

gleichbedeutend mit einer schlichten Abbildung von Realität ist, aber dennoch unverändert wiederholt betrachtet werden können, geben sie Forschenden Zeit, flüchtiges Geschehen in aller Gründlichkeit zu analysieren (Knoblauch, 2004, S. 131). Beobachtungsprotokolle, die im Feld erstellt werden, sind dagegen schon von Beginn an (re-)konstruierende Darstellungen. Mit ihnen hat sich die forschende Person – auch wenn sie meint, sie sei offen für die Phänomene des Feldes – auf eine Sichtweise von Realität festgelegt. Die Möglichkeit der wiederholten Betrachtung von Videomaterial eröffnet zugleich einen forschungsmethodischen Vorteil gegenüber der Beschreibung desselben Vorgangs: Eben weil die Aufzeichnung nur in geringem Maße von Interpretationen beeinflusst ist (natürlich ist bis zum gewissen Maße die Kameraführung, der Bildausschnitt, das selektive Einfangen des Tons etc. schon eine Interpretation), lässt sie es zu, sie unter immer wieder anderen Fragestellungen anzusehen. Dinkelaker (2018) bezeichnet dies mit „reversibler Selektivität“, denn die Auswahl von Beobachtungsschwerpunkten kann jederzeit rückgängig gemacht werden. Ein und dasselbe Videomaterial kann daher für ganz unterschiedliche Forschungsfragen verwendet werden (s. auch schon Scherler, 2002). Während die Videoaufzeichnung selbst sich nicht verändert, würden jedoch immer andere Beschreibungstexte entstehen, die von der Fragestellung abhängen. Insofern sind Beschreibungen „prinzipiell unabschließbar“ (Bergmann, 1991, S. 91) und von der Perspektive der Betrachter\*innen abhängig. Dies lässt sich an einem Beispiel verdeutlichen, das von Loizos (2000, S. 96) stammt. Zeige man verschiedenen Personen das Bild eines Autos, entstünden ganz unterschiedliche sprachliche Beschreibungen: „one observer, looking at a photograph, sees ‚a motor car‘, a second one sees ‚an elderly medium-sized family saloon‘, a third one sees ‚a 1981 Ford Cortina with rally steering wheel and sports wheel-hubs“ (Loizos, 2000, S. 96). Das Bild des Autos ist dasselbe reale Objekt für alle drei Beobachter\*innen, aber sie nehmen es unterschiedlich wahr und wählen unterschiedliche Beschreibungen.

Videoaufnahmen enthalten zudem neben den bewegten Bildern oft auch sprachliche Äußerungen, besonders wenn es sich um Unterrichtsaufnahmen handelt. Man hat in der Simultanstruktur also nicht nur konkurrierende Elemente auf derselben Ebene, z. B. verschiedene einzelne Schüler\*innen, die sich alle unterschiedlich bewegen, sondern zusätzlich noch synchrone Informationen auf unterschiedlichen Ebenen, etwa wenn die sich bewegenden Schüler\*innen gleichzeitig etwas sagen, sich anfeuern oder die Lehrkraft Aufforderungen in die Sporthalle ruft. So erhöht sich die Komplexität der bewegten Bilder noch einmal. Beschreibungen dagegen bleiben immer in dem einen Medium der Sprache. Eng verknüpft damit ist ein weiterer Unterschied zwischen Bildern und Beschreibungen: Während Bilder noch offen für verschiedenste Deutungen sind – Moritz (2010, S. 165) spricht hier von einer „semantischen Unbestimmtheit“ –, wird mit einem Text diese Deutungsoffenheit erheblich eingeschränkt. Man könnte hier von einer „relativen Deutungsgeschlossenheit“ sprechen, die an die Sequenzialität von Texten gebunden ist.

Videoaufzeichnungen dienen häufig als Quelle für Fall(re)konstruktionen. Sie sind damit zunächst „vorwissenschaftliches“ Material, aus dem erst durch die Fragehaltung und Wahrnehmung der Forscher\*innen Daten für ein Forschungsprojekt werden. Als Mittel der wissenschaftlichen Verständigung eignen sich allerdings Videoaufzeichnungen nicht, da sie aufgrund ihrer Komplexität zu

wenig eindeutig sind. Mindestens flankierend ist also immer eine sprachliche Fassung des Inhalts der Videoaufzeichnungen nötig. Die Videotranskripte dienen der Verständigung über die bewegten Bilder, der besseren Übersicht und sie stehen gleichzeitig für eine zusätzliche Perspektive gegenüber dem Ursprungsmaterial (HAMPL, 2010, S. 54-55). Hierbei ist nicht von der Hand zu weisen, dass ein gewisses Sprachvermögen wichtig ist, damit Leser\*innen, die das Ursprungsmaterial nicht kennen, ein nachvollziehbarer, plastischer, ansprechender Text präsentiert werden kann. Oswald (2013, S. 185) weist darauf hin, dass qualitativ Forschende über eine gewisse Begabung zum Verfassen beschreibender Texte verfügen müssen. „Auf qualitativen Daten beruhende Texte kommen normalerweise nicht ohne größere Passagen aus, in denen Realität beschrieben wird. (...) Solche Beschreibungen werden umso beweiskräftiger, je treffender und damit überzeugender die Formulierungen gelingen“ (ebd.). Neben der von der Fragestellung und Erfahrung geleiteten Wahrnehmung als Grundlage der Fallkonstruktion tritt also bei Beschreibungen noch das Sprachvermögen und Sprachempfinden der Forscher\*innen hinzu.

Zusammenfassend muss man bei bewegten Bildern und bei Beschreibungen von zwei sehr unterschiedlichen Modi der Weltbegegnung – um nicht den ungewissen Wirklichkeitsbegriff zu nutzen – ausgehen. Während bewegte Bilder Räume, Personen und deren Handlungen in einer Simultanstruktur wiedergeben und dabei sowohl Visuelles als auch Sprachliches vorkommen kann, sind Beschreibungen rein sprachliche Erzeugnisse, die sich durch Sequenzialität auszeichnen. Eine einfache Entsprechung von Bild und Wort kann es daher nicht geben. Und so bemühen einige Autoren auch Metaphern, also sprachliche Bilder, um sich diesem Problem zu nähern.

„Decoding or translation serves as a bridge between the visual, which in western culture we associate with intuition, art, and implicit knowledge, and the verbal, which we have come to associate with reason, fact, and objective information“ (Collier & Collier, 1986, S. 169-170). Im Folgenden geht es daher um die Konstruktion von Brücken, sprich um die verschiedenen methodischen Vorschläge, wie man von bewegten Bildern zu Beschreibungen gelangt.

### **3 | FORSCHUNGSMETHODISCHE ANGEBOTE ZUR HERSTELLUNG VON BESCHREIBUNGEN**

Knoblauch et al. (2009, S. 13) weisen auf eine Vielfalt von qualitativen und interpretativen Ansätzen hin, die videografisch arbeiten, u. a. nennen sie Ethnomethodologie, Konversationsanalyse, „genre analysis“, Grounded Theory sowie soziologische Hermeneutik. Für dieses Kapitel habe ich diejenigen Ansätze ausgewählt, die dezidierte Aussagen zur Versprachlichung von Visuellem machen, eine gewisse Verbreitung in sozialwissenschaftlichen Studien gefunden haben und die daher auch in aktuellen Studien im sportwissenschaftlichen Kontext (s. Kap. 4) verwendet werden.

#### **3.1 | DOKUMENTARISCHE METHODE**

Die dokumentarische Methode gehört zu den rekonstruktiven Verfahren der empirischen Sozialforschung und ist eng mit dem Namen Bohnsack verknüpft, der sich wiederum auf die Grundbegriffe des Wissenssoziologen Karl Mannheim beruft (z.B. Bohnsack, 2014, S. 58). Grundsätzlich geht die dokumentarische Methode mit zwei Schritten der Interpretation vor (Bohnsack, 2013,

S. 213-14)<sup>2</sup>; der erste Schritt heißt formulierende Interpretation und erfasst die Inhalte und Themen des empirischen Materials. Die Frage nach dem „Was?“ der formulierenden Interpretation wird im zweiten Schritt um die Frage nach dem „Wie?“ ergänzt. Bei dieser reflektierenden Interpretation liegt der Fokus auf dem Rahmen, der Entstehung oder Herstellung bestimmter sozialer Handlungen oder Phänomene. Beispielsweise wird bei Gruppendiskussionen die Art und Weise der Diskursorganisation untersucht. Nach den genannten Interpretationsschritten folgt die Typenbildung, die darauf abzielt, sowohl einzelnen Typen durch Fallvergleich voneinander abzugrenzen als auch eine gesamte Typologie zu entwickeln (ebd., S. 214).

Bohnsack unterscheidet Bild-, Film- oder Videoprodukte, welche von den Erforschten selbst produziert, und solche, die von Forschenden zu Erhebungszwecken hergestellt wurden (Bohnsack, 2011, S. 12). Während Erstere einem gewissen Gestaltungswillen unterliegen und der Selbstdarstellung dienen, entstehen Letztere unter dem Blickwinkel einer bestimmten Forschungsfrage. Bohnsack selbst bezieht sich in seinen Beispielen und Analysen meist auf erstere. Daher wird neben dem zu Sehenden extensiv auch die Machart des Filmmaterials analysiert (Kameraeinstellung, Bildkomposition etc.). Diese oft stark kunsthistorische und bildästhetische Sichtweise soll hier aber nicht weiterverfolgt werden. Ich konzentriere mich auf diejenigen Videoprodukte, die zu Forschungszwecken und ohne bewusste künstlerische Gestaltung hergestellt wurden.

