

6. Unmittelbarkeitserzählung

»Die Mittelbarkeit medialen Gehalts erscheint im Gewande einer Unmittelbarkeit.«

(Krämer 2019, 835)

Als erste Metanarration eint die rekonstruierten Diskurse eine Unmittelbarkeitserzählung, die im Hinblick auf eine in Aussicht gestellte Rezeptionserfahrung sowie in Bezug auf die Inhaltsrealisierung und Übertragung realphysischer Phänomene in das Virtuelle wie folgt in den einzelnen Fächern zum Ausdruck kommt:

Geographie | Im Marketing werden VR-Szenarien als authentische, exakte Abbilder ihrer realphysischen Referenzräume beworben und deren körperlich-leibliche Erfahrbarkeit ins argumentative Zentrum gestellt. Ähnlich verhält es sich mit AR-Navigationsanwendungen, die die Welt in einem Eins-zu-eins-Verhältnis, aber zugleich mit einem individuellen Zuschnitt auf die Nutzer*innenbedürfnisse erfassen sollen. Im Segment ortsungebundener AR-Apps steht die Absicht im Vordergrund, abstrakte geographische Sachverhalte greifbar und ›lebendig‹ erfahrbar zu machen. In Abgrenzung zur passiven Informationsaneignung wird das eigenständige, (inter-)aktive und multisensorische Lernen sowie ein körperlich-leibliches, als grenzenlos ausgelegtes Erkunden der AR- und VR-Inhalte betont. Auch im fachdidaktischen Diskurs steht eine als authentisch und ›direkt‹ umschriebene Erfahrbarkeit von Unterrichtsinhalten im Fokus. Vor allem in Konzepten zum unterrichtlichen Einsatz von Google-Diensten drückt sich – wenn auch nicht ausnahmslos – die Vorstellung einer objektiven, realitätsnahen Abbildhaftigkeit der Welt in 360° aus. Sowohl im Marketing als auch im geographiedidaktischen Diskurs erscheinen Medien der AR und VR primär als neutrale, funktionale Kanäle zur Erfahrung der durch sie hervorgebrachten Inhalte.

Geschichte | Im Marketing wird gemeinhin mit dem Versprechen geworben, mittels AR und VR ein unmittelbares, ›hautnahes‹ (Nach-)Erleben von Geschichte zu ermöglichen. Mit AR könnten kulturelle Artefakte oder Architekturen detailgetreu bzw. originalgetreu nachgebildet und ›in den Raum geholt‹ werden sowie die physi-

sche Präsenz von Zeitzeug*innen derart inszeniert werden, dass die App-Nutzung einer persönlichen Begegnung gleichkäme. Bei VR-Formaten sei es wiederum möglich, ein historisches Ereignis körperlich-leiblich und zuweilen aus der Perspektive einer Person aus jener Zeit nacherleben zu können. Insgesamt werden Medien der AR und VR zu neutralen und ›magischen‹ Informationskanälen erklärt. Im geschichtsdidaktischen Diskurs werden diese Unmittelbarkeitserzählungen und Authentizitätszuschreibungen sowohl reproduziert als auch kritisch kommentiert. Im Zentrum fachdidaktischer Argumentationen steht ein Dekonstruktionsanliegen der Geschichtsdarstellungen als perspektivgebunden, zeitlich und kulturell situiert sowie medienvermittelt. Was genau diese Vermitteltheit bei Medien der AR und VR ausmacht, wird jedoch nicht konkretisiert.

Sport | Im Marketingdiskurs wird die mediale Materialität der VR-Technik als Störfaktor angesehen und das Verschwinden ebendieser als erstrebenswertes entwicklungstechnisches Ziel offensiv beworben. Die Hardware soll leichter, die Datengenerierung präziser und die Handhabung und Schnittstellengestaltung derart ›intuitiv‹ werden, dass sich die Technik unmerklich an die User*innen schmiegt. VR-Anwendungen werden dahingehend vermarktet, dass das VR-System das Bewegungshandeln der Spieler*innen verlustfrei in die virtuelle Umgebung überführe und dadurch ein realitätsnahes Spielerlebnis garantiere (»Play for Real«). Im sportdidaktischen Diskurs bleibt das Marketing einer Eins-zu-eins-Übertragung körperlicher Bewegungsleistung unhinterfragt – im Gegenteil: Das Vorhaben wird als wünschenswert angesehen und dem VR-System eine evidenzproduzierende Objektivität zugeschrieben. Das dahinterliegende, instrumentelle Medienverständnis geht zugleich mit einer werkzeughaft-funktionalistischen Auffassung von Körpern und Körperbewegungen einher. Jenen sportdidaktischen Stimmen, die durch VR bzw. durch digitale Medien allgemein eine künstliche Überformung und Entfremdung ursprünglicher, körperlicher Erfahrungen befürchten, liegt wiederum eine Vorstellung von Unmittelbarkeit zugrunde, die die Medialität von *nicht* digitalen Medien ignoriert.

Unmittelbarkeitserzählungen sind an Medien insofern besonders anschlussfähig, als sich mediale Materialität in ihrem funktionierenden Gebrauch der menschlichen Wahrnehmung entzieht und hinter dem durch sie vermittelten Inhalt zurücktritt:

»Es gibt ein Charakteristikum im Umgang mit Medien, das ›Entzug der Medien in ihrem Vollzug‹ genannt sei: Insoweit Medien in ihrem Gebrauch reibungslos, mithin störungsfrei funktionieren, tendieren sie dazu, etwas zur Erscheinung zu bringen, indem sie selbst sich dabei zurücknehmen und in ihrer eigenen Materialität und Verfasstheit unterhalb der Schwelle des Wahrnehmens verbleiben.« (Krämer 2019, 834)

Diese operative Unsichtbarkeit der Medien (vgl. Burkhardt 2015, 36) ist ein bereits vielfach bearbeiteter medientheoretischer Befund. Weich (2023, 6) versteht ihn »nicht zuletzt [als] Rechtfertigung der Existenz der Medienwissenschaft selbst«, die darum bemüht ist, ebenjene Medienvermitteltheit an die Oberfläche zu bringen (siehe auch S. 28f.). Auch Winkler schreibt in der sechsten Grundthese seiner Mediendefinition: »Medien sind unsichtbar: Je selbstverständlicher wir Medien benutzen, desto mehr haben sie die Tendenz zu verschwinden« (Winkler 2004, 10).

Im Zusammenhang mit Augmented und Virtual Reality muss dieses Unsichtbarwerden der medialen Spuren jedoch nicht nur als unbewusster oder beiläufiger Rezeptionseffekt analytisch an die Oberfläche geholt werden. Stattdessen muss es als explizit ausformulierte Mehrwertzuschreibung für Lernprozesse und damit als Diskursphänomen in den Blick genommen werden. Dass sich eine Unmittelbarkeitserzählung im Zusammenhang mit Medien der Augmented und Virtual Reality derart diskursmächtig zeigt, lässt sich – so die These – wesentlich darin begründen, dass die durch AR und VR hervorgebrachte Bildlichkeit als ein Bild ohne Rahmung diskursiviert und medientechnisch gestaltet wird. Im Folgenden wird ebendieser Bedingungszusammenhang zwischen diskursiven Praktiken und funktionslogischer digitaltechnischer Gestaltung, durch den das Phantasma der »entrahmten Bildlichkeit« hervorgebracht wird, genauer aufgeschlüsselt. Anschließend wird das damit einhergehende Unmerklichwerden der medialen Spuren im Sinne einer kritisch-reflexiven Medienbildung problematisiert.

6.1 Unmittelbarkeit durch Rahmenlosigkeit

Der Rahmen als bildtheoretisches Konzept und optische Bildbegrenzung verweist nicht nur auf die Malerei oder Fotografie, sondern fungiert auch als Leitmetapher im Kontext von Bewegtbildproduktionen. Wie Hasebrink (2024) herausarbeitet, wurde er vermutlich erstmals durch Rudolf Arnheim mit dem Filmbild verknüpft. Arnheim setzte sich 1932 in »Film als Kunst« mit der Notwendigkeit einer filmtechnischen Rahmung der Wirklichkeit durch eine bestimmte Wahl der Perspektive, Größeneinstellung, Brennweite etc. auseinander. Das sichtbar begrenzte Filmbild unterscheide sich damit vom »randlosen« Sehen des menschlichen Auges: »Film ist für Arnheim deshalb niemals direkte Wiedergabe der Wirklichkeit, sondern ihre gezielte ästhetische »Einrahmung« (Hasebrink 2024, 30).¹ Beim televisuellen Bild oder dem Computer wird diese Einrahmung auch als Fenster metaphorisiert – der Verweis auf das Betriebssystem *Microsoft Windows* liegt hier auf der Hand.

1 Zu weiteren filmtheoretischen Ausarbeitungen des Rahmenbegriffs vgl. Hasebrink 2024, 29–33.

In den fachbezogenen Diskursen kommt indes in Sätzen wie: »Und man kann in alle Richtungen gucken. Man kann nach unten gucken, nach oben gucken und man erlebt Geschichte unmittelbar« (aus dem Making-of von ZDF History 360° [ZDF o.J., 0:47-0:53]), keine *Einrahmung*, sondern eine *Entrahmung* als das Fehlen einer optischen Bildbegrenzung zum Ausdruck. Die Kunstaussstellung »Die ungerahmte Welt. Virtuelle Realität als künstlerisches Medium«, die 2017 am Haus der Elektronischen Künste in Basel stattfand, oder das AR-Projekt *The Met Unframed* des Metropolitan Museum of Art von 2021 sind durch die kontrastive Metaphorik der Rahmenlosigkeit im Museumskontext einschlägige Beispiele für diese Vorstellung einer entrahmten Bildlichkeit, die unabhängig von Bildungskontexten im Zusammenhang mit AR und VR Bestand hat (vgl. Przybylka 2022b). Folgende Zitate sollen diesen Befund näher veranschaulichen:

»Ein weiteres Kennzeichen der VR-Metapher ist, dass der Nutzer sich mitten in der simulierten Welt befindet und diese ›von innen‹ erfährt, anstatt wie bei konventionellen PCs die simulierte Welt ›von außen‹ durch ein Fenster zu betrachten.« (Dörner et al. 2019, 17)

»Each of these examples shows you how AR on your phone can be a window to a new world, but what if we wanna transport you to a new world. Well, AR can do that too.« (Barbour; Meta Quest 2020a, 0:45:21-0:45:34)

»I mean, all the media that we watch – television, cinema – they're these windows into these other worlds. [...] But I don't want you in the frame, I don't want you in the window, I want you through the window, I want you on the other side, in the world, inhabiting the world. So that leads me back to Virtual Reality.« (Milk 2015, 04:53-05:10)

Bevor auf die semantische Ausgestaltung dieser zugeschriebenen Rahmenlosigkeit näher eingegangen wird, soll zunächst nachvollzogen werden, wie diese diskursive Entrahmung mit medientechnischen Funktionslogiken von AR und VR ineinandergreift.

