

Robert Schmidt, David Kempf,
Max Weigelin (Hg.)

LEISTUNGS- VERGLEICHE UND EVALUATIVE PRAKTIKEN

Sport als instruktiver Fall
der Soziologie der Bewertung

[transcript] Kulturen der Gesellschaft

Robert Schmidt, David Kempf, Max Weigelin (Hg.)
Leistungsvergleiche und evaluative Praktiken

Robert Schmidt (Dr. phil.), geb. 1964, ist Professor für Soziologie an der Katholischen Universität Eichstätt-Ingolstadt, Leiter des DFG-Forschungsprojekts »Accounting und transformatorische Effekte im Profifußball« und Sprecher des DFG-Graduiertenkollegs »Practicing Place«. Seine Forschungsschwerpunkte sind Praxistheorie, soziologische Ethnographie und Kulturanthropologie des Politischen.

David Kempf (M.A.), geb. 1993, ist wissenschaftlicher Mitarbeiter im DFG-Graduiertenkolleg »Practicing Place« an der Katholischen Universität Eichstätt-Ingolstadt. In seiner Dissertation untersucht er ethnografisch Praktiken des realistischen Darstellens am Fall der Fußballsimulation *Football Manager*. Seine Forschungsschwerpunkte sind Praxistheorie, soziologische Ethnografie und Pragmatismus.

Max Weigelin (M.A.), geb. 1988, ist Mitarbeiter des DFG-Projekts »Accounting und transformatorische Effekte im Profifußball« an der Katholischen Universität Eichstätt-Ingolstadt. In seiner Dissertation untersucht er Schiedsrichter und Schiedsrichterwesen im Fußball organisationsethnografisch. Seine Forschungsschwerpunkte sind soziologische Ethnografie, Kulturosoziologie, Organisationstheorie und Soziologie der Sinne.

Robert Schmidt, David Kempf, Max Weigelin (Hg.)

Leistungsvergleiche und evalutive Praktiken

Sport als instruktiver Fall der Soziologie der Bewertung

[transcript]

Hervorgegangen aus dem von der DFG geförderten Forschungsprojekt »Accounting und transformatorische Effekte im Profifußball« an der Katholischen Universität Eichstätt-Ingolstadt.

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <https://dnb.dnb.de/> abrufbar.



Dieses Werk ist lizenziert unter der Creative Commons Attribution-Non-Commercial 4.0 Lizenz (BY-NC). Diese Lizenz erlaubt unter Voraussetzung der Namensnennung des Urhebers die Bearbeitung, Vervielfältigung und Verbreitung des Materials in jedem Format oder Medium ausschliesslich für nicht-kommerzielle Zwecke.

<https://creativecommons.org/licenses/by-nc/4.0/>

Um Genehmigungen für die Wiederverwendung zu kommerziellen Zwecken einzuholen, wenden Sie sich bitte an rights@transcript-publishing.com

Die Bedingungen der Creative-Commons-Lizenz gelten nur für Originalmaterial. Die Wiederverwendung von Material aus anderen Quellen (gekennzeichnet mit Quellenangabe) wie z.B. Schaubilder, Abbildungen, Fotos und Textauszüge erfordert ggf. weitere Nutzungsgenehmigungen durch den jeweiligen Rechteinhaber.

Erschienen 2024 im transcript Verlag, Bielefeld

© Robert Schmidt, David Kempf, Max Weigelin (Hg.)

Umschlaggestaltung: Maria Arndt, Bielefeld

Korrektorat: Sascha Frischmuth

Druck: Elanders Waiblingen GmbH, Waiblingen

<https://doi.org/10.14361/9783839472361>

Print-ISBN: 978-3-8376-7236-7

PDF-ISBN: 978-3-8394-7236-1

Buchreihen-ISSN: 2703-0040

Buchreihen-eISSN: 2703-0059

Gedruckt auf alterungsbeständigem Papier mit chlorfrei gebleichtem Zellstoff.

Inhalt

Einleitung in den Band

Robert Schmidt, David Kempf, Max Weigelin 7

Leistungsvergleiche und evaluative Praktiken

Sport als instruktiver Fall für eine Soziologie der Bewertung

Robert Schmidt, David Kempf, Max Weigelin 13

Spielerwerte und Spielen/Bewerten

Digitale Fußballspiele und die transmedialen Affekte von Leistungsbewertung(en)

Martin Lüthe 43

Bewerten, um plausibel darzustellen

Researcher und Scouts für Football Manager

David Kempf 61

Die Vermessung von Geschlecht im Sport und der Umgang mit kategorialen

Transgressionen

Dennis Krämer 93

Bahnen ziehen. Schwimmen als Bewertungsspiel

Thorsten Peetz 125

Vergleichen, Differenzieren, Kategorisieren und das Problem der Fairness

(Sport-)Unterricht als instruktiver Fall einer Soziologie der Bewertung

Samira Mummelthey & Kerstin Rabenstein 145

Die Erfindung des modernen Fußballs Praktiken des Vergleichens und Bewertens im <i>Alcock's Football Annual</i> , 1868 bis 1908 <i>Clelia Minnetian & Tobias Werron</i>	173
Die Un/Sichtbarkeit des Regelhüters Zwei Strategien zum Umgang mit der Unterdeterminiertheit von Regeln im Schiedsrichterwesen <i>Max Weigelin</i>	211
Fairer Wettbewerb in Sport und Wissenschaft <i>Frank Meier</i>	241
Autor:inneninformationen	265

Einleitung in den Band

Robert Schmidt, David Kempf, Max Weigelin

Anliegen des Bandes

Die kulturwissenschaftlich-praxeologische Analyse von sozialen Praktiken des Bewertens, aber auch des Kategorisierens, Vergleichens, Quantifizierens und Entscheidens hat sich in den letzten zwei Jahrzehnten zu einem breiten Forschungsfeld soziologischer Grundlagenforschung entwickelt. Studien auf diesem Feld befassen sich hauptsächlich mit der anhaltenden Konjunktur, den Relevanzgewinnen, Operationsweisen, sozialen Logiken und den Effekten von organisational gerahmten Bewertungsverfahren. Überraschenderweise taucht in der Reihe der Schwerpunkte und Modelfälle dieser Forschung der Gegenstandsbereich Sport bislang (noch) nicht auf. Dies überrascht nicht nur, weil Leitbegriffe sportlicher Bewertungsverfahren (zum Beispiel »Fairness«) die Semantiken anderer Bewertungsregime prägen (vgl. etwa Bröckling 2020), sondern auch, weil organisational gerahmtes Bewerten und Quantifizieren zum Wesenskern des modernen Sports gezählt werden (vgl. Guttmann 1978).

- Mit dem vorliegenden Band wollen wir zur Bearbeitung dieser Forschungslücke anregen und beitragen. Bewertungssoziologisch interessant ist der Sport nämlich nicht nur, weil er sich vordergründig auch im Gegenstandsbereich bewertungssoziologischen Interesses befindet. Die Ausbreitung neoliberal geprägter und durch Digitalisierung noch einmal forcierbarer Bewertungsregime findet sich auch hier. Tracking-Devices, Konjunkturschübe sportwissenschaftlicher Quantifizierung und ähnliches stoßen tiefgreifende Transformationen sportlicher Praktiken an. Wir plädieren vielmehr dafür, den Sport noch grundlegender als anregenden Fall für die Soziologie der Bewertung in den Blick zu nehmen. An Fällen aus dem Sport lässt sich bewertungssoziologisch viel gewinnen, weil dieses kulturelle Feld einerseits eigenlogische Bewertungspraktiken und Vergleichsformate (Ligen, Ranglisten etc.) hervorbringt, die von sportfernen Bereichen adaptiert werden.

Andererseits übernimmt der Sport aus anderen Feldern bekannte Bewertungsverfahren, spitzt sie um, de- und rekonstruiert sie und lässt sie dabei immer auch prekär und strittig werden. Bewertungssoziologisch instruktiv ist also zum einen die Frage, wie neuartige (digitale, statistische, neoliberalen etc.) Bewertungsregime das Feld des Sports und einzelne sportliche Praktiken verändern. Zum anderen kann in den Blick gerückt werden, wie sportfremde Bewertungsverfahren durch den Sport angeeignet, dramatisiert und kritisierbar gemacht werden.

Ein besonders anschauliches Beispiel liefert der Fußball. Einerseits lässt sich hier der Sport als prototypischer Fall der Konstruktion moderner Wettbewerbsarrangements untersuchen. Viele zeitgenössische Generationen begegnen Konzepten wie »Wettbewerb«, »Fairness« und »Leistung« biografisch zunächst im Kontext von Sportspielen. Fußball ist dabei ein interessanter Fall, weil er einerseits die sporttypische Veranschaulichung, Zuspitzung und Valorisierung von »Leistungen« leistet, andererseits als »Low Numbers Game« und Sport mit recht viel Spielraum für List und Tücke auch die Ambivalenz sozialer Konkurrenz- und Wettbewerbskonstellationen unterstreicht. Im Fußball scheint sich die moderne Kultur selbst zu beobachten als tief geprägt vom Glauben an den Wettbewerb und zugleich als sich bewusst über die Unzulänglichkeiten konkreter Wettbewerbsformen. Der Fall des Fußballs deutet insofern das Potential für eine, mit dem Blick auf konkrete Bewertungspraktiken empirisch informierte, »vom Spiel ausgehende[n] Soziologie« (Callois, 1982). Daraus ergibt sich eine produktive Neuperspektivierung: Statt – wie in der bewertungssoziologischen Forschung etwa zum Bildungs- und Wissenschaftsfeld üblich – zu untersuchen, wie ökonomische Bewertungsregime nicht-ökonomische Bereiche wie den Sport kolonisieren und transformieren, kann nun umgekehrt danach gefragt werden, wie sportliche Praktiken in ihrem mimetisch-transformativen Spiel mit Materialien der Kultur nicht zuletzt auch Bewertungsverfahren umarbeiten, artikulieren und problematisieren. Durch diese Neuperspektivierung kommen die Eigenlogik und die kulturelle Darstellungsfähigkeit sportlicher Praktiken (Gebauer & Wulf, 1998) in den Blick. Diese Perspektive ist allerdings nur ein Aspekt, der den Sport zu einem instruktiven Fall für die Soziologie der Bewertung macht. Andere Konvergenzonen zwischen der Soziologie des Sports und der Soziologie der Bewertung erscheinen nicht weniger ergiebig. So liefert der Sport Fälle, an denen sich die Bewertung von Körperbewegungen variantenreich studieren (Schmidt et al. 2022) oder die Genese von spezifisch modernen Wettbewerbskonstellationen historisch rekonstruieren lässt (vgl. Minnetan/Werron 2021,

Dies. In diesem Band). Die Beiträge dieses Bandes eint, dass sie auf diesen und anderen Spuren die Idee einer Weiterentwicklung bewertungssoziologischer und angrenzender Konzepte durch empirisch-analytische Bezüge aus dem Kontext des Sports forcieren.

Beiträge des Bandes

In ihrem Beitrag »Bewertungssoziologie und der Fall Sport« identifizieren Robert Schmidt, David Kempf und Max Weigelin das Feld des Sports als Forschungslücke in der Bewertungssoziologie. In methodischer Hinsicht plädieren die Autoren für eine kulturalistische Erweiterung der Bewertungssoziologie, die die spezifische kulturelle Darstellungsfähigkeit des Sports nutzen sollte, um neue Perspektiven auf gesellschaftliche Bewertungspraktiken zu werfen. Der Sport wird somit als ein lebendiges Laboratorium dargestellt, in dem sich die Dynamik von Bewertungsprozessen in besonders prägnanter Weise beobachten lässt. Illustriert wird diese Perspektive an verschiedenen Realitätsverdopplungseffekten im Feld des Fußballs.

Einen empirisch besonders innovativen Zugang zu solchen Phänomenen entwickelt der Amerikanist Martin Lüthe anhand einer transmedialen Untersuchung der sogenannten »Rating Reveal Videos« der Videospielreihe »EA Sports FC« (vormals »FIFA«). In diesen Videos werden Fußballspieler mit ihren quantifizierten Bewertungen innerhalb des Spiels konfrontiert. An diesen Videos zeigt sich, dass digitale Fußballspiele nicht nur Bewertungsmechanismen innerhalb der Spiele darstellen – Spielern werden ihrer »Stärke« entsprechende Zahlenwerte zugeschrieben (»rating«) – sondern auch eine affektive Dimension besitzen. Martin Lüthe arbeitet heraus, dass ihre zentrale Funktion darin besteht, die zunächst kalte quantifizierende Bewertungslogik des Videospiels in ihrer Affektstruktur erfahrbar zu machen.

Das bewertungssoziologisch instruktive Potential der Videospieladaptationen des Fußballs beleuchtet auch der Beitrag »Bewerten, um plausibel darzustellen« von David Kempf. Er fokussiert auf die der Fußballmanagement Simulation »Football Manager« zugrunde liegende Datenerhebung von Spielerstärken. Dabei stößt er auf einen für die Arbeit an einer öffentlich zugänglichen Darstellung spezifischen Bewertungsstil: noch wichtiger als der Wahrheitsgehalt der Darstellungen ist ihre Plausibilität.

Erkennbar wird in diesen Beiträgen also unter anderem die Performativität von spielerisch sportlich relevanten Bewertungen. Die Performativität von

Kategorisierungspraktiken im Kontext von Bewertungsprozessen wiederum beleuchtet unter anderem Dennis Krämer in seinem Beitrag »Die Vermessung von Geschlecht im Sport und der Umgang mit kategorialen Transgressonen«. Der Sport erweist sich hier mit Blick auf Humankategorien als ein voraussetzungreiches Bewertungsregime. Um Vergleichbarkeiten von Leistungen herzustellen und abzugrenzen, mobilisiert er unter anderem die Geschlechterkategorie. Insofern der Sport diese dadurch zum Schlüsselement in Ein- und Ausschließungsprozessen für Wettbewerbe erhebt, transformiert er sie auch. Krämer arbeitet heraus, wie medizinische Klassifizierungspraktiken die Biologisierung des Geschlechts historisch immer wieder erneuern, indem sie den entscheidenden Unterschied sportspezifisch mit Bezug auf »Leistung« fixieren.

Im Beitrag »Bahnhen ziehen: Schwimmen als Bewertungsspiel« untersucht Thorsten Peetz verschiedene Varianten des Schwimmens als »Bewertungsspiele«. Zentral ist die Unterscheidung zwischen unterschiedlichen Formen und Kontexten des Schwimmens, wie Beckenschwimmen und Freiwasserschwimmen, und deren jeweiligen Bewertungsnormen und -praktiken. Durch Rekurs auf sportsoziologische Studien zu diesen Fällen kann er zeigen, dass Schwimmen nicht nur physische Bewegung ist, sondern auch eine soziale Praxis, die durch vielfältige Bewertungsprozesse geprägt ist. Peetz betont die Notwendigkeit, den wertenden Charakter der Praktiken innerhalb des Schwimmsports zu erkennen und zu analysieren. Bewertungen finden nicht nur durch externe Beobachter wie Trainer und Schiedsrichter statt, sondern sind immer auch integraler Bestandteil der Selbsteinschätzung und Selbstregulation der Schwimmerinnen.

Mit einer besonderen Komplexität konfrontieren Samira Mummelthey und Kerstin Rabenstein die Bewertungssoziologie in ihrem Beitrag »Vergleichen, Differenzieren, Kategorisieren und das Problem der Fairness. (Sport-)Unterricht als instruktiver Fall einer Soziologie der Bewertung«: Die Autorinnen argumentieren, dass Sportunterricht sowohl als Verunterrichtlichung von Sport als auch als Versportlichung von Unterricht verstanden werden kann, was zu einer komplexen Amalgamierung spezifischer Bewertungslogiken führt. Dabei heben sie hervor, wie Vergleichbarkeit zwischen Schülern durch bestimmte Praktiken hergestellt wird und welche Kategorisierungen dabei vorgenommen werden.

Damit werfen Mummelthey und Rabenstein auch ein Schlaglicht auf die komplexe kulturelle Konstruktion von Wettbewerben, insofern diese Aushandlungen um den Gehalt von Konzepten wie »Leistung« und »Fairness«

beinhalten. Passend dazu interessieren sich Clelia Minnetian und Tobias Werron für das Zusammenspiel von kulturellen Leistungskonzepten, massen-medialen Bewertungsformaten und den genuinen Leistungssportstrukturen der Liga mit ihrer Tabelle.

Die Untersuchung »Die Erfindung des modernen Fußballs: Praktiken des Vergleichens und Bewertens im Alcock's Football Annual, 1868 bis 1908« zeigt, dass die frühen Formen des Fußballs durch eine eher lokale und unstrukturierte Praxis charakterisiert waren, bei der es weder regelmäßigen Wettbewerbe noch ein einheitliches Regelwerke gab. Mit der Gründung der Football Association begann dann eine Phase der Standardisierung und Institutionalisierung, die durch die Publikation von Ligatabellen und statistischen Jahresberichten ergänzt wurde. Im amerikanischen Baseball behandelte der hierarchische und statistikaffine Leistungsdiskurs des Sportsjournalismus unterschiedliche Spiele als vergleichbar und trug damit zur Entwicklung des Ligaformats bei. Im vorliegenden Fall des Fußballs hingegen führte die umgekehrte Einführung des Ligaformats dazu, dass zuvor auch vom Sportjournalismus kaum miteinander als vergleichbar behandelte Spiele nun in einen institutionellen Vergleichszusammenhang gestellt wurden.

In seinem Beitrag »Die Un/Sichtbarkeit des Regelhüters: Zwei Strategien zum Umgang mit der Unterdeterminiertheit von Regeln im Schiedsrichterwesen« beschäftigt sich Max Weigelin mit der Rolle von Schiedsrichtern im Verbandsfußball aus einer bewertungssoziologischen Perspektive. In den Blick nimmt er dabei die Synergiepotentiale zwischen der Soziologie der Bewertung und der Wettbewerbssoziologie. Zentral ist für ihn die Unterscheidung zwischen zwei Hauptstrategien der Bearbeitung von Regelambiguität durch Schiedsrichter: die Technisierung, die die Kontingenz der Regelanwendung unsichtbar macht, und das Mitspielen, das den Schiedsrichter als aktiven und interpretierenden Akteur im Spiel platziert. Der Artikel argumentiert, dass beide Strategien notwendig sind, um die Glaubwürdigkeit und Fairness des Spiels darzustellen.

Frank Meier schließlich nutzt das Feld des Sports, um Anregungen für bewertungssoziologische Untersuchungen der Wissenschaft zu generieren. Wettbewerbssoziologische Analyseinstrumente erlauben, die beiden Felder in seinem Beitrag »Fairer Wettbewerb in Sport und Wissenschaft« zu vergleichen. Im Skisport etwa findet er kalkulatorische Praktiken, die Unterschiede in den Ausgangsbedingungen (etwa Windstärke) virtuell nivellieren. Hier arbeitet Meier einen zentralen Unterschied zur Wissenschaft heraus: Während ungleiche Ausgangsbedingungen der Leistungserbringung im Sport teilweise

auf kreative Weise in Rechnung gestellt werden, verweist die Wissenschaft auf die Ungleichheit disziplinärer Kontexte.

Literatur

- Bröckling, U. (2020). Wettkampf und Wettbewerb: Konkurrenzordnungen zwischen Sport und Ökonomie. *Konfliktdynamik*, 9(4), 248–255.
- Gebauer, G. & Wulff, C. (1998). Spiel-Ritual-Geste: mimetisches Handeln in der sozialen Welt. Hamburg: Rowohlt Verlag.
- Guttmann, A. (2004). From ritual to record: The nature of modern sports. Columbia university press.
- Minnetian, C., & Werron, T. (2021). Redefining achievement: The emergence of rankings in American baseball. In Ringel et.al. (Hg.) *Worlds of rankings* (127–151). Emerald Publishing Limited.
- Schmidt, R., Weigelin, M., Brümmer, K., Laube, S., & Schäfer, H. (2022). Bodies and embodiment in practices of valuation: Challenging the sociology of valuation with the sociology of the body. *Österreichische Zeitschrift für Soziologie*, 47(3), S. 217–223

Leistungsvergleiche und evaluative Praktiken

Sport als instruktiver Fall für eine Soziologie der Bewertung

Robert Schmidt, David Kempf, Max Weigelin

Die Analyse von sozialen Praktiken des Kategorisierens, Vergleichens, Quantifizierens, Bewertens und Entscheidens hat sich seit rund zwei Jahrzehnten zu einem Forschungsfeld entwickelt, das unter der Bezeichnung *Valuation Studies* bekannt geworden ist.¹ Die überwiegend praxeologisch ausgerichteten Studien auf diesem Feld befassen sich hauptsächlich mit der anhaltenden Konjunktur, den Relevanzgewinnen, Operationsweisen, sozialen Logiken und den Effekten von Bewertungsverfahren.

Die Soziologie der Bewertung ist durch empirische Schwerpunkte und Modelfälle gekennzeichnet, zu denen der Gegenstandsbereich Sport (noch) nicht gerechnet werden kann.² Dies überrascht nicht nur, weil Leitbegriffe

-
- 1 Die Durchsetzung der Bezeichnung und die Etablierung dieses Forschungsfeldes dokumentieren sich nicht zuletzt auch im erfolgreichen Launch des internationalen Journals *Valuation Studies*, das seit 2013 regelmäßig erscheint.
 - 2 Die empirischen Fallbezüge der *Valuation Studies* reichen von Studien zur Konstruktion von Rankings in organisationalen Feldern wie dem der Universitäten (klassisch Espeland & Sauder, 2007) über Analysen makrosoziologisch-semantischer Phänomene wie den »Rechtfertigungsordnungen« (Boltanski & Thevenot, 2007) bis hin zur mikrosoziologischen Entschlüsselung multipler Ontologien, die bei der Bewertung des Anbaus, der Vermarktung und des Einkaufens von Tomaten relevant werden (Mol & Heuts, 2013). Scannt man die Beiträge im Journal *Valuation Studies*, von dem seit 2013 durchschnittlich zwei Ausgaben pro Jahr erschienen sind und dessen Beitragsumfang sich i.d.R. auf vier oder fünf Beiträge beläuft (grob geschätzt ca. 64 Beiträge insgesamt), findet man nicht ein einziges Paper, das den Sport oder eine Sportart im Titel trägt. Das Einführungswerk von Krüger (2022) erwähnt nur eine Studie mit Sportbezug (diese aber gleich mehrfach), nämlich Tobias Werrons Untersuchung zur Entstehung des Ligenformats im Leistungssport (Werron 2010). Wir werden auf diese Arbeit noch zurückkommen.

sportlicher Bewertungsverfahren (zum Beispiel »Fairness«) die Semantiken anderer Bewertungsregime prägen (vgl. etwa Bröckling 2020), sondern auch, weil Bewerten und Quantifizieren zum Wesenskern des modernen Sports gezählt werden (vgl. Guttman 1978). In den Meter-Gramm-Sekunde Sportarten etwa werden sportliche Leistungen gemessen und bewertet, um Athletinnen miteinander zu vergleichen und sie dann in einer Rangfolge als Gewinnerinnen (Erste) bzw. als abgestufte Verlierer (Zweite, Dritte usw.) zu präsentieren. Solche Ergebnisse eines einzelnen Wettkampfs werden wiederum in Ranglisten höherer Ordnung weiterverarbeitet. Man erweitert den Vergleichshorizont durch Saisontabellen und Weltrekordlisten in zeitlicher Hinsicht oder man wechselt das Bezugsobjekt der Bewertung. Letzteres geschieht etwa dann, wenn die Ergebnisse einer Leichtathletik WM in »Medaillenspiegeln« zur Bewertung von Nationen (bzw. von deren Leichtathletikverbänden) aggregiert werden (vgl. Werron 2005).

Als ein bewertungsaffines kulturelles Feld markiert der Sport eine Forschungslücke der *Valuation Studies* – trotz seiner augenfälligen Rolle und Bedeutung im gesellschaftlichen Bewertungsgeschehen, das er fortlaufend mit Bildern, Motiven, Deutungsmustern und Semantiken versorgt. Wir wollen mit dem vorliegenden Band und mit vorausgegangenen Initiativen³ zur Bearbeitung dieser Forschungslücke anregen und beitragen. Bewertungssoziologisch interessant ist der Sport nicht nur, weil sich daran gegenwärtig studieren lässt, wie sich durch digitale Bewertungspraktiken seine evaluativen Infrastrukturen wandeln und wie diese Prozesse Veränderungen sportlicher Praktiken in Gang bringen. An Fällen aus dem Sport lässt sich viel lernen, weil dieses kulturelle Feld einerseits eigenlogische Bewertungspraktiken und Vergleichsformate (Ligen, Ranglisten etc.) hervorbringt, die von sportfernen Bereichen adaptiert werden und weil der Sport andererseits aus anderen Feldern bekannte Bewertungsverfahren übernimmt, umarbeitet, zuspitzt, re- und dekonstruiert, und sie dabei immer auch prekär und strittig werden

3 Der vorliegende Band ist aus einem kollektiven Arbeitsprozess hervorgegangen. Zu diesem Prozess zählen die Forschungen in unserem DFG Projekt »Accounting und transformatorische Effekte im Profifußball«, die Tagung »Accounting von Körperbewegungen – Sport als instruktiver Fall für eine Soziologie der Bewertung«, die wir zusammen mit der DGS Sektion »Soziologie des Körpers und des Sports« und dem DFG Netzwerk »Auf dem Weg in die Bewertungsgesellschaft?« organisiert und ausgerichtet haben (Sommer 2020), sowie zwei Workshops in Eichstätt (Frühjahr und Herbst 2023), in denen wir mit den Autor:innen des Bandes dessen Konzeption und die Beiträge gemeinsam diskutiert haben.

lässt. Bewertungssoziologisch instruktiv ist also zum einen die Frage, wie neuartige (digitale, statistische, neoliberalen etc.) Bewertungsregime das Feld des Sports und einzelne sportliche Praktiken verändern. Zum anderen kann in den Blick gerückt werden, wie sportfremde Bewertungsverfahren durch den Sport angeeignet, dramatisiert und kritisierbar gemacht werden.

Um zwischen bewertungs- und sportsoziologischen Fragestellungen zu vermitteln und Konvergenzonen zu identifizieren, skizzieren wir im Folgenden zunächst thematische Fluchtpunkte, in denen sich Sport- und Bewertungssoziologie treffen (1) und stellen aktuelle bewertungssoziologisch orientierte Studien zu sportlichen Gegenstandsbereichen vor, die bereits dazu beigetragen haben, die genannte bewertungssoziologische Forschungslücke zu schließen (2). Im Anschluss daran umreißen wir eine kulturanalytische Heuristik, die wir in unserer Forschungsarbeit (im DFG-Projekt »Accounting und transformatorische Effekte im Profifußball«) entwickelt haben (3). Ihr Ausgangspunkt ist die Frage, wie der Sport Verfahren und Praktiken des Bewerterns adaptiert, performt und dramatisiert. Wir veranschaulichen analytische Potentiale dieser Perspektive im Rekurs auf unsere empirischen Analysen zu neuen digitalen Bewertungspraktiken im professionellen Gegenwartsfußball (3.1) und zu den Entwicklungen rund um den Transfermarkt (3.2). Zusammenfassend plädieren wir dafür, dass sich eine kulturanalytisch erweiterte Bewertungssoziologie die spezifische kulturelle Darstellungsfähigkeit des Sports zu Nutze machen sollte (4).

1. Konvergenzonen: Die Konstruktion von Wettbewerben und die Bewertung von Körper(bewegungen)

Ein gemeinsames Thema von Sport- und Bewertungssoziologie sind Forschungen zu Konkurrenz- und Wettbewerbsformen. So beschäftigen sich die für die *Valuation Studies* bahnbrechenden Studien von Espeland und Sauder (2007, 2016) mit der Konstruktion von Wettbewerben und Rankings zwischen US-amerikanischen Law Schools und arbeiten die performativen sozialen Effekte (*reactivity*) dieser wettbewerblichen Steuerungsinstrumente heraus. Auch für eine große Zahl anderer Studien, die sich der Bewertungssoziologie zuordnen oder ihnen typischerweise zugeordnet werden, gilt, dass Bewertungsverfahren vor allem deshalb in den Blick geraten, weil sie rekonstruierbar machen, wie Konkurrenzordnungen eines jeweiligen Untersuchungsfeldes wettbewerblich transformiert werden. So beschreibt die Studie von Münch

zu den Bewertungsregimen PISA und den Bologna-Reformen, wie in deren Gefolge neolibrale New Public Management Methoden neue Wettbewerbsformate im Bildungs- und Wissenschaftssektor etablieren (Münch, 2009). Die Studie von Christin untersucht nicht nur die Bedeutung von Seitenaufrufmetriken in Zeiten der Digitalisierung, sondern zeichnet damit auch nach, wie die Relevanz dieser Metriken die Verteilung von Prestige und Gewinn im Feld neu strukturieren (Christin, 2018). Und – um diese relativ beliebige Liste an Beispielen nur noch um einen Fall aus einem weiteren Feld zu ergänzen – Kornberger und Mennicken schlagen vor, die Restrukturierung verschiedener Berufsfelder (etwa der Taxifahrerinnen) durch plattformökonomische Akteure (etwa Uber) und deren digital-algorithmische Bewertungsregime governementalitätstheoretisch zu lesen, so dass u.a. eine Transformation der Konkurrenzordnungen der Arbeitnehmerinnen erkennbar wird (Mennicken & Kornberger, 2021).

Dabei kommt der Sport allerdings kaum in den Blick. Eine der wenigen Ausnahmen stellt eine theoretisch-konzeptionell angelegte Arbeit von Patrik Aspers dar. An ihr zeigt sich, wie aus der empirischen Analyse von sportlichen Konkurrenz- und Wettbewerbsformen theoretische Innovationen für die Bewertungssoziologie generiert werden können. Er schlägt vor, »decision, valuation and contest« begrifflich als »forms of uncertainty reduction« unter der Bedingung der Abwesenheit institutionell fixierter Standards zu fassen (2018, S. 133).⁴ Neben Entscheidungen (*decision*) und Wertungen (*valuation*) führt Aspers als dritte Form den Wettbewerb (*contest*) an und zeigt an der Entstehung von Ligensystemen im Sport, wie die Autorität und Legitimität dieser Rangordnungen durch die direkte Konfrontation im sportlichen Wettkampf verbürgt wird.

Noch deutlicher wird das mit der Verschränkung von sport- und bewertungssoziologischen Perspektiven und ihrer Engführung an den Themen Wettbewerb und Konkurrenz gegebene Innovationspotential mit Blick auf die Arbeiten von Tobias Werron, Clelia Minnetian und Leopold Ringel. In einer Studie zur Entstehung des modernen Sports aus weltgesellschaftstheoretischer und vergleichssoziologischer Perspektive (Werron, 2010) sowie in einigen daran anschließenden Arbeiten (Minnetian & Werron, 2021; dies.

⁴ Dabei perspektiviert Aspers Konkurrenz- und Wettbewerbsformen wirtschaftssoziologisch und gesellschaftstheoretisch als Voraussetzungen und Resultate von Vermarktlichung.

in diesem Band) verknüpfen die Autorinnen sportsoziologische Perspektiven auf die Entstehung genuin sportförmiger Wettbewerbsformen mit bewertungssoziologischen Debatten (Ringel et al., 2021). Sie haben diese Forschungsperspektive mittlerweile zu einer vergleichenden Soziologie des Rankings ausgebaut und erarbeiten mit Bezug auf Fälle aus dem Sport, der Kunst und der Wissenschaft Modelle der Institutionalisierung von Rangfolge-Ordnungen (Ringel & Werron, 2020).

Im Band von Arora et al. (2021) wird der Sport direkt als instruktiver Fall für eine Soziologie der Konkurrenz und des Wettbewerbs in den Blick genommen. Eine vergleichende Soziologie der Konkurrenz muss demnach vor allem nach der historischen Konstruktion von Konkurrenzarrangements wie dem wirtschaftlichen Markt und dem sportlichen Wettbewerb und nach der Aufrechterhaltung und Transformation solcher Arrangements fragen. Sie richtet den Blick damit auf funktional äquivalente Organisationsweisen und stellt die für die Moderne kennzeichnende Ausbreitung und Vermehrung von Konkurrenzregimen und die Faszination des Wettkampfs in der bürgerlichen Kultur heraus.

Der Sport ist darüber hinaus aber nicht nur ein prototypischer Fall der Konstruktion von Konkurrenz und Wettbewerb, sondern auch ein Modelfall für die Analyse der Bewertung von Körpern und Körperbewegungen. Solche Körper-Bewertungen sind ein zentrales Thema der Sportpädagogik, der anwendungsorientierten Sportwissenschaft und nicht zuletzt der sportlichen Leistungsermittlung. Die Soziologie der Bewertung kann hier anschließen, denn sie hat sich in zahlreichen Fallstudien mit Fragen der Körperlichkeit von Bewertungspraktiken beschäftigt (vgl. Schmidt et.al., 2022). Empirisch liegt dieser Fokus nahe, weil auch viele nicht sportbezogene Evaluationspraktiken körperliche Performanz bewerten und den Einsatz des Körpers der Evaluierenden erfordern. So wird von Feinschmeckerinnen ein besonderes körperliches Sensorium erwartet und eingesetzt. Ihre Fähigkeiten, Lebensmittel zu bewerten und sich von Nicht-Gourmets abzugrenzen, beruhen auf einer Sozialisation des Gaumens, die bestimmte Körpertechniken formt (Hennion, 2015; Berli & Alberth, 2022). Aber auch hoch formalisierte oder technisch vermittelte Bewertungsverfahren, die auf den ersten Blick »körperlos« erscheinen mögen, erfordern spezifische Mobilisierungen und Einsätze des Körpers.⁵ Eine Reihe von familienähnlichen Bewertungssituationen (Antal

5 Laube (2019) hat dies am Fall der Beobachtung und Bewertung von Preisentwicklungen im Finanzhandel herausgearbeitet.

et al., 2015) verlangen überdies die körperliche Kopräsenz der zu bewertenden Person und der Bewerter. Dazu zählen u.a. Prüfungen und Vorstellungsgespräche. Interessant ist hierbei, dass viele Verfahren die Anwesenheit und zugleich auch die Abwesenheit der Körper von Bewerteten und Bewerterinnen gezielt einsetzen und als Schritte einer Verfahrenslogik prozessieren.⁶ An die Prüfungs- und Beurteilungssituationen schließt häufig die öffentliche Präsentation der Beurteilungsergebnisse an (Krüger & Hesselmann, 2020). So setzt das Ritual der Preisverleihung die körperliche Anwesenheit der Bewertenden und der Bewerteten voraus. Bewertungsverfahren lassen sich also hinsichtlich der verschiedenen Formen, Grade, Logiken und Strategien der Verkörperung von Leistung und Qualifikation voneinander unterscheiden.⁷

Die körperliche Dimension von Bewertung lässt sich aber auch aus der Perspektive der Bewerteten beleuchten, die in der Bewertungssoziologie oft zu Gunsten der Bewertenden vernachlässigt werden (Lambrix, 2022, S. 229). Für die Entschlüsselung von Bewertungspraktiken, die auf die Körper der bewerteten Akteure abstellen, ist eine solche Perspektive entscheidend. Solche Bewertungspraktiken finden sich nicht nur im Sport, sondern darüber hinaus zum Beispiel in der Mode. Mears und Finlay beobachten, dass die als oberflächlich wahrgenommene Vermarktung eigener körperlicher Attraktivität bei Models zu einem hohen Maß an kompensatorischer emotionaler Arbeit führt (Mears & Finlay 2005).⁸ Diese kompensatorische Arbeit hängt auch damit zusammen,

6 Dies zeigen Untersuchungen zur Praxis der Notenvergabe in Schulen: Kalthoff, 1996; Kalthoff, 2019; Kalthoff & Dittrich, 2016.

7 Auch in beruflichen Auswahlgesprächen (zum Beispiel eines Mitarbeiters in einem Beratungsunternehmen) kann die körperliche Leistungsfähigkeit als »Nebenaspekt« des erwarteten Leistungsspektrums einer zu bewertenden Person thematisch werden. War die beobachtbare und körperlich gezeigte Nervosität der Kandidatin vielleicht ein Hinweis darauf, dass sie etwas zu verbergen hat? Oder dass sie sich unter Druck generell unwohl fühlt (vgl. Rivera, 2017)? Solche Fragen werden besonders dann relevant, wenn die Dichte fachlich qualifizierter Bewerber derart hoch ist, dass fachliche Qualifikation nicht mehr den entscheidenden Unterschied machen können. Rivera (2015) ist in ihren Studien zu Bewerbungsverfahren in US-amerikanischen Elite-Kanzleien und Unternehmensberatungen auf solche Konstellationen gestoßen. Das »cultural matching« avanciert hier zum Leitprinzip der Bewertungslogik bei der Auswahl von Bewerberinnen (Rivera, 2012) und wertet die situative körperliche Kopräsenz von Bewertenden und Bewerteten im Bewerbungsgespräch als Moment des Austestens von »Vibes« auf.

8 Eine Darstellung der körperlichen Vollzugswirklichkeit einer zentralen Bewertungssituation im Modegeschäft hat Hoppe vor kurzem vorgelegt. Seine empirisch und ana-

dass an Menschenkörpern Personen haften (Goffman 1974) und die von den Bewertungsverfahren unterstellte Trennbarkeit von Körper und Person kaum gegeben ist.⁹

2. Bewertungssoziologie im Feld des Sports

Mit den Themen »Bewertung von Körper(bewegungen)« und »Wettbewerb und Konkurrenz« sind zwei Fluchtpunkte für einen produktiven Austausch zwischen Sport- und Bewertungssoziologie angegeben. Wie fruchtbar der auf diese Themen bezogene Austausch sein kann, zeigen nicht zuletzt auch Studien, die im Schnittfeld von Sport- und Bewertungssoziologie arbeiten und zugleich Bewertung von Körpern(und ihren Bewegungen) und Leistungen sowie Bewertungspraktiken in Kontexten der Organisation des sportlichen Wettbewerbs in den Mittelpunkt rücken. So hat etwa Janetzko (2015), die Auswahlverfahren in Nachwuchs-Nationalkadern im (Paar-)Sporttanz untersucht. Sie zeigt, wie im Nebeneinander verschiedener Tanzgruppen während eines Auswahltrainings eine visuelle Bewertungstechnik zum Einsatz kommt, die schon Michel Foucault (1973) in seiner »Geburt der Klinik« beschrieben hat. Körper werden in bestimmten materiellen Arrangements als Vergleichsobjekte behandelt und so nebeneinander angeordnet, dass Unterschiede erkennbar werden (vgl. auch Prinz, 2014).

lytisch dichte Beschreibung ist nicht nur instruktiv im Hinblick auf die Frage, wie Körperlichkeit als Gegenstand bewertungssoziologischer Forschung thematisiert werden kann, sondern auch, weil es ihm gelingt, seine eigene Körperlichkeit als Forschungsinstrument zur explizierenden Beschreibung von bewertungssoziologisch relevanten Körperpraktiken zu mobilisieren. Hoppe lernt den Blick der Jury des Model-Castings als Praktikant und wächst schließlich in die Rolle eines gleichberechtigten Entscheiders im Feld hinein. Er beschreibt, welche »clues« (Körpergröße, Länge der Beine, erkennbare Routiniertheit im Gehen mit sehr hochhackigen Schuhen) und welche Körpertechniken dabei zum Einsatz kommen (Hoppe, 2022, S. 69). Das Model-Casting simuliert die Laufstegsituation, obwohl diese den Models nur sehr kurz Gelegenheit gibt, sich zu zeigen. Das verlangt von Bewerteten und Bewertenden spezifische Kompetenzen: Halten die Models den prüfenden Blicken vom Casting-Tisch stand (S. 70–71)? Kann der Casting-Agent seine Aufmerksamkeit so fokussieren, dass sich zehn Sekunden subjektiv wie 20 anfühlen (S. 71–72)?

9 Dazu auch die Beiträge im Band von Meier und Peetz (2021) im Abschnitt »Die organisierte Bewertung von Personen«.

Die empirische Fallstudie von Kristina Brümmer (2019) zu Videoanalyse-Praktiken im Nachwuchsleistungszentrum eines professionellen Fußballvereins rückt die Bewertung von individueller Leistung und »Teamfähigkeit« in den Mittelpunkt. Sie beschreibt, wie die kostengünstige Verfügbarkeit von Videokameras und Videobearbeitungssoftware die Bewertungspraktiken im Profifußball transformiert (vgl. dazu auch Schmidt, 2015; Schmidt & Hodek, 2019) und wie die spieltaktische Videoanalyse nicht nur die mediale Präsentation des Fußballsports anreichert, sondern auch im Trainingsalltag Einzug hält. Analysesitzungen mit vorgeschnittenen Videomaterialien und teamöffentliche Demonstrationen an der »Taktiktafel¹⁰ werden von Brümmer als Bewertungspraktiken gelesen, in denen die Leistung des Einzelindividuums als Teil des Körperkollektivs der »Mannschaft« avisiert wird. In »wechselseitig aufeinander Bezug nehmenden Spielanalysen und –planungen« werden Spielerkörper »gezielt teamförmig [auseinander] eingestellt.« Sie sollen »befähigt werden [...], sicher(er) zusammenzuspielen« (Brümmer, 2019, S. 271). Es geht dabei nicht nur um die Schulung und Ausbildung eines taktischen Improvisationsvermögens, das als inkorporiertes Wissen *in situ* fungiert, sondern um die organisationale Bearbeitung von »Koordinationsprobleme[n], die im Spiel nicht qua praktischen Sinns ad hoc gelöst werden konnten« (Brümmer, 2019, S. 271). Die von Brümmer beschriebenen Bewertungspraktiken rekurrieren auf eine technisch gestützte »erhöhte« und zugleich sozial übergeordnete Position und Perspektive: Es wird aus der Vogelperspektive gefilmt und zugleich werden die Bewertungen aus der epistemisch wie hierarchisch übergeordneten Position des Trainerstabs vorgenommen, der auf langfristige und kontinuierliche Beobachtungen und Bewertungen zurückgreifen kann. Dadurch entsteht die Tendenz, das »praktische Wissen« in den Spielerkörpern zu »marginalisieren« (S. 293). Gleichzeitig reflektiert das Feld dieses Problem kritisch als einen Hang zur »Überschematik« (S. 294). Man könnte auch von einer Verschiebung im Kräfteverhältnis zwischen individuellem Körperwissen und formalem Organisationswissen sprechen.

Die neuen Videoanalysepraktiken im Feld des Fußballtrainings ermöglichen neue Formen des körperlichen Zusammenspiels, die ohne diese Technik vorher so nicht realisierbar waren. Videoanalysen fungieren als epistemische Praktiken zur Entdeckung und Kreation taktischer Muster durch die

¹⁰ Ein Spielfeldschema in Vogelperspektive auf das Magnete als Repräsentanten für Spieler positioniert werden, um taktische Formationen zu erläutern (vgl. auch Tuma, 2017: 187ff.).

Trainerinnenriege (Hodek, 2021) und sie erweisen sich im Trainingsalltag als Disziplinierungstechniken, vor allem indem sie für Bewertungssituationen in Mannschaftssitzungen mobilisiert werden und dazu verhelfen, taktische Instruktionen zu vermitteln und deren Missachtung zu sanktionieren. Damit wirken diese Beobachtungs- und Bewertungspraktiken nicht nur auf der Ebene der Organisation des Wettkampfs (wer wird wie bewertet?) sondern auch auf der Ebene des Wettkampfinhalts, insofern sie Wettkampfparteien neue Mittel für den Wettkampf (Taktiken, Trainingsmethoden usw.) an die Hand geben.

Videogestützte Bewertungsverfahren spielen im Sport weiterhin auch auf der Seite der neutralen Dritten (Juries, Ringrichter, Schiedsrichter usw.) eine wichtige Rolle. Stacy Lom hat in einer interessanten Studie die Reform der Jury-Praxis im olympischen Eiskunstlauf empirisch untersucht. Ausgangspunkt ihrer Fallkonstruktion ist ein Abspracheskandal zwischen Jury-Mitgliedern während des olympischen Wettbewerbs 2002. Vor dem drohenden Ausscheiden der Sportart aus den olympischen Spielen sah sich der betroffene Sportverband gezwungen, das Regelwerk in Richtung gesteigerter Transparenz und Nachvollziehbarkeit der Entscheidungskriterien zu verändern (Lom, 2016, S. 39–40.). Das Ergebnis dieser Reform beschreibt Lom als Technisierung der Bewertungspraktiken. Videotechnologie fungiert nun als Grundlage zur Urteilsbildung der Jury. Diese bewertet anhand vorgegebener Kriterien und eines Protokolls. Einerseits sorgt diese Standardisierung der Bewertungspraktiken für die gewünschte Transparenz und Berechenbarkeit der Ergebnisse, andererseits schränkt die Technisierung der Bewertungspraktik die sportlich-ästhetischen Ausdrucks- und Bewegungsregister der Sportart aber auch massiv ein.

In ähnlicher Weise haben videogestützte Beobachtungen und Bewertungen auch das Schiedsrichten im Profifußball verändert und auf neue Weise strittig gemacht. Die Bildregie der Fußballübertragungen nutzt multiple Kamerawinkel, Schnitttechnik, Wiederholung und Zeitlupe und unterstützt damit einen kritisch evaluierenden Fehlerblick des Publikums (Colwell, 2004, S. 291ff.) auf die Arbeit des Schiedsrichtergespanns auf dem Rasen. Dies hat zu einer Destabilisierung der »epistemic authority« des Schiedsrichters geführt (Collins et al., 2016). Schmidl hat die verschiedenen Formen des Umgangs mit diesen Entwicklungen und Bereitstellungen von visuellen Expertisen für die Fälle Tour de France und Profifußball vergleichend im Rahmen einer »visual discourse analysis« betrachtet. Der dem Schiedsrichtergespann mittlerweile beigeordnete Video Assistant Referee (VAR) erscheint aus dieser Perspektive

als Reaktion auf ein bewertendes fehlerfixiertes visuelles Beobachtungsdispositiv der Medien und als Versuch einer Aneignung von Diskurshoheit im Umgang mit dem Bildmaterial (Schmidl, 2023).

In den genannten bewertungs- und organisationssoziologischen Studien zur Arbeit von Juries, Kampf- und Schiedsrichterinnen im Sport verbindet sich das Interesse an der Organisation von Wettbewerbsformen mit Fragen nach der Bewertung von Körpern und körperlichen Bewertungspraktiken. Weigelin hat die visuellen Praktiken und Bewertungslogiken rund um die Rolle des Fußballschiedsrichters untersucht. Als Mitläuferin auf dem Spielfeld übt die Schiedsrichterin ihr visuelles Engagement in einem »Positionsspiel« aus, durch das sie sich als Augenzeugin bei Fouls und ähnlichen Spielereignissen zu platzieren versucht (vgl. Weigelin, 2022, und ders. in diesem Band). Schiedsrichter werden demnach von der Logik des Spiels dafür zuständig gemacht, basale Kategorisierungsleistungen zu erbringen, die die situative Organisation der Wettkampfform betreffen: War der Ball nun drin oder nicht? Muss die Spielerin für das Foul vom Platz verwiesen werden (rot) oder reicht eine Verwarnung (gelb)? Je mehr der Schiedsrichterin durch Diskussionen rund um ihre Spielleitung ein problematisches Image erwächst (»zu kleinlich«, »blind«, »überfordert« usw.), desto wahrscheinlicher wird es, dass ihre nächsten Kategorisierungen und Interventionen (»Foul«, »Aus«, »Tor« usw.) durch Bewertungskommunikation als entscheidungsförmig hervorgebracht werden und umstritten sind. Die Schiedsrichterin wird dann von einer tendenziell unsichtbaren Agentin des situativen Wettkampfs zu einer Mitspielerin wider Willen und in ein »Spiel im Spiel« (Gebauer, 2006, S. 146) verwickelt. Durch fußballtypische kommunikative Bewertungspraktiken wie »Reklamationen« und »Proteste« einerseits und durch die der Fußballspielform eigenen Unentscheidbarkeiten (etwa Tor oder kein Tor im Fall des

»Wembley Tors«?¹¹) und Fehlwahrnehmungen (Maradonas »Hand Gottes«¹²) des Schiedsrichters wird das Fußballspiel zu einem Quell dramatischer Wendungen (man denke an »spielentscheidende Fehlentscheidungen«), massiver Frustrationen, Ärgernisse und Diskussionen. Der Fußball dramatisiert die Kontingenz der Bewertungen und Entscheidungen seines neutralen Dritten und stellt die »Fairness« seines Wettbewerbs notorisch unter Vorbehalt. Der Leistungsbegriff des Fußballs ist also gerade nicht »rein«, sondern gekennzeichnet von der Hinnahme des Zufalls (vgl. auch Bromberger, 2003, S. 286) und dem Hinzutreten von List, Durchtriebenheit und Tücke als Bestandteilen der Spielerinnenkompetenz (vgl. Gebauer, 2006, S. 145ff.).

In diesem Zusammenhang zeigen die bewertungsssoziologischen Studien im Feld des Sports nicht zuletzt auch auf, wie der Sport in seine Wettbewerbs- und Konkurrenzarrangements Ambivalenzen und Relativierungsmöglichkeiten einbaut. So hat sich der Fußball möglicherweise gerade deshalb als eine hegemoniale Sportart in der modernen Kultur etabliert, weil sich hier die Fehleranfälligkeit der bürokratischen Regelsysteme und der organisationalen Durchsetzung ihrer Geltung besonders gut dramatisieren, anfechten und in Frage stellen lassen (vgl. Bromberger, 2003). Der Fußball lässt sich als kulturelle Aufführung einer Gesellschaft lesen, die den Glauben an den Sinn des Wettbewerbs und die Geltung seiner Regeln hochhält und zugleich um deren Konstruktionsmängel und ambivalente Folgen weiß. Gerade im Moment der Niederlage lässt sich die Imperfektion der Bürokratie des Spiels besonders leidenschaftlich anprangern. In einer solchen Interpretation des

¹¹ Als Wembley-Tor wird ein Lattentreffer bezeichnet, bei dem der Ball von der Unterkante der Torlatte nach unten springt und anschließend wieder ins Spielfeld springt, wobei jedoch nicht klar ist, ob der Ball nun einmal die Linie überquert hatte oder nicht. Namensgebend für diese Bezeichnung ist ein Tor der englischen Fußballnationalmannschaft in der Verlängerung des Finales der Fußball-Weltmeisterschaft 1966 gegen Deutschland im Wembley-Stadion. Die TV Bilder konnten die Frage – Tor oder nicht Tor? – auch nachträglich nicht aufschlüsseln. Der Legende nach sind die Engländer überzeugt, es handele sich um eine Fehlentscheidung, während man es in Deutschland gerade umgekehrt sehe.

¹² Während der Fußball-Weltmeisterschaft 1986 benutzte der berühmte und kleingewachsene argentinische Spielmacher Diego Maradona (auf sehr offensichtliche Weise) seine Hand in einem Kopfballduell, um ein irreguläres Tor zu erzielen. Der Ausdruck geht auf Maradonas Worte selbst zurück. Nach dem Spiel gab er wenig reuig der Sportpresse zu Protokoll: »Es war ein bisschen Maradonas Kopf und ein bisschen die Hand Gottes«.

Fußballs werden bewertungssoziologische Ansätze und Perspektivierungen des Wettkampfgeschehens kulturanalytisch erweitert.

3. Für eine kulturanalytische Bewertungssoziologie im Feld des Sports

Die kulturanalytische Ausrichtung sportbezogener bewertungssoziologischer Fragestellungen kann an entsprechende Forschungsrichtungen der Sportsoziologie anknüpfen. Sie reichen von videografisch-konversationsanalytischen Untersuchungen von Kinderspielen (Streeck, 2021), über (neo-)phänomenologische und ethnometodologische Mikroanalysen einzelner sportlicher Praktiken (Meyer, 2021) bis hin zu auf den Sport fokussierten kulturtheoretischen Studien zur Kultur der Moderne (Alkemeyer, 2017).¹³ In solchen Studien wird mitunter eine »vom Spiel ausgehenden Soziologie« (Callois, 1982) im Rekurs auf Denkfiguren wie die der Mimesis (Gebauer & Wulf, 1998) entwickelt. Der Sport stellt in dieser Perspektive ein »kulturelles Genre der Aufführung und praktischen Interpretation von außersportlichen sozialen Praxen und Realitäten dar« (Alkemeyer, 1997:, S. 368). Daraus ergibt sich eine produktive Neuperspektivierung: Statt – wie in der bewertungssoziologischen Forschung etwa zum Bildungs- und Wissenschaftsfeld üblich – zu untersuchen, wie ökonomische Bewertungsregime nicht-ökonomische Bereiche wie den Sport kolonisieren und transformieren, kann nun umgekehrt danach gefragt werden, wie sportliche Praktiken in ihrem mimetisch-transformativen Spiel mit Materialien der Kultur nicht zuletzt auch Bewertungsverfahren umarbeiten, artikulieren und problematisieren. Durch diese Neuperspektivierung kommen die Eigenlogik und die kulturelle Darstellungsfähigkeit sportlicher Praktiken (Gebauer & Wulf, 1998) in den Blick.

Mit Bezug auf den Fußball lässt sich diese Neuperspektivierung konkretisieren: Eine mimetische Beziehung zwischen dem Fußball und anderen sozialen Welten anzunehmen impliziert, dass das Fußballfeld anderen sozialen Feldern nicht nachgeordnet ist, sondern – ähnlich wie die Kunst – einen Bereich bildet, in dem durch Noch-Einmal-Machen und Umarbeiten eine eigene Version der Welt erzeugt (Goodman, 2013), oder – mit Luhmann (2017) gesprochen – Realität verdoppelt wird. Als eigenlogische Version der Welt

¹³ Vergleiche für einen Überblick über dieses Spektrum von Ansätzen den Band von Brümmer et al., 2021.

bildet der Fußball ein *social drama*, eine, vielleicht sogar die zentrale *cultural performance* (Singer, 1972) nicht nur der europäischen Gegenwartsgesellschaften. Hier werden nun – unserer analytischen Idee zufolge – insbesondere auch Bewertungspraktiken und -verfahren, auf denen die fußballtypischen Wettbewerbs- und Konkurrenzformate aufbauen, sowie Spiel-, Spielerinnen- und Leistungsanalysen hochgespielt, dramatisiert und angefochten. Der Fußball fungiert wie eine Bühne für solche Dramatisierungen und bietet dafür besondere Voraussetzungen, denn der – im Vergleich mit anderen Sportspielen äußerst große – Stellenwert des Zufalls beim Zustandekommen von Spielergebnissen¹⁴ und die zentrale Rolle eines differenzierten körperlichen fußballerischen Könnens sowie komplexe, auf der Grundlage von praktischem Spielsinn und Spielverständnis mögliche Koordinations- und Kooperationsprozesse konterkarieren und unterminieren die Schematismen des Erfassens, Bewertens und Prognostizierens von Leistung und Erfolg besonders nachdrücklich. Es werden fortlaufend Konzeptualisierungen, Kategorisierungen und Bewertungsformate ins Spiel gebracht, die ›naturgemäß‹ strittig bleiben müssen. Sie werden mit Bezug auf alternative Modi, in denen das körperliche Spielgeschehen beobachtet, erkannt und gewusst wird, in Frage gestellt. In der Welt des Fußballs wird auf diese Weise ein findiges und methodenkritisches Fußballwissen kultiviert.

Wir wollen unsere kulturanalytische Perspektive auf die Dramatisierung von Bewertungsverfahren im Fußball, die wir am Ende des vorangegangenen Abschnitts schon am Beispiel der Organisation des Schiedsrichters angerissen haben, an zwei weiteren Fällen aus unserem Forschungsprojekt zum Profifußball – neuartigen digitalen Spiel- und Leistungsanalysen und Entwicklungen

14 Weil in einem Fußballspiel viel weniger Tore fallen als z.B. Körbe im Basketball (einem Sportspiel, in dem statistische Analysen bereits viel früher entwickelt wurden), ist eine in statistischen *key performance indicators* gemessene Teamleistung häufig relativ unabhängig vom Spielergebnis. KPI-basierte Analysen können im Fußball daher besonders dramatisch fehlgehen. Exemplarisch dafür steht das Champions-League-Spiel Celtic Glasgow gegen den FC Barcelona vom 7. November 2012. Barcelona dominierte die Partie scheinbar nach Belieben: 84 Prozent Ballbesitz, 25:5 Torschüsse, 955 Pässe und eine überragende Passquote von 91 Prozent. Celtic brachte es nur auf 166 Pässe und lieferte 38 Prozent Fehlpässe ab. Doch die Schotten gewannen 2:1 und der Anschlusstreffer von Barcelona fiel erst in der Nachspielzeit (vgl. SZ vom 10. 11. 2017, abgerufen unter <https://www.sueddeutsche.de/sport/fussball-drei-neue-parameter-fuer-fussball-scouts-1.3742095>).

des Transfermarktes – empirisch erläutern. Beide Fälle betreffen Entwicklungen, die auf den ersten Blick als von außen, das heißt durch sportfremde, ökonomische Bewertungsregime, in Gang gesetzte Transformationen des Profifußballs verstanden werden könnten. Eine kulturanalytisch instruierte empirische Perspektive kann jedoch andere Lesarten dieser Neugestaltungen entwickeln.

4. Digitale Spiel- und Leistungsanalysen im Profifußball

Obwohl auch der professionelle Fußball in den letzten zwei Jahrzehnten Verwissenschaftlichungsschübe erfahren hat, unterscheidet er sich diesbezüglich von anderen Sportarten und Sportspielen wie zum Beispiel Baseball oder Basketball und kann – trotz seiner außerordentlich großen ökonomischen Ressourcen – als ein Nachzüglerfeld gelten, das lange durch eine gewisse Reserve in Bezug auf die Nutzung von digitalen Spiel- und Leistungsanalysen gekennzeichnet war. Mittlerweile beschäftigen die Vereine und Verbände des Profifußballs jedoch große Expertenstäbe, die Transferentscheidungen, Trainingssteuerung und auch medizinische Belastungskontrolle kaum noch ohne den Rückgriff auf statistische Daten als plausibel darstellen können. Diese Expertenstäbe bestehen aus Data Scientists und Daten-Scouts, Spielanalytikern und Performance Analysts, Sportmedizinerinnen und Athletiktrainern und bearbeiten breit gestreute Aufgaben rund um die Spiel-, Spieler- und Leistungsanalyse (vgl. Schmidt et al., 2017; Schmidt & Hodek, 2019).¹⁵

Die Produktion und Verwendung von *soccer metrics* im Profifußball erwächst zunächst aus technologischen Innovationen im Fahrwasser der Digitalisierung, die das Sammeln und Aggregieren großer Mengen von Daten ermöglichen (vgl. Memmert, 2016). Wichtig sind dabei vor allem Verfahren der automatisierten Generierung von Echtzeitdaten, sogenannter *positional tracking data*. Diese Daten können entweder GPS basiert erstellt werden, wofür Spieler mit *tracking devices* ausgestattet werden müssen, oder durch die algorithmische Auswertung von Videodaten.¹⁶ Zur Verarbeitung solcher

¹⁵ Auch im Jugend- und Amateurbereich sind zum Beispiel mittlerweile integrierte technische Anwendungen weit verbreitet, die Kamerasysteme für das optische Tracking mit Software für die automatische Ereignisanalyse verknüpfen und so automatisch Daten für die anschließende Mannschaftsbesprechung zusammenstellen.

¹⁶ Das setzt dann freilich voraus, dass Fußballspiele massenhaft videoaufgezeichnet werden. Diese Entwicklung setzt mit der Intensivierung von TV-Übertragungen in der

Daten werden statistische Konzepte entwickelt: Im Feld spricht man betriebswirtschaftlich-managerial von *key performance indicators*. Solche Kennzahlen reichen von einfachen, auf diskrete Ereignisse bezogenen Indizes (Schüsse, Pässe, gewonnene Zweikämpfe etc.) bis hin zu komplexeren Bewertungskonstrukten, die auf taktischen Überlegungen gründen. Ein solches Konstrukt ist z.B. die *packing rate*, d.h. die Konzeptualisierung von Passeeffektivität hinsichtlich der Anzahl mit einem Pass überspielter Gegenspielerinnen oder Indikatoren für das sog. »Umschaltverhalten«¹⁷ einer Mannschaft bis hin zu – insbesondere für die Wettindustrie interessanten – *predictive models* wie *expected goals*. Diese Modellierungen werden getestet, verworfen, neuentwickelt. Sie stehen fußballöffentlich in der Kritik und lösen performativ Effekte aus, d.h. sie führen zum Beispiel zu an ihnen orientierten Trainings-, Spiel- und Rekrutierungsstrategien und Kaderplanungen (die dann wieder in Frage gestellt werden).

Bewertungssoziologisch lassen sich diese Entwicklungen zunächst gut einordnen, insofern sie sich als »Teil des Quantifizierungsschubs, der die (westlichen) Gesellschaften seit dem 18. Jahrhundert erfasst hat« (Heintz, 2021b, S. 170), verstehen lassen. In der Sportgeschichte ist die Sportifizierung westlicher Bewegungskulturen im 19. Jh. als ein Rationalisierungsschub beschrieben worden, in dem sich eine Lust am Quantifizieren artikuliert (Guttmann, 1978, S. 47ff.). Wir kommen jedoch zu einem differenzierten Verständnis des aktuellen Quantifizierungsschubes rund um »Big Data«, wenn wir – orientiert an der skizzierten kulturalistisch-empirischen (Neu-)perspektivierung – untersuchen, wie genau der professionelle Fußball

Nachkriegszeit durch große Sendeanstalten ein und intensiviert sich von dort aus kontinuierlich. Neuerdings hat diese Erfassung von Spielen via Video aber noch einmal einen neuen Schub erhalten, weil die Videotechnik günstiger und einfacher zu handhaben ist, sodass interessierte Experten sich mit Videoaufzeichnungen von Spielen aus Ligen aller Welt und in Ländern wie Deutschland teilweise bis tief ins Amateurniveau hinein versorgen können.

17 Das »Umschalten« ist ein Fokuspunkt der Taktik-Debatten der letzten Jahrzehnte. Ausgangspunkt ist die Annahme, dass Mannschaften insbesondere dann erfolgreich sind, wenn sie gut darin sind, die Momente des eigenen Ballverlusts oder des Ballverlusts des Gegners auszunutzen. Hintergrund ist die Annahme, dass in diesem Moment des »Umschaltens« bzw. der Anpassung von einer Situation (Ballbesitz) an die andere (Verteidigung) eine Unordnung in den taktischen Choreografien der Mannschaftskörper existiert.

diese neuartigen Verfahren der digitalen Leistungserfassung und -analyse aufgreift und transformiert.

Mit der Verbreitung digitaler statistischer Spiel- und Leistungsanalysen hat sich im Profifußball eine Gruppe von Expertinnen herausgebildet, die das Fußballspiel in neuer Weise wissensfähig machen. Sie sind für die sportliche und ökonomische Konkurrenzfähigkeit der Vereine von entscheidender Bedeutung, werden mittlerweile über von privaten Anbietern und Instituten organisierte Fortbildungen, in einem entsprechenden Studiengang an der Sporthochschule Köln ausgebildet oder kommen über ein Engagement im Nachwuchsbereich der Vereine und in den Nachwuchsleistungszentren der Verbände in diese Tätigkeiten.

Wir haben untersucht, wie genau diese Experten arbeiten, und konnten nachzeichnen, wie sie alltäglich von Datendienstleistern eingekauft, automatisiert generierte Spiel- und Spielerdaten aufbereiten und sie mit über proprietäre Internetportale (z.B. *Wyscout*) bereitgestelltem Videomaterial und eigenen Spielbeobachtungen (Live-Scouting) abgleichen. Für uns besonders interessant war, das *Soziologisieren*¹⁸ nachzuvollziehen, durch das die alltägliche Arbeit gekennzeichnet ist: Die Arbeit der Scouts- und Videoanalysten ist in die vereinsinternen Arbeitsprozesse eingebettet. Sie erarbeiten zum Beispiel fortlaufend Gegner-Analysen, stimmen sie im Trainerstab ab, koppeln sie an das vom jeweiligen Cheftrainer eingeführte Sprachspiel und präsentieren sie in den Teamsitzungen. Zudem müssen sie im Rahmen der Kaderplanung ihre Einschätzungen und Bewertungen gegenüber der Geschäftsführung plausibilisieren, geplante Transfers durch entsprechende datengestützte Analysen und Einschätzungen untermauern. Sie müssen *shadow teams* konzipieren und prospektiv, mit Bezug auf Video- und Analysedaten und archivierte Scoutingberichte sowie aus wiederholten Vor-Ort-Beobachtungen von Spielern über längere Zeiträume hinweg Entwicklungspotentiale von in Frage kommenden Neuverpflichtungen ableiten. Dafür ist nicht zuletzt auch eine erfahrungs-basierte Taxonomie und Sozioanalyse entscheidend, die Informationen zur Vertragssituation eines in Frage kommenden Spielers, gerüchteweise kolportierte Hinweise auf die familiäre Situation oder persönliche Macken sowie in eigenen Beobachtungen gewonnene Eindrücke in Bezug auf Verletzungs-anfälligkeit, ›Mentalität‹, ›Charaktereigenschaften‹ und Sekundärtugenden zusammenzieht.

¹⁸ Vgl. zum gesellschaftlichen Teilnehmerinnen-›Soziologisieren‹ Scheffer & Schmidt, 2023.

Scouts und Spielanalystinnen stehen dabei unter Erfolgsdruck. Ihre Arbeit wird nicht nur durch Geschäftsführung und Trainerstab, sondern schon durch den nächsten Spieltag, die Entwicklung des eigenen Teams in der Liga etc. fortlaufend validiert (und gegebenenfalls falsifiziert). Sie müssen zu triftigen Einschätzungen kommen, können aber nicht auf vorgegebene Verfahren zurückgreifen und entwickeln in ihrer alltäglichen Arbeit selbst einen erfahrungsbasierten praktischen Sinn und methodenreflexive Vorgehensweisen, um Geltungsansprüche, Validität und Plausibilität der automatisiert generierten Daten und die Möglichkeiten und Grenzen ihres Gebrauchs kritisch einzuschätzen.¹⁹

Wir wollen das an der Arbeit des Scoutings nun noch ein wenig detaillierter darstellen: Bewertungen auf der Grundlage automatisiert generierter Daten werden im Live-Scouting relativiert und befremdet.

Also Videodaten und so, ist alles super. Aber der Liveeindruck ist unverzichtbar aus meiner Sicht [...]. Aufm Video, klar, du siehst schon viel [...]. Aber wie reagiert der Spieler gerade, im Stadion, in ner Atmosphäre, wo es unruhig wird. Wenn ich die Reaktion der Zuschauer merke, wenn ich dann sehe: Oh, wie verhält sich jetzt der Spieler? Wird er nervös, wenn er nen Fehler gemacht hat? (Bierwagen 2023)

Zugleich werden aber auch die Fußballmetriken genutzt, um das epistemische Privileg der Vor-Ort-Beobachtungen zu irritieren und in den Beobachtungen möglicherweise festgefahrene Eindrücke und (Spieler-)Bilder wieder zu verflüssigen.

Wenn ich fünf, sechs mal nen Spieler schaue, irgendwann hast Du dann halt n Bild [...]. Du liest nen Spielernamen und hast direkt n Bild vor Augen und ich seh nicht noch oder wenig neue Dinge. Durch die Daten habe ich dann aber einfach noch ne andere, ne weitere Einschätzung.

Fragen der Validität und Aussagekraft der digitalen Datentypen sind in der Arbeit der Scouts und Spielanalysten stets virulent. Es wird fortlaufend kritisch

¹⁹ Die Ethnomethoden der reflektierten Generierung und Triangulation heterogener Datentypen, die die Arbeit von Scouts und Analystinnen kennzeichnen hat Jakob Bierwagen in seiner Masterarbeit *Scouting und Videoanalyse im Fußball* (Fachgebiet Soziologie, KU Eichstätt 2023) herausgearbeitet.

reflektiert, wie genau eigentlich die Leistungen und Qualitäten eines (möglicherweise für die zukünftige Kaderplanung in Frage kommenden) Spielers durch die verfügbaren Daten und die mit ihnen verbundenen Konzepte (KPIs) abbildungbar sind.

Hab ich jetzt ne Hunderprozent-Passquote, weil ich nur quer spiele? Weil ich nur ohne Druck zurückspiele? Interessanter ist ja: Spiele ich vertikal nach vorne? Hab ich Risikopässe dabei? [...] n Mittelfeldspieler, der ne Hundertprozent-Passquote hat is ja nicht zwingend besser als einer, der fünfundsiebzig Prozent hat, weil der spielt vielleicht mehr Risikobälle [...] und hat n größeren Impact auf das Offensivspiel.

Die Arbeit der Scouts- (und Analystinnen) ist durch eine kritisch-reflexive und bisweilen experimentelle Findigkeit gekennzeichnet, die in die Entwicklung der Analysesoftware zurückfließt. Dafür lieferte unsere Forschung Beispiele. Dazu zählt etwa der Vorschlag, eine Variable wie *Passquote* in die Subvariablen *vertikale Pässe*, *horizontale Pässe* und *Rückpässe* zu differenzieren, oder die Idee, mit Bezug auf die konkrete Situation und die Notwendigkeiten und Desiderate eines Teams spezifische Positionsprofile zu konzipieren und zu programmieren, mit denen dann in allen verfügbaren Datenbanken und Ligen in Frage kommende Spieler gesucht werden können.

Im empirisch aufgezeigten praktischen Gebrauch der neuartigen Bewertungsverfahren wird ein methodenkritisches, kompetentes und findiges Fußballwissen deutlich, das die Scouts, Spiel- und Videoanalystinnen mit den kompetenten Fans und Nerds, aus deren Reihen sich viele Experten rekrutieren, gemeinsam haben. Die neuen Spiel- und Leistungsanalysen werden kulturanalytisch als Wissensmodi verständlich, in denen sich die Welt des Fußballs selbst thematisiert und dabei zugleich auf andere soziale Welten Bezug nimmt, an denen ihre Mitglieder teilhaben. Es zeigt sich eine Strittigkeit der Formate und ihrer Resultate sowie ein praktischer Experimentalismus, der im Fußball in der Auseinandersetzung mit den neuen datengestützten Analyse- und Bewertungsverfahren kultiviert wird. Der Fußball arbeitet die Bewertungsverfahren rund um »Big Data« um, problematisiert und untermiint sie und führt gleichzeitig ihre immense Bedeutung sowie ihre Begrenzungen vor.

5. Der Transfermarkt als ein Spiel im Spiel

Die Entwicklung des Transfermarktes ist ein weiterer Fall, an dem wir die Trifigkeit unserer kulturanalytisch-empirischen Perspektivierung von Bewertungspraktiken im Profifußball verdeutlichen möchten. Man kann die Entwicklung des Transfermarktes zu einer Dichtezone, in der verschiedene Bewertungslogiken aufeinandertreffen als eine vor allem auch digital getriebene Neubildung, als eine Welt in der (Fußball-)Welt beschreiben, denn durch die Entwicklung neuer, digitalisierter Beobachtungs- und Bewertungsformate hat sich die Kategorisierung und Bewertung von Spielern enorm ausgeweitet und globalisiert. Sie erstreckt sich weit in die unteren Spielklassen und in den Nachwuchsbereich und sie umfasst immer mehr nationale und regionale Ligen.

In der Welt des Fußballs gilt der Transfermarkt einerseits als problematisch: Die Kritik an Vereinen, die mitunter dreistellige Millionenbeträge verschieben, sowie an hochfrequenten Vereinswechseln von als ›Söldnern‹ wahrgenommenen Spieler, die ihren Wechselwunsch zusätzlich medial in Szene setzen und forcieren, ist allgegenwärtig. Die Fußballwelt beobachtet den Transfermarkt einerseits also als eine Deformation des Sportlichen. Andererseits hat sich der Transfermarkt aber auch als eine zusätzliche Spielebene und als eigenständiger sportlicher Wettbewerb etabliert.²⁰ Vereine agieren hier zugleich ökonomisch und sportlich (Magee et al., 2005). Im Transfer-Wettbewerb konkurrieren finanzi starke Eliteclubs mit Traditionsvereinen, die den ›Kommerzclubs‹ Paroli bieten, durch kluges Scouting früh Talente erkennen, an sich binden, ausbilden und gegebenenfalls gewinnbringend verkaufen, um dann zum Beispiel einen spektakulären Transfer finanzieren zu können. Entsprechende Management-Strategien schießen demnach zwar keine Tore, sie machen aber auch bei geringeren Budgets den sportlichen Erfolg wahrscheinlicher. Ein gutes Abschneiden in der Liga gilt entsprechend auch als Resultat eines klugen Managements.

20 In Simulationsspielen wie *Football Manager* wird dieser globalisierte Markt auf der Grundlage von jeweils nach den Transferfenstern aktualisierten Spielerdaten modelliert, ausgearbeitet und für eine Community von Spielern spielbar gemacht, die dazu in die Rolle der sportlichen und geschäftlichen Managerin eines Proficlubs schlüpfen. Vgl. dazu das ethnografisch angelegte Dissertationsprojekt *Doing Realism* von David Kempf.

Sportmedien und Fans beobachten das Geschehen auf dem Transfermarkt entsprechend in der Haltung eines gespannten Sportpublikums: Welcher Verein kauft welchen Spieler? Wie möchte man sich künftig ausrichten, inwiefern vielleicht die eigene Taktik umstellen? Welcher Verein kann aufgrund sportlicher Erfolge in der abgelaufenen Saison oder besonders lukrativer Spielerverkäufe sein Budget kräftig erhöhen – und wird dann eine kluge Reinvestition gelingen? Solche Fragen treiben nicht nur Sport- und insbesondere in der Form des Gerüchts Boulevardmedien um, sondern auch die Fans. Beim Klatsch und in Online-Foren wird mitunter heftig diskutiert: Verhält sich das Management unseres Vereins geschickt? Passen die neu verpflichteten Spieler zur mannschaftstaktischen Ausrichtung? Sind sie jung und vielversprechend oder werden mal wieder nur alte Spieler verpflichtet, die ihren Zenit doch längst überschritten haben – anders ausgedrückt: Priorisiert das Management als Ziel langfristig angelegte Wertsteigerung oder kurzfristigen sportlichen Erfolg? Bewertung ist hier offensichtlich allgegenwärtig: Das Management muss auf die Bewertungen von Scouts zurückgreifen und diese Bewertungen ins eigene ökonomische Feld übersetzen und danach handeln. Medien und Fans wiederum bewerten diese Entscheidungen und greifen dabei auf eigene Bewertungsversuche zurück.

Mit Bezug auf an den Fußball anschließende Videospiele (vgl. dazu die Beiträge von Martin Lüthe und David Kempf in diesem Band) lässt sich nun zeigen, wie die Bewertungspraktiken des Scoutings und der Marktwertanalyse von Fans aufgegriffen, transformiert, dramatisiert und ‚gespielt‘ werden. Sport-Videospiele sind durch einen realistischen Anspruch an Repräsentation gekennzeichnet (Crawford & Gosling, 2009). In besonderem Maße gilt das für die hochkomplexe *Football Manager* Reihe (Crawford, 2006; Hocquet, 2016; Kempf, 2017). Auf dem Markt positionieren die Hersteller (Sports Interactive) das Spiel als besonders realistisch. Mit dem Slogan »Play the Game, Know the Game« forcieren sie dabei die Idee, dass das Spiel als Lernmedium fungiert und weisen stolz darauf hin, dass Profiscouts, Manager und Trainerinnen das Spiel nicht nur spielen, sondern auch seine Datenbank nutzen, teilweise sogar für exklusiven Zugang bezahlen.

Zentral für die Attraktivität des Spiels ist daher nicht nur die Plausibilität der enthaltenen Daten, sondern auch ihre Detailtiefe: In über 40 Attributen bekommen die Fußballspieler jeweils einen numerischen Wert zwischen eins und 20 zugewiesen. Als Videospielerin ist es Aufgabe, die eigene Mannschaft zu managen, also die Taktik einzustellen, Spieler zu kaufen und zu verkaufen, das Training festzulegen etc. Die eigenen Spieler steuern – wie etwa in der po-

pulären *FIFA*-Reihe – kann man indes nicht. Hier bekommt die Simulation ihren prominenten Platz: Die Fußballpartien lassen sich simuliert anschauen, was wichtige Informationen dazu liefert, ob die eigenen Spieler sich auf dem Feld so verhalten, wie man sich das beim Erstellen der Taktik vorgestellt hat – und welche Auswirkungen das im Verhältnis zum Spiel der gegnerischen Mannschaft hat. Über 40 Attribute pro Spieler entwickeln bereits ein erhebliches Komplexitätspotential, das sich angesichts zehn anderer Mitspieler und einer Vielzahl von einzel- undmannschaftstaktischen Einstellungen erheblich steigert. So entsteht eine nicht-triviale Maschine, die es ausgehend vom Zusammenspiel numerisch festgelegter Attribute und der ›zum Leben erweckenden‹ Simulation ermöglicht, im Sinne einer klassisch naturwissenschaftlichen Simulation ›mit den eigenen Theorien zu experimentieren‹ (Dowling, 1999).

Was wäre, wenn mein Herzensverein nur noch junge, talentierte Spieler verpflichtet? Was wäre, wenn die deutsche Nationalmannschaft gänzlich ohne Stürmer spielen würde? Das Fußballfans wohlbekannte Was-wäre-wenn-Spiel erhält so innerhalb des Videospiels Antwort. Das Spielen von *Football Manager* knüpft auf diese Art an eine unter Fans bereits existierende trräumerische Praktik an und führt sie fort (vgl. Pethes, 2018). Aus Interviews, Foren-einträgen und Daten teilnehmender Beobachtung lässt sich eine Vernetzung von Spielpraktiken rekonstruieren, die Grenzen von virtueller Spielrealität am PC und Alltag durchkreuzen. Ausgehend von der Planung beispielsweise einer Taktik, die sich selten auf das direkte Spiel vor dem laufenden PC beschränkt, sondern sich ausweitet auf das Erstellen von handschriftlichen Notizen, Tagbüchern, über die Übersetzung dieser Ideen in die vom Spiel via Interfaces zur Verfügung gestellten Handlungsoptionen und den anschließenden Härte-test in simulierten Partien, bis hin zur Weiterverwendung der dabei gewonnenen Erfahrungen in Diskussionen im Online-Forum des eigenen Vereins: ›Ich empfehle Spieler xy, der hat in meiner Karriere im *Football Manager* super ins Team gepasst.‹ All diese angerissenen Zwischenschritte basieren auf der Übernahme und eigenen Anwendung der zuvor beschriebenen Bewertungspraktiken: Wie selbstverständlich denken *Football-Manager* Spielerinnen über Fußballspieler im Register der im Videospiel vorhandenen Attributskategorien – einschließlich der Möglichkeit, diese quantitativ zu bestimmen. Insofern diese zunächst im Videospiel eingebütteten Denkstile Eingang finden in die Art, wie Fußballspiele angeschaut und darüber in Foren diskutiert wird etc., beginnen sie, performativ Einfluss zu nehmen auf das mimetisch verdoppelte Feld des Profifußballs selbst.

Damit entfalten die aus Quantifizierung und Management stammenden Diskurse und Praktiken eine überraschend kreative, das phantasievolle Spekulieren von Fußballfans über die Zukunft ihrer Vereine anreichernde Wirkung. Gerade ihre scheinbare Eindeutigkeit, die sich im Hinblick auf die bewerteten Fußballspieler durchaus als deterministisch und einengend deuten ließe, ermöglicht im Zusammentreffen von Simulation und Spiel eine offene, kreative Spielerfahrung. Mit den numerischen wie ökonomischen Bewertungsregistern und -praktiken selbst wird so *herumgespielt*: das vermeintlich kalte Netz, auf die Individualität der Fußballspieler gelegt, wird wieder aufgenommen, zerknüllt, verschoben, anders wieder aufgelegt, versuchsweise auf sich selbst und andere angewendet, lächerlich gemacht und ad absurdum geführt, schließlich pedantisch noch zu verbessern gesucht – wie könnte es da am Ende gleichgeblieben sein?

Fazit: Die Darstellungsfähigkeit des Sports und die Unschärfe von Bewertungsarrangements

Der Sport bildet historisch wie gegenwärtig ein Experimentierfeld für die Idee des Wettbewerbs²¹ sowie für soziale und kulturelle Konstruktionen von Konkurrenz und darauf bezogene Bewertungslogiken und -verfahren (Fairness, Leistungserfassung und -differenzierung etc.). Wettbewerbsformate und Bewertungslogiken werden im bzw. durch den Sport ausgestellt, dramatisiert, attraktiv und kritisierbar gemacht. Wir haben in unserem Forschungsprojekt entsprechend auszuleuchten versucht, wie die Welt des Fußballs Praktiken des Bewertens, Kategorisierens, Quantifizierens, Differenzierens und Entscheidens aufgreift, vereinfacht, zuspitzt, re- und dekonstruiert und in eigentümliche und eigenlogische kulturelle Formen bringt. Digitale Fußballmetriken werden auf diese Weise zum Material, an dem sich ein *Soziologisieren* von Trainerinnen, Scouts, Analysten, Journalistinnen, Fans und anderen Teilnehmerinnengruppen entzündet. An den neuartigen Bewertungsverfahren des Profifußballs lassen sich zwar auch ›Eigenlogiken des Digitalen‹ und durch sie in Gang kommende Veränderungen erkennen, entscheidend sind aber die findigen Gebrauchsweisen digitaler Analysetools, die Kritik der Konzepte

²¹ Schon Pioniere der Kulturtheorie des Spiels wie Huizinga und Callois haben den Sport in ähnlicher Weise interpretiert, ihre Beobachtungen aber durch kulturkritische Ambitionen überspitzt.

und Methoden und die daraus hervorgehenden reflexiven, explorativen und spekulativen Wissensformen.

Die kulturelle Darstellungsfähigkeit des Fußballs erkennen wir darin, dass es ihm nicht nur gelingt, die Plausibilität von Konkurrenz und Wettkampf zu beglaubigen, sondern auch spannungsreiche, widersprüchliche und inkonsistente Verbindungen von kulturellen Wertzuschreibung (Zufall, Schicksalhaftigkeit, Ehre etc.) und Leistungsidealen hervorzubringen. Der Fußball propagiert die Konstruktion, Erfassung und Messung von Leistungen *und* den skeptischen und kritischen Umgang mit diesen Bewertungsformaten – insbesondere dann, wenn er den »unverdienten Sieg« dramatisiert. Er affirms seine Bürokratie, seine Regeln und Regularien als unverzichtbare Instrumente zur Gewährleistung des Leistungsvergleichs *und* organisiert gleichzeitig andauernde Empörungsbereitschaft und eine endemische Produktion von »Fehlentscheidungen«.

Unsere kulturanalytische Neuperspektivierung (»Was macht der Fußball mit den Bewertungspraktiken?«) zielte in methodischer Hinsicht darauf ab, dem empirischen Feld maximale Chancen einzuräumen, Konzepte der soziologischen Digitalisierungs-, Accounting- und Bewertungsforschung zu irritieren. Dadurch sind wir auf die Unschärfen von Bewertungsverfahren und ihre Verknüpfungen mit familienähnlichen Praktiken (kategorisieren, quantifizieren, vergleichen, unterscheiden, legitimieren, entscheiden etc.) aufmerksam geworden: Bewertungen schiedsrichterlicher Entscheidungen vollziehen sich in Form von »Gemecker« und »Protesten«; in diese kommunikativen Akte mischen sich Momente des Kategorisierens, des skandalisierenden Vergleichs, die dezisionistisch gekontert werden usw.

Die in unserem Beitrag aufgezeigte empirische Diffusität und Findigkeit des Bewertens in der Welt des Sports ist für die Weiterentwicklung der bewertungsoziologischen Vokabulare auch analytisch relevant: So weist Lamont in ihrem einflussreichen Vorschlag für eine »comparative sociology of (e)valuation« (2012) auf Unschärfen hin, die bereits mit den bewertungsoziologischen Grundbegrifflichkeiten gegeben sind. Lamont zufolge geht es der Bewertungsoziologie einerseits um eine Theorie des Wertens und In-Wert-Setzens (valuation), die auch flüchtige evaluative Vollzüge wie z.B. das Schmecken einbezieht. Andererseits ist das Spektrum empirischer *model cases* (Krause, 2023) nach wie vor sehr eng. Das Forschungsfeld beschäftigt sich vorwiegend mit institutionalisierten evaluativen Verfahren und Techniken. Lamont plädiert deshalb dafür, »processes of evaluation« auch in benachbarten Themenfeldern zu untersuchen. Gut zehn Jahre später hat Bettina Heintz aus der

Perspektive einer Soziologie des Vergleichs heraus ähnlich argumentiert. Es gelte die Soziologie der Kategorisierung, des Vergleichs, der Quantifizierung und der Bewertung als familienähnliche, sich aber gleichzeitig und relativ kontaktarm zueinander entwickelnde Forschungsfelder aufeinander zu beziehen (Heintz, 2021a).²²

Der Fußball und der Sport sind nicht zuletzt deshalb instruktive Fälle für die *Valuation Studies*, weil sich hier charakteristische Verknüpfungen familienähnlicher evaluativer Vollzüge studieren und analytisch produktiv machen lassen. Die Wettbewerbsformate, die organisationalen Praktiken, die Verfahren der Leistungserfassung, die Wissensformen von Fans, Expertinnen und Nerds sind durch eine spannungsreiche Heterogenität bewertungsrelevanter Vollzüge gekennzeichnet und es ist nicht zuletzt diese Heterogenität, die die oben erläuterte kulturelle Darstellungsfähigkeit sportlicher Praktiken gewährleistet. Diese Darstellungsfähigkeit wird erreicht, weil der Sport in seine Bewertungsarrangements Unschärfen, Inkonsistenzen, Ambivalenzen und Möglichkeiten der Relativierung und Anfechtung einbaut. Eine vollständige (zum Beispiel technische) Tilgung von Unschärfen, etwa durch die Automatisierung von Entscheidungen würde diese Darstellungsfähigkeit zunichten machen – seit der Einführung von Goal-Line-Technology ist ein zweites Wembley-Tor nicht mehr möglich. Eine kulturanalytisch ausgerichtete Bewertungssoziologie könnte und sollte sich die kulturelle Fähigkeit des Sports, gesellschaftliches Bewertungsgeschehen aufzuführen, zu pointieren und zu problematisieren für die Weiterentwicklung ihrer analytischen Konzepte zu Nutzen machen.

²² Die klassische Studie von Espeland und Sauder (2007) nimmt ganz gezielt die Verknüpfung von Kategorisieren, Bewerten und Quantifizieren in den Blick. Und auch die jüngst erschienene erste deutschsprachige Einführung von Krüger (2022) versteht es als inhärenten Bestandteil einer »Soziologie des Wertens und Bewertens«, den Austausch zu soziologischen Entwicklungen rund um Konzepte wie »Klassifikation« (23ff.), »Quantifizieren und Vergleichen« (53ff.) zu pflegen. Außerdem wirft die Abgeschlossenheit der Liste von Forschungsfeldern bei Heintz (Vergleichen, Quantifizieren, Bewerten, Kategorisieren) die Frage auf, warum sie nicht deutlich länger ist. Würde es nicht z.B. – um nur ein Beispiel zu nennen – auch Sinn ergeben, neuere kulturanalytische (Pfister et.al. 2019) oder praxeologische (Schmidt 2022) oder auch klassische Ansätze (Simon, Luhmann usw.) zu Praktiken des Entscheidens in die Liste mitaufzunehmen?