Im Gegensatz zu Raab (2008) versteht Bohnsack (2011, S. 21) Bilder oder Videoprodukte nicht als strukturähnlich zu Texten – vielmehr müsse man die Eigengesetzlichkeit von Bildern anerkennen. Bilder und Texte sind je eigene Medien mit für sie charakteristischen Merkmalen der Wirklichkeitsdarstellung (s. auch Kap. 2). Daher verbiete sich das Vorgehen einer Sequenzanalyse, die man bei Texten anwenden könne. Allerdings formuliert auch Bohnsack gemeinsame Standards für die Interpretation von (bewegten) Bildern und Texten. Sie bestehen in dem Wechsel der AnalyseEinstellung vom Was? zum Wie?, die in der Methodologie der dokumentarischen Methode angelegt ist, in einer Rekonstruktion der Formalstruktur (formaler Aufbau eines Textes bzw. formale Komposition eines Bildes) sowie in der komparativen Analyse unterschiedlicher Fälle (Bohnsack, 2011, S. 53).

Zur Beschreibung und Interpretation von bewegten Bildern greift Bohnsack (2011) besonders auf das Vokabular von Panofsky zurück. Panofsky, ein bedeutender deutsch-amerikanischer Kunsthistoriker, hat schon in den 1930er-Jahren ein Interpretationsschema zur Deutung von Kunstwerken entwickelt. Er unterscheidet dabei verschiedene Stufen im Interpretationsprozess (Panofsky, 1991b, S. 223). Die vorikonografische Beschreibung soll als erster Schritt die sichtbaren Objekte bzw. Motive (im kunsthistorischen Sinne), die Konfigurationen, Stimmungen und Gestiken darstellen. Vorrangig geht es um die Erschließung des „Phänomensinns“ (Panofsky, 1991a, S. 188). Im zweiten Schritt, der ikonografische Analyse, werden die Bedeutungen der Motive herausge-

---

2 Da es mir um die Frage geht, wie bewegte Bilder in Sprache gefasst werden können, beschränke ich mich auf knappe Anmerkungen zur Methodologie, die nur dazu dienen, die konkreten Hinweise zur Verschriftlichung von Videomaterial verstehen zu können.

arbeitet sowie Themen und Konzepte des Kunstwerks erschlossen. Hierbei bezieht man sich auf die konventionale Bedeutung des Kunstwerks – dies ist die Ebene des Bedeutungssinns (ebd., S. 188). Die ikonologische Interpretation schließlich – der dritte Schritt – zeichnet sich dadurch aus, dass symbolische Werte angesprochen und ein größerer Zusammenhang durch z. B. nationale, epochale, religiöse oder philosophische Prinzipien hergestellt werden. In diesem letzten Schritt kann die Interpretation durchaus über die bewusste Intention des Künstlers hinausgehen – es wird der „Dokumentsinn“ (Panofsky, 1991a, S. 203) freigelegt. Mit diesem dreistufigen Verfahren hat Panofsky eine grundlegende Methodik geliefert, die deswegen noch heute interessant ist, weil sie zwischen dem Akt der Beschreibung, der Deutung auf der konventionalen Ebene und der Interpretation mit der Perspektive des „Wie?“ unterscheidet.

Als erste Stufe der dokumentarischen Videointerpretation sieht Bohnsack in Anlehnung an Panofsky die ikonografische Beschreibung. „Im Bereich der *ikonographischen* Beschreibung oder Interpretation findet unser sprachlich-textliches Vorwissen also lediglich insoweit Berücksichtigung, als es sich um die kommunikativ-generalisierten Wissensbestände handelt. Dieser bedienen wir uns weitgehend in stereotypisierender Form ...“ (Bohnsack, 2011, S. 57, Hervorhebung i. O.). Im Zuge der reflektierenden Interpretation soll dagegen die formale Komposition der (bewegten) Bilder rekonstruiert werden. Hierzu verwendet Bohnsack das Handwerkszeug der Kunstgeschichte, indem er etwa den Bildaufbau planimetrisch oder perspektivisch untersucht, wobei meist bewusst gestaltete visuelle Medien Gegenstand der Interpretation sind. Da Videomitschnitte, die zu Forschungszwecken beispielsweise Sportunterricht dokumentieren, einen deutlich anderen Charakter haben, gehe ich nicht weiter auf die kompositorische Analyse ein.

Während die wörtlich zitierte Textstelle zur Beschreibung zunächst enttäuschend wirkt, gibt es für die Fragestellung dieses Beitrags jedoch weitere interessante Differenzierungen. So unterscheidet Bohnsack (2011, S. 147) Ebenen, auf denen Bewegungen beschrieben werden können (s. Tab. 1).

Tab. 1: Ebenen von Bewegungsbeschreibungen (modifiziert nach Bohnsack, 2010, S. 280; 2011, S. 147)

Vorikonografische Ebene	Beispiel	Motivkonstruktion
<b>Kineme (Elemente von Gebärden)</b>	Kopf und Schulter gehen nach vorne, Becken nach hinten etc.	-
<b>Gebärden</b>	Beugen des Rumpfs	-
<b>Operative Handlungen</b>	Sich-Setzen	A beugt den Rumpf, um sich zu setzen (Um-zu-Motiv beobachtbar)
Ikonografische Ebene	Beispiel	Motivkonstruktion
<b>Institutionalisierte Handlungen</b>	Lehrerin setzt sich ans Pult	A setzt sich, um den Unterricht zu beginnen (Um-zu-Motiv nicht beobachtbar)

Wie man sieht, teilt Bohnsack den Beschreibungsprozess in eine vorikonografische und in eine ikonografische Ebene ein. Die vor-ikonografische bleibt dabei ganz nah am Beobachtbaren, indem zunächst „Kineme“ – also kleinstmögliche Teile einer Bewegung (im Beispiel: Kopf und Schulter gehen nach vorne, Becken nach hinten etc.) – im Fokus stehen. Etwas allgemeiner werden dann die Gebärden, die einzelne Kineme zusammenfassen (Beugen des Rumpfs). Mit operativen Handlungen bezeichnet Bohnsack Gebärden, denen schon ein „Um-zu-Motiv“ zugeschrieben wird (Sich-Setzen). Auf der ikonografischen Ebene schließlich wird das Um-zu-Motiv auf den institutionalisierten Handlungskontext bezogen (A setzt sich, um den Unterricht zu beginnen), womit die Ebene der reinen Beobachtung verlassen wird, denn das Motiv der Handlung selbst ist nicht mehr zu beobachten, sondern nur aus dem Gesamtzusammenhang zu erschließen.

Man könnte sagen, dass die Beschreibungsschritte sich von der äußerlich zu beobachtenden Bewegung immer weiter entfernen, indem sie interpretiert wird. Ein ähnliches Beispiel nutzen Bohnsack, Fritzsche und Wagner-Willi (2015, S. 18): Wenn man die Kineme „rechter Ober- und Unterarm gehen nach oben, Zeigefinger der Hand wird ausgestreckt, während die anderen Finger sich leicht krümmen“ als operative Handlung deutet, dann wird daraus das schulische „Aufzeigen“. Auf der ikonografischen Ebene lässt sich die Bewegung beschreiben als Sich-Melden im Schulunterricht: Eine Schülerin zeigt auf, um ihren Redewunsch zu signalisieren. Das Verfahren, sich dergestalt vorsichtig an die Beschreibung von Bewegungen heranzutasten, verfolgt das Ziel, vorschnelle Deutungen zu erschweren und so die geforderte methodische Kontrolle im Forschungsprozess zu gewährleisten. Ebenfalls ist zu bedenken, dass die Analyse von Bewegungen es erfordert, die Simultaneität als auch den sequenziellen Bewegungsablauf im Blick zu behalten. „Die tiefere Semantik einer Gebärde (ihr Dokumentsinn) kann in valider Weise nur in Relation zu simultanen anderen Gesten, also in Relation zum gesamten Körper und zur gesamten Interaktionsszene (...) interpretiert werden“ (Bohnsack, 2011, S. 165). Wagner-Willi (2005; 2007; 2009) stellt daher heraus, dass bei Interpretationen von Videoaufnahmen sowohl die simultane als auch die sequenzielle Betrachtung nötig sei. Viele Autor\*innen raten jedoch dazu, nicht die bewegten Bilder als Ausgangspunkt des Interpretationsprozesses zu nehmen, sondern Standbilder (z. B. Raab, 2008, S. 160). Mit diesem Kunstgriff soll der/die Interpret\*in „gebremst“ werden, um nicht im Strudel der Bilder und alltäglichen Seh- und Deutungsgewohnheiten unterzugehen. Vielmehr kann es mit Standbildern gelingen, einen ausreichenden Abstand zum Forschungsgegenstand herzustellen.