6.1.1 Medientechnische Entrahmung

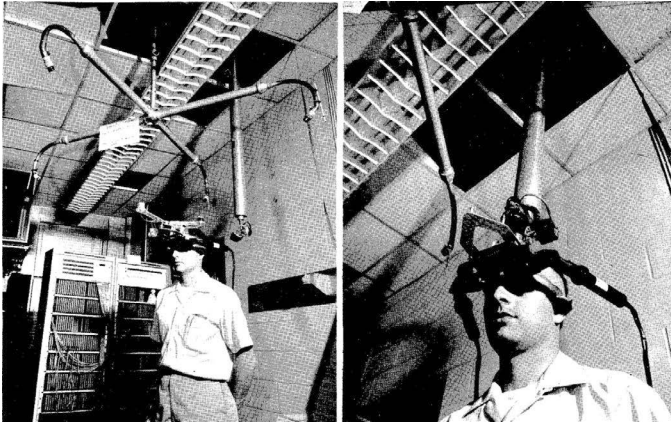
Die Technikgeschichte von Medien der Augmented und Virtual Reality ist von unterschiedlichen medientechnischen und -kulturellen Genealogien durchformt und kann je nach Forschungsinteresse unterschiedlich früh bzw. spät ansetzen. Im Folgenden werden indes zwei entwicklungshistorische Zäsuren fokussiert, mit denen die diskursive Zuschreibung einer Entrahmung insofern plausibilisiert

werden kann, als sie den Grundstein für die ästhetische Verunsichtbarung einer optischen Begrenzung augmentierter und virtueller Bildlichkeiten legten.²

Das *Sword of Damocles* | Im Jahr 1968 legte der amerikanische Informatiker und Computergrafikpionier Ivan E. Sutherland den Grundstein für ein bildtechnisches und bildästhetisches Kernelement von Medien der AR und VR. Sutherland erlangte vor allem mit seiner Arbeit zu Grafical User Interfaces (GUI) und der grafischen Datenverarbeitung Popularität. So stellte er 1962 das Zeichenprogramm *Sketchpad* vor, mit dem mittels eines Lichtgriffels (dem sogenannten Lightpen) erstmals eine direkte Manipulation grafischer Darstellungen auf dem Bildschirm möglich wurde. Das grafische Inputsystem löste als neues Interaktionsparadigma mit dem Computer die bis dato gängige Eingabe von textbasierten Programmanweisungen über Kommandozeilen ab (vgl. Woletz 2016, 47f.). An diese ›Verunmittelbarung‹ der Mensch-Computer-Interaktion anknüpfend legte Sutherland 1968 in »A head-mounted three dimensional display« die Funktionsweise des von seinem Forschungsteam entwickelten, weltweit ersten Head-Mounted Displays dar. Aufgrund der wuchtigen Trackingmechanik zur Erfassung der Kopfbewegung, die bedrohlich über dem Kopf der Proband*innen angebracht wurde, ist es auch unter dem Namen *Sword of Damocles* bekannt (Abb. 47–48).

2 Weiterführende oder auch alternative historische Rückblicke werden etwa bei Lenzen (2020), Schröter (2004), Grau (2001), Braun und Friess (2020) oder Dörner et al. (2019) vorgenommen.

Abb. 47–48: Der mechanische Sensor und Ultraschallsensor des ersten Head-Mounted Displays zur Erfassung der Kopfbewegung und Blickperspektive



Quelle: Sutherland 1968, 760

Die Apparatur konnte noch keine grafisch umfassende Umgebung generieren. Stattdessen wurden über die beiden Displays vor den Augen der Proband*innen dreidimensionale, geometrische, gitterartige Formen in das Sichtfeld projiziert – entweder als im Raum schwebende oder als verankerte, mit physischen Gegenständen in Beziehung stehende Einblendung: »Thus displayed material can be made either to hang disembodied in space or to coincide with maps, desk tops, walls, or the keys of a typewriter« (Sutherland 1968, 759). Obwohl die zweite Variante recht vage umschrieben bleibt und eine optisch stabile Platzierung der grafischen Augmentierung zu jener Zeit noch nicht realisiert werden konnte, sind die dominierenden Darstellungsformen gegenwärtiger AR-Anwendungen, auf die im nachfolgenden Kapitel genauer eingegangen wird, damit bereits 1968 skizziert worden. Die sowohl für spätere AR- als auch VR-Medien ausschlaggebende Innovation war, dass sich die im physischen Raum eingeblendeten Objekte der Kopfbewegung und Blickrichtung der Person unter dem HMD perspektivisch anpassten. Sutherland realisierte mit dem *Sword of Damocles* damit erstmals die Verarbeitung von Bilddaten in »Echtzeit³, die sich um die Rezipient*innen entfalten. Damit stellte er dem ein paar Jahre zuvor durch Sketchpad und Lightpen eingeführten Paradigma der Manipulation von

3 Unter Echtzeit wird zum damaligen Stand der Technik eine durchschnittliche Geschwindigkeit der Datenübertragung von 10 Mikrosekunden pro Zeile verstanden (vgl. Sutherland 1968, 759).

Objekten am Computerbildschirm durch physische Eingabegeräte neue Interaktionsformen zur Seite, die sich durch eine implizite räumlich situierte, körperliche Praxis auszeichneten (vgl. Woletz 2016, 62). Sutherland fasst diese technologische Neuerung wie folgt zusammen:

»The fundamental idea behind the three-dimensional display is to present the user with a perspective image which changes as he moves. [...] The image presented by the three-dimensional display must change in exactly the way that the image of a real object would change for similar motions of the user's head. [...] We can display objects beside the user or behind him which will become visible to him if he turns around. The user is able to move his head three feet off axis in any direction to get a better view of nearby objects. He can turn completely around and can tilt his head up or down thirty or forty degrees.« (Sutherland 1968, 757)

Beim *Sword of Damocles* wird das binokulare Bild der Stereoskopie mit dem raumumgebenden Bild der Panoramamalerei vereint. Im Unterschied zum Panorama, das auch über den fokussierten Bildausschnitt hinaus existiert, synchronisiert sich das Bild im Display mit der entsprechenden Kopfposition und Blickperspektive der Rezipient*innen. Jeder Bildausschnitt ist ein möglicher unter einer Vielzahl anderer Bildausschnitte, die sich um die Rezipient*innen entfalten können. Der dreidimensionale Bildraum ist damit *virtuell* (vgl. Schröter 2004, 191).

Sutherlands HMD gilt als Zäsur in der Entwicklung von Augmented und Virtual Reality, da es den technischen Grundstein für eine fundamentale Funktionslogik beider Technologien legte – nämlich die Synchronisierung der in den Brillendisplays dargestellten Inhalte mit der Körperposition und Blickrichtung der Rezipient*innen im physischen Umgebungsraum. Die blickpunktsensitive Bild(raum)generierung befördert die ästhetische Aufhebung eines Bildrahmens und plausibilisiert damit die besondere Anschlussfähigkeit einer Unmittelbarkeitserzählung an beide Medientechnologien. HMD-basierte Virtual und Augmented Reality lassen Rezipient*innen in alle Richtungen schauen und Objekte im realphysischen Umgebungsraum von allen Seiten betrachten – eine Rahmung als sichtbare Bildgrenze scheint dadurch nicht existent.

»Being able to look anywhere, move freely. It's just a fundamentally different experience from staring at a screen.« (Zuckerberg; Meta 2021, 0:23:31-0:23:37)

»Mit der VR-Brille kann man sich in alle Richtungen umsehen. Ganz anders als beim Fernsehen.« (phoenix 2020, 02:01-02:05)

Bei smartphone- und tabletbasierter AR ist ein Rahmen – nämlich das mobile Endgerät selbst – zwar sichtbar. Nichtsdestotrotz wird auch hier die Aufhebung der Bildrahmung diskursiviert, indem betont wird, dass man sich um virtuelle

Objekte herumbewegen und sie »von allen Seiten« betrachten könne. Zudem wird ihre plastische, dreidimensionale Ästhetik damit umschrieben, als befänden sich die AR-Elemente tatsächlich im physischen Raum: »In diese Umgebung projiziert die App dann ein Interview mit einer Zeitzeugin oder einem Zeitzeugen. Auf dem Bildschirm des Smartphones oder Tablets wirkt es dann so, als befände sich die oder der Interviewte im gleichen Raum« (Planet Schule 2019b). Dieser Eindruck wird vor allem durch ein spezifisches Trackingverfahren technisch perfektioniert, das im Folgenden näher erläutert wird.

Simultaneous Localization and Mapping | Die blickpunktsensitive Bildgenerierung kann mithilfe unterschiedlicher Trackingverfahren realisiert werden. Derzeit ist das Simultaneous Localization and Mapping, kurz SLAM genannt, am verbreitetsten. Mit dem ursprünglich aus der Robotik stammenden Verfahren erstellt ein Roboter sensor- und algorithmenbasiert eine Karte von der ihm unbekannten Umgebung (Mapping) und bestimmt *simultan* seine Position in ebendieser Umgebung (Localization), um darin autonom navigieren zu können. Bei AR und VR wird SLAM basierend auf Kameradaten und der Device-internen Sensorik realisiert. Auch hier »braucht der Bildträger [VR-Headset, Smartphone oder Tablet] eine Karte, um zu lokalisieren, und er muss lokalisieren, um zu kartieren« (Grabbe/van der Veen 2024, 13). Eine sogenannte inertielle Messeinheit (»Inertial Measurement Unit«, kurz IMU) bündelt Gyroskop, Magnetometer und Accelerometer zur Erfassung der Beschleunigung, Orientierung und Winkelraten des digitalen Endgeräts. Für eine akkurate und störungsresistente Positionierung bedarf es über diese interne Lage- und Ausrichtungssensorik hinaus einer kamerabasierten und algorithmisch analysierten Erfassung von sogenannten Feature oder Key Points. Damit sind visuell markante, kontrastreiche Punkte in der kameraerfassten Umgebung gemeint, z.B. Ecken oder Kanten von Fensterrahmen oder Möbelstücken. Diese Feature Points und deren bildanalytisch erfasste Verschiebung von Frame zu Frame⁴ dienen als Referenz für die Bewegung der Kamera und damit auch der Positionierung des digitalen Endgeräts im Raum (vgl. Fehrenbacher 2024a). Das im Wesentlichen auf diesem interdependenten Wechselverhältnis von Kamerabild, Sensordaten und Algorithmen basierende Trackingverfahren wird auch mit der Bezeichnung Visual-Inertial SLAM (viSLAM) präzisiert (vgl. Ranganathan 2022; Servières et al. 2021; Reality Labs 2019).

Schon seit 2008 wurde an Umsetzungsformen von AR auf Smartphones gearbeitet, bei denen ein GPS-Ortungssystem sowie Lage- und Beschleunigungssensoren zur Registrierung der »Blickrichtung« des Endgeräts dienten.⁵ Spätestens mit dem

4 Ein Frame ist ein einzelnes, gerendertes Bild einer Bildsequenz.

5 Dies gelang erstmals beim unter den Namen *HTC Dream* und *T-Mobile G1* vermarkteten Smartphone, das in Kooperation zwischen Google, HTC – einem mittlerweile auch im VR-Bereich führenden taiwanesischen Technologieunternehmen – und T-Mobile entwickelt wurde.

