.

³⁸ See, e.g., DURWOOD MERRILL & JIM DENT, YOU'RE OUT AND YOU'RE UGLY, TOO!: CONFESSIONS OF AN UMPIRE WITH AN ATTITUDE 10 (1998) ("[M]y esteemed [brethren] over in the N.L. have for decades had a larger strike zone than the umps from the A.L."); Murray Chass, *In the Early Returns, American League Hitters Have the Upper Hand*, N.Y. TIMES (Apr. 13, 1997), <http://www.nytimes.com/1997/04/13/sports/in-the-early-returns-american-league-hitters-have-the-upper-hand.html> (discussing differences in the leagues' strike zones).

³⁹ See, e.g., DAN FORMOSA & PAUL HAMBURGER, BASEBALL FIELD GUIDE: AN IN-DEPTH ILLUSTRATED GUIDE TO THE COMPLETE RULES OF BASEBALL 86 (2008); MERRILL & DENT, *supra* note 38, at 201–10.

Regardless of whether that sort of inquiry is possible, the role of tacit knowledge in the constitution of the aesthetic ideal marks a clear distinction between aesthetic and purposive sport. This is, again, because the very nature of aesthetic sports is such that the content of the aesthetic ideal, as tacit knowledge, is not fully articulable in a context-independent way. If it were, the rules of the sport would include such an articulation rather than relying on the sort of general terminology referenced above, and scoring decisions could be evaluated as right or wrong by reference to those rules rather than to some collective, intangible conceptualization. At that point, our aesthetic sport would cease to be an aesthetic sport. Take the example of figure skating: We might imagine rules that specify that the appropriate way to do a certain jump is to achieve a certain rotation speed while rising a certain distance off the ice, with scoring calibrated to nearness to that ideal and with imprecise phrases like *flow* and *effortless glide* removed from the calculus. Having thus specified the components of scoring, everything not enumerated would become part of the strategic calculus—how fast to skate into the jump, arm position, and so on. The goal—performing the jump according to the specified criteria—would be what matters, while the means of achieving it would not. This would be true even under a definition of the goal that took more of the components of a jump into account. The result would be a scoring system that removed its core of tacit knowledge and that could likely be performed by a mechanical system.

We do not see that sort of reductionism in definition, of course, because doing so would leave out what seems to be the most important things. Whatever a standard like *“flow and effortless glide”* includes, it involves more than is measureable in ways that can be reduced to statements in a rule book, and it is something that cannot be fully appreciated by everyone but rather only by those who through some combination of talent and experience have developed appropriately discriminating senses. Its application requires the use of tacit knowledge.

As observers of aesthetic sport recognize, this reliance on tacit knowledge brings with it an associated danger. Phrases like *“flow and effortless glide”* not only stand in for things that we value but cannot more fully articulate, they also allow space for the operation of influences that we do not value but cannot fully separate out. They provide an opportunity, in other words, for biases and prejudices to impact decision making. As we will see in the next part, this danger is real, and it presents a set of subsidiary difficulties because the tacit nature of performative ideals not only makes them impossible to specify in context-independent ways, it also thwarts our ability to identify where the performative ideal ends and bias begins.

III. PERSPECTIVES FROM PSYCHOLOGY

A growing body of research considers the behavior of sports officials and how it deviates from the ideal. For example, a recent study found that umpires in Major League Baseball tended to expand the strike zone in favor of the home team, were influenced by pitch count, and performed more poorly when the game was on the line.⁴⁰ The race of the pitcher also affected whether pitches properly considered balls were called as strikes, though not vice-versa.⁴¹ Being an All-Star helped pitchers get favorable calls, especially for those pitchers with a reputation for having good control.⁴² All of this, of course, happened despite the fact that these umpires work in an environment in which every ball and strike call can almost instantaneously be evaluated for accuracy.

Unsurprisingly, much of the work on sports officiating has examined the behavior of judges in aesthetic sport. Researcher Diane Ste-Marie, considering the role of expertise in sports officiating, makes a distinction between open-skill and closed-skill sports that provides yet another cut at the divide between purposive and aesthetic sports. The role of the official in these two contexts differs in significant ways beyond Suits's basic categories of judges and law-enforcement officials.⁴³ The judge in the closed-skill, subjectively judged aesthetic sport typically remains in a fixed spot relative to the action, and that spot is usually a vantage point outside the field of competition. What is more, the performance environment remains constant for each performer—the gymnastics equipment remains in the same place at the same heights, the ice is periodically resurfaced, and so forth. In contrast, the officials in open-skilled, purposive sports (now more likely to be called referees) typically do their work on the field of play, which they constantly move about so as to follow the action and position themselves in the best vantage point from which to police for rule violations. Perhaps the most significant difference relates to the nature of the decisions being made. Here again, Ste-Marie's distinction closely tracks that of the philosophers. She observes that judges in closed-skill sports "typically analyze

⁴⁰ For a general discussion of this research, see Brayden King & Jerry Kim, *What Umpires Get Wrong*, N.Y. TIMES (Mar. 28, 2014), <http://www.nytimes.com/2014/03/30/opinion/sunday/what-umpires-get-wrong.html>.

⁴¹ *Id.*

⁴² *Id.* See also generally Jerry W. Kim & Brayden G. King, *Seeing Stars: Matthew Effects and Status Bias in Major League Baseball Umpiring*, MGMT. SCI. (May 5, 2014), available at <http://mors.haas.berkeley.edu/papers/Kim%20and%20King%20-%20Matthew%20Effects.pdf>.