Bei aller einleuchtender Methodik gibt es jedoch m. E. drei wichtige Einschränkungen zu beachten. Erstens sind die Beispiele, die zur Veranschaulichung der dokumentarischen Vorgehensweise dienen, beide auf konventionale Bewegungen ausgerichtet, deren Um-zu-Motiv dadurch für Personen, die sich mit den Routinen der Schule auskennen, relativ leicht zu entschlüsseln ist. Wenn man jedoch Bewegungen im sportlichen Kontext betrachtet, sind Um-zu-Motive sehr viel schwieriger zu rekonstruieren, denn gerade sportliche Aktivitäten sind häufig autotelisch (Prohl, 1999, S. 230-231), d. h. sie tragen ihren Sinn in sich selbst. Zweitens wird in den Beispielen Material herangezogen, das die Bewegung einer einzelnen Person zum Gegenstand macht und damit äußerst mikroskopisch vorgeht. Asbrand und Martens (2018, S. 182) wenden ein, dass die dokumenta-

rische Unterrichtsforschung Situationen mit vielen Sprecher\*innen rekonstruieren muss, sodass selbst Fotogramme (sprich: Standbilder) nur sehr selektiv verwendet werden könnten. Ansonsten seien umfangreichere Sequenzen, die benötigt würden, um fachliche Lernprozesse analysieren zu können, gar nicht mehr zu bewältigen. Es ist auffällig, dass Asbrand und Martens sich in ihren Beispielen ausschließlich auf „kognitive“ Schulfächer beziehen. Gleichwohl erstellen sie für die formulierende Interpretation Beschreibungstexte, die auch Bewegungen in Worte fassen (Asbrand & Martens, 2018, z. B. S. 308). Über das Zustandekommen der Beschreibungen reflektiert das Autorenteam nicht, während der Umgang mit der gesprochenen Sprache ausführlich zum Thema gemacht wird (ebd., S. 178-183). Wenn schon „Sitzunterricht“ als zu detailreich erachtet wird, um Bohnsacks Methodik anzuwenden, gilt dies erst recht für den Sportunterricht, denn man hat es in der Regel mit einem viel weiteren Blickwinkel und vielen sich bewegenden Individuen zu tun. In der oben beschriebenen Weise vorzugehen, wäre forschungspraktisch unmöglich. Hampf (2015, S. 441-442) weist denn auch darauf hin, dass die Kameraeinstellung „Totale“ mit zu hohem Detailreichtum einhergehe und daher bei der Datenerhebung Beobachtungsprotokolle zu führen seien, die einen gewissen Überblick über die Situation sichern können (ebd.).

Die dritte Einschränkung betrifft zentral die sprachliche Übersetzung. Wie Bohnsack et al. (2015, S. 21) selbst feststellen, existiert eine eindeutige Beschreibungssprache auf der vorikonografischen Ebene bisher nicht. Selbst wenn man alle Kineme einzeln sprachlich fassen wollte – was zu einem schier unlesbaren Text führen würde –, so ist doch selbst bei den Beispielen, die Bohnsack anführt (s. o.), immer noch Interpretationsspielraum übrig. Nehmen wir die sprachlichen Ausdrücke aus Tab. 1: „Kopf und Schulter gehen nach vorne, Becken nach hinten“. Sie klingen erst einmal präzise, aber wenn wir sie gedankenexperimentell als Bewegungsanweisung verstünden, dann wird klar, dass sie außer für das Sich-Hinsetzen genauso gut für eine Verbeugung oder den Beginn einer Rollbewegung verwendet werden könnten. Eine zweifelsfreie Beschreibungssprache scheint damit tatsächlich noch nicht entwickelt worden zu sein.

### 3.2 | ERZIEHUNGSWISSENSCHAFTLICHE VIDEOGRAFIE

Auch in der Erziehungswissenschaft haben sich forschungsmethodische Ansätze zur Videografie entwickelt. Besonders die Interaktionsanalyse spielt dabei eine Rolle. Sie versucht, die sequenziellen und die simultanen Strukturen von bewegten Bildern gleichermaßen zu berücksichtigen, indem insgesamt vier Verfahren zum Einsatz kommen: die Segmentierungs-, die Konfigurations-, die Sequenz- und die Konstellationsanalyse (s. Tab. 2).

Tab. 2: Verfahren der videografischen Interaktionsanalyse (nach Dinkelaker & Herrle, 2009, S. 53)

	Überblick über Gesamtordnung	Relation einzelner Interaktionselemente
Sequenzielle Strukturen	Segmentierungsanalyse	Sequenzanalyse
Simultane Strukturen	Konfigurationsanalyse	Konstellationsanalyse

Die Segmentierungsanalyse dient dazu, einen Überblick über den sequenziellen Verlauf einer Interaktion herzustellen, während die Konfigurationsanalyse die simultane Raumordnung fokussiert. Bei der Sequenzanalyse werden die sinnhaften Strukturen von Interaktionssequenzen nachvollzogen; die Konstellationsanalyse schließlich zielt auf die Bedeutung eines Raumelements im Kontext aller anderen simultan wahrnehmbaren Raumelemente. Da für meine Fragestellung besonders die Segmentierungsanalyse interessant ist, möchte ich kurz wichtige Aspekte aufführen. Dinkelaker (2018, S. 150-154) nennt vier Ausgangspunkte der Segmentierungsanalyse, nämlich erstens die Konfigurationen der Positionierung der Körper der Beteiligten im Raum, zweitens die Konfigurationen der Ausrichtung der Körper im Raum, drittens die Konfigurationen der Blick- bzw. Kopfrichtungen der Beteiligten und viertens das Muster des Sprecherwechsels. Mit den vier Aspekten gelingt es, die Komplexität des Videomaterials zu reduzieren, da sich Forscher\*innen pro Durchgang auf jeweils einen Aspekt beschränken können. Dieses allgemeine Verfahren sagt allerdings noch nichts darüber aus, wie die bewegten Bilder zu beschreiben sind.

Nach Herrle et al. (2013, S. 604) können Videoaufzeichnungen sehr unterschiedlich aufbereitet werden (s. Abb. 1). Diese Aufbereitungen sollen der Forschungsfrage und dem Analyseverfahren angepasst sein.

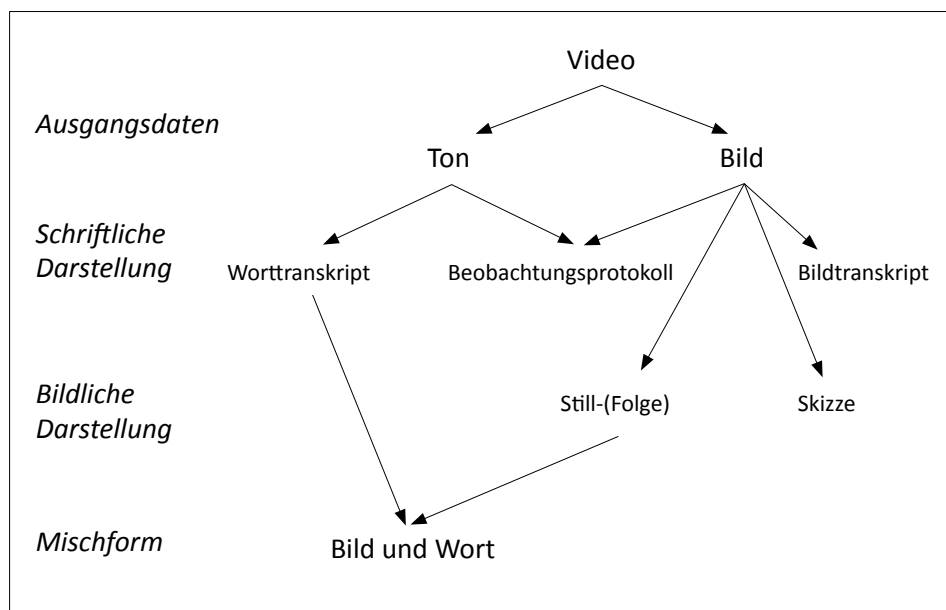


Abb. 1: Aufbereitungsformen für Videomaterial (nach Herrle et al., 2013, S. 605)

Auch wenn dieser Überblick für die grundsätzliche Arbeit mit Videoaufzeichnungen hilfreich ist, so bleibt die Frage nach der Beschreibung offen. Sucht man entsprechende Textstellen in erziehungswissenschaftlichen Veröffentlichungen zur Videografie auf, erhält man eher Probleman-

deutungen als konkrete Hinweise zur Beschreibung: So werde das Gesehene versprachlicht und dabei trete eine Transformation des Wahrgenommenen auf (Dinkelaker & Herrle, 2009, S. 35). Und weiter: „Das Beobachtungsprotokoll ist am stärksten von den Wahrnehmungsschemata und Ausdrucksweisen der Verfasser abhängig; denn damit es erstellt werden kann, müssen Abläufe als relevant ausgewählt und mit Worten charakterisiert werden“ (ebd.). Recht lapidar erläutern Herrle et al. (2013, S. 606), dass Beobachtungsprotokolle in der Regel sowohl die Tonebene als auch die Bildebene berücksichtigten. „Sicht- und hörbare Ereignisse und Abläufe werden unter einem bestimmten Beobachtungsfokus in Worte gefasst.“ Zur Beschreibung von Gestik verweisen sie dann auf Sager (2005), der aus kommunikationstheoretischer Perspektive ein Beschreibungssystem entwickelt hat (s. Tab. 3).