Literatur

- Alkemeyer, T. (1997). Sport als Mimesis der Gesellschaft. Zur Aufführung des Sozialen im symbolischen Raum des Sports. *Zeitschrift für Semiotik*, 19(4), 365–395.
- Antal, A. B., Hutter, M., & Stark, D. (Eds.). (2015). *Moments of valuation: Exploring sites of dissonance*. Oxford University Press.
- Arora-Jonsson, S., Brunsson, N., & Hasse, R. (2021). A new understanding of competition. In Dies (Hg.) *Competition. What It is and Why It Happens* (1–25). Oxford University Press.
- Aspers, P. (2018). Forms of uncertainty reduction: decision, valuation, and contest. *Theory and society*, 47(2), 133–149.
- Berli, O., Nicolae, S., & Schäfer, H. (2021). Bewertungskulturen. Ein Vorschlag für eine vergleichende Soziologie der Bewertung. In Dies. (Hg.) *Bewertungskulturen* (285–302). Springer VS, Wiesbaden.
- Berli, O., & Alberth, L. (2022). Vermittlungsarbeit und Bewertungswissen. *Österreichische Zeitschrift für Soziologie*, 47(3), 291–303.
- Bierwagen, J. (2023): *Scouting und Videoanalyse im Fußball* (Abschlussarbeit am Fachgebiet Soziologie der KU Eichstätt)
- Boltanski, L. & Thévenot, L. (2007). Über die Rechtfertigung. Eine Soziologie der kritischen Urteilskraft. Hamburg: Hamburger Edition.
- Brankovic, J., Ringel, L., & Werron, T. (2018). How rankings produce competition: The case of global university rankings. *Zeitschrift für Soziologie*, 47(4), 270–288.
- Bröckling, U. (2020). Wettkampf und Wettbewerb: Konkurrenzordnungen zwischen Sport und Ökonomie. *Konfliktodynamik*, 9(4), 248–255.
- Bromberger, C. (2003). Fußball als Weltsicht und als Ritual. In: Belliger, A., & Krieger, D. J. (Hg.). *Ritualtheorien: ein einführendes Handbuch* (285–302). Wiesbaden: Springer-Verlag. VS Verlag.
- Brümmer, K. (2019). Spielsysteme, Matchpläne, Spielanalysen: Über Praktiken und Medien des Kontingenzmanagements im gegenwärtigen Fußball. *Sport und Gesellschaft*, 16(3), S. 266–300.
- Brümmer, K., Alkemeyer, & Janetzko, A. (20109) (Hg.). *Ansätze einer Kulturosoziologie des Sports*. Baden-Baden: Nomos Verlagsgesellschaft
- Caillois, R. (1982). *Die Spiele und die Menschen: Maske und Rausch*. Ullstein.
- Christin, A. (2018). Counting clicks: Quantification and variation in web journalism in the United States and France. *American Journal of Sociology*, 123.5, 1382–1415.

- Collins, H., Evans, R. & Higgins, C. (2016). *Bad Call: Technology's Attack on Referees and Umpires and how to Fix it*. Massachusetts: MIT Press.
- Colwell, S. (2004). Elite level refereeing in men's football: A developmental sociological account. Doctoral dissertation, University of Leicester.
- Crawford, G. (2006). The cult of Champ Man: The culture and pleasures of Championship Manager/Football Manager gamers. *Information, Communication & Society*, 9(4), 496–514.
- Crawford, G., & Gosling, V. K. (2009). More than a Game: Sports-Themed Video Games and Player Narratives. *Sociology of Sport Journal*, 26(1).
- Diaz-Bone, R. & Fasson, D.E. (2016): The sociology of quantification – perspectives on an emerging field in the social sciences. In: *Historical Social Research* 41 (2), 7–26.
- Dowling, D. (1999). Experimenting on Theories. *Science in Context*, 12(2), 261–273.
- Elias, N. & Dunning, E. (1982). *Sport im Zivilisationsprozeß: Studien zur Figurationssoziologie*. Münster: LIT Verlag.
- Espeland, W. N., & Sauder, M. (2007). Rankings and reactivity: How public measures recreate social worlds. *American journal of sociology*, 113(1), 1–40.
- Espeland, W. N., Sauder, M., & Espeland, W. (2016). Engines of anxiety: Academic rankings, reputation, and accountability. Russell Sage Foundation.
- Foucault, M. (1973). *Die Geburt der Klinik: Eine Archäologie des ärztlichen Blicks*. Frankfurt a.M.: SuhrkampVerlag
- Gebauer, G. (2006). *Poetik des Fußballs*. Frankfurt a.M.: Campus Verlag.
- Gebauer, G. & Wulff, C. (1998). *Spiel-Ritual-Geste: mimetisches Handeln in der sozialen Welt*. Hamburg: Rowohlt Verlag.
- Geertz, C. (1974). Deep play: Notes on the Balinese cockfight. In: Ders. *The Interpretation of Cultures. Selected Essays*, Basic Books Publishers: New York
- Goffman, E (1974). *Das Individuum im öffentlichen Austausch: Mikrostudien zur öffentlichen Ordnung*. Frankfurt a.M.: Suhrkamp Verlag.
- Goffman, E. (1982 [1967]). *Interaction ritual: Essays in face to face behavior*. Al-dine Transaction.
- Goodman, N. (1976). *Languages of art: An approach to a theory of symbols*. Hackett.
- Goodman, N. (2013). *Ways of worldmaking*. Hackett.

- Gugutzer, R. (2015): Der body turn in der Soziologie. Eine programmatische Einführung. In: Ders. (Hg.): Body turn. Perspektiven der Soziologie des Körpers und des Sports (9–54). transcript.
- Guttmann, A. (2004). From ritual to record: The nature of modern sports. Columbia university press.
- Heintz, B. (2016): »Wir leben im Zeitalter der Vergleichung«. Perspektiven einer Soziologie des Vergleichs. In: Zeitschrift für Soziologie 45 (5), S. 305–323.
- Heintz, B. (2021a). Kategorisieren, Vergleichen, Bewerten und Quantifizieren im Spiegel sozialer Beobachtungsformate. KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie, 73, 5–47.
- Heintz, B. (2021b). Big Observation—Ein Vergleich moderner Beobachtungsformate am Beispiel von amtlicher Statistik und Recommendersystemen. KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie, 73, 137–167.
- Hennion, A. (2015): Paying attention: What is tasting wine about. In Antal et.al. (Hg.) Moments of valuation: Exploring sites of dissonance (37–56).
- Hirschauer, S., & Boll, T. (2018). Un/doing Differences. Zur Theorie und Empirie eines Forschungsprogramms. In Dies. (Hg.) *Un/doing differences* (7–26). Weilerswist:Velbrück Wissenschaft.
- Hocquet, A. (2016). Football Manager: Mutual Shaping between Game, Sport, and Community. Journal of Media Studies and Popular Culture = Revue d'études Des Médias et de Culture Populaire, 6(Special Issue). <https://shs.hal.science/halshs-01312229>
- Hodek, F. (2018). Spielanalysen und Sportwetten: Strategien der Quantifizierung im Profifußball. *Berliner Debatte Initial: sozial- und geisteswissenschaftliches Journal*, 29(1), 147–163.
- Hoppe, A. D. (2021). The Microsociology of Aesthetic Evaluation: Selecting Runway Fashion Models. Qualitative Sociology, 45, 63–87.
- Janetzko, A. (2015). Professionelles Sehen. Sichtungspraktiken im Tanzsport. Sport und Gesellschaft, 12(2), 105–132.
- Kalthoff, H. (1996). Das Zensurenpanoptikum: Eine ethnographische Studie zur schulischen Bewertungspraxis. Zeitschrift für Soziologie, 25 (2), 106–124.
- Kalthoff, H. (2019) Organisierte Humanevaluation. In: Nicolae/Endreß/Berli/Bischur (Hg.) (Be)Werten. Beiträge zur sozialen Konstruktion von Wertigkeit (221–247). Wiesbaden: Springer Fachmedien.
- Kalthoff, H. & Dittrich, T. (2016) Unterscheidung und Härtung. Bewertungs- und Notenkommunikation in Lehrerzimmer und Zeugniskonferenz. Berliner Journal für Soziologie. 26(3-4), 459–483.

- Kempf, D. (2017). Play the Game. Know the Game – Rekursive Wissenstransfers zwischen Football Manager und Fußball. *Paidia – Zeitschrift für Computerspielforschung*. <https://www.paidia.de/play-the-game-know-the-game-rekursive-wissenstransfers-zwischen-football-manager-und-fussball/>
- Krause, M. (2023). Von Mäusen, Menschen und Revolutionen: Modellfälle in der Sozialforschung. Hamburger Edition HIS.
- Krüger, A. K. (2022). *Soziologie des Wertens und Bewertens*. Stuttgart: UTB Verlag.
- Krüger, A. K., & Hesselmann, F. (2020). Sichtbarkeit und Bewertung. *Zeitschrift für Soziologie*, 49(2-3), S. 145–163.
- Lambrix, P. (2022). Erstrebenswerte Unselbständigkeit? Die Ambivalenz der Pflegebegutachtung. *Zeitschrift für Soziologie*, 51.3, 227–243.
- Lamont, M. (2012). Toward a comparative sociology of valuation and evaluation. *Annual review of sociology*, 38., 201–221.
- Kropf, J., & Laser, S. (2019) (Hg.). *Digitale Bewertungspraktiken*. Wiesbaden: Springer Fachmedien.
- Laube, S. (2019). Der Markt Im Körper. Emotionales Beobachten und Bewerten Im digitalisierten Finanzhandel. *Zeitschrift für Soziologie*, 48(4), 263–78.
- Lom, S. E. (2016). Changing rules, changing practices: The direct and indirect effects of tight coupling in figure skating. *Organization Science*, 27(1), 36–52.
- Luhmann, N. (2017). *Die Realität der Massenmedien* (5. Auflage). Springer VS.
- Magee, J., Bairner, A., & Tomlinson, A. (Hg.). (2005). *The bountiful game? Football identities and finances*. Meyer und Meyer Sport.
- Mears, A. & Finlay, W. (2005): Not just a paper doll: How models manage bodily capital and why they perform emotional labor. *Journal of Contemporary Ethnography*, 34(3), 317–343.
- Meier, F. & Peetz, T. (2021). Organisation und Bewertung. Wiesbaden : VS Springer Verlag.
- Memmert, D. & Raabe, D. (2017). *Revolution im Profifußball*. Wiesbaden: Springer VS.
- Mennicken, A. (2011): Sociology of Accounting. In: Beckert, J. & Zafirovski, M. (Hg.): *International encyclopedia of economic sociology* (1–5). London, New York: Routledge.
- Mennicken, A., & Kornberger, M. (2021). Von Performativität zu Generativität: Bewertung und ihre Folgen im Kontext der Digitalisierung. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 73(1), 451–478.
- Meyer, C. (2021). Interkorporalität. In: Brümmer et.al. *Ansätze einer Kultursoziologie des Sports* (153–174). Stuttgart: Nomos Verlag.

- Minnetian, C., & Werron, T. (2021). Redefining achievement: The emergence of rankings in American baseball. In Ringel et.al. (Hg.) *Worlds of rankings* (127–151). Emerald Publishing Limited.
- Münch, R. (2009). Globale Eliten, lokale Autoritäten. Bildung und Wissenschaft unter dem Regime von PISA, McKinsey & Co. Frankfurt a.M.: Suhrkamp Verlag.
- Mol, A. & Heuts, F. (2013). What is a good tomato? A case of valuing in practice. *Valuation Studies*, 1(2), 125–146.
- Pethes, N. (2018). Sportsimulationen. In: Pop. *Kultur und Kritik* 7(1)
- Prinz, S. (2014). Die Praxis des Sehens: Über das Zusammenspiel von Körpern, Artefakten und visueller Ordnung. Bielefeld: transcript.
- Ringel, L. & Werron, T. (2020): Where do Rankings Come From. A Historical-sociological Perspective on the History of Modern Rankings. In: Epple, Annelika et.al. (Hg.): *Practices of Comparing. Towards a New Understanding of a Fundamental Human Practice*, 137–170.
- Ringel, L., Espeland, W., Sauder, M., & Werron, T. (2021). *Worlds of rankings*. Dies. (Hg.) *Worlds of rankings* (1–23) Emerald Publishing Limited.
- Rivera, L. A. (2012). Hiring as cultural matching: The case of elite professional service firms. *American sociological review*, 77(6), 999–1022.
- Rivera, L. A. (2015). Go with your gut: Emotion and evaluation in job interviews. *American journal of sociology*, 120(5), 1339–1389.
- Rivera, L. A. (2017). When two bodies are (not) a problem: Gender and relationship status discrimination in academic hiring. *American Sociological Review*, 82(6), 1111–1138.
- Schatzki, T. (2010). *Site of the social: a philosophical account of the constitution of social life and change*. Pennsylvania State University.
- Scheffer, T., & Schmidt, R. (2023). Öffentliche Soziologie und gesellschaftliches »Soziologisieren«. In: Selke et.al. (Hg.). *Handbuch Öffentliche Soziologie* (351–362). Wiesbaden: Springer Fachmedien.
- Schindler, L. (2016). Kampffertigkeit: Eine Soziologie praktischen Wissens (Vol. 13). Walter de Gruyter GmbH & Co KG.
- Schmidl, A. (2023). Visual discourses in sport. A sociological analysis of the implementation of the video evidence in cycling and football. *Visual Studies* 38.3-4 (2023): 366–376
- Schmidt, R. (2022): *Toward a Culture-Analytical and Praxeological Perspective on Decision-Making*. *Human Studies* 45(4), 653–671
- Schmidt, R., & Hodek, F. (2019). Spielanalysen und Transformationen der Figuration Profifußball. In Ernst et.al. (Hg.) *Transformationen der Ar-*

- beitsgesellschaft: Prozess- und figurationstheoretische Beiträge (321–344), Wiesbaden: VS Verlag
- Schmidt, R., Hodek, F., & Bickl, A. (2017). Organisationale Praktiken und Effekte quantifizierenden Berichtsbarmachens. Lessenich, S.(Hg.). Geschlossene Gesellschaften – 38. Kongress der Deutschen Gesellschaft für Soziologie
- Schmidt, R., Weigelin, M., Brümmer, K., Laube, S., & Schäfer, H. (2022). Bodies and embodiment in practices of valuation: Challenging the sociology of valuation with the sociology of the body. Österreichische Zeitschrift für Soziologie, 47(3), S. 217–223.
- Singer, M. (1972). *When a Great Tradition Modernizes: An Anthropological Approach to Modern Civilization*. New York, Washington, London: Praeger Publishers.
- Streeck, J. (2021). Interaktion. In Brümmer et.al. (Hg.). Ansätze einer Kultursociologie des Sports (87–112). Baden-Baden: Nomos Verlagsgesellschaft
- Tuma, R. (2017). Videoprofis im Alltag: Die kommunikative Vielfalt der Videanalyse. Wiesbaden: Springer Fachmedien.
- Vormbusch, U. (2004): Accounting. Die Macht der Zahlen im gegenwärtigen Kapitalismus. In: Berliner Journal für Soziologie 14 (1), 33–50.
- Weigelin, M. (2022). Entscheidungen und ihre Bewertungen – Zur Mikrosociologie des Schiedsrichter-Pfiffs. Österreichische Zeitschrift für Soziologie, 47(3), 225–246.
- Werron, T. (2005). Quantifizierung in der Welt des Sports. Gesellschaftstheoretische Überlegungen. *Soziale Systeme*, 11(2), 199–235.
- Werron, T. (2010). Der Weltsport und sein Publikum. Zur Autonomie und Entstehung des modernen Sports. Weilerswist: Velbrück.
- Wunderlich, A. (2023). Die verrückte Entwicklung der Bundesliga-Transfersummen. br24, <https://www.br.de/nachrichten/sport/die-verrueckte-entwicklung-der-bundesliga-transfersummen,TneNAma> (zuletzt aufgerufen am 23.1.2024)
- Yarrow, D., & Kranke, M. (2016). The performativity of sports statistics: towards a research agenda. *Journal of Cultural Economy*, 9(5), 445–457.

Spielerwerte und Spielen/Bewerten

Digitale Fußballspiele und die transmedialen Affekte von Leistungsbewertung(en)

Martin Lüthe

Diogo Jota, portugiesischer Fußballnationalspieler in Diensten des FC Liverpool aus dem Norden Englands, grinst in die Kamera und verkündet, »Hallo Leute, wir sind wieder in dieser Zeit des Jahres angekommen!«. Und kurze Zeit später im gleichen Video wissen wir als Publikum, was er damit meint: Electronic Arts (Kurz: EA) hat auch Mitte September 2023 wieder ein neues digitales Fußballspiel für die gängigen Systeme veröffentlicht. Obwohl das Spiel dieses Jahr zum ersten Mal unter neuem Namen, nämlich als EA FC 24 ins Rennen geht, schließt es in seiner Spielweise, seiner digitalen Ästhetik (Graphik, Atmosphäre, Menüführung, Spielmodi) aber auch seiner Marketingstrategie nahtlos an seinen Vorgänger, das weltweit berühmte und global gespielte *FIFA* Franchise, an. Ein zentrales Element dieser Strategie der Vermarktung stellen hierbei die auf der Videoplattform YouTube.com veröffentlichten Rating Reveal-Videos dar. Diese erscheinen kurz vor der offiziellen Veröffentlichung des Spiels und zeigen die Reaktionen der Fußballprofis auf ihre statistische Bewertung nach Leistungsparametern in der Spielwelt (game world) von EA FC — und dieses Jahr moderiert eben jener Diogo Jota das entsprechende Video für seinen Verein.¹

Wir begegnen in diesen Videos, so argumentiere ich im Beitrag, verschiedenen bemerkenswerten Ebenen und Affekten von Bewertungen innerhalb und im Umfeld von digitalen Fußballspielen. Diese Ebenen und Affekte stehen

¹ Meine Übersetzung von: »OK, guys, it is that time of the year again!« <https://www.youtube.com/watch?v=Gn-9eMQBZMI>, 18.10.2023. Das Video hatte 14 Tage nach seiner Veröffentlichung zwei Millionen Klicks generiert; es stellt sich also durchaus die Frage, ob es sich noch um ein Nischengenre handelt oder ob diese Videos gerade dabei sind, sich zu einer eigenen Franchise zu entwickeln. So oder so, stellen Rating Reveal-

im Folgenden im Zentrum der Analyse der Videos. Hierbei geht es dezidiert nicht um die Frage nach der vermeintlichen Richtigkeit der Bewertungen, sondern zum einen um die Gründe für den bemerkenswerten Willen zur Bewertung innerhalb digitaler Fußballspiele und zum anderen um die Affekte, die Bewertungen und deren spielerische »Sichtbarmachung« zu mobilisieren im Stande sind. Als spielerisch verstehe ich hier sowohl die Praxis der Bewertung innerhalb der Spielwelten besagter Spiele (der Fokus liegt hierbei auf Konamis efootball PES 2021) als auch den Umgang mit den Bewertungen in den Rating Reveal-Videos, die vor allem von EA ins Feld geschickt werden. Im Anschluss an eine sehr kurze Einführung in digitale Fußballspiele, analysiert der Beitrag zunächst die Bedeutung der Bewertung von Spielermodellen und Spielertypen für die Spielpraxis innerhalb der Spiele (und deren Affekte), bevor in einem weiteren Abschnitt die YouTube Videos im Zentrum der Analyse stehen. Die Kernthese des Beitrages lautet dabei, dass die Rating Reveal-Videos als ›Umschlagplatz‹ von kaltem Bewertungsregime der in Statistiktabellen zusammengefassten Spielerabbildern in heiße Affekte fungieren. Somit beantwortet der Beitrag – wie andere der hier gesammelten auch – die dem Band zugrunde liegende Frage danach, was der Sport mit den ihn kolonisierenden Bewertungsformen zu machen im Stande ist. Zentral für das Phänomen des ›affektiven Erhitzens‹ in den Rating Reveal-Videos sind entsprechend Formen der Zuspitzung und Dramatisierung, die Affekte über das Spiel und Spielen hinaus sicht- und erfahrbar machen oder transmedial rückkoppeln. Digitale (Sport-)Spiele im Allgemeinen und digitale Fußballspiele im Besonderen verstehe ich mit Steve Shaviro hierbei als »machines for generating

Videos ein eigenes Genre auf YouTube dar, bei der namhafte Fußballer*innen erfolgreicher Profimannschaften durch die Gebäude und Anlagen besagter Mannschaften führen und dabei den Spieler*innen, denen sie im Zuge der Videos begegnen, deren jeweils aktuellen bzw. neuen Bewertungen im jüngsten FIFA-Spiel mitteilen (oder sie diese erraten lassen) und die Reaktion beobachten und begutachten – dies eröffnet einen Blick auf Bewertungen und die Affekte, die Bewertungen dieser Art im Stande sind zu mobilisieren. Interessant ist hierbei überdies, wie jenseits der Videos die Kommentarspalten von User*innen genutzt werden, um die Wertungen selbst zu bewerten – entlang der Parameter von Richtigkeit und Realismus – mit der ebenfalls für die Kommentarspalten nicht untypischen affektiven Sprache und Verve. In Länge und affektiver Intensität liegen die Rating Reveal-Videos mit ihrer für YouTube typischen Länge von zwischen fünf und zwölf Minuten zwischen den sehr kurzen Videos bei Instagram oder TikTok und den mehrere Stunden langen Sportdokumentationen unserer post-TV Ära mit ihrer Vielzahl an Streaming-Plattformen.

affect, and for capitalizing upon, or extracting value from, this affect» (2009, 3; Hervorhebung im Original); also als affektgenerierende Maschinen mit dem Ziel der (kapitalistischen oder monetären) Wertschöpfung. Was die digitalen Sportspiele in diesem Punkt besonders macht, ist, dass sie via Bewertungen und Bewertungslogiken kapitalistische Wertschöpfung innerhalb einzelner Spielmodi spiel- und »affizierbar« machen.

1. Digitale Fußballspiele als Bewertungsspiele

Spätestens im Zuge der digitalen Fußballspiele, die Anfang der 90er Jahre für die damals vertriebenen Spielekonsolen und PCs in Serie aufgelegt wurden, unter anderem mit Kick Off (bis 1990), Sensible Soccer (von 1991 bis 1993, Renegade), Konamis International Superstar Soccer (seit 1994 als Pro Evolution Soccer in Serie gegangen) und natürlich EAs FIFA Franchise, begannen die Spiele im Dialog mit den gleichzeitig beliebt werdenden Fußballmanager-Spielen umfassend zu bewerten: Spielerfähigkeiten und -eigenschaften, aufgeschlüsselt in Parameterlisten, von den auf den Interfaces repräsentierten Fußballern des real life sports wurden im Zuge der 1990er Jahre zu einem offensichtlichen Standard für digitale Fußballspiele, die – so darf man unter anderem wohl annehmen – zum einen dabei helfen sollten die Statistik-affinen Manager-spiel-Spieler*innen zu gewinnen und dabei gleichzeitig die in der »Fußballwelt« tatsächlich stattfindende detaillierte Bewertung von Leistungsparametern für Sportler*innen (zunächst für lange Zeit allerdings lediglich für männliche Profifußballer) abzubilden. Begann dies zunächst mit der Vergabe von, zum Beispiel, Sternen für besonders exzellente Spielerabbilder, finden wir bei denen in der jüngeren Zeit veröffentlichten digitalen Fußballspielen, Konamis PES (efootball) Franchise und EAs FIFA bzw. EA FC 24, umfangreiche Listen mit über 25 Parametern pro individuellem Spielerabbild (siehe Abb Ia und Ib).

Zwei Fragen schließen sich unmittelbar an diese Entwicklung zu einer progressiv detaillierter werdenden Aufschlüsselung von Leistungscharakteristika an und sollen in Folge diskutiert werden: 1. Die Frage nach der Bewertung von Parametern, die auf die Körperlichkeit der Spielerabbilder abzielen entlang einer rassistischen und kolonialen Bewertungslogik. 2. Der Diskurs und die Freude an der Bewertung von Spieler-Bewertungen (Meta-Bewertung) und deren Spielbarkeit. Die Freude an der Bewertung erklärt gleichzeitig die Sinnhaftigkeit und den Erfolg des Genres der Rating Reveal-Videos, die EA auf You-

Tube mit etlichen Vereinen und Protagonist*innen zur Verfügung stellt und verbreitet.²

Abb. Ia



Abb. Ib



Jenseits der Frage nach der Mentalität von Kollektiven und deren Bedeutung als Teil rassistischer und neokolonialer Diskurse (Fußnote 2), entfaltet die

2 Da ich an anderer Stelle schon hierzu geschrieben habe, soll der Fokus dieses Beitrages nicht auf der rassistischen und neo-kolonialen Bewertungslogik der beiden bekanntesten digitalen Fußballspiele liegen; nur soviel: die Bewertungslogik der Spiele macht dieses Phänomen rassistischer Diskurse sichtbar (in den numerischen Tabellen, die, zum Beispiel, das taktische Verständnis und den so genannten Teamgeist afrikanischer Nationalmannschaften niedrig bezeichnen) und spielbar (indem die Gamer*in jene niedrige Bewertung spürt, zum Beispiel durch weniger zielführende Laufwege seitens der KI). In der niedrigeren Bewertung der Nationalmannschaften afrikanischer Länder in der Kategorie Teamgeist verdichtet sich also ein einflussreicher und andauernder kolonialer und rassistischer Diskurs des *othering*, bei dem die Aufteilung in »Geist« und Körper immer schon rassisches Überformt war und ist. Der Effekt dieser rassistischen Annahme wird in den digitalen Fußballspielen auf diese Art und Weise im Spiel erfahrbar gemacht. Insgesamt also eignen sich diese digitalen Fußballspiele nicht zuletzt aufgrund ihres Willen zur Bewertung als Archiv dieser spätkolonialen oder neukolonialen Imaginations von körperlicher und mentaler Differenz — jenem Grundbauelement aufklärerischer Rassendiskurse in Deutschland und der Welt; siehe hierzu (Lüthe 2013). Zur generellen Bedeutung von Bewertungslogiken in den Rasse-diskursen und Kolonialdiskursen, siehe (Weheliye 2014).

detaillierte Bewertung innerhalb der Spiele weitere relevante in-game Effekte auf der Ebene der einzelnen Spieler. Gewissermaßen im Gegensatz zu der KI-geführten Gesamtsituation, schlägt sich die Bewertung einzelner Spieler besonders in jenen Situationen spürbar nieder, in denen wir als Gamer*innen am Controller einzelne Spieler auf dem digitalen Spielfeld selbst steuern. Die Steuerbarkeit einzelner Spieler auf dem Feld folgt hierbei im Angriff dem eigenen Ballbesitz, also wir spielen jeweils den ballführenden Spieler, in der Defensive denjenigen, der dem Spielgeschehen mit Ball am nächsten ist; in der Regel werden die zehn anderen Spieler entsprechend jeweils von der KI gesteuert.³ Während im Spiel, hier beispielhaft Konamis PES 2021, die Werte in einer Liste für jeden einzelnen Spieler aufgeführt werden oder alternativ in der visualisierten Leistungsmatrix des Sechsecks aus Schießen, Passen, Stärke, Verteidigung, Geschwindigkeit und Dribbling, finden sich im Internet auf den entsprechenden Fanseiten, hier beispielhaft bei pesmaster.com, die Parameter in Spielerdatenbanken nochmal detailliert aufgearbeitet und nebeneinander gestellt. Diese im Internet verfügbaren Datenbanken basieren also auf den Bewertungsdatenbanken in den Spielen und exportieren sie zur schnellen, visuellen Verfügbarkeit jenseits des Spiels. Dies ist Ausdruck der sinnstiftenden Bedeutung dieser Bewertungen für die Spiele und unterstreicht, welch zentrale Rolle den Bewertungen für das Spielemachen zukommt.

Im Spiel wird diese statistische Bewertung dann tatsächlich spiel- und erfahrbar: ein zentrales Bewertungspaar stellen hierbei Geschwindigkeit und Beschleunigung dar, die dafür sorgen, dass die unterschiedlichen Sprintgeschwindigkeiten von Fußballprofis sich auf den digitalen Spielfeldern und den Stadien niederschlagen. Wie die Rating Reveal-Videos unterstreichen, kommt diesem Paar eine besondere Bedeutung zu, wohl auch deshalb, weil es eine spezifische und spezifisch messbare Eigenschaft von Sportler*innen darstellt — hierzu zum Ende des Beitrages mehr. Dies gilt wohl weniger für die Bewertungen im Bereich Technik, Finesse und Ballgefühl oder auch Passgenauigkeit und Passgeschwindigkeit/-stärke. Jedoch entwickeln alle diese Bewertungsgruppen (also Geschwindigkeit, Beweglichkeit, Passspiel, Technik, Torabschluss, Schusseigenschaften etc.) jeweils einen eigenen Ein-

3 Der Beitrag setzt sich hierbei nicht zum Ziel die in den Game Studies ausführlich geführte Debatte im Kontext von »Ludologie« und »Narratologie« zu durchleuchten; zur Praxis des Spielens und zu verschiedenen Arten der Befehls-Gebung am Joystick, Controller, oder anderen Eingabeinstrumenten siehe: (Blomberg 2018).

fluss auf das Spielverhalten einzelner Spieler und somit auf den geforderten Umgang mit den einzelnen Spielern seitens der Gamer*in.

Die Spielwelt der digitalen Fußballspiele sorgt via Bewertungen einzelner abgebildeter Spieler also dafür, ob es Sinn macht, mit einem bestimmten Spieler aus 35 virtuellen Metern Torenfernung einen Abschluss zu wagen oder gegen einen bestimmten Typ Verteidiger mit einem spezifischen Angreifer ins Dribbling zu gehen. Die Bewertung der Schussstärke und -präzision des Spielers determiniert also die Wahrscheinlichkeit eines Torerfolges bereits lange bevor das Spiel auf dem digitalen Rasen angepfiffen worden ist. Die Einflussnahme der Gamer*in ist hier limitiert, jedoch gibt es in den narrativen Modi Trainingseinstellungen, um spezifische Parameter der Spieler langfristig zu verbessern. Bei einem bereits angepfiffenen Spiel gilt es indes immer, die Spieler auf Grundlage der eigenen Kenntnis der bewerteten Fähigkeiten einzusetzen: Ist ein Spieler physisch schwach, etwas klein, dafür aber wendig und »ballstark«, dann macht es mehr Sinn, mit diesem zu dribbeln und zu passen, als aus 30 Metern digitaler Tordistanz den Torabschluss zu suchen.

Dies halte ich daher aus drei Gründen für relevant im Kontext dieses Beitrages und der Frage nach der Bedeutung von Bewertungen in digitalen Fußballspielen und ihren Affekten.

- 1) Die Bewertungen sind maßgeblich für das Spiel-Empfinden der Gamer*in, besonders im Hinblick auf die Frage des spielerischen Realismus oder der Authentizität der Spiele. Anders formuliert: Die Stadionatmosphäre, die graphische Darstellung und die Kameraführung mag noch so sehr an Fußballübertragungen erinnern, also tele-realistisch anmuten, die Simulation des Spielgeschehens wird jedoch an zentraler Stelle in die Bewertung des Spiels seitens der Gamer*innen einfließen. Hier erfolgt oft — teils implizit, teils explizit — die Bewertung der Bewertungen durch die Gamer*innen, weil diese für das digitale Fußballspiel fundamental sind. Dies findet sich auch in den Online-Foren und Diskussionen zu den Spielen (PES und FIFA), zum Beispiel bei der Frage nach dem »Über-Powern« einzelner Spieler und/oder Spieleigenschaften und/oder gewisser Tricks (also durch Tastenkombinationen hervorgerufene Bewegungsabfolgen der Spieler wie die Marseille-Roulette oder dem Rainbow-Flick). Genau in diesem Sinne fungieren gegenwärtige digitale Fußballspiele an zentraler Stelle als Bewertungsspiele oder gar als Meta-Bewertungsspiele. Gamer*innen bewerten fortlaufend Spielsituationen auf der Grundlage von ihnen bekannten Bewertungen der einzelnen Spieler auf dem Feld.

- 2) In der Bewertung der Bewertung durch die Gamer*in liegt ein großes Potential der »Verbesserung« des eigenen Spielens. Je bewusster die Gamer*innen mit den Eigenschaften der Spieler auf dem Platz umgehen, desto höher die Wahrscheinlichkeit erfolgreich zu sein und gute Ergebnisse zu erzielen. Diese Kalkulation wird in den »narrativen« Spielmodi der digitalen Fußballspiele entsprechend noch zentraler: Beim Zusammenstellen der Mannschaften muss der Gamer*in klar sein, dass die Spieler entsprechend ihren Bewertungsprofilen in den Kader und die Mannschaft passen müssen. Hier leihen sich die digitalen Fußballspiele also Ideen, die in datenbasierten Fußball- und Sportmanagerspielen schon lange und ausführlich spielerisch erfahrbar gemacht worden sind und die es sogar geschafft haben, den sprichwörtlichen Ball (des Wissens) in den professionellen Betrieb des Fußballs rekursiv zurückzuspielen, wie David Kempf (2017) zeigt.⁴
- 3) Die Effekte dieser detaillierten Bewertung einzelner Spieler sind wirkmächtig für die affektive Logik des Spielens seitens der Gamer*innen. Spielerbewertungen werden in ihrer Erfahrbarkeit auf dem digitalen Fußballfeld auch affektiv mobilisiert. Jenseits der eher kalten, statistisch-detaillierten Darstellung einzelner Spieler als Ausdruck der (vielleicht offensichtlichen) Simulations-Aspiration, wird in einem gelungenen Dribbling mit einem besonders dribbelstarken oder gar besonders dribbelschwachen Spieler die Logik der Zahlen affektiv »aufgeladen«. Frustration und Freude stellen hierbei wohl die offensichtliche affektive Klammer für das Spielen digitaler Sportspiele im Allgemeinen dar. Der mittlerweile zum kulturellen Klischee avancierte »Controller-Wurf« als Quintessenz des performten Wutausbruchs von Gamer*innen repräsentiert hier den einen Pol des Spektrums, während der doppelte Torjubel vor und auf dem digitalen Interface den anderen Pol markiert. Hierfür — also für beide Extreme — hält die detailgetreue Abbildung von Spielertypen Potenzierungseffekte bereit. Trifft eine Gamer*in also einen Distanzschuss aus 35 Metern auf dem digitalen Spielfeld mit einem Spieler, der dies eigentlich

4) Kempf diskutiert diese Prozesse der doppelten Bewertung in Fußballmanagernspielen als »Realitätsverdoppelung«, während ich mit dem Begriff des »Telerealismus« auf die transmediale Inszenierungslogik digitaler Fußballspiele rekurriere. Diese beiden Zugriffe ähneln sich meines Erachtens jedoch, da sie dazu dienen die radikale Banalität von »realem« Außen und »spielerischem/simuliertem« Innen aufzuweichen oder aufzulösen.

nicht können dürfte, oder dribbeln wir mit einem Innenverteidiger (oder gar Torwart) bis ins gegnerische Tor, so verschieben sich die affektiven Pole in beide Richtungen nach außen. Zwar sind diese Momente selten, aber sie sollen es auch sein und stellen genau daher affektive Höhepunkte der jeweils individuellen digitalen Sportgeschichte der Gamer*innen dar.⁵

Neben dieser spielinhärenten Logik und der Realismus-Aspiration der Spiele hat die statistische Bewertung der Spieler noch weitere Effekte, vor allem für die abgebildeten Spieler selbst. In einer betriebswirtschaftlichen Studie zu Spielergehältern in der italienischen Serie A legen Giovanni Bernardo, Massimo Ruberti und Roberto Verona dar, dass der von EA veranschlagte *International Reputation Index* des FIFA (EA FC) Spiels mit der Höhe der tatsächlichen Spielergehälter genauso korrespondiert wie mit der Anzahl der Suche nach individuellen Spielern in Suchmaschinen wie google.com (2022, 598ff). Hier findet sich also einmal mehr jene Verkomplizierung des Außen und Innen der Welten digitaler Spiele ausgedrückt, die Kempf als »Realitätsverdoppelung« (2017) beschreibt. Es bleibt offen, ob Spieler potentiell ein höheres Jahresgehalt erhalten, weil sie sich laut EA einer großen internationalen Reputation erfreuen oder ob sie von EA höher eingestuft werden, weil sie mehr verdienen.⁶

2. Player Ratings und Rating Reveal Videos

Die Rating Reveal-Videos folgen einem wiederkehrenden Muster und sind als Genre daher gut identifizierbar. Sie werden typischerweise auf den YouTube-Kanälen der Vereine veröffentlicht und als so genannte »reaction videos« verschlagwortet. »Reaction videos« stellen somit die nächstgrößere Genre-Klammer dar: In diesen Videos sehen wir die Protagonist*innen der Videos auf verschiedene Phänomene reagieren und wir erleben somit sekundär die Affekte der Personen in Bezug auf die jeweils spezifische Reaktion. Bei den Rating

5 Zur Affekttheorie, siehe: Berlant (2011); zu den Affekten digitalen Spielens, siehe vor allem: Anable (2018).

6 Dass allen diesen Bewertungen eine meritokratische Idee von sportlicher Leistungsfähigkeit zugrunde liegt, davon gehen wir im Profisport oft reflexartig aus; jedoch sollten wir an dieser Stelle ernst nehmen, dass »Reputation« schwer statistisch messbar ist und dass, zum Beispiel, Social Media Reichweite eine nicht nur auf sportlicher Leistung beruhende Einheit darstellt, die für das Veranschlagen von internationaler Reputation eine nicht unerhebliche Relevanz darstellen könnte.

Reveal-Videos geschieht dies, wie der Titel bereits verdeutlicht, in dem Moment, in dem die im digitalen Fußballspiel abgebildeten Spieler ihre jeweils aktuelle Bewertung im neuen Teil der EA-Serie anhand einer Spielerkarte erfahren. Diese Spielerkarte wird überdimensioniert auf Pappe gedruckt und im Video dargestellt. Meistens finden diese Videos auf den Trainingsgeländen der Profisportvereine statt und jeweils ein Spieler, oft der Kapitän oder ein hoch bewerteter Spieler, lauert anderen Spielern des Teams auf und zeigt ihnen die mit Spannung erwartete neue Bewertung. Die Spieler reagieren dann, nicht selten auf heftige Art und Weise, auf die aktuellen Bewertungen und es folgt die für Männerbünde typischerweise imaginierte »Frotzelei«, eine kulturelle Praxis die im Englischen als »banter« bekannt ist und in afrikanisch-amerikanischen Kontexten manchmal als »playing the dozens« (Vognar 2011). Die Videos sind, wie für die Plattform YouTube typisch, ungefähr zwischen fünf Minuten und 15 Minuten lang. Die Rating Reveal-Videos greifen das in der Spielkulturebene schon vorhandene affektive Erwärmungspotential von Bewertungen auf und machen dies einer breiten Öffentlichkeit bei YouTube zugänglich. In den Videos, die nicht im Wortsinn spielerisch sind, aber doch figurativ, werden die Bewertungen der Spieler nochmals oder aufs Neue mit affektivem Wert aufgeladen. Außerdem veranschaulichen die YouTube-Clips die transmediale Einbettung der Spiele-Franchise von EA und anderer vergleichbarer digitaler Sportspiele in unsere gegenwärtige medienökologische Landschaft. Zwar bucht EA auch konventionelle Werbeflächen zur Promotion der jeweils erscheinenden Teile der Videospiel-Serie, so zum Beispiel im Fernsehen oder auf den Banden in designierten Stadien, jedoch stellen die Rating Reveal-Videos längst eine wichtige Gattung für das Marketing neuer Installationen der Franchises dar.

Dies macht aus zweierlei Gründen »ästhetisch« Sinn: Zum einen sind die Adressat*innen der Clips definitiv Teil einer digitalen Welt, sonst könnten sie schlachtweg nicht spielen; zum anderen sind diese quasi-Guerilla-Marketing-Strategien längst ein effektiver Teil einer Medienökologie, für die Authentizität ein zentrales Kriterium oder Leitbild darstellen. Und zwar besonders deshalb, weil im Kontext unserer gegenwärtigen Obsession mit »Reality« und allen Spielformen des vermeintlich oder eben nicht Realen, Nähe und Nahbarkeit wichtige Faktoren für die Steuerung von Marken-Images geworden sind. Schließlich profitieren von den Clips nicht nur die Macher bei Electronic Arts und die Verantwortlichen für die FIFA-Serie, sondern auch die Vereine und Spieler, die in den kurzen Clips vermeintlich tiefe Einblicke gewähren: in die Kantinen der Trainingszentren, das Essen der Spieler, die Trainingsräume und

Trainingsplätze, die hausinternen Produktionsstätten und natürlich den Alltag und das vermeintliche »Seelenleben« der Spieler als Protagonist*innen der Videos. Hier lassen die aufwendig produzierten Langzeit-Sportdokumentationen wie »Hard Knocks« und »All or Nothing« als wichtige Kontexte grüßen. Joshua Malitsky hat die Sportdokumentation als audiovisuelles Erzählgenre eindrucksvoll analysiert. Er etabliert hierbei die Sportdokumentation als generisch »imbricated with capital«, als »visually spectacular«, als ritualisierte Affirmation individuellen Ausdrucks, als immer schon spezifisch »narrativized« und als Beweis dafür, dass »thinking about sport always entails thinking about and through media« (2014, 205–206). Diese fünf Analysepunkte gelten für die digitalen Sportspiele in gleichem Maße und beide medialen Formen sind daher Teil der transmedialen Inszenierung gegenwärtiger Mediensportarten. Es lässt sich an dieser Stelle ergänzen, dass diese medialen Formen nicht nur immer schon spezifisch Sport erzählen (wollen) oder audiovisuell erzählbar machen, sondern eben auch immer schon erlebbar und affizierbar machen wollen.

Auch die *social media* Präsenzen der einzelnen Sportler*innen, deren Erfolg ja auf einer vergleichbaren Idee und ähnlichen Affekten fußt, bebildern nachhaltig den Profifußballsport gewissermaßen von innen. Diese Dokumentationen, Social Media-Accounts und YouTube-Videos stellen zusammen mit den digitalen Fußballspielen somit ein transmediales Projekt der Ästhetisierung von Profifußball dar: Gamer*innen bewerten Sportspiele oft mit einem Fokus auf die erfahrbare Atmosphäre, zum Beispiel der Spiele im Stadion, aber eben auch bezogen auf die Atmosphäre des Innenlebens des Clubs. So lassen sich die Entwickler*innen der Spiele zum Beispiel immer neue Cut-Scenes einfalten, bei denen transferierte Spieler in den Trainingszentren der Vereine – in den entsprechenden Farben und mit den Vereinswappen als Tapete – händeschüttelnd Interviews geben. Auch die aus der Vogelperspektive mit Drohnen gefilmten saftig-grünen Sportplätze der Trainingsgelände der gegenwärtigen Sportdokumentationsästhetik lässt sich als Teil dieses transmedialen Projektes verstehen. Nicht selten diskutieren Gamer*innen in den Foren zu den Spielen über die Farbe des Spielfeldes und dessen Beschaffenheit. Dass diese Formen und Spielarten der Sichtbarkeit bei YouTube für die Spieler lukrativ und reizvoll sein können, deutet die bereits zitierte Studie von Giovanni Bernardo, Massimo Ruberti und Roberto Verona eindrucksvoll an.

Zusätzlich drücken die Clips aus, wie sehr das Bewerten der Leistungsfähigkeit von Profifußballern (und Profifußballerinnen) einen zentralen Teil der Faszination der digitalen Sportspiele selbst darstellt. Die unterschiedlichen

Farben der Karten, die wir in den Videos gezeigt bekommen, nämlich die olympischen Medaillenfarben Bronze, Silber und Gold, sind ihrerseits schon Ausdruck des primären Bewertungsschemas durch das Spiel. Die Größe der zum Marketingzweck ausgehändigten Spielerkarten und die schelmenhafte Freude (sowie die allgegenwärtige, offensichtliche Frustration) der Protagonist*innen der Videos unterstreichen, dass die Bewertung der Spieler anhand von Leistungsdaten ein zentrales — wenn nicht das zentrale — Element der FIFA-Serie (EA FC) und vergleichbarer Franchises darstellt. Zumindest gilt dies für die doppelt Betroffenen, also die Spieler die wir in den Videos sehen und kurz begleiten. Innerhalb der Videos wird somit außerdem die nüchterne, kalte Statistiklogik übersetzt in das affektive oder emotionale Feld mithilfe der statistisch charakterisierten Spieler*innen als Menschen. Wenn man sich also fragt, wie so die ziemlich verbreitete Kritik an der Bewertung in diesen Videos Bestand hat und nicht geschnitten oder aus diesen verbannt wird, dann natürlich einerseits aufgrund der durch den Frust erzielten Authentizitätseffekte für die kleine Produktion, zum anderen aber auch, weil der Frust über die Bewertung sich wunderbar reimt mit dem weit verbreiteten Frust der Spieler*innen beim bereits beschriebenen Controllerwurf aus Wut.

Abb. II: Still Frame aus dem Rating Reveal Video für Bayer Leverkusen in FIFA 23



Anders gesagt, macht hier das Spiel sogar seine kalte Seite heiß: Die kühle Bewertung, die rationale Statistik und die »nackte Zahl« übersetzen die Videos in den Drang zur Flucht einzelner Spieler, in kindliches Kichern, lautes Geläch-

ter, Staunen, Stöhnen und die ganze Palette intensiver, nonverbaler menschlicher Kommunikation. Hier verdoppeln die Videos also nicht nur die Realität des Fußballspielens digital, sondern vielmehr die Affekte der Gamer*innen in den »heißen« Momenten der digitalen Fußballspiele selbst.

Abbildung II veranschaulicht exemplarisch dreierlei: Zum einen wird über das Logo (und Wasserzeichen des Videos) Bayer Leverkusens auch das Branding der jeweiligen Vereine mithilfe dieser Videos vorangetrieben. Dies wird in der schwarz-roten Farbgebung des Studios noch verdeutlicht; im vergleichbaren Video der Wolverhampton Wanderers findet sich entsprechend viel Orange und Schwarz. Zweitens sehen wir die beiden Spieler Frimpong und Hudson-Odoi in den üblichen Trainsingsoutfits, der auch die Lifestylisierung des so genannten »Athleisure« im Video zentral inszeniert: Sportartikelmarken sind längst Alltagsmodehersteller und hier wird diese Verbindung von sportlich aktiv sein und Alltag hergestellt. Drittens sehen wir in diesem Clip, wie die beiden Spieler lächelnd Richtung Studiotür blicken, weil der eigentlich dritte Mann des Videos soeben frustriert oder aus Protest dasselbe verlassen hat: Tim Fosu-Mensah fand seine Bewertung schlachtweg despektierlich und daher verlässt er den Frame des Videos. Dies führt in der Folge dazu, dass Frimpong und Hudson-Odoi vor Lachen von den Barstühlen fallen und zu Boden sinken. Ein Spieler flüchtet frustriert, während zwei andere ihn auslachen. Das Video trägt den zusätzlichen Titel, »Don't disrespect my guys!« (deutsch: Disrespektiere meine Jungs nicht!) und es inszeniert auf kraftvolle Weise die affektive Bedeutung der Bewertung für die so bewerteten. Die exaltierte Performanz des Auslachens verdoppelt zusätzlich die gängige Praxis der Hämme bei als Gesellschaftsspielen gespielten digitalen Sportspielen, zum Beispiel wenn wir Tore erzielen gegen unsere Gegner.

Eine weitere Ebene digitaler Fußballspiele als Bewertungsspiele findet sich in der Art und Weise, wie diese auch die Ökonomie des Fußballs spielbar und affizierbar machen. Diese Ökonomie der Bewertung informiert die Spiele auf unterschiedliche Weise. Wie bereits erwähnt, besteht eine wechselseitige Beziehung zwischen den Bewertungen der Spieler im Spiel und der Vergütung der Sportler des Profisports Fußball. Obiges Kammerspiel der Rating Reveal-Videos macht auf der Affektebene die Zahlen erfahrbar jenseits ihrer statistisch-organisatorischen Relevanz für das Spiel und den Fußballsport selbst. So erfüllen die Werte und die Bewertung ebenfalls den Zweck der Leistungsdagnostik und haben daher de facto einen Effekt bezüglich der Spielbarkeit der Abbilder auf dem Rasen. Im Allgemeinen ist es also so, dass die Bewertung entlang der einzelnen Parameter dabei hilft, Spieler in Typen einzuteilen und

erfassbar zu machen, mit dem stetig gestiegenen Anspruch, das Spiel insgesamt und mit den einzelnen Spielerabbildern »realistischer« zu machen. Realismus stellt im Diskurs zu digitalen Spielen ein immens komplexes Problem aber auch ein hohes Gut dar, das auch die Spieleforschung als Interessengebiet durchzieht: Spielphilosophie, -psychologie, und -narratologie haben allesamt ein Interesse am Verhältnis von Spiel und Umwelt, dem ich mit der Idee von Tele-Realismus begegne.

Hierbei folge ich dem Selbstanspruch der Spiele und dem Diskurs der Spieler*innen-Communities, die Realismus ständig und vehement zu einer wichtigen Kategorie der Beschreibung von Sportspielen machen und gemacht haben. Dies zeigt sich sehr deutlich auch in den vermeintlichen Abgrenzungen zwischen digitalen Sportspielen und digitalen Simulationen der Sportarten. In diesem Kontext stellt also die kleinteilige Erfassung von einzelnen Leistungswerten einen Beitrag zum Markt- und Markenwert der Spieler dar. Überdies stehen die Leistungsparameter in einem direkten Verhältnis zur ökonomischen Logik innerhalb der Spiele — also in-game — da die zentralen Modi dieser Franchises immer auch den Fußballmarkt abbilden und spielbar machen wollen. Schon bei Fußballspielen in den 1990er Jahren wie *Sensible World of Soccer* finden wir die Zuspitzung dieser Logik: Hier ersetzte der in-game Marktwert schlichtweg die Leistungsparameter der Spieler-Abbilder und es galt: je höher der Marktwert, desto besser und schneller der Spieler, auch wenn dies nicht weiter statistisch erfasst war (entlang der einzelnen Parameter). Es blieb aber im Spiel spürbar. In den heutigen Spielen sind die Bewertungsparameter ebenfalls mit dem Marktwert gekoppelt und dies findet sich meistens berechnet aus der Gesamtstärke der Spieler-Abbilder und dem Alter. Vereinfacht ausgedrückt heißt das: Je älter der Spieler, desto geringer sein Marktwert bei gleicher Gesamtstärke. Aber die Beziehung ist da und auch in diesem Sinne positivistisch: Ein teurer Spieler ist ein guter Spieler, oder anders formuliert: Der in-game Transfermarkt regelt sich im besten Sinne klassischer ökonomischer Theorie selbst. Der Transfermarkt der Spiele etabliert als zentrales Spielement wiederum eine Beziehung zum Transfermarkt »außerhalb« der Spiele, verdoppelt also auch die betriebswirtschaftliche Logik der Spiele und macht diese spielerisch erfahrbar. Der Transfermarkt im Spiel dient also einerseits explizit dem Ausdruck von Realismus und beeinflusst den Markt außerhalb des Spieles dabei gleichzeitig; in diesem Sinne verdoppeln die Spiele in der Tat die Realität des Wertes einzelner Spieler, wie Bernardo, Ruberti und Verona gezeigt haben. Zwar dürften sich die wenigsten Gamer*innen über die Verpflichtung eines guten Spielers un-

ter Marktpreis so freuen wie über ein Tor im WM-Finale in der Nachspielzeit der digitalen Spiele und dennoch werden in den Spielen auch Transfers von Spielern mit Affekten aufgeladen, denen wiederum eine Bewertung zugrunde liegt (wie in etwa über Wert oder unter Wert gekauft).

Jedoch endet die Wichtigkeit des Marktes und die ökonomische Logik der Spiele insgesamt nicht an dieser Stelle. Vielmehr haben die Spiele und besonders FIFA seit circa 2013 begonnen, einen weiteren Marktplatz einzupflegen, nämlich mithilfe des Internets und des online Modus *FIFA Ultimate Team* (kurz FUT genannt). Konami setzt bei dem Nachfolger der Pro Evo Serie sogar komplett auf ein online-only Model, das auf einer *free to play* Logik basiert. Hierbei ist die Idee tatsächlich, nicht das Spiel als Einheit selbst zu verkaufen, sondern Spieler-Abbilder innerhalb der Spiele zu verkaufen, zu verlosen oder frei spielbar zu machen. Spieler innerhalb der Spiele können also gekauft werden mit erspielten Spieldingen (*game points*), denen die Logik eines Bonusprogrammes für Vielflieger*innen oder Stammkund*innen zugrunde liegt oder aber direkt mit echtem Geld(Euro und Dollar), die Gamer*innen über die entsprechende Gamingplattform in das Spiel bringen können. So findet aktuell eine Debatte und ein juristischer Vorstoß seitens der in Deutschland staatlich organisierten Glücksspiel-Provider gegen das in digitalen Sportspielen und anderen digitalen betriebene so genannte *lootboxing* statt: Als Konsequenz ist EA FC im Jahr 2023 zum ersten Mal erst ab 12 Jahren freigegeben worden (Fritsche 2023).

Der Argwohn, den die Rating Reveal-Videos teilweise über die Bewertung der betroffenen Spieler*innen ausdrücken, gesellt sich zu jenem über die Monetarisierung der Spiele in ihren online Formaten und Modi — Gamer*innen, besorgte Eltern und politische Akteure bilden hier derzeit eine etwas kuriose Allianz. Die Punkte, *game points* oder *coins* und/oder die Euros, die wir als Spielende investieren, um dann im Lotterie-System Spieler zu erwerben, spiegeln gleichzeitig jedoch das Geschäftsmodell eines Transfermarktes, hier allerdings zusätzlich geprägt von der Logik, dass immer wieder neue Best-Bewertete Spielerkarten in den Markt gespült werden, für die es sich zu bieten lohnt. Anders gesagt, der Transfermarkt der online Modi bei EA FC und Konamis Efootball fußt auf der Idee eines Transfermarktes, jedoch wird dieser kontinuierlich mit neuen besten Spielerkarten im Stile einer ständigen Inflation übersättigt. Hinzu kommt, dass der Zugriff auf die online-Transfermärkte der beiden großen digitalen Fußballspiele durch die Glücksspiel-Komponente eines Lotterie-Glücksrades erfolgt. Die Entwicklung innerhalb der digitalen Sportspiele ist hierbei deutlich: Die Anreize, das zweite Spiel im Spiel zu spie-

len, also das Lotteriespiel am Transfermarkt, steigen stetig, nicht zuletzt, weil das ursprüngliche Primärspiel immer umfangreicher an die *in-game purchase* Elemente und seine Glücksspielkomponente gekoppelt wird.

Als Investitionsanreiz fungieren hier die mit 98 oder 99 gesamtbewerteten Abbilder von Messi, Maradona, Cruyff, Cristiano Ronaldo und dergleichen. Insofern nutzen FIFA Ultimate Team und E-Football die affektive Wärme und Nähe der Rating Reveal-Videos auch als Teaser für ein hiervon mittlerweile entkoppeltes Model der Monetarisierung. Was dabei auffällt ist, dass der ursprüngliche Anspruch einer »realistischen« statistischen Erfassung von Spielertypen und die allgemeinen, komplexen Austauschprozesse zwischen digitalem Fußballspiel und Fußball-Profisport dem exzessiven Ausloten der Grenzen von Super-Spielern mit immer höheren Gesamtbewertungen gewichen ist. Das vielschichtige Bewertungsspiel und die detailgetreue statistische Erfassung der einzelnen Spielertypen und deren Deutungshoheit, die in den Rating Reveal-Videos spürbar wird, tritt in den Hintergrund, wenn der Online-Transfermarkt Spielerabbilder regelmäßig und in großem Stile *boostet*, also aufbläht. Diese Glücksspiele-im-Spiel führen einen ursprünglichen Wesenskern digitaler Fußballspiele, die auf einer detaillierten und kenntnisreichen Bewertung der Spieler im Spiel aufbaute und deren Anspruch es war, Realitäten des Fußballs als Sport und Business im Spiel zu verdoppeln, ad absurdum. Wenn also EA selbstbewusst und mit Nachdruck verlauten lässt, »It's in the game!«, bin ich — zugegeben polemisch — geneigt nachzufragen, um welches Spiel es denn hier nunmehr geht?

Literatur

- Anable, Audre. Playing with Feelings: Video Games and Affect. U of Minnesota P, 2018.
- Berlant, Lauren. Cruel Optimism. Duke UP, 2011.
- Bernardi, Giovanni; Massimo Ruberti & Roberto Verona. »Image is everything! Professional football players' visibility and wages: evidence from the Italian Serie A«, Applied Economics (54, 5): 595–614, DOI: 10.1080/00036846.2021.1967863
- Blomberg, Johan. »The Semiotics of the Game Controller«, Game Studies (18, 2): no pg. <https://gamestudies.org/1802/articles/blomberg>, 14.02.2024.
- »Don't disrespect my guys!« FIFA 23 Prediction», <https://www.youtube.com/watch?v=CSVrgwot4ho>, 14.02.2024.

- Fritzsche, Lima. »FIFA-Nachfolger erst ab zwölf Jahren freigegeben«, Frankfurter Allgemeine Zeitung, 18.08.2023; <https://www.faz.net/pro/d-economy/ea-sports-fc-24-fifa-nachfolger-erst-ab-zwoelf-jahren-freigegeben-19112253.html>, 14.02.2024.
- Jenkins, Henry. *Convergence Culture: Where Old and New Media Collide*. New York: New York UP, 2006.
- Kempf, David. »Play the Game, Know the Game: Rekursive Wissenstransfers zwischen Football Manager und Fußball« in: Paidia: Zeitschrift für Computerspielforschung, 27.07.2017. <https://paidia.de/play-the-game-know-the-game-rekursive-wissenstransfers-zwischen-football-manager-und-fussball/>. [18.10.2023 – 07:42]
- kicker.de, »Nichts mit Glücksspiel gemein!« Game-Verband gegen weitere Lootbox Regularien«, <https://www.kicker.de/nichts-mit-gluecksspiel-ge mein-game-verband-gegen-weitere-lootboxen-regularien-973935/artikel>, 14.02.2024.
- Leonard, David J. *Playing While White: Privilege and Power on and off the Field* (U of Washington P, 2017).
- »Liverpool players react! HILARIOUS Jota & Konate partnership | EA SPORTS FC 24 ratings revealed:« <https://www.youtube.com/watch?v=Gn-9eMQBZMI>, 14.02.2024.
- Lüthe, Martin. »Playing on Fields: Seasonal Seriality, Tele-Realism, and the Bio-Politics of Digital Sports Games« in: Pöhlmann, Sascha (Ed.) *Playing the Field: Video Games and American Studies*. Berlin: DeGruyter, 2019: 97–112.
- Lüthe, Martin. »(Re-)producing the Body: Motion Capture and the Meaning of Physicality in Digital Soccer Games« in: Just, Peter & Inderst, Rudolf (eds.) *Build >em Up – Shoot >em Down: Körperlichkeit in digitalen Spielen*. Glückstadt: Verlag Werner Hülsbusch, 2013: 25–41.
- Malitsky, Joshua. »Knowing Sports: The Logic of the Contemporary Sports Documentary«, *Journal of Sports History* (41,2), 2014: 205–214, DOI: <http://doi.org/10.5406/jsporthistory.41.2.205>.
- Mittell, Jason. *Complex TV: The Poetics of Contemporary Television Storytelling*. New York, 2015.
- pes.neoseeker.com, https://pes.neoseeker.com/wiki/Team_Spirit, 14.02.2024.
- [pesmaster.com](https://www.pesmaster.com), <https://www.pesmaster.com/pes-2021-mobile/player/featured?date=2022-03-31>, 14.02.2024.
- Shaviro, Stephen. *Post-Cinematic Affect*. London: Zero Books, 2010.

Weheliye, Alexander G. *Habeas Viscus: Racialising Assemblages, Biopolitics, and Black Feminist Theories of the Human*. Duke UP, 2014.

Bewerten, um plausibel darzustellen

Researcher und Scouts für Football Manager

David Kempf

Fußballscouts haben die Aufgabe, die Stärken und Schwächen von Spielern möglichst präzise zu bewerten und damit über Management-Entscheidungen von (Profi-)Vereinen zu informieren. Dieser Beitrag beschreibt ethnografisch die auf den ersten Blick sehr ähnliche und lediglich graduell verschiedene Praktik der sogenannten ›Researcher‹, die ebenfalls Fußballspieler bewerten, dabei aber nicht auf Entscheidungen seitens eines Vereins abzielen, sondern auf die möglichst realistische Darstellung innerhalb eines Videospiels, der Fußballsimulation *Football Manager*. Die dabei herausgearbeitete Leitthese lautet, dass es sich hier um zwei zwar eng benachbarte, auf den zweiten Blick aber *qualitativ verschiedene* Bewertungspraktiken handelt: Zielt die Bewertung auf eine – öffentlich zugängliche – Darstellung ab, so wird ihr *Plausibilität* und nicht Reliabilität zum wichtigsten Gütekriterium.

An einem Fußball-Videospiel – einer in den Game Studies weitestgehend ignorierten Spielform (Leonard, 2006; Frome & Martin, 2020) – lässt sich dem potentiell besonders gut nachspüren, weil Sport-Spiele im Vergleich zu anderen Videospielen auszeichnet, eine besonders enge Korrespondenz zwischen realer und virtueller Realität (Esposito, 2019) anzustreben (Crawford & Gosling, 2009). Ferner ist diese Korrespondenz selbst für die Feldteilnehmer zentral und dezidiert thematisch (Crawford, 2006; Hocquet, 2016). Nicht die Gegebenheit dieser zwei verschiedenen Bewertungspraktiken zeichnet damit den hier untersuchten Fall aus dem Bereich des Sports aus, sondern dass deren *Verhältnis* zueinander innerhalb des Feldes (für die Teilnehmerinnen) Relevanz hat, problematisiert, theoretisiert und damit letztlich also *accountable* gemacht wird (Garfinkel, 1984).

1. Der Fall Football Manager¹

Football Manager erscheint – wie für Sportspiele nicht unüblich – alljährlich in einer neuen Version und wird vom relativ kleinen britischen Entwicklerstudio *Sports Interactive* produziert. Anders als beim Branchenprimus der Fußballvideospiele *FIFA* geht es in *Football Manager* nicht darum, einzelne Spieler² – vorzugsweise an einer Konsole – selbst zu steuern. Stattdessen nimmt man die Rolle von Trainer und Manager ein und hat die Aufgabe, Spieler für seine Mannschaft zu verpflichten, das Training zu steuern und eine Taktik zu entwerfen: Die Spieler werden dann gewissermaßen vom Computer gelenkt, ausgehend von den Attributen eines jeden Akteurs sowie den Einstellungen, die die Videospielerin vorher getroffen hat. Doch die eigene Partie macht an einem Spieltag bei *Football Manager* nur einen winzigen Teil aus: In zig verschiedenen Ländern und Ligen werden Partien simuliert, verletzen sich Spieler etc.:

We had spent quite a lot of time playing football games but they all essentially made you the centre of the universe [...] We wanted to make a functioning football world, and then put you into that world. (Paul Collyer, einer der beiden Gründer des Vorläufers *Championship Manager*, zitiert nach Parakin, 2015)

Diese Welt hat ihren Ausgangspunkt in einer ungemein detaillierten Annäherung an die reale Realität (Esposito, 2019) des Fußballs. So sind etwa 320.000 aktive Spieler aus 116 Ligen verzeichnet – mit Werten in bis zu 250 Attributen pro Spieler. Darunter fallen extrem diverse Aussagen etwa über Sprintgeschwindigkeit und Größe oder auch Aggressivität und Konzentration (Craw-

1 Den hier angestellten Beschreibungen und Überlegungen zum empirischen Fall *Football Manager* liegt die bald dreijährige ethnografische Forschung am empirischen Fall des Fußballsimulationsspiels *Football Manager* im Rahmen meines Dissertationsprojekts *Doing Realism. Praktiken des realistischen Darstellens am Fall der Fußballsimulation Football Manager* zugrunde (Veröffentlichung voraussichtlich Ende 2024). Im Sinne der theoretischen Empirie (Kalthoff et al., 2008) schließe ich keine gesonderte Theoriediskussion ein; im Geiste einer ethnografisch praxeologisch orientierten Forschung (Breidenstein et al., 2013; Schmidt, 2012) und angesichts begrenzten Platzes springe ich direkt in den empirischen Fall.

2 Bislang sind im Spiel ausschließlich männliche Fußballspieler simuliert – auch wenn sich das laut eigener Ankündigung seitens Sports Interactive (Jacobson, 2021) in absehbarer Zukunft ändern soll. Deshalb verwende ich in Bezug auf simulierte Fußballspieler ausschließlich das Maskulinum.

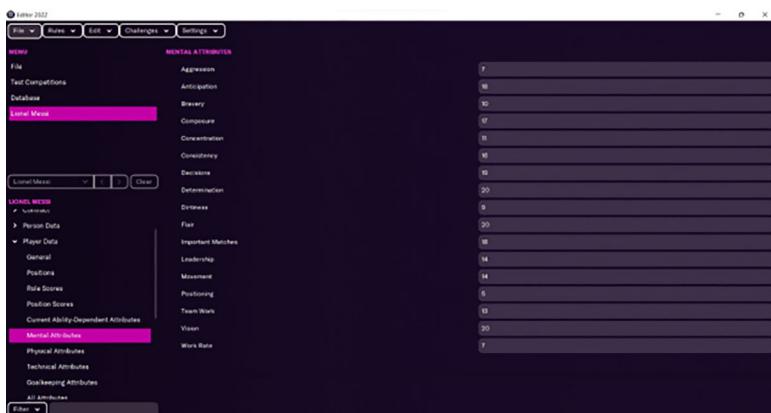
ford & Gosling, 2009). Mehr noch: Jeder Person ist ein gewisses Potential zugeordnet – darin steckt ebenfalls eine Aussage über ihre gegenwärtige Zukunft. Spieler entwickeln sich extrem unterschiedlich je nach ihrem Potential und zufälligen Faktoren wie etwa Verletzungen. So vermittelt sich dem Video-Spieler von *Football Manager* nicht nur ein detailliertes Bild aktueller Vereinskader, sondern durch das spielerische Betreiben der Simulation zunehmend auch eine Aussage darüber, wie es in Zukunft weitergehen könnte, bzw. mit einer gewissen Wahrscheinlichkeit weitergehen wird. Auf diese Verquickung sind die Entwickler besonders stolz und heben sie explizit immer wieder in der Bewerbung der Reihe hervor: *Play the game. Know the game.* lautete vor einigen Jahren etwa der zentrale Werbespruch für die frisch herausgebrachte Version. ›Game‹ und ›game‹ referiert hier offensichtlich zunächst auf die virtuelle und daraufhin auf die reale Realität (Kempf, 2017). Diese zwei Sphären immer enger miteinander zu verschweißen – und dabei in der virtuellen Simulation die Zukunft des realen Geschehens auf den Fußballplätzen vorherzusagen – ist also Anspruch der Reihe. Dass das beeindruckend gut gelingt, zeigt sich daran, dass die erhobenen Daten auf vielfältige Art und Weise performativ zu wirken beginnen: Professionelle Fußballvereine greifen zum Teil für ihr Scouting auf die Datenbanken von *Football Manager* zurück und bezahlen für exklusive Zugangsrechte; es besteht mittlerweile eine Kooperation zwischen dem Entwicklerstudio der Reihe *Sports Interactive* und einem professionellen Scoutingtool; Fußballspieler, Trainer und Fachjournalisten bekunden zum Teil öffentlich, das Spiel zu spielen und so unter anderem ihr taktisches Wissen auszubauen; Fans entwickeln andere Erwartungshaltungen an bis dato vielleicht vollkommen unbekannte, von der Datenbank aber als hoch talentiert eingestufte Spieler. Es entsteht ein dichtes Geflecht gegenseitiger Auswirkungen (Crawford, 2006; Hocquet, 2016).

Die Welt in den Computer holen: Researcher

Zentral für *Sport Interactives* alljährliche Bemühungen, eine möglichst realistische, dem Profifußball nahekommende Fußballsimulation zu entwickeln, sind die sogenannten *Researcher*. Ihre Aufgabe besteht darin, die erforderlichen Daten zu erheben, also ›die Welt in den Computer zu holen‹ (Gugerli, 2018). Konkret schlägt sich das darin nieder, eine von der zentralen Datenbank des Entwicklerstudios vorgegebene Datenmaske auszufüllen. Das beinhaltet grundlegende biographische Daten über Geburtsdatum, Größe, Gewicht, derzeitigen Verein inklusive der gültigen Vertragslaufzeit, eine Schätzung der Höhe

des Gehalts, Aussagen über die derzeitigen fußballerischen Fähigkeiten, die in Werten zwischen 1 und 20 in gut 50 unterschiedlichen Attributen wie Beschleunigung, Stärke, Passen, Konzentration etc. dargestellt werden, darüber hinaus eine Einschätzung der potentiellen Fähigkeiten des Spielers in der Zukunft sowie eine Zuordnung zu verschiedenen vorgegebenen Charaktereigenschaften. Ein großer Teil dieser Attribute ist im Spiel einsehbar, ein noch größerer im Editor, der beim Kauf des Spiels enthalten ist. Damit wird es den Videospielern ermöglicht, Daten ihren eigenen Vorstellungen entsprechend zu modifizieren.³

Abb. 1: Mentale Attribute von Lionel Messi im Editor. Still aus Football Manager 22.



Organisiert sind die Researcher nach Ländern: Für jedes im *Football Manager* dargestellte Land gibt es einen *Head-Researcher*, dem in vielen Fällen ein oder mehrere *Assistant-Researcher* unterstellt sind.⁴ Abhängig von der fußballerischen Prominenz der Länder sowie persönlichen Vorlieben der Head-Researcher variieren Anzahl und Einfluss der Assistant-Researcher beträchtlich: In England ist grundsätzlich für jeden Verein in der ersten Liga, der *Premier*

-
- 3 Ich beziehe mich hier auch auf sogenannte »Hidden-Attributes«, die in der mir zugänglichen, aber nicht zur Veröffentlichung autorisierten internen Datenbank von *Sports Interactive* bestehen. Darunter zählt beispielsweise »Versatility« oder »Sportmanship«.
 - 4 Manche Head-Researcher sind für mehrere Länder gleichzeitig verantwortlich; in mindestens einem mir bekannten Fall teilen sich zwei Head-Researcher, die jeweils eigenverantwortlich bereits ein Land verantworten, gemeinsam ein weiteres Land.

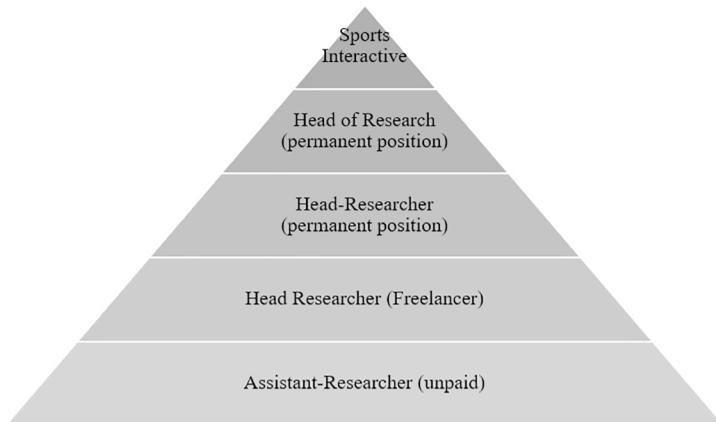
League, ein eigener Assistant-Researcher zuständig, in kleineren und fußballerisch weniger wichtigen Ländern gibt es manchmal nur einen oder gar keinen Assistant. Mir ist von einem Fall bekannt, in dem ein Head-Researcher eines kleineren Landes nur einen Mitarbeiter hat und dem lediglich die Einpflegung der zumeist vollkommen eindeutig abfragbaren Daten wie Alter, Vertragslaufzeit etc. überlässt. In anderen Fällen haben die Assistants bezüglich der ihnen anvertrauten Vereine bzw. deren Spielern und Mitarbeitern beinahe vollkommen freie Hand. Ferner werden Daten über Mitarbeiter wie Trainer, Sportdirektoren, Scouts etc. gesammelt.

Während die Assistants unbezahlt sind, erhalten die Head-Researcher als Freelancer ein Gehalt, das zum Leben allerdings nicht reicht. Der Zeitaufwand ist individuell sehr verschieden, je nachdem wie viele Assistants es gibt, wie viel Verantwortung denen wiederum überantwortet wird und nicht zuletzt wie ernst die Head-Researcher ihre Aufgabe nehmen und wie stark sie ins Detail gehen. Grundsätzlich sei ihre Arbeit nie abgeschlossen, gebe es eben immer noch mehr Spieler noch detaillierter zu sichten. Es wird so bereits erkennbar, dass die Head-Researcher über große Freiheiten verfügen. Unterstellt sind sie dem Head of Research und seinem Team, angesiedelt in der Londoner Zentrale von *Sports Interactive*. Der Head of Research bekleidet diese Rolle ebenso wie der Chef des Unternehmens, Miles Jacobsen, schon seit Jahrzehnten. Er ist seit 28 Jahren bei der Firma angestellt. Neben dem Head of Research gibt es seit jüngstem einige Head-Researcher, die fest angestellt werden, »internal« sind, wie es der Head of Research, Martin, im Gespräch mit mir ausdrückt. Diese fest angestellten Head-Researcher dienen gewissermaßen als Verbindungsstück zwischen der Londoner Zentrale und den Freelancern. Sie sind ca. acht bis zehn anderen Head-Researchern zugewiesen, zu denen sie regelmäßig Kontakt halten, als erster Ansprechpartner bei Problemen fungieren und indirekt zumindest ein Stück weit kontrollieren, ob die Arbeit erledigt wird.

Die formale Unterstellung führt im Arbeitsalltag allerdings keine direkte Kontrolle mit sich: Keiner der Head-Researcher, mit denen ich gesprochen habe, hat laut eigener Aussage jemals direkte Kritik oder sonstiges konkretes Feedback für eigene Bewertungen vom Head of Research erhalten. Alexandre Hocquet (2016, S. 42) bezeichnet die Datengewinnung für *Football Manager* als »crowdsourced«. Es stellt sich also die Frage, wie eine solche Heerschar an national verankerten Head Researchern – Stand Anfang 2023 sind es ca. 100 – mit jeweils einer weiteren Vielzahl an Assistants koordiniert werden kann. Wie wird konkret eine grundsätzliche Kohärenz der Bewertungen sichergestellt, so

dass nicht beispielsweise die österreichischen Spieler viel besser bewertet werden als die dänischen?

Abb. 2: Organigramm der Organisation der Researcher bei Sports Interactive



Um die grundsätzlich offensichtlich erfolgreiche organisationale Koordination und Sicherstellung eines gemeinsam geteilten Bewertungsrahmens erklären zu können, gibt es mehrere Erklärungsansätze. Zum Ersten beruht die Disziplinierung der Head-Researcher auf einem großen Maß an Identifikation sowohl mit der eigenen Rolle als auch mit dem Arbeitgeber und dem aus der Arbeit resultierendem Produkt, also dem Spiel. Über die unbezahlten Assistants – »volunteers« – sagt der Head of Research im Interview: »They want to do it, because they love the game, they love their club, they want to see their club represented correctly in-game.« Disziplinierung ist also primär Selbstdisziplinierung (Bröckling, 2017). Ausnahmslos alle Head-Researcher, mit denen ich gesprochen habe, waren vor ihrer Tätigkeit leidenschaftliche Spieler von *Football Manager*, die spielerisch die jeweiligen Datenbanken im Editor modifiziert und diese Änderungen mit *Sports Interactive* geteilt haben (vgl. zum Verhältnis von Spiel und Arbeit unter anderem Stevens, 1980; T. L. Taylor, 2012). Oft ist dieser Austausch über das offizielle Online-Forum zum Spiel gestaltet. Einer meiner Gesprächspartner erzählt mir aber auch davon, wie er vor weit über 15 Jahren als ›playtester‹ ein paar Wochen bei *Sports Interactive* in London gearbeitet und in dem Rahmen dem Head of Research eine CD mit seinem ei-

genen aktualisierten Datenset präsentiert habe. Der Wechsel in die Rolle des Head-Researchers war für diese Männer insofern ein vergleichsweise nahtloser Übergang innerhalb des eigenen Hobbys: von Spielern, die in ihrer Freizeit an Überarbeitungen der Datenbanken arbeiten – Axel Bruns (2006) spricht von »produsern« – zu formal Angestellten, deren Arbeit zunächst nur kaum stärker vorgegebenen Regeln folgte und nicht entlohnt wurde. Bis heute verbindet diese Head-Researcher eine enge affektiv aufgeladene Beziehung zu *Football Manager*. Daraus leitet sich ein hoher Anspruch ab, diesem so innig geliebten Spiel in ihrer Arbeit gerecht zu werden. Besonders deutlich wird das an einer Stelle aus meinem Gespräch mit Werner, der darüber berichtet, mal ein finanziell äußerst lukratives Angebot abgelehnt zu haben, für EA Sports' Spielreihe *FIFA* zu arbeiten:

W: Also ich hab von EA Sports Angebote bekommen, damals, für FIFA zu arbeiten . da hätt ich das verdient im Monat, was ich da im Jahr hab.

D: Echt?!

W: Ja.

D: Und warum haste das nicht gemacht?

W: Moraleische Ansichten.

D: ›Moralisch‹, weil... erklär!

W: Moraleische Ansichten aus dem Grund, weil ich erstens mal mit FM aufgewachsen bin, den seit meiner Kindheit spiele. Weil ich FIFA absolut nicht ausstehen kann vom Aufbau her. Vom Spieler her, jeder Spieler kann alles. Ich kann mit jedem Spieler, kann ich alles machen. Unrealistisch. Ist nicht dabei. Äh. Die Bewertungen bei FIFA [unverständlich: Beispiel?] mit diesen 1 bis 100 macht dieses Spiel schwach. Weil zwischen 96 und 93 ist kein Unterschied. Beim FM hät man da schon Unterschiede zwischen 180 und 190. Das ist eben das moralische, wo ich sag: Dieses Spiel ist einfach nur zur Belustigung der Menschheit gedacht.

[...]

Drum sag ich ja, also unglaubliches Geld, was bei EA Sports da eigentlich machbar gewesen wäre, aber für mich ist Geld nicht wichtig in dem Sinne. Weil für mich muss alles zusammenpassen, das ganze Konzept muss- ich muss auch meine eigenen Werte dahinterstehen können und sagen ja, ich als [unverständlich] Werner möcht mich mit diesem Spiel identifizieren. Für das Arbeiten halt ich meinen Kopf hin und sag ja, das ist mein Gesicht, ich hab das gemacht. Das ist bei FIFA nicht der Fall. (eigenes Interviewmaterial)

Werner macht »moralische Ansichten« dafür verantwortlich, das finanziell attraktive Angebot ausgeschlagen zu haben. Dabei geht es ihm nicht primär um die eigentliche Tätigkeit im Berufsalltag, sondern um das dabei entstehende Gesamt- und Endprodukt. *FIFA* sei bloß zur »Belustigung der Menschheit« gedacht und eben nicht realistisch. Moralische Bedenken bezüglich der Spielreihe *FIFA* würden sich erstmal eher in Bezug auf deren System der ›Loot-Boxen‹ anbieten, mit denen Hersteller *EA Sports* neben dem Spielverkauf über in-game Käufe enorme Summen an Geld verdient und deshalb unter anderem dem Vorwurf ausgesetzt ist, eine Form des – zudem Minderjährigen zugänglichen – Glücksspiels zu betreiben (Xiao et al., 2022). Darüber spricht Wolfgang aber überhaupt nicht. Es ist für unseren Gesamtkontext äußerst interessant, dass Werner seine moralischen Bedenken mit dem Anspruch an Realismus und der Abgrenzung von Bespaßung begründet. Sogar die affektive Bindung an *Sports Interactive* und deren Produkt ist vom Anspruch an Realismus nicht zu trennen.

Dieser Anspruch wird zusätzlich getragen durch ein hohes Maß an informellem Austausch zwischen den Head-Researchern. *Sports Interactive* ermöglicht diesen Austausch – wie auch die kontinuierliche Identifikation – mit alljährlich organisierten Treffen für alle Head-Researcher in der Londo-ner Zentrale. Innerhalb einer knappen Woche wird den Head-Researchern dabei ein buntes und vor allem spaßiges Programm geboten: Treffen mit einem Premier-League Manager, gemeinsame Kneipentouren, Stadionbesu-che usw. Die Head-Researcher, mit denen ich sprechen konnte, arbeiten in westeuropäischen Ländern und kennen einander gut, stehen beispielsweise in regelmäßigerem Austausch über WhatsApp. Einige bezeichnen sich gegenseitig als Freunde. Darüber hinaus hat sich für alle meine Gesprächspartner seit ihrem Einstieg als Head-Researcher einiges getan: Aus einem Hobby wurde erst ein Nebenjob und dann der Übergang in ein Berufsleben, das im weiteren Feld des Profifußballs verortet ist. Diese Entwicklung verläuft gewissermaßen parallel zu der des Entwicklerstudios *Sports Interactive*, das in den letzten Jahren seine Mitarbeiterzahl verdoppelt hat und sich laut Aussage des Head of Research kontinuierlich professionalisiert, ein ›klassischeres‹ Unternehmen mit festen Kennzahlen, Zielen etc. wird. Das weite Feld des Profifußballs beinhaltet die Arbeit für Wettanbieter, Spielerberatungsagenturen oder auch direkt Profivereine. Vom Wechsel eines ehemaligen Head-Researchers in die Festanstellung bei einem größeren deutschen Fußballverein erzählen mir gleich mehrere Gesprächspartner bewundernd und etwas schwärmerisch – die Anekdote gilt ihnen als Beweis: Die Tätigkeit als Head-Researcher ermöglicht es, beruflich ein Teil der so begehrten Welt des Profifußballs zu werden –

diesen nicht mehr nur in Form einer Simulation darzustellen, sondern aktiv mitzugestalten. So hat für meine Gesprächspartner mit zunehmendem Alter die hobby-bedingte Motivation zwar etwas nachgelassen, wurde aber durch ein professionelles Streben ersetzt. Die Arbeit wird so auch Mittel zum Zweck. Diese zwei Leitmotive fasst *Sports Interactive* explizit als Pole, die die Arbeit als Researcher attraktiv machen. So steht beispielsweise im offiziellen Forum der Entwickler in einem Thread, der darum wirbt, sich zu bewerben:

Our researchers are fans of the series, just like us, and many of them took on their roles because they were passionate about their clubs, domestic leagues and national football setups and they wanted to ensure they were represented accurately in the game. Indeed, many of those who have worked as researchers for Football Manager have moved into jobs in the footballing world as a result of their knowledge and commitment, so it can be a pathway to the job of your dreams. (Brock, 2021)

Neben dem regen Austausch unter Kollegen stellt der Austausch mit Spielerinnen von *Football Manager* eine wichtige soziale Kontroll- und Feedbackfunktion dar. So wie die amtierenden Head-Researcher einst selbst, kommentieren unzählige Spieler der Simulation die Bewertungen von Fußballspielern, üben direkte Kritik daran oder machen Gegenvorschläge. Beispielhaft ist folgender Post aus dem offiziellen Forum von User tbm_cooper (2022):

Moin zusammen,
dies ist mein erster Beitrag hier im Forum. Ich würde gerne etwas loswerten [sic!] zur Bewertung einiger Spieler von Werder Bremen. Ich hoffe dies ist hier der Richtige Platz dafür. Im Prinzip würde ich gerne einfach nur in den kurzen Austausch mit dem deutschen research team gehen, bzw. einfach eine kurze Anregungen weitergeben. [...]

Ich denke einige Spieler sollten überarbeitet werden. Und generell sich vielleicht Gedanken über die Gesamtstärke des Team gemacht werden. So wie es aktuell in der beta [frei zugängliche unfertige Version des Spiels, die wenige Wochen vor offiziellem ›Release‹ herauskommt, D.K.] aussieht, steigt Werder mit diesem Team und diesen Spielern in 9 von 10 Fällen ab.

Das [sic!] Werder sich im Bereich Abstiegskampf befindet ist in Ordnung; allerdings sind sie aktuell zu deutlich Abstiegskandidat Nummer 1.

Tbm_cooper (2022) geht dann noch einige Einzelspieler durch und schreibt unter anderem:

Man hat bei Friedl im FM den Eindruck, dass seit 3 Jahren kein Blick auf ihn geworfen wurde. So wie er aktuell aussieht, sieht er seit Jahren aus. Er hat sich allerdings [sic!] besonders letzte Saison deutlich entwickelt in der Realität. Was die Position und die Persönlichkeit angeht.

Die Head-Researcher verstehen sich nach wie vor zumindest teilweise als Teil dieser Community und nehmen deren Meinung insofern als wichtiges Verbindungsstück zur im Blogpost angesprochenen »Realität« ernst. Das betont auch der Head of Research. Von mir darauf angesprochen, wie es bei über 100 als Freelancer arbeitenden Head Researchern möglich sei, eine Form der Qualitätskontrolle durchzuführen, verweist er sofort auf die aktive Community und deren Feedback. Er spricht in diesem Kontext von »external checks and balances«. Letztlich beruht die gesamte in Jahrzehnten aufgebaute Datenakquise über Fußballspieler auf einer ungemein engen und vertrauensvollen Einbindung der aktiven Spieler und Fans der Simulation. Das zeigt sich anhand des Werdegangs der Head-Researcher von leidenschaftlichen Hobbyisten hin zu Freelancern, die ihre Tätigkeit als Chance zum Sprung in ein neues Berufsleben begreifen, dabei aber in inniger Verbundenheit mit ihrem Arbeitgeber und dem von ihnen geschaffenen Produkt, dem Spiel, verbunden bleiben. Darauf fußt das Vertrauen, das *Sports Interactive* diesen in ihrer Arbeit kaum zu kontrollierenden Menschen entgegenbringt. Emblematisch für dieses Vorgehen ist eine Anekdote aus den Anfängen des Spiels in den 90ern – damals noch *Championship Manager* – die mir der Head of Research erzählt: Als das damals noch winzig kleine Team neben England und Schottland weitere Länder in die Datenbank integrieren wollte, sei allen Beteiligten klar gewesen, weder über die Zeit noch über das nötige Hintergrundwissen zu verfügen, um Spieler aus anderen Ländern wie beispielsweise Deutschland ausreichend gut bewerten zu können. Die Lösung: *Sports Interactive* kontaktierte *Fanzines* von Fußballvereinen aus den entsprechenden Ländern und bat diese darum, ihrer Leserschaft einen Aufruf zur Bewertung der Spieler des eigenen Vereins zukommen zu lassen. In einigen wenigen Kernattributen sollten die Fans Bewertungen abgeben zwischen – – und + +, also in insgesamt fünf Zwischenschritten. Die Resonanz sei mit einer Antwortrate von ca. 80 % ungemein positiv gewesen – die Bewertungen wurden so mehr oder weniger übereinandergelegt und dann übernommen. So stark auf eine Heerschar an leidenschaftlichen Fans der Serie als Researcher zu setzen, lässt sich als professionalisierte Fortführung eines solchen Vorgehens verstehen. *Sports Interactive* bindet Begeisterte in die Produktion des Spiels ein, so dass diese Begeisterten zu Beteiligten werden und sich folglich

noch enger mit dem Spiel verbunden fühlen. Das gilt auch dafür, die Bewertungen von Spielerinnen und Spielern wie Tbm_cooper ernstzunehmen – und zwar als Bewertungen der eigenen Bewertungen, als Ressource also, um sich zu verbessern.

Gleichzeitig gibt es vonseiten der Head-Researcher aber auch Abgrenzungsbewegungen zur Feedback-Kultur aus dem Internet, etwa wenn sich Head-Researcher Werner mir gegenüber im Gespräch impulsiv darüber beschwert, Kritik werde zum Teil persönlich beleidigend oder beziehe sich auf aus seiner Sicht irrelevante Kleinigkeiten: »Wenn vielleicht *passing* um zwei Punkte nicht passt, oder *finishing* um zwei Punkte nicht passt. Naja, gut. Erschlag mich.«⁵ Diese gewissermaßen öffentlich geäußerten Kommentare stellen die Head-Researcher vor die Aufgabe, selektiv im Sinne der Membran Goffmans Ein- und Ausschlüsse organisieren zu müssen. Instruktiv ist dafür insbesondere in Form von Forderungen geäußerte Kritik. So erzählt mir Werner: Vor einigen Jahren habe es erhebliche Aufregung darum gegeben, dass der Dortmunder Youngster Youssoufa Moukoko nicht Teil der Datenbank des damals neuen *Football Managers* war. Moukoko galt damals als weltweit bekanntes Toptalent, wurde bereits im Alter von 15 Jahren Teil des Dortmunder Profikaders und als 16-jähriger zum jüngsten Bundesligadebütanten aller Zeiten. Ein Spieler also, den die Videospieler der Simulation nur allzu gerne verpflichtet und aufgrund seines offensichtlich großen Talents weiterentwickelt hätten. Ein Spieler aber, der gemäß den – insbesondere juristisch motivierten – Vorgaben *Sports Interactives*' eigentlich noch zu jung war, um ihn in die Datenbank zu integrieren. Doch der vor der Veröffentlichung der nächsten Version des Spiels aufgebaute Druck zeigte Wirkung. Werner:

Moukoko war der erste Spieler, der es geschafft hat, ohne die Altersbeschränkung für das Spiel zu haben, in das Spiel zu kommen. Auf Twitter wurde ja ne mörderische Stimmung gegenüber *Sports Interactive* gemacht, dass Moukoko im Spiel sein muss. Da muss [unverständlich] eine offizielle Anfrage machen, damit der Spieler im Spiel sein darf und bla bla bla und hin und her. (Eigenes Interviewmaterial)

5 »[...] here is a basic similarity between gaming encounters and other types of focused gatherings: both can be taken too seriously, both not seriously enough.« (Goffman, 2013, S. 40) Werner verweist hier offensichtlich auf Goffmans ersten Punkt – darauf, dass das Spiel seines Erachtens hier zu ernst genommen und damit als solches verfehlt wird.