⁴³ Diane Ste-Marie, *Expertise in Sport Judges and Referees: Circumventing Information-Processing Limitations*, in EXPERT PERFORMANCE IN SPORTS: ADVANCES IN RESEARCH ON SPORT EXPERTISE 169, 176 (Janet L. Starkes & K. Anders Ericsson eds., 2003).

whether the motor skills that are performed match with a ‘perfect’ template for the skill in question; thus, the quality of the movement becomes a key criteria.”⁴⁴ Officials in open-skill sports, in contrast, “focus on whether the movements fit within given boundaries (e.g., the hit is not too low) and whether players adhere to the rules of the game.”⁴⁵

The assessments in both contexts are complex, and the task is challenging. Officials in both types of sport must continually draw on their accumulated knowledge. “For example, within gymnastics, judges need to store information concerning the symbol code, level of difficulty, and deduction values for associated errors. Similarly, referees need to store information concerning the rules of the game, consequences of given behaviors, and the various hand signals [attached] to decisions made.”⁴⁶ This, of course, must occur at the same time as the official is devoting attention to the ongoing play or performance.

Staying on top of all this information is extraordinarily difficult and perhaps even impossible. As Clare MacMahon and Bill Mildenhall characterize the literature, “Researchers have pointed out that the perceptual-cognitive tasks in officiating create demands that surpass human information processing limits.”⁴⁷ Experienced officials ameliorate the effects of these deficits through the development of an expertise. They are able to more efficiently draw inferences from and otherwise attach significance to a broader array of inputs as compared to the novice.

It is worth pausing to consider the nature of the expertise that experienced officials deploy, which opens another window into the distinction between purposive and aesthetic sports. The insight arises from the simple fact that systematic utilization of expertise creates a need to determine who counts as an expert. For rule-driven activities that generate easily measured results, the task of identifying experts is sometimes no more difficult than finding the people who are most accomplished at the task. Expert chess players are those who win the most games. In similar fashion, the most expert officials in purposive sports are those who best meet Russell’s test of accurately determining whether the antecedent, rule-triggering state of affairs exists. Being the best at calling balls and strikes may not make one the best umpire, but it is certainly a large part of the mix. Of course, not all purposive sports provide such easy metrics. The significance of one individual’s contribution to a team effort can be hard to quantify, and so determining who, for example, is the best linebacker or small forward requires

⁴⁴ *Id.* at 177.

⁴⁵ *Id.*

⁴⁶ *Id.*

⁴⁷ Clare MacMahon & Bill Mildenhall, *A Practical Perspective on Decision Making Influences in Sports Officiating*, 7 INT’L J. SPORTS SCI. & COACHING 153, 154 (2012) (citations omitted).

the application of judgment. And even where there are clear bases for individual comparison, we might resist expertise, as opposed to raw physical ability, as the appropriate label for what distinguishes one player from another. Greg Maddux seems more amenable to description as an expert than does Nolan Ryan. The point remains, though, that however we choose to account for (or discount for) natural talent, there are some endeavors for which the ability to measure results makes it relatively easy to determine who counts as an expert. This holds for officials as well as participants.

In aesthetic sports, it is also possible to assess the expertise of athletes and coaches by reference to results. But, as we have seen, those results can never be measured independently of the determinations of judges. Aesthetic sports are, through and through, the sort of domain in which the only way to determine who counts as an expert is to ask another expert.⁴⁸ The best figure skaters are those who the judges say are the best. The same is true of the best gymnasts, snowboarders, and equestrians. It is also true of judges, who must be experts in both the sport and the task of judging the sport. But their own status as experts is also a product of prior judges' determinations because the measure of whether one is appropriately expert to become a judge is whether one's assessments are sufficiently convergent with those of other experts. Indeed, as we have seen, this process is incorporated into the selection mechanisms, as aspiring judges must go through a period of apprenticeship and can advance through the ranks only by demonstrating that they have internalized the norms of the existing set of experts.

Whatever its ontology, expertise in the officiating context involves the utilization of strategies to help maximize one's ability to take in the vast amount of information presented.⁴⁹ Consistent with this, studies have demonstrated that expert judges outperform novices in detecting subtle variations in movement and position, due in part to the experts' tendency to use different and more effective visual search strategies, including placing greater reliance on peripheral vision.⁵⁰ Another aspect of experts' superiority stems from their ability to fill in gaps in information. Because it is impossible for humans to fully attend to

⁴⁸ "In some domains there are objective criteria for finding experts, who are consistently able to exhibit superior performance for representative tasks in a domain In some domains it is difficult for non-experts to identify experts, and consequently researchers rely on peer-nominations by professionals in the same domain." K. Anders Ericsson, *Introduction to THE CAMBRIDGE HANDBOOK OF EXPERTISE AND EXPERT PERFORMANCE* 3, 3–4 (K. Anders Ericsson et al. eds., 2006).

⁴⁹ "At best, judges know how to allocate their attention. For instance, expert judges in gymnastics have been shown to differ from novices in their visual search strategies By and large, this research shows that expert judges in sports develop effective anticipatory strategies that help to improve their decision making" Henning Plessner & Thomas Haar, *Sports Performance Judgments from a Social Cognitive Perspective*, 7 *PSYCH. SPORT & EXERCISE* 555, 560 (2006).

⁵⁰ Ste-Marie, *supra* note 43, at 178–79.

all the information necessary to make a truly comprehensive assessment of game play or a performance (e.g., it is impossible for a single person to direct his visual focus to multiple players, or to the hands and feet of a single player, at the same time), officials are left to work with less than complete information. They accordingly rely on pattern recognition and various other kinds of short cuts. “For example, if asked to judge whether a player is offside, not having viewed the player’s entire movements, a football referee may use the availability heuristic, to search her memory for any experience of a player at that level of play moving with such speed from an onside position.”⁵¹ These sorts of strategies often become part of the formal and informal tools and “tricks of the trade” that are passed along via training and apprenticeship.⁵² Significantly, some research suggests that it is experience as an official, rather than simply experience as a participant in the sport, that is critical to the development of these heuristics.⁵³

Unsurprisingly, the need to fill in gaps in information can also lead to perceived distortions in the officiating process. Research has demonstrated that officials’ determinations are skewed from the ideal, even in situations involving the application of clear rules in purposive sports. For example, as discussed above,⁵⁴ umpires in Major League Baseball have exhibited systematic deviations in their calling of balls and strikes.