Tab. 3: Beschreibungssystem für Gesten nach Sager (2005) (Auszug)

Schultergelenk: 1. Freiheitsgrad	Elevation Delevation	Angehoben Gesenkt
Schultergelenk: 2. Freiheitsgrad	Abduktion Adduktion	Nach vorn gespreizt Nach hinten gezogen
Handgelenk: 1. Freiheitsgrad	Flexion Extension	Gebeugt Gestreckt
Handgelenk: 2. Freiheitsgrad	Radiale Abduktion Ulnare Abduktion	Gespreizt (der Daumen zeigt zum Unterarm) Gespreizt (der kleine Finger zeigt zum Unterarm)

Das Beschreibungssystem nun wirkt sehr „technisch“, da es vom Sinngehalt der jeweiligen Gestik abstrahiert. Die dokumentarische Methode würde hier wahrscheinlich von Kinemen sprechen, die auf der vorikonografischen Ebene angesiedelt sind, aber noch keine Beschreibung sind, die Bedeutungen in der jeweiligen Situation mittransportieren. Ganzkörperbewegungen oder gar sportliche Bewegungen kommen nicht vor. Die Reduktion auf Gestiken und Raumanordnungen zeigen, dass der Körper nicht ganz ausgespart wird, aber überwiegend als Indikator für eine bestimmte soziale Ordnung angesehen wird, wie etwa auch in der Studie zum Klassenmanagement von Ophardt und Thiel (2007) oder den Untersuchungen von Dinkelaker zur Aufmerksamkeit im Unterricht (2015; 2017).

### 3.3 | ETHNOGRAFISCHE VIDEOANALYSE

Grundsätzlich erfordert ethnografische Forschung eine Haltung der Neugier und Unvoreingenommenheit (Breidenstein et al., 2013, S. 7). „Die besondere Leistung der Ethnografie besteht (...) in einer analytischen Beschreibung fremder (oder eigener) Praktiken, mit dem Ziel, diese so zu präsentieren, dass die Leserschaft ein Bild von diesen Praktiken oder kulturellen Lebensformen gewinnen kann“ (ebd.). Besonders diskursive und stumme Wissens- und Praxisformen sind dabei von großem Interesse, sodass gerade auch Bewegungen zum Gegenstand der Ethnografie werden können. Sind also in der ethnografischen Vorgehensweise Antworten auf die Frage dieses Beitrags zu erwarten?



Leider nicht direkt, denn die Ethnografie „ist kein Regelwerk, das für einen bestimmten Datentyp Verfahrensschritte vorschreibt, deren korrekte Befolgung valide wissenschaftliche Aussagen in Aussicht stellt“ (Breidenstein et al., 2013, S. 34; ähnlich schon Geertz, 2019, S. 10). Betont wird allerdings, dass alle erhobenen Daten immer eingebettet sind in eine länger andauernde teilnehmende Beobachtung (ebd., S. 35). Die Methodenfrage wird in der Ethnografie daher gedeutet als eine Frage nach der kommunikativen Kompetenz der Forscher\*innen (ebd., S. 37). Gleichwohl fehlen methodologische Hinweise für die Ethnografie nicht gänzlich. So geben Breidenstein et al. (2013, S. 37-44) Grundsätze an, die ethnografische Forschung leiten: Neben der teilnehmenden Beobachtung sind Schritte nötig, die die Beobachtungskompetenz erhöhen, die während der Teilnahme an sozialen Situationen vom Handlungsdruck entlasten, um beobachten zu können. Während des Aufenthalts im Feld muss man laufend Verschriftlichungen anfertigen und sich zur Analyse Distanz zum Feld verschaffen. Während Videoaufzeichnungen nach Breidenstein et al. (2013, S. 91) detaillierte Beobachtungen von Handlungsabläufen zulassen, raten die Autoren zugleich dazu, den Aufwand und den Nutzen von Transkriptionen im Auge zu behalten. „Es gilt nur so genau zu transkribieren, wie man nachher auch analysieren will“ (ebd.). Da Videoaufzeichnungen das Geschehen konservieren, könne man später immer noch mehr ins Detail gehen, wenn es sich als nötig herausstellen sollte. Selbstverständlich betrachtet auch die Ethnografie den Schritt der Verschriftlichung schon als interpretativen Akt.

Die ethnografische Vorgehensweise unterscheidet zudem zwei Textsorten im Forschungsprozess: die Feldnotizen, die während oder kurz nach der teilnehmenden Beobachtung geschrieben werden, und die Protokolle. Feldnotizen werden schnell und ohne Rücksicht auf weitere Gestaltungsprinzipien oder sprachliche Richtigkeit verfasst, da sie hauptsächlich der „Speicherung“ von flüchtigen Erfahrungen dienen. Protokolle sind demgegenüber Beschreibungen, die schon mit Blick auf eine Explikation geschrieben werden (ebd., S. 97). „Die Explikation steht dagegen für wichtige analytische Funktionen des Schreibens. Das beginnt mit der Verlangsamung, nicht nur durch Schreib- und Lesetempo, sondern auch durch Detaillierungszwänge. Erst wenn man etwas aufschreibt, merkt man, wie viel man sagen muss, damit ein anonymer Leser folgen kann. So wie der Erzähler durch die Reaktionen eines konkreten Zuhörers zur Vervollständigung von Details angehalten wird, so muss sich der einsame Schreiber durch die Imagination eines anonymen oder beliebigen Lesers selbst zur Explikation von Details anhalten und wortweise feststellen, was es ist, das hier gerade geschieht“ (ebd.). Eine Beschreibung sei daher kein Geniestreich, sondern eine handwerkliche Leistung, die sich allmählich entwickle. Generell zeichnet sich die ethnografische Haltung dadurch aus, dass sie die kommunikative Funktion von Beschreibungen betont, dass Beschreibungen also für ein bestimmtes Publikum verfasst werden. Man könnte noch weiter gehen und sagen, dass Beschreibungen im ethnografischen Sinne immer etwas zeigen wollen, indem sie Einzelfälle als Fälle oder Illustrationen von etwas präsentieren. Knoblauch (2009) macht in seinem Ansatz der Videografie, einer ausdrücklichen Verknüpfung von Ethnografie und Videoanalyse, deutlich, dass dazu auch ein umfangreiches Kontextwissen nötig sei, ohne das beobachtbares Verhalten nicht zu verstehen sei. Reichertz (1992, S. 334) betont, dass Autor\*innen ethnografischer Berichte mit ihren Texten ihr Publikum überzeugen müssen. Dabei ist es gewollt, dass eine Art „Simulation

von Erfahrungsqualitäten erreicht wird, die neben dem Verstand auch die Sinne der Leserin mit anspricht“ (Breidenstein et al., 2013, S. 105). Daher gibt es in dieser Forschungsrichtung durchaus die Auffassung, dass Bilder zwar einen eigenen Charakter haben, aber in sprachliche *Bedeutung* umgemünzt werden können (Reichert & Englert, 2011, S. 22). Es sei weniger die Frage, ob visuelle Daten verlustfrei in sprachliche Daten umgewandelt werden können, sondern vielmehr, ob man eine beobachtbare Handlung *als* eine sozial typisierte Handlung erkennen könne. Reichert und Englert (2011) halten sich daher in ihrer „Einführung in die qualitative Videoanalyse“ nicht damit auf, sich über die Wortwahl in Beschreibungen Gedanken zu machen. Ist die Ethnografie für die Frage dieses Beitrags vielleicht die falsche Adresse? Ein Blick auf ältere ethnografische Literatur kann allerdings etwas mehr Aufschluss geben.

Geertz (2019) prägte im Rahmen der Ethnografie den Begriff der „dichten Beschreibung“. Am Beispiel des Zwinkerns verdeutlicht er, was er damit meint: Die Augenbewegung kann ein ungewolltes Zucken sein, sie kann aber auch gemeint sein als einvernehmliches Zuzwinkern zwischen Freunden in der Öffentlichkeit. Als Bewegung ist beides identisch – eine bildliche Darstellung könne beides nicht unterscheiden (ebd., S. 10). Während eine Verschriftlichung des rein äußerlich sichtbaren Vorgangs keinen Unterschied zwischen Zucken und Zwinkern machen würde (Geertz spricht hier von einer „dünnen“ Beschreibung) schließe die dichte Beschreibung die Bedeutung der Bewegung mit ein (ebd., S. 11-12). Eine gute Beschreibung lasse sich daher von einer schlechten unterscheiden, ob man ihrer Hilfe zwischen Zucken und Zwinkern unterscheiden könne (ebd., S. 24).

Lüders (1995, S. 330) verweist darauf, dass ethnografische Texte keine klare Trennung zwischen Wissenschaft, Literatur und Alltag vornähmen. Das zugrunde liegende Wissenschaftsverständnis lasse zu, dass die Grenzen zwischen Realität und Fiktion, Literatur und Wissenschaft nicht mehr eindeutig gezogen werden. Die schlichte Idee, man könne Wirklichkeit abbilden, wenn man nur die richtigen Methoden fände, sei nicht mehr haltbar. Lüders (1995, S. 331) bezeichnet Ethnografien dann auch als „true fictions“. Breidenstein et al. (2013, S. 101) warnen allerdings davor, sich nur von narrativen Konventionen leiten zu lassen, denn die soziale Wirklichkeit sei oft weniger strukturiert und zufälliger als dramatische Geschichten. Statt einer technologischen Lösung für das Beschreibungsproblem setzt die Ethnografie also auf einen kommunikativen Prozess, der zwischen Forscher\*in und dem Publikum durch Beschreibungen ausgelöst wird. Einerseits kann man von dieser Lösung enttäuscht sein, andererseits eröffnet sie kreative Möglichkeiten des qualitativen Forschens, die dem Gegenstand angemessen sind.

