

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 Bewerterns 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 bildungssoziologische 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 (Berdelmann 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

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, Bewertens, 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, Bewertens, 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 Bedeutungsnahme 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 hät-ten – 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-entierungsnorm 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 Leistungserwartung 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 abbildbar 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. Differente 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 Diskurs-ethnographie. 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. <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-Teppe, 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/s42278-020-00098-0>
- Schmidt, R. (2017). Praxistheorie. In R. Gugutzer, G. Klein & M. Meuser (Hg.), *Handbuch Körpersoziologie: 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>