Wenn ein Fußballverein einen 15-jährigen als Teil des Profikaders führt, dann muss die nach Realismus strebende Fußballsimulation nachziehen. Das Realismusgebot wurde in diesem Fall gewissermaßen öffentlich eingefordert, die Forderung vonseiten *Sports Interactives* inkludiert. Doch Werner berichtet mir auch von Fällen, in denen junge Spieler seines Erachtens grundlos »gehyped« würden und davon, wie er jahrelang mit der Kritik konfrontiert war, einen jungen, medial recht bekannten Spieler angeblich in Bezug auf dessen Potential zu schlecht zu bewerten – bis der Spieler schließlich wie von ihm vorhergesagt »gefloppt« sei, so dass er Recht behalten habe. Offenbar im Internet geäußerte Kritik der Community von Spielerinnen und Spielern dient also als direktes Feedback für die Head-Researcher – die den mit diesem Feedback verbundenen Forderungen allerdings selektiv begegnen und nicht einfach blind vertrauen können.

Schließlich wird eine letzte, zentrale Koordinationsfunktion von der virtuellen Datenbank selbst ausgeübt. Koordinierend ist diese Datenbank auf zwei verschiedenen Ebenen. Erstens legt sie grundsätzlich fest, in welchen Kategorien was überhaupt zu erfassen ist – und was nicht. Damit verankert die Datenbank als Infrastruktur »rules of inclusion and exclusion« (Waibel et al., 2021, S. 42). Sie legt für alle Researcher verbindlich fest, dass beispielsweise Konzentration als Attribut Teil der realistischen Darstellung eines Fußballspielers ist, Fairness hingegen nicht. Damit bestimmt sie, worauf die Researcher achten, in welchen Kategorien sie schlussendlich also bewertend zu sehen lernen (Hoppe, 2022; basaler zum Erlernen eines spezifisch geschulten Sehens Fleck, 2019). Neben der spezifischen Auswahl von Attributen als Kategorien dient das in allen Attributen gleichermaßen geltende metrische System zur Koordination und Bildung von Kohärenz. Ob physische Attribute wie Beschleunigung, mentale Attribute wie Konzentration oder gar Attribute von Akteuren, die nicht Fußballspieler sind, sondern etwa Scouts – beispielsweise »Judging Potential Ability«: Stets wird ein numerischer Wert zwischen Eins und 20 eingetragen. So entsteht prinzipiell Vergleichbarkeit – zwischen den verschiedenen Spielern, aber auch zwischen den unterschiedlichen Attributen selbst. Qualitativ grundverschiedene Attribute werden so quantitativ miteinander in Relation gesetzt und dadurch kommensurabel (Espeland & Stevens, 1998).

Die Datenbank als Infrastruktur koordiniert also erstens, indem sie die zu bewertenden Kategorien und das metrische Bewertungssystem setzt. Das strukturiert die Bewertungspraktiken der Researcher fundamental und nimmt eine gewissermaßen transzendentale Stellung ein – der Blick der

Researcher ist immer schon gelenkt durch die feststehenden Attributkategorien und die damit einhergehende Affordanz (Gibson, 2015), stets einen Wert zwischen Eins und 20 dafür zu finden.

Zweitens existiert eine etwas direkter gelenkte Art der Koordination, die sich an von der Datenbank errechneten Durchschnittswerten von Spielern, Mannschaften und Ligen festmacht. Diese Durchschnittswerte zu errechnen, ist nicht trivial, da nicht einfach ein Durchschnitt aller Attribute gebildet wird, sondern je nach Position und Rolle verschiedene Attribute unterschiedlich stark gewichtet werden. In der Folge gilt, dass sich Werte nur durchs Vergleichen relativ bestimmen lassen. Ein erster grober Anhaltspunkt für den Researcher ist also ein stetiges Vergleichen. Ist der Durchschnittswert des soeben bewerteten Spielers aus der 2. Bundesliga einem Stammspieler von Real Madrid ähnlich, so ist offensichtlich etwas schiefgelaufen. Für die Durchschnittswerte bestehen darüber hinaus aber auch konkrete Vorgaben seitens *Sports Interactive*: So werden beispielsweise die unterschiedlichen Ligen in eine Hierarchie gebracht – die dann bei den Durchschnittswerten eingehalten werden muss. In der Folge entstehen Richtkorridore an Durchschnittswerten für Ligen, die hinuntergebrochen werden auf die einzelnen Mannschaften – hier wiederum muss beispielsweise derzeit der FC Bayern München über den stärksten Durchschnittswert der Bundesliga verfügen – die wiederum die Bewertungen der einzelnen Spieler prägen. Diese sehr grob gesetzten Vorgaben sickern so gewissermaßen immer tiefer hinab und bestimmen letztlich zum Teil maßgeblich, wie ein Spieler bewertet wird. Zwar stellt der Head of Research im Gespräch mit mir überzeugend klar, dass diese Vorgaben nie absolut, mit einer guten Begründung jederzeit Abweichungen davon möglich seien. Die Arbeitspraxis der Researcher scheint von diesen Vorgaben aber doch stark beeinflusst zu sein. So berichtet ein Head-Researcher, dass die vorgegebenen Durchschnittswerte insbesondere seine Bewertungen in den weniger relevanten Kategorien eines Spielers prägen: Zunächst trägt er Werte in den Attributen ein, die ihm bei dem Spieler besonders aufgefallen sind, ihn seines Erachtens auszeichnen. Dann bleiben aber mitunter noch unzählige Attribute frei, die er nach einer nicht allzu detaillierten Sichtung kaum fundiert festlegen kann (»Ist der Defensivzweikampf des Stürmers jetzt acht oder elf?«). Es entstehen Unterschiede, die aus Sicht mehrerer Researcher dann auch nicht wirklich einen Unterschied machen, gerade bei weniger prominenten Spielern aus niedrigeren Ligen. So werden sehr frei und unbedacht Werte eingetragen und dann anhand des daraus entstehenden Durchschnittswerts nachkorrigiert, so dass am Ende ein Profil des Spielers entsteht, das

einerseits die aus Sicht des Researchers besonders hervorstechenden Merkmale berücksichtigt, andererseits einen Durchschnittswert ergibt, der in der errechneten Gesamtstärke die Stellung des Spielers in der Mannschaft den Erwartungen entsprechend wiedergibt. So übernimmt die ‚Zentrale‘ durch Hierarchisierung von Ligen und vorgegebenen Korridoren für Durchschnittsstärken eine indirekte – mit Foucault (2009) gesprochen: *gouvernementale* – Lenkungsfunktion, die sich sehr konkret in den Bewertungen einzelner Spieler niederschlägt. Eine offensichtliche Folge dieser Strategie besteht in einer strukturell angelegten performativen Konservativität: Ergäbe sich aus der höchst aufwendigen und detaillierten Bewertung aller einzelnen Spieler vonseiten des Head-Researchers Deutschlands, dass die Durchschnittsstärke der Mannschaft von Borussia Dortmund dieses Jahr stärker sei als vom FC Bayern München, so müsste er das mindestens der Zentrale gegenüber rechtfertigen – denn Bayern München ist in den letzten Jahren ausnahmslos jede Saison Meister geworden, verfügt also zumindest in der historischen Rückschau über den erfolgreichsten Kader der Liga. Aus den Gesprächen, die ich geführt habe, geht aber eher hervor, dass die Head-Researcher sich diesen Vorgaben zumeist fügen, ohne sie infrage zu stellen, die Vorgaben folglich als sinnvoll anerkennen. Dabei bricht eine solche Setzung dezidiert mit einem Ideal offenen, vorurteilsfreien Bewertens. Beispielsweise wäre es denkbar, dass Dortmund tatsächlich über die besseren Einzelspieler verfügt, diese aber als Kollektiv weniger erfolgreich zusammenarbeiten – beispielsweise, weil der Trainer schlechter ist. *Sports Interactive* folgt hier aber einer anderen Logik, die eben nicht offenes Entdecken vorrangig setzt, sondern realistische, der Realität treue *Darstellung*. Die Realität soll nicht anders und neu beschrieben, sondern möglichst überzeugend *nachgebildet* werden. Dieser Unterschied zwischen den Ansprüchen an Entdecken und Darstellung wird noch deutlicher, wenn man rekonstruiert, weshalb *Sports Interactive Researcher* beschäftigt und keine *Scouts*.

2. Researcher und Scouts

Scouts, zu Deutsch etwa Späher oder Kundschafter, werden von Fußballvereinen entsandt, um entweder die nächste gegnerische Mannschaft auszukundschaften oder um als Verstärkungen für die eigene Mannschaft in Frage kommende Spieler zu sichten. Von besonderer Bedeutung ist dabei das Scouting junger, noch Entwicklungsfähiger Spieler, stets verbunden mit der

Hoffnung, einen Talentschatz zu heben. Die auf eine immer größere Bedeutung von Wissen, Methodik und Einsatz von Technik und Quantifizierung basierende Transformation des Profifußballs der letzten Jahre (Schmidt, 2015) hat auch auf die Arbeit der Fußballscouts fundamentale Auswirkungen.⁶ Zum einen hat die voranschreitende Kommerzialisierung und damit einhergehende ökonomische Konkurrenz zwischen den Profivereinen zu wachsender Bedeutung des Scoutings beigetragen: »Jugendfußballer waren immer einfach da gewesen, wie der Sand am Meer. Nun wurden sie das gehegte, umkämpfte Kapital der Profivereine.« (Reng, 2015, S. 91) In der Folge beschäftigen Vereine mehr und zuweilen habituell andere Scouts – war es früher gang und gäbe, diese Aufgabe schlachtweg ehemaligen Spielern anzutrauen, so gibt es mittlerweile einige Quereinsteiger – insbesondere gilt das für Fachgebiete wie das »Datenscouting«, ist aber keineswegs darauf beschränkt. Doch auch die praktische Vorgehensweise wurde fundamental transformiert: Vereine bauen große Datenbanken mit Spielern auf, fordern von Scouts, nicht nur den einen interessanten Spieler zu sichteten, sondern weniger tiefenscharf alle an einem Spiel beteiligten Spieler. Die Ergebnisse werden zumeist in vorgegebenen Kategorien quantitativ festgehalten, um von da aus auch in die Datenbank eingeschleust werden zu können. Zum anderen wurde der ›bloße Blick‹ des Scouts technisch um eine Vielzahl an statistischen Daten und Werten sowie umfassendem Videomaterial erweitert. War die Arbeit des Scouts vor zwanzig Jahren noch zutiefst vom unablässigen Reisen geprägt (Reng, 2015, S. 177), so findet mittlerweile zumindest ein großer Teil der Vorauswahl per Videosichtung vor dem Bildschirm im Büro statt. Das Videoscouting entbindet dabei nicht nur teilweise von der zeitintensiven Tätigkeit des Reisens, es bietet auch eigens zugeschnittene Highlight-Sequenzen – etwa alle von einem festgelegten Spieler geschlagenen Flanken innerhalb einer Saison – und per ›Scouting-Feed‹ eine stets das ganze Feld überblickende Vogelperspektive, die in manchen Gesichtspunkten, etwa in Bezug auf mannschaftstaktisches Agieren ohne Ball, sogar der physischen Anwesenheit auf einem Nebenplatz ohne erhöhte Tribüne überlegen ist (Brümmer, 2019). Das hat unter anderem zur

6 Die folgenden Ausführungen beziehen sich unter anderem implizit auf ethnografische Daten, die ich gemeinsam mit Robert Schmidt und Max Weigelin im an der KU Eichstätt angesiedelten und von Robert Schmidt geleiteten DFG-Projekt ›Accounting und transformatorische Effekte im Profifußball‹ erhoben und interpretiert habe. Von besonderer Bedeutung ist dabei eine Reihe an Interviews, die wir mit angehenden, jungen Scouts geführt haben.

Folge, dass es mittlerweile schon in niedrigen Altersklassen ab einem gewissen Niveau quasi unmöglich ist, einen noch vollkommen unbekannten Spieler zu entdecken – und das weltweit, bis hin zu fußballerisch exotischen Ländern. Die Scouts bringt das untereinander in eine merkwürdige Lage: Einerseits ist das Geschäft von Konkurrenz geprägt, davon, *besser* zu sichtzen, mehr Wissen zu beschaffen als die anderen Vereine in ähnlicher sportlicher oder ökonomischer Lage, um daraus einen Marktvorteil zu generieren. Andererseits sorgt die Zunahme von vorhandenem Videomaterial und gemeinsam geteilter und auf technischen Infrastrukturen beruhender Methodik beinahe für Deckungsgleichheit bezüglich des tatsächlich vorhandenen Wissens. Das wird anschaulich an der folgenden Stelle aus *Mroskos Talente*, der journalistischen Biografie eines ehemaligen Scouts:

- ›Ich finde, auch für einen Jugendspieler wie Kroos müsste ein Klub einfach mal zwei Millionen hinlegen. Warum nicht?‹
- ›Ach, Kroos. Den habe ich schon vor über einem Jahr entdeckt.‹
- ›An dem bin ich schon viel länger dran.‹
- ›Was meint ihr, was ich mit dem Vater Kroos schon geredet und getan habe, sagte der große Rehberg. ›Ich habe bei uns auf Schalke den Verantwortlichen gesagt: Da können wir die ganze Familie holen, der Vater Kroos ist ein guter Trainer, beide Söhne können kicken, da kannst du gar nichts verkehrt machen. Aber dann kostet der Junge natürlich 'ne gewisse Summe, und für einen 15-Jährigen zu zahlen, kann sich unser Manager Rudi Assauer nicht vorstellen, habe ich gehört.‹
- ›Noch 'ne Flasche Rotwein bitte!‹
- ›Die Bayern sind doch schon längst an Kroos dran, wusste Mrosko. ›Die haben ihn im Februar zum Champions-League-Spiel gegen Arsenal eingeladen, Hotel Vier Jahreszeiten, VIP-Karten, und am nächsten Tag hat ihn Uli Hoeneß in seinem Büro empfangen – einen 15-Jährigen. So macht man das heute.‹
- ›Natürlich wird's für uns nahezu unmöglich, wenn die Bayern aufkreuzen. Aber du musst es trotzdem weiter probieren, wenn es sich um ein Ausnahmetalent wie Kroos handelt, sagte der große Rehberg. ›Ich erzähle euch mal, wie lange wir gebraucht haben, einen kleinen Türken zu kriegen, 88er-Jahrgang, für mich das größte Talent im ganzen Ruhrpott.‹
- ›Wer? Der Özil, oder was? Den kenne ich natürlich schon lange.‹
- ›Den kennt doch jeder hier.‹ (Reng, 2015, S. 93–94)

In kollegialer abendlicher Runde im Rahmen eines wichtigen deutschen Jugendturniers fachsimpelt ein Großteil der damals wichtigsten Fußballscouts des Landes miteinander. Deutlich nachvollziehbar wird dabei die Spannung zwischen dem Bild des einsamen, wertvolles Geheimwissen generierenden Kundschafters einerseits und einem Kollegen unter Kollegen, die zumindest in Bezug auf die Top-Talente längst ohnehin auf demselben Stand sind. »Den [Özil] kennt doch jeder hier.«⁷ Das darf allerdings nicht darüber hinwegtäuschen, dass es sich hier um eine seltene Ausnahmesituation handelt, Konkurrenz und Geheimniskrämerei vorherrschen (Darr & Mears, 2017, S. 7).

Einige Gemeinsamkeiten zwischen Researchern und Scouts drängen sich so bereits auf: Zunächst einmal sind die Alltagspraktiken extrem eng miteinander verwandt: Fußballspielern bei der Arbeit zusehen und ihre Fähigkeiten – in Gegenwart wie antizipierter Zukunft in Form von Potential – bewerten. Jonas, einer der Head-Researcher, den ich interviewt habe, sagt dazu:

Mhh, I bumped into a [anonymisierte Nationalität] scout, who gave us scouting training, scouting courses, there was one near where I lived, so I decided to go there and he was giving a course and there were couple of other people there, scouts, want to be scouts, people on the football business, and that moment I noticed: Hey, I'm seeing the same things as these guys are seeing – so I can do this. [Hervorhebung D.K.] (eigenes Interviewmaterial)

Jonas leitet die empfundene Nähe zwischen den beiden Aufgaben aus der geteilten Fähigkeit ab, richtig zu sehen und folgert daraus im weiteren Verlauf, inwiefern er seines Erachtens dazu befähigt sei, auch im professionellen Fußball zu arbeiten. Mittlerweile ist dieses Sehen beider Gruppen in metrisch organisierte (und organisierende) Infrastrukturen der Bewertung eingelassen und auf die Integration in digitale Datenbanken ausgerichtet. Sichtet ein Scout einen Spieler hingegen detaillierter, greift er auch auf qualitative Bewertungen zurück und erarbeitet ‚Wert-Urteile‘. Konkret bedeutet das beispielsweise, ›Waffen‹ eines Spielers hervorzuheben und schlussendlich eine Empfehlung abzugeben: Sollte der Verein diesen Spieler verpflichten oder nicht? Mehrere Scouts erzählen mir dazu in unseren Gesprächen, dass es wichtig sei, eindeutige Empfehlungen zu generieren: Ein Interviewpartner

7 Auf die Bedeutung solcher kollegialen Zusammenkünfte, die immer auch ein Teilen der eigenen, eigentlich geheimen, Daten beinhalten, verweisen Darr und Mears (2017, S. 11) in Bezug sowohl auf Fußball-, als auch in Bezug auf für die Modebranche nach neuen Modells suchenden Scouts.

erklärt, es sei eine Qualität von Scouts, »angreifbar« zu sein – »ieber drei Mal falsch liegen als nichts gesagt haben«. Als Scout dürfe man sich eben nie zu Unsicherheiten und schwammigen Aussagen verleiten lassen, sondern müsse am Ende klar etwas aussagen – auch und gerade auf die Gefahr hin, falsch zu liegen. Gerade hier komme quantitativen Daten eine entscheidende Bedeutung zu, nämlich nicht als Such- oder Analysewerkzeug, sondern im Sinne einer Fundierung der vorab getroffenen eigenen Einschätzung. Letztlich ist der Scout in der Organisationsstruktur »Vorarbeiter, nicht Entscheider« (Interview mit dem bereits zuvor zitierten Scout). Zwei Spieler als ungefähr gleich stark zu bewerten, entlaste den Entscheidungsträger keineswegs. Deshalb sei eine möglichst eindeutige Bewertung und daraus resultierende Empfehlung gefordert, auch wenn sich diese irgendwann als falsch herausstellen sollte, weil der entsprechende Spieler nach der Verpflichtung den Erwartungen nicht gerecht wird. Dazu passt – ganz im Sinne der weiter oben zitierten Passage –, dass die alltägliche Arbeit der Scouts nicht darin besteht, vollkommen unbekannte Spieler zu ›finden‹, sondern aus einem großen Pool potentiell in Frage kommender Spieler Kandidaten auszuschließen, den Pool also zu verkleinern.⁸ Insofern besteht eine fundamentale Aufgabe des Scouts darin, Kontingenz und Komplexität auf der Entscheidungsebene zu reduzieren. Die damit verbundene Forderung nach Eindeutigkeit trotz faktischer Unmöglichkeit absoluter Sicherheit erinnert stark an Elena Espositos (2019) Charakterisierung von quantitativen Marktprognosen nicht als Zukunftsvorhersagen einer »zukünftigen Gegenwart«, sondern als Orientierungspunkte im Hier und Jetzt, die in einer »gegenwärtigen Zukunft« operieren. Diese Aufgabe haben die Researcher nicht. Entscheidungen zu fällen, ist Aufgabe der Videospieler. Empfehlungen dazu geben die vom Spiel simulierten Scouts ab. Die Aufgabe, aus einem Datenkorpus Empfehlungen abzuleiten, fällt der Simulation zu. Die Arbeit von Scouts und Researchern konvergiert also tendenziell in Bezug auf das Anlegen eines Pools an Spielern, differiert aber entscheidend bezüglich detaillierterer Bewertungen von Einzelspielern und den daran geknüpften spezifischen Anforderungen. Scouts müssen eine weitere, gewissermaßen tiefenschärtere Ebene bedienen können – ein rein quantitativer, granularer Unterschied. Daneben existiert allerdings ein weiterer, tiefgreifend qualitativer Unterschied, der sich durch die Fokussierung auf die sich beiden Gruppen stellende Herausforderung, auch und gera-

8 Ein Pool, den dieselben Scouts freilich selbst aufbauen und pflegen.

die ›weichen‹ Faktoren wie Charaktereigenschaften der Spieler zu bewerten, herauskristallisieren lässt.

3. Charaktereigenschaften bewerten: der qualitative Unterschied

Nichts ist für einen Fußballverein gewonnen, wenn ein aus einem fernen Land für viel Geld verpflichteter Spieler sich zwar als genauso technisch beschlagen herausstellt wie erhofft, jedoch mit den neuen Lebensumständen in einem ihm vollkommen fremden Land nicht zurechtkommt. Ein unglücklicher Spieler kann technisch noch so herausragend sein – die Chancen, diese Fähigkeiten auf den Platz zu bringen, sind gering. Folglich besteht eine der wichtigsten Aufgaben der Scouts darin, etwas über den ›Charakter‹ des infrage kommenden Spielers herauszufinden. Gerade dieses Aufgabenfeld ist technisch nicht oder nur eingeschränkt abzudecken und erfordert die physische Anwesenheit des Scouts, der beispielsweise bei den Aufwärmübungen oder beim Torjubel genau hinsieht. Auch dem jeweiligen Netzwerk kommt eine entscheidende Bedeutung zu:

As part of their engagement in recognitionwork, coaches and scouts also tried to learn if the player drank or gambled, if he demonstrated loyalty to his previous teams, and if he was ready to care about others and not only about himself. To obtain this type of local knowledge they tried to speak with previous coaches and teammates of the candidates, sometimes in remote locations. At time they even tried to gain such information from Israeli sportsjournalists who have ties with other journalists in the player's home-country or city. (Darr & Mears, 2017, S. 10)

Die Bewertung solcher weichen Faktoren entzieht sich einem Großteil der tendenziell ›gleich-schaltenden‹ Methodik und technisch getriebenen Transformation. Folglich sind hier auch die größten Distinktionsgewinne zwischen den Vereinen zu erzielen: ein richtig guter Scout hat auch für derlei potentielle Probleme das ›feine Näschen‹.

Für die Scouts ist es vollkommen evident, jederlei verlässlich wirkende Informationen oder auch nur plausiblen Gerüchte in die eigene Bewertung miteinzubeziehen. Für die Researcher hingegen entfaltet sich hier eine eigene Form der Komplexität. Das möchte ich an einem langen Auszug aus meinem Gespräch mit Jonas festmachen. Jonas ist Anfang 40, Head-Researcher

und neben dieser Tätigkeit freiberuflich im Profifußball tätig. Insbesondere arbeitet er scoutend für Wettanbieter und immer mal wieder für einzelne Vereine, allerdings bislang stets auf Honorarbasis und nicht fest angestellt. Der Übersichtlichkeit halber ist der Auszug in mehrere Teile gegliedert und von kurzen Erläuterungen gerahmt.

J: I mean I've had, I've had a coach enter information a few years ago and he actually did performance reviews with the players.

D: Ok (lacht).

J: And then we were talking, we were e-mailing and it was like: Okay, I can actually put this in and I was like: you know, you don't need to put that in. You're welcome to, but you don't have to put it in. Maybe if you have the few players where you're like: Wow! This kid is- has such a fantastic attitude- okay, put that in. But if you have two or three players where you just know, ok, this kid has a shit attitude, he will not make it. You know what? We don't really have to put that in, so we just- leave it blank, you know? Putting the information in where you yourself are still comfortable with putting that information in. I personally, I speak to scouts, I've spoken to directors of football and they personally give me information where I'm like: wow. Why are you telling me this? I will not put that in. (eigenes Interviewmaterial)

Jonas berichtet hier von einem Fall, in dem einer seiner Assistant-Researcher zugleich als Trainer bei einer von ihm bewerteten Mannschaft gearbeitet hat. Dieser Assistent habe angeboten, die von ihm als Trainer ausgestellten »performance reviews« der Spieler in die spielegene Bewertung einfließen zu lassen. Jonas weist ihn darauf hin, er »müsste diese Informationen nicht einfließen lassen«, sei dazu aber eingeladen. Nachdem der Assistent das aber bereits vorhatte, ist das als höflich indirekte Abweisung des Vorhabens zu verstehen. Stattdessen schlägt Jonas vor, so nur in Bezug auf besonders positive Bewertungen einzelner Spieler vorzugehen und den Rest außen vor zu lassen. Schließlich überführt Jonas die Anekdote zur allgemeiner gefassten Aussage, diverse Akteure aus dem Profifußball hätten ihm schon Informationen zukommen lassen, die er nicht einfließen lasse. Sein Verständnis für diese Indiskretionen fasst er in der rhetorischen frage »Why are you telling me this?« zusammen.

D: And that's kind of information, that, for instance, concerns the behavior of the players besides the pitch, probably-

J: Sometimes, yes. I, I know of players that have really, really bad attitudes,

usually they are also already a little bit in the media, so it's more a confirmation, but I've been told certain situations. And also about people at clubs, like directors, where they were just- one person tells me: that guy is incompetent. A second guy tells me, oh, that guy is incompetent. There are situations where I've been told by about ten or fifteen people about the same person and then it's like, okay, maybe, if one or two people tell me [unverständlich], ok somebody just likes somebody, that happens, but if it's fifteen people telling you from all different situations, different countries, like, ok, it's probably true.

D: Ok, but you would still not incorporate this into the database, or slightly?
J: No. Mhh. If it's out there in the media, if there are situations where it's like, okay, this player has been late five times, he has had a fight with another player, then we can put it in. If it's not public, I don't put it in. (eigenes Interviewmaterial)

Einer Vielzahl deckungsgleicher Gerüchte sei tendenziell Glauben zu schenken – so, wie wenn beispielsweise 15 unterschiedliche Kontakte demselben Sportdirektoren Inkompotenz bescheinigen. Nichtsdestotrotz würde Jonas auch ei- ne derlei für ihn als bestätigte geltende Aussage nicht für seine Bewertung ver- wenden. Das entscheidende Kriterium sei vielmehr, ob das Gerücht in den Me- dien aufgegriffen werde, also ›öffentlicht‹ sei.

D: And that's due to, let's say moral concerns, right?

J: -yes-[...]

J: A very, very long time ago, in the early days of Football Manager, it was still Championship Manager, we had a certain code in the game, that you could put in there or leave out which was called »ARSE«. So, the famous Arse Com- mand, if you put it in the code then you will have situations where a player, for example, will headbutt a referee.

D: Okay, ya.

J: And then he will get a suspension for like 90 days or six months or what- ever. But they took it out, because they were scared – or maybe it even hap- pened, I'm not precisely fully sure – but there were situations where a player or a club would say: Hey, come on! This is something negative about the player, this is too much, [unverständlich] discredit – we will sue you.

D: So it's moral and legal concerns.

J: Yes. I think it's more legal than moral. For me, personally, I think it as a researcher and I think, you know, this is maybe something that somebody tells me in confidence- it could even be a lie. And then I put it in the game and then that person will go: Ha – what an idiot, I told him a lie and he put

it in the game. That guy is not serious. Even that could happen.

D:Ja.

J: But for me, it's like, you know, someone is telling me something, even in confidence, there's no reason for me to put it in the game. Sometimes I will make notes; of course, I will know it, maybe I will even write it down and then in two years there will be: ›Oh, there's a big fight at the club!‹ or you will see a player fighting with another player where I knew, hey, wait a minute, these two players don't like each other. And then it's like, okay, you know what?, we can put this in. (eigenes Interviewmaterial)

Als weiteren Grund, Gerüchte aus seinem Netzwerk zurückzuhalten, gibt Jonas die Möglichkeit an, in die Irre geführt, belogen zu werden. Der imaginierte Täuscher würde dann entdecken, dass es seine Lüge ins Spiel geschafft hätte und das als Ausweis dafür nehmen, Jonas sei ein »idiot« und »not serious«. In dieser fiktiven Darstellung handelt es sich also um eine Art Test seiner Seriosität und Integrität. Die besondere Bedeutung, die das für seine Rolle hat, streicht er dadurch heraus, dass er einleitend einschiebt: »I think it as a researcher«. Aus Gesprächen mit anderen Researchern ist mir bekannt, dass es immer mal Versuche insbesondere seitens Spielerberatern gebe, positiv Einfluss auf Bewertungen der von ihnen vertretenen Spielern zu nehmen. Das schwingt in Jonas' Darstellung sicherlich implizit mit. Schließlich räumt er ein, sich solche Gerüchte natürlich zu merken, manchmal auch aufzuschreiben. Komme es dann irgendwann später zu einem sich damit deckenden journalistischen Bericht – wird das Gerücht also öffentlich –, greife er darauf zurück und lasse es einfließen. Medien nehmen in dieser so von Jonas geschilderten Vorgehensweise also eine nicht primär informierende, sondern absichernde Rolle ein.

Massenmedien sind also nicht in dem Sinne Medien, daß [sic!] sie Informationen von Wissenden auf Nichtwissende übertragen. Sie sind Medien insofern, als sie ein Hintergrundwissen bereitstellen und jeweils fortschreiben, von dem man in der Kommunikation ausgehen kann. Die konstituierende Unterscheidung ist nicht Wissen/Nichtwissen, sondern Medium und Form. Das Medium stellt einen riesigen, aber gleichwohl eingeschränkten Bereich von Möglichkeiten bereit, aus dem die Kommunikation Formen auswählen kann, wenn sie sich temporär auf bestimmte Inhalte festlegt. (Luhmann, 2017, S. 84)

Luhmanns Konzeption der Rolle von Massenmedien möchte ich nicht eins zu eins auf unseren Fall zu übertragen versuchen. Entscheidend und außeror-

dentlich fruchtbar ist allerdings die quasi-transzendentale Stellung der Medien, die nicht auf einer »Wissen/Nichtwissen Unterscheidung« beruht. So läge es auf den ersten Blick nahe, im Fall von Jonas die Medien als klar nachgelagert zu betrachten: *Erst* hört Jonas mehr oder weniger verlässliche Gerüchte, *dann* wartet er ab, ob sie medial bestätigt werden. Diese Hierarchie ist aber – mit Luhmann – in der auf Wissen abzielenden Unterscheidung konzipiert. Tatsächlich stellt sich ja aber gerade heraus, dass für Jonas das Wissen im engen Sinne *zweitrangig* ist: Ist dieses Wissen nicht medial gedeckt, nicht Teil des »riesigen, aber gleichwohl eingeschränkten Bereich[s] von Möglichkeiten«, so kann es nicht verwendet werden. Es wird für die Darstellung im Videospiel wertlos.

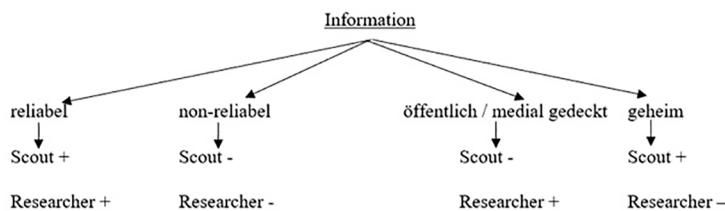
Die Beschränkung auf veröffentlichte Quellen kann darüber hinaus auch als weitere organisationale Antwort auf das weiter oben diskutierte Problem der Koordination einer Vielzahl kaum direkt kontrollierter Head-Researcher interpretiert werden: So findet sich unter dem Begriff »No Original Research« eine höchst ähnliche Maxime auch in Bezug auf die Leitplanken für Editorinnen bei Wikipedia, wo es offiziell heißt: »all material in Wikipedia must be attributable to a reliable, published source« (Wikipedia, 2023, zitiert nach König, 2013, S. 163). Auch eine renommierte Wissenschaftlerin könnte ihre Forschungsergebnisse also nicht mit Berufung auf ihre fachliche Anerkennung direkt bei Wikipedia publizieren, sondern müsste die Ergebnisse erst in einer klassischen Publikation unterbringen, um diese dann als Quelle angeben zu können. Sowohl *Sports Interactive* als auch Wikipedia versuchen in diesem Sinne gar nicht erst, aufgrund der jeweiligen Kredibilität ihrer Zuarbeiterinnen deren Bewertungen zu valorisieren, sondern setzen auf indirekte Kontrolle durch Rekurs auf bereits bestehende Publikationsorgane (vgl. zum Vergleich der Datengewinnung zwischen Wikipedia und *Football Manager* auch Hocquet, 2016): Eingang findet nur, was von einer zweiten Instanz medial gedeckt ist – in Bezug auf *Football Manager* gilt das allerdings natürlich nur für die hier diskutierten ›weichen‹ Charaktereigenschaften und nicht für einen Wert wie ›acceleration‹.

Dass ihre Bewertungen in ein frei zugängliches Computerspiel fließen und damit potentiell selbst öffentlich werden, hat fundamentale Auswirkungen auf die Arbeit der Researcher. Auch ein Scout muss natürlich seine Quellen prüfen, er muss allerdings keinerlei Rücksicht darauf nehmen, ob diese Informationen öffentlich zugänglich sind. Im Gegenteil: Informationen, die nur ihm zur Verfügung stehen, sind besonders wertvoll, da sie seinem Verein einen Wettbewerbsvorteil den Mitbewerbern gegenüber verschaffen. Die Be-

wertungen eines Scouts führen potentiell zu einer Entscheidung seitens seines Vereins – sie geraten aber nicht an die Öffentlichkeit. Im Fall der Researcher ist das entscheidend anders. Das führt Jonas dazu, die Leistungsbewertung seines gleichzeitig als Trainer arbeitenden Assistenten nur in besonders positiven Fällen zuzulassen. Das Problem dieser Bewertungen ist dabei gerade, dass sie vollkommen verlässlich wirken *und* direkt aus einer internen Quelle stammen. Was für einen Scout absolut wünschenswert wäre, wird für den Researcher zur problematischen Verletzung der Intimsphäre der bewerteten Spieler. Auch wenn Ungenauigkeiten der Bewertungen und Gerüchte möglich sind, spielt diese Unsicherheit nicht die eigentlich entscheidende Rolle. Überdeutlich wird das an der Anekdote zum »arse command« aus früheren Zeiten: dem Algorithmus, der nach dem Zufallsprinzip zu groben Unsportlichkeiten einzelner Spieler führte. Hier ging es offensichtlich gar nicht um Bewertungen von Researchern. Nichtsdestotrotz erschien den Entwicklern – ob aufgrund tatsächlicher Klagen oder nicht – das Risiko zu groß, damit Unmut seitens Spielern oder Vereinen zu erwecken. Im Videospiel getroffene Bewertungen sind nicht nur einer breiten Öffentlichkeit zugänglich, sondern natürlich auch den Bewerteten selbst. Das erweitert die Sphäre moralischer Bedenken um die potentielle Sorge vor direkten Beschwerden oder gar juristischen Klagen.

Folglich gilt grundsätzlich, dass in Bezug auf ›weiche‹ Attribute Scouts tendenziell geheimes Wissen besonders wertschätzen, das für Researcher nicht verwendbar und deshalb nur von geringem Wert ist. Andersherum hat öffentlich medial gedecktes Wissen für Researcher den höchsten Wert, da es sich nutzen lässt. Das lässt sich wie folgt visualisieren:

Abb. 3: Visualisierung der unterschiedlichen Bedeutungsgrade von Information für Scouts und Researcher



Die Reliabilität der gewonnenen Information spielt für Scouts und Researcher grundsätzlich dieselbe Rolle: für beide sind möglichst reliable Daten wünschenswert – was für die Scouts sicherlich nochmal wichtiger ist. Anders verhält sich das in Bezug darauf, ob die jeweilige Information geheim oder öffentlich zugänglich ist – hier stehen Scouts und Researcher einander diametral gegenüber.

Für Scouts ist die Reliabilität notwendige Bedingung, wohingegen der Grad an Zugänglichkeit den relativen Wert der Information bestimmt: Wird der jeweilige Spieler als absoluter Vorbildcharakter eingeschätzt, stellt sich nach der Verpflichtung aber als spielsüchtig und zunehmend unzuverlässig heraus, war die Arbeit des Scouts nicht nur wertlos, sondern potentiell schwer vereinsschädigend. Ist die Spielsucht des Spielers dagegen ein offenes Geheimnis, kann der Verein zwar danach handeln, die Information hat für den Scout aber bloß geringen Wert. Für die Researcher wiederum ist die mediale Deckung der Information notwendige Bedingung – und die Reliabilität bestimmt den relativen Wert: Bestehen Medienberichte über die Trainingsfaulheit eines Spielers, so kann der Researcher das in die spielinterne Datenbank aufnehmen; stellt sich über die Zeit hinweg jedoch heraus, dass die Medien hier ein falsches Bild gezeichnet haben, verliert diese Information stark an Wert.

Beide Gruppen bewerten Fußballspieler; bei Scouts dienen diese Bewertungen aber dazu, innerhalb der Organisation des jeweiligen Vereins intern eine Entscheidung zu informieren, wohingegen die Bewertungen der Researcher der prinzipiell öffentlich zugänglichen Darstellung dienen. Damit wird *Plausibilität* zum entscheidenden Qualitätskriterium der getroffenen Bewertung.

Fazit

Der Fokus auf Plausibilität anstelle von möglichst absoluter Exaktheit erhellt die Überzeugung, mit der Werner kleinteilige Kritik an seinen Bewertungen zurückweist: »Wenn vielleicht *passing* um zwei Punkte nicht passt, oder *finishing* um zwei Punkte nicht passt. Naja, gut. Erschlag mich.« Diese Zurückweisung läuft in diesem Lichte betrachtet nicht der Maxime zuwider, das Feedback der spielenden Community ernst zu nehmen. Vielmehr ist sie als Diagnose eines Kategorienfehlers zu verstehen: Hier hat jemand spielerische mit professioneller Darstellung verwechselt – es geht nämlich eben nicht

darum, eine Entscheidung auf Management-Ebene zu treffen, sondern eine für das realistische Spielerlebnis ausreichend plausible Darstellung zu bieten. Das Feedback von Tbm-cooper, der mit der Bewertung des Fußballspielers Friedl so gar nicht mehr zufrieden ist (»Man hat bei Friedl im FM den Eindruck, dass seit 3 Jahren kein Blick auf ihn geworfen wurde.«), passt dagegen ins Schema und kann so interpretiert werden, dass es gewissermaßen das richtige Bewertungsregime ansteuert. In dieses Schema passt folglich auch sowohl die starke, »crowdsourcende« Einbindung begeisterter Spieler als researchende »prosumer« als auch die statistischen ›Erwartungshorizonte‹, in die durchschnittliche Attributstärken von Spielern, Mannschaften und ganzen Ligen passen sollen: Für Scouts würden solche Verfahren die erhebliche Gefahr bergen, gerade auch die blinden Flecke der allgemeinen Einschätzung zu replizieren und damit die Generierung einzigartiger und damit wirklich wertvoller Daten zu verunmöglichen. Was aus dieser Perspektive zutiefst nachteilig ist, wird für die Researcher nicht mal zum notwendigen Übel, sondern *zum eigentlichen Trumpf*, da Datengenerierung und Erwartungshaltung eng und konstant miteinander verschaltet werden.

Gleichzeitig muss betont werden, dass diese Trennung von Bewertungsregimen idealtypisch zwei Pole aufzeigt, die in der Arbeitspraxis von Scouts und Researchern längst nicht immer derart trennscharf markiert sind. Wichtiger Bestandteil des Arbeitsalltags von Scouts ist es, die eigenen Arbeitsergebnisse einer organisational höheren Führungs- und Entscheidungsebene zu präsentieren – und dann spielt die Plausibilität der eigenen Bewertungen eine entscheidende Rolle. Dazu passt die bereits zitierte Aussage eines Scouts, es sei wichtiger, konkrete und entschiedene Aussagen zu treffen als schwammige oder gar keine. Dazu passt aber auch der im Feld gewonnene Eindruck, dass quantitative, »moderne« Bewertungsregister offensichtlich in vielen Fällen vor allem eingesetzt werden, um dem qualitativ bereits gewonnenen Eindruck mehr Autorität zu verschaffen, ihn zu *plausibilisieren*. Auf der anderen Seite stehen die Researcher, die wiederum unter anderem das Ziel verfolgen, die Stars der Zukunft ›richtig vorherzusagen‹ und sich dezidiert von so manchem Hype eines jungen Spielers distanzieren. Wie gesagt handelt es sich also um zwei idealtypische Pole, die sich anhand des Vergleichs der Bewertung derart ›weicher‹ und intimer Attribute wie Charaktereigenschaften herauskristallisieren. Offensichtlich wird dabei auch, dass Plausibilität immer dann an Relevanz gewinnt, wenn es darum geht, Bewertungen für andere *anschlussfähig* zu machen. Sportvideospiele – darauf wurde bereits in der Einleitung verwiesen – bieten für solche Beobachtungen einen besonders geeigneten Rahmen

insofern das Verhältnis von Darstellung und Dargestelltem hier für das Feld relevant wird und in besonderer, dramatisierter Form zum Ausdruck kommt.

Abschließend möchte ich drei weitere mögliche Fälle vorstellen. In zwei Fällen stehen auf Plausibilität abzielende Praktiken ganz nah neben auf direktes Entscheiden abzielenden Praktiken. Dieser Ausblick dient vor allem dazu, zumindest anzudeuten, dass es sich hier um eine weiter verallgemeinerbare Beobachtung handelt und wie diese sich empirisch auf ein breiteres Fundament setzen ließe.

Ausblick: Drei weitere mögliche Fälle

Über plausible Darstellungen lässt sich erstens im Feld der Künste lernen. Eingängig ist hier die klassisch realistische Literatur. So erörtert Barrish (2011) die Bedeutung, die der Nennung spezifischer Preise im Roman *Sister Carrie* zukommen:

That Carrie habitually notes prices does tell us something about her character (as well as her need to watch pennies at this early point in her career). And of course the steak and chicken prices must remain within a certain range of historical plausibility. Still for the purposes of the plot, for conveying character, or even for suggesting the sort of restaurant the Windsor was in 1889 (the year in which Dreiser tells us the book is set), it would not matter if the chicken cost sixty-five cents, eighty-five cents, or even ninety cents. In other words, the precise content of the detail, the number itself, is less significant than the sheer fact that Dreiser provides it. The very presence of such specific prices in Sister Carrie connotes actuality, regardless of whether or not these prices are truly identical with those on the Windsor's 1889 menu. In and of itself, the dollars-and-cents price signifies concrete reality to the reader. (Barrish, 2011: S. 45–46)

Die Spezifität, mit der Preise auf den Cent genau genannt werden, dient hier als Stilmittel, um das im Roman angestrebte hohe Maß an Realismus zu plausibilisieren.

Ein zweiter weiterer Fall stellt das von Waibel et al. (2021) beschriebene *Signaling* bei Bewertungsprozessen in akademischen Auswahlverfahren geeigneter Bewerber dar:

In competitive higher education systems, for instance, university leaders tend to use criteria in their internal evaluation that can also be used for signaling the scientific success of their universities to relevant others – like funding agencies, private donors, government departments of science, or students. Therefore, researchers must signal qualities that their valued valuator can use for their own signaling of quality. As a consequence, researchers do not need to convince the university leadership that they are great scientists, but rather that they can help to convince relevant others that the university excels at research. (Waibel et al., 2021, S. 52)

Wissenschaftler müssten folglich Universitäten nicht davon überzeugen, »greatest scientists« zu sein, vielmehr stellen sie darauf ab, überzeugend auf die Beurterinnen ihrer eigenen Bewerter, also beispielsweise Geberinnen von Drittmitteln oder Studierende, zu wirken. In diesem Sinne kommuniziere eine Bewerbung nicht, inwiefern der sich Bewerbende ›tatsächlich‹ gut für die Anforderungen der Stelle geeignet sei, sondern in welchem Maße seine Berufung es der Universität erlaubt, dadurch wissenschaftliche Qualität an relevante Dritte zu ›signalisieren‹. Eine solche Bewerbung antizipiert also die Bewertungen, denen die Universität als eigentlich bewertende Arbeitgeberin selbst ausgesetzt ist und adressiert sie direkt, wobei zumindest der Tendenz nach die Gefahr besteht, dass die mittlere Bewertung der Universität schlicht wegfällt. Nicht gute Forschung und Lehre stehen hier im Fokus, sondern die Bewertung Dritter – die mehr oder weniger weit außerhalb des eigentlichen Feldes stehen. Was hier also primär angesteuert wird, ist eine *plausible* Bewertung für diese relevanten Dritten.

Ähnlich gelagert ist eine Herausforderung, die sich für professionell arbeitende Investorinnen stellt, die mit dem Geld anderer agieren. Diese müssen einerseits in Bezug auf eine fundamental offene und damit unsichere Zukunft Entscheidungen treffen und sich dafür andererseits retrospektiv rechtfertigen, wenn diese jeweiligen Zukünfte sich geschlossen haben und Gegenwart geworden sind. Dieses im Feld unter dem gängigen Begriff des »Monday Quarterback risk« subsumierte Problem führt Richard Zeckhauser (2013) als wichtigen Grund dafür an, warum institutionelle Investoren im Großteil vor Situationen zurück scheuen, in denen der Ausgang zwar vollkommen ungewiss, das Verhältnis von Risiko und potentiellem Gewinn aber eigentlich außerordentlich attraktiv ist. Zu groß sei die Gefahr, bei negativem Ausgang blöd dazustehen und von den im übertragenen Sinne am Montagmorgen das Spiel vom vorherigen Wochenende perfekt verstehenden Pseudo-Expertinnen zerrissen

zu werden.⁹ Folglich böten solche für alle Marktteilnehmerinnen essenziell unsicheren, aber durchaus attraktiven Gelegenheiten einen signifikanten Vorteil für Laien-Investoren, die nur ihr eigenes Geld investieren und folglich niemandem Rechenschaft schuldig sind.¹⁰ Anders ausgedrückt: Nur Laien können es sich hier leisten, vollständig auf die auf Plausibilität abzielende Bewertungsform zu verzichten.

Insbesondere die letzten beiden Fälle verdeutlichen, wie nah die beiden Bewertungsregime in vielen Fällen beieinander liegen, so dass ihr grundsätzlicher Unterschied auf den ersten Blick leicht übersehen werden kann. Umso entscheidender ist, dass hier gerade aus dieser großen Nähe eine für das Feld offenbar relevante Spannung entsteht: Die benachbarten Bewertungsregime geraten in Konflikt miteinander. Dieses grundsätzliche Potential zum Konflikt wurde hier anhand des Falls von Jonas herausgearbeitet, für den mediale Deckung der Bewertungen von Spielern – gerade im Hinblick auf derart sensible Bereiche wie Charaktereigenschaften – Vorrang hat vor der Reliabilität der dahinterliegenden Information. Dass sich der feine, aber im Feld potentiell bedeutsame Unterschied dieser beiden hier vorgestellten Bewertungsregimes am Fall von *Football Manager* gut herausarbeiten lässt, verwundert nicht: Im (Video-)Spiel über das (Sport-)Spiel wird das Verhältnis von Darstellung und Dargestelltem *selbst* explizit und zum Gegenstand sowohl für Spielerinnen als auch die hier beschriebenen Researcher.

Literatur

- Barrish, P. (2011). *The Cambridge introduction to American literary realism*. Cambridge University Press.
- Breidenstein, G., Hirschauer, S., Kalthoff, H., & Nieswand, B. (Hg.). (2013). *Ethnografie: Die Praxis der Feldforschung*. UVK Verlagsgesellschaft.
- Brock, N. (2021, April 28). The Researcher Interviews. <https://community.sigames.com/forums/topic/552247-the-researcher-interviews/>

⁹ Es erscheint mir hier – ausgehend von der eingangs kurz erwähnten und in der Einleitung zu diesem Band näher ausgeführten besonderen Darstellungsfähigkeit des Sports – keinesfalls *zufällig*, dass das Feld dazu eine Metapher aus dem Sport, hier aus dem in den USA so populären American Football, benutzt.

¹⁰ Ganz ähnlich, aber grundsätzlicher und noch weniger wissenschaftlich, sondern populär Laieninvestorinnen ansprechend argumentiert auch (Lynch & Rothchild, 2000).

- Bröckling, U. (2017). Gute Hirten führen sanft: Über Menschenregierungs-künste (Erste Auflage, Originalausgabe). Suhrkamp.
- Brümmer, K. (2019). Coordination in sports teams – ethnographic insights into practices of video analysis in football. European Journal for Sport and Society, 16(1), 27–43. <https://doi.org/10.1080/16138171.2019.1603682>
- Bruns, A. (2006). Towards Produsage: Futures for User-Led Content Production. Proceedings: Cultural Attitudes Towards Communication and Technology 2006.
- Crawford, G. (2006). The cult of Champ Man: The culture and pleasures of Championship Manager/Football Manager gamers. Information, Communication & Society, 9(4), 496–514. <https://doi.org/10.1080/13691180600858721>
- Crawford, G., & Gosling, V. K. (2009). More than a Game: Sports-Themed Video Games and Player Narratives. Sociology of Sport Journal, 26(1), 50–66. <https://doi.org/10.1123/ssj.26.1.50>
- Darr, A., & Mears, A. (2017). Locating local knowledge in global networks: Scouting in fashion and football. Poetics, 62, 1–14. <https://doi.org/10.1016/j.poetic.2017.03.001>
- Espeland, W. N., & Stevens, M. L. (1998). Commensuration as a Social Process. Annual Review of Sociology, 24(1), 313–343. <https://doi.org/10.1146/annurev.soc.24.1.313>
- Esposito, E. (2019). Die Fiktion der wahrscheinlichen Realität (N. Reinhardt, Übers.; 4. Auflage, Originalausgabe). Suhrkamp.
- Fleck, L. (2019). Entstehung und Entwicklung einer wissenschaftlichen Tatsache: Einführung in die Lehre vom Denkstil und Denkkollektiv (L. Schäfer & T. Schnelle, Hg.; 12. Auflage). Suhrkamp.
- Foucault, M. (2009). Geschichte der Gouvernementalität (2. Dr.). Suhrkamp.
- Frome, J., & Martin, P. (2020, Januar 1). Describing the Game Studies Canon: A Game Citation Analysis.
- Garfinkel, H. (1984). Studies in Ethnomethodology (1. Aufl.). Polity.
- Gibson, J. J. (2015). The ecological approach to visual perception: Classic edition. Psychology Press.
- Goffman, E. (2013). Fun in Games. In E. Goffman, Encounters; Two Studies in the Sociology of Interaction (S. 17–84). Martino Fine Books.
- Gugerli, D. (2018). Wie die Welt in den Computer kam: Zur Entstehung digitaler Wirklichkeit. S. FISCHER.
- Hocquet, A. (2016). Football Manager: Mutual Shaping between Game, Sport, and Community. Journal of Media Studies and Popular Culture = Revue

- d'Études Des Médias et de Culture Populaire, 6(Special Issue). <https://shs.hal.science/halshs-01312229>
- Hoppe, A. D. (2022). The Microsociology of Aesthetic Evaluation: Selecting Runway Fashion Models. *Qualitative Sociology*, 45(1), 63–87. <https://doi.org/10.1007/s11133-021-09496-x>
- Jacobson, M. (2021, Juli 22). How we're introducing women's football into Football Manager. [footballmanager.com](https://www.footballmanager.com/nnews/how-were-introducing-womens-football-football-manager). <https://www.footballmanager.com/nnews/how-were-introducing-womens-football-football-manager>
- Kalthoff, H., Hirschauer, S., Lindemann, G., Universität Bielefeld, & Universität München (Hg.). (2008). *Theoretische Empirie: Zur Relevanz qualitativer Forschung* (1. Aufl.). Suhrkamp.
- Kempf, D. (2017, Juli 27). Play the Game. Know the Game – Rekursive Wissenstransfers zwischen Football Manager und Fußball. *Paidia – Zeitschrift für Computerspielforschung*. <https://www.paidia.de/play-the-game-know-the-game-rekursive-wissenstransfers-zwischen-football-manager-und-fussball/>
- König, R. (2013). WIKIPEDIA: Between lay participation and elite knowledge representation. *Information, Communication & Society*, 16(2), 160–177. <https://doi.org/10.1080/1369118X.2012.734319>
- Leonard, D. (2006). An Untapped Field: Exploring the World of Virtual Sports Gaming. In *Handbook of Sports and Media* (S. 426–441). Routledge.
- Luhmann, N. (2017). *Die Realität der Massenmedien* (5. Auflage). Springer VS.
- Lynch, P., & Rothchild, J. (2000). *One up on Wall Street: How to use what you already know to make money in the market* (1st Fireside ed). Simon & Schuster.
- Parkin, S. (2015, Dezember 5). How Football Manager changed the game. *The Observer*. <https://www.theguardian.com/technology/2015/dec/05/how-football-manager-changed-the-game>
- Reng, R. (2015). *Mroskos Talente: Das erstaunliche Leben eines Bundesliga-Scouts* (2. Auflage). Piper.
- Schmidt, R. (2012). *Soziologie der Praktiken: Konzeptionelle Studien und empirische Analysen* (Erste Auflage). Suhrkamp.
- Schmidt, R. (2015). Neue Analyse- und Wissenspraktiken im Profifußball. *Sport und Gesellschaft: Zeitschrift für Sportsoziologie, Sportphilosophie, Sportökonomie, Sportgeschichte = Sport and society*, 12(2 Schwerpunktheft: Sport als kulturelle Praxis), 171–186.

- Stevens, P., Jr. (1980). Play and Work: A False Dichotomy. In H. B. Schwartzman (Hg.), *Play and culture: 1978 proceedings of the Association for the Anthropological Study of Play*. Leisure Press.
- Taylor, T. L. (Hg.). (2012). *Raising the stakes: E-sports and the professionalization of computer gaming*. Mit Press.
- tbm_cooper. (2022, Oktober 25). Bewertung Spieler Werder Bremen [@german_researcher]. <https://community.sigames.com/forums/topic/568742-bewertung-spieler-werder-bremen-german-researcher/>
- Waibel, D., Peetz, T., & Meier, F. (2021). Valuation Constellations. *Valuation Studies*, 8(1), 33–66. <https://doi.org/10.3384/VS.2001-5992.2021.8.1.33-66>
- Wikipedia. (2023). Wikipedia: Core content policies. https://en.wikipedia.org/wiki/Wikipedia:No_original_research
- Xiao, L. Y., Henderson, L. L., Nielsen, R. K. L., Grabarczyk, P., & Newall, P. W. S. (2022). Loot Boxes: Gambling-Like Mechanics in Video Games. In N. Lee (Hg.), *Encyclopedia of Computer Graphics and Games* (S. 1–7). Springer International Publishing. https://doi.org/10.1007/978-3-319-08234-9_459-1
- Zeckhauser, R. J. (2013). Investing in the Unknown and Unknowable. *Capitalism and Society*, 1(2). <https://papers.ssrn.com/abstract=2205821>

Die Vermessung von Geschlecht im Sport und der Umgang mit kategorialen Transgressionen

Dennis Krämer

1. Geschlechterdifferenzierung im Sport

Seit den frühen wissenssoziologischen Studien von Durkheim und Mauss (Durkheim & Mauss, 1903/1963) über das weite Feld der bis zu Simmel zurückreichenden Differenzierungsforschung (Simmel, 1890) bis hin zur neueren Formation der »Valuation Studies« richtet sich ein soziologisches Augenmerk darauf, wie Kategorien und Bewertungen zu einem »making up people« (Hacking, 1999) und »sorting things out« (Bowker & Star, 2000) beitragen. Diese Vorgänge werden besonders dann heikel, wenn soziale Zugehörigkeiten und Zugänge mit *natürlichen Unterschieden* begründet werden.

Ich möchte diesen Punkt im Folgenden an den Geschlechtstests näher erläutern, die in der modernen Leichtathletik seit dem frühen 20. Jahrhundert durchgeführt werden. Bei diesen handelt es sich um medizinisch-formalisierte Differenzierungsweisen, mit denen Sportverbände auf der Grundlage kontingenter Körperindizien und -bewertungen Menschen in zwei Geschlechterkategorien unterscheiden und die Unterscheidung mit naturalistischen Fairnessargumenten begründen (Krämer, 2023, 2020; Krämer & Shyvinck, 2024; Müller, 2006, 2017; Pape, 2017). Ihre Funktionsweise dramatisiert die anthropologische Grundannahme des »nature doesn't have categories; people do« (Stone, 1988, S. 307), indem sie verdeutlicht, wie das kulturgeschichtlich alte Klassifikationsmodell der Geschlechterunterscheidung durch ein professionalisiertes »category making« (Yanow, 2003, S. 9) szientifiziert wird und Vorstellungen über eine *natürliche Zweigeschlechtlichkeit* durch biomedizinische Körperbewertungen naturalisiert werden.

Zugleich ist die Art und Weise der Differenzierung charakteristisch für den Sport: Während Menschen in Schulen nach ihrer schulischen Leistung

segriert, im Recht kategorial getrennt, durch Religion purifiziert und im Beruf nach ihrer beruflichen Position hierarchisiert werden, ist im modernen Wettkampfsport ein Ordnungsmodell dominant, das Mitgliedschaften anhand von objektivierten biologischen Indikatoren wie Alter, Gewicht oder Geschlecht festlegt. Im Gegensatz zur Altersverifikation, die im Sport auf einer Prüfung der Ausweisdokumente beruht, oder dem Gewichtscheck, der auf einer punktuellen Messung des Körpergewichts basiert, erfordert die Geschlechtertrennung ein aufwendigeres Verfahren: Um Geschlechtstests durchzuführen, müssen Körper untersucht, das heißt, sie müssen entkleidet, Genitalien müssen entblößt, geprüft und vermessen, Blut-, Urin- und Gewebe-proben entnommen und im Labor analysiert, und es müssen medizinische Taxonomien gebildet und Grenzwerte festgelegt werden, um den Bewertungs- und Entscheidungsprozess zu formalisieren. Die Durchführung dieser Tests und das Warten auf die Testergebnisse kann ebenso belasten wie das Erhalten falscher Resultate. Daneben sind diese Tests für viele schambesetzt, können soziale Ächtungen nach sich ziehen und mit kulturellen, religiösen und persönlichen Überzeugungen in Konflikt geraten, so dass anzunehmen ist, dass bereits ihre Androhung bei vielen zu einem Rückzug aus dem Sport führen kann (Hoad, 2010).¹

Daneben existieren verschiedene Parallelen, wie andere Zugehörigkeiten im Sport indiziert werden: So wie die Differenzierung der Athlet:innen nach ihrer sportlichen Leistung über evaluative Praktiken der Quantifizierung erfolgt und mittels Distanz-, Gewichts- und Zeitangaben, Torschüssen, Medaillenspiegeln oder Ranglistenplätzen objektiviert wird, so wird auch die Geschlechtszuordnung, ähnlich einer Ranglistenlogik folgend, über numerische Skalierungen ausgedrückt. Die Funktion der Quantifizierung besteht darin, Mitgliedschaften zu einer reifizierten Geschlechterkategorie über objektive Indikatoren nachzuweisen und somit potenzielle »category advantages« (Parry & Martíková, 2021, S. 5) durch medizinische Vermessungen von Genitallängen und Testosteronwerten zu vermeiden.

Damit ist eine Besonderheit angesprochen, die den Sport von anderen sozialen Bereichen abgrenzt. Während es sich bei der Geschlechtertrennung um den kulturgeschichtlich vermutlich ältesten »Fall von Humandifferenzierung« (Hirschauer, 2014, S. 171) handelt, besteht die Spezifität im Wettkampfsport

¹ So zum Beispiel im Falle der erfolgreichen Schwestern Irina und Tamara Press, die ihre Kandidatur für die Leichtathletik Europameisterschaft 1966 in Budapest nach Androhung einer Genitaluntersuchung zurückzogen (Krämer, 2017).

in seiner meritokratischen Ausrichtung. Einerseits wird in den leistungsorientierten Bereichen des Sports angenommen (im Gegensatz zum Beispiel zum Gesundheits- oder Freizeitsport), dass nur die Leistung zählt und jede:r selbst für diese verantwortlich ist. In dieser Hinsicht funktioniert die Leistungsbeurteilung wie ein »großer sozialer Gleichmacher« (Hirschauer, 2014, S. 171), der dazu auffordert, von anderen Faktoren abzusehen. Andererseits wird unterstellt, dass es nur zwei Kategorien bedarf, um den evaluativen Rahmen abzubilden, in dem alle Teilnehmenden sowie ihre erbrachte Leistung gerecht beurteilt werden können. Neben den genannten Grenzüberschreitungen, die eine Kategorisierung von Geschlecht voraussetzt, besteht eine weitere Brisanz darin, dass Geschlechtstests nur zur Regulierung der Frauenkategorie durchgeführt werden. Es ist exakt dieser Punkt, warum Geschlechtstests seit einigen Jahren im Konflikt mit jenen inklusionspolitischen Positionen geraten, die die Wahrung der Persönlichkeitsrechte, den Wert der Selbstbestimmung bei der Geschlechtszuordnung sowie die Vielfalt der Geschlechter betonen (Sykes, 2006). Vor diesem Hintergrund besteht die aktuelle Herausforderung darin, zwei unversöhnliche Leitprinzipien in Einklang zu bringen, die den modernen Sport prägen: das Prinzip der Chancengleichheit und das Prinzip der Inklusion.

Zugleich handelt es sich bei Geschlechtstests um situierte Differenzierungsweisen, die nicht nur Kategorien reifizieren, sondern mit weitreichenden kollektiven und subjektiven Sinngebungen verwoben sind: Athlet:innen gehen davon aus, dass sie eine Chance gegen die Konkurrenz haben und allein ihre Leistung den Unterschied ausmacht, das Publikum erhofft sich einen spannenden Wettbewerb und Organisation und Sponsoring sind auf das positive Image des Sports angewiesen. Vor diesem Hintergrund fungiert die gesicherte Differenzierung der Athlet:innen als zentrale Weichenstellung, um die Glaubwürdigkeit an eine unparteiische Leistungsbewertung abzusichern, indem zwei Gruppen reifiziert und Menschen auf der Grundlage askriptiver Unterscheidungsmerkmale in diese eingeordnet werden.

Kultursoziologisch betrachtet knüpfen Geschlechtstests an ein Grundmuster der Moderne an: Autoren wie Horkheimer, Adorno, Baumann und Foucault haben gezeigt, dass sich die Moderne durch einen epistemischen Paradigmenwechsel auszeichnet, der sich dahingehend bemerkbar macht, wie Phänomene bewertet und kategorisiert werden und eine *instrumentelle Vernunft*, als professionalisierte Praxis der Vereinzelung, vorangetrieben wird (Baumann, 1995; Foucault, 1966). Es ist ein sich ausbreitendes Denken, das die Welt und die in ihr vorhandenen Weltphänomene als bestimbar und zu

bestimmend begreift und dabei in ein latentes Spannungsverhältnis zwischen »kontingenzenoffenen (Kultur) und kontingenzaversiven (Natur)« (Hirschauer, 2014, S. 173) Annahmen gerät. Der *moderne* Sport schließt an dieses Kulturparadigma einer *Ordnung der Dinge* an und überträgt das charakteristische Kategorisieren auf einen spezifischen Umgang mit Geschlecht: Im Kern basiert dieses auf einem binär gelagerten *Differenzierungzwang*, der keine geschlechtlichen Ambiguitäten, Kontinuitäten oder Nichtzugehörigkeiten toleriert und diese als illegitimate Transitionen labelt. Während »kulturelle Phänomene – anders als naturhaft gegebene Unterschiede – aus kontingenten sinnhaften Unterscheidungen bestehen« (Hirschauer, 2014, S. 170), scheint es sich bei Geschlechtstests um jene Art von Kulturphänomenen zu handeln, die ihre eigene Kontingenz verschleiern und durch diese Verschleierung an ihre Grenzen stoßen. Wie bei anderen kulturellen Kategorien liegt auch ihr Ziel, mit Baumann gesprochen (1995), in einem infiniten Vorgang der Ambiguitätsvermeidung, welche als kulturelle Praxis ein Charakteristikum der Moderne und damit auch des modernen Sports darstellt. Vor diesem Hintergrund sind Kategorisierungen als zivilisatorische Ordnungsleistung durch einen fundamentalen Antagonismus geprägt: Sie haben eine »polarisierende Vorderseite und eine homogenisierende Kehrseite« (Hirschauer, 2014, S. 174), mit der sie zwar egalisieren und differenzieren, aber auch Dritte ausschließen können.

Die Tragweite dieser Ambiguitätsvermeidung wird von Seiten der sozial- und sportwissenschaftlichen Forschung seit einigen Jahren intensiv am Umgang mit jenen Personengruppen diskutiert, die in besonderer Weise davon bedroht sind, vom Wettkampfsport ausgeschlossen und als »unvergleichbar« (Müller, 2017) eingestuft zu werden. In der binären Unterscheidungswelt des Sports betrifft dies jene Sportler:innen, deren Zugehörigkeit entweder qua Biologie eine geschlechterbinäre Segregation irritiert (Intergeschlechtlichkeit, Intersexualität, »Differences of Sex Development«), oder die einer alltäglichen Kontinuitätsannahme widersprechen, indem sie eine geschlechtsanpassende Transition (Transgeschlechtlichkeit, Transidentität, Transsexualität) vollziehen (Kessler, 1990, 1998; Hirschauer, 1993; Lindemann, 2011). Der binären Unterscheidungslogik des modernen Sports folgend werden Personen, die geschlechtliche Variationen oder Transgressionen anzeigen, in vielen Fällen als »abnormal, unnatural or deviant« (Linghede 2018, S. 571) und »in need of treatment or fixing« (Holmes, 2009, S. 3) eingestuft. Denn während sich geschlechtliche Identitätsformen nicht nur in westlichen Gesellschaften auspluralisieren (Voß, 2023) und in vielen Kulturen nie ein rein bipolares

Geschlechtsmodell existierte (Baumann, 1955), hält der Sport auch weiterhin an einer dualistischen Sortierung der Menschen fest.

Der Beitrag will die evaluative Infrastruktur genauer untersuchen, die der Sortierung von Menschen nach ihrem Geschlecht im Sport zugrunde liegt. Am Beispiel der professionellen Leichtathletik werde ich danach fragen, wie Geschlechter kategorisiert und hierzu Körper objektiviert werden, und wie sich die Formalisierung von Kategorien und die Praxis der Kategorisierung im Laufe der Zeit verändern. Damit wende ich mich explizit gegen eine voluntaristische Sichtweise, die Kategorien einseitig als »basis of eligibility« (Parry & Martíková, 2021, S. 1) thematisiert und lasse mich stattdessen auf eine »Kontinenzperspektive« (Reckwitz, 2008, S. 17) ein.

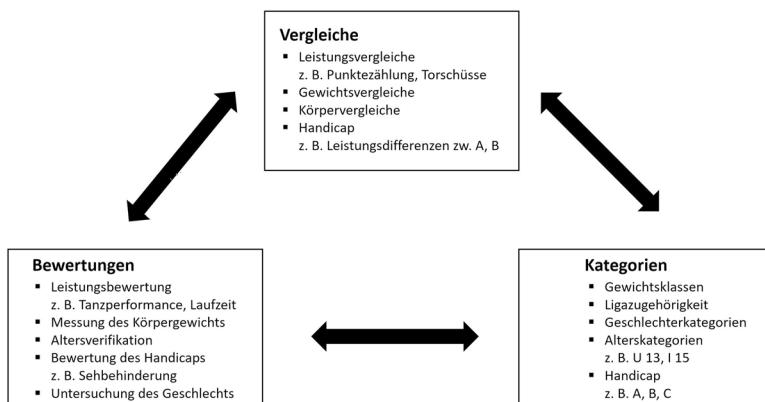
Hierzu gehe ich zunächst auf die Funktion von Kategorien und Leistungsklassen im Sport ein und beschreibe ihre Funktion als grundlegendes Kategorisierungsprinzip des Wettkampfsports (2.). In diesem Zusammenhang stelle ich die Relevanz der Geschlechtertrennung heraus (2.1) und verdeutliche, wie geschlechtsbasierte Leistungsklassen mit »category errors« (Yanow, 2003, S. 14) umgehen (2.2). Anschließend konkretisiere ich die Geschlechterdifferenzierung im Sport und unterscheide hierzu in die zwei Ebenen der Formalisierung (3.1) und medizinischen Objektivierung der Geschlechterdifferenzierung (3.2). Dabei gehe ich auch auf neuere Gender-Regularien der Sportverbände ein und hebe zwei für die Geschlechterunterscheidung zentrale Differenzierungsweisen hervor: Quantifizierungen und Laboratisierungen. Ein Schlussteil fasst den wesentlichen Erkenntnisprozess zusammen (4.).

2. Kategorien und Leistungsklassen

Um den Eindruck der Willkür zu vermeiden und das Vertrauen in die Repräsentativität der Leistungsbewertung sicherzustellen, greift der Wettkampfsport auf eine geregelte Bildung von Kategorien zurück, die als »Leistungsklassen« bezeichnet werden (Müller, 2006, 2017; Heckemeyer, 2018). Leistungsklassen überführen Sportler:innen mithilfe von Unterscheidungen in eine gemeinsame Struktur, in der diese als *un/gleich* und – hiermit zusammenhängend – für die anschließende Leistungsbewertung als *un/vergleichbar* erscheinen (Heintz, 2010, S. 164). Die trianguläre Beziehung zwischen Kategorien, Bewertungen und Vergleichen zeichnet sich dann dadurch aus, dass über spezifische Bewertungsweisen (zum Beispiel Messungen des Körpergewichts) die Grundlage für Kategorien geschaffen werden (zum Beispiel Gewichts-

klassen), innerhalb derer anschließend weitere Vergleiche durchgeführt werden können (zum Beispiel Leistungsvergleiche im Fliegengewicht). Dieser Vorgang kann aber auch impliziter sein und umgekehrt ablaufen, so dass die Trennung der Athlet:innen einen Vergleich der Bewertungsweisen (zum Beispiel verschiedener Messverfahren) oder Kategorien das Vergleichen der Unterscheidungskriterien (zum Beispiel Gewichtsschwellen) voraussetzen. Vor diesem Hintergrund beruht die Kategorienbildung im Sport auf einer Kombination von Verfahren der Objektivierung sowie der Bewertung der objektivierten Kriterien, die zur Kategorisierung in eine Beziehung gesetzt werden.

Abbildung 1: Leistungsklassen als trianguläres Ordnungsprinzip



Dabei können die Kategorien ineinander verschachtelt vorliegen und sich weiter aufsplitten: So basiert der Leistungsvergleich im Boxsport auf der Zuordnung der Athlet:innen zu einer Leistungsklasse (»Gewichtsklasse«), die sich aus dem Zusammenspiel von Gewichtskategorien (Schwergewicht, Fliegengewicht usw.), Geschlechterkategorien (m/w) und Alterskategorien (U13, U15 usw.) ergibt; im Fußball resultiert sie aus der Verbindung von Geschlecht (m/w), Alter (Junioren, Mannschaften) und Ligazugehörigkeit (1. Bundesliga, Regionalliga, Kreisliga usw.), und im Reitsport ergibt sie sich aus dem jeweiligen Rang, der in erfolgsbasierten Jahresturnierlizenzen festgehalten wird – ähnlich dem Ligasystem im Fußball, wo in jeder Saison eine Wahrscheinlich-

keit besteht, auf- oder abzusteigen (abgesehen von Spitzenpositionen wie der 1. Bundesliga, aus der man nur absteigen kann).

Gleichzeitig sind mit den Kategorien verschiedene Erwartungen verknüpft: Ohne die Trennung nach Gewicht würden Boxkämpfe zwischen Personen aus dem Schwergewicht und Fliegengewicht wahrscheinlich genauso vorhersehbar ausfallen wie Fußballspiele zwischen Kreis- und Bundesliga-mannschaften oder Tanzwettbewerbe zwischen Junior:innen und Senior:innen. Zudem ermöglicht die Trennung die gezielte Austragung spezifischer Veranstaltungen, wie Altherren-, Jugend- oder Mädchenturnieren. Auf diese Weise trägt die Kategorisierung »zu einer Pluralisierung der Wettbewerbe sowie zu einer höheren Attraktivität des Konkurrenzkampfs für das Publikum [bei], da Spannung und Ergebniskontingenz auf diese Weise aufrechterhalten werden« (Müller, 2006, S. 395).

Darüber hinaus können bestimmte Kategorien im Spannungsverhältnis zu zeitgenössischen Wertevorstellungen stehen. So wird in den meisten Sportarten einer Kategorienbildung basierend auf Alter und Geschlecht zugestimmt, jedoch nicht nach Hautfarbe, Nationalität oder Ethnizität (Steuwerwald, 2017). Allerdings kann die fehlende Beachtung bestimmter Faktoren zu einer impliziten Benachteiligung führen. So wird zum Beispiel in keiner Sportart nach dem sozioökonomischen Status differenziert. Dennoch ist weithin bekannt, dass diese einen erheblichen Einfluss darauf hat, wer überhaupt welche Sportart betreibt, wem eine Sportdisziplin eher liegt und wer sich bestimmte Hobbys leisten kann (Bourdieu, 1991; Gebauer, 1972). Wenig überraschend ist daher, dass Sportarten wie Polo, Yachting oder Golf eher von ökonomisch privilegierten Bevölkerungsgruppen praktiziert werden (Gilbert & Gillett, 2013). Ebenso ist bekannt, dass diese vergleichsweise viel Sport treiben (Burrmann, 2021; Rohrer & Haller, 2015). Die stillschweigende Akzeptanz dieser Voraussetzungen legt nahe, dass Leistungsvergleiche im Sport mit einem neoliberalen Wertesystem verbunden sind und die sozioökonomische Leistungsfähigkeit von Menschen als impliziter Selektionsvorgang toleriert wird.

Ferner wird die Kategorienbildung von größeren historischen Schwankungen beeinflusst. So rückte erst die Öffnung der Olympischen Spiele für Frauen im Jahr 1900 die Notwendigkeit einer geregelten Geschlechtertrennung auf die Agenda der Sportverbände, die Einführung der UEFA Champions League 1955 (ehemals »Europapokal der Landesmeister«) initiierte den Leistungsvergleich unter den besten Teams im europäischen Männerfußball und in der Geschichte des Behindertensports führte der Weggang von einer Unterscheidung

nach »Behinderungsarten in verschiedene Schadensklassen« (Bennani, 2022, S. 254) zu einer Kategorisierung, die sich auf die funktionellen Auswirkungen im Kontext von »environmental factors« konzentrierte.

2.1 Geschlechterkategorien und Leistungsklassen

Vor diesem Hintergrund zeichnet sich der organisierte Wettkampfsport durch einen hohen Formalisierungsgrad aus, der neben geschlechtsspezifischen Regelwerken die Differenzierung der Athlet:innen nach ihrer Geschlechtszugehörigkeit umfasst. In fast allen Sportarten hat sich somit eine Kategorienbildung etabliert, die die Geschlechtszugehörigkeit als grundlegendes Differenzkriterium betrachtet und die Geschlechtertrennung mit naturalistischen Leistungsannahmen begründet. Müller spricht in diesem Zusammenhang auch von »präsumptiven Leistungsklassen« (Müller, 2006, S. 305), um zu unterstreichen, dass die Zuordnung mit kulturellen Vorstellungen über angeborene Leistungsfähigkeiten begründet wird: Im direkten Vergleich der (zwei) Geschlechter gelten Männer als größer, kräftiger und schneller als Frauen, so dass gemäß binaristischem Kausalschluss Männer von der Frauenkategorie ferngehalten werden sollen. So werden aus diesem Grund seit jeher auch keine Geschlechts-tests zur Kontrolle des Zugangs zur Männerkategorie durchgeführt, da davon ausgegangen wird, dass im *Leistungsvergleich der Leistungsstärksten* keine kompetitiven Verzerrungen entstehen können und diese schlimmstenfalls durch Doping verursacht werden könnten.

Geschlecht wird dabei wie Alter als ein horizontales Differenzmerkmal behandelt, welches, im Gegensatz zu vertikalen Merkmalen wie dem Gewicht im Boxsport oder dem sich in Gürtaufnahmen ausdrückenden Rang im Judo oder Karate, die Kategorienbildung transdisziplinär bedingt. Ausnahmen finden sich dagegen in weniger kompetitiven Sportsphären sowie in Veranstaltungen, die explizit als Kategorienüberschreitungen praktiziert werden, zum Beispiel gemischtgeschlechtliche Teams im Tennis und Hockey (Martinkova, 2020). Hier reihen sich auch jene Disziplinen ein, die auf neuen Formen der Kategorienbildung beruhen und einen anderen Umgang mit (analogen) Kriterien vorsehen, wie etwa Ladder-Systeme im digitalen Sport (Quandt & Kröger, 2013). Sonderfälle bilden zudem komplementär praktizierte Sportarten wie Tanzen oder Eiskunstlauf, in denen die Geschlechter gemeinsam antreten und hier mitunter traditionelle Rollenbilder aufgreifen – so etwa als »Untermann«, der Bewegungsimpulse vorgibt und die Frau führt, hebt und wirft. Exzeptionelle Regelungen finden sich auch in Sportarten, in denen es um den *instrumentellen Um-*

gang mit größeren Geräten oder Tieren geht (Müller, 2006, S. 409). Vor diesem Hintergrund existiert keine Geschlechtertrennung in Sportarten wie Springreiten, Pferderennen, Motor- und Segelsport. Stattdessen werden Trennungen über Vorgaben für Sportgeräte (Segelboote, Motorräder, Autos) sowie Dopingregularien für Tiere (zum Beispiel für Pferde) realisiert.

Ferner nimmt die professionelle Leichtathletik eine Sonderstellung ein. Bei internationalen Sportturnieren wie den Olympischen Spielen wird zwar keine Trennung nach dem Alter vorgenommen. Jedoch führte das Internationale Olympische Komitee Ende der 1990er Jahre ein Mindestalter von 16 Jahren ein, um Kinder und Jugendliche vor möglichen Entwicklungsschäden infolge des exzessiven Trainings zu schützen. Zusätzlich sind die Längen von Sprintstrecken nach Alter gestaffelt (z.B. 50m, 75m, 100m), und der Zugang zu als besonders anspruchsvoll eingestuften Disziplinen wie dem Dreisprung, Stabhochsprung oder Hammerwurf wird über ein zusätzliches Mindestalter reguliert.

Daneben wird die Geschlechtertrennung über sportartspezifische Richtlinien formalisiert. Ein Beispiel hierfür findet sich in den unterschiedlichen Regelwerken für Männer und Frauen. So sind im Frauensport kleinere und leichtere Wurfgeräte (der Frauen-Diskus ist 1 kg leichter), kürzere Streckenlängen (100m statt 110m Hürden) sowie teils andere Sportarten vorgesehen (z.B. Soft statt Baseball). Auch werden Frauenwettbewerbe in einigen Sportarten in weniger aggressiven Settings ausgetragen, zum Beispiel durch Einrichtung eines Bodycheck-Verbots im Eishockey (Heckemeyer, 2018). Neben diesen graduellen Unterschieden existieren grundsätzliche Exklusionen: So treten nach wie vor ausschließlich Frauen in ästhetischen Sportarten wie Rhythmische Sportgymnastik und Synchronschwimmen an, und bei den Olympischen Spielen ist ihnen der Schwebebalken und Stufenbarren vorbehalten, während Männer an Pauschenpferd, Ringen und Reck konkurrieren.

2.2 Geschlechterkategorien und geschlechtliche Ambiguität

Kulturelle Ordnungssysteme beruhen auf dem Prinzip, dass Unterschiede den Unterschiedenen nicht immanent sind, sondern der kollektiven Verständigung über und der praktischen Herausstellung von Unterschiedlichkeiten bedürfen. Als historisch und topographisch kontextualisierte Phänomene sind sie somit contingent und in ihrer sozialen Wirksamkeit zeitlich und kulturell limitiert.

Mit Blick auf den Sport, wird die Geschlechterdifferenzierung zwar mit *natürlichen Unterschieden* begründet. Soziologisch gesehen lässt sich der proklamierte Unterschied jedoch nicht ohne weiteres als naturhaft hinnehmen, sondern setzt eine soziale Unterscheidungspraxis, eine Bourdieusche »Klassifikation der Klassifizierer« (Hirschauer & Boll, 2017, S. 7), voraus. Um Geschlechterkategorien zu bilden, müssen Unterschiede ebenso erkennbar gemacht wie erkannt werden: Sie bedürfen einer evaluativen Infrastruktur inklusive der Etablierung von »biomedizinischen Grenzregimen« (Lindemann, 2002, 2009), die Mitgliedschaften begründet und diese in einem sozialen System verstetigt.

Ein charakteristisches Merkmal bei der Kategorienbildung im modernen Sport zeigt sich in der Reduktion von Komplexität, die zur Konstitution einer klaren Differenzlinie zwischen den Geschlechtern vorgenommen wird. Im Gegensatz zur sozialen Komplexität, in der die Geschlechtszugehörigkeit mit zahlreichen Dimensionen wie der Geschlechtsidentität, Alltagsdarstellung, personenstandsrechtlichen Zuordnung, Kleidungs- und Berufswahl sowie Sexuierung und Partnerschaft verworben ist, wird die Kategorisierung im Sport von persönlichen und sozialen Einflüssen befreit und als ausschließlich binäre Ausprägung verhandelt. Die Grenzen dieser Kategorisierung lassen sich an den Maßnahmen verdeutlichen, die ergriffen werden, um mit geschlechtlichen Grenzüberschreitungen umzugehen. Sie werden im binär-segregierten Sport wie »category errors« (Yanow 2003, S. 14) behandelt.²

Zu den in den vergangenen Jahrzehnten bei der Geschlechterdifferenzierung am häufigsten problematisierten Personengruppen zählen trans- und intersexuelle Athlet:innen. Als Personengruppe »that cannot be categorised and placed in neat binaries« (Linghede, 2018, S. 580), verweisen sie auf die Irritationen sowie regulativen Normalisierungsmaßnahmen, die im gegenwärtigen

2 Yanows Konzept der »category errors« beschreibt Fehler, die bei der Zuordnung zu Kategorien entstehen können und eines besonderen Kategorienmanagements bedürfen. Die Entstehung dieser Fehler führt sie nicht auf die kategorisierten Phänomene, sondern auf die Praxis der Kategorisierung zurück. Demnach treten Fehler erst dann auf, wenn komplexe Sozialphänomene zu starr oder reduktionistisch behandelt werden. Um diesem entgegenzuwirken, schlägt sie eine reflexive Kategorienbildung vor, die die Kategorisierung durch einen fortlaufenden Reflexionsprozess aktualisiert. Jenkins hat diesen Prozess auch als »internal-external dialectic« (Jenkins, 1996) der Kategorienbildung beschrieben und gezeigt, dass Kategorisierungen auf dem reflexiven Erkennen dessen beruhen, dass »similarity and difference are implicit in the other« (Jenkins, 1996, S. 3).

Sport mit einer bipolaren Kategorisierung zusammenhängen. Sie stellen in der Geschichte der modernen Leichtathletik einen zentralen Grund dafür dar, wie die Erschaffung von »boxes in which all individuals fit once and only once« (Yanow 2003, 12) ein kontinuierliches »boundary making« (Lamont & Molnar, 2002; Hirschauer, 2014) erforderlich macht. Zugleich fungieren sie auf organisatorischer Ebene als eine zentrale Referenz, um ein bipolares System über seine Außenbezirke zu festigen. Ich möchte dies in den folgenden Abschnitten am Umgang mit intersexuellen Körpern näher verdeutlichen.

Der vom Biologen Richard Goldschmidt geprägte Begriff »Intersexualität« steht seinerseits für eine westliche Tradition der medizinischen Kategorisierung von nichtbinären Geschlechtskörpern, die von binären Normkörpern abgegrenzt und hierzu in über 80 »Intersex-Conditions« unterteilt werden (Goldschmidt, 1917; Cools et al., 2018; Klöppel, 2010). Während sich die Ambiguitätsvermeidung in der medizinischen Praxis in einem normativen Umgang mit Zweigeschlechtlichkeit und der Etablierung pathologischer Syndrome inklusive entsprechender Normalisierungsmaßnahmen manifestiert, besteht die Ordnungsleistung im Sport darin, Geschlecht messbar zu machen und geschlechtliche Entwicklungsvielfalt durch entsprechende Vermessungen auf ein bipolares System abzustimmen.

Wie sehr es sich hierbei keine naturgemäße Zuordnung, sondern eine Kulturleistung handelt, zeigt sich an verschiedenen Aspekten: Die meisten Menschen haben im Alltag keine Kenntnis darüber, wer intersexuell ist; häufig nicht einmal Intersexuelle selbst. Das Bewusstsein für die eigene Geschlechtszugehörigkeit muss in vielen Fällen aufwendig durch eine medizinische Expertise festgestellt werden (Lang, 2006). In dieser Hinsicht ist die Selbstzuordnung sowie auch interaktive Geschlechtsdarstellung im Alltag durch eine sinnlich-soziale Zuordnung strukturiert, die, wie bereits Garfinkel an Transsexuellen gezeigt hat, eine binäre Zugehörigkeit insinuiert und *nicht-binäre Andersheiten* aufgrund fehlender Alltagsbezüge tendenziell übersieht (Garfinkel, 1967; auch Hirschauer, 1993). Gleichzeitig ermöglicht dieses interaktive *Geschlechterspiel*, dass man im Alltag als das jeweils andere Geschlecht in Erscheinung treten und entsprechende Verhaltenscodes erwerben kann.

Medizinisch betrachtet sieht das anders aus: So kann eine als intersexuell kategorisierte Person, die im Alltag als Frau gelesen wird, einen *männlichen Karyotyp* (46,XY), einen *zu hohen Testosteronwert* und ein *äußeres männliches Genitale* haben. Intersexuelle können, je nach medizinischer Nomenklatur, chromosomale Varianten, Abweichungen oder Anomalien wie 47-XXY (med. zum Beispiel Klinefelter-Syndrom) oder 45,X/46,XY-Mosaike (zum Beispiel

Turner-Syndrom) aufweisen, ebenso wie weitere Kombinationen und Transgressionen, zum Beispiel XY-Chromosomen bei einem weiblichen Phänotyp (med. zum Beispiel Partielle/Komplette Androgeninsensitivität).³

Es ist anzunehmen, dass daher streng genommen auch viele intersexuelle Athlet:innen im Sport unsichtbar bleiben müssten. Dass das nicht der Fall ist, liegt daran, dass der Sport eng mit der medizinischen Forschung kooperiert, um gemeinsam Kriterien und Kategorien zu ihrer Sichtbarmachung zu etablieren.

3. Geschlechterdifferenzierung und Geschlechtstests

Vor dem dargestellten Hintergrund werden Kategorienüberschreitungen als »illegal transgressions« (Günter, 2015) eingestuft und die Kategorienbildung entsprechend eines kulturellen Unterscheidungsmodells des *tertium non datur* (Wacke, 1989; Dreger 2000) vollzogen.

3.1 Formalisierung der Geschlechterdifferenzierung

Bis hierhin können wir festhalten, dass die Differenzierung der Geschlechter dem Prinzip der Leistungsklassenbildung folgt und mit nationalistischen Leistungsannahmen begründet wird. Wie nun zu zeigen sein wird, besteht die Naturalisierung nicht nur in einer vagen Vermutung über Zugehörigkeiten und Potenziale, sondern wird, kennzeichnend für eine moderne Institution, über entsprechende schriftliche Formalisierungen abgesichert. Sie verstetigen als organisationale »Härtungen« (Kathoff & Dittrich, 2016) Entscheidungsstrukturen im Sport und machen Athlet:innen »berechenbar und zurechenbar« (Wettmann, 2020, S. 122). Damit sind sie ein wichtiger »Transmissionsriemen«, um transsituative Entscheidungen kommunizierbar zu machen (Wagner et al., 2023).

Kulturtheoretisch lassen sich Naturalisierungen als Verfahren betrachten, durch die Phänomene, Dinge oder Eigenschaften als Bestandteile einer *natürlichen Umwelt* intelligibel gemacht und in einem Status der *quasi vorsozialen*

³ Zugegeben tut sich meine Beschreibung etwas schwer, da mir das semantische Rüstzeug fehlt, um bei der Beschreibung des sozialen Phänomens aus der Sogwirkung medizinischer Suggestionen herauszutreten. Ich belasse es an dieser Stelle dennoch bei den medizinischen Terminen, um der Leserschaft den Grundgedanken zu vermitteln.