### 3.4 | BEWERTUNG DER FORSCHUNGSMETHODISCHEN ANGEBOTE

Ich möchte hier knapp sowohl Erträge der forschungsmethodischen Angebote zusammenfassen als auch auf Lücken hinweisen, die nach wie vor für die Beschreibung von bewegten Bildern auszumachen sind.

Die dokumentarische Methode sucht Anschluss an kunsthistorische Begriffe und Stufen, indem sie mit vorikonografischen, ikonografischen und ikonologischen Analysen arbeitet. Auf der vorikonografischen Stufe geht es um die genaue sprachliche Fassung des Bild- und Videomaterials.

Ebenso wie in der wissenssoziologischen Videografie nutzt man Standbilder als Ausgangspunkt. Der Übergang von rein äußerlich, „technisch“ anmutenden Bewegungsbeschreibungen auf der vorikonografischen Stufe zur bedeutungsbezogenen Darstellung wird bei Bohnsack (2011) genauer betrachtet. Indem Um-zu-Motive unterstellt werden, kann auch die Bedeutung von Bewegungen geschildert werden, die immer im Kontext der Gesamtsituation zu sehen ist. In dieser methodischen Ausrichtung wird das Beschreibungsproblem also reflexiv behandelt, wenngleich es nach Bohnsack (2011) bisher nicht gelöst werden kann, da eine „Beschreibungssprache“ fehle.

In methodischen Veröffentlichungen zur erziehungswissenschaftlichen Videografie findet man hilfreiche Hinweise, wie Videomaterial überhaupt zur weiteren Bearbeitung aufbereitet werden kann (s. Abb. 1). Weiterhin wird der Blick geschärft für Körper-Raum-Konstellationen, die als Anzeiger für soziale Ordnungen interpretiert werden. Allerdings wird Bewegung kaum zum Thema gemacht, sodass hier das Beschreibungsproblem weitgehend ausgeblendet bleibt.

Die Ethnografie arbeitet durchgehend mit Beschreibungen, gerade auch solcher sozialen Phänomene, die sich unterhalb der sprachlichen Schwelle abspielen. Ethnografische Hinweise zum Beschreiben betonen besonders, dass Schreiben immer ein interpretativer Akt ist und dass er dem Zeigen dient. Insofern bleibt das Urteil über die Güte einer Beschreibung in den Händen der Rezipienten.

Aus der Sicht der qualitativen Forschung in der Sportpädagogik lassen sich in den gesichteten methodischen Angeboten einige Lücken identifizieren. Die Idee, Standbilder statt bewegter Bilder zum Ausgangspunkt zu nehmen, hat sicherlich seine nachvollziehbaren Gründe. Jedoch erkaufte man sich die Vorteile damit, dass die Dynamik von Bewegungen verloren geht (Corsten, 2010, S. 16; Singh, 2018, S. 204-205). Alle hier behandelten Ansätze beziehen sich auf Bewegung nur in dem Sinne, dass sie als Indikatoren für bestimmte soziale Prozesse und Ordnungen gedeutet werden. Der autotelische Charakter von sportlichen Bewegungen lässt jedoch nicht immer eine Konstruktion eines „Um-zu-Motivs“ zu. Ferner ist Bewegung in den oben dargestellten Methodologien prinzipiell nicht als (Unterrichts-)Gegenstand und damit als Forschungsfokus gedacht, sodass beispielsweise solche für den Sport und die Vermittlung von Sport zentralen Vorgänge wie Bewegungslernen und das Üben und Trainieren mit den jeweiligen qualitativen und quantitativen Veränderungen in der Bewegung nicht thematisiert werden. Allen Ansätzen ist außerdem gemeinsam, dass sie das Problem der Versprachlichung durchaus benennen, aber nicht lösen können. Insofern gilt auch heute noch, was Wolters (2005, S. 36) als Fazit eines Aufsatzes festgehalten hat: „Was mit welchen Worten wie detailliert dargestellt wird, hängt maßgeblich von der Forschungsfrage und dem Wissenschaftsverständnis ab, die der jeweiligen Untersuchung zugrunde liegen.“ Sich darüber bewusst zu sein, kann durchaus ein methodischer Gewinn sein, denn Beschreibungen von bewegten Bildern sind kein Selbstzweck, sondern stehen im Dienst des Erkenntnisinteresses. Im letzten Schritt dieses Aufsatzes soll nun anhand von drei Beispielen gezeigt werden, wie aktuelle videografische Studien, die sich mit der Vermittlung von Sport und Bewegung beschäftigen, das Beschreibungsproblem bewältigen.

## 4 | DAS BESCHREIBUNGSPROBLEM IN AKTUELLEN VIDEOGRAFISCHEN ARBEITEN

Die Vermittlung von Bewegungen ist eine zentrale Praxis in sportpädagogischen Handlungsfeldern. Somit muss sich auch die sportpädagogische Forschung mit dem Gegenstand der Bewegung auseinandersetzen. Es ist jedoch durchaus umstritten, was genau unter „Bewegung“ und zumal „sportlicher Bewegung“ verstanden werden soll (z. B. Prohl, 1999, S. 369). Laging (2013) macht sich auf die Suche nach dem fachlichen Gegenstand des Sportunterrichts, wobei er die Auseinandersetzung der Sportpädagogik mit dem Bewegungsbegriff nachzeichnet. Er analysiert, dass nach dem Paradigmenwechsel von der Theorie der Leibeserziehung zur Sportpädagogik das Gegenstandsverständnis theoretisch nicht mehr angemessen diskutiert worden sei, sondern eher der Macht des Faktischen – des organisierten Sports – überlassen wurde (ebd., S. 66). Laging selbst plädiert für eine bildungstheoretische Fundierung von Bewegung, die als Fokus für eine Gegenstandsbestimmung der Sportpädagogik taugt (ebd., S. 74-78). Ausgangspunkt hierbei solle insbesondere das Bewegungshandeln sein. Lagings Auffassung ist durchaus geeignet, um sie zum Ansatz für videografische Studien zu machen – wie ja auch durch Hartmann (2019) geschehen. Allerdings vertritt Laging eben nur eine Position innerhalb der Sportpädagogik, sodass man nicht von einem Konsens über den Gegenstand der Sportpädagogik oder über den Bewegungsbegriff ausgehen kann. Tatsächlich ist eher festzustellen, dass empirische sportpädagogische Untersuchungen oftmals ganz andere Aspekte in den Vordergrund rücken als Bewegungshandlungen im Sinne von Unterrichtsgegenständen. Insofern beruhen die im Folgenden dargestellten Werke – Singh (2016), Wolff (2017) und Hartmann (2019) – auf sehr unterschiedlichen theoretischen Voraussetzungen. Dass sie damit auch zu unterschiedlichen Lösungen für die Versprachlichung bewegter Bilder kommen, ist verständlich.

Singh (2018) stellt anhand seiner videografischen Forschungsarbeiten dar, wie räumliche Bewegungsabläufe von Trampolinturner\*innen und die oftmals simultanen verbalen Äußerungen von Trainer\*innen dargestellt werden können. Die Studien sind ethnomethodologisch angelegt und sollen die performativen Ordnungen in der Trainer\*in-Athlet\*in-Interaktion im (Nachwuchs-)Leistungssport rekonstruieren, wobei dies am Beispiel der Trainingskommunikation im Trampolinturnen geschieht (Singh, 2016). Singh geht es um die habitualisierten und impliziten körperlichen Wissensbestände, die im Nachwuchstraining kommunikativ hergestellt, expliziert und durch die Sportler\*innen in einem langen Übungsprozess angeeignet werden.

Nach Singh (2018, S. 199-200) durchlaufen bewegte Bilder in Videos bis zur Darstellung der Forschungsergebnisse verschiedene Transformationsprozesse, bei denen sicherzustellen ist, dass der Gehalt der Bilder nicht durch die Versprachlichung „verloren geht“. Während des Forschungsprozesses stellte Singh fest, dass Verbaltranskripte mit gelegentlichen Standbildern unzureichend für seine Fragestellung waren. Da die Arbeit mit Standbildern sich als nicht befriedigend erwies, weil die zeitlich-räumliche Verschränkung der Bewegungen der Athletin und der verbalen Äußerungen der Trainerin nicht dargestellt werden konnten, entwickelte Singh (2016) folgende Notationsweise (s. Abb. 2).