Accusations of bias and other improper influences are even more prevalent in aesthetic sport. Most prominent, but perhaps least interesting, are instances in which judges have been revealed to have engaged in (or were at least strongly suspected of) outright corruption, in the sense that they intentionally overlooked the merits of the judged performances to reward participants based on other considerations, most often some version of nationalism. Indeed, as noted in the *New York Times*, “Rarely does a Winter Olympics pass without something happening that invites intense scrutiny of the figure skating judges.”⁵⁵ The 2002 Winter Olympics featured an instance of outright collusion,⁵⁶ and there were suspicions in 2014.⁵⁷

⁵¹ MacMahon & Mildenhall, *supra* note 47, at 160.

⁵² *Id.* at 161.

⁵³ *Id.* at 162 (discussing Fabrice Dosseville et al., *Contextual and Personal Motor Experience Effects in Judo Referees’ Decisions*, 25 SPORT PSYCHOLOGIST 67, 67 (2011)).

⁵⁴ See King & Kim, *supra* notes 40–42 and accompanying text.

⁵⁵ Mary Pilon, *Who Were the Figure Skating Judges?*, N.Y. TIMES (Feb. 20, 2014), <http://www.nytimes.com/2014/02/21/sports/olympics/who-were-the-figure-skating-judges.html>.

⁵⁶ See Christopher Clarey, *FIGURE SKATING; 2 French Officials Suspended 3 Years in Skating Scandal*, N.Y. TIMES (May 1, 2002), <http://www.nytimes.com/2002/05/01/sports/figure-skating-2-french-officials-suspended-3-years-in-skating-scandal.html>.

⁵⁷ See Pilon, *supra* note 55; Mary Pilon & Jeré Longman, *Despite Revamp, Figure Skating Gets*

More interesting are situations in which judges are potentially subject to unconscious influences. A personal example: I am the parent of daughters who participate in equestrian sports. Most classes within a horse show involve sequential competition, and as a result, participants wonder and have theories about whether and to what extent it matters if one rides at the beginning, middle, or end of a class. I can further report that there is a strong sense within the equestrian community that leaving a good or bad impression on a judge in an early stage of competition will work to one's benefit or detriment as things progress. And even more generally, there is a sense of arbitrariness, that one can have the ride of one's life and finish fourth, and the next time out, have a mediocre ride and win.

Research supports the conclusion that there is merit to these intuitions, and that a host of unconscious influences manifest themselves in judges' scoring. One of these is the *affiliation bias*. Judges are likely to react more favorably to performances by those with whom they share an affiliation, broadly defined to include factors like race, gender, and nationality.⁵⁸ Consider, for example, a recent study of figure skating judges undertaken by Cheryl Litman and Thomas Stratmann.⁵⁹ They examined whether skating judges give more favorable scores to competitors from their home skating club. Judges, like skaters, are required to be members of a skating club, and skaters' club affiliations are announced as they take the ice. Litman and Stratmann found that judges did in fact give higher scores to skaters from their home club, and that this effect held "for singles skating and synchronized skating, senior and non-senior levels of skating ability, and for national, sectional and regional events."⁶⁰ Perhaps significantly, they found that the effect disappeared when events were televised. "The finding suggests that judges either intentionally assign higher marks for skaters from one's own club when there is no TV coverage, or, that under the glare of TV cameras try to think harder as to how to score fairly."⁶¹

Judges also manifest the related phenomena of *reputation bias* and the *halo effect*. The basic dynamic in both is that prior information about the skills of a

Mixed Marks, N.Y. TIMES (Feb. 5, 2014), <http://www.nytimes.com/2014/02/06/sports/olympics/despite-revamped-system-for-judging-figure-skating-gets-mixed-marks.html>.

⁵⁸ Litman & Stratmann, *supra* note 34, at 12–13. See also John W. Emerson, Miki Seltzer & David Lin, *Assessing Judging Bias: An Example from the 2000 Olympic Games*, 63 AM. STATISTICIAN 124, 124 (2009); Robert Whissell & Scott Lyons, *National Bias in Judgments of Olympic-Level Skating*, 77 PERCEPTUAL & MOTOR SKILLS 355, 355 (1993); Eric Zitzewitz, *Nationalism in Winter Sports Judging and Its Lessons for Organizational Decision Making* (October 2002) (Stanford GSB Research Paper No. 1796), available at http://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID319801_code020809670.pdf?abstractid=319801&mirid=2.

⁵⁹ See generally Litman & Stratmann, *supra* note 34.

⁶⁰ *Id.* at 4.

⁶¹ *Id.*

particular competitor, based in either the knowledge of the reputation of the athlete or based on having witnessed a specific prior performance of the athlete, leads a judge to score subsequent performances more (or less) favorably than would be the case without the prior information. For example, Findlay and Ste-Marie had two groups of figure skating judges score performances by a group of skaters made up of some who were known to them to have a good reputation and some who were from the other judges' geographical area and thus unknown to them.⁶² They found that "[s]katers' ordinal rankings, which are used to determine their final placement in competition, were better when skaters were evaluated by judges who knew of their positive reputation versus when they were evaluated by judges who did not recognize their name."⁶³ Iain Greenlees and his colleagues showed videos of soccer players performing a passing task, with the order of the quality of execution manipulated so that for some viewers, the players started strongly and made mistakes later, while for others, the order was reversed.⁶⁴ They found that those who started strongly were judged as more skillful than those who started poorly, such that athletes "may bias a judge into giving more favourable scores on the basis of their initial performances."⁶⁵

There are also biases that manifest themselves in unique ways in the context of aesthetic sports due to the fact that competitors in such sports often compete in a sequence, rather than at the same time, and are scored after each athlete has competed rather than after all athletes have competed. One is the product of an *order effect*, pursuant to which contestants who perform later in the sequence of competitors tend to receive higher scores than those who perform earlier even in situations in which order is randomly assigned rather than being a product of a score in a prior round or some other measures of competitive ability.⁶⁶ This may be a product of judges' understandable tendency to refrain from giving high scores early, so as to leave room for potentially better performances to come, or it may be that later competitors (assuming they know their predecessors' scores) actually turn in better performances due to heightened motivation. A third possible explanation is that judges in such an environment can make comparisons

⁶² Leanne C. Findlay & Diane M. Ste-Marie, *A Reputation Bias in Figure Skating Judging*, 26 J. SPORT & EXERCISE PSYCH. 154, 158–60 (2004).