Dauerhaftigkeit verankert werden (Douglas, 1992). Als kontinuitätsgetriebene Stabilisierungsintentionen negieren sie die Kontingenz sozialer Aushandlungen, indem sie so tun, als seien sie Teil einer immer schon gegebenen natürlichen Normalität. Naturalisierungen können sich insofern mit Formalisierungen verschränken, als die Negation einer kulturellen Realität über ihre sprachliche, symbolische oder numerische Kodifizierung verstärkt und zugleich kommunizierbar sowie auf andere Phänomene übertragbar wird. Im Umgang mit geschlechtlicher Ambiguität verstetigen Formalisierungen ein *professionnelles Accomplishment*, indem sie ein standardisiertes Behandlungsprozedere etablieren und klare Richtlinien für die Behandlung von *ähnlichen Fällen* festlegen.

In diesem Zusammenhang hat auch Knorr-Cetina die Bedeutung wissenschaftlicher Prozesse für die Stabilisierung sozialer Kategorien untersucht (sie bezieht sich dabei auf frühe Studien von Mauss und Durkheim, die von »classifications« sprachen). Sie argumentiert, dass moderne Gesellschaften auf eine wissenschaftlich fundierte Formalisierung von Kategorien angewiesen sind, um den Eindruck von »socially contrived arrangements« zu vermeiden (Knorr-Cetina, 1994, S. 9). Erst dieser Prozess verwandle kontingente Kulturphänomene in dauerhafte Dispositionen. Das Paradoxe bestehe demnach darin, dass sich die Forschung – bei ihr vor allem die Technik- und Naturwissenschaften – selbst symbolischer Kategorien bedienten, um den Anschein von Stabilität zu erzeugen; »Science, a paradigmatic modern institution, also enlivens its universe through symbolic classifications« (Knorr-Cetina, 1994, S. 10).

Analog dazu wird auch die Geschlechtertrennung im Sport verwissenschaftlicht und formalisiert. Im Gegensatz zu archaischen Sportformen zielt die Formalisierung darauf ab, die für den modernen Sport charakteristische Bestrebung nach Schaffung einer »formalen Ausgangsgleichheit« herzustellen (Bette, 2010, S. 95). In der Leichtathletik findet sich das entsprechende Wissen in den Gender Regularien (»Gender Regulations«), die von den Medizinkommissionen der Weltsportverbände (World Athletics, Internationales Olympisches Komitee) herausgegeben werden.

Diskontinuitäten in der Geschlechterdifferenzierung

In der Leichtathletik ist die Formalisierung der Geschlechterdifferenzierung von zwei historischen Diskontinuitäten geprägt, die sich in der Kategorisierung der Athlet:innen sowie im Umgang mit ihren Körpern zeigen. Anfänglich resultiert die erkannte Notwendigkeit der Formalisierung aus der Erweiterung des Teilnehmendenkreises um ein weiteres Geschlecht: Analog zur Öffnung

der Olympischen Spiele für Frauen im Jahr 1900 kam erstmals der Verdacht auf, dass es sich bei einigen Teilnehmerinnen um verkleidete Männer handelt (Wiederkehr, 2008; Cole, 2000). Jedoch beschränkte sich der organisierte Sport in der ersten Hälfte des 20. Jahrhunderts noch darauf, solche mutmaßlichen Transitionen nicht selbst aufzuklären, sondern an nationale Einrichtungen wie Polizeidienststellen zu übergeben (Krämer, 2020; Bahro, 2009). Dies änderte sich zur Mitte des 20. Jahrhunderts. Wenngleich die Sportverbände schon zuvor entsprechende Vorkehrungen für eine obligatorische Testung getroffen hatten, wurden diese bedingt durch den Zweiten Weltkrieg unterbrochen (Cahn, 1994). Ab 1946 waren alle Personen verpflichtet, die beabsichtigten, an internationalen Frauen-Leichtathletikturnieren teilzunehmen, ein ärztliches Attest vorzulegen, das auf einer gynäkologischen Untersuchung beruhte. Als die Sportverbände die Integrität des Sports während der Ost-West-Konflikte zusätzlich durch das Problem des Staatsdopings bedroht sahen, richteten sie 1966 Medizinkommissionen ein. Neben der Durchführung von Dopingtests bestand ihre Aufgabe fortan darin, die Geschlechtszugehörigkeit der Teilnehmenden vor Turnierbeginn zu verifizieren und hierzu ihre Geschlechtsteile zu überprüfen, zunächst visuell.

Die öffentlich auch als »nude parade« (Blithe & Hankey, 2015) kritisierte Praxis wurde bereits 1967 ausgesetzt und durch das diskretere Testverfahren der Chromosomentests ersetzt. Statt Geschlechtsorgane zu prüfen, wurden fortan Schleimhautabstriche und Blutentnahmen durchgeführt. Personen, die den Test bestanden, erhielten einen entsprechenden Weiblichkeitsnachweis, der ihnen »no abnormality« bescheinigte.

Infolge der öffentlichen Kritik an den Tests, die unter anderem intersextuelle Athlet:innen wie die spanischen Hürdenläufer:in María Martínez-Patiño äußerten, wurde die obligatorische Testung in den 1990er Jahren ganz ausgesetzt. Die Kritik lautete in diesem Fall, dass gemäß des Stands der biomedizinischen Forschung ein XY-Chromosom weder per se ein männliches Geschlecht noch eine biologische Überlegenheit nachweise (Martínez-Patiño, 2005; Fausto-Sterling, 2012; Krämer, 2020). Damit bestand einer der Hauptgründe für die Beendigung der Chromosomentests darin, »dass dadurch wieder das Problem uneindeutiger Kategorisierungen gelöst noch Chancengleichheit erreicht wurde« (Müller, 2017, S. 217).

*Abbildung 2: Weiblichkeitssnachweis von den Olympischen Spielen
1968*

A L'OCASION DES JEUX DE LA XIXE OLYMPIADE ON OCCASION OF THE XIX OLYMPIC GAMES ON:	II. A ETE PROCEDE SUR:
L'ATHLETE: <u>THE ATHLETE</u>	
SPORT: <u>SPORT</u>	CARTE D'IDENTITE: <u>IDENTITY CARD</u>
NATIONALITE: <u>NATIONALITY</u>	
A UN EXAMEN DE CHROMATINE SEXUELLE QUI N'A DONNE AUCUNE ANOMALIE. THE SEXUAL CHROMATINE INVESTIGATION WAS MADE. THE RESULT PROVED NO ABNORMALITY.	
COMMISSION MEDICALE DU C.I.O. MEDICAL COMMISSION OF THE I.O.C.	
LE PRESIDENT DE LA COMMISSION MEDICALE THE CHAIRMAN OF THE MEDICAL COMMISSION	MEMBRE DE LA COMMISSION MEDICALE MEMBER OF THE MEDICAL COMMISSION
MEXICO LE _____ 1968	

Bundesarchiv 510/862, zit.n. Wiederkehr, 2008, S. 269

Ein erneuter Strategiewchsel vollzog sich nach dem Goldmedaillensieg von Caster Semenya bei der Berliner Leichtathletik-WM 2009, so dass die Sportverbände seit 2011 wieder Geschlechtstests durchführen. Statt Geschlechtsorgane und Chromosomen werden fortan Hormontests zur Bestimmung des körpereigenen Testosteronwerts durchgeführt. Diese unterscheiden sich von den vorherigen Testformen in mehreren Hinsichten: Einerseits wird seitdem nicht mehr obligatorisch, sondern selektiv getestet, d.h. erst bei Verdacht auf das Vorliegen einer geschlechtlichen Variation oder Transition. Die Operationalisierung solcher Verdächtigungen setzt die Formalisierung spezifischer Kriterien des Erkennens von Ambiguität voraus, die über entsprechende Hinweise auf eine *Virilisierung* (wie Gesichts- oder Körperbehaarung) in medizinischen Leitfäden festgelegt werden (Müller, 2017, S. 219). Andererseits stehen seitdem nicht mehr geschlechtliche Grenzüberschreitungen im Fokus, zumindest nicht mehr formell, sondern Teilnehmende in der Frauenkategorien, deren Körper einen erhöht eingestuften Androgenwert produzieren, was medizinisch als »Hyperandrogenismus« eingestuft wird. Die Starterlaubnis wird an die Höhe des jeweiligen Androgenspiegels

geknüpft. Der ermittelte Wert fungiert seitdem als zentrale Referenz, um das Prinzip der Chancengleichheit über eine singuläre biochemische Komponente abzusichern, die als hauptverantwortlich für die sportliche Leistung angesehen wird.

Die Kategorie des »Relevant Athlete«

Angesichts aktueller gesellschaftlicher Inklusionsforderungen, queerer Emanzipationsbewegungen und Diskriminierungsverbote sind die Sportverbände wie nie zuvor dazu angehalten, ihre Testverfahren, Kriterien und Motive fortlaufend zu überprüfen (Krämer, 2022). Dies macht sich einerseits in einer beschleunigten Kommunikation über Differenzierungs-Standards bemerkbar, so dass die Gender Regularien seit 2011 in regelmäßigen Abständen aktualisiert werden. Es zeigt sich aber auch darin, dass diese ein breites Spektrum an gesellschaftlichen Faktoren berücksichtigen müssen und daher nicht mehr bestimmte Geschlechter prinzipiell ausschließen können.

Vor diesem Hintergrund besteht ein politisches Novum darin, dass, korrespondierend mit der zunehmenden gesellschaftlichen Sensitivität für Diversitäts- und Inklusionsfragen, Inter- und Transsexuelle nicht mehr grundsätzlich ausgeschlossen werden. Wie oben bereits angedeutet, wird ihre Teilnahme indirekt über neue Kategorien und Bewertungsweisen reguliert, die vor einigen Jahren in den binär-segregierten Sport implementiert wurden. Vor diesem Hintergrund zeichnet sich der aktuelle Umgang dadurch aus, dass weiterhin an einer zweigliedrigen Differenzierung festgehalten wird, während die Teilnahme über eine komplexe Ordnungsleistung koordiniert wird. Diese beruht im Kern auf zwei verknüpften Differenzierungslogiken:

Schaffung zusätzlicher intra-kategorialer Bereiche, mit denen il/legitime Kriterien innerhalb einer Kategorie durch Anschluss an externe Kategorien begründet werden, zum Beispiel Implementierung medizinischer Syndrome zur Ermittlung der Teilnahme(un)fähigkeit in der Frauenkategorie; Durchführung medizinischer Eingriffe, mit denen die Zugehörigkeit zu einer Geschlechterkategorie nicht mehr als unumkehrbare Feststellung (»Weiblichkeitssertifikat«, siehe oben), sondern als *erwerbbar* angesehen wird, zum Beispiel durch pharmakologische Eingriffe zur Senkung des Testosteronspiegels.

Das formalisierte Zusammenspiel lässt sich an der kürzlich von der World Athletics eingeführten Kategorie des »relevant athlete« konkretisieren. Als »rele-

vant athletes« werden jene Athlet:innen kategorisiert, bei denen sich drei Kriterien überlappen (World Athletics, 2023, S. 5f.):

- (1) »they have one of the following DSDs [med. ›Differences of Sex Development, D.K.]: 5α-reductase type 2 deficiency, partial androgen insensitivity syndrome (aka PAIS); 17β-hydroxysteroid dehydrogenase type 3 (17β-HSD3) deficiency; ovotesticular DSD; or any other genetic disorder involving disordered gonadal steroidogenesis; and
- (2) as a result, they have a concentration of testosterone of 2.5 nmol/L or more in their serum; and
- (3) they have sufficient androgen sensitivity for that testosterone to have a material androgenising effect.«

Um eine binäre Kategorisierung zu realisieren werden aktuell jene Personen als behandlungsbedürftig eingestuft, die nach medizinischen Maßstäben eine bestimmte *DSD-Diagnose* in Kombination mit einem als *erhöht eingestuften Testosteronwert* aufweisen und bei denen sich eine *körperliche Sensitivität* für männliche Hormone nachweisen lässt. Analog zur Regulierung der Transsexualität wird im Falle der Intersexualität bzw. bei DSD-Syndromen die körperliche Sensitivität für Testosteron als eigentliche Ursache der Leistungsverzerrung eingestuft. Die oben angeführten Punkte (1) und (3) hängen insofern zusammen, als dass die unter Punkt (1) genannten DSD-Diagnosen medizinisch betrachtet dafür bekannt sind, eine Virilisierung des Körpers durch Androgene zuzulassen (im Gegensatz zum Beispiel zu Syndromen wie dem »complete androgen insensitivity syndrome«; CAIS).

Zugleich wird das Management der Kategorien nicht mehr einseitig als grundsätzliche Frage der Zugehörigkeit behandelt. Stattdessen werden in den aktuellen Regularien verschiedene Voraussetzungen angeführt, wie auch weiterhin in der Frauenkategorie angetreten werden kann. Diese gestalten sich analog zum Umgang mit Transsexuellen und werden im Falle des Umgangs mit Intersexuellen als »DSD Eligibility Conditions« bezeichnet:

- (1) »they must be recognised at law (for example, in a birth certificate or passport) either as female or as intersex;
- (2) they must have continuously maintained the concentration of testosterone in their serum below 2.5 nmol/L for a period of at least 24 months; and
- (3) they must continue to maintain the concentration of testosterone in their serum below 2.5 nmol/L at all times (i.e., whether they are in competition or

out of competition) for so long as they wish to retain eligibility to compete in the female classification [...].«

Theoretisch betrachtet insinuiert eine solche Formalisierung von Zugehörigkeit den Eindruck einer »sachbezogenen Leistungsbeurteilung« (Heintz, 2023, S. 176). Analog zur Kategorienbildung im Behindertensport sind es nicht Personen, die bewertet und kategorisiert werden, sondern Menschen, die als »Besitzer*innen ihres Körpers adressiert« werden (Bennani, 2022, S. 260). Die Brisanz dieser Kategorisierung resultiert nicht zuletzt aus dem Umstand, dass an »Menschenkörpern Personen haften und die von den Bewertungsverfahren unterstellte Trennbarkeit von Körper und Person kaum gegeben ist« (Schmidt et al. 2024, in diesem Band).

Vor diesem Hintergrund verdeutlichen die drei oben genannten Punkte aus den DSD Eligibility Conditions, dass es sich bei der Kategorisierung, wie ich im Folgenden ausführen werde, nicht um eine externalisierte Form der Geschlechterdifferenzierung handelt: Es geht nicht mehr darum, was medizinische Expertisen gewissermaßen an körperlichen Befunden vorfinden und wie sie ihre Entscheidungsfindung daran ausrichten. Eine Besonderheit der heutigen Kategorisierung besteht darin, dass sie als ein *submissiver Akt* behandelt wird, indem die disjunktive Kategorisierung zwar von außen gerahmt, aber zugleich als eigenverantwortlicher Vorgang der Selbstzuordnung behandelt wird. Dies wird dadurch realisiert, dass die Regularien nicht nur Mitgliedschaften festlegen, sondern auch Bedingungen benennen, wie man Mitglied wird (durch einen legitimen Testosteronwert), Mitglied bleibt (durch kontinuierliche Senkung des körpereigenen Testosteronwerts) oder die Mitgliedschaft wieder verliert (durch Unterbrechung der testosteronenenkenden Therapie).

3.2 Medizinische Objektivierung der Geschlechterdifferenzierung

Als »vielgestaltige, historisch und sozial wandelbare Praktiken« (Epple & Erhart, 2015, S. 17) müssen die Differenzlinien, die Kategorien zugrunde liegen, nicht immer direkt ersichtlich sein. Obwohl die Geschlechtertrennung im Sport mit biologischer Leistungsdifferenz begründet wird, bleibt für Außenstehende unklar, wie Geschlechtszugehörigkeiten verifiziert werden und welche Kriterien einen Unterschied machen. Das Wissen, dass in den meisten Sportdisziplinen nach Geschlecht getrennt wird und dass es bestimmte Gründe für die Trennung gibt, beruht daher weniger auf »Verständnis« als auf »Einverständnis« (Heintz, 2010, S. 172).

Kontingente Körper

Charakteristisch für ein medizinisch fundiertes Kategoriensystem beruht die Geschlechtertrennung im Sport auf einer polaristischen Heuristik, die diskursiven Entwicklungen unterliegt und sich an kontingenaten Bewertungs- und Behandlungsweisen orientiert. Die Kontingenz ist der beste Beleg dafür, dass die Herausstellung von *natürlichen Unterschieden* auf sinnhaften Unterscheidungen beruht, die ihrerseits »gezogen oder zurückgezogen, aufrechterhalten oder unterlaufen, und bei der Begegnung mit anderen Unterscheidungen verstärkt oder verdrängt werden können« (Hirschauer, 2014, S. 181). Der wissenschaftliche Erkenntnisgewinn lässt sich dann im Spannungsverhältnis erkennen, das zwischen den Vorstellungen einer *natürlichen Kontinuität* und ihrer nachgewiesenen *kulturellen Relativität* besteht.

Wie bereits gezeigt wurde, wird die Geschlechterdifferenzierung medizinisch abgesichert und gemäß einer im okzidentalen Kulturraum etablierten Episteme der anatomischen *Inkommensurabilität der Geschlechter* objektiviert (Laqueur, 1992). Die Funktion der Objektivierung besteht darin, Handlungs- und Orientierungshilfen zu etablieren, indem die Grenzen von Gruppen abgesichert und Geschlechtszugehörigkeiten durch ein medizinisches »Wissen vom Körper« (Keller & Meuser, 2011) in eine bivalente Struktur überführt werden. Die körperfixierte Zuordnung von Geschlecht beruht dann darauf, dass bestimmte Kriterien innerhalb eines Sozialsystems ontologisiert und Zugänge in diesem über Bewertungen von Körpern festgelegt werden. Dieses Zusammenspiel von Kriterien und Bewertungen ist integraler Bestandteil einer evaluativen Infrastruktur, die an konkrete Weisen des Untersuchens, Sichtbarmachens und Taxieren von Organen gekoppelt ist (Fourcade & Healy, 2017).

Abbildung 3 ordnet die Geschlechtstests chronologisch an, die seit ihrer offiziellen Einführung im Jahr 1946 in der professionellen Leichtathletik durchgeführt werden (Krämer, 2023, S. 199; Krämer & Schyvinck, 2024, S. 355). Wenngleich eine solche kompakte Darstellung zu reduktionistischen Fehlannahmen verführen und über die gesellschaftliche Einbettung der jeweiligen Tests hinweg täuschen kann, so zeigt das Arrangement, wie die Segregationslinie immer auch mit einem medizinischen *deconstructing sex* bzw. *undoing categorization* verbunden ist (Massa et al., 2023). Erkennen lässt sich so, wie in der Geschichte der Tests ein körperlicher Polymorphismus einerseits einem »ärztlichen Blick« (Foucault, 1963) unterworfen und andererseits über die Einrichtung von pathologischen Rangfolgen, Kombinationen und Höchstwerten in eine bivalente Struktur überführt wird. Im Sinne eines evaluativen »what

is counted counts» (Miller, 2001, S. 386), wird bestimmten Körpermerkmalen eine geschlechts- und damit kategorienanzeigende Qualität zugesprochen, die über medizinische Metriken und Skalierungen objektiviert wird.

Abbildung 3: Polaristische Kategorisierung der Geschlechter im Sport

genital based classification (1946-1966)	Normal ♀	I	II	III	IV	V	Normal ♂
chromosome based classification (1967-1998)							
	Male (XX)	Female syndrome (XO)	Turner syndrome (XXX)	Klinefelter syndrome (XY)	Complete AIS (XXY)	Partial AIS (XY)	5-RD (XX)
Anatomical	fail	pass	pass	fail	pass	p/f/ND	p/f/ND
Y-fluorescence	fail	pass	pass	fail	fail	fail	pass
Barr body	fail	pass	fail	pass	fail	fail	fail
SRY/DYZ1+*	fail	pass	pass	fail	fail	fail	pass
Eligibility based on testosterone level							
Legitimate Testosterone		Tested Athletes		Reference			
0-10 nmol		All Athletes		2011 Hyperandrogenism Regulations (IAAF)	2012 IOC Regulations on Female Hyperandrogenism (IOC)	2019 Eligibility Regulations for the Female Classification (IAAF)	2019 Eligibility Regulations for the Female Classification (WA)
0-5 nmol		46,XX & 46,XY Intersex Athletes		2021 IOC Framework on Fairness, Inclusion and Non-Discrimination on the Basis of Gender Identity and Sex Variations (IOC)			
0-5 nmol		46,XF Intersex Athletes		2023 Eligibility Regulations for the Female Classification (WA)			
0-2,5 nmol		5α-reductase, partial androgen insensitivity, 17β-hydroxysteroid, evotestosterone DSD					

Analog zum jeweiligen medizinischen Forschungsstand haben sich über die Jahrzehnte Differenzierungsformen etabliert, die Genitalien, Chromosomen und Hormone als zentrale Differenzkkriterien heranziehen und dabei andere leistungsrelevante anthropometrische Faktoren wie Körpergröße, -gewicht, kardiovaskuläre Faktoren oder den Gesamtmuskelanteil bestenfalls indirekt einschließen (Hackbart, 2020).

- (1) Im Zeitraum zwischen 1946 und 1966 wurde die kategoriale Zugehörigkeit über eine visuelle Bewertung der äußereren Genitalien ermittelt. Diese wurden innerhalb eines polaristischen Modells hinsichtlich ihrer dichotomen Ausprägungen eingestuft. Die Teilnahmefähigkeit wurde innerhalb einer pathologischen Rangfolge beruhend auf der in der Medizin gebräuchlichen »Praderskala« festgestellt. Damit wurde die Teilnahmefähigkeit von einer ästhetischen Bewertung des Äußeren abhängig gemacht, für die weitere Anhaltspunkte, wie das *tatsächliche Funktionieren* der Organe (zum

Beispiel die Fortpflanzungs-, Urinier- oder Erektionsfähigkeit), keine Rolle spielten.

- (2) Die zwischen 1967 und 1998 durchgeführten Tests wiederum machten die Teilnahmefähigkeit von der Chromosomenkombination abhängig. Je nach Test konnte der Zugang zur Frauenkategorie davon abhängen, ob sich ein Y- oder ein zweites X-Chromosom nachweisen ließ. Dies konnte rückblickend insofern zu Verwirrungen führen, da einige Personen den einen Chromosomentest bestehen und von einem anderen disqualifiziert werden konnten, wie zum Beispiel jene mit einem medizinischen Klinefelter-Syndrom, mit chromosomal Mosaiken oder Turner-Syndrom (Rupert, 2011).
- (3) Die 2011 eingeführten Testosterontests begründen die Segregationslinie zwischen Männer- und Frauenkategorie mit hormonellen Grenzwerten ($\leq 2,5 \text{ nmol/L}$ vs. $\geq 2,5 \text{ nmol/L}$), die in den vergangenen Jahren zunehmend verschärft und in Bezug zu weiteren medizinischen Diagnosen gesetzt wurden. Testosteron wird dabei eine probabilistische Qualität zugedacht, indem es »für Mannsein verantwortlich gemacht« wird (Schmitz & Degele, 2016, S. 8) und ein Überschuss entsprechend medizinisch etikettiert (»Hyperandrogenismus«, s.o.) sowie als Bedrohung für die Chancengleichheit eingestuft wird.

Daneben unterscheiden sich die beiden früheren Testformen ganz wesentlich von den Testosterontests, als ihnen eine *Kontinuitätsannahme* zugrunde gelegt wird: Sowohl Genitalien als auch Chromosomen werden ähnlich wie die Kategorisierung von Beeinträchtigungen im Behindertensport als »lebenslange Konstanten« (Hirschauer, 2014, S. 171) bewertet. Personen, die beabsichtigten in der Frauenkategorie anzutreten und denen ein *ambiges Genitale* oder ein *falscher Karyotyp* nachgewiesen wurde, blieben ihre weitere Sportkarriere von der Frauenkategorie ausgeschlossen.

Im Gegensatz hierzu drückt sich bereits im Begriff des Testosteron-Werts bzw. »Legitimate Testosterone« eine Kontinuitätsannahme aus: Testosteron lässt sich nicht nur geschlechtsübergreifend nachweisen und unterliegt in den jeweiligen Körpern kontinuierlichen Schwankungen (zum Beispiel je nach Uhrzeit, Alter und Tagesform), sondern der Wert lässt sich durch medikamentöse oder operative Eingriffe auch gezielt verändern. In dieser Hinsicht verweisen Testosterontests auf ein Management von »complicated things in ways that make it possible to intervene« (Castaño, 2021, S. 683), das darauf beruht, Körper nicht nur zu testen, sondern auch medizinisch zu normalisieren,

indem diese über die Modifikation eines symbolischen Geschlechtsmarkers an die strukturellen Gegebenheiten eines sozialen Systems angepasst werden. Es ist diese zynische Fokussierung auf ein fluides Differenzkriterium, die gegenwärtig vom organisierten Sport als Nachweis der Vermeidung von »discrimination on grounds of gender identity« (World Athletics, 2023, S. 5) betrachtet wird und zugleich als Bezugspunkt für ein invasives Eingreifen in die Körper der Athlet:innen herangezogen wird.

Quantifizierung und Laboratorisierung

Im Gegensatz zu früheren Genitaluntersuchungen, die den Zugang zur Frau-enkategorie von einer visuellen Prüfung abhängig machten und dabei ein vertrautes Geschlechtermodell aus der Medizin aufgriffen (Mildenberger, 2005), basieren die neueren Tests auf einer Unterscheidungslogik, die Zugehörigkeiten nicht mehr allein aus der Einschätzung einer ärztlichen Expertise (wie zum Beispiel bei gynäkologischen Untersuchungen) ableiten, sondern stattdessen mit Laborergebnissen begründen. Aus wissenschaftsgeschichtlicher Perspektive war die Praxis der visuellen Genitalinspektion in einen indifferenten Medizindiskurs eingebettet, der zuweilen noch verschiedene nicht-heteronormative Phänomene unter widersprüchlichen Begriffen wie dem des »intersexuellen Homosexuellen« zusammenfasste (Mildenberger, 2005, S. 278). In dieser Hinsicht markieren die neueren Testmethoden ein generelles Umdenken in der medizinischen Geschlechtszuordnung, das sich zur Mitte des 20. Jahrhunderts vollzog und sich in neueren Konzepten wie das des »Chromosomengeschlechts« zeigte (Bleuler & Wiedemann, 1956). Der organisierte Sport schließt an diese Diskursgeschichte an und nutzt die jeweiligen Deutungshoheiten, Unterscheidungspraktiken und Grenzwerte zur Etablierung einer bivalenten Struktur. Ausgehend hiervon haben sich zwei grundlegende Differenzierungsweisen im Sport etabliert: Quantifizierungen und Laboratorisierungen.

Quantifizierungen lassen sich mit Porter als »technologies of persuasion« (Porter, 1995) verstehen, die Phänomene der Kultur durch numerische Abstraktionen in eine universelle Metasprache überführen. Bewertungs- und kategorisierungssoziologisch sind sie insofern relevant, als durch Quantifizierungen »klassifizierende Vorgänge des Bestimmens, Bewertens und Einordnens durchgesetzt und der Wertigkeitsstatus einer Person oder Sache dabei in Zahlen ausgedrückt« (Mau, 2017, S. 24) wird. Zahlen fungieren dabei als eine Art »lingua franca« (Heintz, 2010, S. 173), die in der Lage ist, kulturelle Unschärfen kategorisch zu konturieren und damit Unvergleichbares ver-

gleichbar zu machen (Heintz 2021, S. 174). In Bezug auf die Vermessung von Geschlecht im Sport besteht eine zentrale Funktion der Quantifizierung darin, die Teilnahmefähigkeit der Athlet:innen über zahlenbasierte Normbereiche wie Testosteronlimits oder die Prader-Skala zu regulieren. Quantifizierungen machen damit die Heterogenität körperlicher Erscheinungen an konkreten Organen ablesbar und realisieren damit implizit eine *kulturelle De-Lokalisierung* von Geschlecht. Auf diese Weise lässt sich, bedingt durch die geringe Indekikalität von Zahlen, eine Opposition der Geschlechter durch den Transfer in Zahlen gegen kulturelle Andersheiten behaupten: Unabhängig davon, wie viele Geschlechter im Kulturkreis der jeweiligen Sportler:innen, in queeren Communities sowie im Personenstand anderer Länder anerkannt oder in wissenschaftlichen Kreisen außerhalb der Biomedizin verhandelt werden, suggeriert eine zahlenbasierte Festlegung, dass es sich hierbei um kulturübergreifend gültige Messgrößen handelt.

Daneben basieren Testmethoden wie Hormon- und Chromosomentests auf einer laboratorischen Wissensgenese. Im Gegensatz zu den empirischen Feldwissenschaften, die sich auf eine Erforschung dessen konzentrieren, was sie an Phänomenen in der Umwelt vorfinden, zielen Laboratisierungen darauf ab, diese Phänomene in experimentellen Settings selbst zu erzeugen (Stichweh, 1994). Laboratisierungen lassen sich als wissenschaftliche Praktiken betrachten, die »Beobachtungsgegenstände [...] unter kontrollierbaren Bedingungen« erzeugen und sie dadurch für »jedermann und jederzeit zur Verfügung« (Heintz & Werron, 2011, S. 372) stellen. Damit sind sie in der Lage, empirische Fakten von Zeit, Raum und Sache zu dekontextualisieren.⁴ Analog besteht die Besonderheit im Sport darin, dass kontingente Vorstellungen

4 In den Sozialwissenschaften werden Labore und Laboratisierungen von den Laborstudien untersucht. Für sie zentrale Autor:innen wie Latour, Lynch, Woolgar und Knorr-Cetina weisen anhand verschiedener Studien in Laboren der Biowissenschaften nach, dass die dort erzeugten hard facts soziale Konstruktionen darstellen, in die verschiedene Einflussfaktoren wie technische Möglichkeiten, Hierarchisierung von Institutionen, Trends im Experimentieren oder Karriereinteressen von Forschender einfließen (Latour & Woolgar, 1986; Knorr-Cetina, 1991; Graumann & Lindemann, 2010). Aus ihrer Perspektive handelt es sich bei Laboren um soziale Mikrokosmen, in denen Menschen, Diskurse, Praktiken und Instrumente kontinuierlich neu konfiguriert werden. Dabei werden nicht nur diskursive Maßnahmen eingesetzt, sondern auch »semiotische Mechanismen, durch die materielle Substanzen in Diagramme, Tabellen oder Sätze überführt werden und so zu Repräsentanten der vergänglichen Natur im Experiment werden« (Kaiser & Maasen, 2010, S. 691).

von Geschlecht im Labor isoliert und diese Isolationen anschließend in das soziale System des Sports (zurück-)übertragen werden. Als externalisiertes Verfahren der Geschlechterverifikation fungiert die Laboratorisierung auf diese Weise als ein »Härtungsprozess des Wissens« (Kaiser & Maasen, 2010, S. 692), der die Teilnahmefähigkeit der Athlet:innen als »kontinuierlichen und erwartbaren Prozess« (Heintz & Werron, 2011, S. 372) und damit als eine Art »self-vindication« (Hacking, 1992) modelliert.

Schluss

Das Anliegen des vorliegenden Beitrags war es, die Kategorisierung von Geschlecht im Wettkampfsport der Leichtathletik entlang ihrer kategorialen Irritationen zu untersuchen. Ich habe gezeigt, dass die Geschlechtertrennung dem Prinzip der Leistungsklassenbildung folgt und die Unterscheidung in Frauen und Männer mit nationalistischen Fairnessargumenten begründet wird. In diesem Zusammenhang wurde verdeutlicht, dass die binäre Segregation von intersexuellen Athlet:innen herausfordert wird, deren nichtbinäre Körper als genuiner Ausdruck einer zu regulierenden Geschlechterambiguität im Sport betrachtet werden und die zugleich in den jeweiligen Formalisierungen der Gender Regularien als Referenz für ein kontinuierliches »boundary making« fungieren.

Organisatorisch wird die Segregationslinie über Geschlechtstests abgesichert: Sie basieren auf einer spezifischen Bewertung und Kategorisierung von kontingenten Körperindizien wie Genitalien, Geschlechtschromosomen und Sexualhormonen. Vor diesem Hintergrund zeichnet sich die Geschlechterdifferenzierung durch zwei Diskontinuitäten aus, die sich dahingehend bemerkbar machen, wie einerseits Kategorien (Mann/Frau) aufrechterhalten sowie die Aufrechterhaltung über die Schaffung von Intra-Kategorien (»relevant athlete«) stabilisiert werden.

Es bleibt schließlich noch darauf hinzuweisen, dass eine solche strukturelle Perspektive bislang im organisierten Sport kaum eine Rolle spielt. Eine gesellschafts- und kultursensible Kontingenzer- oder Kontextperspektive vermisst man in den Stellungnahmen der Sportverbände, obwohl es sich bei ihnen um global und damit kulturübergreifend agierende Institutionen handelt. Diese kommen bei ihrer bisherigen biologistischen Sortierung der Menschen bislang gewissermaßen ohne diese Perspektiven aus und thematisieren die binäre Kategorisierung aus den gesetzten Strukturen heraus.

Vor diesem Hintergrund besteht im aktuellen Sport eine Herausforderung darin, Inklusion und Differenz zusammenzudenken. Fourcade beschreibt dies als einen generellen Konflikt, mit dem sich modernen Gesellschaften aktuell konfrontiert sehen (Fourcade, 2016). Auf der einen Seite sind sie dazu angehalten, kulturelle Differenzen anzuerkennen und diese in entsprechenden Leitbildern wie der »multikulturellen Gesellschaft« mit entsprechenden politischen Vorkehrungen zu verankern. Auf der anderen Seite setzt eine sich als meritokratisch verstehende Gesellschaft voraus, dass Personen nicht nach askriptiven Merkmalen, sondern nach ihrer Leistung beurteilt werden. Beide Prinzipien stehen im Widerspruch zueinander.

Zugleich zeichnet sich jedoch seit einigen Jahren, gewissermaßen als epistemischer Hoffnungsschimmer, eine Gegenbewegung zur Kolonisierung des Sports mit binären Geschlechtermodellen ab, die mit der internationalen Anerkennung dritter bzw. diverser Geschlechter in Gesellschaftsbereichen wie Alltag, Kunst und Recht korrespondiert und sich mit Inklusion- und Antidiskriminierungsdebatten verschränkt (Voß, 2023). Als Belege hierfür lassen sich die öffentlichen Stimmen von Athlet:innen wie Caster Semenya und Laurel Hubbard anführen, aber ebenso die Zunahme sozial- und sportwissenschaftlicher Beiträge zum Thema sowie eine semantische Diskontinuität auf Seiten der Sportverbände. So sind Einrichtungen wie der Deutsche Fußballbund (DFB) oder der Deutsche Leichtathletikverband (DLV) vor kurzem dazu übergegangen, ein geschlechterinklusives Sportumfeld zu schaffen und die geschlechtliche Selbstzuordnung als Entscheidungskriterium für die Kategorisierung heranzuziehen, zumindest im Amateurbereich (DLV, 2023).

Literatur

- Bahro, B. (1995). Ein Schweres Schicksal: Dora Ratjen. In B. Bahro & J. Braun (Hg.) Berlin 36. Das Buch zum Film (S. 63–98). VBB Verlag.
- Baumann, H. (1955). Das doppelte Geschlecht: ethnologische Studien zur Bisexualität in Ritus und Mythos. Reimer.
- Baumann, Z. (1995). Moderne und Ambivalenz: Das Ende der Eindeutigkeit. Hamburg Edition.
- Bennani, H. (2022). Behinderung klassifizieren. Zur Kontingenz und Normativität von Körperbewertungen in der International Classification of Functioning, Disability and Health. *Öster Z Soziol*, 47(3), 247–268. DOI: 10.1007/s11614-022-00498-9.

- Bette, K.-H. (2010). Sportsoziologie. transcript.
- Bleuler, M., & Wiedemann, H. R. (1956). Chromosomengeschlecht und Psychosexualität. *Arch f Psych u Z f Neurol* 195, 14–19.
- Blithe, S. J., & Hanchey, J. N. (2015). The Discursive Emergence of Gendered Physiological Discrimination in Sex Verification Testing. *Wom Stud Comm* 38(4), 486–506.
- Bourdieu, P. (1991). Sport and social class. *Soc Sci Inf* 17(6), 819–840.
- Bowker, G. C., & Star, S. L. (2000). Sorting Things Out: Classification and Its Consequences. MIT Press.
- Burrmann, U. (2021). Sportbezogene Sozialisation. In A. Güllich & M. Krüger (Hg.), *Sport in Kultur und Gesellschaft* (S. 196–216). Springer Nature.
- Cahn, S. (1994). Coming on Strong: Gender and Sexuality in Twentieth Century Women's Sport. Free Press.
- Castaño, P. (2021). From Value to Valuation: Pragmatist and Hermeneutic Orientations for Assessing Science on the International Space Station. *Am Soc* 52 (4), 671–701.
- Cole, C. L. (2000). One Chromosome Too Many? In K. Schaffer & S. Smith (Hg.), *The Olympics at the Millennium. Power, Politics and the Games* (128–146). Rutgers.
- Cools, M. et al. (2018). Caring for individuals with a difference of sex development (DSD): A Consensus Statement. *Nature* 14, 415–429.
- DLV (2023, 18. Nov.). Deutsche Leichtathletik-Ordnung (DLO). [https://www.leichtathletik.de/fileadmin/user_upload/o06_Wir-im-DLV/o3_Struktur/](https://www.leichtathletik.de/fileadmin/user_upload/o06_Wir-im-DLV/o3_Struktur/DLV_Satzung_Ordnungen/Deutsche_Leichtathletik-Ordnung.pdf) DLV_Satzung_Ordnungen/Deutsche_Leichtathletik-Ordnung.pdf
- Dreger, A. D. (2000). Hermaphrodites and the Medical Invention of Sex. Harvard UP.
- Douglas, M. (1992). Reinheit und Gefährdung. Suhrkamp.
- Durkheim, E., & Mauss, M. (1963). Primitive Classification. Univ of Chic Press.
- Epple, A., & Erhart, W. (2015). Die Welt beobachten. Praktiken des Vergleichens. Campus.
- Fausto-Sterling, A. (2012). Sex/Gender: Biology in a Social World. Routledge.
- Foucault, M. (1963). Die Geburt der Klinik: Eine Archäologie des ärztlichen Blicks. Hanser.
- Foucault, M. (1966). Les Mots et les choses. Une archéologie des sciences humaines. Gallimard.
- Fourcade, M. (2016). Ordinalization. Lewis A. Coser Memorial Award for Theoretical Agenda Setting. *Soc Theor* 34, 175–195.
- Fourcade, M., & Healy, K. (2017). Seeing like a market. *Soc Econ Rev* 15, 9–29.

- Garfinkel, H. (1967). *Studies in Ethnomethodology*. Prentice Hall.
- Gebauer, G. (1972). Leistung als Aktion und Präsentation. *Sportwiss* 2(2), 182–203.
- Gilbert, M., & Gillett, J. (2013). Women in Equestrian Polo: Cultural Capital and Sport Trajectories. In M. Adelman, & J. Knijnik (Hg.), *Gender and Equestrian Sport. Riding Around the World* (S. 111–125). Springer.
- Graumann, S., & Lindemann, G. (2010). Medizinsoziologie. In G. Kneer & M. Schröer (Hg.), *Handbuch spezielle Soziologien* (S. 295 – 308). VS.
- Goldschmidt, R. (1917). Intersexuality and the endocrine aspect of sex. *Endocr* 1, 433–456.
- Günter, S. (2015). The illegal transgression: discourse analysis of the media perception of the transgressive aesthetic of performance and display in top-level sports. *Spo Soc* (11), 1–14.
- Hacking, I. (1992). The Self-Vindication of the Laboratory Sciences. In A. Pickering (Hg.), *Science as Practice and Culture* (S. 29–64). Chicago Univ Press.
- Hacking, I. (1999). *The Social Construction of What?* Harvard Univ Press.
- Heckemeyer, K. (2018). Leistungsklassen und Geschlechtertests. Die heteronormative Logik des Sports. transcript.
- Heintz, B. (2010). Numerische Differenz. Überlegungen zu einer Soziologie des (quantitativen) Vergleichs. *Z Soziol* 39(3), S 162–181.
- Heintz, B. (2021). Kategorisieren, Vergleichen, Bewerten und Quantifizieren im Spiegel sozialer Beobachtungsformate. *Köln Z Soziol* 73(S1), S 5–47.
- Heintz, B., & Werron, T. (2011). Wie ist Globalisierung möglich? Zur Entstehung globaler Vergleichshorizonte am Beispiel von Wissenschaft und Sport. *Köl Z f Soziol* 63(3), 359–394.
- Hirschauer, S. (1993). *Die Soziale Konstruktion der Transsexualität. Über die Medizin und den Geschlechtswechsel*. Suhrkamp.
- Hirschauer, S. (2014). Un/doing Differences. Die Kontingenz sozialer Zugehörigkeit. *Z f Soz* 43(3), 170–191.
- Hirschauer, S., & Boll, T. (2017). Un/doing Differences. Zur Theorie und Empirie eines Forschungsprogramms. In Dies. (Hg.), *Un/doing Differences. Praktiken der Humandifferenzierung* (S. 7–26). Velbrück.
- Hoad, N. (2010). »Run, Caster Semenya, Run!« Nativism and the Translations of Gender Variance. *Safundi* 11(4), 397–405.
- Holmes, Morgan (2009). *Straddling Past, Present and Future*. In Dies. (Hg.), *Critical Intersex* (S. 1–12). Ashgate.
- Jenkins, R. (1996). *Social Identity*. Routledge.

- Kaiser, M., & Maasen, S. (2010). Wissenschaftssoziologie. In G. Kneer & M. Schroer (Hg.), *Handbuch spezielle Soziologien* (S. 685–706). VS.
- Kalthoff, H., & T. Dittrich (2016). Unterscheidung und Härtung. Bewertungs- und Notenkommunikation in Lehrerzimmer und Zeugniskonferenz. *Berl J f Soz* 26(3-4), 459–483.
- Keller, R., & Meuser, M. (2011). Wissen des Körpers – Wissen vom Körper. In Dies. (Hg.), *Körperwissen* (S. 9–27). VS.
- Kessler, S. J. (1998). *Lessons from the Intersexed*. Rutgers University Press.
- Kessler, S. J. (1990). The medical construction of gender: Case management of intersexed infants. *Signs* 16(1), S 3–26.
- Klöppel, U. (2010). *XXOXY ungelöst. Hermaphroditismus, Sex und Gender in der deutschen Medizin; eine historische Studie zur Intersexualität*. transcript.
- Knorr-Cetina, K. (1991). Die Fabrikation von Erkenntnis. Zur Anthropologie der Naturwissenschaft. Suhrkamp.
- Knorr-Cetina, K. (1994). Primitive Classification and Postmodernity: Towards a Sociological Notion of Fiction. *Theo, Cult & Soc* 11, 1–22.
- Krämer, D. (2017). Mediale Praktiken des Gendering. Die Geschwister Press im westlichen Sportdiskurs zu Zeiten des Kalten Kriegs. G. Klein & H. Göbel (Hg.), *Die Performance der Praxis – Die Praxis der Performance* (S. 191–209). transcript.
- Krämer, D. (2020). Intersexualität im Sport: Mediale und medizinische Körperpolitiken. transcript.
- Krämer, D. (2022). Queer Studies. In R. Gugutzer, G. Klein & M. Meuser (Hg.), *Handbuch Körpersozialogie Band 1* (S. 395–409). Springer.
- Krämer, D. (2023). Time to abolish gender boundaries in elite sports? A plea for structural reflection. *Spo u Gesell* 20(2), 195–206.
- Krämer, D., & Shyvinck, C. (2024). Challenging the Binary: Gender, Fraud, and the Complexities of Categorization in Elite Sports. In R. Kühl, D. Link, & L. Heiberger (Hg.): *Sexualitäten und Geschlechter* (S. 347–365). transcript.
- Lamont, M., & Molnar, V. (2002). The Study of Boundaries in the Social Sciences. *Ann Rev Soc* 28, 167–195.
- Latour, B., & Woolgar, S. (1986). *Laboratory Life. The Construction of Scientific Facts*. Princeton University Press.
- Lang, C. (2006). Intersexualität. Menschen zwischen den Geschlechtern. Campus.
- Laqueur, T. W. (1992). Auf den Leib geschrieben: Die Inszenierung der Geschlechter von der Antike bis Freud. Campus.

- Lindemann, G. (2002). *Die Grenzen des Sozialen. Zur sozio-technischen Konstruktion von Leben und Tod in der Intensivmedizin.* Wilhelm Fink.
- Lindemann, G. (2009). *Das Soziale von seinen Grenzen her denken.* Velbrück.
- Lindemann, G. (2011). *Das paradoxe Geschlecht.* VS.
- Linghede, E. (2018). The promise of glitching bodies in sport: a posthumanist exploration of an intersex phenomenon. *Quale Res in Sport, Exerc and Heal* 10(5), 570–584.
- Martinez-Patino, M. J. (2005). Personal account: a woman tried and tested, *Lanc* 366, 38.
- Massa, M. G., Aghi, K., & Hill M. J. (2023). Deconstructing sex: Strategies for undoing binary thinking in neuroendocrinology and behavior. *Horm and Beh* 156, 105441.
- Mau, S. (2017). *Das metrische Wir. Über die Quantifizierung des Sozialen.* Suhrkamp.
- Mildenberger, F. (2005). Diskursive Deckungsgleichheiten. *Hermaphroditismus und Homosexualität im medizinischen Diskurs (1850–1960).* In F. Stahnisch & F. Steger (Hg.), *Medizin, Geschichte und Geschlecht. Körperhistorische Rekonstruktionen von Identitäten und Differenzen* (S. 259–283). Franz Steiner.
- Miller, P. (2001). Governing by numbers. Why calculative practices matter. *Soc Res* 28, 379–390.
- Müller, M. (2006). Geschlecht als Leistungsklasse. Der kleine Unterschied und seine großen Folgen am Beispiel der gender verifications im Leistungssport. *Z f Soz* 35(5), 392–412.
- Müller, M. (2017). Unvergleichbarkeitskonstruktionen im Sport. Von Frauen mit Hyperandrogenismus und Männern mit Carbonprothesen. In S. Hirschauer (Hg.), *Un/doing differences. Praktiken der Humandifferenzierung* (S. 205–233). Velbrück.
- Pape, M. (2017). The Fairest of them All: Gender-Determining Institutions and the Science of Sex Testing. In V. P. Demos & M. T. (Hg.), *Gender panic, Gender policy* (S. 177–200). Emerald.
- Parry, J., & Martíneková, I. (2021). The logic of categorisation in sport. *Europ J Spor Sci* 21(11), 1485–1491.
- Porter, T. W. (1995). *Trust in Numbers. The Pursuit of Objectivity in Science and Public Life.* Princeton UP.
- Quandt, T., & Kröger, S. (2014). *Multiplayer. The Social Aspects of Digital Gaming.* Routledge.

- Reckwitz, A. (2008). Unscharfe Grenzen. Perspektiven der Kultursoziologie. transcript.
- Rohrer, T., & Haller, M. (2015). Sport und soziale Ungleichheit – Neue Befunde aus dem internationalen Vergleich. *Köl Z f Soz u Sozialpsych* 67(1), 57–82.
- Rupert, J. L. (2011). Genitals to Genes: The History and Biology of Gender Verification in the Olympics. *CBMH/BCHM* 28(2), 339–365.
- Schmidt, R., Weigelin, M., & Kempf, D. (2024). Einführung: Leistungsvergleiche und evaluative Praktiken. Sport als instruktiver Fall für eine Soziologie der Bewertung. In Dies. (Hg.), Leistungsvergleiche und evaluative Praktiken. Sport als instruktiver Fall für eine Soziologie der Bewertung. transcript.
- Schmitz, S., & Degele, N. (2016). Testosteron als bio-soziales Konzept: Zur wechselseitigen Transformation von Hormonen, Körpern und Sozialem. https://www.soziologie.uni-freiburg.de/personen/degele/dokumente-publikationen/Schmitz_Degele_Testosteron%20als%20bio-soziales%20Konzept_formatiert.pdf
- Simmel, G. (1890). Über sociale Differenzierung. Duncker & Humblot.
- Steuerwald, Ch. (2017). Warum machen Unterschiede einen Unterschied? In M. Müller & Ch. Steuerwald (Hg.), *Gender, Race und Disability im Sport* (S. 143–176). transcript.
- Stichweh, R. (1994). Wissenschaft, Universität, Professionen. Soziologische Analysen. Suhrkamp.
- Stone, D. (1988). Policy paradox and political reason. Brown College Division.
- Sykes, H. (2006). Transsexual and transgender policies in sport. *Wom Phys Activ* J 15, 3 –13.
- Voß, H. J. (2023). Geschlechtliche und sexuelle Pluralisierung. In K. Weiler & H. J. Voß (Hg.), *Sexualität und Partnerschaft der Deutschen* (S. 191–212). Psychosozial.
- Wacke, A. (1989). Vom Hermaphroditen zum Transsexuellen. Zur Stellung von Zwittern in der Rechtsgeschichte. In E. Heinz, W. Odersky, F. Säcker & F. Jürgen (Hg.), *Festschrift für Kurt Rebmann zum 65. Geburtstag* (S. 861–903). C.H. Beck.
- Wagner, G., J. S. Guse, M. Hasenbruch (2023). »Eigentlich war es immer sonnenklar.« Zur Invisibilisierung von formaler Organisation in Bewertungspraktiken. *Berl J f Soz.* Online: <https://link.springer.com/content/pdf/10.1007/s11609-023-00492-2.pdf>

- Wiederkehr, S. (2008): »Mit zweifelsfreier Sicherheit ...keine Frau«. Geschlechtertests im Spitzensport zwischen medizinischer Expertise und Technik euphorie der Funktionäre. *Technikgesch* 75(3), S 253–270.
- World Athletics (2023): Eligibility Regulations for DSD. <https://worldathletics.org/download/download?filename=2ffb8b1a-59e3-4cea-bb0c-5af8b690d089.pdf&urlslug=C3.6A%20%E2%80%93%20Eligibility%20Regulations%20for%20the%20Female%20Classification%20%E2%80%93%20effective%2031%20March%202023>
- Yanow, D. (2003). Constructing race and ethnicity in America: Category-making in public policy and administration. M. E. Sharpe.

Bahnen ziehen. Schwimmen als Bewertungsspiel

Thorsten Peetz

Einleitung

Bewertungen findet man im Sport an allen Ecken und Enden. Der Schiedsrichter*innenpfiff im Fußball etwa sortiert Ereignisse auf dem Platz in gewertete Kategorien (regelkonform/nicht regelkonform) und stellt die Pfeifende vor das Problem, dass er ihr als Entscheidung zugerechnet werden kann (Weigelin, 2022). Im Eiskunstlauf bewerten Juries die künstlerische wie technische Qualität einer Darbietung (Lom, 2016). In manchen, stark auf Quantifizierung setzenden Sportarten, ist Leistungsbewertung eindeutig (Chambliss, 1988, S. 97). Aber auch dort finden sich immer wieder Bewertungsprobleme, die bearbeitet werden müssen: In der Leichtathletik, in der Messungen vermeintlich unproblematisch vermeintlich objektive Ergebnisse produzieren, betten Auswahltrainer*innen gemessene Leistungen in den Kontext der Leistungserbringung ein, um die Qualität (und erwartbare zukünftige Leistungsfähigkeit) einer Langsprinterin zu bestimmen (Janetzko, 2021).¹ Der Sport bietet sich deshalb als Untersuchungsfeld für eine an Praktiken und Prozessen der Bewertung interessierte Soziologie an. Wenn darüber hinaus die Verstrickung von Körpern in Bewertungsprozesse und die Materialität der Bewertungspraxis interessieren (Schmidt et al., 2022), führt am Sport als einem Pool von »model cases« (Krause, 2021) kein Weg vorbei.

Das »Spiel« hat seinen Weg in die Soziologie der Bewertung weniger als Forschungsgegenstand denn als ein Konzept gefunden, das eine neue Perspek-

¹ Im Kontrast zur Leichtathletik oder dem Schwimmsport, die auf Haltungsnoten verzichten, zeigt das Skispringen, dass Quantifizierbarkeit im Sport keine hinreichende Bedingung für den Verzicht auf qualitative Bewertungen ist. Dass das »schöne« Spiel im Fußball ebensowenig offiziell gewürdigt wird, wie ein ästhetisch ansprechendes Delphin im Schwimmen ist also nicht selbstverständlich sondern erklärmgsbedürftig.

tive auf Bewertungsgeschehen eröffnet. Als Bewertungsspiele (Hutter, 2015) kann man nicht nur Spaß und »unernstes« Spiel und auch nicht nur sportliche Spiele fassen, man kann den Begriff vielmehr als ein sozialtheoretisches Werkzeug verwenden, das die Differenziertheit sozialer Zusammenhänge sowie die differenzierende und konstituierende Leistung von Bewertungspraktiken hervorhebt. Unter »Bewertungsspiele« in diesem Sinn fallen auch »ernste« Spiele, wie zum Beispiel wirtschaftliche und künstlerische Spiele (Hutter, 2015) oder intime Spiele (Peetz, 2022, 2023).

Es liegt nahe, Sport und (Bewertungs-)Spiele zusammenzudenken und genau das möchte ich in diesem konzeptuellen Beitrag tun. Ich werde mich dabei auf die sportsoziologische Literatur zu Bewertungen beziehen, zur Veranschaulichung der Leistungsfähigkeit der Theorie der Bewertungsspiele aber auch auf einen Gegenstandsbereich aus dem Feld des Sports eingehen, dessen Bewertungspraktiken mir zwar nicht als empirischer Forschungsgegenstand bekannt sind, alltäglich aber wohl: dem Schwimmsport.

1. Bewertungsspiele

Der Begriff der Bewertungsspiele (Hutter, 2015) stellt eine begriffliche Alternative zu Konzepten wie Feld, System oder Existenzweise dar. Zentral für den *Spielbegriff* ist die Annahme, dass Spiele sich von ihrer Umwelt unterscheiden und die Spielenden sich in ihrer Sicht der Welt darauf einstellen und ihre Handlungen entsprechend orientieren (Bateson, 1955; Huizinga, 1938). Die Umwelt wird im Spiel zunächst ausgeklammert, und wenn sie sich bemerkbar macht, dann als Irritation oder Störung, auf die sich die Spielenden ggf. einstellen müssen. Spiele in diesem Sinn sind abstrakte Sinnzusammenhänge bzw. »play«, die von den konkreten Spielzusammenhängen bzw. »games« unterschieden werden können – eine »Partie« als eine Realisierung des »Schachspiels« als einer Form des Spiels – ihre Grundeigenschaft, die Grenzziehung, aber teilen (Hutter, 2021).²

² Dass die Autonomie von Spielen nur eine »relative« ist (Bourdieu), sich klar voneinander unterscheidende Sozialzusammenhänge »gekoppelt« sind (Luhmann) ist ein soziologischer Gemeinplatz. Zur Unterscheidung von Spiel und Welt bei Huizinga und daran anschließenden Debatten in den Game Studies vgl. Consalvo (2009). Anders als hier vorgeschlagen argumentiert Geertz (1980) mit Bezug auf Goffman, dass der Begriff des Spiels den strategischen Charakter sozialen Handelns hervorhebt.

Mit dem Begriff des Spiels kommen die »real differences« (Latour, 2013, S. 35), die die Leute trotz aller Unschärfen und Vernetzungen in ihrem sozialen Leben machen in den Fokus rekonstruktiver Gesellschaftsanalysen (Peetz, 2020). Das theoriearchitektonisch interessante Moment dieses Begriffs besteht darin, dass er es erlaubt sowohl einen hohen Abstraktionsgrad zu erreichen – indem etwa Spiele als gesellschaftliche Sinnzusammenhänge unterschieden werden – und gleichzeitig hohe Anschaulichkeit zuzulassen. Das theoriepolitisch interessante Moment besteht darin, dass der Begriff in den unterschiedlichsten Theorietraditionen und sozialwissenschaftlichen Disziplinen anschlussfähig zu sein scheint, von der »Spieltheorie« (Neumann, 1928) über Praxis- (Bourdieu & Wacquant, 2006) und Systemtheorie (Hutter, 2015) bis hin zu eher interaktionistischen Ansätzen (Goffman, 1972). Der Begriff des Spiels hat das Potenzial zum »boundary object« (Star & Griesemer, 1989), das unterschiedliche sozialtheoretische Traditionen zur Koordination ihrer Theoriearbeit verwenden könnten.

Konstitutiv für Spiele ist eine an Regeln orientierte Praxis. Zum Spiel wird Verhalten dann, wenn Leute sich an Regeln orientieren (Caillois, 1958, S. 10; Huizinga, 1938, S. 20) – sei es durch ihre strikte Anwendung, den kreativen Umgang mit ihnen oder den mehr oder weniger bewussten Regelbruch (vgl. dazu Ortmann, 2003) – oder von anderen nach Maßgabe solcher Regeln beobachtet werden. Dieses Verständnis von Regeln lehnt sich einerseits stark an Überlegungen Ludwig Wittgensteins (1953) an. Regeln in diesem Sinn zeichnen sich dadurch aus, dass sie häufig als selbstverständlich vorausgesetzt werden, in und durch die soziale Praxis artikuliert werden und Regelmäßigkeiten des Verhaltens produzieren, ohne es zu determinieren (vgl. auch Puhl, 2002; Puhl, 2010; Schäfer, 2013, S. 27ff.). Andererseits nimmt es die Anregung Niklas Luhmanns auf, zur Erklärung des Handelns weniger auf Interessen oder Motive der Handelnden – die ohnehin nur vermittelt beobachtbar sind – als auf die Beobachtung von Ereignissen und die Konsequenzen ihres Verstehens als Handlungen zu setzen. Natürlich macht die »Motivation« von Leuten, an Spielen teilzunehmen einen Unterschied für die Spiele selbst – Arbeitsorganisationen mit ihrem ganzen Apparat von Anreiz- und Kontrollstrukturen machen das nur zu deutlich. Wie das Verhalten innerhalb des Spiels wirksam wird, ist aber stark abhängig von den Beobachtungsschemata und Anslüssen der weiteren Spielenden. Die in den geteilten Regeln des Spiels eingearbeiteten Annahmen über die Wirklichkeit – ontologische Regeln (Waibel et al., 2021) –

können Effekte ganz unabhängig von den tatsächlichen Beiträgen der Spieler*innen produzieren (Frye, 2017).³

Bewertungen werden in der Praxis des Spiels in zwei Hinsichten relevant. Zum einen haben Praktiken einen zumindest latenten *Wertsinn* – sie können nicht nur mit Blick auf das sachliche Thema, die implizierten sozialen Beziehungen und ihre zeitlichen Bezüge hin beobachtet werden (Luhmann, 1971), sondern auch mit Blick auf Fragen der Bedeutsamkeit (Weber, 1904, S. 180f.). In diesem Sinn sind Praktiken überschattet von potenziellen Wertfragen, können immer als wertende beobachtet werden, selbst dann, wenn ihnen keine Bewertungsprobleme (Dewey, 1939) zugrunde liegen mögen. In *Bewertungskommunikation* wird der Wertsinn der Praxis selbst thematisch. Sie wird als eine explizit wertende Praxis ausgeflaggt: Notenvergabe, moralisches Werturteil, Rezension, Beleidigung sind Prototypen.

In beiden Hinsichten kann Bewertung als eine komplexe soziale Leistung analysiert werden, in der mindestens vier soziale Praktiken aufeinander bezogen werden: Bewertungen identifizieren, kategorisieren, relationieren und valorisieren Bewertungsobjekte (vgl. hierzu Peetz, 2021):⁴

- (1) Bewertungen haben Teil an der Konstruktion des Bewertungsobjekts indem sie dessen Grenzen definieren (vgl. z.B. Schäfer, 2016, S. 367ff.). Die Identifizierung eines Bewertungsobjektes kann in manchen Kontexten unproblematisch sein oder zumindest so erscheinen, in anderen ist sie problematisch und gegebenenfalls sogar umstritten. Sind die Grenzen eines zu bewertenden Kunstwerks z.B. koextensiv mit dem materiellen Rahmen des Objektes oder gehen sie darüber hinaus und schließen zum Beispiel auch politische Äußerungen der Künstler*in mit ein?
- (2) Bewertungen kategorisieren Bewertungsobjekte, indem sie sie in mehr oder weniger gewichtete Kategorien sortieren. Während in der Soziologie

3 Frye (2017) zeigt in ihrer Arbeit wie im Malawischen Bildungssystem Annahmen über die Unvereinbarkeit von Liebesbeziehungen mit Bildungserfolg dazu führen, dass Lehrer*innen Schüler*innen in romantischen Beziehungen bestrafen und Angehörige ihnen die (auch: finanzielle) Unterstützung entziehen und so zum Abbruch der Bildungslaufbahn bei entsprechenden Jugendlichen beitragen. Ohne dass ein entsprechender Leistungsabfall empirisch feststellbar wäre.

4 Alternative Vorschläge zur Unterscheidung und Relationierung von Elementen, Dimensionen oder Praktiken des Bewertens geben Heintz und Wobbe (2021) sowie Lamont (2012).

häufig in der Tradition von Durkheim und Mauss (1993) davon ausgegangen wird, dass Kategorien zwangsläufig unterschiedliche Wertigkeiten zuschreiben (z.B. Hirschauer, 2014, S. 174), Kategorisierung mehr oder weniger identisch mit Bewertung ist, argumentiere ich im Anschluss an kognitionswissenschaftliche Arbeiten (z.B. Rosch, 1978), dass Kategorien als Praktiken des Unterscheidens zwar Unterschiede produzieren, diese aber nicht zwangsläufig Ungleichwertigkeiten implizieren. Ungleichwertige Kategorien sind Resultate vergangener Bewertungsarbeit, die die Ungleichwertigkeiten durch Bewertung von Kategorien produziert haben.

- (3) Bewertungen setzen Bewertungsobjekte zudem in Beziehungen zu anderen Identitäten – zum Beispiel, indem sie Vergleichsoperationen anstellen (Heintz, 2016) oder soziale Nähe behaupten. Letzteres z.B. indem sie die Biographie eines Objekts (Appadurai, 1986) mit der Biographie von Personen, in deren Besitz sie sich befanden, verbinden (Boltanski & Esquerre, 2018).
- (4) Schließlich setzen Bewertungen Bewertungsobjekte narrativ in Wert, sie valorisieren indem sie eine (noch so rudimentäre) Geschichte erzählen, die ihren Wert *accountable* macht. In diesen Geschichten werden Identitäten mit Kategorien und Relationen verknüpft und unter Bezug auf Bewertungskriterien als mehr oder weniger wertig qualifiziert. Ganz deutlich ist das bei den unterschiedlichsten Varianten qualitativer Bewertungen, die, zum Beispiel in der Form von Lebenserzählungen in Nachrufen (Hamann, 2016), die Qualität von Personen nachvollziehbar begründen. Aber auch quantifizierende Bewertungen – die Anzahl von Pferdestärken zum Beispiel – müssen durch eine Einbettung in Narrative Hinweise darauf geben, wie die Zahlen gedeutet werden sollen: als Zeichen von fortschrittlicher Technik oder als Nachhaltigkeitsproblem?

Mit dem Begriff der Bewertungsspiele kommen also die Regeln in den Blick, die soziale Praktiken in einer spezifischen Sphäre des gesellschaftlichen Lebens orientieren und dadurch die Sphäre als solche von anderen gesellschaftlichen Sphären unterscheiden; mit der Heuristik der Bewertungspraktiken wird genauer ausbuchstabiert, wie in Bewertungsspielen Bewertungsobjekte identifiziert und mit Wertigkeiten ausgewiesen werden. Ich werde im Folgenden auf diese theoretischen Werkzeuge selektiv zurückgreifen, um den Schwimmsport als ein Bewertungsspiel zu verstehen und das Bewertungsgeschehen in diesem Spiel zu identifizieren. Zur Vorbereitung und Einordnung dieser Ana-

lyse werfe ich zuvor noch einen kurzen Blick auf die Literatur zu Bewertungspraktiken im Sport.

2. Bewertungspraktiken im Sport

Aus der Perspektive einer Theorie der Bewertungsspiele handelt es sich beim Sport genauso wie bei der Kunst, der Wirtschaft, der Intimität oder der Religion um ein Bewertungsspiel, in dem die unterschiedlichsten Spielformen tagtäglich in unzähligen Partien oder Episoden zur Aufführung gebracht werden. Besonders deutlich wird beim Sport, dass Spiele in und durch Bewertungen realisiert werden: Bewertungen werden häufig expliziert, es wird sowohl durch als auch über Bewertung kommuniziert (vgl. Schmidt, Kempf & Schmidt 2024, in diesem Band). Der eingangs schon angesprochene »Fußball« stellt das deutlich unter Beweis. Das Publikum kann Bewertungen in der Form des Schiedsrichter*innenpfiffs beobachten, an den – wie man schon am heimischen Bildschirm erkennen kann – unter Umständen recht deutliche Bewertungen der Schiedsrichter*innenleistung selbst anschließen (Weigelin, 2022). Während des Spiels verfolgen Analyst*innen das Spiel am Bildschirm, kategorisieren jedes Ereignis und liefern so Daten, die in journalistischen wie in fußball-professionellen Bewertungen des Spiels genutzt werden können (Hodek, 2021). Die Videoaufzeichnungen der Spiele werden von Trainer*innen verwendet, um die Leistung des eigenen Teams einzuschätzen sowie Stärken und Schwächen gegnerischer Teams zu identifizieren und darauf das Training sowie die Spieldtaktik einzustellen (Brümmer, 2021). Schiedsrichter*innenpfiffe, Tabellen über Torschüsse und Ballbesitzanteile oder Videobesprechungen im Team stellen also Situationen expliziter Bewertungskommunikation her.

Neben diesen, das Spiel begleitenden und aufarbeitenden Bewertungspraktiken nehmen Bewertungen im Sport wichtige Rollen in der Karriere von Sportler*innen ein. Über Bewertung wird der Grad der »Mitspielfähigkeit« (Brümmer, 2015) von Spieler*innen und ihrer Professionalisierbarkeit entschieden wenn z.B. Mitglieder für Leistungskader ausgewählt werden (Janetzko, 2021). Am anderen Ende der Karriere stehen Praktiken der Konsekration, die – in den U.S.A. zum Beispiel institutionalisiert in *Halls of Fame* (Allen & Parsons, 2006) – ehemalige Mitspieler*innen identifizieren,

deren Praxis und Karriere als vorbildlich identifiziert und der Nachahmung empfohlen wird.⁵

Bewertungspraktiken im Sport, dienen also u.a. der Durchsetzung der Regeln des Spiels, der Kommunikation über das Spiel, seiner Verbesserung und der Auswahl und dem Lob verdienter Spieler*innen. An dieser Liste fällt auf, dass die untersuchten Bewertungspraktiken zwar im Spiel, aber doch am Rand der genuin sportlichen Praxis lokalisiert sind. Beobachtet wird die Bewertungspraxis von »engaged spectators« und »producers«, nicht die der »creators« der sportlichen Praxis (Hutter, 2021, S. 89, Herv. weggel.). Das ist in gewisser Weise erstaunlich, da gerade in professionalisierten sportlichen Kontexten die Praxis selbst Situationen – im Training wie im Wettkampf – bereithalten wird, die aufgrund der ihnen inhärenten Unsicherheit Bewertungen provozieren, ja unabdingbar machen werden.⁶ Die sportlichen Bewertungsspiele werden so durch eine Differenzierung von Zentrum (der fokalen sportlichen Praxis) und Peripherie (den diese Praxis ermöglichen, regulierenden, auswertenden Praktiken) charakterisiert.⁷ Die explizite Bewertung sportlicher Praxis wird bevorzugt beobachtet, ihr Wertsinn demgegenüber vernachlässigt.

3. »Schwimmen« als instruktiver Fall

Im Folgenden werde ich Bewertungen und sportliche Praxis am Beispiel des Schwimmsports genauer in den Blick nehmen. Dazu kann ich mich nicht auf eigens erhobene Daten beziehen – eine Auswertung der Literatur zur Soziologie des Schwimmens muss vorerst genügen. Dass das Schwimmen als Fall für eine am Körper interessierte Soziologie der Bewertung interessant sein könnte, hat durchaus systematische Gründe. Theoriegeschichtlich bildet die

5 Der Sport steht dabei in direkter Nachfolge zur Römisch-Katholischen Kirche, deren Verfahren der Kanonisierung (allerdings: bereits verstorbener) Mitspieler*innen wohl prototypisch für organisierte Personenbewertung in Bewertungsspielen ist. Vgl. dazu Meier und Peetz (2021).

6 Zu »Unsicherheit« als Ausgangspunkt von Bewertungen vgl. Dewey (1939).

7 Überlegenswert wäre, ob diese Beobachtung über sportliche Bewertungsspiele hinaus generalisiert werden kann. Die These würde entsprechend lauten, dass die Soziologie der Bewertung sich stark auf peripherie, aber explizite Bewertungsphänomene konzentriert, die Kunstkritik und nicht künstlerische Prozesse, Peer Review und nicht die wissenschaftliche Arbeit, Partnerinnenwahl aber nicht das Liebesleben untersucht.

Auseinandersetzung mit dem Schwimmen den »Ausgangspunkt« der soziologischen Thematisierung von »Techniken des Körpers« (Mauss, 1935, S. 200). Marcel Mauss verweist auf seine eigenen Erfahrungen des Wandels der Technik des Schwimmens – die Verbreitung des Kraulens – um auf die Bedeutung von Lernen für den Erwerb von Körpertechniken aber auch auf ihre Geschichtlichkeit und soziale Einbettung hinzuweisen. Kulturgeschichtlich »gehört das Schwimmen zu den immer wiederkehrenden Formen der politischen Körperkonologie« (Bredekamp, 2014, S. 17). Bildlich überliefert sind Schwimmszenierungen u.a. von Mao und Mussolini (vgl. Bredekamp, 2014, S. 13–17).⁸ In den USA wurde der Praxis des Schwimmens diskursiv Verantwortung für die Produktion genuin moderner Körper zugemutet (Stieglitz, 2018):

»Zugespitzt könnte man insbesondere im Zusammenhang mit dem Kraulschwimmen sagen, dass sich Diskurs und Praktiken hier zu einem in seinen intersektionalen Bezügen übergeordnet aufgeladenen Bewegungsmuster vereinigten, das man mit Doing Modernity umschreiben könnte« (Stieglitz, 2018, S. 77).

Der schwimmende moderne Körper war freilich – trotz des Inklusionspotenzials des Schwimmsports (aufgrund relativ geringer Einstiegskosten) – vor allem ein weißer Mittelklassekörper, die Geschichte des Schwimmens in den USA durch Exklusionspolitiken gekennzeichnet (Wiltse, 2014), deren lebensbedrohliche Konsequenzen man an Statistiken zu Ertrinkenden nachvollziehen kann (Hastings et al., 2016).

Ich werde im Folgenden vier Aspekte von Bewertungsspielen am Beispiel des Schwimmens hervorheben: Die Arbeit an der Grenze des Bewertungsspiels »Schwimmen«; die Regelgeleitetheit der Praxis; die Nutzung körperbezogener Kategorisierungen; und die Bewertungspraxis, die Schwimmer*innen im Wasser selbst ausüben, das »doing valuation while doing swimming«, wenn man so will.

⁸ Für Mussolini vgl. <https://youtu.be/O1T1y1LP6BA?si=Kx2ZWkZFw7lvOwyN> (ca. 01:00), für Mao https://youtu.be/xN1P2DHE26g?si=wG_cypgSh5EVCzUY (ca. 00:14). Den statistischen Systemwettbewerb gewinnt Ronald Reagan für die USA mit einem annehmbaren Kraul, s. <https://youtu.be/ivlpwlJ5v5g?si=H8R6rcFnrok-N1Z> (ca. 00:40).

4.1 Schwimmen oder Baden? Die Grenzen des Schwimmspiels

Grenzen sind konstitutiv für soziale Spiele – wo liegen die Grenzen des Schwimmspiels? Ist das »Passiv-Schwimmen« im Rhein bei Basel, das John von Düffel (2007, S. 20) literarisch beschreibt, Teil des Schwimmspiels? Wassergymnastik? Sicher ist, dass Schwimmer*innen an den Grenzen ihres Spiels arbeiten und das nicht nur auf die externe Grenze des Schwimmspiels bezogen, sondern auch intern. Es existieren unterschiedliche Welten des Schwimmens (Chambliss, 1989, S. 76) bzw. Schwimmstile. Die Theorie der Bewertungsspiele legt dabei nahe, dass die Grenzziehung durch Bewertungspraktiken vollzogen oder zumindest unterstützt wird.

Vom Beckenschwimmen unterschiedet sich das Freiwasserschwimmen schon durch den materiellen Kontext der Praxis, vor allem aber auch durch die Orientierungen, die sich die Schwimmenden wechselseitig unterstellen. Das wird zum Beispiel dann deutlich, wenn Bates und Moles (2022, S. 9) von Abgrenzungen von Freiwasserschwimmer*innen (»wild swimming«) von Triathlet*innen berichten, deren kompetitive Orientierung und Betonung von Leistung und Technik sie zu »[a]lmost another tribe altogether« machen würden. Das unterscheidet gemeine Freiwasserschwimmer*innen auch von Langstreckenschwimmer*innen bzw. »marathon swimmers« (Throsby, 2013).

»For example, there is no dispute that marathon swimming and organised events are sports, but when people dip and bob the practice becomes more informal and the rules, regulations and expectations less clear. Swimmers can invest in expensive kit, or they can make-do with the basics. They can be identified in the water by their swimming stroke, the line they take, the duration of their swim. These different ways of being in the water create different tribes, gathering families together under one roof and helping people come together through shared passions, while also sowing the seeds of division (Lay, 2022).« (Bates/Moles 2022: 2)

Ganz ähnlich zeigen Lenneis und Ko-Autorinnen (Lenneis et al., 2021) in ihrer Studie von »women only« Zeiten in Dänischen Schwimmbädern oder Caudwells (2021) Analyse von »private-hire swimming pool sessions« von Gruppen nicht-binärer und transsexueller Personen dass ganz eigene Spielformen des Aufenthalts menschlicher Körper im Schwimmbecken eingerichtet werden können.

Die Grenzziehungen werden durch Selbst- und Fremdbewertungen unterstützt. Eine dieser Bewertungen, in diesem Fall eine Abwertung des Spiels der anderen, bezieht sich auf die Welt des Beckenschwimmens. Bates und Moles (2022, S. 7) argumentieren zum Beispiel, dass »outdoor swimmers, [...] often decry pools as dead spaces, full of chemicals and lacking the romance, release and community of outdoor swimming (Griffiths, 2017). Rebecca Olive, (in press) describes pool swimming is ›Chlorinated, sanitised, regulated, bounded.‹« In den zitierten Beispielen werden die Grenzen des eigenen Schwimmspiels durch die Assoziiierung des Schwimmens der anderen mit negativ konnotierten Begriffen – der Pool ist begrenzend, sein Wasser nicht natürlich, sondern chemisch behandelt – gezogen. Im Gegensatz dazu – und wie man vermuten kann: in Kenntnis entsprechender Abwertungen – stehen affirmative Selbstbeschreibungen von Beckenschwimmer*innen als »chlorine junkie« (Scott, 2009, S. 125, 143). Auf den »Geschmack von Chlor« (Vivès, 2010) gekommen zu sein zeigt damit an, dass die Schwimmende am Bewertungsspiel des Beckenschwimmens teilnimmt, das sie in ihrer Schwimmpraxis konstituiert.

4.2 Die Regeln des Schwimmspiels

Wenn man nun Schwimmen als ein (intern in Teilspeile differenziertes) Spiel verstehen kann, dann stellt sich die Frage nach den Regeln, die die Praxis des Spiels orientieren. Soziologische Beobachterinnen des Schwimmens haben die Regelhaftigkeit dieses Bewertungsspiels auf der Grundlage autoethnographisch orientierter Beobachtungen für das Beckenschwimmen deutlich herausgearbeitet (Nixon, 1986; Scott, 2009). Howard Nixon (1986, S. 323) identifiziert neben der Erwartung, dass sich Körper auf Schwimmbahnen kreisförmig bewegen ein ganzes Set an Regeln des Schwimmbadaufenthalts:

»swim in the section where others tend to share your skill and intensity levels; if you rest or converse at the shallow end of the pool, stay out of the way of swimmers who are touching the wall to turn around; slower swimmers stay to the right and allow faster swimmers to pass on the left (in the center of the section); yield to faster swimmers who finish a lap and touch the wall at the same time as you; use strokes that do not interfere with the progress of others in the section; and do not start a lap just before someone else in your section is about to touch the wall for a turn.« (Nixon, 1986, S. 323)

Susie Scott (2009, S. 128) identifiziert drei Gruppen von Erwartungen, die an Schwimmer*innen gerichtet werden: »respect for personal space, respect for disciplinary regimes, and the desexualization of encounters.« Insbesondere der zweite Regelkomplex, der »respect for disciplinary regimes« erfordert dabei Praktiken der Selbst- und Fremdbewertung, da Schwimmer*innen ihre eigene Leistungsfähigkeit entsprechend einschätzen müssen. Ein nicht einfaches Unterfangen, da »because the speeds are defined relatively, there is some degree of ambiguity over what constitutes fast, slow, and especially medium swimming, and so the absolute speed of each lane is open to negotiation.« (Scott, 2009, S. 129)

Die von Nixon und Scott identifizierten Regeln beziehen sich vor allem auf die Regelung von Abläufen und die Koordination der schwimmenden Körper. Es handelt sich um Verfahrensregeln (Waibel et al., 2021, S. 41f.). Ontologien als weiterer Typ von Regeln der Bewertung (Waibel et al., 2021, S. 42f.) zeigen sich zum Beispiel in Formen der Anerkennung, die Scott bei Beckenschwimmerinnen beobachtet hat. Obwohl andere Schwimmende aufgrund des Desexualisierungsimperativs und eines »taboo on eye contact« (Scott, 2009, S. 130) nur zurückhaltend beobachtet werden können, ist die Beobachtung der Übereinstimmung ihrer Praxis mit allgemeineren Vorstellungen vom Sinn und Zweck des Schwimmens doch verbreitet und eine Grundlage der Bewertung von Personen. »Respect is [...] given to those who embody the official purpose of the swimming pool – fitness – and remind everyone of their common values.« (Scott, 2009, S. 130) Hier wird »fitness« als ein die Praxis des Schwimmens orientierendes Bewertungskriterium identifiziert, das Grundlage der Sinn- und Wertigkeitszuschreibungen der Spielenden ist.

Ganz deutlich wird die ontologische Dimension von Regeln des Schwimmspiels im Kontrast zweier literarischer Beschreibungen des Schwimmens. John v. Düffel (2007, S. 8f.), bekennender Freiwasserschwimmer, beschreibt den Wechsel von einem »tristen Hallenbadwinter auf den immergleichen Bahnen« zum Schwimmen in offenen Gewässern als eine besondere Erfahrung, es gebe »kaum etwas Beglückenderes als dieses freie, offene Wasser und seine unverbrauchte Frische«. Leanne Shapton (2012, S. 272), ehemalige Leistungsschwimmerin, vergleicht demgegenüber ihr »Wohlgefühl im Becken mit [ihrem] Unwohlgefühl im offenen Wasser« und zitiert die Reaktion eines ihrer ehemaligen Trainer, der, gefragt nach seinem Verhältnis zum offenen Wasser, unter anderem antwortet: »Was zur Hölle ist da unten?«. Diese Beschreibungen sind natürlich anekdotisch, sie verweisen aber dennoch darauf, dass innerhalb des Schwimmspiels über die Bewertungskriterien hinaus weitere

Vorstellungen über den Charakter der Welt und die Angemessenheit von Bewegungen in ihr das Schwimmen orientieren.

4.3 Kategorisieren und Bewerten

In der schönen Literatur und ihrer kulturwissenschaftlichen Reflexion wird dem Schwimmen ein besonderes Verhältnis zur Kategorisierung nachgesagt:

»Der Schwimmer bewegt sich wie die Literatur in einem Zwischenreich, in dem sich feste Kategorisierungen auflösen, eine Begegnung und Anverwandlung des Fremden sowie eine Auseinandersetzung mit dem Erinnerten möglich wird. [...] Sein Schwellenerlebnis bietet ihm eine wertvolle anthropologische Erkenntnismöglichkeit, die ihn im besten Falle mit den Grundlagen seiner eigenen Existenz bekannt machen und ihm alternative Existenzmöglichkeiten eröffnen kann [...].« (Bartl, 2009, S. 494)

In der sozialwissenschaftlichen Forschung zum Kategoriengebrauch im Schwimmen findet man dann tatsächlich auch Hinweise darauf, dass am Beckenrand manche Umweltkategorien eingeklammert werden können. Musto (2013, S. 368) zeigt zum Beispiel, dass jugendliche Wettkampfschwimmer*innen im Training bei der Beobachtung der Leistung ihrer Mitschwimmer*innen Geschlechterkategorien nicht aktualisieren:

»While racing »for times,« Sophia could have teased Nick and Jon for losing to »a girl,« or Jon could have told Lesley that »girls suck« at swimming. However, when coaches or athletes directly compared girls' and boys' performances during SVST practices, I *never* heard athletes use these comparisons as an opportunity to evoke antagonistic interactions.«

In ähnlicher Weise zeigt Grahn (2015) am Beispiel gemischter Trainingsgruppen schwedischer Leistungsschwimmer*innen, dass im Kontext des Schwimmsports im Alltagsleben verbreitete vergeschlechtlichte Körperideale kritisch reflektiert werden. Die Schwimmenden arbeiteten symbolisch insbesondere an der Inklusion von Muskeln in die »aesthetics of a female body« (Grahn, 2015, S. 9).⁹ Eine ganz andere Form der Umwertung gesellschaftlicher Körperfaktoren findet man beim »marathon swimming« –

⁹ Siehe aber die autoethnographischen Arbeiten von Jenny McMahon, die am Beispiel des Australischen Leistungsschwimmens von einer Ideologie des »slim to win« McMa-

dem Freiwasserschwimmens über lange Distanzen wie es zum Beispiel Kanalschwimmer*innen betreiben –, in denen Körperfett aufgrund seiner Isolierfunktion gewertschätzt wird (Throsby, 2013, S. 771). Wie Karen Throsby (2013) hervorhebt, wird die Aufhebung der negativen Valenz von »fatness« in diesem Kontext aber deutlich an seine Funktion im Kontext des Freiwasserschwimmens geknüpft und bleibt damit sowohl Ausnahme als auch temporär.

Die wertende Verwendung von Kategorien kann man abschließend auch an Praktiken nachvollziehen, in denen Schwimmer*innen die Kategorie »Geschlecht« relevant machen. In den autoethnographischen Beobachtungen der sozialen Ordnung des Schwimmbeckens von Susie Scott tritt die Figur des »hairy torpedo« auf, die eindeutig geschlechtsspezifisch typisiert wird: »typically a middle-aged man who charges up and down the lanes, tunnel visioned and sending people flying—is regarded with frustration and annoyance by other swimmers. ›Who does he think he is?‹ they murmur to each other. ›This is a public pool, not the Olympics!‹« (Scott, 2009, S. 138) Unklar bleibt bei dem Beispiel, ob es sich bei dem »hairy torpedo« um eine im Feld übliche Kategorie handelt, oder um eine der Forscherin, da sie in einem Auszug aus ihren Feldnotizen vorkommt (Scott, 2009, S. 141).

Die soziale Welt des Schwimmens zu einer Welt der Unschärfe und Kategorienlosigkeit zu erklären, schießt also über das Ziel hinaus. Auch wenn im Schwimmspiel Umweltkategorien nicht zwangsläufig aktualisiert werden müssen, finden sozialwissenschaftliche Studien zum Schwimmen doch eine vielfältige Kategorisierungspraxis vor. Darin liegt ein Aspekte der Grenzen von Bewertungsspielen, dass sie den Gebrauch von Kategorien nach eigenen Relevanzkriterien regulieren.

4.4 Bewertungen der Schwimmpraxis

Dass die Praxis des Schwimmens selbst, die Auseinandersetzung zwischen Schwimmer*in und Wasser, von unterschiedlichen Bewertungsproblemen durchzogen sein kann, kann man schließlich am Beispiel von Wettkampfschwimmer*innen sehen. Bewertungen beginnen hier mit dem ersten Sprung ins Wasser des Trainingspools, auf den eine Bewertung der Wassertemperatur folgt. Abweichungen von der Normaltemperatur, zu kaltes oder zu warmes

hon und McCannon (2017) spricht und damit verbundene Praktiken der Selbstbeobachtung und -quantifizierung (McMahon und Dinan-Thompson 2011) beschreibt.

Wasser, werden von expliziter Bewertungskommunikation begleitet: »If the water was felt to be especially cold, this was often conveyed to other swimmers via some form of expletive such as: ›Jesus Christ, that's cold!‹ or ›fuck me, that's fresh!‹« (McNarry et al., 2020, S. 214).

Dann bezieht sich die Bewertungspraxis aber vor allem auch auf die körperlichen Abläufe der Schwimmenden selbst. McNarry u.a. (2021, S. 4) heben hervor, dass zur Herstellung von Wettkampftauglichkeit nicht stumpfes Bahnenziehen, sondern »practical experimentation, discovery and the ability to adapt constantly to changes in the environment and the swimmer's own corporeality« erforderlich sind. Entsprechend sei die Praxis des Wettkampfschwimmens, das »doing competitive swimming« durch »constant monitoring and (re)evaluation of skilled movements and techniques« gekennzeichnet (McNarry et al., 2021, S. 4). Sie weisen damit darauf hin, dass Schwimmende in der Praxis des Schwimmens selbst vor Bewertungsprobleme gestellt werden können, die sie in ihrer Trainingspraxis zu bearbeiten anstreben. Voraussetzung dieser Bewertungspraxis scheint eine erhöhte Aufmerksamkeit zu sein, eine institutionalisierte Dauerreflexion auf Position und Bewegung des menschlichen Körpers im Medium des Wassers.

Damit kann man Praktiken der Bewertung auch im Kern der Praxis, die das Bewertungsspiel des Schwimmens als Spiel ausmacht, identifizieren. Dass es gerade Spielteilnehmer*innen sind, die den Großteil ihrer Zeit mit spielbezogenen Praktiken verbringen, an deren Beispiel man die Bewertungsdimension der Praxis selbst herausarbeiten kann, ist nicht verwunderlich. Es deutet darauf hin, dass die Involviertheit der Spielenden ins Spiel, ihr Professionalisierungsgrad bzw. das Ausmaß ihrer amateurialen Leidenschaft, sie mit unterschiedlichen Sensibilitäten für die Praxis und unterschiedlichen Bewertungskompetenzen ausstattet. Dass der Körper im Schwimmspiel als ein Bewertungsinstrument verwendet werden kann (vgl. Meier & Peetz, 2022, S. 280), setzt Übung voraus.

Schlussbemerkung

Wie ich zeigen konnte, sind Schwimmer*innen nicht nur den Bewertungen von »engaged spectators« (Hutter, 2021) – den Kampfrichter*innen, Trainer*innen, Sportjournalist*innen etc. – ausgesetzt. Bewertungen sind auch ein konstitutiver Teil der Praxis des Schwimmens: Der Produktion der Grenzen des eigenen Spiels wie auch der Auseinandersetzung mit dem Wasser.

Auffallend ist, dass die Position der Forschenden im Schwimmspiel einen Einfluss auf die untersuchten Bewertungsspiele zu haben scheint. Während Schwimmer*innen, die ihre Freizeitbeschäftigung zum Forschungsgegenstand gemacht haben, die internen Abgrenzungspraktiken des Schwimmspiels thematisieren, kommt die Bewertungspraxis des Schwimmens, die Suche nach dem Gefühl für das Wasser, in der Arbeit eines ehemals professionellen Schwimmers in den Blick. Im Anschluss an diese Beobachtung könnte man vermuten, dass Amateure – Nicht-Professionelle, die die Fähigkeit entwickelt haben, ihre eigene Praxis reflexiv zu beobachten (Hennion, 2005, S. 136) – eine privilegierte Beobachtungsposition einnehmen könnten. Die Praxis der Selbstamateurisierung könnte einen Zugang zu einer integrierten Perspektive auf das Schwimmen als Bewertungsspiel eröffnen. Ob dies auch für andere Bewertungsspiele gilt, wäre zu prüfen.