Erfolg des Mobile Games *Pokémon Go* im Jahr 2016 erfolgte eine vorläufige Abkehr von der bei Sutherland konzipierten, jedoch in der konkreten Umsetzung bis dato gescheiterten Wiedergabe von AR über ein HMD. Im gleichen Jahr präsentierte der Beitrag »SLAM: Bringing art to life through technology«⁶ in Metas Forschungsblog die erfolgreiche Implementierung des algorithmischen Kameratrackings als Grundlage für die Realisierung von Augmented Reality auf mobilen Endgeräten. Im Vergleich zu den frühen Experimenten mit AR-Applikationen auf dem Smartphone liegt das entscheidende Novum hierbei in der Ermöglichung einer stabilen Verankerung des augmentierten Inhalts in dem von der Kamera erfassten Raum: »Facebook previously released a version of SLAM for AR on mobile devices which uses a single camera and inertial measurement unit (IMU) to track a phone's position and enable world-locked content — content that's visually anchored to real objects in the world« (Romea/Meyer 2016). Veranschaulicht wurde dies anhand einer künstlerischen Arbeit von Heather Day, die am Gebäude der Hauptzentrale von Meta virtuell »angebracht« wurde (Abb. 49). Während bei *Pokémon Go* die animierten Monster noch eingeschränkt und recht eigenwillig auf die Bewegungen des Smartphones reagierten, verbleibt die augmentierte Kunstinstallation stabil an den dafür vorgesehenen Bereichen der realphysischen Architektur. Die vorherigen Ausführungen zur blickpunktabhängigen Bildgenerierung aufgreifend wurde damit das von Sutherland bereits 1968 erdachte Gestaltungsprinzip einer stabilen Objektverankerung verwirklicht. Allerdings führte erst ein Jahr später⁷ ein paradigmatischer Use-Case die augmentierte Objektverankerung zur Marktreife – als nämlich Apple das SLAM-Verfahren in seine Programmierschnittstelle *ARKit* integrierte, deren Funktionalität bzw. »Massentauglichkeit« mit der Applikation *IKEA Places* prominent vermarktet wurde (vgl. Fehrenbacher 2024a, 90).⁸ Mit der AR-App können Einrichtungsgegenstände des schwedischen Möbelherstellers maßstabsgetreu in den eigenen vier Wänden virtuell platziert und deren größentechnische sowie ästhetische Passung getestet werden. Die potenziellen Käufer*innen können sich um das Möbelstück mit dem Smartphone herumbewegen und es von verschiedenen Blickwinkeln aus

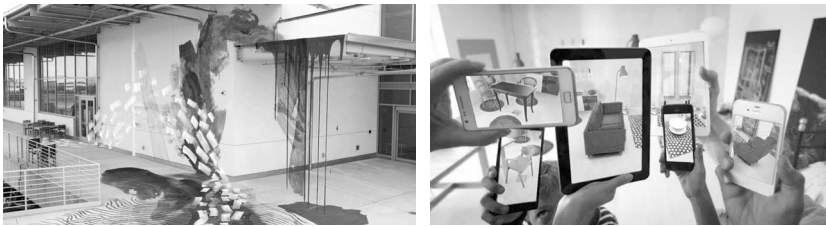
2008 war das *HTC Dream* in den USA und Großbritannien und ab 2009 europaweit erhältlich. 2009 realisierte auch Apple die GPS-Ortung in Smartphones (vgl. Lenzen 2020, 272f.).

- 6 Auf das Sprachbild der (Wieder-)Belegung wurde auch im Diskurs über AR und VR zur Geschichtsvermittlung vielfach zurückgegriffen.
- 7 Der zuvor besprochene Beitrag in Metas Forschungsblog zur virtuellen Verankerung der künstlerischen Arbeit von Heather Day von 2016 ist mittlerweile nur noch über das Internet Archive abrufbar. Der Artikel wurde jedoch unter einer anderen URL mit einem auf das Jahr 2017 datierten Erstelldatum erneut veröffentlicht. Demnach wäre Meta die AR-Verankerung im selben Jahr wie Apple gelungen und nicht bereits ein Jahr früher (vgl. Collet/Meyer 2017).
- 8 Im Jahr 2018 zog Google mit seiner Programmierschnittstelle *ARCore* nach, die ebenfalls auf SLAM basiert.

betrachten. Neben dem Mobile Game *Pokémon Go* fungiert die IKEA-App mittlerweile als prototypisches Beispiel, um AR-Technologie und ihre Funktionalität zu illustrieren – so auch in den für die vorangegangene Diskursanalyse gesichteten Beiträgen (z. B. in Seitz/Kerber/Bernsen 2017, 129; Stiftung Lesen 2019, 6). Bei Buchner (2019) dient die eng mit der Vermarktung von Produkten verschränkte IKEA-App bzw. eine Anlehnung an ebendiese sogar als bildlicher Aufmacher für einen Beitrag zum unterrichtlichen Einsatz von AR (Abb. 50).

Abb. 49: Künstlerische AR-Arbeit von Heather Day (links)

Abb. 50: Bildlicher Aufmacher zum Artikel von Buchner (2019); originale Bildunterschrift: »Mit Augmented Reality kann der Lernort Klassenzimmer erweitert werden« (rechts)



Quellen: Romea/Meyer 2016 (links); Buchner 2019 (rechts)

Ein AR-Objekt, das von allen Seiten betrachtet werden kann, erscheint nicht länger als »dynamisches Artefakt« (Grabbe/van der Veen 2024, 3), sondern als stabil im Raum verankert. Allerdings kann es dies nur auf Grundlage der zuvor beschriebenen algorithmischen Datenprozessierungen, Echtzeitabtastungen des Kamerabil-des und Verarbeitung der internen Sensorik des Devices. Auf rein datentechnischer Ebene ist es deshalb nicht stabil: »Es dreht, verschiebt und skaliert sich abhängig von Bewegung und Drehung des Geräts« (Fehrenbacher 2024a, 86). Zugunsten eines realistischen Eindrucks wird mit der stabilen Objektverankerung, wie der Medienwissenschaftler und Medienkünstler Jens Fehrenbacher herausarbeitet, auf ästhetischer Ebene dieses relationale Geschehen und damit die Medialität der Augmentierung verunsichtbart.

Die Implementierung von SLAM respektive Visual-Inertial SLAM in Metas Virtual Reality Headsets, die damals noch unter dem Tochterunternehmen Oculus firmierten, bedeutete ebenfalls eine entwicklungstechnische Zäsur. Mit der Einführung der Oculus Go im Jahr 2018 gelang es erstmals, ein mobiles Standalone-Headset für den Consumer-Markt bereitzustellen. Auf externe Tracker und kabelgebundene Lösungen mit einem Standcomputer, die bei Vorgängermodellen oder Produkten der Konkurrenz aufgrund anderer Trackingverfahren (z. B. Lighthouse-Tracking) notwen-

dig waren und sind, kann seitdem verzichtet werden. Das in Metas VR-Technik integrierte algorithmische Kameratracking kann zudem als Möglichkeitsbedingung dafür in den Blick genommen werden, dass vier Jahre später mit der als Mixed Reality Brille gelabelten Meta Quest 3 sowohl Virtual Reality als auch Augmented Reality Inhalte wiedergegeben werden können. Mit der auf der Entwicklerkonferenz 2023 prominent auf- und eingeführten Geste des ›Double Tap‹ ist ein schneller Wechsel zwischen Virtual Reality und dem sogenannten Passthrough-Modus als ›Live-Bild‹ des physischen Umgebungsraums möglich. In diese kamerabasierte Ansicht können nun virtuelle Objekte als *video see-through* AR (siehe Fußnote 13, S. 18) eingeblendet werden. Beim Start einer AR-Anwendung wird dazu im Head-Mounted Display die sensorische Raumerkennung als ein sich sukzessiv aufbauendes Polygonnetz sichtbar. Zusätzlich erfolgt eine algorithmische Interpretation der registrierten Flächen als Wände, Decken, Fenster und Möbel. Die Implementierung des Visual-Inertial SLAM-Verfahrens machte VR-Brillen damit nicht nur mobiler einsetzbar und deren Materialität unmerklicher. Auch wurden nach einer entwicklungstechnischen Trennung der Ausgabemedien von AR (Smartphone und Tablet) und VR (HMD) beide Medientechnologien hinsichtlich ihrer Hardware und auch Software wieder zusammengeführt.

Durch das Trackingverfahren SLAM wird zusammenfassend eine Verunsichtbarung relationaler Datenprozessierungen durch eine wahrgenommene, stabil im Raum verankerte Objekthaftigkeit sowie eine Verringerung materieller Komponenten als potenzielle Störfaktoren eines unmittelbaren Rezeptionseindrucks vorangetrieben. Neben der blickpunktabhängigen Bild(raum)generierung als funktionslogischer Grundstein von Medien der AR und VR bedingt und befördert somit auch SLAM die Erzählung rahmenloser Bildlichkeit und damit zusammenhängender Unmittelbarkeit.

6.1.2 Implikationen phantasmatischer Entrahmung

Der Rahmen als das im Bildformat eingeschlossene Bildfeld ist mehr als eine visuell-materielle Grenze zwischen dem Bild und einem nicht-bildlichen Außen. Zum einen erfordert er von den Bildproduzent*innen die Einhaltung eines bestimmten Formats und die Konzeption einer innerbildlichen Ordnung. Gleichzeitig wirkt er auf das schauende Subjekt und organisiert dessen Wahrnehmung und Deutung des Gezeigten mit. Entsprechend argumentiert die Filmschaffende Julia Bruton, dass der filmische Bildrahmen für sie Inhalte vereindeutige: »In einem geschichtlichen und gesellschaftlichen Kontext wird eingeschränkt und vorgegeben, was wir zu sehen und zu fühlen haben – und dieser Kontext beruht damit zugleich auf Macht, Dominanz und auch Gewalt« (Bruton 2022, 108). Demgegenüber sieht Bruton ihre in VR realisierten Arbeiten von ebendieser rahmenden Beschränkung befreit. Diese

Argumentation weist auf eine der zentralen Implikationen der diskursiven Entrahmung hin – nämlich, dass virtuelle und augmentierte Bildlichkeit mit dem Wegfallen einer als fremdbestimmt empfundenen Begrenzung (z.B. einer festgelegten Bildkomposition, dem Kamerablick und einer vorgegebenen Narration) assoziiert und semantisch mit der souveränen Kontrolle über eben diesen Bildraum aufgeladen wird. In diesem Sinne sei die Wahrnehmung bei der Rezeption von AR- und VR-Inhalten lediglich auf die Physiologie des menschlichen Auges zurückgeworfen. Der Blick erscheint grenzenlos, unvermittelt und unmittelbar, was in folgenden Zitate exemplarisch zum Ausdruck kommt:

»Möglich macht das die besondere Technik der 360°-Aufnahmen. Bei diesen Videos ist das Blickfeld der Betrachter nicht auf eine vorgegebene Kameraperspektive festgelegt. Stattdessen kann man sich rundum frei umsehen. Wie in der realen Situation kann man sein Smartphone in alle Richtungen wenden und so beispielsweise zum Himmel schauen oder auf den Boden blicken. Dadurch entsteht der Eindruck, man sei tatsächlich vor Ort und könne die Umgebung mit eigenen Augen entdecken.« (Planet Schule 2019a)

»Motiviert durch das Ziel, eine authentische mediale Umgebung zu schaffen, war der nächste logische Schritt in dieser Entwicklung der Versuch, den besagten Rahmen der Darstellung zu verlassen.« (Stiftung Lesen 2018, 4)

Diese Implikationen spiegeln sich in den fachspezifischen Diskursen insofern wider, als AR- und VR-Inhalten medienvergleichend eine gesteigerte Authentizität, Realitätsnähe und Wahrhaftigkeit zugeschrieben wird. Auch in der Rede vom Hineinversetzen in vergangene Zeiten, fremde Körper oder beliebige Orte der Welt mittels VR und vom Herholen von Personen und Objekten in den »persönlichen Raum« mittels AR wird das Unmittelbare und Grenzenlose – unter, wie noch zu zeigen sein wird, Ausblendung der tatsächlich medialen Bedingtheit – der Rezeptionserfahrung diskursiviert. Andrejevic bringt diese Ausdeutungen in seinen Überlegungen zur »framelessness«, die er im Kontext digitaltechnischer Automatisierungsprozesse entwickelt, auf den Punkt:

»Consider the developments in digital imagery that promise to dispense with the limit of the frame: virtual reality, augmented reality, and 360-degree cameras. The goal is to free the image from the selectivity that has shaped and constricted it – in part to fulfill the promise of total information capture and to overcome the biases of partiality.« (Andrejevic 2019, 117)

Das blickpunktsensitive Verarbeiten von Bilddaten und performative Erzeugen ebendieser durch die eigene Körperpositionierung und -bewegung im physischen Raum befördert zudem Beschreibungen, in denen die Rezipient*innen nicht länger

als passive und stillgelegte Zuschauende, sondern als selbstbestimmte Bildproduzent*innen und »aktive Gestalter« (Buchner 2018, 132) adressiert werden:

»Diese maßgebliche Eigenschaft der 360° Videos drängt die Autorschaft eines Kameramannes gänzlich aus dem Herstellungsprozess. Mit der kleinsten Kopfbewegung wird der passive Rezipient zum aktiven Bildgestalter und übernimmt die Rolle des Kameramannes. Als 360° Video Rezipient erhält man nicht nur vollständige Kontrolle über den sichtbaren Raum, sondern auch darüber, wann welcher Inhalt gezeigt wird.« (Neußl 2018, 81)

»Auf diese Weise ist er nicht mehr auf den vom Kameramann gewählten Bildausschnitt festgelegt, sondern kann aus der Perspektive der Kamera das gesamte Geschehen überblicken. VR macht also aus einem passiven Beobachter einen aktiven Teilhaber des Geschehens.« (Stiftung Lesen 2018, 6)

Entsprechende Mehrwertzuschreibungen vom selbstständigen und selbstbestimmten, aktiven bzw. interaktiven, entdeckenden und erlebnisorientierten Lernen mit AR und VR wurden in allen drei Fächern weitestgehend pauschal ausgesprochen, ohne eine genauere Differenzierung hinsichtlich der tatsächlichen technischen Möglichkeitsräume für (Inter-)Aktionen nachzuliefern.

Wenn Sybille Krämer die für Medien charakteristische Verunsichtbarung medialer Bedingtheiten mit den Worten umschreibt: »Medien wirken wie Fensterscheiben: Sie werden ihrer Aufgabe umso besser gerecht, je unauffälliger sie unterhalb der Schwelle unserer Aufmerksamkeit verharren« (Krämer 1998, 74), ist diese Feststellung in Bezug auf AR und VR umso brisanter – ist es doch der Fensterrahmen oder gar das Fenster an sich, das in der diskursiven Aufladung beider Medientechnologien affirmativ zu Disposition gestellt wird.

6.2 Materielle Vermitteltheit

»Medien sind Mittler und bilden eine Sphäre der Vermittlung. Medien treten dazwischen. Zwischen die Kommunizierenden, und zwischen sie und die Welt. Wie alle Mittler sind die Medien freundlich-verbindliche Diener und unüberwindliche Trennung/Barriere. Sphäre der Moderation, der Verständigung und des Ausgleichs, machtvoll-unumgängliche Zwischeninstanz, Ort der Verfälschung, Umleitung, des Mithörens und der Zensur. ›Unmittelbarkeit‹ gibt es in den Medien nicht.«

(Winkler 2008, 39)

Es wurde deutlich, wie sich das Phantasma einer Entrahmung augmentierter und virtueller Bildlichkeit als Zusammenspiel aus medientechnischen Möglichkeitsbedingungen und diskursiven Praktiken beschreiben lässt. In besonderem Maße wird dabei eine Unmittelbarkeitserzählung bedient, die mit Implikationen eines selbstbestimmten, aktiven, allumfassenden Bildzugriffs sowie mit Authentizität und Objektivität der AR- und VR-Inhalte einhergeht – und deren grundsätzliche Medienvermitteltheit ausgeblendet. In der Rahmenmetapher verbleibend kann mit der filmtheoretischen Unterscheidung zweier Arten von Einrahmung bzw. Kadrierung – dem sogenannten *hors-champ* und dem *hors-cadre* – dieser Befund im Hinblick auf den Einsatz von AR und VR in Bildungskontexten problematisiert werden. Das *hors-champ* bezeichnet das diegetische Off, also den durch den Bildrahmen ausgeschlossenen Raum der fiktionalen Welt, der von den Zuschauer*innen imaginativ fortgesetzt wird. Demgegenüber benennt das *hors-cadre* das durch den Bildrahmen ausgeschlossene, nicht diegetische produktionsbezogene Außen. Damit gemeint ist das profilmische Setting der Dreharbeiten und der Filmproduktion (z.B. das Filmequipment, die Kulissenbauten, die am Set tätigen Personen), das zur Aufrechterhaltung der Diegese eben nicht in das Filmbild gelangen soll. Wie Hasebrink (2024, 28) argumentiert, weist das *hors-cadre* jedoch nicht nur eine räumliche, sondern auch eine zeitliche Dimension auf, indem Herstellungsprozesse des Bildes der Rezeption zeitlich vorgelagert und damit ebenso aus dem Filmbild ausgeklammert bzw. gänzlich unverfügbar sind.

Wohlwissend, dass sich filmtheoretische Begrifflichkeiten nicht umstandslos auf AR und VR übertragen lassen, sensibilisiert die Differenzierung beider Kadrierungen dafür, dass in der diskursiven und zumeist auch medientechni-

schen Ausgestaltung der rahmenlosen Bildlichkeit mehr aus dem Blick gerät als lediglich eine sichtbare Begrenzung der durch AR und VR hervorgebrachten Inhalte. So wird nicht nur das *hors-champ* als diegetischer, »unsichtbare[r] Raum im Off« (ebd., 27) ausgeblendet, sondern es wird gleichzeitig auch eine Kadrierung im Sinne des *hors-cadre* und damit das gesamte soziotechnische Außen als nicht-diegetisches Off negiert. Dadurch treten die Ermöglichungsbedingungen der Bildproduktion selbst systematisch in den Hintergrund, womit einer Diskursivierung von AR- und VR-Medien als neutrale Informationskanäle Vorschub geleistet wird. Wie Hasebrink (2024, 60) darlegt, kann bei 360°-Bildern zwar von keiner klassischen, räumlichen Produktionsrealität im Off, also einem Filmset »hinter der Kamera«, gesprochen werden. Dies gelte allerdings nicht für zeitlich vor- und nachgelagerte Produktionsprozesse, die für ihn durchaus als für Rezipient*innen unverfügbares Außen verstanden werden können. Übertragen auf AR- und VR-Anwendungen entspräche dies etwa vorgelagerten redaktionellen Entscheidungen, dem Programmieren von VR-Szenarien oder Filmen eines 360°-Videos und dessen Nachbearbeitung. Auf Grundlage der vorherigen Ausführungen zur performativen Bildproduktion durch Körper und Datenprozessierungen wird an dieser Stelle dafür argumentiert, unter dem *hors-cadre* auch die zugrundeliegenden Datenprozesse und Datenpraktiken (Kapitel 6.2.1) sowie die mit dem medialen Gefüge aus Hard- und Software verschränkten, »im Off« agierenden Körper (Kapitel 6.2.2) zu fassen. Inwiefern beides zu einer *Rahmung* der virtuellen und augmentierten Bildlichkeiten beiträgt und den Unmittelbarkeitszuschreibungen insgesamt zuwiderläuft, wird im Folgenden aufgezeigt.

6.2.1 Datenprozesse und Datenpraktiken

»Digitale Welten sind Zeichenwelten respektive Ansammlungen von Daten in Form von Bits (*binary digit*) [Herv. i.O.] ähnlich den Molekülsammlungen der physisch-realen Welt. Natürlich basieren die Bitansammlungen auf den Zuständen der Molekülsammlungen der Hardware; den Siliziumkristallen und deren Elektronenanordnungen. Das macht das Digitale so interessant, da es binär-informatisierte Materie ist [...].«
(Gramelsberger 2023, 141)

Digitale Daten sind auf technologisch mannigfaltige Weise physikalisch materialisiert (vgl. Burkhardt et al. 2021). Sie können auf unterschiedlichen Trägermedien gespeichert und in vielfältigen Zuständen übertragen werden. Digitale Daten

lassen sich als ausschließlich für den Computer lesbare Zeichen nicht physisch fassen, riechen, schmecken, hören oder betrachten. Es braucht daher einer instantanen Umwandlung von computergelesenen Zeichen in für die menschliche Sensorik wahrnehmbare Zeichen bzw. Informationen, z.B. alphanumerisch in Zahlen oder Buchstaben, in Grafiken, (Bewegt-)Bilder, Ton oder Klang. Sie sind damit »Produkt von Berechnung und Relationieren, von algorithmischen Prozessen, die sich der sinnlichen Wahrnehmung entziehen« (Leistert 2015, 176). Die objekt-hafte Erscheinungshaftigkeit der durch digitale Medien vermittelten Inhalte, die Gabriele Gramelsberger als »phänomenale Dichte« bezeichnet, »kaschiert die Ambivalenz des Digitalen, autooperative Maschinenausführbarkeit und zugängliche Darstellbarkeit zugleich zu sein« (Gramelsberger 2023, 150). Ein analytisches Verweilen auf der phänomenalen Dichte lässt die zugrundeliegende datentechnische Basis und mit ihr die semiotischen Übersetzungsprozesse und damit einhergehenden Verfremdungen, Verschiebungen und Abstraktionen unberücksichtigt. Dementsprechend versteht Gramelsberger die von ihr entworfene Philosophie des Digitalen als Digitalitätskritik, »die die Selbstverständlichkeiten der qualitativen, lebensweltlichen Seite der Digitalisierung, also der Signatur der digitalen Lebenswelt respektive Wirklichkeit« (ebd., 226) hinterfragt. In diesem Sinne fordert auch Missomelius eine »Entselbstverständlichung« binärlogischer Strukturen und verfolgt in ihrer Monographie »Bildung – Medien – Mensch. Mündigkeit im Digitalen« (2022) damit nicht nur ein medienwissenschaftliches Anliegen, sondern ein dezidiert bildungstheoretisches Programm. Vor diesem Hintergrund gilt es, die Unmittelbarkeitserzählung datenkritisch zu befragen.