⁶³ *Id.* at 163.

⁶⁴ Iain Greenlees et al., *Order Effects in Sport: Examining the Impact of Order of Information Presentation on Attributions of Ability*, 8 PSYCH. SPORT & EXERCISE 477, 481–82 (2007).

⁶⁵ *Id.* at 485.

⁶⁶ E.g., Wändi Bruine de Bruin, *Save the Last Dance II: Unwanted Serial Position Effects in Figure Skating Judgments*, 123 ACTA PSYCHOLOGICA 299, 307–08 (2006) ("Despite being randomly assigned to their starting number in the initial round, figure skaters who perform later receive better scores in the first round, and in the second round, in which figure skaters with better scores in the first round are invited to skate later . . .").

in only one direction (i.e., they are able to evaluate any given performance relatively only to those that have come before and not relative to those yet to come) and tend to give outsized influence to ways in which a new performance is unique compared to those that came before.⁶⁷

A second bias that arises in serially judged competitions involves *conformity effects*. This arises in situations where there is a panel of judges and refers to the tendency of individual judges to adjust their scoring to be in greater conformity with the scores of their fellow panelists.⁶⁸ Thus, a panel of judges will evolve toward greater consensus in its evaluations over the course of a competition. The influence here may have two sources:

Social-psychological research has identified two basic reasons for conformity: informational and normative influence Informational influence implies that people conform to the group norm because they want to make a correct judgment and because they are more certain about the judgment of others than about their own judgment. Normative influence implies that people conform to the group norm because they want to make a good impression on others or because they fear to be rejected by others when their judgment stands out negatively.⁶⁹

Whether this effect is properly characterized as a bias or not is debatable. On the one hand, the effect has real implications for competitors, since those who perform later in a competition will, in a very real sense, be performing before a differently oriented panel than their predecessors were. On the other hand, striving toward consensus is at the very core of judging in aesthetic sports, where consensus with other experts provides the only metric by which inexperienced and prospective judges are assessed.

This last point highlights an unstated assumption underlying this research, namely that the impact of these unconscious influences is a bad thing. The intuition behind that assumption is easy to understand. Competitors, one imagines, regard the bargain as being that they will be judged on the specific performance they turn in on a given occasion, and that nothing else should count. The mediocre athlete having a great day should triumph over the great athlete having a mediocre day, and more generally, a better performance should receive a

⁶⁷ *Id.* at 300, 308.

⁶⁸ See Filip Boen et al., *The Impact of Open Feedback on Conformity Among Judges in Rope Skipping*, 7 PSYCH. SPORT & EXERCISE 577, 578–79 (2006) (discussing prior research on conformity effects).

⁶⁹ *Id.* at 580.

higher score than a lesser one.

That it is self-evidently true that better performances should prevail, however, does not necessarily mean that we should want to close officials off from all information not specific to the performance they are judging. As MacMahon and Mildenhall point out, there is a danger associated with screening judges off from all information that has been associated with bias in the research outlined above. Because their cognitive limitations render judges unable to process and evaluate all of the information relevant to assessing a given performance, there will necessarily be gaps in their knowledge. As a consequence, “[r]emoval of the supposedly undesirable information . . . would not decrease the difficulty of the task, nor prevent the judge from consciously or subconsciously searching for other information to compensate for the now even bigger information gaps.”⁷⁰ Given this, it might at least sometimes be better for judges to fill in the gaps with the sort of information portrayed as leading to bias, on the grounds that doing so is preferable to filling those gaps with less appropriate information or not addressing them at all.

This logic certainly cannot excuse all unconscious influences, but one could tell a story in which it is appropriate for a judge to take an athlete’s reputation into account. The judge’s goal is to identify the athlete whose performance most closely approximates the ideal. That ideal is a product of a consensus within the relevant sporting community. So, too, is an athlete’s reputation a product of that same community or a subset of it. In cases of doubt, then, one could argue that reliance on reputation as a tiebreaker is consistent with the judge’s task. Of course, even if one accepts that argument, there remains the difficulty, which we will explore below, of limiting the use of the information so that it is a supplement to, rather than a substitute for, the performance itself.

IV. JUDGES AS JUDGES: THE FORMS AND LIMITS OF AN ANALOGY

Perhaps the primary objection to the “judge as umpire” analogy is that it fails to acknowledge that judges play an active role in making the rules they are to apply. Aesthetic sports may provide a better, if perhaps less alluring, analogy. The notion of a broodingly omnipresent ideal performance fits just as well at first blush, and ultimately falls prey to the same sorts of critiques, as the brooding omnipresence of the common law.⁷¹ Both sorts of judges shape rather than

⁷⁰ MacMahon & Mildenhall, *supra* note 47, at 155.

⁷¹ See *Southern Pac. Co. v. Jensen*, 244 U.S. 205, 222 (1917) (Holmes, J., dissenting) (“The common law is not a brooding omnipresence in the sky but the articulate voice of some sovereign or [quasi-sovereign] that can be identified; although some decisions with which I have disagreed seem to me to

simply discover the standards by which they judge. And although the specific mechanisms differ—most obviously in that the aesthetic judge does not generate written opinions with precedential effect—both do so by resorting to and shaping the conventions on which those standards are based. Judges in aesthetic sport and common-law judges alike draw on their sense of a collectively held ideal they are all striving to attain.