Für die Bewertungssoziologie des Sports bzw. die bewertungsinteressierte Sportsoziologie zeigt der Fall des Schwimmens, dass es lohnt, Praktiken der Bewertung auch jenseits der großen Publikumssportarten und selbst in Sportarten mit eindeutiger Leistungsbewertung nachzuzeichnen. Vielleicht sind es gerade Sportarten, die – anders als der Fußball (siehe Einleitung) – weniger massenmedial beobachtet werden und vom Publikum geringere Aufmerksamkeit erfahren, die es erlauben, den Blick von Bewertungen *über das Spiel* und die Spielenden zu Bewertungen *in der Spielpraxis* zu richten. Ein Perspektivwechsel auf Bewertungen *in der Praxis* unterschiedlicher Spiele ist schließlich auch die Empfehlung, die von hier ausgehend der Soziologie der Bewertung gegeben werden kann. Neben der Untersuchung von expliziter Bewertungskommunikation, von Preisvergaben und Preisauszeichnungen, Peer Reviews und Rankings, Konsekrations- und Kanonisation ist die Aktualisierung von Wertsinn in der alltäglichen Praxis ein Forschungsfeld, dass es noch zu erschließen gilt.

Literatur

- Allen, M. P., & Parsons, N. L. (2006). The institutionalization of fame: Achievement, recognition, and cultural consecration in Baseball. *American Sociological Review*, 71, 808–825.
- Appadurai, A. (1986). Introduction: Commodities and the politics of value. In A. Appadurai (Hg.), *The social life of things* (S. 3–63). Cambridge University Press.

- Bartl, A. (2009). »Der Wechsel von einem vertrauten Element in das andere, fremde«: Das Schwimm-Motiv in der deutschen Gegenwartsliteratur. *German Life and Letters*, 62(4), 482–495.
- Bates, C., & Moles, K. (2022). Bobbing in the park: Wild swimming, conviviality and belonging. *Leisure Studies*, 1–13. <https://doi.org/10.1080/02614367.2022.2085774>
- Bateson, G. (1955/1972). A theory of play and fantasy. In G. Bateson, *Steps to an ecology of mind* (S. 177–193). Ballantine.
- Boltanski, L., & Esquerre, A. (2018). *Bereicherung. Eine Kritik der Ware*. Suhrkamp.
- Bourdieu, P., & Wacquant, L. J. D. (2006). *Reflexive Anthropologie*. Suhrkamp.
- Bredenkamp, H. (2014). *Der schwimmende Souverän. Karl der Große und die Bildpolitik des Körpers*. Wagenbach.
- Brümmer, K. (2015). *Mitspielfähigkeit. Sportliches Training als formative Praxis*. transcript.
- Brümmer, K. (2021). Lernen am Video – Analysemedien im (spitzen-)sportlichen Training. In K. Brümmer, A. Janetzko, & T. Alkemeyer (Hg.), *Ansätze einer Kultursoziologie des Sports* (S. 249–268). Nomos. <https://doi.org/10.5771/9783845288796-249>
- Caillois, R. (1958/2017). *Die Spiele und die Menschen. Maske und Rausch*. Matthes & Seitz.
- Caudwell, J. (2021). Queering indoor swimming in the UK: Transgender and non-binary wellbeing. *Journal of Sport and Social Issues*, 46(4), 338–362. <https://doi.org/10.1177/01937235211043648>
- Chambliss, D. F. (1988). *The making of olympic swimmers*. William Morrow.
- Chambliss, D. F. (1989). The mundanity of excellence: An ethnographic report on stratification and olympic swimmers. *Sociological Theory*, 7(1), 70–86.
- Dewey, J. (1939). Theory of valuation. In O. Neurath (Hg.), *International Encyclopedia of Unified Science* (Bd. II(4), S. 1–67). University of Chicago Press.
- Düffel, J. v. (2007). *Schwimmen* (4. Aufl.). dtv.
- Durkheim, E., & Mauss, M. (1993). Über einige primitive Formen von Klassifikation. Ein Beitrag zur Erforschung der kollektiven Vorstellungen. In E. Durkheim, *Schriften zur Soziologie der Erkenntnis* (S. 169–256). Suhrkamp.
- Frye, M. (2017). Cultural meanings and the aggregation of actions: The case of sex and schooling in Malawi. *American Sociological Review*, 82(5), 945–976. h <https://doi.org/10.1177/0003122417720466>

- Goffman, E. (1972). Fun in games. In E. Goffman, Encounters (S. 15–72). Penguin Books.
- Grahn, K. (2015). Gendered body ideals in Swedish competitive youth swimming: Negotiating and shifting symbolic boundaries. *Sport in Society*, 19(5), 680–694. <https://doi.org/10.1080/17430437.2015.1073947>
- Hamann, J. (2016). Peer Review post mortem. Bewertungen in akademischen Nachrufen. *Berliner Journal für Soziologie*, 26(3–4), 433–457.
- Hastings, D. W., Zahran, S., & Cable, S. (2016). Drowning in Inequalities. *Journal of Black Studies*, 36(6), 894–917. <https://doi.org/10.1177/0021934705283903>
- Heintz, B. (2016). »Wir leben im Zeitalter der Vergleichung.« Perspektiven einer Soziologie des Vergleichs. *Zeitschrift für Soziologie*, 45(5), 305–323.
- Hennion, A. (2005). Pragmatics of taste. In M. D. Jacobs & N. W. Hanrahan (Hg.), *The Blackwell companion to the sociology of culture* (S. 131–144). Blackwell.
- Hirschauer, S. (2014). Un/doing Differences. Die Kontingenz sozialer Zugehörigkeiten. *Zeitschrift für Soziologie*, 43(3), 170–191.
- Hodek, F. (2021). Von der Handnotation zum ›Live-Scouting‹ am Bildschirm. Quantifizierung (Be)Wertung und Medialität im Profifußball. In O. Berli, S. Nicolae, & H. Schäfer (Hg.), *Bewertungskulturen* (S. 165–194). Springer VS.
- Huizinga, J. (1938/2017). *Homoludens. Vom Ursprung der Kultur im Spiel* (25. Aufl.). Rowohlt.
- Hutter, M. (2015). *Ernste Spiele. Geschichten vom Aufstieg des ästhetischen Kapitalismus*. Fink.
- Hutter, M. (2021). Three modes of valuation practices in art games. *Valuation Studies*, 8(1), 85–119.
- Janetzko, A. (2021). (Be)Wertungen sportlichen Talents – Praktiken und ihre Zusammenhänge. In O. Berli, S. Nicolae, & H. Schäfer (Hg.), *Bewertungskulturen* (S. 139–164). Springer VS.
- Krause, M. (2021). *Model cases. On canonical research objects and sites*. University of Chicago Press.
- Latour, B. (2013). *An inquiry into modes of existence. An anthropology of the moderns*. Harvard University Press.
- Lenneis, V., Evans, A. B., & Agergaard, S. (2021). Swimming as self-care – A Foucauldian analysis of swimming for Danish Muslim women. *International Review for the Sociology of Sport*, 57(3), 401–420. <https://doi.org/10.1177/10126902211016843>
- Lom, S. E. (2016). Changing rules, changing practices: The direct and indirect effects of tight coupling in figure skating. *Organization Science*, 27(1), 36–52.

- Luhmann, N. (1971). Sinn als Grundbegriff der Soziologie. In N. Luhmann & J. Habermas, *Theorie der Gesellschaft oder Sozialtechnologie – Was leistet die Systemforschung?* (S. 25–100). Suhrkamp.
- Mauss, M. (1935/1978). Die Techniken des Körpers. In M. Mauss, Soziologie und Anthropologie (S. 197–220). Ullstein.
- McNarry, G., Allen-Collinson, J., & Evans, A. B. (2020). Sensory sociological phenomenology, somatic learning and ›lived‹ temperature in competitive pool swimming. *The Sociological Review*, 69(1), 206–222. <https://doi.org/10.1177/0038026120915149>
- McNarry, G., Allen-Collinson, J., & Evans, A. B. (2021). ›Doing‹ competitive swimming: Exploring the skilled practices of the competitive swimming lifeworld. *International Review for the Sociology of Sport*, 56(1), 3–19.
- Meier, F., & Peetz, T. (2022). Der stigmatisierte Körper. Bewertungspraktiken und -konstellationen im Fall »Konnersreuth«. *Österreichische Zeitschrift für Soziologie*, 47(3), 269–290. <https://doi.org/10.1007/s11614-022-00499-8>
- Musto, M. (2013). Athletes in the pool, girls and boys on deck. *Gender & Society*, 28(3), 359–380. <https://doi.org/10.1177/0891243213515945>
- Neumann, J. v. (1928). Zur Theorie der Gesellschaftsspiele. *Mathematische Annalen*, 100(1), 295–320.
- Nixon, H. L. (1986). Social order in a leisure setting: The case of recreational swimmers in a pool. *Sociology of Sport Journal*, 3(4), 320–332.
- Ortmann, G. (2003). *Regel und Ausnahme. Paradoxien sozialer Ordnung*. Suhrkamp.
- Peetz, T. (2020). Der Sinn der Ökonomisierung. *Sozialer Sinn*, 21(1), 1–23.
- Peetz, T. (2021). Elemente einer Soziologie der Bewertung. In F. Meier & T. Peetz (Hg.), *Organisation und Bewertung* (S. 25–47). Springer VS.
- Peetz, T. (2022). Theorie der Bewertungsspiele: Digitale intime Bewertung. In K. Block, A. Deremetz, A. Henkel, & M. Rehbein (Hg.), *10 Minuten Soziologie: Digitalisierung* (S. 75–86). transcript.
- Peetz, T. (2023). Intimate valuation devices. Doing valuation while doing dating in Tinder texts. *Journal of Cultural Economy*. <https://doi.org/10.1080/17530350.2023.2246987>
- Puhl, K. (2002). Die List der Regel. Zur retroaktiven Konstitution sozialer Praxis. In U. Baltzer & G. Schönrich (Hg.), *Institutionen und Regelfolgen* (S. 81–99). mentis.
- Puhl, K. (2010). Regelfolgen. In E. v. Savigny (Hg.), *Ludwig Wittgenstein: Philosophische Untersuchungen* (S. 119–142). Akademie Verlag.

- Rosch, E. (1978). Principles of categorization. In E. Margolis & S. Laurence (Hg.), *Concepts. Core readings* (S. 189–206). MIT Press.
- Schäfer, H. (2013). *Die Instabilität der Praxis. Reproduktion und Transformation des Sozialen in der Praxistheorie*. Velbrück.
- Schäfer, H. (2016). »Outstanding universal value«. Die Arbeit an der Universalisierung des Wertvollen im UNESCO-Welterbe. *Berliner Journal für Soziologie*, 26(3/4), 353–375.
- Schmidt, R., Weigelin, M., Brümmer, K., Laube, S., & Schäfer, H. (2022). Bodies and embodiment in practices of valuation: Challenging the sociology of valuation with the sociology of the body. *Österreichische Zeitschrift für Soziologie*, 47(3), 217–223. <https://doi.org/10.1007/s11614-022-00502-2>
- Scott, S. (2009). Re-clothing the emperor: The swimming pool as a negotiated order. *Symbolic Interaction*, 32(2), 123–145. <https://doi.org/10.1525/si.2009.32.2.123>
- Shapton, L. (2012). *Bahnen ziehen*. Suhrkamp.
- Star, S. L., & Griesemer, J. R. (1989). Institutional ecology, ›translations‹ and boundary objects: Amateurs and professionals in Berkleys museum of vertebrate zoology, 1907–1939. *Social Studies of Science*, 19, 387–420.
- Stieglitz, O. (2018). The American Crawl – Praktiken von Geschlecht und Moderne in US-amerikanischen Schwimmbecken, 1900–1940. *GENDER – Zeitschrift für Geschlecht, Kultur und Gesellschaft*, 10(1), 63–80. <https://doi.org/10.3224/gender.v10i1.05>
- Throsby, K. (2013). ›You can't be too vain to gain if you want to swim the Channel: Marathon swimming and the construction of heroic fatness. *International Review for the Sociology of Sport*, 50(7), 769–784. <https://doi.org/10.1177/1012690213494080>
- Vivès, B. (2010). *Der Geschmack von Chlor*. Reproduct.
- Waibel, D., Peetz, T., & Meier, F. (2021). Valuation constellations. *Valuation Studies*, 8(1), 33–66. <https://doi.org/10.3384/VS.2001-5992.2021.8.1.33-66>
- Weber, M. (1904/1988). Die ›Objektivität‹ sozialwissenschaftlicher und sozialpolitischer Erkenntnis. In M. Weber, *Gesammelte Aufsätze zur Wissenschaftslehre* (7. Aufl., S. 146–214). Mohr.
- Weigelin, M. (2022). Entscheidungen und ihre Bewertungen – Zur Mikrosoziologie des Schiedsrichter-Pfiffs. *Österreichische Zeitschrift für Soziologie*, 47(3), 225–246. <https://doi.org/10.1007/s11614-022-00500-4>
- Wiltse, J. (2014). The black–white swimming disparity in America. *Journal of Sport and Social Issues*, 38(4), 366–389. <https://doi.org/10.1177/0193723513520553>

Wittgenstein, L. (1953/2003). *Philosophische Untersuchungen*. Suhrkamp.

Vergleichen, Differenzieren, Kategorisieren und das Problem der Fairness

(Sport-)Unterricht als instruktiver Fall einer Soziologie der Bewertung

Samira Mummelthey & Kerstin Rabenstein

Einleitung¹

Wir beschäftigen uns im Folgenden mit auf sportliche Leistung bezogenen Bewertungen im schulischen Sportunterricht. Dabei gehen wir für unseren Fall (Sport-)Unterricht sowohl von einer Verunterrichtlichung von Sport als auch einer Versportlichung von Unterricht aus. Das heißt, wir rechnen aus praxistheoretischer Perspektive sowohl mit eigenlogischen Formen der Bewertung im (außerschulischen) Sport, etwa bestimmten Vergleichs- und Kategorisierungspraktiken, als auch mit eigenlogischen Formen der Bewertung im Unterricht, die im Vollzug des Sportunterrichts miteinander amalgamieren. So lässt sich für den Sportunterricht beispielsweise beobachten, dass mit von im außerunterrichtlichen Sport (bislang) als legitim geltenden Kategorisierungen nach Geschlecht auf körperliche Dispositionen geschlossen wird (Wilm, 2021; Krämer in diesem Band), die mit sportlicher Leistung in einen Zusammenhang gebracht werden, und im Sportunterricht zu Praktiken des Differenzierens, Kategorisierens und Bewertern führen, die so in anderen Schulfächern nicht zu beobachten sind. Wir fragen insgesamt danach, wie sich sportbezogene und unterrichtliche Logiken in Praktiken des Differenzierens, Kategorisie-

¹ Wir danken Robert Schmidt, David Kempf und Max Weigelin herzlich für die Einladung zur Mitwirkung an diesem Band und ihre äußerst wertvollen und weiterführenden Anregungen zum ersten Entwurf unserer Argumentation sowie allen Autor:innen des Bandes für hilfreiche und konstruktive Anmerkungen.

rens und des Vergleichens im Sportunterricht verbinden und was dabei prekär zu werden scheint.

Mit diesem Anliegen begegnen wir gleich einem dreifachen Desiderat: Schulische Bewertungsprozesse werden im Vergleich zu der in den letzten 10 Jahren sich entwickelnden, äußerst umfangreichen ethnografischen Forschung zu verschiedenen Verfahren und Settings unterrichtlicher Bewertungen (Bräu & Fuhrmann, 2015; Kalthoff, 1996; Meister & Hollstein, 2018; Rabenstein, 2021; Rabenstein et al., 2015; Reh & Ricken, 2018; Strauß, 2023; Zaborowski et al., 2011) innerhalb einer Soziologie der Bewertung noch wenig thematisiert (Hirschauer, 2019; Kalthoff, 2019; Kalthoff & Dittrich, 2016; F. Meier et al., 2016). Umgekehrt nimmt auch die umfangreiche bildungsssoziologische und erziehungswissenschaftliche ethnografische Unterrichtsforschung zu Differenzierungen, Kategorisierungen und/oder Bewertungen noch kaum Bezug auf bewertungssoziologische Ansätze. Drittens – und dies erscheint uns für diesen Band in besonderer Weise relevant – spielt der Sportunterricht auch dort, wo unterrichtliche Bewertungen ethnografisch beobachtet werden, so gut wie keine Rolle. Anders gesagt, Bewertungsverfahren im Sportunterricht werden durch das häufige Ausklammern von Sportunterricht auch in der sich für Differenzierungen, Kategorisierungen und Bewertungen interessierenden Unterrichtsforschung noch kaum berücksichtigt. Während einerseits die Frage schulform- und schulstufenspezifischer Bewertungskulturen (Gellert & Hümmer, 2008; Zaborowski et al., 2011) und andererseits Fragen der Bewertung im Unterricht mit und ohne Noten (Kalthoff & Dittrich, 2016; Rabenstein et al., 2015; Strauß, 2023) behandelt werden, wird noch kaum gefragt, inwiefern sich Bewertungspraktiken in fachunterrichtlichen Kontexten unterscheiden (Grein et al., 2024 zu Französisch- und Spanischunterricht; Heberle, 2019 zu Musikunterricht).

Ausgehend von der Verwobenheit von Praktiken der Leistungsbewertung mit Praktiken des Vergleichens und Kategorisierens fragen wir in diesem Beitrag danach, wie im Sportunterricht und im Sprechen von Lehrkräften über Sportunterricht an der Herstellung von Vergleichbarkeit zwischen den Schüler:innen gearbeitet wird und wie dabei welche Differenzierungen und Kategorisierungen vorgenommen werden (Mummelthey, 2024; zur methodologischen Ausrichtung Mummelthey et al., 2023)². Dabei zeigte sich in den Aus-

² Die Studie ist von 2020 bis 2022 im Handlungsbereich C>Diversität gerecht werden< im Schröder Programm Lehrerbildung der Universität Göttingen entstanden, das im Rahmen der gemeinsamen »Qualitätsoffensive Lehrerbildung« von Bund und Ländern mit

handlungen von Vergleichbarkeit immer wieder ein Bezug zur Norm der Fairness, die sich sowohl schulisch wie sportbezogen interpretieren lässt.

Wir gehen wie folgt vor: Zunächst nehmen wir eine Kontextualisierung unseres Vorhabens vor, indem wir erziehungswissenschaftliche und bewertungssoziologische Thematisierungen von Praktiken des Vergleichens und Bewertens aufeinander beziehen. Im nächsten Schritt legen wir methodisch-methodologische Prämissen der diskursethnografischen Studie offen, aus der wir empirisches Material verwenden, um zu zeigen, wie an der Herstellung von Vergleichbarkeit zwischen den Schüler:innen gearbeitet wird. Abschließend resümieren wir den Ertrag unserer Analysen im Sinne einer Betrachtung von (Sport-)Unterricht als einen instruktiven Fall einer ‚Soziologie der Bewertung‘.

1. Kontextualisierungen: Bewertungen im schulischen (Sport-)Unterricht in bewertungssoziologischer und erziehungswissenschaftlicher Perspektive

Die bewertungssoziologische Forschung speist ihre Relevanz u.a. aus den sozialen Praktiken, die mit der Neo-Liberalisierung des Arbeitsalltags der 1990er Jahre sowie der seitdem zunehmend Fahrt aufnehmenden Digitalisierung gesellschaftlicher Prozesse in Verbindung gebracht werden (Berli et al., 2021; Kropf & Laser, 2019; A. K. Krüger, 2022; F. Meier & Peetz, 2021). Praktiken des Bewertens und Bewertet-Werdens sind demnach mittlerweile allgegenwärtig im Alltag (post-)moderner Gesellschaften. Zwar trifft dies für Bewertungen im schulischen Unterricht schon länger zu (Berdemann et al., 2018; Kalthoff, 1996; Zaborowski et al., 2011), im Zuge der *New Governance* ist jedoch von einer Dynamisierung schulischer Bewertungspraktiken auszugehen (z.B. Bellmann et al., 2016). So werden beispielsweise Lehrkräfte seit einiger Zeit angehalten, ihren Unterricht durch Feedback von Schüler:innen bewerten zu lassen (z.B. Kuhlmann & Sotzek, 2019). Eltern bewerten angesichts der eröffneten Schulwahlmöglichkeiten verstärkt Schulen, bevor sie sich für eine Schule für ihr Kind entscheiden (z.B. J. O. Krüger, 2014). Bewertungsmaßstäbe in Bezug auf die Kompetenzentwicklung zukünftiger Lehrkräfte sollen durch mehr

Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen 01JA1917 gefördert wurde. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autorinnen.

Kohärenz zwischen der ersten und zweiten Phase der Lehrer:innenausbildung schlagkräftiger gemacht werden (z.B. Hellmann et al., 2019). Diese Liste von Beispielen ließe sich leicht fortführen. Vor dem Hintergrund, dass diese Dynamisierung von Bewertungspraktiken in Schule, Unterricht bzw. Lehrkräftebildung in dem sich interdisziplinär verstehenden Forschungsfeld einer Soziologie der Bewertung (A. K. Krüger & Reinhart, 2016) noch eher selten untersucht wird (Kalthoff, 2019; Kalthoff & Dittrich, 2016, 2017), skizzieren wir im Folgenden für unsere Analysen des Sportunterrichts produktive Anschlüsse zwischen erziehungswissenschaftlichen und bewertungssoziologischen Perspektivierungen von Bewertungspraktiken.

Angesichts der schulpädagogischen und -didaktischen Dominanz eines psychologischen Verständnisses unterrichtlicher Bewertung von Leistung als Disposition von Individuen ist es ein wesentliches Anliegen ethnografischer Unterrichtsforschung, unterrichtliche Bewertungsprozesse als soziales Phänomen zu theoretisieren und in ihrem Vollzug zu analysieren. Das Interesse gilt damit der sozialen Konstruktion von Leistung. Die bewertungssoziologisch vorgenommenen begrifflichen Differenzierungen können indes hilfreich für die ethnografische Beobachtung unterrichtlicher Bewertungsprozesse sein. Dort, wo es uns zielführend erscheint, weisen wir auf die sich in der sportpädagogischen Literatur andeutenden Besonderheiten des Sportunterrichts hin, die aus bewertungssoziologischer Perspektive zu erforschen gleichzeitig Anliegen der hier dargestellten Studie ist. Zwei thematische Fokussierungen stellen wir dafür im Folgenden heraus:

Bewertungssoziologisch formuliert, wird *erstens* für unsere Untersuchung nicht nur deutlich, dass Differenzierungen, Kategorisierungen und Vergleiche konstitutiv für Schule und Unterricht sind, sondern auch, dass mit Bewertungen immer auch Vergleiche vollzogen werden, Vergleiche und Bewertungen also konstitutiv aufeinander bezogen sind (A. K. Krüger & Reinhart, 2016, S. 494 nach Heintz 2010, 2016). Der bewertungssoziologisch vorgeschlagenen Ausdifferenzierung von Praktiken des Vergleichens, Bewerterns, Kategorisierens zu folgen, scheint vielversprechend, um die (In-)Stabilität der sportunterrichtlichen Leistungsordnung und aktuelle (sport-)pädagogische Diskussionen um Leistungsbewertung im inklusiven Unterricht zu verstehen. Um empirisch die verschiedenen Praktiken des Vergleichens, Bewerterns, Kategorisierens (und Quantifizierens) in ihrer Verwobenheit miteinander zu betrachten, nehmen wir im Folgenden die Überlegungen von Bettina Heintz (2021) auf.

Heintz versteht Vergleiche als »eine Grundoperation des sozialen Geschehens« (Heintz, 2016, S. 306) und konzeptualisiert diese als »Beobach-

tungsinstrumente, die zwischen Einheiten oder Ereignissen eine Beziehung herstellen« (Heintz, 2010, S. 164). Etwas zu kategorisieren, ist zentrales Element von Vergleichen, gleichzeitig bedarf eine Kategorisierung eines vorangehenden Vergleichs (Heintz, 2016, 2021). Das Ziel von Vergleichen ist es dementsprechend etwas/jemanden mithilfe einer dritten Größe »auf ihre Unterschiede und Ähnlichkeiten hin zu beobachten« und so zu relationieren (Heintz, 2016, S. 307). Um dies zu tun, müssen die zu Vergleichenden zwei Merkmale erfüllen: Sie müssen als vergleichbar, also der gleichen Kategorie zuordbar hervorgebracht werden, und es müssen Vergleichskriterien und Verfahren zur Verfügung stehen, um Differenzen ermitteln zu können. Diese Verbindung von »*Gleichheitsunterstellung*« und »*Differenzbeobachtung*« ist dabei als spezifisches Merkmal von Vergleichen zu verstehen (Heintz, 2016, S. 307).

So gesehen, lässt sich auch der Einbezug aller Heranwachsenden in den schulischen Bildungsanspruch im 19. Jahrhundert – im Sinne Stichwehs (2016) als Inklusion zu bezeichnen – als ein Vergleichbar-Machen aller Kinder und Jugendlichen verstehen.

Im Zuge der Durchsetzung der Schulpflicht und des Ausbaus des öffentlichen Schulwesens wurde überhaupt erst die Bandbreite der unterschiedlich inklusionswilligen und inklusionsfähigen Schüler*innen sichtbar, die als Gruppen identifiziert und zusammengefasst auf unterschiedliche Plätze im sich entwickelnden Schulsystem verteilt werden sollten (Reh et al., 2021, S. 9).

Erst diese Entwicklungen führten dazu, dass im Unterricht »eigene neue« (Reh et al., 2021, S. 9) Probleme entstanden, auf die wiederum (inner-)schulisch mit Vergleichen, Differenzierungen und Kategorisierungen reagiert wurde (Reh et al., 2021). Inklusion als gesellschaftspolitische Aufgabe wurde also »schulisch «umgebaut», »umformuliert« (Reh et al., 2021, S. 11); sie »erzeugte schließlich im Schulwesen eine pädagogisch legitimierte Differenzierung der Schülerschaft, nämlich in gut oder weniger gut integrierbare Gruppen von Schüler*innen« (Reh et al., 2021, S. 11). Die Einführung der Jahrgangsklasse als ein sozialer Verbund von alters- und damit (so die Annahme) entwicklungs-gleichen Heranwachsenden war – im Unterschied zum sogenannten System der Fachklassen – ein zentrales Moment für die Entstehung von Praktiken des Vergleichens von Schüler:innen und ihren Leistungen. Die Schule leistete damit im 19. Jahrhundert einen Beitrag zur Entstehung des Leistungsdispositivs (Berdelmann et al., 2018). Darin zeigt sich auch empirisch, inwiefern

die Herstellung von Gleichheit und damit Vergleichbarkeit als konstitutives Element von Praktiken der Differenzierung zu verstehen ist.

Wie Heintz (2010) fasst auch Riepe (2021) dieses Verhältnis von Vergleichen und Differenzieren für die Schule in der Gegenwart, wie folgt:

Bevor Abstände gemessen, Niveaus bestimmt, Besonderheiten fixiert und somit Unterschiede benannt und zueinander in ein (hierarchisiertes) Verhältnis gesetzt werden können, besteht jedoch erst einmal die Notwendigkeit der Herstellung eines »gleichen« Bezugspunktes – einer Norm, von der ausgehend überhaupt erst differenziert werden kann (Riepe, 2021, S. 14).

Riepe fordert dementsprechend *Doing-Difference* im Zusammenhang mit »*doing-sameness*« zu analysieren (Riepe, 2021, S. 14). Dies bedeutet aber auch, dass nie alle vermeintlichen Eigenschaften eines Sachverhalts in einen Vergleich eingehen – vielmehr werden bestimmte Merkmale dadurch auch unsichtbar gemacht (Heintz, 2010, 2016).

Mit der gesetzlichen Verankerung inklusiven Unterrichts seit 2009³ ist für die (sport-)pädagogische Diskussion die Frage nach Verfahren der Leistungsbewertung in einem formal inklusiv gerahmten Unterricht virulent, der für sich in Anspruch nimmt Diversitätssensibel auf Unterschiede von Lernenden zu reagieren. Auf ordinalen Rangfolgen beruhende Bewertungen werden seitdem im Sinne von (un-)passenden, ableistischen Leistungsanforderungen (sport-)pädagogisch zunehmend kritisiert (S. Meier, 2023; S. Meier et al., 2016; S. Meier & Ruin, 2015; Ruin, 2017; Schiller et al., 2021). Bewertungssoziologisch lässt sich davon sprechen, dass ein besser/schlechter Vergleich von Personen – der bis dato in der schulischen Leistungsbewertung »übliche« Vergleichsmodus –, der sich auf individuelle Merkmale bezieht (Heintz, 2021, S. 12), in die Kritik geraten ist. Mit der gesellschaftlichen Bedeutungszunahme dieser Vergleichsform und der damit einhergehenden Vorstellung einer grundsätzlichen Vergleichbarkeit beliebiger Personen entwickelte sich auch eine Kritik daran. Die (sport-)pädagogisch schon lange tradierte Kritik an ordinalen Rangfolgen erhält vor dem Hintergrund reformpädagogischer Entwürfe inklusiven Unterrichts in den letzten 10 Jahren neuen Aufwind. Eine neuere Form des Vergleichs, die Heintz mit Bezug auf Steinmetz als »simply-different-Vergleiche« bezeichnet (Heintz, 2021, S. 12), setzt keine kategoriale

³ Deutschland hat sich 2009 durch die Ratifizierung der UN-BRK zu einem inklusiven Schulsystem verpflichtet (Monitor Lehrerbildung, 2022).

Einheitlichkeit von Individuen voraus, sondern plädiert im Gegenteil für die Annahme einer Unvergleichbarkeit dieser. Diese Forderung findet sich auch in (sport-)pädagogischen Vorschlägen für inklusiven Unterricht. Menschen würden sich demnach zwar unterscheiden, aber nicht vergleichen lassen (Heintz, 2021).

Aber auch Behauptungen einer grundsätzlichen Andersartigkeit beruhen letztlich auf einem Vergleich, denn anders oder sogar unvergleichbar kann man nur im Vergleich zu anderen sein (Heintz, 2021, S. 13).

Andere Stimmen ordnen diesen Konflikt zwischen nominalen und ordinalen Vergleichen als Dilemma moderner Gesellschaften ein: Zum einen sollen als benachteiligt markierten Gruppen besondere Rechte eingeräumt werden, zum anderen ist das Selbstbild einer meritokratischen Gesellschaft daran gebunden, Personen allein daran zu beurteilen, was als Leistung verstanden wird. Das Dilemma ist nun, dass diese Ansprüche sich nicht gleichzeitig realisieren lassen (vgl. Fourcade 2016 nach Heintz, 2021, S. 19). Insofern beide Vergleichsformen einander ausschließen, kann dies analog als eine trilemmatische Struktur des Inklusionsdiskurses gefasst werden (Boger, 2017).

Eine zweite thematische Fokussierung knüpft daran an, dass unterrichtliche Bewertungen durch Ziffernnoten (sport-)pädagogisch schon lange wiederkehrend problematisiert werden (Schröder, 2024). Die Legitimität von Ziffernnoten scheint dabei in besonderer Weise umstritten: So wird seit Jahrzehnten wiederkehrend Zweifel an der Objektivität, der Gültigkeit und der Sinnhaftigkeit der Notengebung formuliert (Ingenkamp, 1971). Unterrichtliche Bewertungen durch Ziffernnoten werden (mittlerweile) auch im Unterricht angefochten (Rotter & Bressler, 2022). So betont Feth (2014) in ihrer Interviewstudie mit Sportlehrkräften, dass es sich aus deren Perspektive bei der Notengebung deshalb um ein komplexes »Handlungsproblem« (Feth, 2014, S. 53) handelt, weil Noten auch vor der Klasse legitimiert werden müssten. Bewertungssoziologisch interessant ist, wie schulisch an der Legitimität von Bewertungen gearbeitet wird. Praxistheoretische, an Michel Foucault angelehnte Arbeiten zeigen auf, wie Zweifel an der Legitimität von Noten über Strategien ihrer Objektivierung, wie zum Beispiel der Quantifizierung von Bewertungen über eine Punktevergabe, und ihrer Subjektivierung, wie zum Beispiel der pädagogischen Kommentierung von Bewertungen zur Herstellung subjektiver Relevanzen, bearbeitet werden (Breidenstein & Thompson, 2014; auch Kalthoff & Dittrich, 2016). In den letzten 20 Jahren ist

dabei eine Dynamisierung von (sport-)unterrichtlichen Bewertungspraktiken zu beobachten, die auf eine neue Allianz, wie Bellmann und Waldow (2007) formulieren, eine »merkwürdige Ehe« aus reformpädagogischen Entwürfen selbstregulierten Lernens und einer technokratischen Form der Schulreform zurückzuführen ist, deren Legitimierung sich durch die Allianz wechselseitig verstärkte. In Folge der anhaltenden Problematisierung von Ziffernnoten wird mittlerweile (sport-)pädagogisch breit für die Substitution durch sogenannte alternative Formate der Leistungsbewertung plädiert. Reformpädagogische Begründungsmuster für einen Verzicht auf Noten schlagen sich auch in praxisinstructiven Beiträgen der Sportpädagogik nieder (Balz, 2009; Scherler, 2000; Tillmann, 2001; Wolters & Lüsebrink, 2017), mit denen u.a. vorgegeben wird, Schüler:innen besser motivieren zu können. Mit Verschiebungen in den Verfahren der Bewertung wird nicht nur die Legitimität der Bewertungen zu stützen versucht, bewertungssoziologisch betrachtet, verändert sich auch die Konstellation aus Bewertetem, Bewertenden und Publikum (F. Meier et al., 2016).

Dies zeigt sich beispielsweise dort, wo Verfahren, in denen Schüler:innen aufgefordert sind, die eigene Leistung zu bewerten, auch in den (Sport-)Unterricht einsickern, in dem weiterhin Ziffernnoten vergeben werden. So ist über alle Schulstufen und Schulformen hinweg im Unterricht mittlerweile zu beobachten, wie nach Präsentationen Mitschüler:innen Feedback geben und wie das ›richtige‹ Feedbackgeben der Beurteilung durch Lehrkräfte unterliegt und wie sich Schüler:innen in sogenannten Lernberatungs- bzw. Lernentwicklungsgeräten vor Lehrkräften und Eltern selbst bewerten (müssen). In den Studien wird deutlich, dass die Anzahl an Dokumenten (Portfolios, Lernentwicklungsberichte und Berichtszeugnisse) gewachsen ist, in denen schriftlich über die Lernfortschritte durch Schüler:innen selbst und durch Lehrkräfte Auskunft gegeben wird (Bossen, 2020; Strauß, 2023), dass das Publikum aus Mitschüler:innen oder Eltern nicht selten zur Mitbewertung aufgefordert wird (Bossen, 2020; Strauß, 2023) und dass sich das, was als Leistung verstanden wird, ausdifferenziert, umcodiert und vom Anspruch her, was damit als Leistung erfasst werden soll, gesteigert hat (Rabenstein, 2021; Rabenstein et al., 2015).

Zusammenfassend gesagt, ermöglicht es ein Verständnis von Vergleichen und (Be-)Werten bzw. Vergleichen und Differenzieren als miteinander verschrankten Praktiken, (sportliche) Leistung als eine Differenz zu verstehen, die konstitutiv mit der Herstellung von Vergleichbarkeit verknüpft ist. Als analytische Perspektivierung kann so in den Blick geraten, wie im Detail

durch die Arbeit an einer Vergleichbarkeit der Schüler:innen auch an deren Differenzierung und Kategorisierung anhand der Schulnoten gearbeitet wird. In unseren Daten hat sich, wie wir im Kapitel zu den empirischen Rekonstruktionen zeigen werden, dabei prominent die Frage nach der Fairness im Kontext sportunterrichtlicher Praktiken des Vergleichens, Differenzierens und Bewertens herausgestellt, die sowohl bewertungssoziologisch als auch erziehungswissenschaftlich im Zusammenhang mit unterrichtlichen Bewertungspraktiken noch kaum aufgegriffen ist. In diesem Zusammenhang wird in unseren Daten auch die Frage nach der Legitimität (sport-)unterrichtlicher Bewertungen sowohl aus Schüler:innen- als auch Lehrkräfteperspektive vi-rulent, auch wenn wir in den folgenden empirischen Rekonstruktionen nicht im engeren Sinne auf die hier angedeutete Dynamisierung (sport-)unterrichtlicher Bewertungspraktiken (etwa im Zusammenhang mit Verfahren der Schüler:innenselbstbewertung) eingehen können.

2. (Diskursive) Praktiken des Vergleichens, Kategorisierens, Bewertens und Entscheidens

Im Folgenden analysieren wir Daten aus einer diskursethnographischen Studie⁴ zur Konstruktion von Leistung als Differenzordnung im Sportunterricht (zur methodologischen Ausrichtung Mummelthey et al., 2023). Aus einer praxis- und diskurstheoretischen Perspektive wird Bewertung dabei als ein sozial situierter Prozess untersucht (Kalthoff & Dittrich, 2016, S. 463–464). Die verschiedenen Materialien, die den Diskurs abbilden, in diesem Fall ethnografische Protokolle und Interviews, werden in der Diskursethnografie nicht als zwei unterschiedliche voneinander getrennte Gegenstandsbereiche, sondern im Foucaultschen Sinne als miteinander verwoben verstanden (Ott et al., 2012). Auch die in Texten vorhandenen Diskurse werden dabei als Praktiken des Sprechens/Schreibens über etwas konzeptualisiert, die wiederum

4 Die Studie bezieht im Sinne einer Diskursethnografie sowohl in-situ-Praktiken der Leistungskonstruktion als auch Leistungsdiskurse des Sportunterrichts mit ein. Die hier verfolgte diskurs- und praxistheoretisch inspirierte Perspektive auf Sportunterricht ermöglicht diesen als Zusammenspiel »körperlich befähigte[r]« (Schmidt 2017, S. 338) Akteur:innen zu verstehen, die situativ spezifische Vorstellungen von Sportunterricht bzw. sportunterrichtlichen Gegenständen hervorbringen, die wiederum mit diskursiv hervorgebrachten Vorstellungen schulischer Normen, etwa zu sportiven Körpern, verbunden sind.

Gegenstände und Akteur:innen konstituieren (Ott et al., 2012). Während sich anhand der Protokolle aus der teilnehmenden Beobachtung alltägliche Formen des Vergleichens und der Differenzierung im Sportunterricht rekonstruieren lassen und der Fokus hierbei nicht nur auf Gesprochenem, sondern auch auf körperlichen Praktiken liegt, lassen sich aus den problemzentrierten Interviews zusätzlich normative Orientierungen von pädagogischen Professionellen zu Sportunterricht, Leistung und Benotung herausarbeiten, die im Vollzug des alltäglichen Sportunterrichts unausgesprochen bleiben. Insgesamt rücken wir damit die lokale Praxis des sportunterrichtlichen Bewertens in den Blick und verstehen Leistung nicht als einer Bewertung vorgelagert, sondern als Ergebnis solcher Praktiken (Kalthoff, 1996).

In den folgenden empirischen Rekonstruktionen werden wir zum einen Ausschnitte aus ethnografischen Protokollen einer mehrwöchigen teilnehmenden Beobachtung (Breidenstein et al., 2020) aus dem Sportunterricht eines Gymnasiums verwenden und zum anderen Interviewsequenzen aus problemzentrierten Interviews (Reinders, 2012; Witzel, 2000) mit Sportlehrkräften zu ihrem Erleben von Sportunterricht und der dortigen Bewertung. Die beiden ersten Protokolle zeigen Situationen der Anbahnung von Wettkampfspielen und damit zusammenhängende Praktiken der Herstellung von in Bezug auf sportliche Leistungsfähigkeit vergleichbare Schüler:innen(-gruppen). Das dritte Protokoll zeigt eine Situation, in der die Leistung von Schüler:innen anhand einer Tabelle zu sportlicher Leistung im Rahmen des Sportabzeichens in eine ordinale Logik eingeordnet wird, an der sich ein sportunterrichtliches Verständnis sportlicher Leistung beobachten lässt. In den Interviewauszügen werden anhand von Positionierungen von Lehrkräften zu dem Problem der Bewertung bzw. der Frage »wie werde ich dem Einzelnen gerecht«, »fairer Wettkampf« und »gerechte Bewertung« im Sportunterricht als fragile Konstrukte rekonstruiert.

An den Interpretationen der Daten zeigen wir im Folgenden im Detail,

1. wie Vergleichen und Differenzieren in alltäglichen Praktiken des Sportunterrichts miteinander verwoben sind und wie diese eine Arbeit an dem Konzept des ›fairen‹ im Sinne eines chancengleichen Wettkampfs darstellen.
2. wie – stärker auf schulische Benotung als institutionelle Bewertung bezogen – Vergleichbarkeit im Zuge einer Differenzierung nach sportunterrichtlicher Leistung ebenfalls mit der Frage nach einer ›fairen‹ bzw.

allen Schüler:innen gerecht werdenden Bewertung zu einem Entscheidungsproblem von Lehrkräften wird.

3. wie sich hieran für den Sportunterricht bzw. die Legitimierung von Bewertungen sportlicher Leistung im Sportunterricht über ordinale Logiken insgesamt die konstitutive Verschränkung der Herstellung von Vergleichbarkeit und Differenzierung bzw. die Bedingung einer festzustellenden Vergleichbarkeit für eine als legitim angesehene Differenzierung von Schüler:innen nach Leistung nachvollziehen lässt.

a) Vergleichen und Differenzieren in alltäglichen Praktiken des Sportunterrichts

Wie ›Fairness‹ im Zusammenhang mit der Herstellung von Vergleichbarkeit zwischen den Schüler:innen aufgerufen wird, zeigen wir im ersten Schritt. Das erste Protokoll verweist dafür auf ein außerschulisch festverankertes vergeschlechtlichtes Verständnis sportlicher Leistung, dass in Bezug auf eine zu erreichende Chancengleichheit für sportliche Leistung im Sportunterricht relevant gemacht wird (Wilm, 2021). Es erscheint im unterrichtlichen Kontext dort (bzw. nur dort) legitim, wo nach Geschlecht differenzierte, unterschiedlich leistungsfähige Schüler:innen angenommen werden (etwa Zeiten und Wurf- und Sprungweiten im Kontext Leichtathletik), d.h. »implizite Annahmen über körperliche Voraussetzungen ersichtlich werden« (Wilm, 2021, S. 279). Die protokolierte Situation verweist auf eine im Sportunterricht (nicht nur dieser Klasse) etablierte Praktik zur Herstellung von ›fairen‹ Teams im Sportunterricht in Bezug auf eine angenommene vergeschlechtlichte körperliche Leistungsfähigkeit von Schüler:innen.

Es werden erst zwei Jungen und dann zwei Mädchen von der Lehrkraft ausgewählt. Eines der Mädchen soll zwei Mädchenteams und einer der Jungen zwei Jungenteams zusammenstellen. Die jeweils andere Person des gleichen Geschlechts, bestimmt dann Mädchen bei den Mädchenteams, Junge bei den Jungenteams – wer der beiden welches Team bekommt. Bei den Mädchen lässt sich ein Zuteilen á la ›du hier und du da‹ beobachten. Dann werden jedoch ein paar Entscheidungen mit Buh-Rufen kommentiert, was die Aufmerksamkeit der Ethnographin weckt. Aus der Jungengruppe lassen sich nun ebenfalls Unstimmigkeiten vernehmen, die jedoch direkter auf zugeschriebene Merkmale einzelner Personen verweisen: »der ist aber schlechter als der«, »wir haben aber dafür den«, »aber der kann nur fangen

nicht werfen«. Schließlich haben sich sowohl die Mädchen als auch die Jungen in je zwei Gruppen geteilt. Nachdem in einem Sitzkreis die Regeln rekapituliert und das Spielfeld markiert wurden, weist Frau Herring⁵ den Teams Seiten zu, sodass je ein Jungen- und ein Mädchenteam auf einer Seite zu einem großen Team verschmelzen. (Fieldnotes KU1_3)

In der Situation wird die binär verstandene Differenzdimension Geschlecht für die Mannschaftseinteilung relevant gemacht. Dabei wird nicht bedeutsam gemacht, welche Mädchen mit welchen Jungen in ein Team kommen, sondern nur, dass in beiden Teams am Ende etwa gleich viele sind. Erst vor dem Hintergrund dieser Vergleichbarkeit der Leistung der Teams erscheint der Wettbewerb, in den sie treten sollen, sinnvoll, insofern das Ergebnis des Wettkampfs nur unter diesen Voraussetzungen offen scheint, was wiederum als eine Grundbedingung (»fairen«) sportlichen Wettkampfs diskutiert wird (Schürmann, 2023; Minnetian und Werron in diesem Band). Teams mit einer möglichst gleichen Verteilung von – geschlechtlich unterschiedlich und damit körperlich unterschiedlich leistungsfähig gelesener – Schüler:innen wurden dabei als Voraussetzung für ein »faires« Spiel bzw. eine »faire« Differenzierung in Gewinner:innen und Verlierer:innen hervorgebracht. Durch die Konstruktion einer Gleichverteilung (geschlechtlich) unterschiedlicher Schüler:innenkörper entsteht eine Vergleichbarkeit der Leistungen der Teams für das kommende Spiel als Vorstellung und als Anforderung.

Indem nun eine solche Einteilung von Schüler:innen, wie sich hier ebenfalls zeigt, kritisiert oder in Frage gestellt wird – im Protokoll wird beschrieben, wie die Entscheidungen bei der Mannschaftswahl ablehnend oder zustimmend mit Reinrufen von Schüler:innen bewertend kommentiert werden –, wird in der Wahrnehmung der Ethnografin eine Norm der »gerechten« Zuteilung der Schüler:innen zu den Mannschaften aufgerufen. Somit werden auch im Zuge der Mannschaftseinteilung Schüler:innen schulklassenöffentlich (erst) als solche hervorgebracht, von denen mehr oder weniger sportliche Leistung in dem kommenden Spiel erwartet werden kann.

Ein solches In-Frage-Stellen bzw. Abwägen von Vergleichbarkeit im Kontext der Anbahnung eines sportlichen Wettkampfs zeigt auch das folgende Protokoll zu einem Ranglistenturnier. Während der Erklärung der Spielmodalitäten des »Wäscheleinenspiels« kommt es hier seitens der Schüler:innen zu Kritik in Bezug auf die »Fairness« der Startbedingungen für das Spiel.

⁵ Alle Namen in den Protokollen wurden pseudonymisiert.

Zunächst wurden an eine Wäscleine, die als Wettkampftableau dient, ohne weitere Anweisung und ohne, dass ein Anfang oder Ende markiert wurde, von den Schüler:innen Namenskärtchen platziert.

Dann erklärt Herr Simon das »Wäscleinenispiel«. Die Wäscleine samt Namenskarten dient dabei als Wettkampftableau und die Schüler:innen sollen sich gegenseitig zu Miniwettkämpfen an einer der Stationen auffordern. Die Person die aufgefordert wird, darf maximal fünf Plätze von dem eigenen Namenskärtchen weg sein. Das Schild ganz vorn markiert den ersten, das ganz rechts den letzten Platz. (Fieldnotes KU4_1)

Mit dem Wissen um die Aufgabe, sich mit den Personen, deren Namenskarte vor der eigenen Karte hängt, zu duellieren und auf diese Weise zu versuchen, auf Plätze weiter vorn zu gelangen, lässt sich die zufällige Positionierung der Namenszettel an der Leine als eine Aufforderung verstehen, sich gegenüber bestimmten Mitschüler:innen in Bezug auf bestimmte Bewegungen zu beweisen. Dies erinnert an den im Anfang der modernen Schule im Klassenzimmer bei jeder Lehrerfrage (m!) körperlich ausgeführten Wettstreit der Schüler (m!) um ihren oder einen besseren Rangplatz in der Sitzordnung, was allerdings im Laufe der Zeit schulisch zunehmend unerwünscht war und schließlich verboten wurde (Lindenhayn, 2013). In den Äußerungen, in denen die Ausgangsbedingungen des Spiels kritisch kommentiert werden, zeigt sich, wie die eigenen Chancen angesichts der Positionierung der eigenen Namenskarte zu taxieren versucht werden.

Herr Simon verkündet auch, dass er selbst mitspielen wird. Woraufhin ich aus den Reihen der Schüler:innen den Einwand höre, dass das unfair sei, er müsse sein Namenskärtchen »schon ganz hinten hinhängen«. Phillip wechselt seinen Platz im Sitzkreis und ich höre wie er mehrmals wiederholt »ich bin sechster, ich bin sechster. Vor mir sind nur Mädchen aber die sind gut im Seilspringen« (Seilspringen ist eine der möglichen Stationen). Ziel des Spiels ist es durch die verschiedenen Duelle möglichst weit vorn im Tableau zu landen bzw. sein Namenskärtchen möglichst weit nach rechts auf der Wäscleine zu platzieren. Es melden sich vier Schüler:innen, die auch drangenommen werden. Maja merkt Bedenken am Start des Spiels an: »ist das nicht unfair, weil jetzt einige weiter vorne hängen?« »Nein, das ist Pech« antwortet Herr Simon. (Fieldnotes KU4_1)

In den Einwänden und der Kritik an der Anordnung der Karten auf der Wä-scheleine tritt die Norm einer anzustrebenden Chancengleichheit für die Be-teiligung an dem Spiel immer wieder in den Vordergrund. Dies zeigt sich vor allem in der Reaktion auf die Teilnahme Herr Simons am Spiel sowie der In-teraktion zwischen Maja und Herr Simon.

Die zuerst aufgerufene Kritik an den Ausgangsbedingungen für das Spiel fokussiert die Vergleichbarkeit von Lehrkraft und Schüler:innen für den Wettkampf. Herrn Simon wird dabei von den Schüler:innen ein ›unfairer‹ Vorteil gegenüber ihnen unterstellt. Dieser Vorteil könne, wenn überhaupt, durch die schlechteste Startposition (›schon ganz hinten‹) ausgeglichen werden. An die-ser Stelle wird deutlich, dass die Verteilung der Namenskärtchen auf der Wä-scheleine für den weiteren Verlauf des Wettkampfs als bedeutend angesehen wird. Auch die zweite Kritik bezieht sich auf die Startpositionierungen: Durch die unterschiedlichen und zufälligen Platzierungen der Zettel an der Leine hätten – so Maja – nicht alle Schüler:innen die gleichen Ausgangspositionen, um auf einen Platz vorne an der Leine zu gelangen. Von einem Platz weiter hinten an der Leine zu starten, wird dabei – vor dem Hintergrund des Aufforderungs-charakters zur Verbesserung der eigenen Positionierung während des Spiels – als negativ hervorgehoben. Die als ungleich markierten Startbedingungen werden dabei als unfair kommentiert. Maja nimmt damit Bezug auf die Ori-en-tierungsnorm der Fairness, die auch im außerschulischen Sport eine wichtige Rolle spielt, und fordert diese auch für den Schulsport ein. So wird Fairness etwa als Maßstab in der Olympischen Charta, sprich Verfassung der Olympi-schen Bewegung, deklariert (Schürmann, 2023, S. 336). Herr Simon weist den Vorwurf, das Spiel sei nicht fair organisiert und würde damit spezifischen An-sprüchen an einen sportlichen Wettkampf nicht gerecht werden, mit einem ›Nein‹ entschieden zurück und bringt mit dem Begriff ›Pech‹ stattdessen Zu-fall als Größe ein. Einem zufälligen Vorgehen kann nur schwer vorgeworfen werden, Personen ungleich zu behandeln, da es zufällig ist, sprich keiner be-stimmten Logik – also auch keiner Differenzierungslogik – folgt. Die Aussage ›das ist Pech‹ immunisiert sich zudem selbst, in dem sie keine weiteren Mög-lichkeiten des gegenstandsbezogenen Anknüpfens ermöglicht.

In der nun folgenden Beobachtung zum dritten Protokoll ist die Frage der Fairness eng mit einer formalisierten Form der Bewertung im Unterricht, nämlich dem Sportabzeichen verbunden. Zwei Schüler:innen sollen/möchten ihre Disziplinen für das Sportabzeichen um die Ausweichdisziplin Seilspringen ergänzen. Beide stehen vor der durch die Leistungstabellen des Sportab-zeichens gesetzten Anforderung bestenfalls 30 Sprünge zu erreichen, ohne zu

stolpern. Die Lehrkraft übernimmt dabei die Aufgabe eines Protokollanten, der durch lautes Mitzählen jeden einzelnen Sprung für das Sportabzeichen verifiziert. Die Prüfungen von Maja und Johann unterscheiden sich dabei im Umgang Herrn Simons mit der jeweiligen Anzahl an Sprüngen bei der Überführung in die Skalierung des Sportabzeichens.

Johann springt und Herr Simon zählt laut mit und stoppt bei 29. »Ach was, ich schreib 30 auf. Du hast Gold« verkündet Herr Simon und Johann nickt das Ergebnis ab. »Er war ja auch so lange nicht da« stimmt eine Schülerin zu. Dann ist Maja dran und springt. Sie stolpert bei der laut mitgezählten 26. »Ach was du machts gleich nochmal« fordert Herr Simon sie zu einem neuen Versuch auf. In der Zwischenzeit springt noch ein weiterer Schüler, dann ist Maja nochmal. In der Halle sind jetzt nur noch Maja, die Lehrkraft und ich. Auch dieses Mal werden es keine 30. Dann verlassen wir alle die Halle. (Fieldnotes KU4_4)

Die Prüfung von Johann verläuft dabei unerwartet: Johann erhält die bestmögliche Zuordnung (Gold), indem Herr Simon ihm mehr Sprünge zuschreibt als gezählt wurden, nämlich 30 statt 29. Johann wird somit in Bezug auf sportliche Leistung als starker Schüler anerkannt – auch wenn er dies in der konkreten Situation nicht vollumfänglich aufführen konnte. Es ließe sich hier von einer Kulanzpraktik sprechen (M. Meier, 2011): Angesichts eines zugeschriebenen grundsätzlichen Könnens wird von der genauen Erfüllung der Leistungs-erwartung Abstand genommen und ungeachtet dessen volle Punktzahl gegeben. Herr Simon arbeitet damit explizit daran, Johann als sehr guten Sportler in Bezug auf die Kriterien des Sportabzeichens zu erhalten. Er positioniert sich als nachsichtige Lehrkraft, die bei einem so knappen »Scheitern« an der höchsten Auszeichnung auch mal bereit ist, einen Wert anzupassen. Eine Legitimation hierfür gibt dabei interessanterweise nicht er selbst, sondern eine Schülerin, die zustimmend feststellt, dass Johann ja auch länger nicht dagewesen ist. Zu vermuten ist, dass dies impliziert, dass er weniger Möglichkeiten hatte, Werte für das Sportabzeichen zu sammeln als die anderen Schüler:innen, und das Aufrunden deshalb gerecht sei. Auch bei Maja wird nicht die erste erzielte Anzahl an Sprüngen für das Sportabzeichen übernommen. Anstatt dass der Wert von der Lehrkraft angepasst wird, verbleibt der Zugzwang aber bei der Schülerin, die zu einem weiteren Versuch aufgefordert wird. Damit markiert er, im Gegensatz zur Prüfung von Johann, keine Differenz zwischen der gera-

de beobachteten ›Leistung‹ Majas im Vergleich zu einem globaleren Eindruck in Bezug auf ihr sonstiges Abschneiden im Sportunterricht.

Insgesamt konnten wir zeigen, wie in Praktiken des Sportunterrichts daran gearbeitet wird, einen ›fairen Wettkampf‹ im Sinne einer Chancengleichheit zwischen den Schüler:innen oder Teams herzustellen und zu erhalten. ›Fairness‹/Chancengleichheit wurden dabei als fragile normative Konstrukte deutlich, die durch Beobachtung und Kritik der Schüler:innen im Vollzug sportunterrichtlicher Praktiken mithergestellt und aktualisiert werden, aber in ihrer Bedeutung auch immer in Gefahr sind. Aus bewertungssoziologischer Sicht wird deutlich, dass die Herstellung vereinheitlichter Startbedingungen der in Bezug auf ihre Leistung zu vergleichenden Schüler:innen erst ermöglicht, einen Wettkampf als fair im Sinne gleicher Chancen für alle (Reinold & Kalthoff, 2021) zu verstehen.

b) Positionierung von Lehrkräften zu Bewertung sportlicher Leistung im Unterricht als Problem

In den folgenden Interviewsequenzen wird eine ›faire Benotung‹ als die Möglichkeit thematisiert, dass alle eine gute Benotung erhalten könnten. Als zentrales Dilemma wird dabei die Passung von geforderten Werten (z.B. Laufzeiten, Sprungweiten) für gute Noten und die körperlichen Merkmale einzelner Schüler:innen hervorgehoben.

I: welche Rolle [spielt] der Körper oder eben die Körperlichkeit der Schüler:innen für den Sportunterricht

LK2: Ok also ist im endeffekt schon ne zentrale rolle weil einfach wie schon gesagt wir müssen sportarten unterrichten wir haben notentabellen wo einfach drinsteht wer die hundertmeter in zwölf sekunden läuft kriegt so und so ne note wer die in fünfzehn sekunden läuft kriegt die note und wenn dann jemand mit deutlichem Übergewicht in die hundertmeterbahn startet dann kann der einfach nicht so schnell laufen wie jemand der vom vom bmi her oder von anderen/vom körperlichen her einfach andere voraussetzungen hat daher für die/für diese reinen leistungsnoten spielt das für die schüler natürlich ne riesen rolle und für mich auch daher ist es für mich auch immer ganz wichtig bei der benotung nicht nur diese reinen leistungsnoten sondern eben auch techniknoten zu vergeben dass ich eben da auch son bisschen die entwicklung und den einsatz und das engagement im unterricht bewerten kann und da fließt dann eben auch mit ein ob man ne entwicklung überhaupt sehen kann also die leute die einfach

weniger vorerfahrung haben und das müssen ja jetzt nicht unbedingt ich sag mal dickere oder dünnere leute sein sondern das ist einfach jeder der eben keine vorerfahrung in der sportart hat der kann sich entwickeln muss aber nicht und je nachdem wie da das eigene engagement eben ist das fließt dann auch in diese techniknoten ein bisschen mit ein wie war die entwicklung über die dauer und von wo ist er gestartet und wo ist er jetzt rausgekommen und da würd ich sagen spielt jetzt die körperlichkeit jetzt nicht sone übergeordnete rolle aber für diese reinen leistungsnoten denke ich macht das schon einen unterschied ja und wie war dann/das wissen die schüler natürlich auch und sind dann meistens auch sehr froh über diese techniknoten dass sie eben da auch ein bisschen was zeigen können und eben nicht eben einfach nur über ihre voraussetzungen des körpers da manche noten einfach nicht mehr erreichbar sind. (Interview LK2)

Die sprechende Lehrkraft positioniert sich angesichts der an sie gerichteten Frage nach der Rolle von Körper und Körperlichkeit von Schüler:innen für Bewertungen im Sportunterricht als in ihrer Auffassung über Bewertung im Sportunterricht an bestimmte Vorgaben gebunden: Es existieren Tabellen, die messbare sportliche Leistungen in einer bestimmten Sportart in bestimmter Weise – nach Alter und Geschlecht kategorisierte Gruppen von Schüler:innen – als Ordinalskala abbilden und bestimmte Werte mit bestimmten Notenbezeichnungen äquivalent setzen. Angesichts dieser Vorgaben positioniert sich die Lehrkraft als kritisch gegenüber der damit suggerierten Vergleichbarkeit der sportlichen Leistungen von Schüler:innen im Sportunterricht, da sie aufgrund bestimmter hier herausgestellter körperlicher Voraussetzungen nur eingeschränkt vergleichbar seien: Angesichts eines implizit vorauszusetzenden trainierten/schlanken Körpers für schnelles Laufen werden Körpermaße und Gewicht zu Kriterien, anhand derer Schüler:innen taxiert und ihre Leistungs(un)fähigkeit kommentiert werden kann. Die Kategorisierungen der Schüler:innenkörpern ermöglichen von sportlicher Leistungsfähigkeit abweichende Schüler:innenkörper auszumachen. Einen Ausgleich zu schaffen sei möglich, so stellt es die Lehrkraft dar, wenn neben der vorgesehenen Messung sportlicher Leistung in Form von Schnelligkeit etc. zusätzlich Aspekte wie Technik und Entwicklung für wohlgesonnene Bewertungen genutzt werden würden. Die Bewertungen von Technik und Entwicklung werden in der Interviewsequenz mit Aspekten, wie dem Zeigen von Motivation, Entwicklungsbereitschaft und -anstrengung im Allgemeinen, welche sich auch als klassische pädagogische Kriterien schulischer Leistung benennen ließen, in einen Zu-

sammenhang gebracht. Diesen Aspekten wird das Potenzial zugeschrieben, damit kompensatorisch auf ungleiche körperliche Voraussetzungen reagieren zu können.

Eine ähnliche Auseinandersetzung findet sich auch bei LK1, der ›Fairness‹ und Vergleichbarkeit vor allem in Bezug auf körperliche Voraussetzungen und ihre Bedeutung für sportunterrichtliche Bewertung thematisiert.

I: [...] wie stellst du da sicher dass die Schüler:innen alle gut mitmachen können [...] gibts da ne Möglichkeit das sicher zu stellen

LK1: dass sie gut mitmachen können oder dass sie gut abschneiden

I: Mhm beides vielleicht *lacht

LK1: *lacht* weil das ist ein Unterschied im Unterricht selbst kann jeder im Rahmen seiner Möglichkeiten teilnehmen [...] wenn ich jetzt ein stark adipösen Schüler oder Schülerin hätte [...] die einfach Hochsprung nicht leisten können dann kann ich das auch nicht abprüfen genauso mit körperlichen Behinderungen wenn es einfach physisch nicht möglich ist dann kann/darf ich finde ich ethisch auch keine Note darauf geben einfach weil [...] eigentlich es gibt dieses schöne Bild wo verschiedene Tiere vor einem Baum stehen also Giraffe Affe Elefant Goldfisch in einem Glas und dann wird die Aufgabe gestellt sollen alle auf den Baum klettern also dass haben wir (unv.) häufig als Sinnbild dafür auch gehabt und es funktioniert nicht ähm genauso kann ich nem Rollstuhlfahrer nicht sagen dass er Hochsprung machen soll und dann muss ich andere Kriterien dann finden die Frage ist nur ab wo zieht man die Grenze ab wo sagen also ab welchem/ab welcher Einschränkung letztlich sagt man die Leistung und die Leistungsüberprüfung ist nicht ablegbar [...] grundsätzlich sehe ich aber das im Normalfall alle Schüler an der Arbeit mit dem Inhalt teilnehmen können und auch eine Prüfung ablegen können [...] genau sodass es da prinzipiell erstmal jeder eine Leistung erbringen kann ob die Leistung dann in den vorgeschriebenen Tabellen abzubilden ist oder abbildungbar ist das ist dann nochmal eine zweite Frage (Interview LK1)

LK1 rahmt angesichts einer Norm der Berücksichtigung spezifischer körperlicher Besonderheiten und daraus resultierender Teilnahmemöglichkeiten am Sportunterricht die Entscheidung darüber, wann ein:e Schüler:in in einer sportlichen Disziplin unter der vorgegebenen Rahmung prüfbar sei, als eine ethische Frage. Damit wird diese Frage mit Bedeutung aufgeladen. Ob eine Prüfung ethisch vertretbar ist oder nicht, wird an die physische Möglichkeit gebunden, die geforderten Bewegungsabläufe zu absolvieren (z.B. Hochsprung als Rollstuhlfahrer). Das »ethische« Problem wird zu einem

Entscheidungsproblem der Lehrkraft gemacht. Darin wird ihre Verantwortlichkeit für Unterricht und Bewertungen von Schüler:innen hervorgebracht und in der Frage nach der Grenzziehung zwischen prüfbar und nicht prüfbar verortet. Interessant ist hier auch LK1 Verweis auf den »Normalfall«: Als Normalität des Sportunterrichts wird hier entgegen der vorherigen Ausführungen angenommen, dass alle Schüler:innen die Prüfung, also die geforderten Bewegungsabläufe, absolvieren können – sprich diesbezüglich gleich sind – und die Frage der Prüfbarkeit sich erst in der Überführung der erzielten Werte in die Tabellen ergibt, welche die Werte in Noten umwandelt.

Beide Lehrkräfte nehmen in ihrem Sprechen über – aus ihrer Perspektive – (un)gerechte Bewertung Bezug auf mithilfe von festgelegten Tabellen organisierte Benotungspraktiken. Diese Tabellen würden dabei eine Vergleichbarkeit der Schüler:innen unterstellen, die aus Sicht der Lehrkräfte – aufgrund unterschiedlicher körperlicher Voraussetzungen der Schüler:innen – nicht zwingend gegeben ist. Als prekär verstandene Schüler:innenkörper werden dabei von den Lehrkräften als Ausgangspunkt verstanden weitere Benotungspraktiken vorzuschlagen, die eine attestierte Unterschiedlichkeit der Schüler:innen in Bezug auf ihr sportliches Potenzial ausgleicht.

c) Zur Bedeutung von Fairness im Sport und seine Verunterrichtlichung. Relationierung der Rekonstruktionen

Angesichts der im Call für dieses Publikationsprojekt von den Herausgebern dieses Bandes aufgeworfenen Frage (vgl. Schmidt et al. in diesem Band), inwiefern »instruktive Fälle aus dem Sport, die die praktische Verwobenheit des Kategorisierens, Vergleichens, Quantifizierens, Entscheidens und anderer, familienähnlicher Praktiken zeigen«, bestehende Konzepte, Theorien und Begriffe in der Soziologie der Bewertung präzisieren oder erweitern könnten, regen unsere Analysen an, das Konzept einer Verwobenheit von Vergleichen, Differenzieren/Kategorisieren und Bewerten um die Frage nach dem in der sportpädagogischen Diskussion mitgeführten Konstrukt eines ›fairen Wettkampfs‹ zu erweitern. Der schüler:innen- und lehrer:innenseitige Bezug auf ›Fairness‹ taucht in den hier herangezogenen Protokollausschnitten als ein zentraler Bezugspunkt sportunterrichtlicher Bewertungslogik auf. Deutlich wurde, dass in alltäglichen Praktiken des Sportunterrichts stetig auch daran gearbeitet wird einen ›fairen Wettkampf im Sinne einer Chancengleichheit zwischen den Schüler:innen oder Teams herzustellen und zu erhalten. Fair Play im Sinne von Chancengleichheit lässt sich als moralische Grundlage

für sportliche Leistungsvergleiche verstehen (Reinold & Kalthoff, 2021). Die Relevanz von Fairness im Sportunterricht sieht Schürmann (2023) durch die starke Orientierung des Sportunterrichts am Wettkampfsport begründet. An sportliche Wettkämpfe wird dabei der Anspruch erhoben, dass sein Ausgang offen sein soll, also nicht vorher feststehen darf, wer gewinnt. Hieran sieht Schürmann »ein minimales Verständnis von Fairness« (2023, S. 335). ›Fairness‹ zu gewährleisten, stellt sich allerdings in diesen Daten insofern als fragil dar, als dass sie von den Akteur:innen des Sportunterrichts auch immer wieder erst angemahnt und eingefordert wird. Die Frage von Chancengleichheit, gerade auch in Bezug auf vergleichbare Ausgangsbedingungen für einen Wettkampf/ein Spiel, ist dabei eng verbunden mit der Aussicht auf das antizipierte Ergebnis einer Einordnung als besser/schlechter als andere.

In den Interviewsequenzen wird von beiden Lehrkräften eine ›gerechte Beurteilung‹ als Möglichkeit thematisiert, dass alle Schüler:innen – trotz ihrer körperlichen Voraussetzungen – eine gute Benotung erhalten könnten. Als zentrales Dilemma wird die Passung von (z.B. durch Leistungstabellen) geforderten Werten und den als defizitär beschriebenen körperlichen Merkmalen einzelner Schüler:innen hervorgehoben. Als prekär verstandene Schüler:innen-körper werden dabei als Ausgangspunkt genommen, zahlenwertbasierte Benotungspraktiken zu erweitern, um so die attestierte Unterschiedlichkeit der Schüler:innen in Bezug auf ihr sportliches Potenzial auszugleichen.

Zudem lässt sich die Arbeit an der Vergleichbarkeit der sportlichen Leistungen unterschiedlicher Schüler:innenkörper bzw. der den Schüler:innen zugeschriebenen Leistung angesichts der Forderung nach einem inklusiven Sportunterricht als Bearbeitung des Spannungsfelds zwischen der sportpädagogischen Forderung nach einem inklusiven, allen Schüler:innen gerecht werdenden Unterricht und der sich durch den Alltag von (Sport)Unterricht konstitutiv ziehenden Anforderungen an Leistung deuten, oder, abstrahierter gedacht, als Bearbeitung der Unvereinbarkeit von Einfach-anders- und Besser/schlechter-Vergleichen (Heintz, 2021).

Angesichts dessen, dass sich, wie wir auf Basis der Analysen zeigen können, ein ›fairer Wettbewerb/Wettkampf‹ im Sportunterricht einerseits als fragil beschreiben lässt, andererseits viel Aufwand betrieben wird, durch entsprechende Vorkehrung für die Vergleichbarkeit von sportlicher Leistung unterschiedlicher Schüler:innenkörper den ›Glauben an ihn‹ aufrechtzuerhalten, kommen wir zu der These, dass sich die Legitimität von Kategorisierungen von Schüler:innen in Bezug auf sportliche Leistung als besser/schlechter im Sportunterricht in doppelter Weise stabilisiert: einerseits durch die gerade

nicht von außen (nur) als gesellschaftlicher Auftrag kommende, sondern für Unterricht konstitutive Anforderung, angesichts einer gestellten Aufgabe und deren Bearbeitung, etwas als Leistung zu zeigen (Dietrich, 2019; Mechan, 1979), und andererseits durch ordinale Logiken des außerschulischen (Wettkampf)Sports.

Fazit

In diesem Beitrag haben wir nach den sozialen Logiken und Effekten von Bewertungspraktiken im Sportunterricht mit Blick auf Praktiken des Vergleichens, Differenzierens und Kategorisierens sowie Entscheidens im Alltag des Sportunterrichts gefragt. Ein Verständnis von Vergleichen und (Be-)Werten bzw. Vergleichen und Differenzieren als miteinander verschränkte Praktiken kann hierbei dazu beitragen (sportunterrichtliche) Leistung als eine Differenz zu verstehen, die konstitutiv mit der Herstellung von Vergleichbarkeit verknüpft ist.

Zusammenfassend lässt sich *erstens* festhalten, dass dem Erkennen, Messen und Beurteilen von Leistung die Annahme und Herstellung gleicher bzw. vergleichbarer Schüler:innen zugrunde liegt. Diese angenommene Gleichheit wird jedoch immer wieder brüchig und bedarf einer stetigen Aktualisierung oder Modifikation, um Bedeutung zu behalten. *Zweitens* erhält die Frage nach Vergleichbarkeit als Grundlage einer Differenzierung nach Leistung für den Sportunterricht durch die sportliche Norm der Fairness und das meritokratische Versprechen von Schule eine (ver-)doppelte Relevanz. Hinzu käme noch, nimmt man Schürmann (2023) zur Kenntnis, eine pädagogisch moralische Dimension im Sinne des Einübens eines gerechten Miteinanders.

Mit Rückbezug auf die eingangs formulierte Verunterrichtlichung sportlicher Leistung und Versportlichung von Unterricht – legen unsere Überlegungen nahe, dass das, was sportlich als Wettbewerb markiert wird, sich in der unterrichtlichen Logik als Inszenierung eines Wettbewerbs – und damit gerade nicht als Wettbewerb darstellt –, insofern der Ausgang des Wettbewerbs – welches Team, gewinnt – als folgenlos für die Leistungsbewertung im Sportunterricht angenommen werden kann. Nichtsdestotrotz braucht auch diese Inszenierung den festen Glauben an vergleichbare Leistungen beider Mannschaften. Die sport(unterricht)liche Leistung der Schüler:innen wird hingegen angesichts fixierter zu erwartender (Mindest-)Werte (ordinale Tabelle) oder angesichts einer in Bewertungspraktiken immer wieder zu justierenden bzw.

justierten sozialen Bezugsnorm in der Lerngruppe vergeben. Die Benotungspraxis im Sportunterricht zu untersuchen, wäre aber das Thema eines zukünftigen Forschungsprojekts.

Literatur

- Balz, E. (2009). Fachdidaktische Konzepte *update* oder: Woran soll sich der Schulsport orientieren? *Sportpädagogik*, 33(1), 25–32.
- Bellmann, J., Duzevic, D., Schweizer, S. & Thiel, C. (2016). Nebenfolgen Neuer Steuerung und die Rekonstruktion ihrer Genese. Differentielle Orientierungsmuster schulischer Akteure im Umgang mit neuen Steuerungsinstrumenten. *Zeitschrift für Pädagogik*, 62(3), 381–402. <https://doi.org/10.2565/01:16764>
- Bellmann, J. & Waldow, F. (2007). Die merkwürdige Ehe zwischen technokratischer Bildungsreform und empathischer Reformpädagogik. *Bildung und Erziehung*, 60(4), 481–503. <https://doi.org/10.7788/bue.2007.60.4.481>
- Berdelmann, K., Reh, S. & Scholz, J. (2018). Wettbewerb und Ehrtrieb. In S. Reh & N. Ricken (Hg.), *Leistung als Paradigma: Zur Entstehung und Transformation eines pädagogischen Konzepts* (S. 137–163). Springer VS. https://doi.org/10.1007/978-3-658-15799-9_6
- Berli, O., Nicolae, S. & Schäfer, H. (Hg.). (2021). *Bewertungskulturen*. Springer VS. <https://doi.org/10.1007/978-3-658-33409-3>
- Boger, M.-A. (2017). Theorien der Inklusion – eine Übersicht. *Zeitschrift für Inklusion*, 11(1). <https://www.inklusion-online.net/index.php/inklusion-online/article/view/413>
- Bossen, A. (2020). *Das Artefakt als materialisiertes Netzwerk Neuer Lernkulturen: Ein Ikonographisch-Ikonologisch-netzwerktheoretischer Zugang zu stillgestellten Praktiken Neuer Leistungskulturen Im Portfolio*. Peter Lang. <https://doi.org/10.3726/b16630>
- Bräu, K. & Fuhrmann, L. (2015). Die soziale Konstruktion von Leistung und Leistungsbewertung. In K. Bräu & C. Schlickum (Hg.), *Soziale Konstruktionen in Schule und Unterricht: zu den Kategorien Leistung, Migration, Geschlecht, Behinderung, soziale Herkunft und deren Interdependenzen* (S. 49–65). Verlag Barbara Budrich.
- Breidenstein, G., Hirschauer, S., Kalthoff, H. & Nieswand, B. (2020). *Ethnografie: Die Praxis der Feldforschung* (3. Aufl.). UVK.

- Breidenstein, G. & Thompson, C. (2014). Schulische Leistungsbewertung als Praxis der Subjektivierung. In C. Thompson, K. Jergus & G. Breidenstein (Hg.), *Interferenzen: Perspektiven kulturwissenschaftlicher Bildungsforschung* (1. Aufl., S. 89–109). Velbrück Wissenschaft.
- Dietrich, F. (2019). Inklusion und Leistung – Rekonstruktionen zum Verhältnis von Programmatik, gesellschaftlicher Bestimmtheit und Eigenlogik des Schulischen. In E. von Stechow, P. Hackstein, K. Müller, M. Esefeld & B. Klocke (Hg.), *Inklusion im Spannungsfeld von Normalität und Diversität. Band I: Grundfragen der Bildung und Erziehung* (S. 195–206). Verlag Julius Klinkhardt.
- Feth, C. (2014). *Benotungspraxis von Sportlehrkräften – Rekonstruktion von Benotungsstrategien im Sportunterricht* [Dissertation]. Ruhr Universität, Bochum.
- Gellert, U. & Hümmer, A.-M. (2008). Soziale Konstruktion von Leistung im Unterricht. *Zeitschrift für Erziehungswissenschaft*, 11(2), 288–311. <https://doi.org/10.1007/s11618-008-0019-1>
- Grein, M., Vernaci, D., Stöbel, L. & Tesch, B. (2024). Uneindeutige Differenzkonstruktionen und Fachlichkeit im alltäglichen Spanisch- und Französischunterricht sowie der fachdidaktischen Unterrichtsforschung. In K. Rabenstein, C. Bräuer, D. Hülsmann, S. Mummelthey & S. Strauß (Hg.), *Differenzkonstruktionen in fachunterrichtlichen Kontexten. Forschungsansätze und Erträge zu Differenz(de)konstruktion aus Fachdidaktik, Erziehungswissenschaft und Diversitätsforschung*. (S. 55–75). Verlag Julius Klinkhardt.
- Heberle, K. (2019). *Zur Konstruktion von Leistungsdifferenz im Rahmen musikpädagogischer Unterrichtspraxis: eine Videostudie zum instrumentalen Gruppenunterricht in der Grundschule*. Waxmann.
- Heintz, B. (2010). Numerische Differenz. Überlegungen zu einer Soziologie des (quantitativen) Vergleichs. *Zeitschrift für Soziologie*, 39(3), 162–181. <https://doi.org/10.1515/zfsoz-2010-0301>
- Heintz, B. (2016). »Wir leben im Zeitalter der Vergleichung.« Perspektiven einer Soziologie des Vergleichs. *Zeitschrift für Soziologie*, 45(5), 305–323. <https://doi.org/10.1515/zfsoz-2015-1018>
- Heintz, B. (2021). Kategorisieren, Vergleichen, Bewerten und Quantifizieren im Spiegel sozialer Beobachtungsformate. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 73(S1), 5–47. <https://doi.org/10.1007/s11577-021-00741-3>
- Hellmann, K. A., Kreutz, J., Schwichow, M. G. & Zaki, K. (Hg.). (2019). *Kohärenz in der Lehrerbildung: Theorien, Modelle und empirische Befunde*. Springer VS. <https://doi.org/10.1007/978-3-658-23940-4>

- Hirschauer, S. (2019). Urteilen unter Beobachtung. In S. Nicolae, M. Endreß, O. Berli & D. Bischur (Hg.), *(Be)Werten. Beiträge zur sozialen Konstruktion von Wertigkeit* (S. 275–298). Springer VS. https://doi.org/10.1007/978-3-658-21763-1_12
- Ingenkamp, K. (Hg.). (1971). *Die Fragwürdigkeit der Zensurengebung: Texte und Untersuchungsberichte*. Beltz.
- Kalthoff, H. (1996). Das Zensurenpanoptikum: Eine ethnographische Studie zur schulischen Bewertungspraxis. *Zeitschrift für Soziologie*, 25(2), 106–124. <https://doi.org/10.1515/zfsoz-1996-0202>
- Kalthoff, H. (2019). Organisierte Humanevaluation. In S. Nicolae, M. Endreß, O. Berli & D. Bischur (Hg.), *(Be)Werten. Beiträge zur sozialen Konstruktion von Wertigkeit* (S. 221–247). Springer VS. https://doi.org/10.1007/978-3-658-21763-1_10
- Kalthoff, H. & Dittrich, T. (2016). Unterscheidung und Härtung. Bewertungs- und Notenkommunikation in Lehrerzimmer und Zeugniskonferenz. *Berliner Journal für Soziologie*, 26(3–4), 459–483. <https://doi.org/10.1007/s11609-017-0324-8>
- Kalthoff, H. & Dittrich, T. (2017). Unterricht ohne Urteil? Zur Bewertungspraxis reformpädagogischer Schulen. In T.-S. Idel & H. Ullrich (Hg.), *Handbuch Reformpädagogik* (S. 305–323). Beltz.
- Kropf, J. & Laser, S. (Hg.). (2019). *Digitale Bewertungspraktiken: Für eine Bewertungssoziologie des Digitalen*. Springer VS. <https://doi.org/10.1007/978-3-658-21165-3>
- Krüger, A. K. (2022). *Soziologie des Wertens und Bewertens*. transcript. <https://doi.org/10.36198/9783838557229>
- Krüger, A. K. & Reinhart, M. (2016). Wert, Werte und (Be)Wertungen. Eine erste begriffs- und prozesstheoretische Sondierung der aktuellen Soziologie der Bewertung. *Berliner Journal für Soziologie*, 26(3–4), 485–500. <https://doi.org/10.1007/s11609-017-0330-x>
- Krüger, J. O. (2014). Vom Hörensagen. Die Bedeutung von Gerüchten im elterlichen Diskurs zur Grundschulwahl. *Zeitschrift für Pädagogik*, 60(3), 390–408. <https://doi.org/10.25656/01:14663>
- Kuhlmann, N. & Sotzek, J. (2019). Situierte (Geschichte der) Subjektivierung im Unterricht: Ein Gespräch zwischen adressierungsanalytischer und dokumentarischer Interpretation. In A. Geimer, S. Amling & S. Bosančić (Hg.), *Subjekt und Subjektivierung: Empirische und theoretische Perspektiven auf Subjektivierungsprozesse* (S. 113–142). Springer. https://doi.org/10.1007/978-3-658-22313-7_6

- Lindenhayn, N. (2013). Rangordnung/Raumordnung.: Die kompetitive Logik schulischer Praxis im 19. Jahrhundert. In M. Tauschek (Hg.), *Kulturen des Wettbewerbs: Formationen kompetitiver Logiken* (S. 269–288). Waxmann.
- Mehan, H. (1979). *Learning Lessons: Social Organization in the Classroom*. Harvard University Press. <https://doi.org/10.4159/harvard.9780674420106>
- Meier, F. & Peetz, T. (Hg.). (2021). *Organisation und Bewertung*. Springer VS. <https://doi.org/10.1007/978-3-658-31549-8>
- Meier, F., Peetz, T. & Waibel, D. (2016). Bewertungskonstellationen. Theoretische Überlegungen zur Soziologie der Bewertung. *Berliner Journal für Soziologie*, 26(3–4), 307–328. <https://doi.org/10.1007/s11609-017-0325-7>
- Meier, M. (2011). Die Praktiken des Schulerfolgs. In K. U. Zaborowski, M. Meier & G. Breidenstein (Hg.), *Leistungsbewertung und Unterricht: Ethnographische Studien zur Bewertungspraxis in Gymnasium und Sekundarschule* (S. 39–161). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-93218-7_3
- Meier, S. (2023). Leisten, Leistung und Leistungsbewertung – Auslegungen in pädagogischem Interesse. In S. Ruin & G. Stibbe (Hg.), *Sportdidaktik und Schulsport: Zentrale Themen einer diversitätssensiblen Fachdidaktik* (S. 171–186). Hofmann-Verlag.
- Meier, S., Haut, J. & Ruin, S. (2016). Leistung als Selbstverständlichkeit? Eine kritische Reflexion divergierender Leistungsverständnisse (im Sport) vor dem Hintergrund des Inklusionsdiskurses. *Zeitschrift für Inklusion*, 10(3). <https://www.inklusion-online.net/index.php/inklusion-online/article/view/379>
- Meier, S. & Ruin, S. (2015). Ist ein Wandel nötig? Körper und Leistung im Kontext von inklusivem Sportunterricht. In S. Meier & S. Ruin (Hg.), *Inklusion als Herausforderung, Aufgabe und Chance für den Schulsport* (S. 81–99). Logos-Verlag.
- Meister, N. & Hollstein, O. (2018). Leistung bewerten. In M. Proske & K. Rabenstein (Hg.), *Kompendium Qualitative Unterrichtsforschung* (S. 123–137). Verlag Julius Klinkhardt.
- Monitor Lehrerbildung. (2022). *Inklusiv lehren lernen – Alle angehenden Lehrkräfte müssen auf Inklusion vorbereitet werden: Policy Brief*. https://www.monitor-lehrerbildung.de/wp-content/uploads/2022/11/Monitor-Lehrerbildung_Inklusion_Policy-Brief-2022.pdf
- Mummelthey, S. (2024). Konstruktionen von Leistung als Differenzordnung im Sportunterricht. Erste Ergebnisse einer praxistheoretischen Diskursethnographie. In K. Rabenstein, C. Bräuer, D. Hülsmann, S. Mummelthey

- & S. Strauß (Hg.), *Differenzkonstruktionen in fachunterrichtlichen Kontexten: Forschungsansätze und Erträge zu Differenz(de)konstruktion aus Fachdidaktik, Erziehungswissenschaft und Diversitätsforschung*. (S. 39–55). Verlag Julius Klinkhardt.
- Mummelthey, S., Rabenstein, K., Droepe, T. & Hunger, I. (2023). Relationen von Leistung und Körper im Sportunterricht. Einblicke in ein diskursethno-graphisches Forschungsdesign. In B. Zander, D. Rode, D. Schiller & D. Wolff (Hg.), *Qualitatives Forschen in der Sportpädagogik: Beiträge zu einer reflexiven Methodologie* (S. 385–405). Springer VS. https://doi.org/10.1007/978-3-658-38038-0_17
- Ott, M., Langer, A. & Rabenstein, K. (2012). Integrative Forschungsstrategien – Ethnographie und Diskursanalyse verbinden. In B. Friebertshäuser, H. Kelle, H. Boller, S. Bollig, C. Huf, A. Langer, M. Ott & S. Richter (Hg.), *Feld und Theorie: Herausforderungen erziehungswissenschaftlicher Ethnographie* (S. 169–184). Verlag Barbara Budrich.
- Rabenstein, K. (2021). Leistung (be)werten. In J. Budde & T. Eckermann (Hg.), *Studienbuch pädagogische Praktiken* (S. 78–99). Verlag Julius Klinkhardt.
- Rabenstein, K., Idel, T.-S. & Ricken, N. (2015). Zur Verschiebung von Leistung im individualisierten Unterricht. Empirische und theoretische Befunde zur schulischen Leitdifferenz. In J. Budde, N. Blasse, A. Bossen & G. Rißler (Hg.), *Heterogenitätsforschung – Empirische und theoretische Perspektiven* (S. 241–258). Beltz Juventa.
- Reh, S., Bühler, P., Hofmann, M. & Moser, V. (2021). Einleitung. Prüfen, Testen, Auslesen und Zuweisen. Zum Inklusions-Paradox des Schulsystems. In S. Reh, P. Bühler, M. Hofmann & V. Moser (Hg.), *Schülerauslese, schulische Beurteilung und Schülertests 1880 – 1980* (S. 7–28). Verlag Julius Klinkhardt. h <https://doi.org/10.25656/01:22267>
- Reh, S. & Ricken, N. (Hg.). (2018). *Leistung als Paradigma: Zur Entstehung und Transformation eines pädagogischen Konzepts*. Springer VS.
- Reinders, H. (2012). *Qualitative Interviews mit Jugendlichen führen: Ein Leitfaden* (2. Aufl.). De Gruyter. <https://doi.org/10.1524/9783486717600>
- Reinold, M. & Kalthoff, J. (2021). Fair Play – ethische Grundlagen und pädagogische Potenziale. In S. Pietsch, G. Doll-Teppler, G. Fritz & M. Lämmer (Hg.), *Bildungspotenziale der Olympischen Idee* (S. 191–212). Academia. <https://doi.org/10.5771/9783896659552-191>
- Riepe, V. (2021). *Choreographien der Homogenisierung: Zur Verkörperung von Gleichheiten in der Grundschule*. transcript. <https://doi.org/10.1515/9783839455425>

- Rotter, C. & Bressler, C. (2022). Immer Ärger mit Noten? Bearbeitung von Herausforderungen der Leistungsbewertung. *Empirische Pädagogik*, 36(3), 327–342.
- Ruin, S. (2017). Vielfältige Körper? *German Journal of Exercise and Sport Research*, 47(3), 221–231. <https://doi.org/10.1007/s12662-017-0452-5>
- Scherler, K.-H. (2000). Messen und Bewerten. In P. Wolters, H. Ehni, J. Kretschmer, K.-H. Scherler & W. Weichert (Hg.), *Didaktik des Schulsports* (S. 167–186). Hofmann Verlag.
- Schiller, D., Rode, D., Zander, B. & Wolff, D. (2021). Orientierungen und Praktiken sportunterrichtlicher Differenzkonstruktionen. Perspektiven praxistheoretischer Unterrichtsforschung im formal inklusiven Grundschulsport. *Zeitschrift für Grundschulforschung*, 14, 67–81. <https://doi.org/10.1007/s42278-020-00098-o>
- Schmidt, R. (2017). Praxistheorie. In R. Gugutzer, G. Klein & M. Meuser (Hg.), *Handbuch Körpersozialologie: Band 1: Grundbegriffe und theoretische Perspektiven* (S. 335–344). Springer VS. https://doi.org/10.1007/978-3-658-04136-6_36
- Schröder, S. (2024). *Prüfung als pädagogisches Problem* [Dissertationsmanuskript]. Goethe-Universität, Frankfurt.
- Schürmann, V. (2023). Fairness außerhalb des Wettkampfsports – gemessen am Wettkampfsport. *German Journal of Exercise and Sport Research*, 53(3), 333–343. <https://doi.org/10.1007/s12662-022-00861-8>
- Stichweh, R. (2016). *Inklusion und Exklusion: Studien zur Gesellschaftstheorie* (2., erweiterte Aufl.). transcript. <https://doi.org/10.1515/9783839422946>
- Strauß, S. (2023). *Unterricht als Bewertungsregime.: Eine adressierungsanalytische Untersuchung von Bewertungspraktiken ohne Noten in der Sekundarstufe* [Dissertation]. Georg-August-Universität, Göttingen.
- Tillmann, K.-J. (2001). Leistungsbewertung und Zensierung im Fach Sport. *Pädagogik*, 53(2), 44–48.
- Wilm, G. (2021). *Geschlecht als kontingente Praxis im Sportunterricht: Eine videobasierte Praxeographie*. transcript. <https://doi.org/10.1515/9783839459515>
- Witzel, A. (2000). Das problemzentrierte Interview. *Forum Qualitative Sozialforschung*, 1(1). <https://doi.org/10.17169/fqs-1.1.1132>
- Wolters, P. & Lüsebrink, I. (2017). Unterrichtsforschung im Kontext aktueller sportdidaktischer Ansätze. In H. Aschebrock & G. Stibbe (Hg.), *Schulsportforschung: Wissenschaftstheoretische und methodologische Reflexionen* (S. 57–77). Waxmann.
- Zaborowski, K. U., Meier, M. & Breidenstein, G. (Hg.). (2011). *Leistungsbewertung und Unterricht: Ethnographische Studien zur Bewertungspraxis in Gymnasien*.

um und Sekundarschule. VS Verlag für Sozialwissenschaften. <https://doi.org/10.1007/978-3-531-93218-7>

Die Erfindung des modernen Fußballs

Praktiken des Vergleichens und Bewertens im Alcock's Football Annual, 1868 bis 1908

Clelia Minnetian & Tobias Werron

Einleitung

Der moderne Fußball entstand in der zweiten Hälfte des 19. Jahrhunderts. Seine zentralen Merkmale – wie translokal einheitliche Regeln; jährlich wiederkehrende, über die gesamte Saison ausgetragene Wettbewerbe (Ligasysteme, Pokalwettbewerbe usw.); Fußballverbände, die die Einhaltung der Regeln überwachen und Wettbewerbe organisieren; eine Sportpublizistik, die alle Wettbewerbe aufmerksam verfolgt und unablässig kommentiert – prägten sich zuerst in Großbritannien zwischen 1860 und 1900 aus. Die heute noch existierende englische »Football Association« (FA) wurde 1863 gegründet, die ersten Pokalwettbewerbe fanden in den späten 1860er- und frühen 1870er-Jahren statt, und 1887/88 wurde die »Football League« gegründet, die in ähnlicher Form ebenfalls noch immer existiert. Wie innovativ diese Merkmale sind bzw. waren, fällt im Vergleich mit älteren Formen des Fußballs auf, die in Großbritannien in großer Zahl existierten, aber bis ins mittlere 19. Jahrhundert hinein lokal verstreut waren, je eigenen Regeln folgten und keine regelmäßige wiederkehrenden Wettbewerbe kannten (Harvey, 2005; Taylor, 2008).