Unsichere und situierte Daten | Im Marketingdiskurs schlägt sich die Vorstellung einer allumfassenden, objektiven Abbildhaftigkeit und eines verlustfreien Übertrags realweltlicher Phänomene in die computergenerierten oder in 360° videografierten Audiovisionen nieder. Dies zeigt sich...

- in der Bewerbung virtueller Räume als Reiseersatz und Kopie ihrer realweltlichen Referenz sowie in entsprechend umfassenden Kartierungsoffensiven »to create a digital 1:1 map of the entire world« (Carter/Egliston 2023, 492), wie sie neben Google auch Forschungsinitiativen wie *Project Aria* von Meta für die Entwicklung von AR-Brillen verfolgen,
- in dem Paradigma einer original- bzw. detailgetreuen Rekonstruktion historischer Artefakte oder Architekturen und der körperlich-leiblichen Nacherlebbarkheit von eindeutigen historischen Ereignissen im Stile eines »so ist es gewesen«,
- in der Bewerbung einer exakten Vermessung, Analyse und Simulation von Bewegungsausführungen im Sport.

Wenn der (arbeits-)technische Produktionsaufwand von AR- und VR-Formaten im Marketingdiskurs thematisiert wird, geschieht dies nicht, um eine kritische Auseinandersetzung mit der Konstruiertheit der Darstellungen anzuregen. Vielmehr fungiert diese Erwähnung als Strategie, um Authentizität und Realitätsnähe zu unterstreichen. Diese Zuschreibungen bleiben in den entsprechenden fachdidaktischen Diskursen zum Teil unwidersprochen oder werden reproduziert. Kritische Stimmen, die sich vor allem in der Geschichtsdidaktik finden ließen, fordern zwar eine Reflexion oder Offenlegung der unumgänglichen Partikularität und Unvollständigkeit von Abbildungen der Vergangenheit. Jedoch fehlt es an einer expliziten Auseinandersetzung mit den Darstellungslogiken von AR und VR sowie mit dahinterstehenden Datenspuren und -praktiken.

Gramelsberger zeichnet die binärlogischen Repräsentationsmöglichkeiten als Transformationskette nach, die »vom Qualitativen [realweltlicher Phänomene] ins Formal-Quantitative und vom Formal-Quantitativen zurück ins qualitativ Lebenseweltliche« (Gramelsberger 2023, 225) erfolgt. Hierbei plädiert sie für eine kritische Auseinandersetzung mit den Übersetzungsprozessen und den damit einhergehenden Verschiebungen, Abstraktionen und Verzerrungen, die den Anspruch auf objektive Abbildhaftigkeit eben nicht erfüllen. In vergleichbarer Weise formulieren auch Carter und Egliston (2023, 494) ihre Kritik an einer mit VR verbundenen »fantasy of perfect data«. Darüber hinaus sind die zugrundeliegenden Daten an sich niemals objektiv. Stattdessen ist ihre Generierung immer machtdurchgesetzt, produziert Sichtbarkeiten und geht zugleich mit Unsichtbarkeiten einher; Daten sind »situiert, sozial und politisch« (Macgilchrist 2024, 237). Carter und Egliston betonen in diesem Zusammenhang:

»But data is never a neutral representation. It is always collected for a specific purpose, making visible something that was previously concealed and constructing a new view about the world. In other words, as Lisa Gitelman and Virginia Jackson put it, data is never raw. It is always ›cooked‹ — collected, stored, and circulated with particular aims and logics in mind. Despite this, predictive analytics are increasingly employed to shape our lives on the basis of the pervasive belief in data's inherent objectivity.« (Carter/Egliston 2020b)

An vorherige Ausführungen anschließend ließe sich also fragen, welche Übersetzungsprozesse und in Datensätze eingeschriebenen Perspektiven durch eine vermeintliche Ungerahmtheit und unmittelbare Erfahrbarkeit der AR- und VR-Inhalte verunsichtbart werden. Das Bewegungstracking von VR-Systemen im Sport ist eben nicht verlustfrei und liefert keine neutralen Daten. Was unter einer reibungslosen, störungsfreien und umgangsfreundlichen Nutzeroberfläche verschwindet, sind normative Vorannahmen über mit dem System interagierende Körper, eine Selektion von erfassten und nicht-erfassten, richtigen und falschen Bewegungen, die

Sichtbarkeiten und Unsichtbarkeiten sowie Inklusion und Exklusion von Subjekten organisieren. Ebenso sind die virtuellen und augmentierten Rekonstruktionen historischer Artefakte, Sachlagen oder geographischer Räume und Phänomene, die Original- bzw. Detailtreue versprechen, nicht vollumfänglich, nicht ungerahmt und unmarkiert. Stattdessen sind sie von fragmentarischen und unsicheren Datenlagen durchzogen. Darüber hinaus sind sie, wie die Diskursanalysen der Fächer Geographie und Geschichte verdeutlichen und wie in Kapitel 9 weiter vertieft wird, von neokolonialen Perspektiven gekennzeichnet, die zugunsten eines Realitäts- und Unmittelbarkeitseindrucks unsichtbar gemacht werden. Insbesondere Forschung aus den (feministischen) Critical Data Studies (vgl. D'Ignazio/Klein; Macgilchrist/Jarke 2025) sowie den Postcolonial und Critical Heritage Studies (vgl. Dekker 2025; Lex 2025) beschäftigt sich mit dem Marginalisierten und damit den diskriminierenden Momenten und machtförmigen Ausgestaltungen binärlogischer Repräsentationen. Im Raum stehen somit Fragen nach den medientechnischen sowie medienkulturell verfestigten Formen der Wissens(re)produktion und Wissensvermittlung: Welche in Datensätze eingeschriebenen Perspektiven bleiben unsichtbar und welche werden reproduziert? Welche historischen Ereignisse, kulturellen Artefakte, Natur- und Kulturräume oder auch Sportarten wird es nie als AR- oder VR-Anwendung geben, weil entsprechende Daten aus geopolitischen Gründen, privatwirtschaftlichem Desinteresse oder aufgrund von historischen Gegebenheiten nicht existent sind, nicht archiviert oder zerstört wurden und den Anforderungen an eine detailgetreue Rekonstruktion damit nicht entsprechen (vgl. Fehrenbacher/Przybylka 2025)?

Gesicherte Datenspuren für Konzern XY | Ein weiterer Aspekt drängt sich bei der Reflexion der datentechnisch materialisierten Bedingungen von Medien der AR und VR mit Blick auf Lehr- und Lernkontexte auf: das Sammeln von Daten und deren Verquickung mit datenkapitalistischen Strukturen. Wie die vorherigen Ausführungen zur blickpunktsensitiven Bildgenerierung und zur Trackingmethode SLAM gezeigt haben, werden die Position und (Blick-)Bewegung des menschlichen Körpers im realphysischen Raum permanent getrackt und in die virtuelle, dreidimensionale Umgebung übersetzt. Das am Kopf getragene Headset dient dabei nicht nur der Positionsbestimmung, sondern kann je nach Modell auch Augen- und Mundbewegungen sowie Gesichtsregungen erfassen. Neuere VR-Systeme sind zudem in der Lage, mittels Handtracking die Controller zu ersetzen. Bei VR-Brillen des Konzerns Meta wird beim Aktivieren dieser Funktion darauf hingewiesen, dass dabei Daten über Hand- und Körperhaltung sowie Hand- und Körpermaße erfasst werden. Überdies können durch optionale Tracker die Bewegungen weiterer Körperglieder nachverfolgt und sogar Herzfrequenzen gemessen werden. Während beim sogenannten Lighthouse-Tracking⁹ die Positionsverfolgung der Körper der Rezipi-

9 Dieses Trackingverfahren wird z.B. bei den VR-Brillen der HTC Corporation eingesetzt.

ent*innen durch im physischen Raum installierte Basisstationen erfolgt, die mittels Infrarotabtastung die Lokalisierung des HMD und der Controller vornehmen – Informationen über die Beschaffenheit des physischen Raums werden somit nicht gesammelt –, werden beim zuvor beschriebenen SLAM-Verfahren permanent umgebungsbezogene Daten kamera- und sensorbasiert generiert, ausgewertet und zur Wiedererkennung der Räume gespeichert. Zusammenfassend lässt sich festhalten, dass bei VR durch eine körpernahe Schnittstellengestaltung sowohl biometrische als auch umgebungsbezogene Daten kontinuierlich erfasst, verarbeitet und miteinander in Beziehung gesetzt werden.

Auch bei AR-Apps für Smartphones und Tablets erfolgt auf Grundlage der zuvor erläuterten technischen Kernelemente eine »dauerhafte und aufwendige Analyse der Umgebung und der Bewegung des Endgeräts durch Kamera, Sensoren und darauf aufbauende Algorithmen« (Fehrenbacher 2024a, 89).¹⁰ Darüber hinaus werden bei merkmalsbasierten Face Filter Apps zum Testen von Brillen, Frisuren, Kosmetik etc. Gesichter getrackt und analysiert. Bei GPS-basierten Anwendungen wird darüber hinaus der Standort der Nutzer*innen per Geolokation erfasst. AR-Navigationsapps für Innen- und Außenräume sowie die Entwicklung und Forschung an AR-Brillen potenzieren zudem das orts- und kontextsensible Sammeln, Verarbeiten und Archivieren von Datenspuren (vgl. Carter/Egliston 2020a, 18). So heißt es auf der Homepage von Metas *Project Aria*: »Sensors on the Project Aria glasses capture the wearer's video and audio, as well as their eye tracking and location information. [...] Helping to build a better understanding of our world, together« (Meta 2024b) – wobei zu fragen wäre, welchen Akteuren diese Mensch-Maschine-Kooperation primär zugutekommt. Um die Präzision des AR-Features in Apples Navigationsapp *Maps* zu verbessern, bittet der Konzern seine Nutzer*innen, Informationen über ihre Umgebung zu teilen. Betont wird, dass dabei keine Fotos, sondern lediglich die für das SLAM-Verfahren obligatorischen Feature Points erfasst würden: »No photos are sent to Apple or stored on your device when you use augmented reality features« (Apple Inc. 2024). Diese Hervorhebung scheint darauf abzielen, Datenschutzbedenken zu zerstreuen. Doch nur weil Feature Points für die Nutzer*innen unsichtbar sind, sind sie dennoch hochgradig aussagekräftig hinsichtlich der Umgebungs- und Bewegungssignatur der User*innen.