Complaints and concerns about judging in aesthetic sport likewise mirror those directed at the judicial system. At the heart of critiques of judging in aesthetic sport is a concern that is often voiced in terms of the perceived “subjectivity” of the process, a term that is typically used as a shorthand for a sense that decisions are unduly influenced by extraneous factors and thus arbitrary. So, too, with law, observers are tempted to conclude that politics, in some form, is the true driver of judicial decisions.⁷² In both instances the ultimate concern is with a perceived lack of what Karl Llewellyn called “reckonability.”⁷³ Llewellyn never offered a precise definition of the phrase, but it is clear that he had in mind a concept akin to “predictability.” He spoke of whether there is “any real stability of footing for the lawyer, be it in appellate litigation or in counseling, whether therefore there is any effective craftsmanship for him to bring to bear to serve his client and to justify his being.”⁷⁴ He sketched it as a middle ground between an outcome “foredoomed in logic” and “the product of uncontrolled will which is as good as wayward.”⁷⁵ “The true ideal is *reasonable regularity* of decision.”⁷⁶

One reason these sorts of concerns ring true to so many observers in both domains is that judges in the two settings rely heavily on tacit knowledge in reaching their decisions. This is considerably more obvious in the case of the aesthetic judge, the core of whose job is to apply a standard that will be, by its very nature, one that defies precise definition. The law judge, in contrast, is engaged in an activity that celebrates the offering of reasoned—which is to say articulated in terms of explicit knowledge—justifications. Legal standards, we are tempted to imagine, are by their very nature subject to definition. On this superficial view, law and aesthetic sport could not be more different. To find the parallel, we must look deeper, to what stands behind these written justifications and to what they cannot say. We must recognize that words cannot do all

have forgotten the fact.”).

⁷² For a recent overview of this and related debates, see Lawrence B. Solum, Book Review, *The Positive Foundations of Formalism: False Necessity and American Legal Realism*, 127 HARV. L. REV. 2464, 2464–65 (2014).

⁷³ KARL N. LLEWELLYN, *THE COMMON LAW TRADITION: DECIDING APPEALS* 3–7 (1960).

⁷⁴ *Id.* at 3–4.

⁷⁵ *Id.* at 4. See also *id.* at 17–18.

⁷⁶ *Id.* at 216.

the work we in law ask of them,⁷⁷ and that legal standards are underdeterminate.⁷⁸

None of this is to discount the existence of easy cases. They are present in law, and they are present in aesthetic sport, including not only situations such as where the gymnast falls mid-routine, but also those in which even a moderately knowledgeable lay observer could make the requisite quality distinctions. But it is important to recognize that the judge facing the more difficult case has options. These range from choices with effects at the broad level of who wins and who loses to those with implications at narrower levels, such as what the specific grounds for decision are, as well as whether they are formulated broadly or narrowly as a rule or a standard, and so forth. It is here that what guides the judge is less precise, more the product of norms of practice, “situation sense,” or tacit knowledge. Michael Polanyi expressly drew the connection between tacit knowledge and judging in the context of the common law and its reliance on precedent, which, he contended, “recognizes the principle of all traditionalism that practical wisdom is more truly embodied in action than expressed in rules of action.”⁷⁹

Polanyi is not alone in that regard. In his commencement address at Yale Law School in 2006, Dan Kahan likened the training students receive at Yale to that of a professional chick sexer.⁸⁰ Chick sexers are a frequently used example of tacit knowledge. It turns out to be important to be able to determine the sex of a baby chick as soon as possible after they hatch. It also turns out to be something that is not easily done, for there are no evident markers of chick gender. And yet, those with a sufficient period of training, in which they get instant feedback on their identifications from expert chicken sexers, develop the ability to identify the sex of a chick to a rate of accuracy greater than ninety-five percent.⁸¹ But they will not be able to articulate how they did it.

Kahan suggests that it is the same with law and legal training:

⁷⁷ Chad M. Oldfather, *Writing, Cognition, and the Nature of the Judicial Function*, 96 GEO. L.J. 1283, 1285 (2008).

⁷⁸ See Lawrence B. Solum, *On the Indeterminacy Crisis: Critiquing Critical Dogma*, 54 U. CHI. L. REV. 462, 473 (1987) (“The law is *underdeterminate* with respect to a given case if and only if the set of results in the case that can be squared with the legal materials is a nonidentical subset of all imaginable results.”).

⁷⁹ POLANYI, *supra* note 28, at 54.

⁸⁰ See Dan M. Kahan, Yale Law School Commencement Remarks, Address at the Yale University Law School Commencement, *available at* <http://www.law.yale.edu/documents/pdf/kahanREVISED.pdf> (last amended May 25, 2006).

⁸¹ See, e.g., *id.*; See also Richard Horsey, *The Art of Chicken Sexing* (2002), *available at* <http://www.phon.ucl.ac.uk/publications/WPL/02papers/horsey.pdf>.

The formal doctrines and rules that make up the law – unconscionability, proximate causation, character propensity, unreasonable restraints of trade – are just as fuzzy and indeterminate as the genitalia [sic] of day-old chicks. And yet just as the trained chick sexer can accurately distinguish female from male, so the trained lawyer can accurately distinguish good decision from bad, persuasive argument from weak.⁸²

And the lawyer, too—who has likewise attained a great deal of her training through exposure to the habits and reactions of lawyers, judges, and professors with more experience—will be unable to provide a complete description of how she makes those distinctions.

The point is illustrated by an anecdote passed along by Karl Llewellyn, who wrote of asking seven of his fellow legal academics who had been appointed to the bench to put in writing their decision-making processes during their first year as judges.⁸³ All seven agreed to do so; none actually followed through.⁸⁴ Judge Henry Friendly's reaction to this was to offer, as a partial explanation, the fact "that the new judge soon learns that each judge judges [differently] from every other judge and that any one judge judges differently in each case."⁸⁵ Friendly's answer sheds some additional light but ultimately dodges the question. When we appreciate the role of tacit knowledge in judging, we come to understand that judges judge differently from one case to the next because the professional norms that they have absorbed instruct them that it is right to do so. They may not be able to articulate fully why it is that one case is different from the next, but they sense that it is so and that it is appropriate for them to act accordingly.