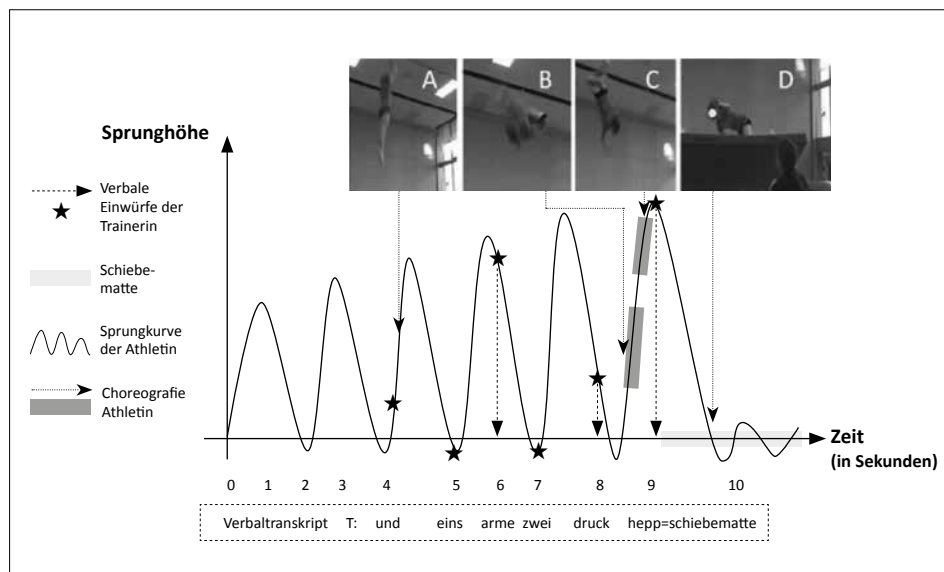


Abb. 2: Visualisierung und Transkript einer Trainingssituation (Singh, 2018, S. 209)

„Die an einem spezifischen Gegenstand entwickelte Darstellung dient demzufolge der Sichtbarmachung der hervorgebrachten Handlungskordinierungen, die im Video wie auch anhand konventioneller Transkripte nur schwer rekonstruierbar sind“ (Singh, 2018, S. 212). Durch diese Art der Darstellung umgeht Singh allerdings das Problem der Versprachlichung von Bewegung. Die jeweiligen Sprünge der Athletin sind mit Landung und Höhepunkt als Kurve verbildlicht, während die Ausführung der Bewegungen in der Luft über die Standbilder repräsentiert wird. Singh beschreibt die analysierten Bewegungen im Transkript gar nicht mit Worten. Erst bei der Interpretation finden wir folgende Beschreibung der obigen Szene: „... wird nun vor allem die Phase betrachtet, die sich zwischen den Sekunden acht bis zehn ereignet. (...) Im Verlauf des zweiten Saltos, genauer gesagt, als sie den ersten gerade beendet hat und in den zweiten übergeht – der Körper ist zu diesem Zeitpunkt nahezu waagerecht in der Luft – kommt von der Trainerin das Signal „hepp“ (roter Stern bei neun Sekunden). (...) Nun wird jedoch ersichtlich, dass etwas nicht stimmt: Die Athletin scheint die Kontrolle über ihren Körper zu verlieren. Ihre Beine schwingen haltlos, d. h. ohne sichtbar intendiert ‚gesteuert‘ zu werden, durch die Luft, während die Arme vom Körper abstehen und unkontrolliert umherrudern (Bild C)“ (Singh, 2018, S. 210).

Inwieweit kann man aus Singhs Vorgehensweise etwas für sportpädagogische Fragestellungen lernen? Die wichtigste Konsequenz scheint mir, dass an diesem Forschungsbeispiel sehr deutlich wird, dass audiovisuelle Daten nicht vorwiegend standardisiert zu versprachlichen, sondern diese im Sinne eines „unique-adequacy-requirement“ (Garfinkel & Wieder, 1992) reflexiv an die Forschungskontexte zurückzubinden sind. Mit der visualisierenden Darstellung hat er seinen Gegenstand so dokumentiert, dass eine hohe Nachvollziehbarkeit gegeben ist.

Trotz der innovativen Notation wird spätestens bei der Interpretation klar, dass auch hier Bewegung in Sprache gefasst werden muss. An den Adjektiven und Adverbien (z. B. „haltlos“, „unkontrolliert“) erkennt man, dass die Bewegung schon bewertet wird vor dem Hintergrund eines (im Wettkampfsinne) gelungenen Sprungs. Im Prinzip erfasst diese qualitative Bewegungsbeschreibung die Abweichung des Istwerts vom Sollwert. Aufgrund der Fragestellung kann man die Beschreibung sicherlich als angemessen ansehen; jedoch verwendet Singh im Vergleich zur sehr elaborierten Darstellung und Begründung seines Notationssystems überhaupt keine methodischen Überlegungen darauf, wie der Interpretationstext zustande kommt. Wie kommt er zu dem sprachlichen Ausdruck: „Ihre Beine schwingen haltlos“? Insofern ist die visualisierende Notation ein Zwischenschritt, aber doch keine Lösung des Problems, wie Bewegung in Sprache zu fassen ist. Zudem bezieht sich Singh auf einen sehr speziellen Kontext, der durch eine 1:1-Betreuung einer Athletin durch eine Trainerin gekennzeichnet ist und zudem Interaktionen von wenigen Sekunden Länge umfasst. Daher ist Singhs Vorgehen sicherlich inspirierend, aber kaum übertragbar auf die typischen Bedingungen, unter denen etwa Sportunterricht stattfindet.

In jüngerer Zeit hat Wolff sich mit der sozialen Ordnung im Sportunterricht (2017) beschäftigt. Wolff gibt seiner Dissertation den Untertitel „Eine Praxeographie“ und verdeutlicht damit seinen theoretischen und methodologischen Hintergrund. Die Arbeit zielt auf eine scheinbar einfache Frage ab, nämlich wie soziale Mikroordnungen im Sportunterricht entstehen und wie Sportunterricht von den Akteuren am Laufen gehalten wird (Wolff, 2017, S. 8). Mithilfe von 24 Videograferten Stunden rekonstruiert der Autor über die Auswahl verdichteter Interaktionssituationen und die Bildung von Kategorien einen „Spielplan“ für Sportstunden.

Der von Lehrpersonen wie von Schülerinnen und Schülern hergestellte Ablauf ordnet sich im Spielplan in die vier Kategorien Koordination, Konstitution, Spiel und Wandlung (ebd., S. 168-169). Deutlich wird, dass das Wie? der Unterrichtskonstruktion im Fokus steht, weniger dagegen die Unterrichtsinhalte oder Lernprozesse der Schüler\*innen. Wie löst Wolff nun das mich interessierende Problem? Der Autor macht sich Gedanken über das Verhältnis von Bildern und Sprache (ebd., S. 135). Dabei formuliert er den Einwand, dass Bilder schon eine andere Wirklichkeit konstruieren als eine laufende Videoaufzeichnung; die Verschriftlichung wiederum stehe für eine weitere Repräsentation der Realität: „Im Video vorhandene Körper<sup>3</sup> und Materialitäten können nicht mit einer Syntax oder einer Semantik deskriptiv übertragen werden, da andere Bedeutungszuweisungen hergestellt werden“ (Wolff, 2017, S. 135).

Der Autor schlägt vor, „mit dem sich offenbarenden Sinn des Videoausschnitts zu arbeiten“ (ebd.) und ordnet sich damit in ethnografische Argumentationen ein (s. Kap. 3.3). Die verwendeten Unterrichtsszenen stellt Wolff mithilfe eines vereinfachten konversationsanalytischen Notationssystems nach Hecht (2009) dar, das er durch Standbilder und Anmerkungen zum

---

3 Auffällig ist, dass der Autor hier von Körpern – und nicht von Bewegungen oder sich bewegenden Akteuren – spricht, was wohl seiner praxeologischen Sichtweise zu verdanken ist.

Handlungskontext und damit auch durch Bewegungsbeschreibungen ergänzt. Folgend ein Beispiel:

Tab. 4: Auszug eines Transkripts zur Kategorie Körper-Ordnungen (Wolff, 2017, S. 220)

	00:00	00:02	00:08
Lm v		(stöhnt)	
Lm nv	Geht zu seinen Unterlagen, die auf dem Boden liegen, und setzt sich im Schneidersitz neben seine Unterlagen auf den Boden.	Schiebt seine Unterlagen vor sich und schaut in die Halle, anschließend legt er seine Arme auf seinen Knien ab und macht eine kreisende Handbewegung mit beiden Händen. Dabei blickt er zu einer SchülerInnengruppe am Hallenrand.	Blickt weiterhin in die Halle und zur SchülerInnengruppe.
Sx v	(Gespräche)		
Sx nv	Laufen in der Halle umher, spielen mit einem Ball oder stehen in Kleingruppen und unterhalten sich.	Erste SchülerInnen orientieren sich zu Lm.	Die SchülerInnengruppe schaut zu Lm und geht in dessen Richtung. Gleichzeitig kommen weitere SchülerInnen von der rechten Seite zu Lm und beginnen, sich hinzusetzen.

Für die Akteure des Sportunterrichts unterscheidet Wolff jeweils die verbale und die nonverbale Ebene (s. linke Spalte). Oftmals zeigt er zudem die korrespondierenden Standbilder zu dem Zeitpunkt der Videoaufnahme (s. erste Zeile). So wird das Problem der simultanen Struktur der bewegten Bilder gelöst. Lassen Sie uns kurz ein Beispiel aus Tab. 4 betrachten: „Erste SchülerInnen orientieren sich zu Lm.“ Was tun diese Schüler und Schülerinnen genau? Blicken Sie zur Lehrkraft? Wenden sie sich ihm mit ihrer Körperachse zu? Machen Sie ein paar Schritte in seine Richtung? Auf der Ebene des Beobachtbaren verbleibt „sich orientieren“ nicht, sodass schon eine zusammenfassend-interpretierende Sicht in diese Darstellung eingeflossen ist. Der Detaillierungsgrad ist hier deutlich geringer als in Singhs Studien, allein deswegen schon, weil sehr viele Sportstunden analysiert werden. Das ist für die Fragestellung, die sich für das „Funktionieren“ von Sportunterricht als sozialer Interaktion interessiert, angemessen – für die konkrete Beschreibung von Bewegungshandlungen liefert Wolffs Vorgehen jedoch keine „Blaupause“.