Bei der Nutzung von Medien der AR und VR werden demnach personenbezogene Daten gesammelt, die nicht länger nur den Zugriff auf Kontakte, Aufenthaltsorte

10 Arbeiten aus der künstlerischen Forschung wie *Gold Environment* (2022) von Chiara Passa oder die App *AR-Brotkrumen*, die von Fehrenbacher seit 2022 sukzessive erweitert wird, zeigen Möglichkeiten auf, unsichtbare Datenpraktiken (hier das Umgebungs- und User*innen-Tracking) sichtbar zu machen. Eine Dokumentation der Entstehung und Weiterentwicklung der App *AR-Brotkrumen* findet sich in Fehrenbacher 2022.

oder Browserinformationen (Tagesrhythmen und Routinen, Verweildauer auf Webseiten etc.) betreffen, sondern auch umgebungs- und körperbezogene, z.T. biometrische Informationen in die Profilbildung mit einfließen lassen. Das Themenfeld der Datafizierung fand in den untersuchten Diskursen jedoch mit Ausnahme von vereinzelten Hinweisen auf erforderliche Accounts beim Nutzen einer AR-App keine Erwähnung. Dieser Befund spiegelt sich auch in Ausführungen von Carter und Egliston wider, die eine Zurückhaltung datenkritischer Kommentierungen in Auseinandersetzungen mit Medien der VR in Bildungskontexten konstatieren:

»As we have recently argued elsewhere, VR devices represent one of ›the most data-extractive digital sensors we're likely to invite into our homes in the next decade‹ [...], yet we are not aware of any critical discussions of the application of learning analytics to VR in higher education or elsewhere, or the necessary speculation (or anticipation [...]) regarding the dangers that VR may present in further accelerating the datafication of education and learning.« (Carter/Egliston 2023, 486)

Die Autoren führen weiter aus, wie sich auf Grundlage der kontinuierlichen, sensorischen Abtastung des Kopfes, der Hände und Körperpositionierung spezifische Bewegungsrelationen und charakteristische Verhaltensmuster ableiten und so User*innen identifizieren lassen:

»One reason that enterprise and higher education institutions should be militant about the ownership and trade o[f] VR data is the potential for biometric identification of users based on their behaviour in VR. Both Pfeuffer et al. (2019) and Miller et al. (2020) have recently demonstrated the potential to identify specific individual users based on unique behavioural biometric markers via movements of the hand, head and eye – data that can be captured by current VR devices. The implications of this are profound, as it suggests that the de-anonymisation of individual training sessions may always be possible.« (Carter/Egliston 2023, 498)

In den beiden angeführten Studien wurde nachgewiesen, dass eine Identifikation von Personen beim Anschauen von 360°-Videos (vgl. Miller et al. 2020) sowie bei der Ausführung banaler Tätigkeiten in virtuellen Umgebungen, z.B. beim Greifen eines Objekts oder bei der Eingabe eines Satzes über eine virtuelle Tastatur (vgl. Pfeuffer et al. 2019), möglich ist, wodurch explizite Authentifizierungsverfahren umgangen werden können. Durch die seitdem zu verzeichnende Weiterentwicklung von Machine-Learning-Verfahren gewinnt diese Problematik an Brisanz, da mit einer steigenden Genauigkeit solcher Identifikationsmöglichkeiten zu rechnen ist.

Auch bei AR können über standort- und körperbezogene Daten personalisierte Profile angelegt werden, was Heemsbergen (2023,19) pointiert als »biospacial surveillance« bezeichnet. Diese Daten verbleiben nicht im Klassenraum, sondern wer-

den durch technologisch-infrastrukturelle Abhängigkeiten von Drittanbietern auf intransparente Weise gesammelt, verarbeitet, als Trainingsdaten genutzt und kommodifiziert. Lernende bzw. ihre in Datenform gegossenen Profile werden somit unbewusst und unfreiwillig für »kommerziell-kapitalistische [...] Verwertungsinteressen von Konzernen« (Niesyto 2021, 5) außerhalb der Schule »auf unsichtbare Weise sichtbar« (Jornitz/Macgilchrist 2021, 118).¹¹ In Anbetracht der aufgezeigten Menge und Sensibilität der durch AR- und VR-Medien erhobenen Daten sowie der Vormachtstellung profitorientierter Konzerne wie Meta in diesem Bereich, der eine äußerst problematische datenschutzrechtlichen Historie aufweist (vgl. Leistert 2015), ist eine kritische Inblicknahme dieser Daten(ab)flüsse erforderlich (vgl. Zender et al. 2022, 38).¹² An welche Akteure welche Daten weitergeleitet werden, lässt sich für Außenstehende nur schwer nachvollziehen: »[They are] hidden behind non-disclosure agreements and within technological black boxes« (Carter/Egliston 2023, 500). Anstatt das Ideal informationeller Selbstbestimmung mit individualisierbaren Lösungen und selbstverantwortlichen Kompetenzschulungen aufrecht zu erhalten, bedarf es daher rechtlich-administrativer Richtlinien und politischer Vorgaben. Diese Forderung findet sich auch in dem Papier »Welche Daten? Welche Literacy?« von der Gesellschaft für Medienwissenschaft:

»Die Zuspitzung des Datenbegriffs als Problem der Privatheit verharmlost die Problematik, die mit der zunehmenden Digitalisierung des Lebens einhergeht, als eine einzelner Individuen. Diese isolierte Betrachtung der Benutzer:in als Subjekt verliert das Gemeinschaftlich-Gesellschaftliche aus dem Blick. Welche Verantwortung tragen gesellschaftliche Institutionen? Welche Verantwortung kommt wirtschaftlichen Sektoren zu? Welche Aufgaben fallen in den politischen Bereich (z.B. der Regulierung)?« (Burkhardt et al. 2021; vgl. auch Missomelius 2022, 98)

11 Die Verquickung der Lernenden und übergeordnet der Bildungsinstitutionen selbst mit wirtschaftlichen Interessen und global agierenden Technologieunternehmen drängt sich nicht erst mit der Einführung von Medien der AR und VR auf (Stichwort: »iPad-Klassen«, Schulbuchverlage). Der bildungsindustriellen Verflechtungen bei der Konzeption, Entwicklung und dem Vertrieb von lerntechnologischen Anwendungen (»bildungsindustrieller Komplex«) sowie der Präsenz global agierender Technologieunternehmen und -plattformen im Bildungssektor wird sich im Rahmen der Critical EdTech Studies umfassend gewidmet (vgl. Macgilchrist 2019/2023; Niesyto 2021; Förtscher 2018). Hingewiesen sei zudem auf die Initiative *Critical Data Lab* und das Positionspapier der Initiative *Bildung und digitaler Kapitalismus* aus dem Jahr 2023.

12 Mit Dander (2024, 117) soll an dieser Stelle betont werden, dass Vermessungspraktiken und Datafizierung nicht zwangsläufig mit »Formen von Instrumentalisierung, Kapitalisierung, Disziplinierung, Regierung etc.« verbunden sein müssen. Auch können sie machtkritisch eingesetzt werden, um gesamtgesellschaftliche und strukturelle Ungleichheiten sichtbar zu machen.

Angesichts gegenwärtiger Datafizierungstendenzen in Lehr- und Lernkontexten (vgl. Schiefner-Rohs/Hofhues/Breiter 2024; Dander et al. 2024) stellt sich die Frage, ob die beim Lernen mit AR und VR generierten, mess- und auswertbaren Daten perspektivisch in schulische oder universitäre Bewertungssysteme eingebunden werden. Im Hinblick auf die in diesem Kapitel thematisierte Unmittelbarkeitserzählung und damit einhergehende Glaubwürdigkeitszuschreibungen an die durch AR und VR hervorgebrachten Daten bzw. Inhalte wird eine kritische Begleitung dieser Entwicklungen unerlässlich. Dies gilt umso mehr angesichts der durch VR ermöglichten Rückschlüsse auf emotionale Zustände und Einstellungen anhand von erfassten Augenbewegungen, Herzfrequenzen, Gesten und Mimiken.

Programmierte Standardisierungen | Das hier entwickelte Argument für die Relevanz einer Thematisierung der Datenprozesse und -praktiken als materielles, rahmendes Dazwischen zielt nicht zuletzt auch auf programmierte Standardisierungen ab, die den Möglichkeits- und auch Vorstellungsraum dafür abstecken, welche Inhalte wie realisiert werden können. Die Nutzung des SLAM-Verfahrens zur Darstellung augmentierter Inhalte auf Smartphones ist somit nicht ausschließlich als medientechnische Zäsur zu bewerten. Die AR-Programmierschnittstellen von Apple (ARKit) und Google (ARCore) sowie darauf aufbauende Autorentools zum niedrigschwelligen Erstellen von AR-Anwendungen, z.B. das auf den schulischen Unterricht ausgelegte Programm *Delightex Edu* (ehemals *CoSpaces Edu*), legen die stabile Positionierung von virtuellen Objekten im Raum als programmiertechnischen Standard fest. Dadurch wird nicht nur eine bestimmte Gestaltungsnorm etabliert, sondern es werden auch spezifische Anwendungsfälle und Rezeptionsformen verfestigt, wie Fehrenbacher (2024a, 90) darlegt. So zeichnet sich das prototypische Einsatzkonzept von Augmented Reality Apps dadurch aus, ein virtuelles Objekt auf einer horizontalen Fläche im Raum (z.B. auf dem Boden oder auf einem Tisch) zu platzieren bzw. zu verankern, sodass es anschließend von allen Seiten betrachtet werden kann. Dieser Standard lässt sich auch mit den in Kapitel 3–4 beschriebenen Anwendungen belegen.¹³ Das Betrachten der Objekte wird dabei als unmarkierter (Über-)Blick von außen inszeniert, denn: Durch die zuvor beschriebene Verunsichtbarung der kontinuierlichen algorithmischen Relationierung zwischen AR-Inhalt, Umgebungsraum und Smartphone respektive dem menschlichen Körper wird

13 In Begriffsdefinitionen von AR wird zum Teil ausschließlich von dieser festen Verankerung ausgegangen, was besonders im Fach Geographie deutlich wurde. Alternative Realisierungsformen werden dadurch kategorisch ausgeklammert, wie auch das folgende Zitat aus einem Beitrag zum Einsatz von AR und VR im Unterricht veranschaulicht: »Das Video wird mit virtuellen Inhalten überlagert, die durch eine integrierte Positionsbestimmung der mobilen Geräte fest in der Welt verankert sind. Entsprechend bleiben Inhalte an der gleichen Stelle in der Welt überlagert, auch wenn sich das Gerät bewegt« (Buchner et al. 2023, 214f.).

der*die Rezipient*in von dem augmentierten Objekt rezeptionsästhetisch entkoppelt (ebd., siehe auch Kapitel 8.2.2). Die Umsetzung alternativer AR-Formen, die das relationale Verhältnis zwischen Körper, Raum und Technik sichtbar machen, ausloten und damit experimentieren, wird durch die technologische Standardsetzung der marktführenden Programmierschnittstellen erschwert und verlangt informatisches Wissen über die im Backend praktizierte Raum-, Flächen- und Bildanalyse. In seiner Dokumentation des Forschungsprojekts *AR in Public* stellt Fehrenbacher zudem fest, wie in Googles *ARCore* durch das automatisierte und nicht konfigurierbare Erkennen von Ebenen und Oberflächen bestimmte Vorstellungen von Architekturen und Räumen eingeschrieben sind. So werden zur Platzierung von AR-Inhalten von der Software horizontale und vertikale Flächen priorisiert: »Die AR-Schnittstelle ›ARCore‹ von Google [...] ist also offensichtlich auf urbane, möglichst rechtwinklige Räume ausgelegt« (Fehrenbacher 2022). Diese einprogrammierte Bevorzugung spezifischer Raumstrukturen ist im Kontext ortsgebundener Augmented Reality Anwendungen, insbesondere bei Navigationsapps, von Brisanz.