This idea is one that Llewellyn referred to using phrases like "situation sense" and "horse sense," and by invoking the conception of appellate judging as a craft. Thus,

the rules not only fail to tell the full tale, taken literally they tell much of it wrong; and while words can set forth such facts and needs as ideals, craft-conscience, and morale, these things are bodied forth, they live and work, primarily in ways and attitudes which are much more and better felt and done than they

⁸² Kahan, *supra* note 80.

⁸³ LLEWELLYN, *supra* note 73, at 264–65.

⁸⁴ *Id.* at 265.

⁸⁵ Henry J. Friendly, *Reactions of a Lawyer—Newly Become Judge*, 71 YALE L.J. 218, 229 (1961).

are said.⁸⁶

A danger attends subjecting such a practice to “self-conscious intellectual analysis [T]he problem goes to whether articulate [principles] or rules for doing, phrasings for the inculcation or [transmission] of knowhow, will not cripple or kill, rather than further and better the doing of the job.”⁸⁷

Llewellyn is hardly alone in taking this position. Anthony Kronman wrote of law as “a craft demanding a cultivated subtlety of judgment whose possession constitutes a valuable trait of character,”⁸⁸ and of the judge who “will be guided in his deliberations by what might be called the *ethos* of his office, by a certain ideal of judicial craftsmanship, and by the habits that a devotion to this ideal and long experience in [attempting] to achieve it tend to instill.”⁸⁹ More recently, Dan Kahan has called for renewed attention to “the systematic study of the profession's situation sense, of the cognitive mechanisms through which it operates, and of its interaction with broader social and political dynamics.”⁹⁰

As Brett Scharffs notes, there is something quaint and anachronistic-seeming in giving this sort of account of judging.⁹¹ Ours is a world that tends to discount that which cannot be easily identified,⁹² and one in which judges were long ago deprived of the ability to engage in the sort of leisurely consideration of their cases that the notion of judge-as-craftsperson suggests.⁹³ Yet while the circumstances in which judges operate have changed, the need to rely on tacit knowledge to do the job has not. Indeed, the increased time pressures that most judges feel have probably led to a situation in which that reliance plays a greater role.

My point, then, is not to assert that judges ought to regard themselves as engaged in a craft. (Though I also do not mean to assert that there is anything unattractive about such an idea.) It is, instead, to underscore the extent to which judges must apply tacit knowledge in a context in which they face increasing cognitive demands. In this sense, they have become even more like aesthetic

⁸⁶ LLEWELLYN, *supra* note 73, at 214.

⁸⁷ *Id.* at 264.

⁸⁸ ANTHONY T. KRONMAN, *THE LOST LAWYER: FAILING IDEALS OF THE LEGAL PROFESSION* 295 (1993).

⁸⁹ *Id.* at 214.

⁹⁰ Dan M. Kahan, *Foreword: Neutral Principles, Motivated Cognition, and Some Problems for Constitutional Law*, 125 HARV. L. REV. 1, 77 (2011).

⁹¹ Brett G. Scharffs, *Law as Craft*, 54 VAND. L. REV. 2245, 2250 (2001).

⁹² *Id.* at 2287 (outlining the ways in which the legal culture has drifted away from the values of craft and toward the values of theory).

⁹³ See, e.g., William M. Richman & William L. Reynolds, *Elitism, Expediency, and the New Certiorari: Requiem for the Learned Hand Tradition*, 81 CORNELL L. REV. 273, 281–86 (1996).

judges, charged with the task of drawing on a set of partly inarticulable norms and ideals as they face a steady stream of cases they must judge with little time for reflection.

As the psychological research outlined in Part II reveals, there is a potential downside to heavy reliance on tacit knowledge, which is that it can easily serve to hide, if not more actively facilitate, the influence of biases.⁹⁴ This is a concern in the context of legal judging, as well.⁹⁵ There is a line—which exists in both contexts—that separates out the things that ought to count in the analysis from those that ought not. What is more, while most everyone will agree that such a line exists, they will disagree about many of the particulars of its location. In law, this will be true in the sense that there will be disagreement about the outcome of cases or the substance of legal standards. It will also be true as to methodology. As Adrian Vermeule puts it, “The history of interpretive theory in American courts is, above all, a history of persistent and deep disagreements among judges and courts about the proper methods and sources of legal interpretation.”⁹⁶ So, too, we can appreciate the existence of “camps” within the judging of aesthetic sport, as different judges have different ideas about what to look for and perhaps even how to look for it. There are shared assumptions in both contexts, and the shared space is by most measures undoubtedly larger than the area of disagreement. But disagreement is unavoidable, and to some degree desirable.

The point has become especially salient over the last several decades, which have witnessed a steady accumulation of evidence that judicial decision making is correlated with, if not influenced by, factors that are not properly “legal” in the sense that we usually think of that term. The most prominent of this work is, of course, that demonstrating the connection between political ideology and decision-making. Even when measured in a very reductionist way—with judicial ideology determined by the party of the president who appointed a judge

⁹⁴ Scharffs, *supra* note 91, at 2286 (“One danger associated with ‘tacit’ knowledge is that what we ‘know’ might only mask our prejudices.”).