Wie sind diese neuen Strukturen entstanden und was kann eine Untersuchung ihre Entstehungsgeschichte zum Verständnis heutiger Vergleichs- und Bewertungspraktiken des Fußballs beitragen? Wir interessieren uns für diese Fragen im Rahmen eines Forschungsprojekts zur Institutionalisierung der Rankings, in dem wir untersuchen, wann Rankings – verstanden als quantitative, tabellarisch visualisierte und seriell veröffentlichte Praktiken des Leistungsvergleichs – in verschiedenen Gesellschaftsbereichen aufgetaucht

und institutionalisiert worden sind (zum Beispiel Ringel et al., 2021).¹ Unser empirisches Interesse ist breit angelegt und reicht von Rankings im Sport über Kunst und Wissenschaft bzw. Universitäten bis zur internationalen Politik. Im Vergleich dieser Felder kann der moderne Sport als eine Art Pionierkontext gelten, denn Ligatabellen und andere Leistungsstatistiken setzten sich im Diskurs mancher Sportarten schon im 19. Jahrhundert durch – und damit früher als in anderen Feldern (Ringel & Werron, 2020). Eine pauschale Zurechnung auf »den Sport« ist gleichwohl problematisch, weil verschiedene Sportarten unterschiedlichen Entwicklungspfaden folgten, sich insbesondere die Vergleichs- und Bewertungspraktiken zwischen den Sportarten unterschieden und selbst in die institutionelle Formation dieser Sportarten eingriffen. Die Präferenz für bestimmte Leistungsvergleiche – wie Rankings – ist daher nicht ohne Weiteres auf eine inhärente »Logik des modernen Sports« zu reduzieren. Vielmehr gilt es zu untersuchen, welche Praktiken des Leistungsvergleichs sich in unterschiedlichen Sportarten etablierten und wie sie sich in deren institutionelles Arrangement einfügten.

Wir sind solchen Fragen zunächst an der Geschichte des US-amerikanischen Baseballs nachgegangen, das in einem ähnlichen Zeitraum entstanden ist. Dabei zeigte sich, dass die Frage nach dem besten Team bzw. danach, wer als »Champion« gelten kann, ein entscheidender Treiber des öffentlichen Leistungsdiskurses war, denn sie motivierte die Entwicklung immer neuer Statistiken und Tabellen und – nicht zuletzt – die Einführung des Ligasystems in den 1870er-Jahren. Das Ligasystem, in dessen Rahmen die besten Clubs regelmäßig und wiederholt gegeneinander antreten, kam nicht nur dem Bedürfnis nach Identifikation des »Champions« entgegen. Es entsprach auch einem neuen, im Kern statistischen Leistungsverständnis, das Leistungsfähigkeit an konstanten – statt singulären – Erfolgen festmachte (Minnetian & Werron, 2021). Ligatabellen und andere statistische Praktiken des Leistungsvergleichs begleiteten also nicht lediglich den Wandel dieser Sportart, sie ermöglichten ihn und wurden zu einem integralen Bestandteil des neuen institutionellen Arrangements.

Vor dem Hintergrund dieser Forschungserfahrung möchten wir nun analoge Fragen an die frühe Geschichte des britischen Fußballs richten: Wel-

¹ Wir danken der Deutschen Forschungsgemeinschaft, ohne deren Unterstützung des Projekts »Zur Institutionalisierung der Rankings: Diskurskarrieren tabellarischer Leistungsvergleiche zwischen 1850 und 1980« die Datenerhebung für diesen Artikel sowie die gemeinsame Arbeit am vorliegenden Text nicht möglich gewesen wäre.

che Praktiken des Leistungsvergleichs und der Leistungsbewertung lassen sich hier beobachten, wie fügten sie sich in das institutionelle Arrangement des modernen Fußballs ein? Welche quantitativen Vergleichs- und Bewertungspraktiken kamen auf, wie korrelierten sie mit dem damals üblichen Leistungsverständnis bzw. formten dieses mit? Empirisch stützen wir uns primär auf eine jährliche Publikation, *Alcock's Football Annual*, das die frühe Geschichte des Fußballs durchgängig begleitete und an dem sich der Wandel zentraler Vergleichs- und Bewertungspraktiken repräsentativ ablesen lässt.

Der Beitrag ist wie folgt strukturiert: Wir stellen zunächst unsere Hauptquelle, das *Alcock's Football Annual*, kurz vor, erläutern ihren Entstehungskontext und begründen die Auswahl des empirischen Materials (1). Anschließend diskutieren wir zentrale Motive des in den Jahrbüchern sichtbar werdenden Leistungsdiskurses. Dabei geht es um vorherrschende Praktiken des Vergleichens und Bewertens von Clubs und Spielern sowie um die Identifizierung der wichtigsten Spiele bzw. Wettkämpfe (2), um das Schaffen von Historizität (3) sowie um die Entwicklung von Tabellen und anderen statistischen Vergleichs- und Bewertungspraktiken (4). Sodann thematisieren wir den Wandel der Wettbewerbsformate in diesem Zeitraum und analysieren, wie sich die Entwicklung der Vergleichs- und Bewertungspraktiken sowie der Wettbewerbsformate zueinander verhielten (5). Der Fall des US-amerikanischen Baseballs dient uns dabei gelegentlich als Vergleichsfolie, um die Besonderheiten des Fußballs genauer herauszuarbeiten. Wir schließen mit einer zusammenfassenden Charakterisierung der Rolle, die Praktiken des Leistungsvergleichs und der Leistungsbewertung in der Entstehungsgeschichte des britischen Fußballs spielten.

1. Aus den Annalen des modernen Leistungsvergleichs. Jahrbücher des britischen Fußballs im 19. Jahrhundert

Die Entwicklung des modernen Sports und des modernen Pressewesens gingen in der zweiten Hälfte des 19. Jahrhunderts Hand in Hand (Mason, 1986: 184–185; Mason, 1993, S. 3). Der mediale Fußballdiskurs setzt sich aus verschiedenen Publikationsformaten zusammen: zum einen aus der spezialisierten Sportpresse in Form von täglichen, wöchentlichen und monatlichen Sportzeitschriften – sowie der für den Fußball typischen *Evening Papers* am Abend des Spieltages – und den *Football Annuals*, die zum Saisonende erschie-

nen. Zum anderen begann auch die allgemeine Tagespresse zum Ende der zweiten Hälfte des 19. Jahrhunderts über Sport zu berichten (Mason, 1986).

Nach breiter Sichtung dieser Quellen möchten wir uns im Folgenden auf die *Football Annuals* – die Fußballjahrbücher – konzentrieren. Wir analysieren das *Alcock's Football Annual* im Zeitraum von seiner Ersterscheinung 1868 bis zur Etablierung der Liga (Einführung 1888) bis Mitte der 1890er-Jahre. Nach einer ersten Phase der Etablierung des Fußballs mit seinem zunehmenden Organisationsgrad, der Gründung der FA 1863 und der zunehmenden Vereinheitlichung der Regeln umfasst die von uns untersuchte Zeitspanne die Phase des Wachstums des Fußballs – inklusive zunehmender gesellschaftlicher Relevanz –, die Gründung regionaler *Associations*, die Entstehung der Cup-Wettbewerbe, bis hin zur Einführung und Etablierung der Liga. Die jährlich zum Saisonende erscheinenden Jahrbücher erhoben den Anspruch, einen Gesamtüberblick über den Stand des Fußballs und insbesondere über die aktuelle Saison zu bieten. Sie traten also bereits durch ihre Konzeption einen Schritt hinter einzelne Ereignisse des Fußballs zurück und boten eine resümierende Perspektive.

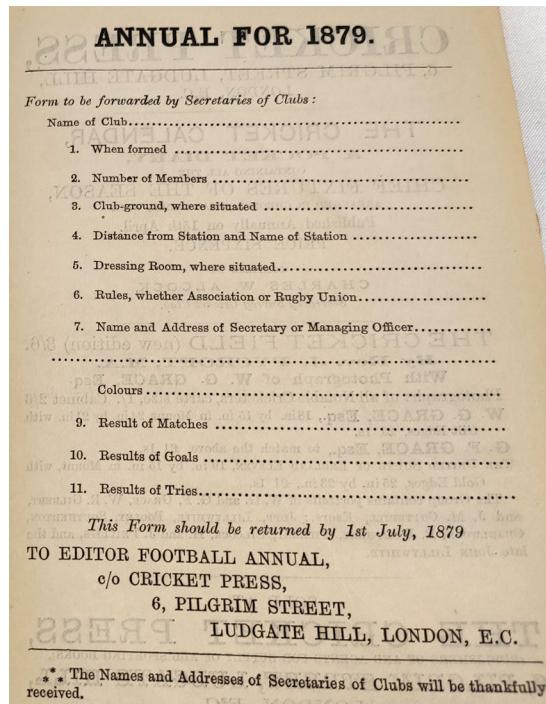
In diesen jährlichen Rückblicken nahmen Vergleichs- und Bewertungspraktiken einen prominenten Stellenwert ein, sie bieten daher für das hier verfolgte Forschungsanliegen einen besonders geeigneten Zugang. Zudem können die Jahrbücher als Ausdruck einer (Selbst-)Reflexion des Feldes verstanden werden. Die Einführung dieses Mediums in den späten 1860er-Jahren stand für die Entwicklung einer Community, die ein inklusives Verständnis ihrer selbst entwickelte und dieses auf spezifische Weise in einem Jahrbuch (re-)produzierte. Nicht untypisch für den Sportjournalismus dieser Zeit wurden die Jahrbücher von zentralen Figuren des Fußballs selbst herausgegeben, die ihre Community adressierten und zugleich als Journalisten ein wachsendes Publikum ansprachen. Man könnte nun davon ausgehen, dass die Nähe der Produzenten zum Feld eine Verzerrung der Berichterstattung etwa hinsichtlich der Kritikfähigkeit mit sich bringen würde. Eine solche Tendenz bestätigt sich aber vor dem Hintergrund anderer Quellen nicht; vielmehr prägten sich die Vergleichs- und Bewertungspraktiken in den Jahrbüchern auf für den Fußballdiskurs typische Weise aus.

Das hier untersuchte Jahrbuch *Alcock's Football Annual* erschien von 1868 bis 1908 und war das erste seiner Art im Fußball.² Es wurde von Charles W. Alcock herausgegeben und thematisiert die heute klar voneinander differenzierten Sportarten des Fußballs und des Rugbys, die damals aber noch zur gemeinsamen Football Community gehörten. Diese umfasste *Association Football* und *Rugby Football* sowie weitere Formen des Footballs mit zwar ähnlichen, aber leicht unterschiedlichen Regeln (wie etwa den *Sheffield Football*). Das Jahrbuch bot eine umfassende Sammlung von Informationen zum Fußball. Sein Anspruch war, die gesamte Community zu kartografieren und einen allgemeinen Überblick über die Geschichte des Sports sowie über die aktuellen Geschehnisse der Saison zu bieten. Die Sammlung beruhte unter anderem auf Einsendungen von den jeweiligen Organisationen (vor allem Public Schools und Clubs), wozu im Jahrbuch immer wieder aufgerufen wurde. Schulen und Clubs, die das Einsenden der Informationen versäumt hatten, wurden gelegentlich öffentlich dafür getadelt, so etwa in der Ausgabe von 1873: »Frequent applications were made to Eton College for statistics, but none have been received.« (*Alcock's Football Annual*, 1873, S. 25). Um solchen Versäumnissen entgegenzuwirken, wurde späteren Ausgaben ein standardisiertes Formular beigelegt (Abbildung 1).

Die Jahrbücher umfassten zu Beginn ca. 70 Seiten und wurden im Lauf der Jahre immer umfangreicher, bis ca. 160 Seiten lang. Sie behandelten die Dachorganisationen der *Football Association* und der *Rugby Union* mit ihrer jeweiligen Geschichte, ihrem Regelwerk und ihren Funktionären. Dazu gehörten auch allgemeine Beschreibungen des Spiels, Strategiehinweise und zeitgenössische Moralvorstellungen. Oft ähnelten sich diese Teile stark in den jeweiligen Jahresausgaben. Es wurden aber auch immer aktuelle Informationen über zur Community gehörende *Public Schools* und Clubs gesammelt und präsentiert, inklusive Spielerinformationen und Listen von ausgetragenen Spielen mit den jeweiligen Ergebnissen. Eine weitere Kategorie war eine Zusammenfassung der letzten Saison. Zu Beginn noch relativ knapp, wird diese immer ausführlicher und enthält Beschreibungen zu den zentralen Ereignissen und Entwicklungen sowie den als am wichtigsten eingestuften Spielen mit ihren Ergebnissen.

² Die erste Ausgabe erschien 1868 unter dem Titel *John Lillywhite's Football Annual*, und die letzten beiden Ausgaben 1907 und 1908 wurden nach Alcocks Tod von anderen herausgegeben.

Abbildung 1: Formular im Football Annual zur Informationseinreichung für Clubs



Quelle: Alcock's Football Annual, 1878, S. 157.

Das Jahrbuch ist in erster Linie eine englische Publikation und ist in London verortet. Das merkt man insbesondere daran, dass zwar Schottland mit abgedeckt wird, aber ein eigenes, eher nachgelagertes Kapitel erhält; später werden auch Fußball in Irland und Wales sowie Fußball im Ausland behandelt, womit insbesondere Fußball in den Kolonien gemeint ist. Auch wird an verschiedenen Stellen zwischen Fußball in London und in der Provinz unterschieden – eine Differenz, die sich mehrfach als signifikant für die Fußballgeschichte herausstellen sollte. Obwohl *Association Football* und *Rugby Football* in einigen Kapiteln getrennt behandelt werden, gibt es auch Kapitel, in denen sie in einer gemeinsamen Kategorie zusammengefasst werden, etwa wenn es um die *Public Schools* geht oder um die Sammlung aller *Football Clubs*. Dort scheint

der Inklusionsanspruch dominanter und/oder das Distinktionsmerkmal ein anderes zu sein.

Der Herausgeber Charles William Alcock (1842–1907) war eine zentrale Figur des frühen Fußballs. Er war aktiver Sportler sowohl im Fußball als auch im Cricket, besetzte zentrale administrative Positionen und war erfolgreicher Journalist und Autor. Er war Captain der Wanderers F.C., mit denen er in den ersten beiden Jahren 1872 und 1873 den FA-Cup-Titel gewann, und Captain des Englischen Teams in fünf Spielen gegen Schottland zwischen 1870 und 1872. Mit Gründung der FA 1863 wurde Alcock ins Komitee gewählt (bis 1866) und war ab 1870 für 25 Jahre *Honorary Secretary* und Schatzmeister der FA. In seine Amtszeit fielen unter anderem die Etablierung von Länderspielen (1872), die Einführung des FA Cups (1871–72), die Regelvereinheitlichung zwischen den Regelwerken der *Sheffield Association* und der *Football Association* (1877) sowie die Legalisierung des Professionalismus (1885). Neben seinen vielfältigen Rollen im Fußball spielte er auch Cricket bei führenden Clubs, wo er ebenfalls administrative Posten besetzte. Als Journalist arbeitete er in den späten 1860er-Jahren bei *The Sportsman* als Reporter für Leichtathletik, Fußball und Cricket sowie bei einer Cricket-Sportzeitung. Neben dem Fußballjahrbuch gab er auch eines für Cricket heraus und war Autor weiterer Bücher – unter anderem zum Fußball. Diese Verbindung zwischen dem Journalismus und der Sportcommunity war typisch für den frühen Fußball, als Fußballjournalisten meist als Enthusiasten des Sports galten (Tate, 2007, S. 18).

Alcock kann somit als Insider angesehen werden, der über seine Community und für seine Community schrieb. Daraus ergibt sich auch eine spezifische Form der Wissensvermittlung, die – wie wir zeigen werden – eine stark implizite Form aufwies. In den Jahrbüchern lief die Unterstellung stets mit, dass den Leser*innen die Fußballteams und ihre Rollen und Relevanzen in der Community vertraut waren. Die ausführliche Sammlung von Informationen in den Jahrbüchern präsentierte sich daher in der Regel weniger als explizite Vergleichs- und Bewertungspraxis denn als implizites *Angebot* an die Leser*innen, ihre eigenen Vergleiche und Bewertungen vorzunehmen. Allerdings konstruierte die Zusammenführung dieser Informationen bereits einen gemeinsamen Horizont einer Football-Community, zu der all diese gesammelten Elemente gehören: *Public Schools* und *Clubs*; *Rugby Football* und *Association Football*; und die Zusammenführung singulärer Spiele. Wir wollen nun genauer zeigen, wie sich im Laufe der Jahre die Vergleichs- und Bewertungspraktiken veränderten, neue eingeführt und weiterentwickelt wurden, und wie sich diese

Praktiken zum Wandel des institutionellen Arrangements – insbesondere der Wettbewerbsformate – verhielten.

2. Wissen um die »Besten«: eine inklusiv-multipolare Ordnung

Bei der Entwicklung des Baseballs hin zu einem professionellen Sport konnten wir sehen, dass die Suche nach angemessenen Kriterien zur Bestimmung des »Champions« und der besten individuellen Spieler einer der zentralen Treiber hin zur Entstehung des Ligasystems war. Das Ligasystem wurde dem neuen, im Kern statistischen Leistungsverständnis gerecht, das von den Besten konstante Leistungen über die gesamte Saison erwartete (Minnetian & Werron, 2021). Doch wie verhält es sich im Fußball des 19. Jahrhunderts? Wir zeigen in diesem Abschnitt anhand von drei Kategorien – den *Public Schools* und Clubs, den Spielern und den Spielen bzw. Wettbewerben –, ob und wie die »Besten« bestimmt, bewertet und verglichen wurden.

In den hier untersuchten Jahrbüchern vom Ende der 1860er- bis in die 1890er-Jahre findet man zunächst einmal eine Informationssammlung von *Public Schools* und Fußballclubs, die dem Inklusivitätsprinzip folgt: Möglichst alle Clubs (bei Verfügbarkeit von Angaben) werden aufgeführt, und zwar über alle Formen des Fußballs (inklusive Rugby) hinweg. Seit 1863 – infolge der Gründung der *Football Association* – gab es eine zunehmend stärkere Differenzierung zwischen den verschiedenen Fußballregeln (»Codes«), insbesondere zwischen Association und Rugby Football. Trotz dieser Entwicklung gibt es ein Verständnis, zu einer gemeinsamen Community zu gehören. Dies zeigt sich nicht nur hier in den Jahrbüchern, die ohne weitere Erläuterung beide Formen umfassen; auch in (Sport-)Zeitungen fungiert die Überschrift *Football* für beide bzw. alle Formen des Fußballs (so etwa in der zwischen 1822 und 1886 erscheinenden populären *Bell's Life in London, and Sporting Chronicle*).

Anscheinend gibt es dabei ein implizites Wissen von »guten« Teams, die sich in der Überschrift der Jahrbücher zur Aufzählung der Schulen und Clubs wiederfindet. So wird dort von »Chief Schools und Leading Clubs« gesprochen. Was dies genau bedeutete, welchen Kriterien diese Differenzierung folgte und wer alles dazu gehörte, wird nicht explizit erläutert. Dies lässt auf ein feldinternes implizites Wissen schließen sowie darauf, dass die Zielgruppe der Publikation in erster Linie eine Fußballcommunity war, bei der dieses Wissen vorausgesetzt werden konnte.

Generell kann hinsichtlich der Frage, »wer der Beste ist«, im Fußball der 1870er- und 1880er-Jahre – vor Einführung der Liga 1888 – festgehalten werden, dass sie nicht eine solch zentrale Rolle spielte wie im Baseball in den USA in einer vergleichbaren Entwicklungsphase der Professionalisierung des Sports. Man kann vielmehr von einer multipolaren, lokal und regional fragmentierten Ordnung sprechen, bei der es ein implizites Wissen um »gute« Teams gab, das Bemühen um diskursive Identifikation der »Besten« aber nicht im Mittelpunkt stand. Im frühen Fußball wurden primär Freundschaftsspiele (*Friendlies*) und Herausforderungsspiele (*Challenges*) gespielt, 1-gegen-1-Spiele, bei denen (nur) Gewinner dieses direkten Vergleichs ermittelt wurden. Diese Ergebnisse wurden im öffentlichen Diskurs *nicht* in einen Vergleichszusammenhang gestellt, um eine Hierarchisierung der Teams vorzunehmen. Die Ergebnisse standen jeweils für sich, so dass auch kein Vergleich aller Clubs vorgenommen werden konnte – wie dies im Baseball schon früh versucht wurde. Entsprechend zielte die Einführung der ersten Cup-Wettbewerbe (1867 Youdan Cup im *Sheffield Football*; 1872 FA Cup) nicht primär auf Ermittlung des besten Teams, sondern auf Aufmerksamkeitssteigerung für den Fußball insgesamt. Andererseits ist auch festzustellen, dass den Cup-Spielen mit ihrer Einführung sogleich hohe Relevanz zugeschrieben wird, insbesondere im Kontrast zu den nun zunehmend an Bedeutung verlierenden *Friendlies*.

Der Sieg eines Cupwettbewerbes ging durchaus mit Reputationsgewinnen einher, jedoch ist einschränkend festzuhalten, dass zu Beginn noch nicht sonderlich viele Teams am Cup teilnahmen (im ersten Jahr des FA Cups waren es 15) und noch keinesfalls von einem repräsentativen oder gar nationalen Wettbewerb gesprochen werden kann (Taylor, 2008: 41). Zudem nahm die Zahl der Cup-Wettbewerbe Ende der 1870er- und Anfang der 1880er-Jahre deutlich zu, mit einer entsprechend zunehmenden Zahl an Gewinnern, die im Fußballdiskurs nicht weiter hierarchisiert wurden (nur über den FA Cup wurde hervorgehoben berichtet). Im von uns gesichteten Material wird in dieser Zeit rückblickend auf eine Saison selten von »den besten« oder gar »dem besten« Team gesprochen. Man wusste stattdessen von guten Teams – den »Chief Schools und Leading Clubs«. Auch nach Einführung der Cup-Wettbewerbe ergibt sich also ein fragmentiertes, multipolares, wenig hierarchisiertes Bild. Auch spielte traditionell zugeschriebene Bedeutung innerhalb der Community, die nur bedingt durch Leistungen verdient werden musste, weiterhin eine wichtige Rolle. In diese Kategorie fielen renommierte *Public Schools* – der Herausgeber Alcock besuchte selbst eine –, und alle früh gegründeten Fußballclubs. Es han-

delte sich also auch oder sogar primär um eine Kategorie, die auf Seniorität (und Verfügbarkeit von statistischem Material) abstellte, weniger auf aktuelle Leistung. Dies veränderte sich allmählich im Laufe der Jahre und deutlich nach Einführung der Liga.

Eine zweite Ebene der Identifizierung der »Besten« – was auch hier in der Regel zunächst meint: der in der Community bekannten »Guten« – ist die Bewertung einzelner Spieler. Ähnlich wie bei den Clubs gab es eine Auflistung von »Chief Players« (Alcock's Football Annual, 1872, S. 68), wobei an manchen Stellen explizit auch auf deren Leistungsfähigkeit – »their merits« (Alcock's Football Annual, 1876, Inhaltsverzeichnis) – Bezug genommen wurde. Ein klares Differenzierungsmerkmal für herausragende Leistungsfähigkeit ist die Wahl in die Nationalmannschaft – »played in internationals« (ab 1879 wird auch von »caps³ gesprochen). Oft wurden Spieler durch einen vorab bestimmten *Captain* für die Nationalmannschaft ausgewählt, also durch einen Peer bestimmt. Manchmal findet sich eine erweiterte Liste von Spielern, deren Inklusion nicht weiter begründet wurde. Da es sich um ein vornehmlich englisches Medium handelte, wurden die englischen Spieler aufgelistet, während die Spieler anderer Nationen zum Teil in separaten Listen – insbesondere für Schottland – aufgeführt wurden. Diese Listen wurden wiederum alphabetisch geordnet, ohne weitere Hierarchisierung.

Die Spieler wurden darüber hinaus auch *qualitativen Bewertungen* unterzogen. Insbesondere wurden Spieler sowohl in der Kategorie der »Chief Schools« als auch bei den »Chief Players« narrativ in ein bis drei Sätzen beschrieben. Meist handelte es sich um Beschreibungen ihrer Stärken – manchmal auch in einer übersteigernden Form –, vereinzelt wurden auch kritische Punkte aufgeführt und Schwächen benannt.

Lubbock, E., (Eton,) West Kent. The finest kick anywhere, his aptitude for kicking the ball in the most difficult position being unequalled, would be better still was he faster. (Alcock's Football Annual, 1872, S. 69)

Nepean, C. E. (Charterhouse,) Oxford University. May be fairly classed as the best player of the day, works the ball with surprising skill, and never misses a shot at goal; can take any position with credit to himself. (Alcock's Football Annual, 1872, S. 69)

³ Dies bezieht sich auf die Praxis der Vergabe von Mützen (Caps) bei der Teilnahme an einem Länderspiel.

Eine weitere Bewertungspraxis war die Nennung von Torschützen bei Saisonstatistiken oder als Ergänzung zu Spielresultaten.

LONDON V. SHEFFIELD.—(3rd match) played at Sheffield, on Sat. Mar. 2, 1872, one half according to rules of Sheffield Association, and one half according to rules of the Football Association. Sheffield won by 2 goals (W. Horton and J. C. Clegg) to 1 (C. J. Chenery.) (Alcock's Football Annual, 1872, S. 42)

Die qualitativen Leistungsbewertungen folgten weder festgelegten Kriterien, noch gingen sie mit expliziten Vergleichen mehrerer Spieler einher. Eine klare Differenz erscheint nur zwischen Nationalspielern und Nichtnationalspielern bzw. einer zusätzlichen Kategorie, der erweiterten Liste von »Chief«-Spielern, wobei die Gründe für diese Differenzierungen nicht expliziert wurden und von den kundigen Lesern – den Mitgliedern der Football Community – »zwischen den Zeilen« herausgelesen werden mussten. Auch benannte der Diskurs gute Spieler nur in Form einer elitären Gruppe oder kategorisch erzeugten »Vielzahl« – im Gegensatz zu dem *einen besten* Spieler oder dem *einen besten Team*, wie man ihn/es etwa in einem Ranking hervorheben könnte. Kurz: Das Bewerten erfolgt explizit, das Vergleichen bleibt weitgehend implizit. Diese implizite, kategoriale und mit Blick auf die Begründung von Leistungsdifferenzierungen vage Vergleichspraxis unterscheidet sich deutlich vom Baseball, wo es in einer vergleichbaren Phase der institutionellen Stabilisierung in den 1860er- und 1870er-Jahren sowohl für Teams als auch für einzelne Spieler bereits klare quantitative Leistungsindikatoren – wie etwa den »batting average« (für Angriffsspieler/*Batter*) oder den »earned run average« (für Werfer/*Pitcher*) – gab.

Eine dritte Kategorie, die in der Bewertungs- und Vergleichspraxis des frühen britischen Fußballs eine Rolle spielte, ist die Hervorhebung der wichtigsten Spiele: die »important matches of the season« oder die »chief matches«. Mit diesen Begriffen wurden bestimmte Spiele als besonders relevant markiert und von den restlichen Spielen unterschieden. Die Bewertung erfolgte durch die *Nennung* – in den ersten Jahren der Veröffentlichung – und – spätestens ab 1876 – die *Beschreibung* einiger Spiele. Im ersten Jahrbuch wurden nur die »internationals« – die Länderspiele England gegen Schottland – und ein Regionalspiel – *North* gegen *South* genannt. Doch schon in den Folgejahren wurden auch *Inter-Association-* und *County-Matches*, Spiele zwischen traditionellen Rivalen (sogenannte *Old Rivalries*), zwischen herausragenden »leading« Clubs

sowie Cupspiele – insbesondere das Finale oder spätere Runden – hervorgehoben.

Der Umfang dieser Spielbeschreibungen nahm über die Jahre deutlich zu und bald die Form seitenlanger Aufzählungen an (1871 ca. eine Seite, 1878 ca. acht Seiten, 1885 ca. 17 Seiten).⁴ Eine besondere Bedeutung weisen erstens Spiele von Auswahlmannschaften auf, die zum einen an eine nationale oder regionale Identität anknüpfen und zugleich durch ihre leistungsorientierte Selektion der Spieler eine attraktive Begegnung versprechen. Zweitens wird an traditionelle Rivalitäten angeknüpft, etwa beim Spiel Oxford gegen Cambridge. Und drittens sind leistungsbezogene Spiele zwischen erfolgreichen Mannschaften sowie (Leistungs-)Wettbewerbe – die einzelne Spiele in einer übergeordnete »competition« integrieren (wie anfangs insbesondere Cup-Wettbewerbe) – von Bedeutung. Es spielen also sowohl lokale, regionale und nationale Kollektividentitäten und historisch gewachsene Rivalitäten als auch erste Tendenzen hin zu einer Integration aller Teams in einer übergreifenden Leistungshierarchie eine Rolle. In der Berichterstattung zu diesen Spielen werden oft auch die einzelnen Spieler aufgezählt, mit wechselseitig qualifizierendem Effekt: Gute Spieler verleihen den Spielen Relevanz, und zugleich dient die Wahl in eine Auswahlmannschaft – wie zuvor gezeigt – als Qualitätsmerkmal für die ausgewählten Spieler.

Die zugrunde liegende Praxis des Vergleichens blieb jedoch in allen diesen Fällen implizit: Wir finden explizite Bewertungen der Leistungsfähigkeit bestimmter Teams und Spieler sowie der Bedeutung einzelner Spiele, die auf Vergleichen beruhen (müssen), ohne dass die Vergleichskriterien benannt und begründet würden. Im Ganzen lässt sich das Resultat als eine *inklusiv-multipolare Leistungsordnung mit weitgehend impliziter Vergleichspraxis* bezeichnen. Es ging in diesem Diskurs offenbar nicht um die Erzeugung und Begründung eindeutiger Leistungshierarchien, sondern um die Inklusion aller Mitglieder einer Community und die weitgehend implizite Auszeichnung besonders leistungsfähiger Mitglieder. Eine solche Ordnung lässt explizite *Bewertungen* durchaus zu, unterdrückt aber zugleich explizite *Vergleiche*. Zugleich lädt sie das Lesepublikum dazu ein, eigene Vergleiche vorzunehmen oder eigene Vergleichskriterien in die Texte hineinzulesen. Diese explizite Bewertungspraxis mit implizitem Vergleichsangebot ist charakteristisch für die britische Fußballkultur der zweiten Hälfte des 19. Jahrhunderts und dominiert in den

4 Einiges an Platz nehmen zunehmend auch Historiendarstellungen ein, auf die wir in Kapitel 3 eingehen.

Jahrbüchern noch bis in die 1890er-Jahre hinein, als die Football League bereits existierte und der Fußball an der Spitze bereits weitgehend professionalisiert war.

3. Das Schaffen von Historizität

Das Jahrbuch als spezifisches Publikationsformat ist bereits selbst eine journalistische Praxis, die auf die Imagination von Historizität ausgerichtet ist. Es sammelt und archiviert Informationen für die Fußballcommunity und produziert dadurch ein schriftliches Gedächtnis, das für gegenwärtige und künftige Ereignisse einen Vergleichs- und Bewertungskontext zur Verfügung stellt (zur spezifischen Memory Culture des modernen Sports siehe Stauff, 2019; Werron, 2015). Dabei fallen zwei Entwicklungen besonders auf: erstens die Etablierung der Saison als übergreifender Bewertungskontext, zweitens das Aufkommen explizit historisierender Vergleichs- und Bewertungspraktiken. Beide seien im Folgenden an einigen Beispielen illustriert.

Von Publikationsbeginn an präsentieren die Jahrbücher unter dem Titel *Fixtures* (Abbildung 3) einen rückblickenden Spielkalender der jeweils porträtierten *Public Schools* und Clubs. Die *Fixtures* geben einen Überblick über alle gespielten Spiele, wobei die Zusammengehörigkeit dieser einzelnen Spiele zu einer jährlich wiederkehrenden *Season* (Saison) markiert wird. Die Saison als jährlicher Vergleichs- und Bewertungskontext ist zu diesem Zeitpunkt bereits bekannt, doch sind es solche Praktiken, die diese zeitliche Einheit reproduzieren und stabilisieren. Ebenfalls von Beginn an – allerdings nur vereinzelt – gibt es Saisonstatistiken, die alle gespielten Spiele umfassen. Dabei wird angegeben, wie viele Spiele gewonnen, verloren und unentschieden gespielt wurden, meist mit Angabe der erzielten und erhaltenen Tore.

Abbildung 2: Saisonstatistik über gespielte Spiele mit Torschützen

Matches played, 35—won, 20 ; lost, 6 ; drawn, 9. Goals obtained, 54 ; lost, 20. Of the 54 goals, C. W. Alcock obtained 17 ; T. C. Hooman, 6 ; A. F. Kinnaird, 4 ; A. W. Howard, 3 ; A. Borwick, 2 ; Hon. T. Pelham, 2 ; R. W. S. Vidal, 2 ; C. L. Huggins, 2 ; W. P. Crake, 2 ; and 1 each by D. Allport, A. Baker, F. Barnett, E. Bowen, C. J. Chenery, J. A. Cruikshank, E. Hall, Q. Hogg, A. M. Jones, C. F. Reid, H. Rushworth, R. S. Walker, and W. Wallace.

Quelle: Alcock's Football Annual 1871, S. 33.

Abbildung 3: Fixtures mit vorangestellter Saisonstatistik

76	
NOTTINGHAM CASTLE CLUB.	
Matches played,	16—won, 10 ; drawn, 6 ; Lost, 0. Goals obtained, 33 ; lost, 3.
1875.	
Nov. 6, — v. E. H. Greenhalgh's Team.	Won by 1 goal to nothing.
20, — v. Nottingham Lace Club.	Drawn ; 1 goal each.
Dec. 4, — v. Robin Hood Club, Burton-on-Trent.	Drawn ; no goals.
11, — v. Ashbourne.	Won by 3 goals to nothing.
27, — v. Grantham.	Drawn ; no goals.
1876.	
Jan. 1, — v. Grey Friars.	Won by 4 goals to nothing.
8, — v. Mansfield.	Won by 8 goals to nothing.
15, — v. Robin Hood Club, Burton-on-Trent.	Drawn ; no goals.
22, — v. Ashbourne.	Won by 1 goal to nothing.
29, — v. Long Eaton.	Won by 4 goals to nothing.
Feb. 5, — v. Long Eaton.	Won by 1 goal to nothing.
19, — v. Nottingham Lace.	Won by 3 goals to nothing.
29, — v. Mansfield.	Drawn ; no goals.
21, — v. Sheffiled Wednesday Club.	Drawn ; 1 goal each.
Mar. 7, — v. Grantham.	Won by 3 goals to nothing.
25, — v. Mr. E. H. Greenhalgh's Team.	Won by 3 goals to 1.

Quelle: Alcock's Football Annual 1876, S. 76.

Diese Saisonstatistiken verbreiteten sich nach und nach und wurden ab 1876 den *Fixtures* nicht mehr nach-, sondern vorangestellt (was einen Bedeutungsgewinn vermuten lässt). Fast alle Clubs führen nun eine solche Statistik, auch bei den Public Schools kommen sie häufiger vor. Ab 1873 wurden sie im Kapitel der Sammlung aller Clubs am Ende jedes Jahrbuches (bei vorliegender Information) angeführt (siehe Abbildung 7 weiter unten) und ergänzten damit die ansonsten eher administrativen Informationsangaben (zu Gründungsjahren, Spiellokalität usw.) um leistungsbezogene Kategorien. Die Saisonstatistiken erweiterten den Bewertungshorizont vom einzelnen Spiel auf die Saison und legten diese als maßgebenden Zeitraum für die Bewertung der Gesamtleistung fest. Erneut lässt sich hierin ein – implizites – Angebot sehen, die verschiedenen Teams miteinander zu vergleichen, erneut jedoch ohne explizite Vergleiche vorzunehmen. Das Vergleichen wird so ermöglicht, aber auch erschwert – zum einen dadurch, dass die Informationen im Jahrbuch verstreut zu finden sind, aber auch dadurch, dass sich die Voraussetzungen für die Leistungserbringung zwischen den Teams markant unterscheiden, etwa durch die unterschiedliche Anzahl an absolvierten Spielen und die Unterschiedlichkeit der Gegner. Hinzu kommt, dass diese Statistiken – ebenso wie die Übersichten über die *Public Schools* und *Leading Clubs* – nicht zwischen *Association Football* und *Rugby Football* unterschieden (bei Rugby spielenden Vereinen sind nur weitere »Score«-Angaben vorhanden, die sich auf erzielte Tries beziehen), so dass sich die Ergebnisse nicht immer eindeutig einer spezifischen Fußball-Spiel-

form zuordnen lassen. Auch hier dominieren offenbar das Inklusionsprinzip und der Versuch, der Footballcommunity einen Überblick über das Gesamtgeschehen zu verschaffen und nur wenig zwischen den Mitgliedern zu diskriminieren.

Eine komplementäre Praxis zu jener der *Fixtures* ist 1886 die Einführung eines Fußballkalenders, der die rückblickende Perspektive der *Fixtures* um eine zukunftsgewandte ergänzt. Es handelte sich um einen Gesamtkalender aller wichtigen Spiele der bevorstehenden Saison, der ebenfalls der oben beschriebenen Relevanzhierarchie folgte: *Internationals*, *Inter-Association-Matches*, Cup-Spiele und Spiele herausragender Mannschaften. Die Zusammenführung in einer Liste stabilisierte den Eindruck einer – mit Highlights gespickten – Fußballsaison und affinierte zugleich die Zugehörigkeit aller dieser Clubs und Ereignisse zu einer übergreifenden Community.

Damit kommen wir zum zweiten Aspekt der Herstellung von Historizität, den historisierenden Vergleichs- und Bewertungspraktiken. Von Mitte der 1870er-Jahre an ist eine konzeptionelle und grafische Neuerung zu beobachten: Ergebnisse von einzelnen Spielen wurden in einer Reihe »gleicher« Spiele zusammengefasst und somit eine je eigene Historie für bestimmte Typen von Spielen konstruiert. Eine erste solche Darstellung bezieht sich 1876 auf die bisherigen Gewinner des FA Cups.⁵ Das aktuelle FA-Cup-Finale wird somit nicht nur als einzelnes Ereignis gewürdigt, es wird zugleich eingereiht in die Geschichtlichkeit dieses Wettbewerbs. Mehrfache Gewinner – hier die *Wanderers* – werden zudem mit einer solchen tabellarischen Darstellung implizit hervorgehoben: Ihre Leistung – in fünf Jahren drei Mal den FA Cup gewonnen zu haben – übertrifft jene der Oxford University und der Royal Engineers, die je nur einmal gewannen.

Abbildung 4: Gewinner des FA Challenge Cup;

FINAL TIE.	
Wanderers beat Old Etonians by three goals to none, after a drawn game, 0-0.	The winners of the Cup have been:
1872 Wanderers	1875 Royal Engineers
1873 Wanderers	1876 Wanderers
1874 Oxford University	

Quelle: Alcock's Football Annual 1876, S. 41.

5 Die beiden Jahrgänge davor – 1874 und 1875 – fehlten im Archiv des National Football Museums in Preston. 1873 war noch keine solche Historie vorhanden.

Solche historisierenden Darstellungen wurden ab 1879 häufiger und ausführlicher. Zumeist bezogen sie sich auf renommierte Spiele, analog zu den oben diskutierten »chief matches«, das heißt Begegnungen von Nationen, *Associations* oder prominenten Clubs. Aufgrund der wachsenden Zahl an Begegnungen nahm zudem ihr Umfang zu, wie hier an dem regelmäßig stattfindenden Duell zwischen den beiden *Associations* London und Sheffield zu sehen ist.

Abbildung 5: London v. Sheffield – Resultate vergangener Spiele

RESULTS OF PAST MATCHES.	
1871 (Dec. 2).	—Sheffield (Sheffield Rules). Sheffield won by 3 goals to 1.
1872 (Jan. 27).	—London (Association Rules). London won by 1 goal to none.
1872 (Mar. 2).	—Sheffield (mixed Rules). Sheffield won by 2 goals to 1.
1872 (Nov. 2).	—Sheffield (Sheffield Rules). Sheffield won by 4 goals to 1.
1873 (Jan. 4).	—London (Association Rules). Drawn; no goals to either.
1873 (Mar. 15).	—Sheffield (mixed Rules). Sheffield won by 2 goals to 1.
1873 (Nov. 11).	—Sheffield (Sheffield Rules). Sheffield won by 8 goals to 2.
1874 (Jan. 3).	—London (Association Rules). Drawn; 1 goal to each.
1874 (Apr. 4).	—Sheffield (mixed Rules). Sheffield won by 4 goals to 2.
1874 (Nov. 7).	—Sheffield (Sheffield Rules). Sheffield won by 2 goals to none.
1875 (Jan. 16).	—London (Association Rules). London won by 3 goals to 1.
1875 (Mar. 29).	—Sheffield (mixed Rules). London won by 2 goals to none.
1876 (Jan. 1).	—Oval (London Rules). London won by 4 goals to none.
1876 (Mar. 25).	—Sheffield (Sheffield Rules). Sheffield won by 6 goals to none.
1876 (Nov. 11).	—Sheffield (Sheffield Rules). Sheffield won by 5 goals to 1.
1876 (Dec. 30).	—Oval (London Rules). London won by 3 goals to 1.

Quelle: Alcock's Football Annual 1877, S. 34.

Die Bewertungspraxis insgesamt nahm mit solchen Darstellungen immer deutlicher eine historisierende Dimension an. Begegnungen zweier Teams wurden zu einer seriellen Reihe, die sich über eine Jahr für Jahr anwachsende Zahl von Spielzeiten erstreckte. Die Leistungsfähigkeit der Teams konnte nun nicht mehr nur an einzelnen Ergebnissen abgelesen werden, sondern an allen archivierten und wiederholt abgedruckten Ergebnissen – auch wenn inzwischen ganz andere Spieler für diese Clubs aktiv waren. Vergangene Spiele wurden durch ihre stete Wiederholung aktualisiert und im Gedächtnis präsent gehalten; neue Spiele wurden in diesen historischen Kontext eingeordnet und erhielten vor diesem Hintergrund neue, »historische« Bedeutung.

Diese Historisierungstendenz nahm mit der Einführung der Liga 1888 noch einmal erkennbar zu. Wie bei der historisierenden Darstellung der Cup-Gewinner (Abbildung 4) wurden die Leistungen der Clubs über mehrere Jahre hinweg dargestellt, wobei im Unterschied zu den Cup-Statistiken – die auf die Gewinner fokussieren – mehrere oder alle Liga-Teilnehmer dargestellt und nun auch explizit im Rahmen einer tabellarischen Darstellung nach den

gleichen Leistungskriterien miteinander verglichen wurden. So sieht man in der ersten Tabelle der Abbildung 6 die Clubs auf den jeweils ersten drei Positionen der *Football League* von 1888–89 bis 1892–93. Die Tabelle spannte damit einen längerfristigen Vergleichshorizont der drei Erstplatzierten auf. Das Team von Preston North End schnitt in dieser Darstellung besonders gut ab, da es zwei Mal Erster und drei Mal Zweiter geworden war. Die zweite Tabelle der Abbildung 6 steigert die Anzahl der Bewerteten noch weiter und ermöglicht eine vergleichende Bewertung aller Teams der *Football League* über den Verlauf mehrerer Saisons. Alle Teams waren in dieser Tabelle mit ihren Punkten pro Saison repräsentiert. Die Reihenfolge richtet sich – wie von Beginn an typisch für Ligatabellen – nach den erzielten Punkten der aktuellen Saison, doch dahinter sind historisierend auch die erzielten Punkte der Vorsaisons abgebildet. Auch ein Vergleich zwischen den Teams hinsichtlich ihrer Leistungen über den gesamten Zeitraum, um etwa das Beste oder die besten Teams zu identifizieren, wurde dadurch möglich, wenn auch durch die Vielzahl der Angaben und Diskontinuitäten (bezüglich Ligazugehörigkeit) zugleich wieder erschwert.

Abbildung 6: Historien der Liga

The following will shew the three leading clubs in each season since the formation of the League:—					
First.	Second.	Third.			
1888-89 Preston N. End	Aston Villa	Wolverhampton Wandrs.			
1889-90 Preston N. End	Everton	Blackburn Rovers			
1890-91 Everton	Preston North End	(Notts			
1891-92 Sunderland	Preston North End	(Wolverhampton Wand.			
1892-93 Sunderland	Preston North End	Bolton Wanderers			
		Everton			

Club.	1892-3	1891-2	1890-1	1889-90	1888-89
Sunderland	42	42	23	—	—
Preston North End	37	37	27	33	40
Everton	36	22	29	31	20
Aston Villa	35	30	12	19	29
Bolton Wanderers	32	36	25	19	22
Burnley	30	26	21	13	17
Stoke	29	14	—	10	12
Blackburn Rovers	29	26	24	27	26
West Brom. Albion	29	12	12	25	22
Notts Forest	22	—	—	—	—
Wolverhampton W.	22	26	26	25	22
Sheffield Wednesday	27	—	—	—	—
Derby County	27	—	—	—	—
Notts	24	24	15	21	16
Accrington	24	26	26	17	12
Newton Heath	23	20	16	24	20
Darwen	12	—	—	—	—
Highest possible	—	12	—	—	—
	60	52	44	44	44

Quelle: Alcock's Football Annual 1893, S. 49.

Zusammenfassend hielt mit diesen neuen Praktiken eine historisierende Dimension Einzug in den Leistungsdiskurs des Fußballs. Nicht alle diese diskursiven Praktiken sind per se Vergleichs- und Bewertungspraktiken: So nahmen etwa die *Fixtures* aus der vergangenen Saison oder die zukunftsgerichtete Fußballkalender an sich keine Bewertung der und keinen Vergleich zwischen den Clubs und Spielern vor. Gleichwohl sind sie für das Verständnis neuer Bewertungs- und Vergleichspraktiken wichtig, denn sie konstituierten die Saison als zeitlichen Bezugshorizont für Vergleiche und Bewertungen aller Art. Diesen Bezugshorizont machten sich saisonübergreifende Statistiken zunutze, die Leistungen als Resultat aller gespielten Spiele einer Saison definierten. Die zweite Innovation waren Listen und Tabellen, die Leistungen explizit historisierten, indem sie Spiele und Wettbewerbe in Serien gruppierten und Leistungen über mehrere Saisons hinweg miteinander in Beziehung setzten. Durch solche Praktiken entstand ein Gedächtnis des Fußballs, das vergangene Ereignisse aufzurufen und neue in deren Kontext zu verorten erlaubte.

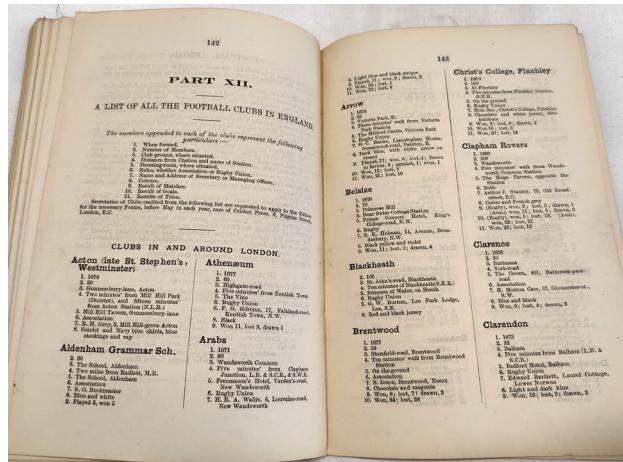
4. Tabellen und statistische Kategorien

Die Einführung und allgemeine Verwendung von Tabellen brachte nicht nur eine grafische Veränderung, sie erweiterte auch das Repertoire der Vergleichs- und Bewertungspraktiken. Im Folgenden wollen wir zum einen auf einige prägnante Beispiele eingehen, die zeigen, wie die Einführung von tabellarischen Formaten die Vergleichs- und Bewertungspraxis transformierte; zum anderen analysieren wir die Einführung neuer statistischer Vergleichskategorien und -kriterien, die im Kontext dieser Praxis aufkamen.

Als erstes Beispiel betrachten wir die bereits oben angeführte Liste von Clubs unter Angabe grundlegender Informationen, die sich am Ende jedes Jahrbuches befand. Von 1871 bis 1880 wurden alle Clubs nacheinander in einer »kästchenförmigen« Sammlung dargestellt (Abbildung 7). Diese beinhaltete Angaben zur Zugehörigkeit zu einem der beiden großen Regelwerke (Association oder Rugby), Clubname, Gründungsjahr, Ortsangabe der Umkleidekabinen, Vereinsfarben, Adresse des Sekretärs sowie die leistungsbezogenen Angaben zu den gespielten Spielen inklusive der gewonnenen, unentschieden gespielten und verlorenen Spiele, der erzielten und erhaltenen Tore sowie der erzielten und erhaltenen Tries im Falle von Rugby spielenden Vereinen. Diese Sammlung diente der Kartografierung des Feldes, sollte einen Gesamtüberblick ermöglichen und spezifische Informationen zu den Vereinen geben,

etwa um Kontakt aufzunehmen oder die Anreise zu gewährleisten. 1881 wird diese Liste dann in die Form einer Tabelle überführt (Abbildung 8).

Abbildung 7: Liste aller Football Clubs in England vor dem Tabellenformat 1881, hier 1879 (Auszug)



Quelle: Alcock's Football Annual 1879, S. 142–143.

Das Tabellenformat enthält weiterhin die gleichen Informationen: Noch immer werden die Clubs alphabetisch sortiert, und weiterhin werden – dem Inklusivitätsprinzip folgend – alle Clubs der verschiedenen Spielformen zusammengeführt. Und doch präsentiert die veränderte Darstellungsform ein *neues implizites Vergleichsangebot*. Vergleiche werden nicht explizit vorgenommen, denn weder wird der Inhalt der Tabelle in einem begleitenden Text erläutert noch findet ein Ranking nach Leistungskriterien – unter Erzeugung einer entsprechenden Leistungshierarchie – statt. Aber durch die grafisch-räumliche Zusammenführung dieser Leistungskategorien wird das In-Beziehung-Setzen vereinfacht, ja nahegelegt. Als Leser*in ertappt man sich dabei, die verschiedenen Vereine hinsichtlich ihrer Leistungskategorien miteinander zu vergleichen und nach besonders erfolgreichen Clubs zu suchen. Obschon Tabellen hier also nicht explizit als Vergleichsinstrumente genutzt werden, laden sie gleichwohl zum Vergleichen ein. Sie tragen damit, so ließe sich vermuten, über die Jahre zur Institutionalisierung eines *impliziten Vergleichsangebots*.

gleichswissens bei, das später – nach Einführung der Liga – dann nicht völlig neu erfunden, sondern nur noch in explizites Wissen transformiert werden musste.

Abbildung 8: Liste aller Football Clubs in England in Tabellenformat 1881 (Auszug)

PART XII.
—
A LIST OF THE CHIEF FOOTBALL CLUBS.

CLUBS IN AND AROUND LONDON.

NAME.	CLUB.	GROUND.	DRESSING ROOM.	COLOURS.	ADDRESS OF HON. SEC.	MATCHES.	WINS.	DRAWS.	LOSSES.	GOALS FOR.	GOALS AGAINST.	POINTS.
a Action	74 Greenwich-lane.	Kings Head	Scarlet & dark blue	R. H. Green, 6, Mill-grove,	21	9	6	10	10	71	—	—
n Arrow	74 Greenwich-lane.	Down Hotel	Navy blue	G. J. Banks, Ivy Bank, New-	16	8	2	6	—	—	—	—
n Arab	71 Wandsworth	Premises	Light blue and black	H. P. Ford, 10, Highfield-lodge,	16	6	9	3	7	—	4	3
a Albermarle Grammar-School	The School	The School	Orange and black	J. M. A. Thompson, Aldershot-School, Midhurst	15	9	2	—	—	—	—	—
a Belmont	77 Clapton	The School	Yellow and black	J. B. Choute, Belmont Home, Chalcots, New-Brentwood	30	20	4	6	—	—	—	—
a Birkenhead	green	—	Chocolate & magenta	H. French, 35, Newgate-street	17	7	4	6	22	22	—	—
a Burlington	73 Stamford-road	On ground	Violet and black	H. French, 35, Newgate-street	17	5	3	1	—	—	—	—
a Buttresses	73 Eton and Mid-Prince Consort	—	Chocolate & magenta	H. French, 35, Newgate-street	17	5	3	1	—	—	—	—
a Buffaloes	70 Primrose-hill	Chalk Farm	Chocolate & yellow	G. Green, 10, Park-oval, N.W.	10	6	2	2	10	13	5	—
n Belize	77 Eton and Mid-Prince Consort	Hotel	blue, yellow and black	H. French, 35, Newgate-street, 188, Adelphi	17	10	5	2	20	9	7	6
n Bellingeoke	70 Wandsworth	Hotel	Green and gold	O. Blakely, 25, Wandsworth-road, 12	18	8	2	8	9	21	4	—
n Beoley United	79 Beoley	Kings Head	Dark blue	T. J. Wardsworth, Elm Cross, 14	9	1	4	—	—	—	—	—
a Barnes	93 Barnes	White Hart	Blue and white	E. Eyes, Elm Cross,	14	9	1	4	—	—	—	—
n Chelsea	93 Battersea-park	White Hart	Blue and yellow	F. J. Jones, 10, Newgate-street, W.C.	19	5	2	1	28	29	—	—
n Croydon	76 Whiteladies-Green	Hotel	Blue and white	H. F. Woods, 129, Broad-street, Croydon	17	5	3	1	—	—	—	—
n Clayton	— Pond-lane	Tavern	Brown and white	H. French, 10, North End, Croydon S.E.	15	8	1	1	11	14	15	—
n Cleveland	80 Stamford-road, Bully-head	Black	Blue and white	C. Clayton, 10, Lord Clyde, Croydon S.E.	10	8	—	3	—	—	—	—
n Carlton United	74 Lambeth	Spread Eagle	Black	Clayton, 10, Lord Clyde, Croydon S.E.	9	3	1	1	—	—	—	—
n Cleveland	79 Brixton	Plough, Lewis	Blue and red	F. Clay, 21, Kepham-road, Croydon S.E.	17	5	2	10	10	13	17	—
n Christ's College, Finchley	70 Finchley	On ground	Chocolate and white	A. L. Chapman, 250, Lewisham High-road, London, S.E.	23	8	6	9	12	10	25	12
n Dulwich College	55 All-At-College	On ground	Blue and black	H. French, F.C., Dulwich Forest Gate	11	3	2	6	—	—	—	—
a Downs	78 Ley Bridge-rd., Down Hotel	—	Navy blue and white	W. R. Davies, 11, Queensdown-Hill, London, S.E.	27	10	1	7	75	44	—	—
a Dreadnought	75 Clapton	Upton Manor	Navy blue and white	F. S. Hayes, 21, Windsor-road, Epsom	19	6	4	5	34	17	—	—
a Excelsior	70 Battersea-park	Tavern	Blue and black	G. French, 1, St. George's-ter-ace, Kennington, S.W.	13	7	3	4	21	32	—	—
n Eaton Rovers	69 Battersea-park	Albert Tavern	Dark and light blue	N. L. Jackson, 23, Barton-street, S.W.	19	6	4	9	—	—	—	—
n Epsom College	27 At College	At College	Navy blue and white	H. French, 11, Queensdown-Hill, London, S.E.	11	5	1	5	—	—	—	—
n Ealing	72 Ealing College Bell Inn	—	Chocolate and white	W. D. Sims, Cheltenham-villa, Somerset-road, Ealing	—	—	—	—	—	—	—	—
a Finchley	76	—	Red and Black	N. L. Jackson, Long-lane, Chancery-lane, Finchley	—	—	—	—	—	—	—	—

Quelle: Alcock's Football Annual 1881, S. 176–177.

Unser zweites Beispiel ist die Ligatabelle von 1888. Mit Einführung des Ligasystems erscheint auch das Darstellungsformat der Ligatabelle, die bereits im Sportdiskurs um den historischen Vorläufer – der Baseballliga in den USA – entwickelt und seit den 1870er-Jahren regelmäßig verwendet wurde.

Abbildung 9: Ligatabelle, Clubs gerankt nach Punkten; Sieg = 1

Punkt, Unentschieden = $\frac{1}{2}$ Punkt

The positions of the various Clubs who formed the League last year will be gathered from the following table:							
	Played.	Won.	Lost.	Drawn.	Points.	Goals for	Goals agst.
1 Preston North End.	22 ..	18 ..	0 ..	4 ..	20 ..	74 ..	15
2 Aston Villa.....	22 ..	15 ..	6 ..	5 ..	14½ ..	61 ..	43
3 Wolverhampton W.	22 ..	12 ..	6 ..	4 ..	13 ..	51 ..	37
4 Blackburn Rovers.....	22 ..	10 ..	6 ..	9 ..	13 ..	66 ..	45
5 { Bolton Wanderers.	22 ..	10 ..	10 ..	2 ..	9 ..	11 ..	63 ..
West Brom. Albion.....	22 ..	10 ..	10 ..	2 ..	11 ..	40 ..	46
6 Accrington.....	22 ..	6 ..	8 ..	8 ..	10 ..	48 ..	48
7 Everton.....	22 ..	9 ..	11 ..	2 ..	10 ..	35 ..	47
8 Burnley.....	22 ..	7 ..	12 ..	3 ..	8½ ..	42 ..	62
10 Derby County.....	22 ..	7 ..	13 ..	2 ..	8 ..	41 ..	61
11 { Notts County.....	22 ..	5 ..	15 ..	2 ..	6 ..	40 ..	73
Stoke.....	22 ..	4 ..	14 ..	4 ..	6 ..	26 ..	61

* * A draw counts as half a win.

Quelle: Alcock's Football Annual, 1889, S. 124.

Die Liga als Wettbewerbsformat führte einen neuartigen Kontext für das explizite Vergleichen von Leistungen ein. Bei der Durchsetzung dieser expliziten Vergleichspraxis spielten Tabellen eine zentrale Rolle. Durch sie können sowohl der tagesaktuelle Stand während der Saison als auch das finale Ergebnis am Saisonende auf effiziente und eindeutige Weise dargestellt werden, indem die einzelnen Clubs durch Zuordnung von Punkten anhand der Ergebnisse aller gespielten Spiele *bewertet* und zugleich in eine hierarchische Reihenfolge gebracht und also explizit *verglichen* werden. Heute erscheint uns diese Form selbstverständlich, doch gab es im britischen Fußball – im Unterschied zum amerikanischen Baseball – vor Einführung der Liga keinerlei Darstellungen dieser Art. Selbst mit Bezug auf die Cup-Wettbewerbe, die ja bereits in den 1870er-Jahren ein saisonübergreifendes Wettbewerbsformat zur Verfügung stellten, wird über die Mitteilung der Ergebnisse hinaus kein Versuch unternommen, die teilnehmenden Mannschaften expliziten tabellarischen Vergleichen zu unterziehen. Meist wird nur der Sieger genannt. Es handelt sich also um eine neue Form der Ergebnisdarstellung, die alle Teilnehmenden umfasst, sie bewertet und sie vergleichend zueinander in Beziehung setzt.

Interessanterweise erscheinen 1888 – also im Jahr der Einführung der Liga – auch Tabellen in Bezug auf *Inter-Association-Matches* (Abbildung 10). Darin werden alle *Associations* unter Angabe ihrer gespielten Spiele und deren Ergebnisse im üblichen Format (played, won, drawn, lost, goals won, goals lost) dargestellt. Getrennt wurden sie nach einer regionalen Kategorie: den *Northern Associations* und den *Southern Associations* (obwohl diese auch gegeneinander spielten, etwa in dem prominenten Spiel von Sheffield gegen London). Wonnach sie in diesem Fall geordnet sind, ist hier unklar. Die übliche alphabeti-

sche Ordnung ist in diesem Fall nicht festzustellen; aber – und das ist ein entscheidender Unterschied zur Ligatabelle – es findet auch kein Ranking anhand leistungsbezogener Kategorien statt. Es scheint jedoch, als sei bereits die Darstellung in der neuen tabellarischen Form ein Ausdruck für die Relevanz der *Associations* und der *Inter-Association-Matches* als Wettbewerbskontext; denn obwohl Tabellen und Listen inzwischen häufiger Verwendung fanden, waren sie doch zumeist für besonders relevante Inhalte des Diskurses reserviert. Diese anhaltende Relevanz der *Associations* ist typisch für die Entwicklung des englischen Fußballs und wird auch nach Einführung der Liga insbesondere für den Amateursport weiter von Bedeutung sein.

Abbildung 10: Tabelle der *Associations*

The following table from <i>Pastime</i> will show the results of the matches played by the different Associations last winter:—								
NORTHERN ASSOCIATIONS.								
	Played.	Won.	Drawn.	Lost.	Goals won.	Goals lost.		
Staffordshire	...	4	3	1	24	4		
Birmingham	...	5	4	0	32	7		
Lancashire	...	2	1	0	4	6		
Derbyshire	...	2	1	0	3	5		
Sheffield	...	2	1	0	6	15		
Liverpool and District	2	1	0	1	6	3		
Lincolnshire	...	3	2	0	8	3		
Nottinghamshire	...	2	0	1	2	11		
Cleveland	...	4	2	0	9	9		
Northumberland	3	1	0	2	7	11		
Durham	...	4	1	0	3	17		
Cheshire	...	3	0	0	5	15		
Walsall	...	1	0	0	2	3		
Scarborough & E.Riding	1	0	0	1	2	4		

SOUTHERN ASSOCIATIONS.								
	Played.	Won.	Drawn.	Lost.	Goals won.	Goals lost.		
London	...	5	2	1	9	10		
Surrey	...	6	2	3	14	13		
Sussex	...	8	2	0	10	26		
Kent	...	2	1	0	7	2		
Northants	...	5	2	1	9	13		
Bucks and Berks	...	6	2	1	10	12		
Middlesex	...	3	2	1	8	8		
Essex	...	4	1	0	5	8		
Somerset	...	3	1	1	20	13		
Hampshire	...	6	4	2	9	8		
Wiltshire	...	6	6	0	19	5		
Dorsetshire	...	5	0	3	4	6		
Suffolk	...	8	4	2	15	8		
Norfolk	...	6	1	4	7	13		
Cambridgeshire	10	6	1	3	24	15		
Hertfordshire	...	6	3	1	20	17		
Oxfordshire	...	1	0	0	0	4		

Quelle: Alcock's Football Annual, 1888, S. 101–102.

Kommen wir nun zur Einführung neuer statistischer Kategorien und ihrer Bedeutung für die Entwicklung neuer Vergleichs- und Bewertungspraktiken. Ein interessantes Experiment findet sich ebenfalls 1888 in einer Tabelle der *Hampshire Association* (Abbildung 11). In dieser werden die Spielergebnisse der dieser *Association* zugehörigen Clubs gesammelt und in tabellarischer Form dargestellt, wobei die Clubs entsprechend ihrer Leistung »gerankt« werden,

und zwar nach einer neu eingeführten statistischen Kategorie: dem »Average«. Dieser Durchschnitt wird berechnet anhand der erzielten Tore, dividiert durch die erhaltenen Tore. Eine denkbare Inspiration hierfür ist die Bewertungspraxis im Baseball oder auch im Cricket, wo *Averages* bereits zu dieser Zeit übliche Leistungskategorien waren. Interessant ist zudem die Orientierung an den Toren – nicht Punkten – als Leistungsmerkmal. Mit dem Durchschnitt aus den geschossenen und erhaltenen Toren soll offenbar einem Vergleichsproblem begegnet werden, nämlich dass die Clubs eine unterschiedliche Anzahl an Spielen absolviert haben, das bloße Zusammenrechnen der Punkte daher keinen leistungsgerechten Vergleich erlauben würde. Auch hier bestätigt sich also der oben an der Ligatabelle erläuterte Trend zum expliziten Vergleichen, wenn auch in einer Form, die sich langfristig nicht durchgesetzt hat.

Abbildung 11: Tabelle der Hampshire Association mit Club Averages

		Matches played.	Won.	Lost.	Drawn.	Goals		
						For	Against	Average
Hilsea Ramblers	..	18	14	1	3	51	7	.28
Winchester	..	22	17	3	2	92	19	.484
Christchurch II.	..	4	5	1	0	17	4	.425
Portsmouth Flower Jun.	..	2	2	1	0	7	2	.35
Winchester II.	..	4	2	1	1	10	3	.33
St. Mary's, Southampton	..	21	15	4	2	46	10	.306
Harriers, Southampton	..	20	9	4	7	39	14	.25
Pokesdown	..	4	4	0	0	11	4	.275
Ringwood Hornets	..	15	10	3	2	61	23	.265
Royal Engineers, Aldershot	..	82	21	5	6	91	96	.252
Woolston Works	..	7	5	1	1	23	10	.23
Fordingbridge Turks	..	11	7	0	0	29	10	.23
Andover	..	18	11	6	1	59	22	.223
Bournemouth Arabs II.	..	17	12	4	1	43	22	.213
Winchester Rovers	..	20	12	4	4	46	23	.2
Havant	..	10	6	1	3	17	9	.188
Boscombe	..	12	8	3	1	33	19	.173
Fordingbridge Pirates	..	9	4	2	3	15	11	.172
Queen's School, Basingstoke	..	12	10	6	1	48	29	.165
Bournemouth Dean Park	..	12	8	4	2	37	28	.160
Portsmouth Sunflowers	..	9	5	2	2	21	16	.131
Portsmouth	..	22	9	8	1	59	27	.196
Freemantle	..	13	8	4	1	24	19	.126
Whitchurch	..	13	8	3	2	26	21	.123
Banister Court, South'ton	..	15	7	6	2	28	24	.116
Bournemouth Arabs	..	21	6	7	8	41	36	.113
Cowes	..	17	9	4	4	40	37	.103
Bournemouth Albion	..	18	5	7	6	38	36	.105
Christchurch	..	19	7	9	3	80	29	.103
Bournemouth East	..	17	6	7	4	36	40	.9
Bournemouth Wanderers	..	23	9	11	3	45	52	.86
Bournemouth Dean Park II.	..	10	3	3	4	24	28	.85
Bournemouth East II.	..	14	5	8	1	27	38	.81
Hayling Island	..	20	5	8	7	19	24	.79
Christchurch School	..	9	4	5	0	25	40	.62
Lymington	..	15	5	6	4	16	25	.60
Totton	..	15	3	11	1	14	51	.27

Quelle: Alcock's Football Annual, 1888, S. 168.

Ein weiteres Beispiel für die Verwendung von *Averages* ist dieser Fall einer Historisierung der Ergebnisse des in den ersten Jahren der Liga besonders erfolgreichen Clubs Preston North End (Abbildung 12). Auch hier wird der *Average* durch die Formel »erzielte Tore« dividiert durch »erhaltene Tore« berechnet. Zusätzlich wird hier – wie oben ausgeführt – eine Historisierung vorgenommen. Die Ergebnisse verschiedener Jahre werden dabei zueinander in Beziehung gesetzt und erhalten dadurch einen relativen Wert. So ist Preston North End etwa entsprechend des *Averages* in der Saison 1884–5 mit einem Wert von »4,19 zu 1« besser als in der Saison 1883–4 mit einem Wert von »2,34 zu 1«.

Abbildung 12: Historientabelle von Preston North End mit saisonalen Ergebnissen

Year.	Played	Won.	Drawn.	Lost.	For.	Against.	Goals	Average,
1883-4	..	44	30	10	4	117	50	2.34 to 1
1884-5	..	52	42	7	6	210	51	4.19 to 1
1885-6	..	63	59	3	2	318	60	4.30 to 1
1886-7	..	63	51	5	7	286	69	4.14 to 1
1887-8	..	69	58	4	7	312	79	3.94 to 1
1888-9	..	60	46	6	8	174	61	2.85 to 1

Quelle: Alcock's Football Annual, 1889, S. 125.

1892 findet sich dann eine weitere statistische Kategorie, die einer ähnlichen Logik folgt: die »Proportion«. Im Gegensatz zum *Average*, der sich auf Tore als Leistungsindikator stützt, berechnet sich die *Proportion* aus den erzielten Punkten im Verhältnis zu den möglichen Punkten (»Pts.«/»Out of«) und der Anzahl gewonnener, verlorener und unentschiedener Spiele. Diese Vergleichsmethode entspricht im Prinzip besser unserem heutigen – vermutlich am Ligasystem geschulten – Verständnis sportlicher Leistungsfähigkeit in Mannschaftssportarten, das Siege höher bewertet als die Tordifferenz (wobei letztere als sekundärer Leistungsindikator Verwendung finden kann). Auch die *Proportion* soll offenbar die den Vergleich erschwerende Tatsache kompensieren, dass unterschiedliche Teams unterschiedlich viele Spiele absolviert haben. Bei genauerem Hinsehen zeigen sich jedoch auch Nachteile dieser Leistungskategorie. Denn da sie von der Zahl der ausgetragenen Spiele abstrahiert, kann es zu merkwürdigen Ergebnissen wie in Abbildung 13 kommen: Cheshire erreicht den optimalen Wert 1,00, war aber nur in einem einzigen – gewonnenen – Spiel angetreten.

Die *Proportion* wird in den Jahrbüchern zur Bewertung und für den Vergleich der Leistungen von *Associations* verwendet. Im Unterschied zur Liga-

tabelle werden die *Associations* aber weiterhin alphabetisch angeordnet, nicht nach Leistung hierarchisiert. Dies scheint bezeichnend zu sein für eine Differenzierung innerhalb des britischen Fußballs, die sich nach Einführung der Liga einspielt: Während im Amateursport noch lange an der inklusiv-multipolaren Bewertungsordnung festgehalten wird, mit lediglich geringen Anteilen expliziter Vergleichselemente (wie den *Averages* und *Proportions*), lässt sich im professionellen Ligabetrieb ein deutlicher Trend hin zum hierarchischen Vergleichen beobachten, der im Prinzip bis heute angehalten hat.

Abbildung 13: Tabelle von Associations und County-Clubs mit statistischer Kategorie Proportion

The following table taken from the <i>Sportsman</i> will, however, show at a glance the doings of the various Associations and County Clubs in 1891–92:—							
	P.	W.	D.	L.	Goals		Proportion.
					For.	Agt.	
Berks and Bucks	4	2	1	1	11	12	.62
*Cambridgeshire	1	0	0	1	1	2	—
Cheshire	1	1	0	0	2	0	1.00
Cleveland	1	0	0	1	3	6	—
Cornwall	2	1	0	1	3	7	.50
Cumberland	2	0	0	2	0	5	.33
Devonshire	3	1	0	2	7	6	.50
Dorsetshire	2	1	0	1	3	2	.40
Essex	4	2	2	0	12	5	.66

23							
	P.	W.	D.	L.	Goals		Proportion.
					For.	Agt.	
Gloucestershire	2	0	0	2	3	10	—
Huntingdonshire	1	1	0	0	2	1	.2
Hampshire	4	3	1	0	19	9	.87
Hertfordshire	3	1	0	2	9	7	.33
§Kent	4	0	0	4	11	17	—
Leicestershire	1	0	0	1	0	7	—
London	3	0	1	2	6	12	.16
Middlesex	5	4	0	1	17	7	.80
Northumberland	2	2	0	0	9	3	1.00
Norfolk	2	0	1	1	2	3	.25
¶Northamptonshire	5	3	2	0	18	8	.80
Oxfordshire	2	0	0	2	6	10	—
Sheffield	2	1	0	1	6	5	.50
Somersetshire	2	0	1	1	4	5	.25
**Suffolk	1	0	1	0	0	0	—
††Surrey	6	3	2	2	22	25	.58
Sussex	5	1	2	2	16	19	.40
Wiltshire	6	4	1	1	20	13	.75

A win equals 2, and a draw 1 point.

* Exclusive of match v. Cambridge University, drawn (2—2).
 † Exclusive of game with Canadians, lost (1—7).
 ‡ Exclusive of match v. Corinthians, lost (0—1).
 § Exclusive of match v. United Hospitals, won (6—3).
 || Exclusive of fixtures with Canadians v. Cambridge University (2), &c.
 ¶ Exclusive of contest (drawn 1—1) with Canadians.
 ** Exclusive of fixtures with Preston North End, lost (0—3), and Cambridge University, drawn (1—1).
 †† Exclusive of match v. United Hospitals, won (4—1).

Quelle: Alcock's Football Annual, 1992, S. 22–23.

Eine weitere statistische Kategorie entsteht im Kontext von Länderspielen. Ab 1894 werden auch die *Internationals* nicht mehr nur als einzelne Spiele oder Spiele einer Serie behandelt, sondern die Nationen werden in eine Rangliste

anhand von »Relative Positions« überführt (Abbildung 14). Auch dadurch wird ein gemeinsamer Horizont für zuvor voneinander getrennt wahrgenommene Spiele aufgespannt.

Abbildung 14: Länderspiele – Übersicht Spiele und Relative Positions

SUMMARY OF INTERNATIONAL MATCHES.						
45						
RELATIVE POSITIONS.						
					GOALS.	
Feb. 24	Swansea ..	Wales	4	beat Ireland 1
Mar. 3	Belfast ..	Ireland	2	drawn England 2
.. 12	Wrexham ..	England	5	beat Wales 1
.. 24	Kilmarnock ..	Scotland	5	beat Wales 2
.. 31	Belfast ..	Scotland	2	beat Ireland 1
April 7	Glasgow ..	Scotland	2	drawn England 2
England's points were credited to J. G. Veitch (3), R. C. Gosling, Devey, Spikesley, Goodall, Reynolds, and one from a rebound off a Welsh back (Parry).						

Quelle: Alcock's Football Annual, 1894, S. 45.

Zusammenfassend kann man seit der Einführung der Liga in den späten 1880er-Jahren einen sich ausweitenden Trend hin zur Einführung statistisch-tabellarischer Vergleichspraktiken festhalten. Tabellen scheinen durchgehend für hervorzuhebende Inhalte genutzt worden zu sein, dienten also der Zuschreibung von Relevanz. Darauf hinaus waren sie ein Instrument für die Institutionalisierung expliziter Leistungsvergleiche, die mit der Einführung der Liga üblich(er) wurden. Der aus den USA importierte Ligawettbewerb verlangte eine hierarchische Anordnung anhand eines Punktesystems und machte sich umgekehrt von tabellarischen Darstellungen abhängig. Doch die Wahrnehmung anderer Wettbewerbe und Vergleichskontexte, seien es Clubs innerhalb einer Association, Inter-Association-Matches oder Internationals, scheinen von der neuen Leistungsorientierung ebenfalls angesteckt worden zu sein – auch auf sie beziehen sich nun explizite Vergleichspraktiken, die statistische Kategorien wie Averages und Proportions in Anschlag bringen. Dieser Wechselbeziehung zwischen Wettbewerbsformaten sowie Bewertungs- und Vergleichspraktiken möchten wir im folgenden Abschnitt noch etwas genauer nachgehen.

5. Wettbewerbsformate und ihre Bewertungs- und Vergleichspraktiken

In der frühen Phase des sich professionalisierenden Fußballs – in den 1850er- und 1860er-Jahren⁶ – waren *Friendlies* das übliche Wettbewerbsformat, das heißt 1-gegen-1-Spiele, die manchmal auch im Rahmen einer Serie gespielt wurden. Diese Spiele arrangierten die Clubs selbstständig untereinander, ohne übergreifenden Wettbewerbsmodus, der viele oder gar alle Spiele zueinander ins Verhältnis setzen würde. Manchmal wurde auch von *Challenges* gesprochen, wenn ein Team ein anderes herausforderte. Beispielsweise schrieben ad hoc gegründete Teams häufig Anzeigen für *Challenges* in der Presse aus, meist ging es dabei (auch) um ein Preisgeld. Die Aufmerksamkeitslenkung bzw. die Hierarchisierung von *Friendlies* erfolgte meist über die Anknüpfung an traditionelle Rivalitäten, zum Beispiel Oxford gegen Cambridge oder Old Etonians gegen die Old Harrovians, oder an lokale *Rivalries*, zum Beispiel bei lokalen Derbys oder zwischen den beiden dominanten Fußballkulturen London gegen Sheffield. Zur Orientierung dienten dabei offenbar primär lokale Identitäten und historisch begründete Rivalitäten.

Im Rahmen des allgemeinen Wachstums der Fußballkultur und im Zuge der zunehmenden Vereinheitlichung der Regeln – nach Gründung der FA 1863 sowie der *Sheffield Football Association* 1867 – wurde 1867 mit dem *Youdan Cup* in Sheffield der erste Cup-Wettbewerb durchgeführt. Cup-Wettbewerbe zeichnen sich durch einen übergreifenden Wettbewerbsmodus aus, in dem die Teilnehmenden im Knock-out-Format gegeneinander antreten, bis im Finalspiel ein Sieger gekürt wird. Das Cup-Format unterscheidet sich also von den *Friendlies* wie auch den *Challenges* dadurch, dass alle Cup-Spiele in einem gemeinsamen Bezugssystem stattfinden und trotz der Vielzahl der Spiele letztlich ein eindeutiger Sieger festgestellt werden kann. Der erste Cup wird von Music-Hall-Besitzer Youdan eingerichtet, der die florierende Sheffield-Fußballkultur⁶ als ökonomische Chance erkannte und Eintrittsgelder zu den Spielen verlangte (Harvey 2001, S. 62–63). Doch dieser in seinen Anfängen primär

6 Der Sheffield Football gilt als die erste größere regionale Fußballkultur in den 1850er- und 1860er-Jahren; bereits vor Gründung der FA. Der Sheffield Football Club entwickelte als hegemonialer Club ein Set einheitlicher Regeln, die von allen Clubs der Region genutzt wurden (Harvey 2001). An diese Fußballcommunity richtete sich der Youdan Cup.

ökonomisch begründete Wettbewerb wurde auch sportlich und medial als großer Erfolg wahrgenommen. An diese Erfahrung knüpfte der 1871–72 erstmals ausgetragene FA Cup an. Er sollte als Anreiz dienen, die Association-Regeln zu verbreiten, zur Regelvereinheitlichung beitragen und die Attraktivität des *Association Footballs* erhöhen.⁷ Cup-Wettbewerbe folgten zumeist einer inklusiven Logik; auch der FA Cup war von Beginn an für alle Vereine offen, die nach den Association-Regeln spielten.

In der medialen Sportberichterstattung führt dieses Wettbewerbsformat zu einer deutlichen Aufmerksamkeitssteigerung. Wie oben bereits angedeutet, lässt sich dies auch an den Jahrbüchern nachvollziehen: Cup-Spiele werden hervorgehoben, erhalten eigene narrative Beschreibungen im Kapitel der »important matches of the season« oder ein eigenes Unterkapitel für die Ziehungen und Ergebnisse des FA Cups. Cup-Spiele werden auch sonst bald medial prominenter herausgestellt als *Friendlies*, erhalten quantitativ mehr Platz und werden intensiver besprochen. Am ersten FA Cup nehmen zwar nur 15 Clubs teil, doch schon schnell gewinnt er an Popularität. Mit dem Erfolg des Formats kommt es auch zu einer Ausweitung. Die sich neu gründenden regionalen *Associations*, die Ende der 1870er- und Anfang der 1880er-Jahren überall in England aus dem Boden schießen, führen schon bald eigene Cup-Wettbewerbe ein (Taylor, 2008, S. 39). Dies führte zu einer Verschiebung der Berichterstattung von *Friendlies* zu Cup-Spielen, aber auch zu einer neuen Unübersichtlichkeit, die mit der schieren Zahl von Cup-Spielen zu tun hatte. Abbildungen 15a-c geben einen Einblick in die Vielzahl der Cup-Wettbewerbe, die in diesen Jahren neu eingeführt wurden.

Der FA Cup wird besonders herausgestellt, ansonsten findet keine klare Hierarchisierung zwischen den Cup-Wettbewerben statt. Dies wurde insofern zum Problem, als viele der Clubs an einer Vielzahl von Cups teilnahmen und sich bei Terminkollisionen gezwungen sahen, Prioritäten zu setzen oder sich gar aus einzelnen Cupwettbewerben zurückzuziehen. Dies führte nicht nur zu terminlicher Unsicherheit, sondern gefährdete auch die Einnahmen durch Eintrittsgelder, die vor allem für die professionellen Clubs immer wichtiger wurden (Taylor 2005, S. 3). In diesem Zusammenhang kam in den 1880er-

⁷ Der Initiator Charles W. Alcock verwies bei seinem Vorschlag eines Cups allerdings auf seine Erfahrung mit einem Knock-out-Wettbewerb von der Public School Harrow. Es ist aber davon auszugehen, dass er den Youdan Cup und dessen Nachfolger – den Cromwell Cup – kannte (Taylor 2008: 40–41).

Jahren der Vorschlag zur Gründung einer professionellen Liga auf – das Format war schon länger durch Kontakte mit dem US-amerikanischen Baseball bekannt (Taylor 2005, S. 4) –, die den Terminkalender für die teilnehmenden Clubs verlässlicher machen und den »leading clubs« Zuschauereinnahmen garantieren sollte.

Abbildung 15a: Finalergebnisse etlicher Cup-Wettbewerbe

The results of the different Cup competitions which follow are reproduced from the <i>Sportsman</i> :				
Football Assoc. Cup..	West B. Albion	Aston Villa	3 0	
London Senior	Old Westminsters	Holland	2 1	
Scottish Cup	Celtic	Queen's Park	5 1	
Welsh	Clark 2nd	Westminster Rovers	2 1	
Irish	Linfield	Black Watch	2 0	
Eton (Lower Boys) ..	Rev. E. Hale's	Mr. R. A. H. Mitchell's	1 1	
Glasgow	Celtic	Clyde	7 0	
Bristol	Mr. R. A. Mitchell's	Mr. A. C. James's	2 0	
Weymouth	Maidstone Norfolkians	Marlow Ramblers	3 1	
Cambridge Inter-Col.	Jesus College	Trinity Harrovians	4 0	
Renfrewshire	Abercorn	St. Mirren	2 0	
Stirlingshire	Campsie	Falkirk	4 1	
Ayrshire	Anbank	Harford	3 1	
Lanarkshire	Airdrieonians	Wishaw Thistle	4 1	

r = rouge.

Quelle: Alcock's Football Annual 1892: 27–29.

Abbildung 15b: Finalergebnisse etlicher Cup-Wettbewerbe
(Fortsetzung)

29				
Middlesex Junior	Millwall 2nd	London Caledonians 2nd ..	1 0	
West Ham	Iford	Upton Park	4 0	
Oxfordshire	Oxford College Servants	Banbury H.	6 0	
Leicestershire	Hugglescote 'Robin Hood's	Leicester Fosse	3 1	
Swindon	Stowes 2nd	Stratford Rovers	6 1	
Norfolk	C. E. M. M. S.	Worthing	6 0	
Shropshire	Ironbridge	Whitchurch	6 1	
London Charity	Crusaders	Millwall Athletic	1 0	
Cleveland Charity	M. Ironopolis	Middlesbrough'	2 1	
Kettering Charity	Gresham Rovers	Kettering Town	1 0	
Staffordshire	Burslem Southgate	Aterton Villa	1 1	
West London	Stanley	Harrow	6 1	
Hampshire Charity	Royal Engineers	St. Mary's, Southampton	3 1	
Manchester	Ardwick	Bolton Wanderers	4 1	
Lancashire	Bury	Blackburn Rovers	2 0	
Birmingham City Charity	Emm Lane Heath	Aston Villa	2 0	
Liverpool	Everton	Southport Central	2 1	
Lincolnshire	Lincoln City	Grimsby Trinity	3 2	
Middlesex	2nd Scots Guards	1st Scots Guards	3 2	
Bath Charity Vase	Derby County	Notts	5 0	
Derbyshire Charity	Derbyshire Works	Claycross	4 0	
Birmingham	Wolverhampton W.	West Bromwich Albion	5 2	
Shropshire Charity	St. George's	Ironbridge	1 0	
Worcestershire Char.	Worcester Rovers	Berwick Rangers	2 0	
Chatham Charity	Chatham	West Kent Regiment	2 2	
Dudley	West Bromwich Alb.	Stourbridge	3 1	

Quelle: Alcock's Football Annual 1892: 27–29.