Es wird deutlich, wie durch derartige Formatierungen eines »technological sensing and sensemaking« (Borbach/Kanderske 2024, 177) in Programmierschnittstellen die Umsetzung bestimmter Inhalte wahrscheinlicher wird als die anderer, wie damit zusammenhängend Rezeptionspraktiken nahegelegt werden und sich ästhetische Gestaltungsprinzipien etablieren. Der technische Möglichkeits-, aber auch Vorstellungsraum für die Gestaltung und Einsatzmöglichkeiten von Apps wird dadurch enggeführt. Auch gehen mit diesen Standardsetzungen gewisse lebensweltliche Implikationen einher, z.B. die technologische An/Erkennung bestimmter Raumstrukturen ebenso wie deren Nicht-Berücksichtigung. Derartige Rahmungen finden sich ebenso in Bezug auf Softwareumgebungen zur Erstellung von VR-Anwendungen, wie Fehrenbacher (i.E.) anhand der *Unreal Engine* darlegt. Besonders wirksam zeigen sie sich zudem in der Erzeugung standardisierter User*innenkörper, wie im Folgenden näher ausgeführt wird.

6.2.2 Der Körper als Schnittstelle

»Andererseits aber bleibt der Körper innerhalb der virtuellen Welt ein Fremdling; als ein tatsächlicher, körperlicher Körper bleibt er an die tatsächliche Welt gebunden.«
(Winkler 1998, 6)

In den 1960er-Jahren revolutionierten Computertechnikpioniere wie Douglas Engelbart mit den sogenannten Grafical User Interfaces (GUI) die Mensch-Computer-Interaktion. Anstatt über schriftliche Befehlseingaben von Maschinencode konnte nun über eine grafische Bedienoberfläche mit dem Computer interagiert werden.

Wenngleich auch bei AR und VR auf grafische Symbole, Icons, Fenster und Menüfelder zurückgegriffen wird, weichen diese zunehmend einem den Körper umgebenden, multisensorischen Interaktionsraum. In diesem werden räumlich verankerte computergenerierte Objekte gegriffen, gedreht, verschoben, geworfen – sei es durch Fingergesten auf dem Display eines Smartphones oder Tablets, mit Controllern oder mit getrackten und in der VR visuell dargestellten Händen. Der Körper ist jedoch nicht nur notwendige Bedingung für diese expliziten Interaktionsformen. Wie zuvor ausgeführt, bedingt die blickpunktabhängige Bildgenerierung zudem, dass AR- und VR-Inhalte konstitutiv mit der Bewegung und Positionierung des Körpers im physischen Raum verknüpft sind und durch diese performativ hervorgebracht werden. Während bei Augmented Reality die Erfassung des Körpers bzw. seiner Bewegungen indirekt über das Tracken des digitalen Endgeräts erfolgt, wird der physische Körper im VR-Setting zur »Aktivierung des Datenkörpers« (Krämer 2002, 66) umfassend »als Datenquelle [...] von der Soft- und Hardware prozessiert« (Braun/Friess 2020, 624).

Michael Abrash, Vice President von Meta und Chief Scientist der Forschungseinrichtung Reality Labs, bezeichnet AR und VR als zweite große Welle des »Human-oriented Computing« (Abrash; Meta Quest 2020a, 1:14:45-1:14:49).¹⁴ Was in seinen alljährlichen Vorstellungen auf Metas Entwicklerkonferenz zur symbiotischen Zusammenführung von Mensch und Computertechnik aufgerufen wird, bedient medienhistorisch tief verankerte transhumanistische Extensionsnarrative, auf die an dieser Stelle nicht näher eingegangen werden soll. Von Interesse ist hingegen das Bestreben, Ein- und Ausgabetechniken zielperspektivisch derart zu gestalten, dass sie sich nahtlos in menschliche Kommunikationsformen einfügen, an körperliche Dispositionen anpassen und unmerklich werden. In einem Blogbeitrag von Reality Labs in der Rubrik »Tech at Meta« heißt es entsprechend:

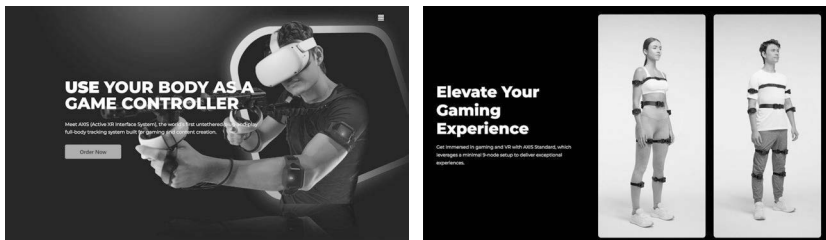
»You might gesture with your hand, make voice commands, or select items from a menu by looking at them — actions enabled by hand-tracking cameras, a microphone array, and eye-tracking technology. But ultimately, you'll need a more *natural, unobtrusive way of controlling* [Herv. durch Verf.] your AR glasses. [...] AR glasses interaction will ultimately benefit from a novel integration of multiple new and/or improved technologies, including neural input, hand tracking and ges-

14 Bei Augmented Reality bezieht sich Abrash auf AR-Brillen und nicht auf smartphone- oder tabletbasierte AR. Als erste Welle des »Human-oriented Computing« versteht Abrash, wie er auf der Entwicklerkonferenz von 2025 erläutert, die Ablösung von der Eingabe von Maschinencode durch grafische Benutzeroberflächen (Graphical User Interfaces). Der wesentliche Unterschied zu am Kopf getragenen AR-Brillen und VR-Headsets besteht für ihn darin, dass uns die 2D-Displays und Bildschirme der ersten Welle zu sehr von unserer situativen Eingebundenheit in die Welt entkoppeln würden (Abrash; Meta Developers 2025b, 03:45-06:53).

ture recognition, voice recognition, computer vision, and several new input technologies [...]» (Reality Labs 2021)

Als Pendant zum Interaktionsparadigma der Graphical User Interfaces wird diese in Aussicht gestellte Schnittstellengestaltung auch unter dem Label der Natural User Interfaces (NUI) verhandelt. Wie sehr das an diese Vorstellung anknüpfende, vor allem im Sport deutlich gewordene Marketingversprechen auf ein realitätsnahes Spielerlebnis in VR (»Play for Real«) und daran geknüpfte Befreiungsoffensiven von störenden, »unnatürlichen« technischen Materialitäten der Realität aktueller VR-Systeme jedoch zuwiderlaufen, lässt sich am Produktportfolio des Unternehmens Refract veranschaulichen. Um den Körper »natürlich« und nahtlos in das virtuelle Geschehen einfügen zu können, werden zwar punktuell physische Steuerungselemente durch kamerabasierte Trackingverfahren ersetzt. Demgegenüber und im Sinne einer multisensorischen Affizierung wird der Körper bzw. am Körper jedoch auch umso umfassender hardwaretechnisch aufgerüstet (Abb. 51–52).¹⁵

Abb. 51–52: Werbung für zusätzliche Trackerkomponenten



Quellen: Refract Technologies 2022 (links); Refract Technologies 2024 (rechts)

In den geographie- und geschichtsbezogenen Diskursen wurde der Körper vor allem im Kontext von sinnlicher Wahrnehmung, Affekt und Emotion thematisiert – sei es, um die Eindrücklichkeit der augmentierten und virtuellen Audiovisionen positiv hervorzuheben oder um eine fehlende kognitive Distanz zu ebendiesen zu problematisieren. In den untersuchten sportbezogenen Diskursen über VR stand wiederum ein motorisch aktiver Körper in seiner werkzeughaften Dimension als Bewegungsapparat im Fokus. Im Rahmen emotionaler und kognitiver Nähe- und Distanzverhältnisse gegenüber den durch AR und VR hervorgebrachten Inhalten und der funktionalen und störungsfreien Einpassung der Lernenden in das technische VR-Setting werden Körper und Leib somit durchaus verhandelt. Was auf eine

15 Die Trackerkomponenten des Unternehmens wurden u.a. für die Taekwondo-Anwendung auf der ersten *Olympic Esports Series* 2023 eingesetzt (siehe S. 161).

Unmittelbarkeitserzählung einzählend jedoch unterschlagen wird, ist die Bedeutung des Körpers als materielles Dazwischen, als im physischen Umgebungsraum handelnder und datafizierter Körper, dessen Einpassung in die medientechnischen Settings von AR und vor allem VR voraussetzungsvoll ist sowie bestimmte Körper schlichtweg ausschließt.¹⁶ Mit einer Semantik der Natürlichkeit, Unmerklichkeit und Störungsfreiheit werden die Materialität der Technik und mit ihr die in Hard- und Software eingeschriebenen Normvorstellungen von Körpern und ihren Funktionalitäten diskursiv ausgeklammert. Wie Kasprovicz (2019, 20) feststellt, dominiert »das Bild einer Übertragung von Daten auf die menschlichen Sinneskanäle und zurück, wobei die Materialität dieser Kommunikation selten berücksichtigt wird.« Um eben diese Materialität soll es im Anschluss an die vorherigen Ausführungen zu Datenprozessen und -praktiken mit Fokus auf den sich im medialen Gefüge aus Hard- und Software bewegenden Rezipient*innenkörper gehen. Da der Körper bei der Nutzung eines VR-Headsets umfassender involviert und mit Technik verschaltet ist, spielt der Aspekt der Verkörperung und Einverleibung informatischer Modellierungen hier eine größere Rolle als bei smartphone- und tabletbasierter AR. Aus diesem Grund beziehen sich die folgenden Ausführungen primär auf VR-Systeme bzw. HMD-basierte AR.