⁹⁵ Indeed, in a series of posts on Prawfsblawg, Jordan Singer used Findlay & Ste-Marie’s study of reputation bias in figure skating, see generally Findlay & Ste-Marie, *supra* note 62, as a jumping off point for consideration of such an effect in the legal system. See Jordan Singer, *Judgment Calls and Reputation, Part One: Figure Skating*, PRAWFSBLAWG (Feb. 4, 2014, 11:03 AM), <http://prawfsblawg.blogs.com/prawfsblawg/2014/02/judgment-calls-and-reputation-part-one-figure-skating.html#more>; Jordan Singer, *Judgment Calls and Reputation, Part Two: Trial Judges*, PRAWFSBLAWG (Feb. 13, 2014, 3:43 PM), <http://prawfsblawg.blogs.com/prawfsblawg/2014/02/judgment-calls-and-reputation-part-two-trial-judges.html>; Jordan Singer, *Judgment Calls and Reputation, Part Three: Attorneys*, PRAWFSBLAWG (Feb. 24, 2014, 11:31 AM), <http://prawfsblawg.blogs.com/prawfsblawg/2014/02/judgment-calls-and-reputation-part-three-attorneys.html>.

⁹⁶ Adrian Vermeule, *The Judiciary is a They, Not an It: Interpretive Theory and the Fallacy of Division*, 14 J. CONTEMP. LEGAL ISSUES 549, 556 (2004).

and decision ideology determined by the identity of the winning party in a case—research has consistently revealed that the votes of “conservative” and “liberal” judges tend to skew in predictable directions.⁹⁷

The dynamic here, too, resembles that in the judging of aesthetic sport, in that what is virtuous in moderation becomes vice when taken too far. Just as there are senses in which it is arguably rational and consistent with the aim of the sport for judges in aesthetic sport to draw on information outside the confines of the specific performance they are observing and required to judge,⁹⁸ so too might it be appropriate for judges in law to do the same. Consider the following observation from Judge Posner:

Because the materials of legalist decision making fail to generate acceptable answers to all the legal questions that American judges are required to decide, judges perforce have occasional—indeed rather frequent—recourse to other sources of judgment, including their own political opinions or policy judgments, even their idiosyncrasies. As a result, law is shot through with politics and with much else besides that does not fit a legalist model of decision-making.

The decision-making freedom that judges have is an *involuntary* freedom. It is the consequence of legalism’s inability in many cases to decide the outcome (or decide it tolerably, a distinction I shall elaborate), and the related difficulty, often impossibility, of verifying the correctness of the outcome, whether by its consequences or its logic. That inability, and that difficulty or impossibility, create an open area in which judges have decisional discretion—a blank slate on which to inscribe their decisions—rather than being compelled to a particular decision by “the law.”⁹⁹

This all maps out fairly nicely onto the template of the aesthetic judge. Because we cannot fully articulate all the contours of the standards under which such a judge is to judge, she must necessarily draw on her tacit knowledge,

⁹⁷ For an overview and critique of this research, see generally Carolyn Shapiro, *Coding Complexity: Bringing Law to the Empirical Analysis of the Supreme Court*, 60 HASTINGS L.J. 477 (2009); Carolyn Shapiro, *The Context of Ideology: Law, Politics, and Empirical Legal Scholarship*, 75 MO. L. REV. 79 (2010).

⁹⁸ See *supra* notes 58–65 and accompanying text.

⁹⁹ POSNER, *supra* note 5, at 9.

which will reflect the preferences of the community that selected her. In doing so, she will be working from a position of cognitive disadvantage because she cannot possibly take in all the information necessary to fully comprehend a given performance, and she cannot possibly hold in mind all the information necessary to comprehensively compare the present performance to those whom have come before. She will, accordingly, be inclined (whether consciously or not) to rely on other available information to aid her in her task. There is nothing inherent in her internalization of the underlying norms that will necessarily prevent these biases, or her idiosyncrasies, from manifesting themselves. What is more, we might not want her to discard all of this information, for the result might be that her decision making will be negatively affected in that it will be based on less information or on information of lower quality. Some controls—the use of panels of judges, dropping the highest and lowest scores, screening for situations in which affiliation bias might arise—are appropriate. Too many could be detrimental.

It is not necessary to change much in the preceding paragraph to turn it into a description of judging in the legal system. The processes of acculturation and paths to the judicial role are undoubtedly more varied in law than in most aesthetic sports, but so, too, are the disagreements over what the content of the law should be and how law should be applied. If indeed “law is shot through with politics,” then the judge whose politics influence her decisions may be acting rationally and appropriately in letting that influence occur. The influence of factors that are not legal, in the strict sense in which the term is typically used, perhaps provides an avenue for an attenuated sort of democratic feedback into the law-making process. One answer to the counter-majoritarian difficulty is to show that judging is not counter-majoritarian, and one way that might be so is for ideology to affect decision-making.¹⁰⁰

The questions of whether and to what extent it is appropriate for judges in either context to draw on what we have traditionally conceived of as extraneous or improper influences raise a host of normative questions. My aim here is to identify the possibilities rather than to attempt to resolve these deeper issues. Of course, even if we recognize that the influence of ideology or other sources of bias on decision-making can be a good thing, we also recognize that there can be too much of a good thing. Both contexts have made use of similar remedial measures to address this. Codes of judicial ethics direct judges to “uphold

¹⁰⁰ I am packing a lot into a single sentence here, and it is a sentence subject to all sorts of qualification. Suffice it to say that what I have in mind is embodied in analyses like that in Jack M. Balkin & Sanford Levinson, *Understanding the Constitutional Revolution*, 87 VA. L. REV. 1045 (2001). Balkin & Levinson argue that the justices “are temporally extended representatives of particular parties, and hence, of popular understandings about public policy and the Constitution.” *Id.* at 1067.

and apply the law,”¹⁰¹ “perform all duties of judicial office fairly and impartially,”¹⁰² and do so “without bias or prejudice.”¹⁰³ These general standards are implemented via more specific provisions such as prohibitions on ex parte communications¹⁰⁴ and recusal requirements.¹⁰⁵ Aesthetic sports tend to have similar codes of ethics,¹⁰⁶ though the particulars vary considerably.¹⁰⁷ Both law and aesthetic sport make use of multiple-judge panels, especially as the stakes get higher. Some commentators in law have reacted to the evidence of the influence of ideology by suggesting that we should do more to counteract it, such as by

¹⁰¹ ABA, MODEL CODE OF JUDICIAL ETHICS R. 2.2 (2011).