Hartmanns Arbeit (2019) trägt den Titel „Bewegungsaufgaben im Horizont von Lernenden“ und lässt damit einen deutlichen fachdidaktischen Hintergrund erkennen. Die Autorin will „Aneignungslogiken von Schülerinnen und Schülern im Umgang mit einem Lerngegenstand des Sportunterrichts“ untersuchen (Hartmann, 2019, S. 9). Dabei setzt sie sich kritisch mit dem Kompetenzdiskurs auseinander und fundiert ihre empirische Studie sowohl bewegungs- als auch bildungstheoretisch. Wie sie zu Recht anmerkt, findet sich dieser Forschungsfokus erstaunlicherweise kaum in

der Sportpädagogik. Methodisch nähert sich Hartmann vor allem videografisch und mit Interviews ihrem Gegenstand (Hartmann, 2019, Kap. 3), sodass man von einer Kombination von Innen- und Außersicht der Bewegung sprechen könnte. Anhand von acht Fällen rekonstruiert Hartmann die „Bewegungshabits“ bei der Auseinandersetzung der Schüler\*innen einer elften Jahrgangsstufe mit dem Springen vom Minitrampolin. Für den vorliegenden Beitrag betrachte ich nun, wie Hartmann das Problem der Versprachlichung von Videomaterial gelöst hat. Hartmann schließt an die dokumentarische Bild- und Videointerpretation an (ebd., S. 118-124), deren Grundzüge sie darstellt, um spezifische Verfahrensweisen für ihr Projekt zu entwickeln. Um die Bewegungen der Videografierten Schüler\*innen zu beschreiben, erwägt sie, Vorgehensweisen der Bewegungsdiagnostik zu übernehmen, z. B. die von Meinel (1972), von Laban (1995) oder von Sherborne (1998). Meinels morphologischer Ansatz lässt sich durchaus analog zu der vorikonografischen Darstellung im Sinn der dokumentarischen Methode verstehen (Hartmann, 2019, S. 130) und ist daher für die Studie relevant, während Labans Notationssystem als zu aufwendig und Sherborns Bewegungsdiagnose zu spezifisch auf soziales Lernen bezogen eingeschätzt werden (ebd., S. 131). Um sich vor vorschnellen Interpretationen zu schützen, werden erste Analysen nur mithilfe der Videobilder (also ohne Ton) vorgenommen (ebd., S. 133), wobei alle Bewegungsvollzüge einer Person formulierend interpretiert werden. Diese Beschreibungen wiederum werden zu typischen Mustern verdichtet (ebd., S. 135). Ergänzend finden sich dazu aus den Videoaufnahmen hergestellte Bildreihen. Während die textlichen Vorarbeiten nicht umfänglich dokumentiert sind (was vermutlich den Rahmen eines Buchs gesprengt hätte), beschreibt Hartmann (2019, S. 147) die typischen Weise des Minitrampolens von „Thorsten“: „(...) Für sein Vom-Boden-Lösen schwingt Thorsten zudem meistens die Arme vor dem Bodenabsprung von hinten nach vorne durch. Nach dem Lösen zieht er die Beine an, neigt den Oberkörper vor und schwingt seine Arme von hinten nach vorne durch (...)“. Wenn man versucht, diese Beschreibung nachzuvollziehen, hat man die Schwierigkeit, sich zu erklären, wie die Arme sich zwischen erstem und zweitem Satz bewegt haben, denn nach Satz 1 müssten sie sich schon vorne befinden. Anhand dieser einen Textstelle wird klar, dass bei aller methodischer Sorgfalt doch immer wieder Fallstricke in der Umsetzung von bewegten Bildern in Sprache drohen. Insgesamt ist es Hartmanns Verdienst, das Verfahren der dokumentarischen Methode auf eine sportpädagogische Fragestellung angewendet und dabei eine Verknüpfung mit bewegungstheoretischen Grundlagen hergestellt zu haben. Was die Verschriftung von Bewegung angeht, sind jedoch auf der Beispielebene keine so detaillierten vorikonografischen Beschreibungen zu finden, dass man daraus eine genuin sportpädagogische Adaption der dokumentarischen Methode erkennen könnte.

## 5 | FAZIT

Mit der Verbreitung von videografischen Arbeiten steigt der Bedarf an methodischem Know-how, um bewegte Bilder in Sprache fassen zu können. Denn trotz der „visuellen Revolution“ (Schnettler, 2007) wird voraussichtlich Wissenschaft weiterhin im Modus der rationalen Argumentation und damit der Sprache funktionieren. Tatsächlich bemühen sich auch verschiedene sozialwissenschaftliche Methodologien, das Problem der Beschreibung von sich bewegenden Menschen in Videoaufnahmen einzugrenzen und zu lösen. Im vorliegenden Text wurden die Ansätze der dokumentarischen Methode (Kap. 3.1), der erziehungswissenschaftlichen Videografie (Kap. 3.2) und



der ethnografischen Videoanalyse (Kap. 3.3) vorgestellt. Nach wie vor bestehen jedoch Lücken in den oftmals sonst sehr elaborierten Forschungsmethodiken, was die Erforschung von sportlichen Bewegungen und damit verknüpften sportpädagogischen Fragestellungen angeht (s. Kap. 3.4). Daher wurde geprüft, wie neuere videografische Studien (Singh, 2016; Wolff, 2017; Hartmann, 2019) das Beschreibungsproblem lösen. Es lässt sich erwartungsgemäß nicht „die“ Lösung finden, sondern je nach Fokus werden eigene Darstellungsformen und Beschreibungspraktiken entwickelt. Die in Kap. 4 dargestellten Beispiele verwenden alle neben der sprachlichen Fassung des Geschehens auch Bilder in verschiedener Form, sodass eine Rückbindung an das ursprüngliche empirische Material gegeben ist. Wie im ethnografischen Ansatz diskutiert, ist jedoch eine standardisierte Vorgehensweise nicht möglich, sondern es ist vielmehr ein kommunikativer Prozess zwischen Autor\*innen solcher Studien und den Rezipient\*innen der daraus entstandenen Texte, der über die Qualität von sprachlichen Darstellungen bewegter Bilder entscheidet.

## LITERATUR

- Asbrand, B., & Martens, M. (2018). *Dokumentarische Unterrichtsforschung*. Springer VS.
- Bergmann, J. R. (1985). Flüchtigkeit und methodische Fixierung sozialer Wirklichkeit. In W. Bonß & H. Hartmann (Hrsg.), *Entzauberte Wissenschaft* (= Soziale Welt, Sonderband 3) (S. 299-230). Otto Schwartz & Co.
- Bergmann, J. R. (1991). Deskriptive Praktiken als Gegenstand und Methode der Ethnomethodologie. In M. Herzog & C. F. Graumann (Hrsg.), *Sinn und Erfahrung. Phänomenologische Methoden in den Humanwissenschaften* (S. 86-102). Asanger.
- Boehm, G. (2007). *Wie Bilder Sinn erzeugen*. Berlin University Press.
- Bohnsack, R. (2010). Zugänge zur Eigenlogik des Visuellen und die dokumentarische Videointerpretation. In M. Corsten, M. Krug & C. Moritz (Hrsg.), *Videographie praktizieren. Herangehensweisen, Möglichkeiten und Grenzen* (S. 271-298). Springer VS.
- Bohnsack, R. (2011). *Qualitative Bild- und Videointerpretation* (2. Aufl.). Barbara Budrich.
- Bohnsack, R. (2013). Gruppendiskussionsverfahren und dokumentarische Methode. In B. Friebertshäuser, A. Langer & A. Prengel (Hrsg.), *Qualitative Forschungsmethoden in der Erziehungswissenschaft* (4. Aufl.) (S. 205-218). Beltz Juventa.
- Bohnsack, R. (2014). *Rekonstruktive Sozialforschung* (9. Aufl.). Barbara Budrich.
- Bohnsack, R., Fritzsche, B., & Wagner-Willi, M. (2015). Dokumentarische Video- und Filminterpretation. In R. Bohnsack, B. Fritzsche & M. Wagner-Willi, M. (Hrsg.), *Dokumentarische Video- und Filminterpretation. Methodologie und Forschungspraxis* (2. Aufl.) (S. 11-41). Barbara Budrich.
- Breidenstein, G., Hirschauer, S., Kalthoff, H., & Nieswand, B. (2013). *Ethnografie. Die Praxis der Feldforschung*. UVK Verlagsgesellschaft.
- Collier, J., & Collier, M. (1986). *Visual Anthropology: Photography as a Research Method*. University of New Mexico Press.
- Corsten, M. (2010). Videographie praktizieren – Ansprüche und Folgen. In M. Corsten, M. Krug & C. Moritz (Hrsg.), *Videographie praktizieren. Herangehensweisen, Möglichkeiten und Grenzen* (S. 7-22). Springer VS.
- Dinkelaker, J. (2015). Varianten der Einbindung von Aufmerksamkeit. In S. Reh, K. Berdelmann & J. Dinkelaker (Hrsg.), *Aufmerksamkeit. Geschichte – Theorie – Empirie* (S. 241-266). Springer VS.
- Dinkelaker, J. (2017). Aufmerksamkeit in pädagogischen Situationen. In A. Kraus, J. Budde, M. Hietzge & C. Wulf (Hrsg.), *Handbuch Schweigendes Wissen: Erzie-*