*Abbildung 15c: Finalergebnisse etlicher Cup-Wettbewerbe
(Fortsetzung)*

		28	
East of Scotland	Heart of Midlothian		
Kent Junior	Highland L. I. 2nd	Cray Wanderers	2 0
Sussex Junior	Surrey	Harrow	3 2
London Junior	Olympic	London Canarians (2nd)	5 3
Brighton	Brighouse Hornets	Brighton Athletics	4 0
Derbyshire	Derby Co. (2nd)	Long Eaton Ramblers	0 0
Hants Junior	Dumbarton	Fremantle	3 2
Peterborough	Weymouth	Ward Athole	6 1
Dorset Junior	Brighton Hornets.	Wardour	4 1
Sussex	1st Div. (res.)	Worthing	5 3
Northumbrian	Sheffield United	Rendel	2 1
Sheffield	71st Highlanders	W. London Wanderers	2 1
Kent	Warmley	Gravesend	1 0
Gloucestershire	Leamington Head Boys	Bedminster	2 1
Herald	St. Mary's, Southampton	Battersea Albion	3 1
Hampshire	Durham	Medical Staff	5 0
	Wiltshire	Bishop Auckland	2 1
Middleton Norfolkian	Swindon Town	Darlington	13 0
Oxford University Inter-Collegiate	Maidenhead 2nd	Swindon 2nd	3 0
Westminster School Shield	Magdalene	Slough	
Suffolk	Rigall's	Trinity	4 0
Lancashire Juniors	Bogart's College	Home Brothers	2 1
Cumberland	Leasby	Leeds	5 1
Portsmouth	Moss Bay Exchange	Southport Central	1 0
Birmingham Juniors	Freemantle	Workington	4 3
Inter-Hospital	Singers, Coventry	Portsmouth Grammar Sch.	2 0
Hants Minor	St. Bart.'s Hospital	Willenhall Pickwick	2 0
	Bournemouth Wanderers	St. Mary's Hospital	2 1
Middleton Charity	Weymouth	Cowes White Star	4 1
West End	Kildare	N. Stafford Regt. Brigade	1 0
Cleveland Priory	M. Ironopolis	West End	1 0
Berks and Bucks	Reading	Middlesbrough	4 0
Cheshire	Crewe Alexandra	Wolverton L. and N. W. R.	2 0
Manchester Junior	Middleton	Northwich Victoria	3 1
Irish League	Linfield	Harrow Ramblers	4 3
Northants	Finedon	Harrow	5 1
Surrey	Reigate Priory	Marlow	3 1
West Middlesex	Uxbridge	Wellington	3 1
Henessey	2nd Scots Guards	Guy's Hospital	3 1
Essex	1st Div.	Harrow Athletic	3 2
Cleveland Amateur	Middlebrough Ironopolis	Marlow	1 0
City of London	Mineavia	Colchester	3 1
Berks and Bucks Juniors	Slough	Port Clarence	3 1
Hertfordshire	West Herts	Olympic	5 2
West Surrey League	Godalming	Windsor and Eton Victoria	3 1
W. London Observer	Stanley	Hoddesdon	5 2
Grantham Charity	Derby Co. Wanderers	Dorking	2 1
Norfolk Juniors	Diss	Fulham	1 0
Working Men's Federation	Eton Mission	Kettering	3 1
Army Cup	2nd Scots Guards	Yarmouth 2nd	7 1
		Oxford House	2 1
		1st Scots Guards	

Quelle: Alcock's Football Annual 1892: 27–29.

Der Kontext der Einführung der Liga im britischen Fußball unterscheidet sich deutlich von jenem des US-amerikanischen Baseballs. Im Baseball war die Liga eine Lösung für ein *diskursiv erzeugtes* Problem: Schon in den 1860er-Jahren kreiste der Diskurs in der Sportpresse zunehmend um die Frage »who is the champion?« und artikulierte das Bedürfnis, einen klaren Sieger – den *Champion* – zu bestimmen. Das im Baseball in der New Yorker Region schon länger praktizierte Challenge-Modell – Cup-Wettbewerbe gab es im Baseball nicht – wurde zunehmend als unzureichend wahrgenommen. Diese veränderte Wahrnehmung reflektierte die Einübung in statistische Vergleichspraktiken, die im Baseballdiskurs seit den 1860er-Jahren üblich waren und es nahelegten, *überlegene Leistungsfähigkeit als Ausdruck konstanter Leistungen über ei-*

ne ganze Saison hinweg zu definieren. Einmalige Siege, und sei es in einer Serie von *Challenges* gegen den bisherigen *Champion*, wurden daher nicht länger als ausreichend angesehen (Minnetian & Werron, 2021). Der Ligawettbewerb löste dieses Problem, indem es denjenigen Club als *Champion* anerkannte, der im Laufe einer Saison die meisten Spiele gewonnen hatte – was im Umkehrschluss bedeutete, dass man als *Champion* auch einzelne Spiele verlieren konnte (Leifer, 1995).

Die Gründung der *Football League* im britischen Fußball – im Kontext der Unübersichtlichkeit der Cupwettbewerbe – war dagegen (primär) ökonomisch motiviert. Den Quellen ist nicht zu entnehmen, dass ihre Einführung auf eine diskursiv erzeugte Nachfrage nach neuen Kriterien zur Bestimmung eines Champions reagierte. Jedoch lassen sich in den Jahren *nach* Einführung der Liga bald zahlreiche Hinweise auf einen Wandel des Leistungsdiskurses beobachten. In den Jahrbüchern wurde nun zum Saisonende rückblickend vielfach auf Konsistenz der Ergebnisse und konstant gute Leistungen verwiesen; so etwa 1892 hinsichtlich der in den Anfangsjahren der Liga besonders erfolgreichen Preston North End (wobei über die saisonale Leistung hinaus auch historisch ausgegriffen und auf vorherige Spielzeiten verwiesen wird):

One of the most noteworthy features of the season was the *consistently brilliant form* shown by the Preston North End team. For the last four or five seasons the North Enders have been in quite the fore-front of Association Football. Since professionalism was legalized they have proved themselves to be an exceptionally powerful combination at all points. [...] Since 1883 the Club has had a most brilliant record, and a summary of its doings during the last six years will be found very interesting. (Alcock's Football Annual, 1889, S. 124, Hervorhebung durch Autor*innen)⁸

Ebenfalls 1892 wurde anlässlich der FA-Cup-Finalteilnahme der Wolverhampton Wanderers über deren »excellent performance throughout the season« berichtet, wobei zugleich auch – für die Wanderers nicht vorteilhafte – Leistungsvergleiche mit anderen Clubs angestellt wurden.

The Wolverhampton Wanderers showed themselves to be a thoroughly useful hard-working lot, and their *excellent performances throughout the season* fully justified them reaching the final stage of the Cup competition. At the

8 Darstellung der dazugehörigen saisonübergreifenden Statistik siehe Abbildung 12.

same time they lacked the science which was so characteristic of the Preston team, and public opinion would hardly adjudge them perhaps to be as dangerous an opposition as the Blackburn Rovers. (Alcock's Football Annual, 1889, S. 125, Hervorhebung durch Autor*innen)

Der positiven Hervorhebung konstanter Leistungen über die ganze Saison oder mehrere Spielzeiten hinweg stand die Kritik am unsteten Spiel gegenüber. So sei der Cup-Erfolg der Wolverhampton Wanderers ein Jahr später (1893) überraschend angesichts ihrer sonst wenig konstanten Leistungen (»uneven play«):

Certainly no one would have ventured to predict, from their *uneven play* in the early rounds of the latter competition [FA-Cup], that the Wolverhampton Wanderers would prove to be the fortunate eleven to secure the highest honours of the season. (Alcock's Football Annual, 1893, S. 47–48, Hervorhebung durch Autor*innen)

Allerdings konnte offenbar – dem hier vorherrschenden Leistungsverständnis zufolge – mangelnde Konstanz durch Erfolge in verschiedenen Wettbewerben kompensiert und damit relativiert werden. Um den Vergleichshorizont entsprechend zu erweitern, wurde narrativ Leistung aus verschiedenen Wettbewerben zueinander in Beziehung gesetzt:

Although there was perhaps some reason for discounting the victory of the Wolverhampton Wanderers in the National Cup Competition they gave an unmistakeable proof of their all-round capacity by their success in their final for the Birmingham Association Cup. (Alcock's Football Annual, 1893, S. 50)

Der neuen Präferenz für konstante Leistung entspricht ein neuer Sinn für seinen Gegensatz: den Glücksfaktor. Er wird explizit bemerkt und betont, wenn die Erfolge eines Teams in der Liga mit ausbleibenden Erfolgen im Cup-Wettbewerben einhergehen:

Luck, though, had not in previous years befriended them in their attempts to carry off the premiership of Association football—The National Cup—and in more than one instance, they had decidedly hard lines in the later stages of the competition. (Alcock's Football Annual, 1889, S. 124)

Oder:

»Down South, the Old Carthusians still maintain the high reputation they have held for several years, and when represented by their best team, are capable of taking their own part with any team in the country, professional or amateur. In the Cup competition they were *singularly unlucky* [...] They are certainly the best of the Southern Clubs at the present time.« (Alcock's Football Annual 1889, S. 125–126, Hervorhebung durch Autor*innen)

Und bereits ab dem ersten Jahr der Koexistenz von FA Cup und Liga gab es bereits die Sonderleistung, beide Wettbewerbe zu gewinnen, für deren Bezeichnung eigene Ausdrücke entwickelt wurden: »double event« (Alcock's Football Annual, 1892, S. 25), »double first« (Alcock's Football Annual, 1892, S. 25) oder »dual feat«:

That they [Preston North End] were the *best team of 1888–89 was proved beyond a doubt*. Their extraordinary record in the League Competition, in which they played twenty-two matches without a reverse, would have been of itself conclusive evidence, had there not been additional confirmation in their subsequent successes for the Association Cup. Their defeat of the Wolverhampton Wanderers in the final tie, at the Oval, was the result of really good all-round play, and their dual feat in carrying off at once the championship of the League and the Association Challenge Cup *placed them unmistakably at the head of Association Clubs*. (Alcock's Football Annual 1889, S. 124–125, Hervorhebung durch Autor*innen)

Diese Zitate zeigen, dass sich nach der Einführung der *Football League* 1888 die Grundorientierung des Fußballdiskurses veränderte: Explizite Leistungsvergleiche wurden üblich, und die Kriterien für überlegene Leistungsfähigkeit wurden auf Konstanz über die gesamte Saison hinweg ausgerichtet. Zugleich zeigte sich ein Bemühen um faire Bewertungen, das sich auf explizite Vergleiche von Leistungen in unterschiedlichen Wettbewerben stützte. In dieser Neuausrichtung des Diskurses spielte eine wichtige Rolle, dass nun auch die Erfolge in Cup- und Ligawettbewerben miteinander verglichen werden konnten, mit der Folge, dass das Zufalls- oder Glücksmoment im Cupwettbewerb – wo jede Niederlage zum sofortigen Ausscheiden führt – besonders auffiel und hervorgehoben wurde. Dieses institutionelle Arrangement aus verschiedenen Wettbewerbsformaten und den zu ihnen passenden Vergleichs- und Bewertungspraktiken hat sich bis heute gehalten und ist zu einem hervorstechenden Merkmal der Vergleichs- und Bewertungskultur des Fußballs geworden.

Schlussbemerkungen: Für eine historische Perspektive auf Vergleichs- und Bewertungspraktiken

Die Sammlungen von Informationen, die in den Jahrbüchern des britischen Fußballs in der zweiten Hälfte des 19. Jahrhunderts präsentiert wurden, artikulieren in den Anfangsjahren eine inklusiv-multipolare Leistungsordnung, die nicht auf Erzeugung klarer Hierarchien ausgerichtet war – wie dies im Baseball bereits früh der Fall war –, sondern auf die gleichberechtigte Ansprache aller Community-Mitglieder. Bis in die 1880er-Jahre überwiegt der Community-Charakter, der Leistungsvergleich bleibt implizit.

Zugleich gibt es Hinweise, dass sich im Fußballdiskurs in den 1870er-Jahren zumindest *Angebote* für Leistungsvergleiche entwickeln, insbesondere in Form von tabellarischen Übersichten von »chief schools« bzw. »leading clubs«, »chief players« und »important matches«. Jedoch werden die Kriterien für die Hervorhebung bestimmter Clubs, Spieler und Spiele nirgends explizit diskutiert. Daneben lassen sich früh bereits historisierende Rück- bzw. Vorausblicke (Fixtures bzw. Saisonkalender) sowie historisierende Bewertungspraktiken beobachten (Saisonstatistiken, Liste von FA-Cup-Gewinnern der letzten Jahre), die »die Saison« als Bewertungseinheit konstituierten und besonders erfolgreiche Clubs hervorhoben. Historische Bezugnahmen und kategoriale Bewertungen erfolgten explizit, die zugrunde liegenden Vergleiche dagegen blieben implizit und der Vorstellungskraft der Leser*innen überlassen. Dies ändert sich in Teilen mit der Einführung der *Football League*. Mit der aus den USA importierten Liga gab es ein Wettbewerbsformat, dem die Idee einer Hierarchisierung der Clubs nach Leistungsfähigkeit eingeschrieben war und das aufgrund seines seriellen Charakters Anlass gab, die Entwicklung dieser Leistungsfähigkeiten über die gesamte Saison hinweg tabellarisch zu beobachten – und also die Leistungsfähigkeit der Clubs und Spieler nicht nur zu bewerten, sondern auch fortlaufend und explizit miteinander zu vergleichen.

Anders als im amerikanischen Baseball, wo der hierarchisierende Leistungsdiskurs das Interesse am Ligaformat weckte und die Erfindung der Liga mit bewirkte, scheint hier also gerade umgekehrt erst die Einführung des Ligaformats dazu geführt zu haben, dass der bis dahin weitgehend implizite Leistungsdiskurs explizit(er) wurde und sich in hierarchisierenden Leistungsvergleichen artikulierte. Die Einführung der Liga war also offenbar mit einer *unbeabsichtigten diskursiven Nebenfolge* verbunden: der Transformation des inklusiv-multipolaren und weitgehend impliziten in einen stärker hierarchisch geprägten und explizit vergleichenden Leistungsdiskurs. Zuge-

spitzt formuliert: Nicht der Leistungsdiskurs motivierte die Einführung des neuen Wettbewerbsformats, das neue Wettbewerbsformat transformierte den Leistungsdiskurs.

Wir verstehen diese neue und, wie wir hoffen, präzisierte Beschreibung der Rolle von Vergleichs- und Bewertungspraktiken in der Entstehungsgeschichte des modernen Fußballs auch als Hinweis darauf, dass die Vergleichs- und Bewertungssoziologie davon profitieren könnte, sich mit der Entstehungsgeschichte verschiedener Sportarten intensiver auseinanderzusetzen und sie vergleichend zu untersuchen. Alle Sportarten verfügen über eigene institutionelle Arrangements, innerhalb derer sich bestimmte Wettbewerbsformate und spezifische Vergleichs- und Bewertungspraktiken gegenseitig abstützen, wobei die Wettbewerbsformate von Vergleichs- und Bewertungspraktiken ermöglicht werden wie auch umgekehrt Vergleichs- und Bewertungspraktiken erst in Reaktion auf die Einführung bestimmter Wettbewerbsformate entstehen können. Diese unterschiedlichen Entstehungslogiken sind nicht nur historische Tatsachen, sie können sich langfristig fortsetzen und der Kultur einer Sportart einprägen. Exemplarisch gilt dies für den Fußball mit der ihn auch heute noch auszeichnenden Koexistenz von Cup- und Ligawettbewerben und der Verschränkung von Amateur- und Profisport unter der Führung von Verbänden. Wer die institutionellen Arrangements einzelner Sportarten verstehen will, ist folglich auf die Untersuchung ihrer historischen Entstehung und der Rolle, die Vergleichs- und Bewertungspraktiken darin spielen, angewiesen.

Literatur

- Harvey, A. (2001). An Epoch in the Annals of National Sport. Football in Sheffield and the Creation of Modern Soccer and Rugby. *The International Journal of the History of Sport*, 18(4), 53–87.
- Harvey, A. (2005). Football: The First Hundred Years. The Untold Story. Routledge, 2005.
- Leifer, E. M. (1995). Making the Majors. The Transformation of Team Sports in America. Harvard University Press.
- Mason, T. (1986). Sporting News, 1860–1914. In M. Harris & A. J. Lee (Hg.), The Press in English society from the seventeenth to nineteenth centuries (S. 168–186). Fairleigh Dickinson University Press & Associated University Presses.

- Mason, T. (1993). All The Winners and the Half Times. *The Sport Historian*, 13(1), 3–13.
- Minnetian, C., & Werron, T. (2021). Redefining Achievement: The Emergence of Rankings in American Baseball. *Research in the Sociology of Organizations*, 74, 127–151.
- Ringel, L., Espeland, W., Sauder, M., & Werron, T. (Hg.) (2021). *Worlds of Rankings*. Emerald.
- Ringel, L., & Werron, T. (2020). Where Do Rankings Come From? A Historical-Sociological Perspective on the History of Modern Rankings. In A. Epple, W. Erhart, & J. Grave (Hg.), *Practices of Comparing. Towards a New Understanding of a Fundamental Human Practice* (S. 137–170). Bielefeld University Press.
- Stauff, M. (2019). „History in the Making“: Sports and the Serialized Production of Collective Memory. *Erdélyi Társadalom*, 17(1), 9–20.
- Tate, S. (2007). *The Professionalisation of Sports Journalism, 1850 to 1939, with Particular Reference to the Career of James Catton*. University of Central Lancashire.
- Taylor, M. (2005). *The Leaguers. The Making of Professional Football in England 1900–1939*. Liverpool University Press.
- Taylor, M. (2008). *The Association Game. A History of British Football*. Pearson Longman.
- Werron, T. (2015). How are Football Games Remembered? Idioms of Memory in Modern Football. In W. Pyta, N. Havemann (Hg.), *European Football and Collective Memory* (S. 18–39). Palgrave.

Quellen

- Alcock, C. W. (Hg.) (1871). *Football Annual. Virtue*.
- Alcock, C. W. (Hg.) (1872). *Football Annual. Virtue*.
- Alcock, C. W. (Hg.) (1873). *Football Annual. Virtue*.
- Alcock, C. W. (Hg.) (1876). *Football Annual. Cricket Press*.
- Alcock, C. W. (Hg.) (1877). *Football Annual. Sharp & Co.*
- Alcock, C. W. (Hg.) (1878). *Football Annual. Sharp & Co.*
- Alcock, C. W. (Hg.) (1879). *Football Annual. Cricket Press*.
- Alcock, C. W. (Hg.) (1881). *Football Annual*.
- Alcock, C. W. (Hg.) (1888). *Football Annual. Wright & Co.*
- Alcock, C. W. (Hg.) (1889). *Football Annual. Wright & Co.*

- Alcock, C. W. (Hg.) (1892). Football Annual. Wright & Co.
- Alcock, C. W. (Hg.) (1993). Football Annual.
- Alcock, C. W. (Hg.) (1994). Football Annual. Wright & Co.

Die Un/Sichtbarkeit des Regelhüters

Zwei Strategien zum Umgang mit der Unterdeterminiertheit von Regeln im Schiedsrichterwesen

Max Weigelin

Einleitung

Der folgende Beitrag befasst sich mit Schiedsrichtern im Verbandsfußball.¹ Ziel ist dabei, dass kulturanalytische Befremdungspotential des bewertungssoziologischen Blicks auf den Fußball zu zeigen. Schaut man nämlich mit der Bewertungssoziologie auf diesen sportlichen Wettkampf, fällt auf, dass eine zentrale Bewertungsfigur der Wettkampfpraxis von der sportsoziologischen Forschung bislang sträflich vernachlässigt wird: der Schiedsrichter. Bewertungssoziologische Instrumente helfen, dessen Praxis in ihrer Instabilität, Ambivalenz und Komplexität ernst zu nehmen. Andererseits adressiert der Beitrag sich andeutende Potentiale in der Bewertungssoziologie, die darauf hinauslaufen, gezielt Synergien zwischen der Soziologie der Bewertung und ähnlichen Forschungsprogrammen zu entwickeln. Meine Überlegungen zum Fall des Fußballschiedsrichters stellen in diesem Sinne eine Verbindung zwischen Soziologie der Bewertung und der Soziologie des Wettbewerbs her. Dieses doppelte Erkenntnisinteresse – Kulturosoziologie des Sports und Soziologie der Bewertung – wird im Folgenden entfaltet, indem das Problem des Hantierens mit formalen Regeln im Fußball empirisch beschrieben wird.

¹ Um den Gebrauch des generischen Maskulinums zu vermeiden, wechsele ich im folgenden Text fließend zwischen männlicher und weiblicher Form des Begriffs »Schiedsrichterin«. Ich spreche durchgehend von »Spielern«, weil es sich beim Frauenfußball – aus der Perspektive einer Ethnografie des Schiedsrichterwesens – um ein spezielles Separatmilieu handelt. Zur Spielleitung berechtigt sind dort nur weibliche Schiedsrichter. Leider hat sich zum Zeitpunkt der Anfertigung dieses Beitrags noch keine Gelegenheit ergeben, auch solche Spiele zu untersuchen.

Dafür gliedere ich meine Darstellung in fünf Schritte. Ich setze mich dazu zunächst theoretisch mit Potentialen des Falls Fußballschiedsrichter für eine Soziologie der Bewertung und des Wettbewerbs auseinander (1). Danach schildere ich kurz das Forschungsdesign der Studie, auf der dieser Aufsatz beruht (2.). Im empirischen Teil beschreibe ich am empirischen Material unterschiedliche Formen der Handhabung formaler Regeln in der Wettbewerbspraxis des Fußballs, indem ich die Rolle der Schiedsrichterin hierbei fokussiere. Zuerst schildere ich eine Reihe von Infrastrukturen, Praktiken und impliziten Verhaltenserwartungen, die ich unter dem Begriff der Technisierungsstrategie zusammenfasse (3.). Anschließend befasse ich mich mit einem Komplex von Praktiken und Verhaltenserwartungen, die ich – von der Schiedsrichterperspektive aus formuliert – als Mitspielen bezeichne. Aus der Perspektive von Spielerinnen und Publikum lässt sich dieser Aspekt des Fußballwettkampfs als »Spiel im Spiel« (Gebauer 2006) bezeichnen, weil es hierbei darum geht, das fortlaufende Operieren des Schiedsrichters mit der Unterdeteminiertheit der Regeln als kontrovers und einzelne Pfiffe als Fehlentscheidungen zu dramatisieren (4.). Im Fazit unterstreiche ich, dass sich die Soziologie des Wettbewerbs nicht nur mit der Institutionalisierung, sondern auch mit der praktisch-situativen Vollzugswirklichkeit des regelgeleiteten Wettkampfgeschehens befassen sollte. Zwischen den Polen der Technisierung und des Mitspielens spannt sich ein breites Spektrum unterschiedlicher Formen des Hantierens mit Regelwerken in Wettkampfpraktiken verschiedener Sportarten auf (5.).

1. Der Schiedsrichter als instruktiver Fall für die Soziologie der Wettbewerbsformate

Die Soziologie der Bewertung konvergiert gegenwärtig mit verschiedenen anderen sozialtheoretischen Trends. Einerseits erscheint es fruchtbar, Bewertung seitwärts zu relationieren mit Phänomenen auf einer ähnlichen begrifflichen Ebene. So schlägt Heintz beispielsweise vor, Bewertung als ein spezifisches modernes »Beobachtungsformat« zu betrachten und es als solches in eine Reihe zu stellen mit anderen solchen Beobachtungsformaten, wie der Kategorisierung, dem Vergleich und der Quantifizierung (Heintz 2021). Eine andere Möglichkeit, Bewertungssoziologie konzeptionell mit anderen soziologischen Forschungstrends zu verknüpfen, besteht dagegen darin, Bewertungsphänomene einzurordnen – oder genauer noch: unterzuordnen – in spezifischere gesellschaftliche Trends, vor allem der Digitalisierung und

der Ausbreitung von New-Public-Management-Regimen (Heintz 2021b). Insbesondere wenn man Letzteres betont, ergibt sich ein konzeptionell enger Zusammenhang zwischen der Forschung über die gesellschaftliche Konstruktion von Konkurrenz- und Wettbewerbsverhältnissen (vgl. Werron 2015; Arora et.al. 2021) und der Soziologie der Bewertung (zum Beispiel bei Ringel/Werron 2020). Im Folgenden arbeite ich deshalb von spezifischen empirischen Bewertungspraktiken ausgehend, aber versetze diese Beobachtungen begrifflich in einen wettbewerbstheoretischen Horizont, indem ich dabei Begriffe und Phänomene wie das Regelwerk, die Fairness und den Wettbewerb selbst zum Thema mache. An dieser Stelle scheint daher ein kurzer Blick auf die Programmatik einer Soziologie der Konkurrenz bzw. des Wettbewerbs geboten.

In einem rezenten programmatischen Aufschlag dazu halten Arora et al. fest, dass die soziologische Theoretisierung der Konzepte von Konkurrenz und Wettbewerb dominiert wird von Beobachtungen, die am Phänomen des ökonomischen Marktes ansetzen bzw. die Begriffe im Grenzfall auch in Eins setzen (vgl. Arora et.al. 2021: S. 5ff., sowie Meier in diesem Band). Demgegenüber gelte es, die Forschung zu öffnen für eine komparative Soziologie der Konkurrenz, welche die Entstehung, Aufrechterhaltung und Formenvielfalt von Konkurrenzen mit ihren je feldspezifischen Kontextuierungen ernst nimmt (vgl. ähnlich: Werron 2015). Forschungen zu diesem Thema durchziehen die Soziologie in Vergangenheit und Gegenwart. Die Etablierung einer systematischen und programmatisch ausformulierten Soziologie der Konkurrenz und des Wettbewerbs nimmt allerdings erst in jüngerer Zeit Fahrt auf (vgl. die Beiträge in Arora et.al. 2021). In Anschluss an das Programm von Arora et al. gehe ich davon aus, dass es sich bei Wettbewerben um eine sozial konstruierte und spezifisch institutionalisierte Form von Konkurrenzverhältnissen handelt, die eine erhebliche Formenvielfalt und -wandelbarkeit aufweist und empirisch offen zu untersuchen ist. Insofern gilt es, nur Minimalkriterien zur Beschreibung der Form des Wettbewerbs und keine theoretische Ausarbeitung zugrunde zu legen, insofern letztere Gefahr läuft, Sonderfälle als Modelfälle zu setzen. Grundlegend sind für einen solchen minimalistischen Begriffsrahmen dabei die Annahmen, dass die Konstruktion eines Konkurrenz- oder Wettbewerbsverhältnisses die Konstruktion eines Gutes als knapp und begehrte beinhaltet und darüber hinaus einen bestimmten Kreis von Akteuren als konkurrenzhaft aufeinander bezogen hervorzubringen versucht. Der Sport konstruiert insofern das Gut des Siegens als knapp, indem es Wettkämpfe nach Nullsummenlogiken organisiert; als begehrte, indem es

diese Siege als Ausdruck von Leistungsfähigkeit interpretierbar macht und als Erfolge valorisiert; und adressiert so Sporttreibende und Zuschauende (als deren »Anhänger«) als mehr oder weniger leistungsbereite und -fähige Subjekte.

Innerhalb einer solchen Soziologie der Konkurrenz kann der organisierte Verbandssport meiner Einschätzung nach deshalb eine in dreierlei Hinsicht besonders instruktive Rolle einnehmen. Erstens besteht die grundlegendste Herausforderung für die Soziologie der Konkurrenz darin, die Vorstellung einer natürlichen Gegebenheit von Konkurrenz zu überwinden. Sport ist hier ein instruktiver Fall, weil die Genese, Ausbreitung und Erzeugung des Konkurrenzregimes Leistungssport unter anderem mit Hilfe seiner verbandsförmigen Organisation hier aufgrund seines historisch jungen Status gut nachvollziehbar sind (Minnetian/Werron 2022; Dies. In diesem Band). Zweitens lässt sich der Sport auch als ein Fall untersuchen, an dem sich eine Art gesellschaftlicher Institutionalisierung und Aufführung eines Idealbildes des Wettbewerbs beobachten lässt (vgl. Schmidt et.al. in diesem Band) und der deshalb auch als Metapher-Vorrat für Konkurrenz- und Wettbewerbssemantiken in anderen Feldern fungiert (vgl. Bröckling 2020). Insofern kann eine vom Fall des Sports ausgehende Soziologie der Konkurrenz und des Wettbewerbs insbesondere auch zu Fragen der Dissemination und Ausbreitung von Konkurrenzsemantiken und –dispositiven (vgl. Arora 2021: S. 3ff., Werron 2015: S. 189ff.) instruktiv sein. Drittens, und dafür scheint mir die Soziologie des Wettbewerbs – wie sie bei Arora et al. oder auch bei Werron (2015) projektiert und betrieben wird – noch wenig sensibilisiert, kann insbesondere eine mikrosoziologisch arbeitende und mit der Bewertungssoziologie für Ambivalenzen von Valorisierungsprozessen und Bewertungspraktiken sensibilisierte Soziologie des Wettbewerbs die Selbstidealisierung von Wettbewerbssemantiken durch eine empirisch reichhaltige Explikation der Widersprüche, Spannungen und Ambivalenzen konkreter Vollzugswirklichkeiten von Wettbewerben bereichern. Insbesondere an dieser Stelle sehe ich den theoretischen Einsatz der hier vorgelegten Fallbeschreibung.

In diesem Sinne interessiert mich an bewertungssoziologischen Arbeiten vor allem deren Potential zur Explikation konkreter Vollzugslogiken von Bewertungspraktiken und Valorisierungsprozessen, wie es sich insbesondere in mikrosoziologischen Beiträgen zeigt (Hoppe 2022, Lambrix 2022, Wagner et.al. 2023). Darüber hinaus nehme ich zwei weitere Anregungen aus jüngeren Debattenschwerpunkten innerhalb der Bewertungssoziologie auf. Erstens hat es sich für meinen Fall als lohnend herausgestellt, die

Begriffe der Kategorisierung, des Vergleichs, der Bewertung und der Entscheidung am empirischen Fall gezielt auf ihre Verknüpfungslogiken hin zu befragen. So stellt sich beispielsweise der Wettkampf zunächst als institutionalisiertes Format zur Konstruktion eines knappen Guts (Siege) und von Wertunterschieden zwischen Konkurrenzteilnehmern (Leistungsunterschiede von Sportlerinnen) dar. Dies gelingt insofern der sportliche Wettbewerb als Vergleichsarrangement angelegt. Und dieses setzt wiederum bestimmte Kategorisierungspraktiken – wie zum Beispiel die Festlegung von Spielzuständen (Tor oder kein Tor?) durch Spielteilnehmer voraus. In einem anderen Paper habe ich außerdem argumentiert, dass der primäre Beitrag der Schiedsrichterin zum Fußballspiel zunächst darin gesehen werden kann, dass sie das Problem dieser für Spielverlauf konstitutiver Herstellung von formalen Spielereignissen durch Kategorisierungspraktiken über eine organisational verbrieft Deutungshoheit (Pfiffe) auflöst. Es ist äußerst typisch für den Fußball, aber auch für viele andere Sportarten, dass dieses formale Monopol in der Praxis auf komplexe Aushandlungsprozesse trifft, die alle Spielteilnehmer involviert (Spielerinnen, Trainern aber auch Zuschauende). Spezifische Formate der Bewertungskommunikation, wie das »Reklamieren« und »Meckern«, laden diese Kategorisierungssituation auf und transformieren sie in eine Entscheidungssituation für den Schiedsrichter (vgl. Weigelin 2022). In Fortführung dieser Überlegungen fokussiere ich im Folgenden die Frage, wie das Fußballspiel mit seinem formalen Regelwerk umgeht. Einige Ausgangspunkte für ein solches Unterfangen finden sich in der sportsoziologischen Forschung.

Mit Blick auf diese Forschung ergibt sich aber zunächst ein Problem. Die soziologische Forschung zu den formalen Regelwerken des Sports fokussiert nämlich sehr stark Fragen der Entstehung solcher Kodifizierungen (vgl. Elias/Dunning 1982), ihrer fortlaufenden Transformation (etwa Kew 1987) und ihrer Feld- oder systemspezifischen Funktionen (vgl. Schimank 2005). Damit liefern diese Arbeiten durchaus einen interessanten Ausgangspunkt, um über die wettbewerbssoziologische Frage nach der Entstehung und Kontingenz von Wettbewerbsformaten nachzudenken. Fokussiert man aber die Frage der Kodifizierung und evolutionären Anpassung von Regelwerken an Spielumstände, geht der Blick für das komplexe Verhältnis von formalem Regelwerk und situativer Wettkampfpraxis verloren. Spiel- oder Wettkampfpraxis sind dann eine Blackbox, deren Intransparenz mit der Common-Sense-Idee überspielt wird, wonach formale Regeln etwas an sich eher Unkompliziertes sind, das richtig oder falsch angewendet werden kann (vgl. dazu Wittgenstein 1984). Im Folgenden versuche ich mit diesem falschen Vertrauen in die regu-

lierende Kraft der Regeln zu brechen, indem ich mit Wittgenstein von einer Unterdeterminiertheit sozialer Praxis durch ihre Regeln und mit Kew (1992) spezifischer noch von der Unterdeterminiertheit des sportlichen Wettkampfs durch seine formalen Regelwerke ausgehe. Damit verschiebt sich aber die empirische Beobachtungsweise des Problems gegenüber geläufigen Feldperspektiven: Es geht nicht mehr um die Frage, was nun die richtige Auslegung der Regeln beinhalten würde (das wären Gegenstände für normative Diskurse der Sportethik oder des Sportrechts), sondern darum zu beschreiben, wie es gelingt, die Wettkampfpraxis als formal reguliert darzustellen, obwohl die Antwort auf die Frage – ›Welches Vorgehen schreiben die formalen Regeln vor? – nicht auf der Hand liegt.

Einige Spielformen und speziell Leistungssportformen mit ihrer Kompetitivität akzentuieren diese Unterdeterminiertheit der Regeln laufend, geradezu als eine Art klaffende Regelungslücke. Man gewinnt in solchen Fällen wie dem Fußball mit seinen hitzigen Diskussionen um Schiedsrichterentscheidungen fast den Eindruck, dass der Umgang mit der Unterdeterminiertheit der Regeln nicht zufällig Bestandteil der Spielpraxis geworden ist, sondern auch positive Funktionen erfüllt. Soziologische Studien zur situativen Enaktierung von Regelwerken in Sportspielen sind zwar relativ rar, liefern aber einige Hinweise. DeLand weist für den Fall des »Pick Up Street Basketball«, eine Spielform, die ohne einen Schiedsrichter auskommen muss, darauf hin, dass der hier typische »litigious way«, indem Spielereigniss-Kategorisierungsfragen verhandelt werden, das Spiel symbolisch als ernste Angelegenheit bzw. »not just playing around« (DeLand 2013: 659) rahmt. Sport wird hier also, so Deland in Anschluss an Goffman, mit kommunikativen Verhandlungspraktiken aus Rechtsrahmen strategisch moduliert. Für den Fall des Spiels, das von Schiedsrichtern geleitet wird, stellt sich die Situation notwendigerweise anders, aber nicht unähnlich dar. Heck weist in einer Studie, die schiedsrichtergeleiteten Verbandsfußball und Fußball ohne Schiedsrichter vergleicht, darauf hin, dass die Delegation der Auflösung von Regelungslücken an den formal-autorisierten neutralen Dritten Spielerinnen von Kooperationsanforderungen entlastet und das Potential einer »professionellen Unfairness« dadurch steigert (Heck 2019: 48). Wenn Spielregeln nicht aus eigener Kraft definieren können was der Fall ist, die Spielkultur ihre Spieler*innen zu quasi-juristischer Streitlust (Deland 2012) und nicht konsensorientierter (Heck 2019: 40ff.) im Umgang mit der Unterdeterminiertheit der Regeln anhält (wie im Verbandsfußball), werden Fragen der Kategorisierung von Spielereignissen (Tore, Fouls usw.) durch Kommunikation sich widersprechender Bewertungen der Spielteilnehmer (Reklam-

mation, Beschwerde usw.) strittig und die Pfiffe der Schiedsrichterin zunehmend entscheidungsförmig (vgl. Weigelin 2022).

Im Folgenden will ich diesen Zusammenhang von Wettkampf, Regeln und Bewertung aber nicht nur auf der situativen Ebene der Vollzugswirklichkeit des Einzelwettkampfs, sondern auch auf der intersituativen Ebene des durch den Verband organisierten Wettkampfbetriebs in den Blick nehmen. Ein interessanter Hinweis dazu stammt aus der Studie zum Schiedsrichterwesen im Profi-Eishockey von Rains. Er hat darauf hingewiesen, dass die situativ erzeugte und in der massenmedialen Nachberichterstattung weiterverarbeitete Strittigkeit von Schiedsrichterentscheidungen den Verband in die Lage bringt, auf der Organisationsschauseite (vor allem gegenüber den Sportvereinen und der Sportpresse) eine professionelle Produktion von Fairness aufzuführen. In den Ergebnissen seiner Feldforschung stellt sich diese Organisationstätigkeit als Darstellen von Verfahrenstechniken zur Gewährleistung von Konsistenz der Regelauslegungen über die Gesamtheit der Einzelwettkämpfe (Rains 1984) dar. Die oben genannte Frage – wie es die dramaturgische Struktur des Fußballspiels ermöglicht, die Wettkampfpraxis als formal reguliert darzustellen – muss also ergänzt werden durch die Frage, in welcher Form diese Struktur es ermöglicht, diese Regulation als »fair« im Sinne intersituativer Konsistenz der Regelauslegungen über verschiedene Wettbewerbssituationen hinweg zu gewährleisten. Bevor ich eine Möglichkeit, diese Frage am empirischen Material zu beantworten, entfalte, möchte ich noch kurz näher auf das Forschungsdesign meiner Studie eingehen.

2. Organisationsethnografie im Schiedsrichterwesen

Die Studie, auf der dieser Beitrag basiert, versteht sich als empirisches Forschungsprojekt zu den Themen »Schiedsrichter und Schiedsrichterwesen im Fußball«, das sich am Anspruch der theoretischen Empirie (Hirschauer 2011) orientiert. Damit einher geht zunächst einmal die Betonung der Offenheit des qualitativen Forschungsprozesses. Konkret bedeutet das, dass die möglichen theoretischen Einsätze der Arbeit nicht vor dem Kontakt mit dem Feld vollständig (oder wenigstens hauptsächlich) antizipiert werden müssen, sondern sich im Prozess der Feldforschung als zirkulär-iterativem Prozess erst sukzessive ergeben. Dieser Empirismus wird aber nicht erkauft mit einem Theorieverzicht. Um sowohl dem empirischen Fall als auch den Ansprüchen einer theoriegenerierenden soziologischen Forschung gerecht zu werden, gilt

es daher den analytischen Tunnelblick auf die theoretisch relevanten Phänomene und den holistischen Blick für das Feld als Ganzes auszubalancieren.² Die flexible und an die Gegenstandseigenschaften anpassbare ethnografische Methode musste insofern im Laufe meines Forschungsprozesses in zweierlei Hinsicht entfaltet werden: Einerseits bin ich einer holistischen kulturellen Neugier für das Feld in möglichst vielen Facetten gefolgt. Ich habe daher nicht nur mit Schiedsrichtern, sondern auch unterschiedlichen Funktionären des Schiedsrichterwesens (Lehrwarte, Ansetzer usw.) genauso wie mit Funktionären angrenzender Organisationsbereiche (zum Beispiel Vorsitzende von Sportgerichten) gesprochen. Andererseits galt es, Zugänge zu sehr spezifischen Phänomenen zu entwickeln, und diese systematisch zu fokussieren. In diesem Sinne habe ich beispielsweise die Begleitung von sogenannten »Schiedsrichterbeobachtern« – Evaluatoren des Schiedsrichterwesens, die Schiedsrichtereinsätze beobachten und auf »Leistung« hin bewerten – fokussiert, welche es ermöglichte, insidertypische Praktiken des Sehens anzueignen (vgl. Schindler/Lieg 2013).

Zu Beginn hat die Studie den Fußballplatz (und das Stadion) mit Goffman vor allem als eine Interaktionsordnung untersucht. Im Laufe der Zeit hat die Arbeit aber auch immer stärker Züge einer Organisationsethnografie angenommen. Der folgende Beitrag verknüpft dieses situationistische Anfangsinteresse mit organisationstheoretischen Anreicherungen, indem Theorieprobleme rund um den Begriff der Regel genutzt werden, um das Material für die Fallbeschreibung zu mobilisieren. Ich greife dabei weniger auf ethnografische Explikationen durch den soziologischen Beobachter zu als auf Selbstbeschreibungen des Feldes, die ich in Interviews und Gesprächen selbst mitproduziere oder frei verfügbaren Felddokumenten entnehme. Ich nutze diese, um normative Strukturen der Interaktionsordnung in diesem Feld zu explizieren, die ich mit Goffman als situative Engagementgebote (2018: 49ff.) auffasse.

Die Darstellung des empirischen Materials erfolgt, wie gesagt, orientiert am Problem der Unterdeteminiertheit der Regeln und dem sich daraus ergebenden Problem der Darstellung der Regelgeleitetheit des Spiels, für die der Schiedsrichter eine zentrale Stellung einnimmt. Ich beschreibe zunächst, wie diese Unterdeteminiertheit mit dem Engagementgebot der unsichtbaren Spielpartizipation und dazugehörigen Verhaltensrepertoires technisch gelöst wird (3.). Daran anknüpfend beschreibe ich, wie das Spiel diese Strategie der

² Vgl. zu diesem Spannungsverhältnis, allerdings mit deutlicher (wie gut begründeter) Präferenz für den analytischen Tunnelblick: Zerubavel 1980.

bürokratischen Technisierung durch die Dramatisierung von Kategorisierungsfragen unterläuft und die Spielregeln so zum Gegenstand eines »Spiels im Spiel« macht, das von Schiedsrichterinnen ein situativ flexibles Mitspielen fordert, was sich im Engagementgebot zum Managen von Konflikten äußert (4.).

3. Das Unsichtbar-Machen der Regelungslücke: Technisierung

Technisierung löst das Problem der Unterdetiniertheit der Regeln durch das unsichtbar Machen der Kontingenz der Regelauslegung. Ich benutze diesen Begriff, weil Schiedsrichter selbst Begriffe wie »technischer Pfiff« – im Unterschied zum »strittigen Pfiff« – verwenden, er sich darüber hinaus aber auch als anschlussfähig an soziologische Begriffsbildung erweist. Mit Goffman etwa lässt sich »Technik« dann als Deutungsinstrument verstehen, das Phänomene einer natürlichen (anstatt einer sozialen) Rahmung zurechnet und damit nicht nach der Darstellung von Handlungen, Tätern und Urteilen, sondern nach Ursachen und Wirkung verlangt (vgl. Goffman 1974: 21ff.). Ich untersuche diese Darstellungsstrategie, indem ich eine Reihe von Infrastrukturen, Praktiken und impliziten Verhaltenserwartungen in den Blick nehme. Diese werden insbesondere erkennbar mit Blick auf die öffentliche Selbstdarstellung des Verbandsfußballspiels in seiner Offiziösität und Interviewaussagen von Schiedsrichterinnen bezüglich des normativen Sinns ihrer Tätigkeit.

Die Strategie der Technisierung drückt sich zunächst in einem zentralen Engagementgebot der Schiedsrichterpraxis aus, das ich Unsichtbarkeits-Gebot nenne. Auf dieses stoße ich in meinem Interviewmaterial unter anderem in Antworten auf die Frage »Was macht deiner Meinung nach einen guten Schiedsrichter aus?«. In rund der Hälfte der Fälle erhalte ich eine Antwort wie diese:

Also ja. Da würde ich jetzt glaub ich tatsächlich mit ner Plättiude antworten. Wenn nach dem Spiel keiner über dich redet, hast du alles richtiggemacht.

Wie aber wird der Schiedsrichter sichtbar? Zunächst einmal darf er sich nicht »angreifbar« machen. Insofern ist der Schiedsrichterhabitus – etwa im Vergleich zum Spielerhabitus – im Fußball mit Restriktionen des Ausdrucksrepertoires belegt. Auf einem »Neulingslehrgang« artikuliert ein Schiedsrichter-

funktionär im Rahmen seines Vortrags diese informelle Regel am Beispiel der Schuhfarbe:

Hannes erläutert jetzt die Liste mit den Utensilien, die ein Schiedsrichter immer zum Einsatz mitzubringen hat. Auf der Power-Point-Folie dazu findet sich eine Liste. Darunter auch das Stichwort »Schuhe«. »Schwarze natürlich!« sagt er mit Nachdruck. Spieler von heute würden ja Fußballschuhe in allen Farben tragen, erläutert er abfällig. Aber das gehe für den Schiedsrichter nicht. Gelbe etwa: »Wie sieht das denn aus?« »Mir sind ja keine Pinguine.« (Ihm schwebt wohl das Bild vom typischen schwarzen Schiedsrichtertrikot mit gelben »Füßchen« vor.) Lachen in der Runde.

Diese expressive Zurückhaltung findet sich auch auf der Ebene von Gefühlausdrücken. Man müsse, so erläutert mir ein Schiedsrichter im Interview, auch Aggressivität, die einem entgegengebracht wird, spiegeln, aber in gezielter Weise:

Natürlich: Wenn der dich anbrüllt, dann gibst du auch mal ein bisschen Pfeffer zurück. Sonst nimmt der dich ja gar nicht ernst in dem Moment. [...] Aber die goldene Regel ist: Niemals lauter als der Spieler!

Auch gestische Repertoires, die ein spielerisches Vergnügen beim Ausüben der Spielteilnahme anzeigen, die bei Fußballspielern bis zu einem gewissen Grad – etwa als »Tricksen« (Schmidt 2012: S. 90–91) – als angemessen gelten, sind für Schiedsrichterinnen eher unangemessen. Ein schlechter Schiedsrichter gilt in diesem Sinne als »Selbstdarsteller«. Ein Beispiel für diese Darstellung des Gebots zur Zurückhaltung anhand der Kritik der Selbstdarstellerei stellt die im Feld recht virulente Anekdote vom »Eifel-Django« dar. Ich zitiere zunächst aus einem Protokoll und dann aus einem Zeitungsartikel:

Den Namen »Eifel-Django« habe ich heute während der Autofahrt (mit drei Schiedsrichtern und einer Schiedsrichterin unterwegs zum Leistungstest) aufgeschnappt. Er war allen geläufig außer mir. Der Spitzname wurde ihm von einer freudig-erregten Sportpresse verliehen, nachdem dieser während eines Drittligaspiele erstmalig für einen DFB-Schiedsrichter beim Vergeben von zwei gelben Karten in einer Situation nicht eine Karte in einer Hand zwei Mal hochgehoben hatte (typisch: nacheinander, jeweils einmal zu dem bestraften Spieler gerichtet), sondern (wie der Western »Django« mit seinen zwei Kanonen) indem er in jeder Hand eine Karte hielt und

sie gleichzeitig beiden Spielern entgegenstreckte. Im Auto hieß es, diese Aktion habe »seine Karriere beendet«. Über den im Schiedsrichterwesen internen Nachhall lese ich in einem Artikel auf Welt.de (2008):

»Beim DFB fürchtet man um den seriösen Ruf der Unparteiischen. Wenn ihm im Spiel ein dicker Fehler unterläuft, redet niemand mehr von einem netten Gag. Jeder fragt sich dann, wie der DFB einen solchen Selbstdarsteller zu einem Bundesligaspield schicken kann«, schimpfte Krug [DFB Offiziel-ler, Anm. MW] nach der Aktion von Metzen, und Top-Schiedsrichter Herbert Fandel machte der Auftritt »sprachlos.«

Bestimmte expressive Zurückhaltungs- bzw. Seriösitätsrepertoires, so sehen es die ranghohen Funktionäre des Schiedsrichterwesens, sind also kein Selbstzweck, sondern dienen vor allem dazu, im Moment des Sichtbar-Werdens des Schiedsrichters durch Dramatisierung von Fehlentscheidungen in einer guten Darstellungsposition zu sein. Das Bemühen um keinerlei unnötige Sichtbarkeit in den Repertoires der Seriösität unterstreicht also gewissermaßen symbolisch, dass Schiedsrichterinnen das Unsichtbar-Bleiben im Sinne eines Fehler-Vermeidens als ihre Hauptaufgabe betrachten. Ein Schiedsrichter erläutert mir im Interview, warum »Fehlentscheidungen« und eine inkonsistente Spielleitung ein Problem darstellen:

Also das Beste ist, wenn die das Spiel unter sich ausmachen. Wenn ich nur im Hintergrund das ermögliche.

Man könnte im Anschluss an diese Begründungslogik sagen: Die Unsichtbarkeit des Schiedsrichters für das Zustandekommen eines Wettkampfergebnisses ist konstitutiv dafür, dass dieses Ergebnis den Spielenden als »Leistung« zugerechnet werden kann.³ Integer bzw. fair ist der Wettbewerb demnach

3 Es ist daher nur konsequent, dass das Regelwerk bis vor Kurzem dem Schiedsrichter zugleich das Privileg einräumte sich auf dem Spielfeld zu bewegen und ihn als virtuell körperlich abwesend behandelte, als »Luft«, wie es im Feld hieß. So konnte es passieren, dass ein Schiedsrichter kausal durch einen Zusammenstoß mit dem Ball ein Tor verursacht, dieses formal aber den Spielenden zugerechnet werden muss. Ein Fall in dem ein Schiedsrichter im Profibereich sich derart ungeschickt bewegte, dass die Presse (man schaue sich die Szene auf Youtube an) nicht völlig unplausibel formulierte, er habe ein Tor »geschossen«, ereignete sich noch im Jahr 2019 wenige Monate vor Inkrafttreten der Regeländerung (vgl. Stuttgarter Zeitung 2019). Die neue Regel sieht eine Spielunterbrechung (»Schiedsrichterball«) bei physischen Kontakt von Schiedsrichter und Spielgerät vor.

dann, wenn der Schiedsrichter in diesem Sinne unsichtbar bleibt.⁴ Wie also lösen Schiedsrichterinnen und Schiedsrichterwesen die Regelungslücke (im Idealfall) fortlaufend in einer Weise auf, die es ermöglicht, dass diese Lücke und damit auch ihr Beitrag zum Spiel in der Praxis gar nicht erst als Problem erkennbar wird? Dazu formatiert die Praxis des Fußballspiels die Ordnungsleistungen des Schiedsrichters als technisierte Wahrnehmung von Tatsachen. Wie genau funktioniert das?

Zunächst geschieht dies, indem auf formaler Ebene diese Kategorisierungsleistungen des Schiedsrichters mit dem eigentümlichen Begriff der »Tatsachenentscheidung« gerahmt werden. Verstanden als Tatsachenentscheidung hat der Pfiff eine irreversible Gültigkeit. Der Begriff der »Tatsachenentscheidung« ist zunächst einmal nicht so sehr Bestandteil eines esoterischen Spezialisten-Vokabulars, sondern im Feld – d.h. unter Spielerinnen, Trainern und Zuschauerinnen – weithin geläufig. Im offiziellen Regelwerk findet er sich zunächst nicht. Eine Formulierung, prominent platziert in »Regel 5 Der Schiedsrichter«:

[...] die Entscheidungen des Schiedsrichters zu Tatsachen sind endgültig [...].
(FIFA 2021: 34)

Praktisch bedingt dieser Passus, dass Klagen gegen die Falschheit der Inhalte von Schiedsrichterentscheidungen gegenüber dem Verband und seinem »Sportgericht« im Nachgang eines Spiels keine Erfolgsschancen haben. Das formale Konstrukt unterstreicht so den performativen ontologischen Status des Schiedsrichterpiffs, im Sinne eines »Foul ist, wenn der Schiedsrichter pfeift«. Seine Performativität wird formal der Frage nachvollziehbarer Referenz völlig enthoben. Man weiß im Feld daher aus Erfahrung, dass der einmal getätigte Pfiff fast ausnahmslos endgültig Bestand hat.⁵ Zwischen Pfiff und

4 I.d.S. auch sportphilosophisch Gebauer (2006): »Er handelt nicht in seinem eigenen Namen; er drückt nicht dem Spiel seine Marke auf wie ein Spielmacher, sondern ist ein Sachwalter der Regeln« (135) Für Gebauer verleiht die Schiedsrichterin den Regeln »Kraft«. Er darf nicht »zu schwach« sein, sonst würden die Regeln nicht befolgt, nicht »zu stark«, sonst tritt die Regel hinter seiner Person zurück (vgl. ebd.: 136).

5 Sehr wohl aber können Vereine aber bei »Formfehlern« von Schiedsrichtern Einspruch gegen die Spielwertung einlegen. Diese sind der Organisation wie der einzelnen Schiedsrichterin besonders peinlich (vgl. auch Rains 1984: 154f.) auch wenn sie keine Implikationen im Sinne einer Anfechtbarkeit des Spielergbnisses haben.

Wieder-Anpfiff (formal »Spielfortsetzung«) gibt es zwar ein kurzes Zeitfenster, das zur Revision des Pfiffs genutzt werden könnte, aber dies geschieht so gut wie nie. Möglich ist eine Ausnahme von dieser Regel, laut meiner Gesprächspartner, wenn zum Beispiel der gefoulte Spieler gesteht, dass er nicht gefoult wurde.⁶

Dramaturgisch reflektiert wird das so entstehende und im Sonderbegriff der »Tatsachenentscheidung« linguistisch deutlich markierte Willkürpotential durch eine bestimmte Dramaturgie des Wahrnehmens. Ich möchte zwei Grundzüge dieser Dramaturgie anhand eines Interviewauszugs herausarbeiten:

Und dann hast du halt nur Millisekunden Zeit. [...] Jede Sekunde, die du länger nachdenkst, und wir reden wirklich über Sekunden, die du dastehst, Foul gepfiffen hast und willst rot geben, und dann nimmst du dir nochmal fünf Sekunden, weil du willst darüber nachdenken, dass sieht nach Außen hin. Ja, schon dann... Dann nimmt dir das keiner mehr. ›Das hat der sich gut überlegt‹ oder so. Sondern ›der ist sich halt unsicher!‹

Erkennbar wird an dieser Beschreibung und den darin explizierten Erwartungserwartungen des Schiedsrichters, genauso allerdings auch an der Beobachtung tatsächlichen Schiedsrichterverhaltens in Spielen, dass die typische Dramaturgie des Schiedsrichter-Pfiffs den Schiedsrichter als Kategorisierer eines Spielereignisses darstellt, der auf *bloße Wahrnehmung* zurückgreift. Das »Zögern« im Pfeifen wird ausgelegt als Moment des Nachdenkens, Abwägens oder sogar kommunikativer Offenheit für Impulse anderer. Tritt ein »spätes« Pfeifen wiederholt auf, tritt typischerweise der Eindruck auf, ein Schiedsrichter pfeife »auf Zuruf«. Die Dramaturgie des Pfiffs als »Tatsachenentscheidung« platziert den einzelnen Kategorisierungsakt des Schiedsrichters aber als kommunikativ unzugänglich – Diskussionen im Anschluss sind üblich,

6 Etwas anders sieht es seit der Einführung des sogenannten Video Assistant Referees in bestimmten Ligen des Profifußballs aus. Dabei handelt es sich um einen »Videoschiedsrichter«, der mit dem »On Field« Schiedsrichter per Headsetfunk verbunden ist und in Echtzeit Videomaterial des Spielgeschehens sichtet. Unter bestimmten Umständen kann er dem Schiedsrichter eine Revision empfehlen und dafür Videomaterial am Spielfeldrand einblenden. Auch hier ist allerdings die Möglichkeit der Korrektur beschränkt auf den Moment vor der nächsten formalen »Spielfortsetzung«. Und auch hier ist nicht vorgesehen, dass ein Schiedsrichter sich im Zeitfenster der möglichen Revision von protestierenden Spielern umstimmen lässt.

aber sollen nicht zur Revision führen – und um jeden weiteren Anschein von Deliberation zu zerstreuen, darf er noch nicht einmal als überlegt dargestellt werden. Insofern hat der Begriff der Schiedsrichterin etwas Irreführendes an sich. Der Pfiff ist nicht eingebettet in eine Verhandlung. Vor dem Pfiff gibt es keine Anhörung. Die Schiedsrichterin schiedst nicht. Die Schiedsrichterin im Sinne der Technisierung ist eine Apparatur zur qualitativen Kodierung. Die Kategorisierung »ergibt« sich für den Schiedsrichter mit »konzentrierter« Wahrnehmung aus dem Ablauf des Spiels und der möglichst »korrekten« Registratur bestimmter entscheidender Spielereignisse.⁷ Diese Auffassung fließt in die Selbstbeschreibung der Schiedsrichterinnen ein. Erkennbar ist das beispielsweise daran, wie ein Schiedsrichter im folgenden Interviewauszug diesen Kodierungsvorgang retrospektiv rekonstruiert:

Und dann kommts einfach auch drauf an. Wie fokussiert bin ich auf ne gewisse Situation? Ich bin da ein bisschen bekloppt. Ich bereite mich auf Situationen vor, ohne dass sie passieren. D.h.: Ein Stürmer läuft allein aufs Tor und ein Verteidiger verfolgt ihn. D.h., da gibt's noch gar nichts zu tun. Aber dann rufe ich schon in meinem Kopf schon ›Rot, Rot, Rot‹ [Anm. M.W.: Szenario ist dabei ein Foul des Abwehrspielers das als »Notbremse« mit Rot zu bestrafen wäre]. Und wenn der Verteidiger dann foults, dann muss ich gar keine Entscheidung mehr treffen.

Neben der körperlichen Darstellung dieser Kodierung als fundiert in regelkundiger und konzentrierter Wahrnehmung, sind auch materielle Strukturen beteiligt an der Hervorbringung der Schiedsrichterpiffe als technisierter Wahrnehmung von Tatsachen. Im Fußball handelt es sich nämlich bei der Kodierungspraktik vor allem um die Zuschreibung kategorialer Spielereignisse zu Positionen des Balls im Verhältnis zur weißen Liniatur auf dem Rasen (»Aus«, »Tor«), Positionen und Bewegungen von Körpern, unterscheidbar nach Trikotfarben und Rückennummern, im Verhältnis zu anderen Körpern (»Foul«) oder zur weißen Liniatur (»Abseits«, »Freistoß«, »Elfmeter«). Man erkennt daran, dass meistens materielle Infrastrukturen daran beteiligt sind, Spielereignisse in einer Weise zu erzeugen, die das Manifest-Werden von

⁷ Vermutlich hängt dieser Wettbewerbsfigur eine Benennung an, die für eine Vorform der heutigen Schiedsrichterrolle noch sehr passend war. Die ersten »Referees« (person to refer to) im Fußball hatten keine Interventionskompetenz, sie wurden lediglich von Umpires (für Regelfragen zuständige Vereinsmitglieder) angerufen, wenn diese sich in einem Streitpunkt nicht einigen konnten.

Regelungslücken unwahrscheinlicher machen können. Das Regelwerk legt deshalb nicht nur Vorgaben zur stofflichen Ausstattung der Spielerkörper fest (Mannschaftsuniformen, Spielernummern usw.), sondern definiert eine ganze baulich-materielle Platzordnung, deren ordnungsgemäßen Zustand Schiedsrichter vor Spielbeginn zu kontrollieren haben (sogenannte »Platzbegehung«). Diese materiellen Strukturen sollen es ermöglichen, dass die für den bewertenden Ausgang des Wettkampfes entscheidenden kategorialen Zustände des Spiels (»Tor für Rot. Eins zu Null.«) möglichst gemeinhin erkennbar werden. Durch sein »Zappeln« im Netz machen Ball, Tor und Netz es mehr oder weniger unverkennbar für alle Beteiligten, dass der Stürmer »getroffen« hat.⁸ Der Schiedsrichter muss das dann nur mit aufmerksamer Wahrnehmung registrieren, sowie durch einen Pfiff formal ratifizieren, was ohnehin schon wettkampföffentlich als Tatsache etabliert ist.

Die materiellen Infrastrukturen der Platzordnung erzeugen also die Darstellung des Pfeifens als bürokratischer Registriertätigkeit mit. Es geht dabei darum, diese Kategorisierungspraktiken, man könnte auch spezifischer von Kodierung sprechen, trotz der widersprüchlichen formalen Erwartung an Endgültigkeit des Pfiffs (»Tatsachenentscheidung«) einerseits und informeller Erwartung an reine Wahrnehmungsbasiertheit (zeitliches Erzwingen von Gedanken- und Kommunikationslosigkeit) andererseits zu ermöglichen.⁹ Das Regelwerk kodifiziert aber nicht nur Vorgaben für die materielle Infrastruktur, sondern ist auch kognitive Ressource zum Darstellen von Entscheidungen als regelbasiert. Diese kognitive Ressource steht dabei jedoch nicht nur Schiedsrichtern, sondern auch Spielern zur Verfügung. Schiedsrichterinnen haben dabei aber sozusagen einen anderen, versierteren und professionelleren Zugang zum Regelwerk als andere Spielteilnehmer. Man kann das Bemühen mit Hilfe von kognitiven Kategorien Schiedsrichterwahrnehmungen zu technisieren gut erkennen, wenn man sich anschaut, wie das Schiedsrichterwesen in bestimmten Fällen mit der Formulierung von Regelauslegshilfen auf

- 8 Aber eben auch nur »mehr oder weniger unverkennbar«. Beispiele für Uneindeutigkeiten finden sich selbst im Profifußball. Am 18.10.2013 z.B. köpfte der Leverkusener Spieler Kießling, im Rahmen eines Bundesligaspiels, eine Flanke von Außen durchs Tornetz. Was auf den Kamerabildern gut erkennbar war, entging dem Schiedsrichter und den Spielern (außer Kiesling selbst, der schwieg jedoch ganz professionell unfair), das Tor zählte (vgl. sportschau.de 2023). Kolleg*innen älteren Semesters erinnern sich vielleicht eher an das »Phantomtor« von Thomas Helmer aus den 90igern.
- 9 Widersprüchlich ist diese Erwartung, weil Wahrnehmung in den Ethnotheorien des Feldes als »immer auch subjektiv« – so ein Lehrwort auf einem Neulingslehrgang – gilt.

wiederkehrend auftretende Enttäuschungen der Idee des Regel-Anwendens als Technik reagiert. Regelauslegungshilfen geben Schiedsrichtern Zusatzkriterien zur konsistenten Auslegung einer Regel an die Hand. Es handelt sich insofern um Meta-Regeln, weil sie den Umgang mit den Regeln regeln. Oft bilden diese Metaregeln auch den Ausgangspunkt für die Spezifizierung schon vorhandener Regeln in Regelreformen. Die Handspielregel im Fußball ist ein gutes Beispiel, an dem sich die Bemühung zeigt, das technische Versagen der Regel durch technische Nachbesserung auf organisationaler Ebene zu reparieren. Welche Probleme sind hier aufgetaucht und wie wurde mit ihnen umgegangen? Schauen wir zunächst einmal in die von 1902 bis 2019 gültige Formulierung der Handregel in den »Laws of the game« (gelten weltweit):

Ein Handspiel liegt vor, wenn ein Spieler den Ball absichtlich mit der Hand oder dem Arm berührt. Folgendes ist zu berücksichtigen: die Bewegung der Hand zum Ball (nicht des Balls zur Hand), die Entfernung zwischen Gegner und Ball (unerwarteter Ball), die Position der Hand (das Berühren des Balls an sich ist noch kein Vergehen).

Konstruiert wird die Differenz zwischen verbotenen und gestatteten Handspielen hier also über die Kategorie der Absicht. Der restliche Regeltext umfasst lediglich drei Spiegelstrichabsätze, jeweils mit einem Halbsatz. Manche davon lesen sich wie hermeneutische Hilfsmittel. Die offensichtlich dichte (im Sinne Ryles) und nur mit komplexer Interpretationsleistung anwendbare Kategorie der Absicht wird zum Beispiel mit einem Anhaltspunkt ergänzt, der sich aber seinerseits ebenso dicht liest (»unerwarteter Ball«). Dicht sind diese Kategorien insofern, als dass sie dem Anwender unterstellen, dass er in sinnvoller Weise Annahmen über die Intentionalität des Spielers treffen kann. Eine andere Formulierung liest sich dünner. Der Anhaltspunkt der »Bewegung der Hand zum Ball« abstrahiert mechanistisch von der Intentionalität des Spielers. Beobachtet man die Frage, wie diese Regel im Spiel praktisch enaktiert wird und welche sprachlichen Kategorien dabei verwendet werden, geht es meistens um eine Unterscheidung zwischen dem »Angeschossen werden« und dem gezielten »Abwehren« eines Schusses. Insbesondere weil die typischen Handspielgrenzfälle oft besonders relevante Spielereigniskategorisierungsfragen aufwerfen (oft geht es um die Frage eines Elfmeterpfiffs), eignet sich die Auslegung dieser Regel weiterhin besonders als Gelegenheit für Spielteilnehmer, die Gegenstrategie zur Technisierung zu entfalten: das Dramatisieren von Schiedsrichterpuffen als Fehlentscheidungen.

Dieser Regeltext scheint aber trotzdem bis in die 80er oder 90er Jahre hinein ein hinreichend unproblematisch nutzbar gewesen zu sein. Wie in anderen Fällen auch gibt es ergänzende Meta- oder Zusatzregeln, die hauptsächlich Insidern des Schiedsrichterwesens geläufig sind. Letztlich ergeben sich aber wieder kontinuierliche Diskussionen um die Frage der Auslegung dieser Regel, noch gibt es Anstrengungen in Richtung einer Regelreform auf der Ebene der Fußballverbände. Seit den 90er-Jahren, in den Nullerjahren dann aber noch einmal mit mehr Fahrt und bis zum jetzigen Zeitpunkt persistierend, ist das »Hand-Problem« ein kontinuierliches Thema in der Fußballberichterstattung, sodass auch öffentliche Forderungen von Spielern, Funktionären und Journalisten nach einer Reform regelmäßig auftauchen (vgl. zum Beispiel sport1.de 2019).

Eine strukturelle Voraussetzung für die Konjunktur dieser Problematisierung der Handspiel-Regelungslücke scheint darin zu bestehen, dass die Dramatisierung von Fehlentscheidungen nicht nur auf dem Platz und im Spiel, sondern (besonders intensiv seit den 90igern) auch in und mit den Mitteln der TV-Übertragung betrieben wird und von dort aus auf den Umgang mit einer Regel-Kategorie wie der »Absicht« auf neue Art und letztlich mit erhöhter visueller Tiefenschärfe problematisiert werden kann.¹⁰ Diese Debatten sind dem Fußballpublikum mittlerweile so geläufig, dass auch das typische exoterische Regelwissen des Publikums mittlerweile um Aneignungsformen der zusätzlichen Hilfskonstrukte aus den verbandseigenen Metaregeln weiß (und in der eigenen Analyse von strittigen Szenen mit ihnen agiert). Die Verbände haben –

¹⁰ Colwell (2001) hat die massenmediale Konstruktion von Regelverstößen und Fehlentscheidungen seit dem Ende des Zweiten Weltkriegs untersucht. Demnach ist das Thema der »Schiedsrichter Fehlentscheidung« erst in den 80iger Jahren und damit relativ spät in den Medien aufgetaucht. Ging es der klassischen TV-Kameraführung noch bloß darum, den Ball ins Bild zu kriegen, kommt nun die Entscheidung-Prüfen als Zeige-Stil des TVs und seiner Sprecher hinzu. Eine ganze Armada technischer Gerätschaften wird nun bemüht, um Pifffe als Fehlentscheidungen zu dramatisieren: »The coverage of live Premier League football now includes a range of ›modern, scientific instruments‹ which provide viewers with a mass of additional information. Viewers and commentators have access to instant statistics about the number of corners, shots on target, fouls conceded or percentages of possession time between teams. On-screen graphics indicate the ›offside line‹, how far from goal a shot or free kick is and how far back the defensive wall should be. Commentators can now access instant, slow motion replays of incidents, taken from any of 20 different camera angles and, as a result, referees and players have their performances scrutinised in more detail than ever before.« (ebd.: 296f.)

mit den Mitteln ihrer Öffentlichkeitsarbeit (zum Beispiel Workshops für Trainer oder für Journalisten) – Zusatzkriterien wie die Kategorie der »Armhaltung« ins Spiel gebracht, wobei diese erst instruktiv gemacht wird durch die Unterscheidung »angelehnter Arm« oder »Vergrößerung der Körperoberfläche durch den Arm«.

Im Jahre 2019 sind diese Zusatzregeln teilweise in eine weltweit gültige Neuformulierung der Handregel aufgenommen worden. Damit wuchs der Regeltext von seinen vorherigen vier Zeilen auf mehr als eine halbe Seite an. Dabei fällt auf, dass das hinsichtlich der klassischen Regelformulierung erwähnte Spannungsverhältnis aus einem interpretativ-offenen Begriff der »Absicht« und dünnen Objektivierungskriterien, nun noch einmal stärker in die Richtung einer Relativierung der situationssensiblen Unterstellbarkeit von Absicht hinzu einer formal durchdefinierten Operationalisierung von »Absicht« verschoben wurde. Ein Absatz aus der Regel, der sich dafür beispielhaft aufrufen lässt, lautet etwa:

den Ball mit der Hand/dem Arm berührt und seinen Körper aufgrund der Hand-/Armhaltung unnatürlich vergrößert (FIFA 2021: 72)

Eine Tendenz dieser Entwicklungen scheint darin zu bestehen, die Kategorie der »Absichtlichkeit« umzudefinieren, von einer explizit interpretativ-offenen Kategorie, welche Schiedsrichterinnen als urteilsfähig adressiert, zu einer technischen Kategorie, die in »dünne«, gewissermaßen behaviouristische Begriffe aufgelöst werden kann. Imaginiert wird die Schiedsrichterin in solchen Metaregeln und Regelspezifizierungen insofern als eine Art Kodier-Personal. Die Konjunktur der Problematisierung des Versagens der Technisierung wird insofern gewissermaßen interpretiert als Ausdruck einer mangelnden Inter-coder-Reliabilität. Reagiert wird darauf mit der Spezifikation der Instruktionen sowie Anstrengungen der Vermittlung dieser Instruktionen (Regelbriefe in der Schiedsrichterzeitung, wiederholte Aufnahme des Themas in die Tagessordnung von sogenannte »Belehrungsabenden« auf denen das Regelwissen der Schiedsrichter aufgefrischt und erweitert werden soll). Es ist interessant, das Scheitern dieser Regelreform – gescheitert ist sie an der Absicht, die fußballöffentliche Unzufriedenheit mit solchen Fällen zu verringern – genauer nachzuverfolgen. Das würde an dieser Stelle aber zu weit führen. Interessanter ist es an dieser Stelle, das Gegenstück zur Strategie der Technisierung in den Blick zu nehmen: das Mitspielen.

4. Mitspielen: Oder wie das Sichtbar-Machen der Regel-Auslegung den Schiedsrichter ins Spiel verwickelt

Die Kultur des Fußballs selbst weiß um Unentscheidbarkeit von Kategorisierungsfragen im Spiel. Schon der Gebrauch der Begriffe »Wahrnehmung« und der »Entscheidung« in Zusammenhang mit dem Schiedsrichter – mit ihrer Nähe zur »Subjektivität« – in Regelwerk und Feldjargon legen nahe, dass eine technische Bearbeitung der Unterdeterminiertheit der Regeln nur als begrenzt möglich angesehen wird. Der Feldjargon kennt darüber hinaus Ausdrücke für Unentscheidbarkeiten, mit denen Schiedsrichter konfrontiert sind. Beispielsweise sogenannte »Fifty-Fifty-Entscheidungen« meinen Situationen, in denen man zwei Kodierungsoptionen hat (etwa Foul/Nicht-Foul), aber aus dem isoliert betrachteten Ereignis heraus keine Anhaltspunkte für eine korrekte Auslegung vorhanden sind. Die Folklore des Feldes tradiert daran anknüpfend mit einer gewissen Faszination Geschichten über die Unentscheidbarkeiten des Spiels in ihren Legenden (zum Beispiel das sogenannte »Wembleytor«). Die soziologische Forschung zum Schiedsrichter bemüht sich deshalb, einen stark dezisionistischen Typ von Schiedsrichterpiffen zu spezifizieren. Rains etwa spricht von »judgement calls« (Rains 1984: S. 153–154), Smith von einem »overriding need for some mechanism imposing ›official closure‹« (Smith 1982: 43) der Vorrang hat vor der Korrektheit der Entscheidung, und Collins in Anschluss an Winch von »judgments« im Unterschied zum bloßen »rule-based decision-making« (Collins 2020). Im Folgenden wende ich mich der Frage zu, welche Repertoires der Schiedsrichterrolle es prinzipiell möglich machen, die Kategorisierung einer Spielsituation im Einzelnen, aber mehr noch die Spielleitung insgesamt, auch noch in Anbetracht manifester Kontingenzen und Regelungslücken, wie in der »Fifty-Fifty-Situation«, mehr oder weniger überzeugend als »fair« darzustellen. Wenn die Bearbeitung der Unterdeterminiertheit der Regeln durch Technisierung der Wahrnehmung des Schiedsrichters darauf zielt, Schiedsrichter-Pfiffe als intersituativ und bürokratisch gesteuert gleichförmig kodierende Wahrnehmung von Tatsachen erscheinen zu lassen, dann zielt die Strategie des Mitspielens darauf ab, die Schiedsrichterin in das situative Wettkampfgeschehen zu verwickeln und von dort aus als urteilsfähig darzustellen. Wie genau wird die Schiedsrichterin also ins Spielgeschehen selbst verwickelt und so auf situative Relevanzen jenseits des Regeltextes hin orientiert?

Die erste Ebene der Antwort auf diese Frage ist einfach: Die Schiedsrichterin darf als einzige Nicht-Spielerin auf dem Feld mitlaufen und kann sich so

immer in der Nähe des aktuellen Spielgeschehens aufzuhalten. Der Insider-Diskurs des Schiedsrichters verwendet große Aufmerksamkeit darauf, das Nutzbar machen dieser Bewegungsfreiheit zu strategisieren. Laufwege werden zum Beispiel in der organisationalen Evaluation von Schiedsrichtereinsätzen genau beobachtet und nach spezifischen Kriterien bewertet. Man bemüht sich um »ständige Seiteneinsicht«, aber muss gleichzeitig aufpassen, nicht zu nah am Geschehen zu sein, sodass es zum »Tänzeln« kommt (spontanen Ausweichbewegungen) oder Situationen in denen der Schiedsrichter dem Ballgeschehen »im Weg steht«. Dabei geht es darum, im Sinne eines sportlich-informierten und im Hinblick auf die Überwachungsfunktion strategischen »Positionsspiels« fortlaufend eine proxemische Position der Zeugenschaft am Ort mit der höchsten Wahrscheinlichkeit für Regelungsbedarf einzunehmen. In den Wörtern eines Schiedsrichters:

Wenn man ein Foulspiel pfeift, schaut sich der Spieler zumeist um, um zu sehen, wo der Schiedsrichter steht. Steht er direkt hinter oder neben ihm, bleibt der Protest aus. (Schröder 2015: 78)

Die schiedsrichterliche Laufpraktik lässt sich insofern als motorische Form eines »visuellen Engagements« (Heimerl 2013: 103) verstehen. Welche Darstellungsprobleme verlangt aber der so durch gute Laufarbeit vorbereitete Moment des Pfiffs, etwa eines Fouls, vom Schiedsrichter dem Schiedsrichter ab? Hier sind unter anderem bestimmte emotionale Engagementgebote bzw. Feeling-Rules (Hochschild 1983) interessant. So sind Schiedsrichter aufgefordert, schon beim Blasen in die Pfeife einen der Konfliktintensität angemessen unaufdringlichen oder scharfen Ton zu erzeugen. So heißt es beispielsweise in einer Ausgabe der DFB-eigenen »Schiedsrichterzeitung«, dass es im Fall eines Fouls gilt, einen »beherzten« nicht aber einen »schüchternen« Pfiff erklingen zu lassen (DFB 2021: S. 7). Der derart »differenzierte Pfiff«, wie ihn der Diskurs der Lehrwärte nennt, kommuniziert insofern schon eine Einschätzung der Konfliktlage und damit zusammenhängend eine Selbstdarstellung des Schiedsrichters als mehr oder weniger stark in die emotional-kommunikative, nicht bloß die formal-regeltechnische Ebene des Geschehens intervierend. Man könnte diesen Punkt noch an einer Reihe weiterer Verhaltensrepertoires rund um Stichworte wie die »Ansprache« eines zu bestrafenden Spielers, den Umgang mit »hitziger Atmosphäre« auf dem Platz usw. darstellen. Mir geht es darum, dass alle diese Repertoires den Pfiff nicht primär als technische Kodierung, sondern als konfliktsensibles Darstellen des Gebrauchs

formaler Regeln darstellen. Noch stärker tritt das hervor, wenn man sich das Gebot zur Darstellung von Entschiedenheit anschaut. Pfiffe müssen nicht nur durch zeitliche Nähe zum Geschehen und passende tonale Variation »Präsenz erzeugen« (DFB 2021: S. 7), sondern auch in verbalen Anfechtungen (Gemecker, Protest usw.) durch Spieler selbstsicher dargestellt werden. Im Grenzfall gilt es sogar, Sicherheit vorzutäuschen, wo sie faktisch nicht gegeben ist:

Man muss Entscheidungen verkaufen. Manchmal ist es wichtiger, so zu tun, als ob man sich sicher ist. Klar, kann man sagen. Also das auch wirklich vor täuschen!

An diesem Grenzfall wird das eigentliche Grundproblem erkennbar: Das Erkennen von Wahrheiten ist nachrangig gegenüber dem schnellen und verlässlichen Schaffen von Tatsachen, die das Weiterspielen ermöglichen. Und man tut dies manchmal, indem man es so aussehen lässt, als tue man konsequent nichts Anderes als Wahrheiten zu erkennen (wo diese gar nicht erkennbar sind). Die Performativität des Pfiffs ist vorrangig gegenüber seiner empirischen Korrespondenz mit der Realität. Aber sie lässt sich nicht stabilisieren, wenn wiederholt der Eindruck entsteht, dass dem so ist. Das Vortäuschen von Überzeugtheit wird daher auch als eine Strategie für seltene Grenzfälle schwer entscheidbarer Situationen und nicht als generell brauchbare Strategie beschrieben. In anderen Fällen gilt es aber auch, sich weniger technisiert wahrnehmend und dafür kommunikativ verfügbarer zu zeigen. Eine Schiedsrichterin, die sich Gesprächen mit Spielern vollständig entzieht und die Möglichkeit einer Relativierung der eigenen Wahrnehmung kategorisch ausschließt, gilt als »arrogant« und »letztlich unsicher«. Ein Schiedsrichter im Interview:

Dann sagst du gerade bei so einer Fifty-Fifty-Sache halt durchaus auch mal: ›Gerade hab ich das für euch gepifffen, jetzt pfeif ich es halt einmal gegen euch.‹ Also dass es sich ausgleicht.

Das Mitspielen relativiert hier das Technisieren mit seiner Dramaturgie der Kodierung auf Grundlage bloßer Wahrnehmung, aber es bricht dessen Darstellungsregeln nicht auf. Die kommunikative Offenheit der Schiedsrichterin führt nicht zu einer Revision der soeben getroffenen »Tatsachenentscheidung«, sondern zu prospektiv eingeräumten Konzessionen für spätere Entscheidungen (bestimmten Typs, nicht sämtlicher!). Noch ein weiterer Punkt

zeigt sich an dieser Aussage: Mitspielen heißt auch zu wissen, wann man das Pfeifen stur als technisch-bürokratisch registrierendes Wahrnehmen inszeniert und wann man dem Frust sich benachteiligt fühlender Spieler auf verbaler Ebene Konzessionen zugestehen kann. Dabei wird das Verhältnis von Technisierung und Mitspielen durch das Schiedsrichterwesen auf unterschiedliche Weise beobachtet. Die Unterscheidung mit Begriffen wie »Kann-« und »Muss-Entscheidungen« hinsichtlich einzelner Vergehenstypen, die sich mehr oder weniger direkt aus dem formalen Regeltext ableiten lassen, liefert relativ klare Anhaltspunkte.¹¹ Der Begriff der »Fifty-Fifty-Situation« geht über die Unterscheidung von »Kann-« und »Muss-Entscheidungen«, im Einräumen von dezisionalen Spielräumen, noch hinaus. Denn ein bestimmter Typ von Spielereignis wird gewissermaßen als zur freien taktischen Nutzung qualifiziert. Für solche Fragen von »Spieltaktik« ist der Begriff »Linie« zentral. Gemeint ist eine sich interaktionsgeschichtlich in jedem Spiel aufs Neue aufbauende Erwartungshaltung: Wenn die Schiedsrichterin von Beginn an auch Vergehen, die den »Kann-Pfiffen« zuzuordnen sind, »streng wegpfifft«, wird von ihr erwartet, dies auch weiterhin zu tun. Besagte Ausgeglichenheit oder das Motto »im Zweifel nicht Pfeifen« (leitet sich ab aus dem Unsichtbarkeitsgebot) können ebenso als Auslegungshilfen herangezogen werden.

Ein anderes, derzeit sehr verbreitetes, Stil-Gebot für das Entwickeln einer Linie suggeriert eine Harmonie zwischen dem Unsichtbarkeitsgebot der Technisierung und dem Gebot zum gezielten Einsatz von Präsenz im Mitspielen. Handreichungen zur Bewertung von Schiedsrichtern – wie sie für die Schiedsrichterbeobachter erstellt werden –, aber auch Porträts in esoterischen (Schiedsrichterzeitung) genauso wie in exoterischen Medien der Fußballöffentlichkeit (Sportpresse) rufen immer wieder das Attribut der »Persönlichkeit« auf. Erfahrene und durchsetzungsstarke Schiedsrichterinnen sind in der Lage, »mit viel Persönlichkeit« zu agieren, insofern sie den Einsatz formaler Sanktionsmittel auf ein Minimum reduzieren. Prävention gelinge diesen Schiedsrichtern durch »aktive Kommunikation«, wie sie sich beispielsweise in der »individuellen Ansprache«¹² eines Spielers in Anschluss

¹¹ Analog dazu reflektiert die Sportjuristerei und -ethik die Differenz als Differenz von Regeln und Standards (Zglinski 2022).

¹² Die Bezeichnung habe ich einem Schiedsrichterbeobachtungsbogen entnommen. Solche Bögen sind sozusagen Zeugnisse für Schiedsrichtereinsätze. Dankenswerterweise haben mir verschiedene Schiedsrichterbeobachter einige anonymisierte Exemplare zugänglich gemacht.

an ein Vergehen zeigt. Der Begriff ist besonders vielsagend im Hinblick auf das Spannungsverhältnis von Sichtbarkeit und Unsichtbarkeit. Er konzidiert der Schiedsrichterin Sichtbarkeit, und das gar als »Person«, die den Spieler als »Person« adressiert und ihm gewissermaßen mit moralischer Zusatzebene des Strafens droht auf der Ebene von Achtungskommunikation, gleichzeitig macht sich die Schiedsrichterin im formalen Sinn unsichtbar, indem seine informelle Präsenz betont wird.

Eine andere Ebene, auf der sich das Partizipieren des Schiedsrichters am Wettbewerbsbetrieb als Mitspielen beschreiben lässt, findet sich ebenfalls auf der Ebene der Evaluation von Schiedsrichtern in für das Schiedsrichterwesen typischen Begriffen. Eine Szene, die ich im Rahmen einer Gespannbegleitung zur Halbzeit in der Kabine beobachten konnte, kann dazu dienen, dies zu erläutern:

»So eine dumme Szene!«, ärgert sich Thomas lautstark. »Ich war einfach zu geil auf diese gelbe Karte. Wenn ich den Vorteil laufen lasse: Tor. So kommt dieser dämliche Konter und es steht nicht 3:1 sondern 2:2! Spieltaktisch war das komplett Banane! Dann kann ich auch verstehen, wenn die komplett bedient sind.« Ich frage nach, ob das eine Szene wäre, die ihm bei Beobachtung die Note versaut. Ja, ein Abzug bei Vorteilen statt einer möglichen Aufwertung. »Und das ärgert mich besonders. Weil Vorteil und Spielmanagement sind eigentlich meine Stärken. Aber ich war zu geil auf diese Gelbe. Der war lange fällig und es war auch Zeit endlich die erste Gelbe zu zeigen.Und dann auch noch in der Nachspielzeit!

Aus Schiedsrichterperspektive ist diese Szene ärgerlich, weil sich der Schiedsrichter hier sichtbar einen großen Einfluss auf den Spielverlauf zuschreibt (»nicht 3:1 sondern 2:2«). Der Zusatz »und dann auch noch in der Nachspielzeit« unterstreicht diesen Einfluss noch einmal: Nachspielzeit gibt es nur nach und durch das Ermessen des Schiedsrichters. Hätte er früher abgepfiffen, wäre es gar nicht erst zu der Szene gekommen. Wenn er weiterhin sagt: »Dann kann ich auch verstehen, wenn die komplett bedient sind«, spielt er auf die lautstarken Proteste im Zuge dieser Szene und die spürbare Unzufriedenheit der Mannschaft beim Gang in die Kabine an. Besonders interessant sind außerdem die Stichworte »spieltaktisch« und »Spielmanagement«. Als spieltaktisch werden Spielereignisse von bloß formalen Spielkategorisierungslässen im Sinne des formalen Regelwortlauts zu situativen Konfliktinterventionen umgedeutet, die zu einem Spielerlauf passen. Typischer ist es dabei, zum

Beispiel »Problemspieler« zu identifizieren, die es mit einem »Anschiss« oder einer »Karte« »einzufangen« gilt. Das Gelingen oder Misslingen solcher Interventionen wird dann in der retrospektiven Reflexion – in der Kabine mit den Kollegen aus dem Gespann oder im Coachinggespräch mit einem Schiedsrichterbeobachter – beschrieben, indem sie als »Schlüsselszenen« für den »Spielverlauf« figuriert werden.

Das Hantieren mit dem formalen Regelwerk wird hier also nicht mehr nur auf Fragen der sachgemäßen Kodierung von Spielereignissen bezogen, sondern primär auf das (Mit-)Gestalten einer situativen Konfliktdynamik. Wenn sich das sture technische Kodieren nach formal-bürokratischen Vorgaben in der Spielleitung durch ein situatives Mitspielen und Einlassen auf die situative Konfliktdynamik des Spiels relativiert, dann wird damit auch mehr oder weniger unverblümmt Spielteilnehmerinnen ohne formale Entscheidungssouveränität (Spielerinnen, Trainerinnen, Publikum) eine informelle Mitgestaltungslizenz am Bearbeiten der Unterdeterminiertheit der Regeln eingeräumt, sodass diese auch mehr oder weniger diffus auf die Regelauslegung einwirken.

Dieser Aspekt des Hantierens mit den Regeln wurde und wird in der Wettkampfkultur des Fußballs stark akzentuiert, insofern informelle Normen – »Feeling Rules« (Hochschild 1983) des Spielens – den Wettkampfteilnehmern (und Zuschauern) große Spielräume im Entfalten einer emotional aufgeladenen und moralisierenden Bewertungskommunikation rund um Probleme des laufenden Bearbeitens der Regelunterdeterminiertheit ermöglichen. Fußballschiedsrichter beklagen deshalb in Interviews den Fußball als eine Sportart, in der ein gewisser Grad an Gemecker und – Fairplay-Idealen entgegenlauender – Theatralik mit dazu gehört. Soziologisch ist das zunächst einmal wertfreier zu betrachten. Justus Heck hat diesen Verhaltensstil in diesem Sinne als »professionelle Unfairness« (Heck 2019: 48) bezeichnet und damit darauf hingewiesen, dass eine gewisse Form von Unfairness im körperlichen Spiel und Opportunismus in der moralischen Kommunikation (Bergmann/Luckmann 1999) um die Regelauslegungen – bis hin zu den bekannten Täuschungsmanövern wie der »Schwalbe« – im Fußball nicht als Abweichung von Normen, sondern als kompetente Spielweise betrachtet wird. Dazu passend zeigt Cunningham in einer Interviewstudie mit Spielern, in der er Fragen der Spieler-Schiedsrichter-Kommunikation nachgeht, dass Spieler mit der Erwartung operieren, Täuschungsverhalten gegenüber den Schiedsrichtern könne ihnen Wettkampfvorteile verschaffen (Cunningham et.al. 2015). Der Sportjournalist Eichler (2000: 180ff.) weist darauf hin, dass einige Feldbegriffe einerseits das Foulen positiv valorisieren (etwa das »kluge Foul«), andererseits

den Begriff der Moral von Fragen der Fairness entkoppeln. So »beweist« eine Mannschaft oder ein Spieler im Jargon des Fußballfeldes dann Moral, wenn er gerade im Moment der wahrscheinlichen Niederlage hartnäckig weiterkämpft. »Moral« ist hier also begrifflich an einen schwer zu brechenden Willen zu leisten/zu siegen gekoppelt und nicht an Fragen der intersubjektiven Verständigungsorientierung und Rücksichtnahme. In diesem Sinne, so schließt er korrekt, erwartet das Publikum im Fußball eine professionelle Unfairness von Spielern.

Gunter Gebauer hat die wettbewerbstheoretischen Pointen dieses Phänomens auf den Punkt gebracht, indem er es aus einer kulturanalytischen Warte beschreibt. Wenn der Schiedsrichter keine Tore schießen oder kassieren kann und er weder gewinnen noch verlieren kann, woran genau spielt er dann eigentlich mit? Mit Gebauer (2006) lässt sich sagen, dass der mitspielende Schiedsrichter dem Wettkampf eine zusätzliche Ebene der Herstellung von Leistungsdifferenzen ermöglicht: »Was macht das Besser-Sein der siegreichen Partei aus? [...] Seitdem die Durchsetzung der Regeldeutung zu einem Bestandteil des Spiels selbst geworden ist, kommt es nicht mehr nur auf die Qualität des Spiels an, sondern auch auf die Listen, mit denen Regeln umgebogen, unterlaufen, umdefiniert, zu eigenen Gunsten eingesetzt werden.« (147) Er spricht von einem »Spiel im Spiel«. Auf dieser Ebene sind formale Regelwerke des Wettkampfs nicht nur Teil des organisationalen Versuchs gleiche Wettbewerbsbedingungen herzustellen, sondern auch Spielressourcen in der Wettkampfpraxis selbst.