¹⁰² *Id.*

¹⁰³ *Id.* at R. 2.3(A).

¹⁰⁴ *Id.* at R. 2.9.

¹⁰⁵ *Id.* at R. 2.11.

¹⁰⁶ For example, the Skate Canada Officials’ Code of Ethics includes among its list of “Obligations to the Skaters and Coaches” the following:

To maintain objectivity and integrity of judging by marking a performance based on sound technical knowledge. When judging, to mark only the skating being performed without bias or prejudice and not to be influenced by audience approval/disapproval, the reputation and/or the past performance of the skater.

When judging, to mark independently and from the commencement to the conclusion of the event not to discuss with any person, except the Referee and/or Assistant Referee of that event, one’s own assessment or marks or the assessment[]or marks of other judges. . . .

. . . .

To declare a conflict of interest on occasions when applicable and to refrain from officiating in situations where the perception of conflict of interest may be present.

Skate Canada, *Officials’ Code of Ethics*, SKATECANADA, <http://www.skatecanada.ns.ca/en-us/coachesofficials/officials/codeofethics.aspx> (last visited Dec. 21, 2014).

¹⁰⁷ The provisions of the rules of the United States Equestrian Federation that related to judicial conduct are comparatively sparse. The most general provision is as follows: “Good judging depends upon a correct observance of the fine points and the selection of best horses for the purpose described by conditions of the class. A judge serves three interests: his own conscience, exhibitors and spectators. He should make it clear that the best horses win.” U.S. EQUESTRIAN FEDERATION, 2014 RULE BOOK GR 1034(1), available at <https://www.usef.org/documents/ruleBook/2014/GeneralRules/GR10-LicensedOfficials.pdf>. And in snowboarding, “an easygoing vibe permeates” competition. John Branch, *Who Needs Stopwatches? From the Shadows, Judges Take Starring Roles*, N.Y. TIMES (Feb. 5, 2014), http://www.nytimes.com/2014/02/06/sports/olympics/who-needs-stopwatches-from-shadows-judges-are-co-stars.html?_r=0. “The judges invite athletes and coaches to talk to them – even in the middle of a competition, maybe between qualifying rounds and the finals. Some sports are inclined to sequester judges, to protect them from the lobbying efforts of persuasive participants. Not snowboarding and freeskiing.” *Id.*

mandating ideological diversity on appellate court panels, which would roughly parallel the concern with nationalistic bias that drives some of the rules relating to the composition of aesthetic judging panels.¹⁰⁸

One last parallel stems, as well, from the reliance on tacit knowledge in both types of judging. It is that efforts to reform the process of judging by making it more rule-governed are, if not doomed to fail, at least likely to create a host of unintended consequences. Perhaps the most prominent legal example is the Federal Sentencing Guidelines, which were created in response to concerns about disparities resulting from the discretionary sentencing regime that preceded them. Without denying that the guidelines have had positive effects, or even that the positives might outweigh the negatives, it is clear that many judges feel that the guidelines have deprived them of the ability to tailor sentences to appropriately match context,¹⁰⁹ and some commentators have decried the guidelines more generally.¹¹⁰ Changes to judging standards and procedures in figure skating have similarly constrained officials, and had a corresponding effect on what gets rewarded in performance. As one commentator put it, while the new standards have engendered consistency, their

rigid technical criteria have promoted a “teaching to the test” mentality that has homogenised performances and squelched individual expression and creativity. Because the scoring system no longer rewards overall aesthetic beauty, skaters, coaches, and choreographers instead devote their efforts to poring over the codes and reviewing slow-motion video, hoping to devise new features that will enhance the base values of required elements or squeeze out an extra GOE point. That leaves little time either during routines or in training sessions for optional acrobatic or artistic showstoppers, like Michelle Kwan’s notorious spirals, in which she flashed a huge smile while speeding down the ice and audiences routinely jumped to their feet.¹¹¹

¹⁰⁸ See Emerson H. Tiller & Frank B. Cross, *A Modest Proposal for Improving American Justice*, 99 COLUM. L. REV. 215, 228–32 (1999).

¹⁰⁹ For a sampling of these complaints, see Deborah E. Dezelan, *Departures from the Federal Sentencing Guidelines After Koon v. United States: More Discretion, Less Direction*, 72 NOTRE DAME L. REV. 1679, 1690 (1997).

¹¹⁰ See, e.g., Melissa Hamilton, *McSentencing: Mass Federal Sentencing and the Law of Unintended Consequences*, 35 CARDOZO L. REV. (forthcoming 2014).

¹¹¹ Chloe Katz, *Judging in Figure Skating: Levelling the Rink*, ECONOMIST BLOG (Feb. 15, 2014, 10:48 PM), <http://www.economist.com/blogs/gametheory/2014/02/judging-figure-skating>.

V. CONCLUSION

The notion of the “judge as judge,” as I have outlined it in this essay, clearly has less rhetorical appeal than the more familiar “judge as umpire.” The reasons are obvious. Judging in aesthetic sport has a known history of corruption and apparent bias and lacks the umpire’s association with a game that has long been central to American identity. As a device for selling one’s role, whether to litigants or to the Senate Judiciary Committee, it accordingly falls short.

But if what we are after is an accurate point of comparison, the “judge as judge” works better. The same is true when the analysis is meant to serve as a means for thinking and generating insights about the judicial role in the legal system. There are, of course, reasons to be skeptical of the sport-law comparison in general, and of any specific analogy in particular. But so long as we are mindful of those limitations, aesthetic sports might serve as laboratories of sorts, generating insights and perhaps even institutions and procedures that can aid in the evolution of the judicial system as it copes with demands both new and old. So, too, might law have insights to offer to governing bodies in sport.

* * *