- hung, Sozialisation, Bildung und Lernen (S. 380-391). Beltz Juventa.
- Dinkelaker, J. (2018). Reversible Selektivität. Zur videobasierten Analyse pädagogischer Interaktionen. In M. Heinrich & A. Wernet (Hrsg.), *Rekonstruktive Bildungsforschung* (S. 141-158). Springer VS.
- Dinkelaker, J., & Herrle, M. (2009). *Erziehungswissenschaftliche Videographie*. Springer VS.
- Garfinkel, H., & Wieder, L.D. (1992). Two incommensurable asymmetrically alternate technologies of social analysis. In W. Graham & R. Seiler (Eds.), *Text in context: Contributions to ethnomethodology* (pp. 175-206). SAGE.
- Geertz, C. (2019). *Dichte Beschreibung. Beiträge zum Verstehen kultureller Systeme*. Suhrkamp. (Original 1973).
- Hampl, S. (2010). Videos interpretieren und darstellen. Die dokumentarische Methode. In M. Corsten, M. Krug & C. Moritz (Hrsg.), *Videographie praktizieren. Herangehensweisen, Möglichkeiten und Grenzen* (S. 53-88). Springer VS.
- Hampl, S. (2015). Videotranskription und das System MoViQ. In R. Bohnsack, B. Fritzsche & M. Wagner-Willi, M. (Hrsg.), *Dokumentarische Video- und Filminterpretation. Methodologie und Forschungspraxis* (2. Aufl.) (S. 441-463). Barbara Budrich.
- Hartmann, M. (2019). *Bewegungsaufgaben im Horizont von Lernenden: Eine rekonstruktive Fallstudie zu Orientierungsmustern beim Bewegungslernen im Sportunterricht*. Schneider Verlag.
- Hecht, M. (2009). *Selbsttätigkeit im Unterricht: Empirische Untersuchungen in Deutschland und Kanada zur Paradoxie pädagogischen Handelns*. Springer VS.
- Herrle, M., Kade, J., & Nolda, S. (2013). Erziehungswissenschaftliche Videographie. In B. Friebertshäuser, A. Langer & A. Prengel (Hrsg.), *Qualitative Forschungsmethoden in der Erziehungswissenschaft* (4. Aufl.) (S. 599-619). Beltz Juventa.
- Imdahl, M. (1996). *Giotto – Arenafresken. Ikonographie – Ikonologie – Ikonik*. Fink.
- Knoblauch, H. (2004). Die Video-Interaktions-Analyse. *sozialersinn*, 5(1), 123-138.
- Knoblauch, H. (2009). Videography. Focus Ethnography and Video Analysis. In H. Knoblauch, B. Schnettler, J. Raab & H.-G. Soeffner (Eds.), *Video Analysis: Methodology and Methods* (2nd edition) (pp. 69-83). Peter Lang.
- Knoblauch, H., Schnettler, B., & Raab, J. (2009). Video-Analysis. Methodological Aspects of Interpretive Audiovisual Analysis in Social Research. In H. Knoblauch, B. Schnettler, J. Raab & H.-G. Soeffner (Eds.), *Video Analysis: Methodology and Methods* (2nd edition) (pp. 9-26). Peter Lang.
- Laban, R. v. (1995). *Kinetographie – Labanotation. Einführung in die Grundbegriffe der Bewegungs- und Tanzschrift*. Noetzel.
- Laging, R. (2013). Auf der Suche nach dem fachlichen Gegenstand des Sportunterrichts – Sportpädagogische Reflexion und Perspektive für eine bewegungsorientierte Didaktik. *Zeitschrift für Sportpädagogische Forschung*, 1(2), 61-82.
- Loizos, P. (2000). Video, Film and Photographs as Research Documents. In M.W. Bauer & G. Gaskell (Eds.), *Qualitative Researching with Text, Image and Sound* (pp. 93-107). Sage.
- Lüders, C. (1995). Von der teilnehmenden Beobachtung zur ethnographischen Beschreibung. In E. König & P. Zedler (Hrsg.), *Bilanz qualitativer Forschung* (S. 311-342). Deutscher Studienverlag.
- Mannheim, K. (1980). *Strukturen des Denkens*. Suhrkamp. (Original 1922-1925).
- Meinel, K. (1972). *Bewegungslehre. Versuch einer Theorie der sportlichen Bewegung unter pädagogischem Aspekt*. Volk und Wissen.
- Moritz, C. (2010). Die Feldpartitur. Mikroprozessuale Transkription von Videodaten. In M. Corsten, M. Krug & C. Moritz (Hrsg.), *Videographie praktizieren. Herangehensweisen, Möglichkeiten und Grenzen* (S. 163-193). Springer VS.
- Ophardt, D., & Thiel, F. (2007). Klassenmanagement als professionelle Gestaltungsleistung. Methodische Überlegungen zur Rekonstruktion von Handlungsmustern des Klassenmanagements anhand von Unterrichtsvideographierungen. In D. Lemmermöhle, M. Rothgangel, S. Bögeholz, M. Hasselhorn & R. Water-

- mann (Hrsg.), *professionell lehren – erfolgreich lernen* (S. 135-145). Waxmann.
- Oswald, H. (2013). Was heißt qualitativ forschen? Warnungen, Fehlerquellen, Möglichkeiten. In B. Frieberthshäuser, A. Langer & A. Prengel (Hrsg.), *Qualitative Forschungsmethoden in der Erziehungswissenschaft* (4. Aufl.) (S. 183-201). Beltz Juventa.
- Panofsky, E. (1991a). Zum Problem der Beschreibung und Inhaltsdeutung von Werken der bildenden Kunst. In E. Kaemmerling (Hrsg.), *Ikongraphie und Ikonologie. Theorien – Entwicklung – Probleme* (5. Aufl.) (S. 185-206). Dumont (Original 1932).
- Panofsky, E. (1991b). Ikonographie und Ikonologie. In E. Kaemmerling (Hrsg.), *Ikongraphie und Ikonologie. Theorien – Entwicklung – Probleme* (5. Aufl.) (S. 207-225). Dumont. (Original 1939)
- Prohl, R. (1999). *Grundriss der Sportpädagogik*. Limpert.
- Raab, J. (2008). *Visuelle Wissenssoziologie*. UVK Verlagsgesellschaft.
- Reichertz, J. (1992). Beschreiben oder Zeigen – Über das Verfassen Ethnographischer Berichte. *Soziale Welt*, 43, 331-350.
- Reichertz, J., & Englert, C. J. (2011). *Einführung in die qualitative Videoanalyse*. Springer VS.
- Sager, S.F. (2005). Ein System zur Beschreibung von Gestik. In K. Bührig & S. F. Sager (Hrsg.), *Nonverbale Kommunikation im Gespräch. Osnabrücker Beiträge zur Sprachtheorie*, 70, 19-47.
- Scherler, K. (2002). Zur Mehrfachverwendung von Unterrichtsdokumenten. In G. Friedrich (Hrsg.), *Sportpädagogische Forschung. Konzepte – Ergebnisse – Perspektiven* (S. 65-70). Czwalina.
- Schnettler, B. (2007). Auf dem Weg zu einer Soziologie visuellen Wissens. *sozialersinn*, 8(2), 189-210.
- Sherborne, V. (1998). *Beziehungsorientierte Bewegungspädagogik*. München: Reinhardt.
- Singh, A. (2016). *Wissenskommunikation im Nachwuchstraining. Eine videographische Studie zur kommunikativen Konstruktion von Körperwissen im Trampolinturnen*. Universität Bayreuth.
- Singh, A. (2018). Zur Transkription und Repräsentation von Handlungskoordinationen in Zeit und Raum. Am Beispiel von Wissenskommunikation im Trampolintaining. In C. Moritz & M. Corsten (Hrsg.), *Handbuch Qualitative Videoanalyse* (S. 197-216). Springer VS.
- Wagner-Willi, M. (2005). *Zwischen Vorder- und Hinterbühne. Rituelle Übergangspraxen bei Kindern von der Hofpause zum Unterricht*. Springer VS.
- Wagner-Willi, M. (2007). Videoanalysen des Schulalltags. Die dokumentarische Interpretation schulischer Übergangsrituale. In R. Bohnsack, I. Nentwig-Gesemann & A.-M. Nohl (Hrsg.), *Die dokumentarische Methode und ihre Forschungspraxis* (2. Aufl.) (S. 121-140). Springer VS.
- Wagner-Willi, M. (2009). On the Multidimensional Analysis of Video-Data. Documentary Interpretation of Interaction in Schools. In H. Knoblauch, B. Schnettler, J. Raab & H.-G. Soeffner (Eds.), *Video Analysis: Methodology and Methods* (2nd edition) (pp. 143-153). Peter Lang.
- Wolff, D. (2017). *Soziale Ordnung im Sportunterricht. Eine Praxeographie*. transkript.
- Walters, P. (2005). Vom Video zum Text. *Spectrum der Sportwissenschaften*, 17(1), 24-37.