5. Wettkampfpraxis und Wettkampftonus

Die beiden beschriebenen Strategien führen zu einer starken Spannung in der Rolle des Schiedsrichters. Die Technisierung lässt den Schiedsrichter als eine qualitative Kodierungsapparatur erscheinen, die durch die Kontextuierung in intersituativ-bürokratischen Verweisungshorizonten für eine situationsübergreifende Stabilität (Intercoder-Reliabilität) von Regeldeutungen Gewähr stehen soll. Das Mitspielen mobilisiert den Schiedsrichter dagegen als Konfliktmoderator, der situationssensible Urteile fällt. Der bewertungssoziologische Blick auf die Wettbewerbskonstellation des Verbandsfußballspiels hat in diesem Sinne sichtbar gemacht, dass und wie Wettbewerbsformen nicht nur auf der Ebene der Institutionalisierung von rahmenden Strukturen (Organisationen, Regelwerke usw.), sondern auch auf der Ebene ihrer situativen Vollzugs-

wirklichkeit eine gewisse Formenvielfalt aufweisen können. Der streitlustige (Deland 2013, Heck 2019) Tonus (Elias/Dunning 1982: 338–339) des Fußballspiels evoziert Bewertungspraktiken, die Regeln von einer bloßen Formbedingung des Wettkampfs zu einem Teil seines Inhalts machen. Es sind erstens Bewertungspraktiken durch Spieler, Trainer und Zuschauer (Dramatisierung von Pfiffen als Fehlentscheidungen), die das Auslegen der Regeln als kontingentes und situatives Fällen von Urteilen durch ihren Anwender sichtbar machen, und es sind die komplexen Bewertungsstrategien im Mitspielen des Schiedsrichters, die auf diese Problemlage antworten.

Die Wettkampfpraxis des Fußballs akzentuiert insofern ihre Situativität in besonderem Maße. Damit relativiert sie selbst auch die Idee, wonach der Sieg eine situations-unabhängige Leistungsfähigkeit zum Ausdruck bringt. Das liegt schon in seiner Form als »low numbers game« begründet, die das Glück zum mitunter entscheidenden Mitspieler macht. Deutlicher als in anderen Sportarten stellt der Fußball die Bedeutung von Zufällen auch anhand der Markierung der Willkürlichkeiten von Schiedsrichterentscheidungen aus. Das Fußballdrama produziert nicht nur sauber-meritokratische Sporthelden (Bette 2007), sondern auch ambivalente Helden (z.B. den beim Publikum oft beliebten »Treter«) bzw. Antihelden (wenn es Gegner sind), die sich mit kleinen Beträgereien einen Vorteil verschaffen. Der Fußball erzeugt auch eine Form von Spannung, die vom »Argwohn« (Bromberger 2003: 289) gegen den Gegner herröhrt und vom Potential, sich über die Bürokratie der Schiedsrichterei zu empören. Betrachtet man den Fußball kulturanalytisch als eine Produktion von Dramen, ist der Schiedsrichter deshalb eine »schwarze Figur« (Bromberger 2003: 289). Das ist aber weniger ein Problem für diese Form des kulturellen Spektakels, als eine attraktive dramaturgische Ressource. Mit den Worten eines Schiedsrichters (Schröder 2015: 185):

Fehlentscheidungen sind, man kann das auch positiv sehen, der Zündstoff, der die Flamme des Mythos Fußball immer wieder anfacht.

Einerseits zeigt sich also ein Potential für die Sportsoziologie, die kulturoziologische Bedeutung ihres Gegenstandes besser zu verstehen, wenn sie Fälle wie den Schiedsrichter oder andere Figuren des neutralen Dritten im Sport in den Blick nimmt. Insofern sich hier im Hinblick auf die Vielfalt der Sportarten sehr unterschiedliche Formen dieser Figuren identifizieren lassen, deutet sich hier aber auch ein Potential an, sehr verschiedene Arten des Hantierens mit formalen Regelwerken in Wettkampfpraktiken und mithin verschiedene Ton-

arten des Wettbewerbs zu identifizieren. Ein weiterer Austausch zwischen der Soziologie des Wettbewerbs und einer mikrosoziologisch orientierten Soziologie des Dritten im Sport könnte also dazu beitragen, für die Formenvielfalt, innere Widersprüchlichkeiten und Auswirkungen komplexer situativer Vollzugslogiken des Hantierens mit Regelwerken auf Wettbewerbskonstellationen aufmerksam zu machen. In diesem Sinne täte die Soziologie des Sports sicher gut daran, Figuren des Dritten im Sport – von der Jury, die ästhetisches Urteilsvermögen beansprucht, bis hin zum »Renndirektor« der Formel 1, der vor allem technisches Know-how vorweisen können muss – systematisch und vergleichend zu untersuchen. Entscheidend ist dabei, solche Figuren nicht vom Standpunkt des Phantasmas »reiner Leistung« aus als mehr oder weniger optimal objektivierende Formen von »Regelhütern« zu bewerten, sondern sie jeweils in ihrer Eigenlogik ernst zu nehmen. Die Unterdeterminiertheit der Regeln ist in diesem Sinne kein Problem, sondern einfach eine unhintergehbar Grundbedingung von Praxis, die auf die eine oder andere Art problematisiert, bearbeitet oder eben auch – wie in der Kultur des Verbandsfußballs – genutzt werden kann.

Andererseits kann die Soziologie der Bewertung mit ihrem Blick auf situierte Vollzugslogiken von Bewertungspraktiken den Gegenstandsbereich der Soziologie des Wettbewerbs um eine wichtige Facette erweitern. Es gilt nicht nur die Institutionalisierung bestimmter Formen von Wettbewerben zu untersuchen, sondern auch deren praktische Hervorbringung und Prägung durch situative Eigenlogiken der Wettkampfpraxis.

Literatur

- Arora-Jonsson, S., Brunsson, N., & Hasse, R. (2021): A new understanding of competition. In Dies (Hg.) *Competition. What It is and Why It Happens* (1–25). Oxford University Press.
- Bergmann, J. & Luckmann, T. (Hg.) (1999): Die Kommunikative Konstruktion von Moral. Opladen: Westdeutscher Verlag.
- Bette, K. H. (2007): Sporthelden. Zur Soziologie sozialer Prominenz/Sports Heroes. *Studies in the Sociology of Social Prominence*. In: Sport und Gesellschaft, 4(3), 243–264.
- Bromberger, C. (2003): Fußball als Weltsicht und als Ritual. In: Belliger, A., & Krieger, D. J. (Hg.). *Ritualtheorien: ein einführendes Handbuch* (285–302). Wiesbaden: Springer-Verlag. VS Verlag.

- Collins, H. (2020): This is not a penalty! What's gone wrong with technology and football in the age of VAR? Draft working paper, Online unter: http://sites.cardiff.ac.uk/harrycollins/files/2020/01/Video_Assistant_Referee_VAR_judgement_versus_measurement.pdf
- Collins, H., Evans, R. & Higgins, C. (2016): Bad Call: Technology's Attack on Referees and Umpires and how to Fix it. Massachusetts: MIT Press.
- Colwell, S. (2004): Elite level refereeing in men's football: A developmental sociological account. Doctoral dissertation, University of Leicester.
- Cunningham, I., Simmons, P., Mascarenhas, D., & Redhead, S. (2015): Exploring player communication in interactions with sport officials. In: Movement & Sport Sciences-Science & Motricité, (87), 79–89.
- DeLand, M. (2013): Basketball in the Key of Law: The Significance of Disputing in Pick-Up Basketball. In: Law & society review, 47(3), 653–685.
- DFB (2021): Schiedsrichter Zeitung. 2/2021 <https://www.dfb.de/schiedsrichter/aktiverschiedsrichterin/artikel/schiedsrichter-zeitung-22021-3319/>
- Eichler, C. (2000): Lexikon der Fußballmythen. Frankfurt a.M.: Eichborn Verlag
- Elias, N. & Dunning, E. (1982): Sport im Zivilisationsprozeß: Studien zur Figurationssoziologie. Münster: LIT Verlag.
- FIFA (2021): Fußball-Regeln 2021/2022. <https://www.dfb.de/verbandservice/publikationen/fussballregeln/>
- Gebauer, Gunter 2006. Poetik des Fußballs. Frankfurt a.M.: Campus Verlag.
- Goffman, E. (1974): Frame analysis: An essay on the organization of experience. Harvard University Press.
- Goffman, E. (2018): Interaktion im öffentlichen Raum. Frankfurt: Campus Verlag.
- Heck, J. (2019): Die Angst des Schiris vor dem Elfmeter. Zur Interaktionssoziologie des Fußballspiels. Sport und Gesellschaft, 16(1), 33–60.
- Heimerl, B. (2013): Die Ultraschallsprechstunde: Eine Ethnografie pränataldiagnostischer Situationen. Bielefeld: transcript.
- Heintz, B. (2021): Kategorisieren, Vergleichen, Bewerten und Quantifizieren im Spiegel sozialer Beobachtungsformate. KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie, 73, 5–47.
- Hirschauer, S. (2009): Die Empiriegeladenheit von Theorien und der Erfindungsreichtum der Praxis.: In: Kalthoff/Hirschauer/Lindemann (Hg.): Theoretische Empirie: Zur Relevanz qualitativer Forschung. Frankfurt a.M.: Suhrkamp, 165–187

- Hochschild, A. R. (1983): *The Managed Heart. Commercialization of Human Feeling*. Berkeley: University of California Press.
- Hoppe, A. D. (2021). The Microsociology of Aesthetic Evaluation: Selecting Runway Fashion Models. *Qualitative Sociology*, 45, 63–87.
- Janetzko, A. (2015). Professionelles Sehen. Sichtungspraktiken im Tanzsport/Professional Seeing. *Talent Identification Practices in Dance. Sport und Gesellschaft*, 12(2), 105–132.
- Janetzko, A. (2021): Talent (be)werten. Eine praxeographische Untersuchung von Talentsichtungen im Leistungssport. Bielefeld, transcript.
- Kew, F. (1987): Contested rules: An explanation of how games change. In: *International Review for the Sociology of Sport*, 22(2), 125–135.
- Lambrix, P. (2022). Erstrebenswerte Unselbständigkeit? Die Ambivalenz der Pflegebegutachtung. *Zeitschrift für Soziologie*, 51.3, 227–243.
- Lamont, Michele. 2012. Toward a comparative sociology of valuation and evaluation. *Annual review of sociology* 38: 201–221.
- Lom, S. E. (2016). Changing rules, changing practices: The direct and indirect effects of tight coupling in figure skating. *Organization Science*, 27(1), 36–52.
- Minnetian, C., & Werron, T. (2021). Redefining achievement: The emergence of rankings in American baseball. In: Ringel/Espeland/Sauder/Werron (Hg.): *Worlds of rankings*. Emerald Publishing Limited, 127–151
- Rains, P. (1984): The production of fairness: officiating in the national hockey league. *Sociology of Sport Journal*, 1(2), 150–162.
- Ringel, Leopold, & Werron, Tobias (2020): Where do Rankings Come From. A Historical-sociological Perspective on the History of Modern Rankings. In: Epple, Angelika et.al. (Hg.): *Practices of Comparing. Towards a New Understanding of a Fundamental Human Practice*, 137–170.
- Schimank, U. (2005): Die Entwicklung des Sports zum gesellschaftlichen Teilsystem. In: Ders.: *Differenzierung und Integration der modernen Gesellschaft: Beiträge zur akteurzentrierten Differenzierungstheorie*. Wiesbaden: Springer Verlag, 111–142.
- Schindler, L., & Liegl, M. (2013): Praxisgeschulte Sehfertigkeit: Zur Fundierung audiovisueller Verfahren in der visuellen Soziologie. In: *Soziale Welt*, 51–67.
- Schröder, C. (2015): Ich Pfeife! Aus dem Leben eines Schiedsrichters. Tropen Verlag

- Schmidl, A. (2023). Visual discourses in sport. A sociological analysis of the implementation of the video evidence in cycling and football. *Visual Studies* 38.3-4 (2023): 366–376
- Sport1.de (2019): Handspiel-Chaos in der Bundesliga. <https://www.sport1.de/news/fussball/bundesliga/2019/05/handspiel-chaos-in-bundesliga-muss-der-fussball-voellig-umdenken>
- Sportschau.de (2023): Stefan Kießlings Phantom-Tor in Hoffenheim <https://www.sportschau.de/fussball/bundesliga/video-sportschau-hoffenheim-l-everkusen-phantom-tor-100.html>
- Stuttgarter Nachrichten (2019): Bizarres Schiedsrichter-Tor geht um die Welt (27.05). <https://www.stuttgarter-nachrichten.de/inhalt.harkemase-boys-gegen-hsv-hoek-bizarres-schiedsrichter-tor-geht-um-die-welt.db4c2d7c-4733-4b84-8771-0dea4259cece.html>
- Weigelin, M. (2022): Entscheidungen und ihre Bewertungen – Zur Mikrosoziologie des Schiedsrichter-Pfiffs. *Österreichische Zeitschrift für Soziologie*, 47(3), 225–246.
- Wagner, G., Guse, J. S., & Hasenbruch, M. (2023): »Eigentlich war es immer sonnenklar.« Zur Invisibilisierung von formaler Organisation in Bewertungspraktiken. *Berliner Journal für Soziologie*, 1–29.
- Welt.de: »DFB-Schiris schießen gegen den »Eifel-Django« <https://www.welt.de/sport/fussball/article2778743/DFB-Schiris-schiessen-gegen-den-Eifel-Django.html>
- Werron, T. (2010): Der Weltsport und sein Publikum. Zur Autonomie und Entstehung des modernen Sports. Weilerswist: Velbrück.
- Werron, T. (2015): Why do we believe in competition? A historical-sociological view of competition as an institutionalized modern imaginary. In: *Distinktion: Journal of Social Theory*, 16(2), 186–210.
- Zerubavel, E. (1980): If Simmel were a fieldworker: On formal sociological theory and analytical field research. *Symbolic Interaction*, 3(2), 25–34.
- Zglinski, J. (2022): Rules, standards, and the video assistant referee in football. *Sport, Ethics and Philosophy*, 16(1), 3–19.

Fairer Wettbewerb in Sport und Wissenschaft

Frank Meier

1. Einleitung¹

Wettbewerb – und Konkurrenz² – erfreuen sich seit einiger Zeit in Sozial- und Kulturwissenschaften gesteigerter Aufmerksamkeit (Rosa, 2006; Jessen, 2014; Kirchhoff, 2015; Werron, 2015; Mayer, 2019; Stark, 2020; Arora-Jonsson et al., 2020, 2021). Auch wenn Wettbewerb in der Literatur oft kurzsätzlich mit Ökonomie, Verwettbewerblichung mit Ökonomisierung assoziiert wird, ist eine solche Verbindung keineswegs zwingend. Wettbewerb kann als ebenso typisch für die Kunst (Preda, 2021) oder eben den Sport gelten.

Dieser Beitrag macht es sich zu Nutze, dass sich der Sport – zunächst das Feld selbst – in besonderer Weise die Fairness des Wettbewerbs zum Problem gemacht hat. Welche Einsichten lassen sich aus dem Sport und seiner soziologischen Untersuchung für die Analyse fairen Wettbewerbs in der Wissenschaft gewinnen? Dieser Frage sollen sich einige vorsichtige Sondierungen annähern. Diese »Probebohrungen« (Foucault 1983, S. 7) sollen moralisch aufgeladene normative Regeln der Angemessenheit in den Blick nehmen, die sich auf die Frage eines adäquaten Leistungsvergleichs beziehen.

Gegenstand dieses Beitrags ist Fairness *des* Wettbewerbs, nicht Fairness *im* Wettbewerb. Es geht also nicht darum, Erwartungen an faires Verhalten zu rekonstruieren, die an Wettbewerbsteilnehmerinnen adressiert werden. Vielmehr steht die Fairness des Wettbewerbs selbst im Fokus. Das schließt nicht

-
- 1 Die Forschung zu diesem Beitrag wurde im Rahmen des Projekts »Fächer im Wettbewerb« (Projektnummer 447967785) im Rahmen der Forschungsgruppe »Multipler Wettbewerb im Hochschulsystem« durch die DFG gefördert.
 - 2 Ich behandle hier Konkurrenz und Wettbewerb als synonym. Differenzierungen der beiden Begriffe sind keineswegs grundsätzlich unplausibel, haben sich im deutschen Sprachgebrauch jedoch nicht klar etabliert. Zudem werden im Englischen beide Begriffe üblicherweise gleichermaßen mit »competition« übersetzt.

aus, dass Wettbewerbe dann als unfair erachtet werden, wenn sie so gestaltet sind, dass sie unfaires Verhalten systematisch belohnen.

Es geht auch nicht um die Frage, ob es fair ist, den Zugang zu knappen Resourcen über Wettbewerb zu regeln. Vielmehr soll gefragt werden, welche Anforderungen an einen konkreten Wettbewerb gerichtet werden, damit dieser als legitim anerkannt wird. Wiederum schließt dies nicht aus, dass konkrete kritische Anfragen an die Fairness den Wettbewerb grundsätzlicher in Frage stellen können und womöglich sogar genau darauf abzielen. Wer das Spezifischere kritisiert, mag das Allgemeinere meinen.

Schließlich geht es nicht darum, eigene Urteile über die Fairness eines Wettbewerbs abzugeben. Vielmehr soll, im Modus der Beobachtung zweiter Ordnung, die soziale Konstruktion von Fairness in den Blick genommen werden. Das freilich sollte sich aus dem Gesagten bereits ergeben und sich im Kontext dieses Bandes ohnehin von selbst verstehen.

Das, was als fairer Wettbewerb gilt, so die Grundannahme dieses Beitrags, ist Leistungswettbewerb. Damit soll nicht gesagt werden, dass jeder Wettbewerb, der Leistungswettbewerb ist, als fair gilt, wohl aber dass Abweichungen vom Leistungsprinzip typischerweise als unfair erachtet werden oder genauer: dass zentrale moralisch aufgeladene Konfliktlinien über die angemessene Gestaltung von Wettbewerb sich auf solche Abweichungen beziehen. Ein Element soll dabei fokussiert werden: Leistung ist Leistung im Rahmen eines als plausibel und legitim anerkannten Vergleichshorizonts.

Der Beitrag kann damit zum einen an neuere organisationswissenschaftliche Beiträge zum Thema Wettbewerb anschließen (Arora-Jonsson et al., 2020, 2021), die, hier an eine der Kerneinsichten der so genannten Neuen Wirtschaftssoziologie zur Konstitution von Märkten erinnernd, Wettbewerb nicht als gegeben oder sich zumindest unter bestimmten strukturellen Bedingungen naturwüchsigt ergebend voraussetzen, sondern nach den Prozessen seiner sozialen Konstruktion fragen. Dabei treten insbesondere Akteure – nicht zuletzt Organisationen – in den Blick, die aktiv an der Wettbewerbskonstruktion beteiligt sind.³

³ Dabei griffe es im Übrigen zu kurz, solche von den Autoren im Anschluss an Simmel (s.u.) als »Vierte« bezeichneten Akteure lediglich als *Organisatorinnen* von Wettbewerben anzusprechen. Wie wichtig diese – nicht zuletzt auch als Adressaten von Kritik – sein mögen: Arora-Jonsson und Kollegen haben hier in einem sehr viel breiteren Sinne Instanzen vor Augen, die daran mitwirken, dass Akteure sich als im Wettbewerb befindlich interpretieren.

Zum anderen kann zumindest in einem losen Sinne an die Soziologie der Kritik (Boltanski & Thévenot, 2007) angeschlossen werden, wenn es darum geht, dass Wettbewerbsordnungen mit Bezug auf übergreifende und gegebenenfalls konfigrierende Regeln der Angemessenheit gestaltet, gerechtfertigt und kritisiert werden.

Die hier vorgestellten Überlegungen zum Leistungswettbewerb stehen offensichtlich in einem Zusammenhang zum Thema Meritokratie, also zu den Debatten darüber, ob gesellschaftliche Positionen, Ressourcen und Lebenschancen in modernen Gesellschaften auf der Grundlage von Meriten vergeben werden, die unterschiedlich verstanden, aber jedenfalls von askriptiven Merkmalen (Linton, 1936) unterschieden werden, und wie das dann zu bewerten wäre (klassisch Young, 1958; Bell, 1972; zuletzt Markovits, 2020; Sandel, 2020). Auch wenn Konflikte um die Fairness von Wettbewerben oft in der einen oder anderen Weise an Diskurse über gesellschaftliche Ungleichheit rückgebunden sind, können sie darauf nicht reduziert werden.

Im Weiteren wird der Beitrag in folgenden Schritten vorgehen: Zunächst soll geklärt werden, welche Phänomene mit den Begriffen Wettbewerb und Leistungswettbewerb gemeint sind und welche analytische Perspektive auf sie eingenommen werden soll (2). Der darauffolgende Abschnitt befasst sich mit differenziellen Voraussetzungen der Leistungserbringung (3). Dann soll die Thematik des fairen Leistungswettbewerbs zunächst im Feld des Sports (4) und daran anschließend und auf den dabei gewonnenen Einsichten aufbauend im Feld der Wissenschaft (5) sondiert werden.

2. Wettbewerb und Leistungswettbewerb

Was ist Wettbewerb? Zu Klärung dieser Frage lässt sich fruchtbar an Georg Simmels Soziologie der Konkurrenz (1903, 2013) anschließen. Dieser Startpunkt ist auch deshalb hilfreich, weil er es erlaubt, Wettbewerb nicht streng mit bestimmten sozialen Feldern zu assoziieren, sondern vielmehr als allgemeine soziale Form zu fassen. Auch Simmel theoretisiert zweifellos von bestimmten Modelfällen (Krause, 2021) her (insbesondere Sport, Wirtschaft, Politik, Liebe), wählt diese jedoch differenzierungstheoretisch betrachtet heterogen.

Nach Simmel vollzieht sich Konkurrenz als Form eines indirekten Konflikts in triadischen Konstellationen. Für unsere Zwecke rekonstruiert kann man sagen: In Wettbewerbskonstellationen sieht sich ein fokales Wett-

bewerbssubjekt auf der Position Ego gegenüber Konkurrierenden auf der Position Alter in einem »Streit« um ein knappes Gut, einen Siegpreis, der von einer Identität auf der Position des Tertius kontrolliert wird. Bei diesem Dritten ist insbesondere an Individuen, Organisationen aber auch an Kollektividentität wie »der Markt«, das »Publikum«, »die Wählerinnen« zu denken. Tertius vergibt nun, so jedenfalls die Erwartung Egos, den Siegpreis auf der Grundlage einer vergleichenden Bewertung der Offerten Egos und Alters. Damit wird Bewertung zum Schlüsselmoment des Wettbewerbs⁴, denn erst die Antizipation einer vergleichenden Bewertung einer Offerte durch Tertius ermöglicht es Ego (und analog den anderen Wettbewerbssubjekten auf der Alter-Position), zu konkurrieren⁵ und nicht etwa die Zuteilung der knappen Ressource nur fatalistisch hinnehmen oder in den direkten Konflikt mit Tertius (und gegebenenfalls Alter) um das Gut eintreten zu müssen.

Was ist nun vor diesem Hintergrund *Leistungswettbewerb*? Der Leistungsbegriff stellt offenbar auf Qualitäten der Offerte ab. Dabei ist zunächst an solche Qualitäten zu denken, die für das fokale Wettbewerbssubjekt durch eigene Anstrengungen beeinflussbar sind. Ego kann konkurrieren, in dem sie die Offerte im Lichte der antizipierten Bewertungskriterien Tertius' verbessert. Die Vorstellung der Leistungssteigerung durch Wettbewerb basiert zumindest zum Teil auf der Annahme, dass Wettbewerbssituationen zu solchem Konkurrieren anreize. Allerdings ist Konkurrieren über Leistungsverbesserung nicht die einzige Option, wie beispielsweise Werner Sombarts (1927) Unterscheidung von Leistungs- und Suggestionskonkurrenz illustriert.

Im Gegensatz zu Simmel, der Konkurrenz wie gesagt als allgemeine soziale Form analysiert, geht es Sombart in seiner Kapitalismusanalyse zunächst nur um ökonomischen Wettbewerb. Leistungs konkurrenz ist für den Autor (1927, S. 557) »Konkurrenz im eigentlichen, engeren, man kann auch sagen idealen Sinne, wenn man unter Konkurrenz dem Wortsinn gemäß Wettbewerb versteht«. Und dieser Wortsinn wiederum entstammt für Sombart, den Theoretiker der ökonomischen Konkurrenz und des Kapitalismus, bezeichnenderweise dem Sport: »Das Bild ist aus der Arena genommen« (Sombart, 1927, S. 557). Es gehe um »ein ›Mit-einander-um-die-Wette-Laufen‹, bei dem

⁴ Wettbewerbskonstellationen lassen sich als Spezifikationen von Bewertungskonstellationen (Meier et al., 2016; Waibel et al., 2021) interpretieren. Vgl. in diesem Sinne auch Meier (i.E.).

⁵ Vgl. zum hier zugrunde gelegten Verständnis von Konkurrieren Kosmützky & Meier, 2023.

einer der Sieger bleibt. [...] Preisrichter ist »das Publikum« (Sombart, 1927, S. 557). Im ökonomischen Zusammenhang ist »[d]ie Leistung, deren Feststellung gilt, [...] die beste und billigste Lieferung von Waren und Diensten. Der Preis ist die Auszeichnung durch den Kauf« (Sombart, 1927, S. 557). In der idealen Leistungskonkurrenz gilt, »daß das Publikum, sachkundig und unbeeinflußt, allein bestimmt, wer Sieger im Wettbewerb auf dem Markte bleibt, ausschließlich den Leistungen gemäß« (Sombart, 1927, S. 558).

Für Sombart ist Leistung also immer Leistung in den Augen des Publikums. Immerhin gilt es verschiedene Dimensionen von Qualität, Lieferung und Preis zueinander zu relationieren. Aber auch wenn man also nicht wird sagen können, dass Leistung eine von der Beobachtung des Publikums unabhängige Eigenschaft der Offerten sei, setzt Leistungskonkurrenz für den Autor doch voraus, dass die vergleichende Leistungsbewertung des Publikums in den objektiven Merkmalen der Offerte ihren Bezugspunkt findet. Was dies bedeutet, wird in der Unterscheidung von *Leistungskonkurrenz* und *Suggestionskonkurrenz* deutlich:

Hier will der Konkurrent auf das Urteil des Kunden nicht nur durch seine Leistungen wirken, sondern strebt, ihn auf andere Weise für sich einzunehmen, indem er dessen selbstständiges Denken, die eigene Überzeugungs- und Entschlussfähigkeit möglichst auszuschalten sucht, indem er zwangsläufig die von ihm beabsichtigten Vorstellungsweisen und Gefühle im andern zu erwecken trachtet, indem er diesem mit einem Wort den Kauf »suggeriert« (Sombart, 1927, S. 558).

In der Suggestionskonkurrenz, Sombart (1927, S. 558) denkt hier – offensichtlich äußerst kritisch gestimmt – an »Reklame«, löst sich demnach die Bindung des Wettbewerbsfolgs an die Leistung. Die Konkurrenz wird so vom normativen Ideal des Leistungswettbewerbs her betrachtet problematisch.⁶ Die Wettbewerbschancen sollen schließlich nicht durch eine verbesserte Leistung, sondern durch eine zweifelhafte Intervention in die Bewertung durch das Publikum beeinflusst werden. Dort wo Leistungswettbewerb herrscht, so muss man Sombart verstehen, belohnt er Leistung im Auge eines Publikums, das sich durch »selbstständiges Denken« und »eigene Überzeugungs- und Entschlussfähigkeit« auszeichnet.

⁶ Vgl. hierzu auch die Unterscheidung von Leistung und Erfolg bei Neckel (2001).

3. Differenzielle Voraussetzungen der Leistungserbringung

Im Wettbewerb der Warenmärkte, die Sombart vor Augen hatte, bezieht sich Leistung auf eine Offerte, deren Qualitäten am angebotenen physischen Objekt selbst ablesbar sind. Das Tuch, das eine Kundin ersteht, hat eine Qualität, eine Menge, einen Preis, Lieferbedingungen. Wie diese Offerte zustande kam, ist zunächst nicht Bestandteil der Leistung selbst. Neuere wirtschaftssoziologische Forschung hat darauf hingewiesen, dass auch kommunizierte Produktionsbedingungen (»nachhaltig«, »bei Vollmond geerntet«) im Sinne Vatins (2013) zur Valorisation, also zur Anreicherung eines Gutes mit Wert (Boltanski & Esquerre, 2020) genutzt werden können (vgl. Arnold & Hasse, 2016) und mithin zu Bestandteilen der erbrachten Leistung avancieren.

Dennoch bleibt es für ökonomischen Güter-Wettbewerb – und zweifellos auch für viele andere Wettbewerbstypen – eher untypisch, Leistungen von differenziellen Voraussetzungen der Leistungserbringung her zu denken. Wir werden sehen, dass das im Sport und zum Teil auch in der Wissenschaft durchaus anders aussieht. Jenseits spezifischer narrativer Inwertsetzungen (Peetz, 2021): Ob es den Produzierenden leichter oder schwerer fällt ihre Produkte herzustellen, bestimmt in der Regel gerade nicht die Leistungen in den Augen der Kundinnen. Dass von solchen Differenzen abgesehen wird, stellt dann für die Beobachterinnen auch nicht in Frage, dass es sich um Leistungswettbewerb handelt. Freilich findet diese Indifferenz gegenüber den Voraussetzungen der Leistungsproduktion dort ihre Grenzen, wo Unternehmen ihre besseren Möglichkeiten vornehmlich zu dem nutzen, was Sombart Gewaltkonkurrenz nennt und als dritte Form neben Leistungskonkurrenz und Suggestionskonkurrenz stellt. Gewaltkonkurrenz wird bestimmt durch Praktiken, die – etwa durch aggressive Preisgestaltung – darauf abzielen, Konkurrierende aus dem Wettbewerb zu drängen. Solche »cutthroat competition« (Reynolds, 1940) gilt geradezu als Inbegriff unfairen Wettbewerbs.⁷

Leistung kann also sowohl auf das reine Ergebnis eines Prozesses der Leistungsproduktion bezogen sein, wie auf diesen Prozess selbst und seine Voraussetzungen. Für eine vergleichende empirische Perspektive, wie sie hier vorgeschlagen wird, ist die Frage nun *nicht*, wie sie es in einer normativen Perspektive wäre, wie Leistung »eigentlich« erfasst werden sollte, sondern wie sie

⁷ Vgl. hierzu auch die auf fairen Wettbewerb abzielende US-amerikanische Industriepolitik der New-Deal-Ära (etwa Taylor, 2007).

praktisch konstruiert wird. Hierauf wird im Weiteren mit Blick auf Sport und Wissenschaft zu achten sein.

4. Leistungswettbewerb im Sport

Leistungsvergleich ist konstitutiv für den modernen Sport (Stichweh, 1990), der als Wettkampfsport durch die Verschränkung von kontinuierlicher Wettkampfpraxis und fortlaufendem öffentlichen Leistungsvergleich unter den Bedingungen vereinheitlichter Regeln gekennzeichnet ist (Werron, 2010). Gerade weil der Sport den Leistungsvergleich zum Selbstzweck erhebt, können plausibel anspruchsvolle normative Anforderungen an seine Ausgestaltung gerichtet werden. Als Zuschauersport ist er zudem auf die Akzeptanz des Publikums angewiesen. Allerdings ist das Publikum nicht nur und nicht immer an Fairness interessiert. Insofern ist zu erwarten, dass auch Belange der Attraktivität für Zuschauerinnen in der Gestaltung von Wettbewerben Berücksichtigung finden, die in Spannung zu Fairnessfragen stehen können.

Theorien des Wettbewerbs lassen sich fruchtbar auf den Sport anwenden (vgl. auch Weigelin in diesem Band). Ob Wettkampfsport in *jedem* Falle im *einzelnen* Wettkampf einem »simmelschen« triadischen Modell entspricht, ist eine interessante Frage. Sie wird insbesondere dann virulent, wenn es um Disziplinen geht, in denen Leistungen nicht wie im paradigmatischen Fall des Wettkampfs (für den wohl auch nicht an Mittelstrecke gedacht wird) parallel zum Gegner, »sowieso ohne ihn zu berühren« (Simmel 2013, S. 226) erbracht werden, sondern in direkter Auseinandersetzung gegen den Gegner: am augenfälligsten vielleicht im Kampfsport, aber auch etwa beim Fußball. Jedenfalls vertritt David Stark (2020, S. 4) die Auffassung, dass es sich beim Halbfinalespiel der Fußball-WM der Herren zwischen Brasilien und Deutschland (das er ins Jahr 2016 verlegt) nicht um triadischen, sondern um davon zu unterscheidenden agonalen Wettbewerb (*head-to-head competition*) gehandelt habe. Ohne dieser Frage hier zu viel Gewicht einzuräumen, sei zumindest darauf hingewiesen, dass es wohl zu kurz greift, das einzelne Wettkampfereignis aus dem Strom des fortlaufenden Leistungsvergleichs (vgl. wiederum und im Weiteren Werron, 2010) herausgelöst zu denken. Direkter und indirekter Wettbewerb können auch gleichzeitig stattfinden. Mit Starks eigener Unterscheidung könnte man von einer Verschränkung von *Competition* und *Competitions* sprechen.

4.1 Regeln

Sportlicher Leistungsvergleich setzt die Vereinheitlichung von Regeln voraus, zumal diese funktional nicht zuletzt auf die Vereinheitlichung durch Regeln bezogen sind. Regeln spezifizieren und vereinheitlichen, was in einem operativen sportlichen Vollzug als Leistung beobachtet und mit anderen Leistungen verglichen werden soll, sie steuern aber ebenso, welche Vollzüge zuallererst in der Wettkampfpraxis als Leistungen erbracht werden, die dann im Leistungsvergleich als äquivalent erscheinen können.

Die Bedeutung der Regeln geht aber noch weiter, zielen diese doch oft darauf ab, die Bedingungen zu vereinheitlichen unter denen Leistungen erbracht werden, etwa indem sie Sportstätten reglementieren oder angeben, unter welchen Bedingungen Wettkämpfe abgebrochen oder erbrachte Leistungen aus dem Vergleich ausgeschlossen werden sollen – zum Beispiel, wenn Leistungen wegen zu viel Rückenwind keine Anerkennung als Rekord finden. Regeln spezifizieren Leistungsvergleiche innerhalb von Sportarten, Vergleiche zwischen Disziplinen sehen sie normalerweise nicht vor.⁸

Wenn, wie unsere obigen Überlegungen nahelegen, differenzielle Bedingungen der Leistungsproduktion mitbestimmen, was als eine legitim zu erbringende Leistung im Rahmen eines Leistungswettbewerbs gilt, lohnt es sich hier genauer hinzuschauen.

Die Grundtendenz des Sports scheint zu sein, die Unterschiede in den äußeren Voraussetzungen der Leistungserbringung zu minimieren. Dass Regeln existieren, die dazu dienen sollen, einen interessanten – weil offenen – Wettkampf zwischen ungleich starken Teilnehmerinnen zu ermöglichen und zu diesem Zweck bewusst *ungleiche* Bedingungen schaffen (Handicap im Golf, Materialvorgabe im Schach), kann eher als Sonderfall gelten. Ebenso ein Sonderfall sind kalkulatorische Praktiken, die notgedrungen in Kauf genommene Differenzen in den äußeren Bedingungen rechnerisch auszugleichen versprechen.

Interessant sind hier vor allem neuere Regeln im Skispringen, einem Sport, in dem es – insbesondere windbedingt – offensichtliche und bedeutsame Unterschiede in den Bedingungen der Leistungserbringung gibt (Voigt, 2010; Hahn, 2013). Hier gibt es zum einen den Windfaktor, der, wie

⁸ Allerdings stellen Punktesysteme in Mehrkämpfen im Prinzip solche Vergleichsmöglichkeiten bereit. Sportartübergreifende Vergleiche können behelfsweise auf das große Kommensurierungsmedium Geld zurückgreifen (Knight 2023).

der Name nahelegt, ungleiche Windbedingungen beim Sprung kompensieren soll. Noch radikaler – wenn auch physikalisch einfacher zu bestimmen und weniger umstritten – erscheint jedoch der so genannte *Gatefaktor*. Dieser reagierte auf ein praktisches Problem. Sollte sich während eines laufenden Wertungsdurchgangs herausstellen, dass zu große Weiten erzielt werden, was die Sportlerinnen in einem problematischen Bereich gefährden könnte, muss der Anlauf verkürzt werden. Durch einen kürzeren Anlauf wird die Anlaufgeschwindigkeit verringert, was im Ergebnis wiederum zu geringeren Weiten führt. Vor der Regelinnovation führte nun eine solche Anlaufveränderung notgedrungen dazu, dass Springerinnen, die ihren Sprung bereits absolviert hatten, diesen nun mit kürzerem Anlauf wiederholen mussten, um nicht einen ungerechtfertigten Vorteil zu erhalten, was aber diese andererseits womöglich gerade dadurch benachteiligte. Der Gatefaktor macht solche Wiederholungen überflüssig, indem die Änderungen des Anlaufs rechnerisch kompensiert werden. Zweifellos spielten bei der Einführung des Gatefaktors daher nicht nur Überlegungen zur Fairness, sondern – sporttypisch – gerade auch zur Publikumsattraktivität eine Rolle.

Solche kalkulatorischen Praktiken transformieren die Leistung in eine ziemlich abstrakte Größe, in der bestimmte Voraussetzungen des Geleisteten (aber nicht andere) mitrepräsentiert werden. Freilich geschieht dies in einem Sport, der eine Vielzahl von Besonderheiten aufweist, die diese Praktiken weniger wegen erscheinen lassen: Schon erwähnt wurde der starke Einfluss nicht kontrollierbarer äußerer Bedingungen. Durch die Haltungsnoten war es im Sport bereits vorher etabliert, dass die Leistung in einer Mischgröße besteht und so nichtzwangsläufig die Person mit den weitesten Sprüngen den Wettbewerb gewonnen hat. Durch die sequenzielle und nicht parallele Durchführung der Sprünge wird dieser Sachverhalt weniger sichtbar. Allerdings zitiert die Süddeutsche Zeitung (Hahn, 2013, hier auch die folgenden Stimmen) den österreichischen Skispringer Gregor Schlierenzauer, von den Haltungsnoten offenbar absehend, wie folgt:

»Skispringen ist eine Sportart mit sehr viel Tradition, in der es darum geht, wer am weitesten springen kann[,] [...] das ist eigentlich der Hauptgedanke. Und wenn dann Situationen herauskommen, dass man kürzer springt als ein anderer und trotzdem vor dem ist, dann ist das nicht ganz in Ordnung für mich.«

Schlierenzauers Kritik richtet sich offenbar gegen die wahrgenommene Kontingenz der Kalkulation: »Es kommt irgendein Windfaktor heraus«. Tatsächlich ist die Berechnung problematisch, denn, wie der Konstrukteur des Compensationsmodells Hans-Heini Gasser einräumt: »Man kann nicht die ganze Natur ausmessen und in ein Rechenmodell packen«. Dennoch hält der Springer Tom Hilde die Kompensationsregel für »die beste Regel, die es je gegeben hat[,] [...] weil sie den Wettbewerb fairer macht«.

Auch wenn die Regel umstritten bleibt: Was im Skispringen geschieht, macht darauf aufmerksam, was in vielen anderen Kontexten nicht geschieht. Es macht darauf aufmerksam, welche Möglichkeiten des Vergleichbarma- chens andernorts gerade nicht genutzt werden.

4.1 Kategorien

Eine besonders aufschlussreiche Klasse von Regeln, stellen solche dar, in denen Kategorien des Leistungsvergleichs festgelegt werden, die diesen Vergleich organisieren, aber auch Grenzen der Vergleichbarkeit abbilden und reproduzie- ren.

Leistungsklassen, wie sie zum Beispiel in Ligen-Systemen vorkommen, sind zunächst Instrumente, die Leistungsvergleiche praktisch ermöglichen. Wenn in Sportarten, in denen Leistungsdifferenzen über direkte Konfrontationen zwischen Gegnerinnen ermittelt werden, aus pragmatischen Gründen nicht jede gegen jede antreten kann, ermöglichen Leistungsklassen ein selektiveres Wettkampfgeschehen, das zugleich die Fiktion eines Gesamtvergleichs plausibel aufrechterhalten lässt. Dies nicht in beliebigen Klassen, sondern Leistungsklassen zu tun, erlaubt zudem, spannendere – weil ergebnisoffene – Begegnungen wahrscheinlich zu machen, was zum einen ein Publikumsinteresse befriedigen und zum anderen die Teilnehmerinnen zu Anstrengungen anhalten mag, was wiederum dem Publikumsinteresse zugutekommen sollte.

Die kategoriale Ordnung des Vergleichs ist jedoch nicht auf solche Praktikabilität begrenzt, auch der Raum des Zulässigen ist kategorial bestimmt. Nun kann es verschiedene Gründe geben, an kategoriale Eigenschaften geknüpfte Vergleichstabus zu etablieren – beispielsweise kann es aus rassistischen, sexistischen oder klassistischen Motiven heraus darum gehen, bestimmte soziale Interaktionen zu vermeiden. Uns interessieren hier jedoch solche Gründe, die unmittelbar mit Erwartungen an sportliche Leistungsfähigkeit verknüpft sind. Alter, Gewicht, Geschlecht, Beeinträchtigung sind im Sport hochgradig institutionalisierte Dimensionen der kategorialen Schließung,

die in dieser Hinsicht an Plausibilitäten des Alltagswissens anschließen (vgl. Müller, 2006, 2018). Es erscheint alltagsweltlich unfair (vielleicht sogar sinnlos) Leistungsvergleiche über solche kategorialen Grenzen zu betreiben, weil die Voraussetzungen der Leistungserbringung als fundamental unterschiedlich eingeschätzt werden. Durch Regeln werden derartige Vergleiche zumindest im Wettkampfgeschehen selbst vermieden, verhindert, weniger sichtbar gemacht. Allerdings spricht Müller (2006, S. 398) in diesem Zusammenhang von »präsumtiven Leistungsklassen« und verweist damit auf die interessante Ambiguität: Im Unterschied zu den – wie man sie vielleicht nennen kann – overten Leistungsklassen ist es in diesen Fällen geradezu die Voraussetzung eines Vergleichs-Tabus, dass wie selbstverständlich unterstellt wird, der vermiedene Vergleich hätte, wäre er denn praktisch vollzogen worden, Leistungsunterschiede hervorgebracht⁹ – oder besser: vermeintliche Leistungsunterschiede, wenn man die differenziellen Bedingungen der Leistungserbringung mitbedenkt.

Kalkulatorische Lösungen wie im Skispringen scheinen wenig Plausibilität zu haben, um etwa über Geschlechtsunterschiede hinweg Leistungen vergleichbar zu machen.¹⁰ Dabei spielt sicher eine Rolle, dass es im Skispringen um äußere Bedingungen der Leistungserbringung geht, nicht um Voraussetzungen der Person selbst. Von diesen vollständig abzusehen, ist wohl nie das Programm des Sports gewesen. Unvergleichbarkeiten werden nur hochgradig selektiv konstruiert. Dabei sind die Grenzziehungen zwischen den Kategorien prekär und zunehmend strittig, wenn sich die vermeintlich natürlichen Kategorien als brüchig erweisen. Im Hinblick auf Altersjahrgänge ist beispielsweise ein stichtagsabhängiger *relative age effect* in Nachwuchskadern nachgewiesen worden (Musch & Grondin, 2001). Im Hinblick auf Geschlecht ist etwa an die Debatten zu denken, die sich um die intersexuelle Mittelstrecklerin Caster Semenya drehten (Cook & Dworkin, 2013), oder auch an die kontroverse Aufnahme der Olympiateilnahme der transgender Gewichtheberin Laurel Hubbard

-
- 9 Man kann hier in Anschluss an Meyer und Rowan (1977) von einem Mythos sprechen, womit lediglich gemeint sein soll, dass eine Annahme in einem sozialen Kontext un hinterfragt und ungeprüft als wahr geltend unterstellt werden kann. Müller (2009, S. 301–306) spricht mit Bezug auf raunende Erzählungen über – seinerzeit noch stärker verheimlichte – Testspiele zwischen Spitzfußballerinnen und Jugendspielern von »modernen Sportlegenden«.
- 10 Was nicht bedeutet, dass nicht gerechnet würde, um eine etwaige kategoriale Ungleichheit festzustellen.

(Panja & Belson, 2021). Namen wie Oscar Pistorius oder Markus Rehm verbinden sich mit Kontroversen um die Frage, ob behinderte – also vermeintlich benachteiligte – Sportler mit Prothesen *unfaire Vorteile* gegenüber Nichtbehinderten genießen (Müller, 2018). Die Konflikte, die sich hier entfalten, sind als Kontroversen über Fairness im Wettkampfsport gehalten, aber es wäre verfehlt anzunehmen, es ginge hier *nur* um Fairness. Vielmehr werden hier im Sport Auseinandersetzungen über gesellschaftliche Ontologien, Differenzierungen und Wertigkeiten geführt.

Was lässt sich nun gewappnet mit diesen Vorüberlegungen im Feld der Wissenschaft sehen?

5. Leistungswettbewerb in der Wissenschaft

5.1 Reputationswettbewerb

Aus der Sicht der »klassischen« Wissenschaftssoziologie ist Wettbewerb um Reputation konstitutiv für die moderne Wissenschaft (Merton, 1957; Bourdieu, 1975). Die Produktion und Weitergabe von neuem Wissen ist demnach nicht nur normativ durch das wissenschaftliche Ethos (Merton 1942) abgesichert, sondern auch motivational durch die Reputation, die derjenigen winkt, die neue Wissensansprüche mitteilt. Konkurrierende wären hier Wissenschaftlerinnen, die um den Siegpreis der Reputation ringen, der ihnen von der *scientific community* als einer kollektiven Dritten abhängig von der Bewertung wissenschaftlicher Beiträge zugewiesen wird.

Ähnlich wie sich im Sport die Frage stellt, ob es sich beim Wettkampfsport um triadischen Wettbewerb handelt, lässt sich in Zweifel ziehen, ob das »simmelsche« Modell den »klassischen« Reputationswettbewerb wirklich trifft. Denn: Ist Reputation eine knappe Ressource? Brunsson und Wedlin (2021) etwa bezweifeln dies und sehen die Funktion von Rankings nicht zuletzt darin, Reputation zuallererst zu verknappen. Unstrittig sollte die Knappheit allerdings im Prioritätswettbewerb (Merton, 1957) sein, also im Wettbewerb darum, eine wissenschaftliche Entdeckung zuerst gemacht und mitgeteilt zu haben. Der Prioritätswettbewerb ist gewiss nicht in allen Disziplinen gleichermaßen relevant, prägt aber zumindest Teile der Forschung und ist auch in neueren Beiträgen Gegenstand der *science studies* (vgl. etwa aus der ökonomischen Wissenschaftsforschung Hill & Stein, 2022).

So essenziell Reputationswettbewerb für die Wissenschaft sein mag, Wettbewerb ist nicht im gleichen Sinne konstitutiv für die Wissenschaft wie für den Sport. Wissenschaft ist eben nicht das Prozessieren von Leistungsvergleichen, sondern von Wissensansprüchen.

Die Wissenschaftssoziologie hat bereits früh auf Abweichungen vom Ideal des Leistungswettbewerbs im empirischen Reputationswettbewerb aufmerksam gemacht. Zu denken ist hier vor allem an das, was Robert K. Merton (1968) – theologisch eher beliebig¹¹ – *Matthäuseffekt* genannt hat. Wie sich am Beispiel von Mehrfachentdeckungen und Co-Publikationen plausibilisieren lässt, entspricht die Belohnung in der Form der Reputationszuweisung nicht einfach nur der erbrachten wissenschaftlichen Leistung, sondern ist ebenso von der schon bestehenden Reputation abhängig. Daraus folgt zwar nicht zwangsläufig, dass sich Leistung für die Einzelne nicht lohnt, wohl aber dass Wissenschaftlerinnen für die gleiche Leistung *unterschiedlich* belohnt werden. Die Konfliktlagen, die sich um das Phänomen der Mehrfachentdeckung ranken, machen zudem deutlich, dass es nicht lediglich die Entdeckung selbst ist, um die es in der Wissenschaft geht, sondern auch die Anerkennung als Entdeckerin durch die (jeweilige) wissenschaftliche Gemeinschaft, womit diese zur Dritten in Simmels triadischem Model avanciert.

Schon Merton selbst (insbesondere 1988) hat mit dem Begriff Matthäuseffekt unterschiedliche Formen der Vorteilsakkumulation zusammengefasst. Soweit man Leistung nur vom Ergebnis, nicht aber von den Voraussetzungen wissenschaftlicher Arbeit abhängig denkt, spricht Vorteilsakkumulation nicht grundsätzlich gegen das Prinzip des Leistungswettbewerbs. Gleichwohl kann eine zu ausgeprägte Ungleichheit in der Verteilung von Ressourcen dazu führen, dass der Reputationswettbewerb als solcher unterlaufen wird, wie etwa Richard Münch (2014) beklagt.

Dessen ungeachtet hat die wissenschaftssoziologische Ungleichheitsforschung die Bedeutung »nicht-meritokratischer« Kriterien – zu denken ist hier insbesondere an klassische Variablen der Ungleichheitsforschung – für wissenschaftliche Karrieren immer wieder herausgearbeitet, aber durchaus auch problematisiert (Gross & Jungbauer-Gans, 2007; van den Brink & Benschop, 2014; Auspurg et al., 2017; Zimmer, 2021). Bei allen Maßnahmen, die praktisch getroffen werden, um Benachteiligungen zu verhindern: Eine kategoriale Schließung der Vergleichshorizonte findet hier gerade nicht statt. Es erschien (heute) absurd zu fordern, Leistungen von Frauen dürften nicht

11 Vgl. die längliche Fußnote 7, Merton 1988, S. 609.

mit denen von Männern, solche von Akademikerkindern nicht mit denen von Arbeiterkindern verglichen werden. Auch kompensatorische Kalkulationen, etwa die Einführung eines Arbeiterinnenfaktors werden zumindest bislang nicht zur Problembehandlung herangezogen.¹²

5.2 Wettbewerbliche Steuerung

Solche Fragen werden noch bedeutsamer, wenn Wettbewerb gezielt als Steuerungsinstrument eingesetzt wird, wie das in vielen Wissenschafts- und Hochschulsystemen in den letzten Jahrzehnten verstärkt geschehen ist. In Deutschland hat sich seit den 1980er Jahren ein wettbewerbszentrierter wissenschaftspolitischer Diskurs (Meier, 2009), aber eben auch eine verwettbewerbliche politische und universitäre Praxis (Mayer, 2019) etabliert. Je expliziter Wettbewerb als ein bewusst zu Steuerungszwecken eingesetztes Instrument markiert wird, statt als ein inhärentes Merkmal der Wissenschaft gewissermaßen naturalisiert zu werden, desto sichtbarer wird die Kontingenz und desto stärker gerät die konkrete Ausgestaltung des Wettbewerbs potenziell unter Rechtferdigungsdruck.

Schaut man etwa in die Wettbewerbsstellungnahme des Wissenschaftsrats von 1985, die zweifellos ein Schlüsseldokument des frühen deutschen Wettbewerbsdiskurses darstellt, dann sieht man, dass der zu entfaltende Wettbewerb in Wissenschaft und Hochschule deutlich als Leistungswettbewerb imaginiert wurde, ging es doch explizit darum, sicherzustellen, dass diejenigen »belohnt werden, die mehr und Besseres leisten als andere« (Wissenschaftsrat 1985: 7). Daher sollten »knappe Güter materieller oder immaterieller Natur als Leistungsprämien denen zufallen, die sich in der Konkurrenz mit anderen als überlegen erweisen« (Wissenschaftsrat, 1985, S. 7).

Dabei war dann allerdings doch eine kategoriale Schließung der Vergleichshorizonte angelegt, denn es war klar, »daß es sinnvollerweise nur Fächer sein können, die beurteilt werden, nicht ganze Hochschulen« (Wissenschaftsrat, 1985, S. 7).¹³ Auch wenn sich die Lage spätestens mit der Exzellenzinitiative (jetzt -strategie) geändert hat (siehe zu dieser Transformation Meier,

¹² Am nächsten kommen dem Versuche, das »akademische Alter« von Bewerberinnen auf Stellen zu errechnen. Hier kommt es auch zu kategorialen Schließungen im Wettbewerb um »Nachwuchs«-Stellen und in bestimmten Drittmittelprogrammen.

¹³ Zur disziplinären Differenzierung in der Wissenschaft siehe zum Beispiel Stichweh 2013.

2017), bleibt die Vorstellung, dass sich Leistungsvergleiche im Rahmen fachbezogener oder – was das gleiche meint – disziplinärer Horizonte zu bewegen habe, wirkmächtig. Dass »[v]on Agrar- bis Wirtschaftswissenschaften: [a]lle über einen Leisten« geschlagen werden können, ist nicht nur für das Evaluationssystem der Wissenschaftlichen Kommission Niedersachsen (2009), der diese Formulierung entnommen ist, ausgeschlossen. Dabei geht es nicht nur darum, fachspezifische Kriterien anzuwenden, wie es in der normativen Evaluationsforschung etwa auch für metrische Bewertungen breit gefordert wird (zum Beispiel Butler & McAllister, 2011), und fachangehörige oder zumindest fachnahe Evaluatoren zu gewinnen, wie es für Drittmittelwettbewerbe üblich ist, sondern auch die Vergleichsobjekte disziplinär zu wählen.

Werfen wir einen Blick auf ein Zitat aus einem programmatischen Buch des langjährigen Generalsekretärs der Volkswagenstiftung Wilhelm Krull (2017). »In der vermessenen Hochschulwelt«, hier spielt Krull (2017, S. 28) auf den wortspielerischen Titel seines Buches, »Die vermessene Universität«, an,

scheinen zudem Wissenschaftspolitiker wie Hochschulleitungen immer mal wieder zu vergessen, dass man Zitations- und Drittmittelquoten eines Historikers nicht mit denen eines Biochemikers oder Ingenieurs [...] und die Anzahl der Promotionen pro Professur an einer medizinischen Fakultät nicht mit denen an einer philosophischen Fakultät vergleichen kann und soll.

Zunächst fällt auf, dass hier Vergleiche als unmöglich gekennzeichnet werden, die, – soweit die Daten vorliegen – sehr leicht möglich sind. Wie so oft beim »Unvergleichlichen« (vgl. Epple & Erhart 2020) heißt »kann und soll« – offenbar vor allem »soll«. Aber warum soll man nicht? Gibt es nicht Situationen, in denen sich solche Vergleiche als nützlich herausstellen könnten? Offenbar ist hier nicht an den Vergleich als solchen gedacht, sondern an den Vergleich als Leistungsvergleich. Differenzen in den genannten Quoten sollen nicht als Leistungsdifferenzen gelesen werden, jedenfalls nicht dann, wenn damit fachübergreifende Leistungsvergleiche beabsichtigt wären. Eine medizinische Fakultät soll nicht einfach deshalb als besser als eine philosophische Fakultät gelten, weil sie mehr Promotionen aufweist. Hintergrund des konkreten Beispiels sind offensichtlich auch die exzeptionellen Promotionsquoten in der Medizin.

Doch es stellen sich sogleich Folgefragen: Darf man innerhalb der philosophischen Fakultät die Promotionszahlen einer Anglistin mit denen einer Skan-

dinavistin vergleichen? Darf man die Zitations- und/oder Drittmittelquoten einer Zeitgeschichtlerin mit denen einer Mittelalterhistorikerin vergleichen?

Mit Blick auf das Skispringen kann man zudem Fragen: Lässt sich die Vergleichbarkeit zwischen Fächern in Hinblick auf Promotionszahlen, Zitations- und Drittmittelquoten nicht ebenso kalkulatorisch herstellen? Die Antwort ist hier selbstverständlich ja. Scientometrisch ist eine feldspezifische Normalisierung nicht nur möglich, sondern auch naheliegend. Inwiefern derartige Vergleiche in der öffentlichen vergleichenden Leistungsbewertung praktiziert werden und auf Anerkennung rechnen können, ist jedoch eine andere Frage.

Solche Spitzfindigkeiten gehen freilich an den Intentionen des Autors vorbei. An anderer Stelle kritisiert er Rankingkriterien, die »einen klaren Bias Richtung Natur-, Lebens- und Technikwissenschaften aufweisen« (Krull 2017, S. 24) und eine »Imbalance in der Berücksichtigung der verschiedenen Fächerkulturen« (Krull 2017, S. 25) zeigen. Und so ist es »insbesondere die Vernachlässigung der Geistes- und Gesellschaftswissenschaften« (Krull 2017, S. 25) als dem ganz Anderen der Wissenschaft, um die es hier geht. Das schließt an eine lange Geschichte der Krisendiskurse der Geisteswissenschaften an, die in den 1980er Jahren durch eine Debatte um die forschungspolitische Benachteiligung (zunächst auch: Vernachlässigung) der Geisteswissenschaften belebt wurde (vgl. hierzu auch Röß, i.E.).

In den Nullerjahren entspann sich eine keineswegs auf die Geisteswissenschaften beschränkte aber besonders mit Blick auf sie geführte Debatte um fachspezifische Leistungskriterien (vgl. etwa Lack & Marksches, 2008). Sie mündete in der ambitionierten Initiative der Wissenschaftsrats zu Pilotstudien zu einem Forschingsrating (Wissenschaftsrat 2008). Hier unterzogen Vertreterinnen eines Faches anhand selbstdefinierter und daher als disziplinspezifisch und disziplinenadäquat verstandener Kriterien Forschungseinheiten des jeweiligen Faches einer deutschlandweiten vergleichenden Bewertung in der Form eines – der Name verrät es – Ratings.¹⁴ Vor dem Hintergrund der Debatten um die Geisteswissenschaften entbehrt es nicht einer gewissen Ironie, dass es just die erste ausgewählte geisteswissenschaftliche Kandidatin – die Geschichte – war, die sich einer Beteiligung verweigerte (Plumpe, 2009). Letztlich folgte aus den aufwendigen Anstrengungen daraus nicht sehr viel. Jedenfalls wurde das Rating nicht wie angedacht auf Dauer

¹⁴ Zur Form des Ratings siehe Heintz, 2019.

gestellt. Ein großes nationales Evaluationssystem entstand in Deutschland anders als in anderen Ländern nicht.¹⁵

Derlei Aktivitäten belegen zwar, wie stark die Vorstellung, dass Leistungsvergleich in der Wissenschaft disziplinspezifisch zu erfolgen habe, zumindest im deutschen System verankert ist, die genaue Spezifizierung der Disziplinen und ihrer Grenzen ist aber keineswegs klar. Und das nimmt nicht Wunder, denn: »Die deutschen Hochschulen verfügen über ein geradezu beängstigend breites Spektrum an Fachgebieten und Spezialdisziplinen« (Krull, 2017, S. 104).

Doch es ist nicht nur eine Frage der Zahl: In der Tat stellt sich hier ein Kernproblem der normativen Idee disziplinärer Vergleichshorizonte (von der spezifischen Problematik interdisziplinärer Forschung ganz abgesehen): Welche Disziplinen es eigentlich gibt und wie ihre Grenzen genau verlaufen, ist alles andere als selbstverständlich oder konsentiert. Die vielfältigen Klassifikationsschemata, etwa die Fächersystematik der DFG, diejenige des Statistischen Bundesamtes oder der Publikationsdatenbanken, stimmen weder miteinander noch zwangsläufig dem Selbstverständnis der Forschenden überein und sind überdies nicht stabil. Epistemische Kulturen werden ohnehin (Knorr Cetina, 2002) als inkongruent zur Ordnung der Disziplinen gedacht.

Wir sehen also mit dem durch die sportsoziologische Forschung geschärften Blick ganz ähnliche Problemstellungen wie im Sport: Zwar scheint es zumindest den akademisch eingeweihten unmittelbar einsichtig, dass Chemikerinnen ganz anderes arbeiten und anderes leisten als Philosophinnen. Im Einzelnen sind die Grenzziehungen jedoch unklar und strittig. Und damit bleibt auch die Frage der legitimen Vergleichshorizonte notwendig prekär.¹⁶

¹⁵ Als funktionales Äquivalent lässt sich die – freilich als Drittmittelwettbewerb ganz anders funktionierende – Exzellenzinitiative verstehen.

¹⁶ Schließlich bleibt auch das praktische Problem, dass es bei Entscheidungen über Resourcenzuweisungen gar nicht ausbleiben kann, Entscheidungen über Disziplingen Grenzen hinweg zu treffen. Die Bewertungsforschung hat sich in dieser Hinsicht insbesondere mit der Frage auseinandergesetzt, welche Regeln der Fairness interdisziplinär besetzte Gutachtendengremien anleiten. Hierbei ist nicht zuletzt die Bedeutung der Anerkennung disziplinärer Autonomie, fachlicher Expertise und Offenheit für methodologischen Pluralismus herausgearbeitet worden (Lamont et al., 2006; Lamont & Huutoniemi, 2011).

Zusammenfassung und Ausblick

In diesem Beitrag ging es darum, zu demonstrieren, dass es sich beim Sport um einen instruktiven Fall handelt, aus dem man viel für die Analyse von Wettbewerb lernen kann. Dies gilt insbesondere auch für die normativen Regeln der Angemessenheit, die mit fairem Wettbewerb verbunden werden. Die sportsoziologische Forschung hat hierzu in aufschlussreichen Untersuchungen wertvolle Einsichten gewonnen, an die anzuschließen sich auch für Forschende, die über Wettbewerb in ganz anders gearteten sozialen Kontexten arbeiten, lohnt.

Was als fairer Wettbewerb gelten soll, so die Grundannahme des Beitrags, muss den Kriterien des Leistungswettbewerbs entsprechen, in dem der Wettbewerberfolg der erbrachten Leistung entspricht. In jedem Fall setzen sich Abweichungen von diesem Prinzip normativer Kritik aus. Fairer Wettbewerb setzt dabei aber einen Leistungsvergleich im Rahmen eines als plausibel und legitim erachteten Vergleichshorizonts voraus. Die Regeln des Sports bearbeiten dieses Problem durch vereinheitlichte und vereinheitlichende Regeln, durch kalkulatorische Praktiken, aber auch durch kategoriale Reglementierungen zulässiger Vergleiche. Es würde sich lohnen, systematischer herauszuarbeiten, wo im Sport genau aus welchen Gründen und mit welchen Folgen welche Problembearbeitungen gewählt werden.

Auch der Wettbewerb in der Wissenschaft soll normativ ein Leistungswettbewerb sein, in dem diejenigen »belohnt werden, die mehr und Besseres leisten als andere« (Wissenschaftsrat, 1985, S. 7). Und auch in der Wissenschaft stellt sich das Problem plausibler und legitimer Vergleichshorizonte. Interessant ist hier das Thema Disziplinarität. Es scheint oftmals selbstverständlich, wissenschaftlichen Wettbewerb als Wettbewerb innerhalb disziplinärer Grenzen zu denken. So natürlich es jedoch zu sein scheint, dass es Physikerinnen, Soziologinnen und Amerikanistinnen gibt, so wenig stabil und eindeutig sind disziplinäre Grenzziehungen, so wenig einheitlich und unumstritten sind disziplinäre Klassifikationsschemata. Die Disziplinen der Wissenschaft sind eben anders als Sportarten nicht als distinkte Wettbewerbsfelder angelegt, vielmehr folgen sie, zumindest der Selbstbeschreibung nach, den Dynamiken der wissenschaftlichen Wissensproduktion und -reproduktion. Zudem lassen sich Ressourcenentscheidungen über wie auch immer gezogene fachliche Grenzen hinweg nicht vermeiden. Auch hier spielen Regeln der Angemessenheit in der sozialen Konstruktion von Fairness eine zentrale Rolle. Wie genau und mit welchen Folgen disziplinäre Vergleichshorizonte in wissenschaftli-

che Wettbewerbe eingeschrieben sind, wie sie gegebenenfalls kritisiert und gerechtfertigt werden, gilt es genauer empirisch zu untersuchen.

Der hier eingeschlagene Pfad untersucht Vergleiche, setzt aber auch methodisch auf Vergleiche. Die Soziologie der Bewertung vergleichend anzugehen, ist vielfach gefordert worden (Lamont, 2012). Was hier nur in ersten Sondierungen angedeutet werden konnte, gilt es nun systematisch empirisch durchzuführen. Vergleiche auch über unterschiedliche soziale Felder hinweg anzulegen, erscheint besonders anspruchsvoll, aber auch besonders vielversprechend. Hier wird der Sport sein Anregungspotenzial entfalten.

Literatur

- Arnold, N., & Hasse, R. (2016). Die Organisation von Wertzuschreibungen: Zur Bedeutung von Drittparteien für die Signalisierung moralischer Qualitäten in Märkten. *Berliner Journal für Soziologie*, 26(3-4), 329–351.
- Arora-Jonsson, S., Brunsson, N., & Hasse, R. (2020). Where does competition come from? The role of organization. *Organization Theory*, 1(1). <https://doi.org/10.1177/263178771988997>
- Arora-Jonsson, S., Brunsson, N., & Hasse, R. (2021). A new understanding of competition. In S. Arora-Jonsson, N. Brunsson, R. Hasse & K. Lagerström (Hg.), *Competition. What it is and why it happens* (S. 1–25). Oxford University Press.
- Auspurg, K., Hinz, T., & Schneck, A. (2017). Berufungsverfahren als Turniere: Berufungschancen von Wissenschaftlerinnen und Wissenschaftlern. *Zeitschrift für Soziologie*, 46(4), 283–302.
- Bell, D. (1972). Meritocracy and equality. *The public interest*, 29, 29–68.
- Boltanski, L., & Thévenot, L. (2007). Über die Rechtfertigung: Eine Soziologie der kritischen Urteilskraft. Hamburg: Hamburger Edition.
- Boltanski, L., & Esquerre, A. (2020). Enrichment: A critique of commodities. Cambridge, Medford: polity.
- Bourdieu, P. (1975). The specificity of the scientific field and the social conditions of the progress of reason. *Sociology of Science*, 14(6), 19–47.
- Brunsson, N., & Wedlin, L. (2021). Constructing competition for status: sports and higher education. In S. Arora-Jonsson, N. Brunsson, R. Hasse & K. Lagerström (Hg.), *Competition. What it is and why it happens* (S. 93–111). Oxford University Press.

- Butler, L., & McAllister, I. (2011). Evaluating university research performance using metrics. *European political science*, 10, 44–58.
- Cooky, C., & Dworkin, S. L. (2013). Policing the Boundaries of Sex: A Critical Examination of Gender Verification and the Caster Semenya Controversy, *The Journal of Sex Research*, 50(2), 103–111.
- Epple, A., & Erhart, W. (2020). Practices of comparing: A new research agenda between typological and historical approaches. In A. Epple, W. Erhart & J. Grave (Hg.), *Practices of comparing: Towards a new understanding of a fundamental human practice* (S. 11–38). Bielefeld: Bielefeld University Press.
- Foucault, M. (1983). *Der Wille zum Wissen. Sexualität und Wahrheit 1*. Frankfurt a.M.: Suhrkamp.
- Gross, C., & Jungbauer-Gans, M. (2007). Erfolg durch Leistung? Ein Forschungsüberblick zum Thema Wissenschaftskarrieren. *Soziale Welt*, 58(4), 453–471.
- Hahn, T. (2013, 27. Dezember). Verzwickte Mathematik des Fliegens. *Süddeutsche Zeitung*. <https://www.sueddeutsche.de/sport/technologie-im-skispringen-verzwickte-mathematik-des-fliegens-1.1851305>
- Heintz, B. (2019). Vom Komparativ zum Superlativ: Eine kleine Soziologie der Rangliste. In S. Nicolae, M. Endreß, O. Berli & D. Bischur (Hg.), (Be)Werten. Beiträge zur sozialen Konstruktion von Wertigkeit (S. 45–79). Wiesbaden: Springer VS.
- Hill, R., & Stein, C. (2022). Race to the bottom: Competition and quality in science. Working Paper.
- Jessen, R. (Hg.) (2014). *Konkurrenz in der Geschichte: Praktiken—Werte—Institutionalisierungen*. Frankfurt a.M.: Campus-Verlag.
- Kirchhoff, T. (Hg.). (2015). *Konkurrenz: Historische, strukturelle und normative Perspektiven*. Bielefeld: transcript.
- Knight, B. (2023, 16. Mai). The world's highest-paid athletes. *Forbes*. <https://www.forbes.com/lists/athletes/>
- Knorr Cetina, K. (2002): *Wissenskulturen. Ein Vergleich naturwissenschaftlicher Wissensformen*. Frankfurt a.M.: Suhrkamp.
- Kosmützky, A. & Meier, F. (2023): Konkurrieren. Ein Beitrag zur Soziologie des Wettbewerbs. In P.-I. Villa (Hg.), *Polarisierte Welten. Verhandlungen des 41. Kongresses der Deutschen Gesellschaft für Soziologie 2022*. https://publikationen.soziologie.de/index.php/kongressband_2022/article/view/1612

- Krause, M. (2021). Model cases. On canonical research objects and sites. Chicago: University of Chicago Press.
- Krull, W. (2017). Die vermessene Universität. Ziel, Wunsch und Wirklichkeit. Wien: Passagen.
- Lack, E., & Marksches, C. J. (2008). What the hell is quality?: Qualitätsstandards in den Geisteswissenschaften. Campus-Verlag.
- Lamont, M. (2012). Toward a comparative sociology of valuation and evaluation. *Annual review of sociology*, 38, 201–221.
- Lamont M., & Huutoniemi, K. (2011). Comparing customary rules of fairness: evaluative practices in various types of peer review panels. In C. Camic., N. Gross & M. Lamont (Hg.), *Social knowledge in the making* (S. 209–32). Chicago: University of Chicago Press.
- Lamont, M., Mallard, G., & Guetzkow, J. (2006). Beyond blind faith: overcoming the obstacles to interdisciplinary evaluation. *Research Evaluation*, 15(1), 43–55.
- Linton, R. (1936). *The study of man*. New York: Appleton-Century-Crofts.
- Markovits, D. (2020). *The Meritocracy Trap*. Penguin.
- Mayer, A. (2019). Universitäten im Wettbewerb. Deutschland von den 1980er Jahren bis zur Exzellenzinitiative. Stuttgart: Franz Steiner.
- Meier, F. (2009). Die Universität als Akteur. Zum institutionellen Wandel der Hochschulorganisation. Wiesbaden: VS.
- Meier, F. (2017). Die Geburt des Wettbewerbs aus dem Geiste der Planung. Schwerpunktbildung, Ungleichheit und die Erfindung des Sonderforschungsbereichs. In J. Hamann, J. Maeße, V. Gengnagel, A. Hirschfeld (Hg.), *Macht in Wissenschaft und Gesellschaft. Diskurs- und feldanalytische Perspektiven* (S. 191–213). Wiesbaden: Springer VS.
- Meier, F. (i. E.). Competition and valuation in science and higher education. In: A. K. Krüger, T. Peetz & H. Schäfer (Hg.), *The Routledge international handbook of valuation and society*. London: Routledge.
- Meier, F., Peetz, T., & Waibel, D. (2016). Bewertungskonstellationen. Theoretische Überlegungen zur Soziologie der Bewertung. *Berliner Journal für Soziologie*, 26, 307–328.
- Merton, R. K. (1942). Science and Technology in a Democratic Order. *Journal of Legal and Political Sociology* 1: 115–26.
- Merton, R. K. (1957). Priorities in scientific discovery: A chapter in the sociology of science. *American Sociological Review*, 22(6), 635–659.
- Merton, R. K. (1968). The matthew effect in science. *Science*, New Series, 159, 56–63.

- Merton, R. K. (1988). The matthew effect in science, II: Cumulative advantage and the symbolism of intellectual property. *Isis*, 79(4), 606–623.
- Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology*, 83(2), 340–363.
- Müller, M. (2006). Geschlecht als Leistungsklasse. Der kleine Unterschied und seine großen Folgen am Beispiel der »gender verifications« im Leistungssport. *Zeitschrift für Soziologie*, 35, 392–412.
- Müller, M. (2009). Fußball als Paradoxon der Moderne. Zur Bedeutung ethnischer, nationaler und geschlechtlicher Differenzen im Profifußball. Wiesbaden: VS.
- Müller, M. (2018). Unvergleichbarkeitskonstruktion im Sport. Von Frauen mit Hyperandrogenismus und Männern mit Carbonprothesen. Stefan Hirschauer (Hg.), *Un/doing differences: Praktiken der Humandifferenzierung* (S. 203–233). Weilerswist: Velbrück.
- Münch, R. (2014). Academic capitalism. Universities in the global struggle for excellence. London/New York: Routledge.
- Musch, J., & Grondin, S. (2001). Unequal competition as an impediment to personal development: A review of the relative age effect in sport. *Developmental Review*, 21(2), 147–167.
- Neckel, S. (2001). »Leistung «und »Erfolg « Die symbolische Ordnung der Marktgemeinschaft. In E. Barlösius, H.-P. Müller & S. Sigmund (Hg.), *Gesellschaftsbilder im Umbruch: Soziologische Perspektiven in Deutschland* (S. 245–265). Wiesbaden: VS.
- Panja T., & Belson K. (2021, 31. Juli). Olympics' first openly transgender woman stakes debate on fairness. New York Times. <https://www.nytimes.com/2021/07/31/sports/laurel-hubbard-trans-weight-lifting.html>
- Peetz, T. (2021). Elemente einer Soziologie der Bewertung. In F. Meier & T. Peetz (Hg.), *Organisation und Bewertung* (S. 25–47). Wiesbaden: Springer VS.
- Plumpe, W. (2009, 18. Mai). Qualitätssmessung: W. Plumpe: Stellungnahme zum Rating des Wissenschaftsrates aus Sicht des Historikerverbandes. H-Soz-Kult. <https://www.hsozkult.de/debate/id/fddebate-132196>
- Preda, A. (2023). Rivalry as a social relationship: conceptualizing the micro-foundations of competition. *Distinktion: Journal of Social Theory*, 24(1), 87–110.
- Rosa, H. (2006). Wettbewerb als Interaktionsmodus. *Leviathan*, 34, 82–104.
- Röß, A. (i. E.). Zur Situation der Geisteswissenschaften im forcierten Hochschulwettbewerb. In: P. Fäßler & R. Pöppinghege (Hg.), *Junge Hochschu-*

- len – Emporkömmlinge zwischen Reform und Beharren? Stuttgart: Franz Steiner Verlag.
- Sandel, M. J. (2020). The tyranny of merit. What's become of the common good. Allen Lane.
- Simmel, G. (2013). Soziologie: Untersuchungen über die Formen der Vergesellschaftung. 7. Auflage. Berlin: Duncker & Humblot.
- Sombart, W. (1927). Der moderne Kapitalismus: Historisch-systematische Darstellung des gesamteuropäischen Wirtschaftslebens von seinen Anfängen bis zur Gegenwart. Band 3, Das Wirtschaftsleben im Zeitalter des Hochkapitalismus—Halbband 2, Der Hergang der hochkapitalistischen Wirtschaft; Die Gesamtwirtschaft. Berlin: Duncker & Humblot.
- Stark, D. (2020). The performance complex. In D. Stark (Hg.), *The performance complex: Competition and competitions in social life* (S. 1–28). Oxford: Oxford University Press.
- Stichweh, R. (1990). Sport – Ausdifferenzierung, Funktion, Code. *Sportwissenschaft* 20: 373–389.
- Stichweh, R. (2013). Wissenschaft, Universität, Professionen. Soziologische Analysen. Bielefeld: transcript.
- Taylor, J. (2007). Buy Now! Buy Here! The Rise and Fall of the Patriotic Blue Eagle Emblem, 1933–1935. *Essays in Economic & Business History*, 25, 117–130.
- van den Brink, M., & Benschop, Y. (2014). Gender in academic networking: The role of gatekeepers in professorial recruitment. *Journal of Management Studies*, 51(3), 460–492.
- Vatin, F. (2013). Valuation as evaluating and valorizing. *Valuation Studies*, 1(1), 31–50.
- Voigt, B. (2010, 28. Dezember). Den Wind eingerechnet: Neue Regeln bei der Vierschanzentournee. *Tagesspiegel*. <https://www.tagesspiegel.de/sport/nue-regeln-bei-der-vierschanzentournee-6464947.html>
- Waibel, D., Peetz, T., & Meier, F. (2021). Valuation constellations. *Valuation Studies*, 8(1), 33–66.
- Werron, T. (2010). Der Weltsport und sein Publikum. Zur Autonomie und Entstehung des modernen Sports. Weilerswist: Velbrück.
- Werron, T. (2015). Why do we believe in competition? A historical-sociological view of competition as an institutionalized modern imaginary. *Distinktion: Scandinavian Journal of Social Theory*, 16(2), 186–210.

- Wissenschaftliche Kommission Niedersachsen (2009). Von Agrar- bis Wirtschaftswissenschaften: Alle über einen Leisten. Zur Forschungsevaluation in Niedersachsen und ihren fachspezifischen Kriterien. Hannover: WKN.
- Wissenschaftsrat (1985). Empfehlungen zum Wettbewerb im deutschen Hochschulsystem. Köln.
- Wissenschaftsrat (2008). Pilotstudie Forschungsrating. Empfehlungen und Dokumentation. Köln.
- Young, M. (1958). *The Rise of the meritocracy 1870–2033. An essay on education and equality*. London: Thames and Hudson.
- Zimmer, L. (2021). Bildungsaufstiege in der Wissenschaft. Zur Nicht-Reproduktion sozialer Ungleichheit beim Übergang von der Junior- auf die Lebenszeitprofessur. *Zeitschrift für Soziologie*, 50(6), 415–433.

Autor:inneninformationen

David Kempf (M.A.) ist wissenschaftlicher Mitarbeiter im DFG-Graduiertenkolleg »Practicing Place« an der Katholischen Universität Eichstätt-Ingolstadt. In seiner Dissertation untersucht er ethnografisch Praktiken des realistischen Darstellens am Fall der Fußballsimulation Football Manager. Seine Forschungsschwerpunkte sind Praxistheorie, soziologische Ethnografie und Pragmatismus.

Ausgewählte Publikationen:

Play the Game. Know the Game – Rekursive WissensTransfers zwischen Football Manager und Fußball. In: Paidia. Zeitschrift für Computerspielforschung (2017).

Different Types of Users, Different Types of Immersion: A User Study of Interaction Design and Immersion in Consumer Virtual Reality. In: IEEE Consumer Electronics Magazine 7(4) (Juni 2018); (mit Jonathan Harth et. al.)

Dennis Krämer (Dr. phil.) ist Soziologe, Postdoc und Forschungsreferent an der Universität Münster. 2020 wurde er für seine Studie zu Intersexualität im Sport mit dem Nachwuchspreis der Sektion Sport- und Körpersozialogie der DGS sowie dem Wissenschaftspreis des DOSB ausgezeichnet (Bronzemedaille). Er ist Gutachter in einer interdisziplinären Kommission, die über das sensible Thema der medizinischen Behandlung von Kindern mit einer Variante der Geschlechtsentwicklung berät (BGB §1631e) und forscht neben dem gesellschaftlichen Umgang mit Geschlechterdiversität zur Bedeutung von Technologien in Krisensituationen.

Ausgewählte Publikationen:

Intersexualität im Sport. Mediale und medizinische Körperpolitiken. tran-

script (2021).

Challenging the Binary: Gender, Fraud, and the Complexities of Categorization in Elite Sports. In R. Kühl, D. Link, L. Heiberger (Hrsg.): Sexualitäten und Geschlechter. Historische Perspektiven im Wandel (S. 347–365). (mit D., Schyvinck, C.) transcript (2024).

Martin Lüthe (Dr.) ist wissenschaftlicher Mitarbeiter am John-F.-Kennedy Institut für Nordamerikastudien der Freien Universität Berlin, wo er zuvor als Juniorprofessor, Einstein Junior Fellow und Gastprofessor tätig war. Zu seinen Forschungsschwerpunkten zählen unter anderem die Kulturgeschichte Nordamerikas, Gender Studies, Kulturtheorie, Digital Game Studies, Popular Culture Studies und African American Studies. Lüthe ist außerdem Herausgeber bei Eludamos: Journal for Computer Game Culture und arbeitet an einer Monographie zur Kulturgeschichte des digitalen Fußballspiels.

Ausgewählte Publikationen:

»We Missed a Lot of Church, So the Music Is Our Confessional«: Rap and Religion, Lit Verlag (2008).

Color-Line and Crossing-Over: Motown and Performances of Blackness in 1960s American Culture, WVT (2011).

Frank Meier (Dr. phil.) ist wissenschaftlicher Mitarbeiter am SOCIUM, Universität Bremen. Er ist Mitglied der DFG-Forschungsgruppe »Multipler Wettbewerb im Hochschulsystem«. Forschungsschwerpunkte sind: Organisationssoziologie, Soziologie der Bewertung, Wissenschafts- und Hochschulforschung, Religionssoziologie.

Ausgewählte Publikationen:

The Agency of Scientific Disciplines. Historical Social Research, 48(3), 277–295 (2023); (mit Peetz, T.).

Der stigmatisierte Körper. Bewertungspraktiken und -konstellationen im Fall »Konnersreuth«. In: Österreichische Zeitschrift für Soziologie, 47(3), 269–290 (2021); (mit Waibel, D. & Peetz, T.)

Valuation constellations. Valuation Studies, 8(1), 33–66 (2021).

Clelia Minnetian (M.A.) ist wissenschaftliche Mitarbeiterin im DFG-Projekt «Zur Institutionalisierung der Rankings. Diskurskarrieren tabellarischer Leistungsvergleiche zwischen 1850 und 1980» an der Universität Bielefeld. Sie

promoviert zu gouvernementalen Techniken der staatlichen Berufsorientierung mit Blick auf Prozesse der Subjektivierung.

Ausgewählte Publikationen:

Redefining Achievement. The Emergence of Rankings in American Baseball. In: *Research in the Sociology of Organizations*, 74, 127–151; (2021); (mit Tobias Werron).

»Soziale Innovationen« für den Fortschritt von morgen. Eine diskursive Betrachtung der deutschen Innovationspolitik. In: Jannis Hergesell, Arne Maibaum, Clelia Minnetian, Ariane Sept (Hrsg.): *Innovationsphänomene. Modus und Effekte der Innovationsgesellschaft*, Wiesbaden, Springer VS, 145–166 (2017); (mit Lilli Braunisch).

Samira Mummelthey (M.A.) ist wissenschaftliche Mitarbeiterin am Institut für Erziehungswissenschaft an der Georg-August-Universität Göttingen. Sie promoviert am Arbeitsbereich von Prof. Kerstin Rabenstein. Ihre Forschungsschwerpunkte sind Differenz und Ungleichheit, Praktiken der (Be)Wertung und Konstruktionen von Leistung.

Ausgewählte Publikationen:

Konstruktionen von Leistung als Differenzordnung im Sportunterricht. Erste Ergebnisse einer praxistheoretischen Diskursethnographie. In: K. Rabenstein et al (Hg.), *Differenzkonstruktionen in fachunterrichtlichen Kontexten: Forschungsansätze und Erträge zu Differenz(de)konstruktion aus Fachdidaktik, Erziehungswissenschaft und Diversitätsforschung*, Verlag Julius Klinkhardt, 39–55 (2023).

Relationen von Leistung und Körper im Sportunterricht. Einblicke in ein diskursethnographisches Forschungsdesign. In: B. Zander, D. Rode, D. Schiller & D. Wolff (Hg.), *Qualitatives Forschen in der Sportpädagogik: Beiträge zu einer reflexiven Methodologie*, Springer VS, 385–405 (2023); (mit Rabenstein, K., Droepe, T. & Hunger, I.).

Thorsten Peetz (PD Dr.), vertritt den Lehrstuhl für Soziologie, insb. Soziologische Theorie an der Universität Bamberg. Zu seinen Forschungsgebieten zählen Soziologische Theorie, Soziologie der Bewertung, Kultur- und Organisationssoziologie sowie Soziologie der Digitalisierung.

Ausgewählte Publikationen:

- Intimate valuation devices. Doing valuation while doing dating in Tinder texts. In: *Journal of Cultural Economy* (2023).
Der stigmatisierte Körper. Bewertungspraktiken und -konstellationen im Fall »Konnersreuth«. In: *Österreichische Zeitschrift für Soziologie* (2022); (mit F. Meier).
Valuation constellations. In: *Valuation Studies* (2021); (mit D. Waibel und F. Meier).

Kerstin Rabenstein (Dr. phil. habil) lehrt Erziehungswissenschaft mit dem Schwerpunkt Schulpädagogik an der Georg-August-Universität Göttingen. Ihre Forschungsschwerpunkte sind die Transformation pädagogischer Ordnungen von Schule und Unterricht, sowie die Weiterentwicklung praxis-theoretischer Schul- und Unterrichtsforschung.

Ausgewählte Publikationen:

- Individualisierender Unterricht und der Wandel der Leistungsordnung. Erträge subjektivierungstheoretischer Forschung zur Responsibilisierung für Leistung in Praktiken. *Zeitschrift für Erziehungswissenschaft*, 26(6), 1491–1509 (2023); (mit Steinwand, J. & Strauß, S.).
Zur Herstellung der Schulkasse in Artefakten. Eine Praxeografie zu ersten Tagen von neuen 5. Klassen. In: *ZISU – Zeitschrift für interpretative Schul- und Unterrichtsforschung*, 12(1), 154–170 (2022).
Interferenzen in digitalen Praktiken der Bereitstellung von unterrichtlichen Aufgaben. In: *Sozialer Sinn*, 23(2), 297–315 (2022); (mit Wagener-Böck, N., Macgilchrist, F. & Bock, A.)

Robert Schmidt (Dr. phil.) ist Professor für Soziologie an der Katholischen Universität Eichstätt-Ingolstadt, Leiter des DFG-Forschungsprojekts »Accounting und transformatorische Effekte im Profifußball« und Sprecher des DFG-Graduiertenkollegs »Practicing Place«. Seine Forschungsschwerpunkte sind Praxistheorie, soziologische Ethnographie und Kulturanthropologie des Politischen.

Ausgewählte Publikationen:

- Toward a Culture-Analytical and Praxeological Perspective on Decision-Making. *Human Studies* (2022).
Bodies and embodiment in practices of valuation: Challenging the sociology

of valuation with the sociology of the body, in: Österreichische Zeitschrift für Soziologie (2021); (zusammen mit Kristina Brümmer, Stefan Laube, Hilmar Schäfer und Max Weigelin).

Situationalität und Verortung des Politischen. Eine praxeologische Forschungsskizze. In: Corsten, Michael (Hg.): Praxis. Ausüben. Begreifen. Weiterswist, Velbrück Wissenschaft (2021).

Max Weigelin (M.A.) ist ehem. Mitarbeiter des DFG-Projekts »Accounting und transformatorische Effekte im Profifußball« an der Katholischen Universität Eichstätt-Ingolstadt und nun Mitarbeiter am Arbeitsbereich für Wirtschafts- und Organisationssoziologie an MLU Halle. In seiner Dissertation untersucht er Schiedsrichter und Schiedsrichterwesen im Fußball organisationsethnografisch. Seine Forschungsschwerpunkte sind soziologische Ethnografie, Kultursoziologie, Organisationstheorie und Soziologie der Sinne.

Ausgewählte Publikationen:

Entscheidungen und ihre Bewertungen – Zur Mikrosoziologie des Schiedsrichter-Pfiffs. In: Österreichische Zeitschrift für Soziologie, 47(3), 225–246 (2022).

Raunen, Singen, Jubeln: Sinnessoziologische Sondierungen im Fußballstadion. Soziopolis: Gesellschaft beobachten (2023).

Tobias Werron ist Professor für soziologische Theorie und allgemeine Soziologie an der Universität Bielefeld. Gemeinsam mit Leopold Ringel leitet er zwei DFG-Forschungsprojekte zur Geschichte und Herstellung von Rankings. Seine Forschungsschwerpunkte sind Globalisierungs- und Weltgesellschaftsforschung, Soziologie der Konkurrenz, historische Soziologie des Nationalismus sowie Praktiken des Theoretisierens in den Geistes- und Sozialwissenschaften.

Ausgewählte Publikationen:

Practices of Global Challenging. A Historical Perspective and Preliminary Typology, in: Alina Isakova, Malte Neuwinger, Robin Schulze-Waltrud, Oday Uraiqat (Hrsg.): Constructing Global Challenges in World Politics, London (2024).

Review Essay: To Build a Future for Social Theory – What Do We Have to Know about Its past? In: American Journal of Sociology 130, No. 2.