Nutzungsgewohnheiten und Medienbiografien | Die Fragilität der Betitelung von Mensch-Computer-Schnittstellen als »natürlich« wurde medienwissenschaftlich bereits hinreichend diskutiert (vgl. Rieger 2019a, hier speziell 57–71; Andreas/Kasprovicz/Rieger 2016). Ist das Natürliche, Reibungslose und Intuitive der Interaktionsformen doch vielmehr Ausdruck einer Einschleifung von Affordanzen und Resultat der Einverleibung standardisierter Schnittstellengestaltungen:

»None of these systems is inherently more natural than the others. The mouse and keyboard are not natural. Speech utterances will have to be learned and gestures carefully developed and standardized through time. The standards don't have to be the best of all possibilities. The keyboard has standardized upon variations of qwerty and azerty throughout the world even though neither is optimal—standards are more important than optimization.« (Norman 2010, 10)

Vor diesem Hintergrund ist die Praxis mit VR-Technik für Lernende eben nicht überindividuell »natürlich«, sondern von medienkulturell geprägten Wissensbeständen durchsetzt und mit spezifischen Nutzungsgewohnheiten und Medienbiografien verknüpft. So sind im Kontext von VR routinierte Umgänge mit digitalen

16 Mit einer anders gelagerten Frage, nämlich wie und als was Körperlichkeit und Körpererfahrungen in virtuellen Räumen – auch hinsichtlich ethisch-moralischer Implikationen – theoretisch beschrieben werden können, setzt sich Breil (2025) aus einer phänomenologischen Perspektive auseinander.

Konsolenspielen von Vorteil, um sich in den virtuellen Welten und mit der Handhabung der Controller zurechtzufinden. Über dieses anwendungsbezogene Wissen hinaus spielen Gewöhnungsprozesse an die visuelle Abschirmung nach außen beim Tragen eines VR-Headsets eine Rolle. Gerade beim erstmaligen Ausprobieren von VR kann dies ein Gefühl von Kontrollverlust, Verunsicherung oder Scham hervorrufen.¹⁷ Anschließen lassen sich diese exemplarischen Befunde zur Voraussetzungshaftigkeit der als ›intuitiv‹ und ›natürlich‹ ausgewiesenen Umgänge an medienpädagogische Forschung zu strukturellen Ungleichheitsverhältnissen bei der Nutzung digitaler Medien, die im Zusammenhang mit intersektional verwobenen Variablen wie sozioökonomischem Status, Alter, *race* oder Geschlecht stehen. Diese drücken sich in ungleich verteilten Zugangsmöglichkeiten zu Medien und ihren Infrastrukturen (First-Level Digital Divide) und damit zusammenhängenden unterschiedlichen Nutzungspraktiken (Second-Level Digital Divide) aus. Mit dem sogenannten Zero-Level Digital Divide geraten zudem die programmiertechnische Ebene, datenbasierte und algorithmische Strukturen von Medientechnologien sowie dadurch bedingte diskriminierende Ungleichheitsverhältnisse in den Blick (vgl. Verständig/Klein/Iske 2016, 50ff.). An die zuvor machtkritisch diskutierten, datentechnischen Prozessierungen anknüpfend werden diese Ausschlüsse mit Fokus auf den Körper nachfolgend stärker in den Blick genommen.

Gestaltungsstandards und normative User*innenmodelle | Der Gestaltung von Hardware sowie dem Programmcode und seinen algorithmischen Strukturen liegen normative Annahmen über Körper und ihre Funktionalitäten zugrunde. Dies beginnt bei der standardisierten und teilweise auch alternativlosen Ausrichtung der Controllersteuerung für Rechtshänder*innen und einer Passform der HMDs, die Träger*innen großer Brillengestelle ausschließt. In einem Erfahrungsbericht merkt die Journalistin Adi Robertson zudem an, dass weibliche Körper tendenziell aus dem Design von VR-Headsets ausgeschlossen werden. So funktioniere Augentracking bei aufgetragener Wimperntusche schlechter¹⁸ und die Passform der Brillen sei auf größere, männliche Köpfe ausgelegt: »It's extremely intimate technology, but it's made for someone else« (Robertson 2016). Daran anschließen lassen sich Studien, die bei Frauen ein häufigeres Auftreten von Schwindel und

17 Diese Annahmen basieren auf Erfahrungen, die von der Autorin bei der Durchführung von Seminaren und Workshops mit VR-Technik gemacht wurden. Ein Blick in jene praxisorientierten Leitfäden zur Nutzung von VR im Unterricht, die im Rahmen des Projekts »Virtual Reality Technologie für das Lehren und Lernen in der digitalen Welt« in Nordrhein-Westfalen herausgegeben wurden, zeigt, dass in Informationen für Lehrpersonen Hinweise auf die Notwendigkeit einer didaktisch und pädagogisch sensibilisierten Begleitung jedoch häufig hinter technische Handhabungs- und Verwaltungshinweise zurückfallen (vgl. auch Bunnenberg et al. 2026).

18 Dieser Aspekt beschränkt sich natürlich nicht zwangsläufig auf weibliche Körper.

Übelkeit (»motion sickness«) während der VR-Rezeption feststellen und dies u.a. auf einen zu weiten Abstand der beiden Miniaturbildschirme in der VR-Brille zurückführen (vgl. Munafo/Diedrick/Stoffregen 2017; Stanney/Fidopiastis/Foster 2020). Eine intersektionale Perspektive wird durch Beobachtungen eröffnet, die daraufhinweisen, dass sich VR-Headsets bei vielen von Frauen getragenen Frisuren sowie bei bestimmten Haartypen und Haarstrukturen nicht gut am Kopf befestigen lassen, etwa bei sehr dichtem, lockigem Haar und damit insbesondere bei Afrohaaren (vgl. Davis 2021; Mboya 2020; Parente 2024, 16). Diese Befunde fügen sich in eine vor allem in den (schwarz-)feministischen Science and Technology Studies und den (Critical) Disability Studies geäußerte Kritik an fehlender Diversität in Testgruppen, Forschungssettings und Produktionsteams, in denen nach wie vor überwiegend männliche, weiße und *able-bodied*¹⁹ Perspektiven vertreten sind (vgl. Klimek 2025).

Besonders eindrücklich wird die fehlende hardwaretechnische sowie sensoralgorithmische Erkennung und damit auch fehlende Anerkennung bestimmter Körper und Bewegungen (»My VR does not acknowledge me as a person«, Eglis-ton/Clark/Carter 2023) beim oftmals als nächster Schritt zur »Naturalisierung« von Schnittstellen ausgerufenen Handtracking. Wie die Screenaufnahme eines Users zur Effektivität des Handtrackings der Oculus Quest bei seiner Hand mit einem Finger zeigt, glitcht die in VR dargestellte Hand zwischen der Darstellung halber Finger, angewinkelter Finger und ausgestreckter Finger diffus hin und her. Sein Resümee auf der Plattform Reddit fällt entsprechend ernüchternd aus: »I was excited to try out the hand tracking, not only because it's cool but to see how it would react to my hand. [...] Unfortunately, I don't think hand tracking will work for me because I lack length in my fingers« (Kommentar von dustindps, 2019). Die Reaktion einer Person, die sich selbst als Programmierer*in bezeichnet, bezeugt diese offensichtliche Leerstelle in Forschung und Entwicklung: »I'm glad you followed up

19 Mit der Zuschreibung *able-bodied* ist keine essentialistische Kategorie gemeint – ebenso wenig, wie die nachfolgenden Argumente zu medial bedingten Teilhabelimitierungen als technikdeterministische Positionierung missverstanden werden sollen. Mit den Worten Schillmeiers (2016, 291) lässt sich vielmehr sagen, dass technische Materialitäten von Medien der VR (und auch AR) in Resonanz mit den User*innen und ihren Körpern »ermöglichende/behindernde [Herv. i.O.] Akteur-Netzwerke« ausbilden. Gemäß einem relationalen Ansatz medialer Teilhabe werden In- und Exklusion sowie Nicht-/Teilhabe, un-/fähige oder in-/kompatible Subjekte oder Objekte nicht als gegebenes Apriori, sondern als Effekte wechselseitiger körperlicher und materieller Mediationsprozesse gedacht (vgl. Ochsner 2023; Spöhrer 2023). Im Kontext von VR werden Momente der Affordanzfindung, wie Dokumaci (2016, 278) jene körperlichen, handlungsmächtigen Praktiken von Menschen mit Behinderungen beschreibt, mit denen sie sich auf andere Arten und Weisen in soziotechnischen Gefügen unserer Welt bewegen, von Initiativen bzw. Projekten wie *Virtuell Barrierefrei*, *WalkingVR Driver* oder *Brecht into the Metaverse* ausgelotet.

with a comment, it's really important for tech types to hear that there is a massive gap between me thinking this is an amazing technical feat, versus your perspective which reads this as unusable« (lyth 2019). Ein Kommentar aus dem Jahr 2024 über die Meta Quest 2 zeigt, dass sich an diesen Umständen bisher nicht viel geändert hat: »I don't have a thumb on my left hand which causes hand tracking to not be so kind to me« (Total_Trash_Otter 2024). An dieser Stelle soll keine Diskussion darüber geführt werden, inwiefern der genaue Transfer der Körper von Menschen mit Behinderungen in die VR ebendiesen ein gewünschtes Identifikationspotenzial bietet oder ob es stattdessen ein ermächtigendes Moment eröffnet, im Virtuellen einen normierten *able-bodied* Avatar verkörpern zu können. Stattdessen sollte veranschaulicht werden, dass die in den sportbezogenen Diskursen geäußerte Vorstellung einer objektiven und verlustfreien Übertragung des Körpers in die VR die zugrundeliegenden, an einem standardisierten User*innenmodell ausgerichteten informatischen Modellierungen ignoriert, die bestimmte Körper(-Charakteristiken) voraussetzen und davon abweichende Körper behindern.

Auf ebensolchen normierten Körpermodellen basieren auch visuelle halb- oder ganzkörperliche Darstellungen und Bewegungsanimationen in VR, die anhand der sensorisch erfassten Position des Head-Mounted Displays und der Controller bzw. Hände interpoliert, also algorithmisch berechnet werden (geläufig ist dieses Verfahren unter dem Begriff »Inverse Kinematics«). Die Interpolation basiert auf einem standardisierten User*innenmodell, das beispielsweise keine Bewegungen von Rollstuhlnutzer*innen oder von Menschen mit Behinderungen enthält, die von »typischen« Bewegungen abweichen. Auch individuelle Bewegungsstile sind nicht oder nur schwer abbildbar. Wenn es in den Systemeinstellungen der Meta Quest Brillen also heißt, dass Dank Hand- und Bodytracking »ein nahtloseres und immersiveres Erlebnis« in der VR ermöglicht werde, so bleibt dieses Erlebnis einer bestimmten User*innenschaft verwehrt. Darüber hinaus ermöglichen Medien der VR durch multimodale und multisensorische Darstellungsmöglichkeiten zwar neue Rezeptionszugänge. Nichtsdestotrotz privilegieren gegenwärtige VR-Anwendungen und Nutzeroberflächen gemäß der in westlichen Kulturen verbreiteten visuellen, sozialen Ordnung (vgl. Schillmeier 2016, 295) den visuellen Sinn – was sich in dieser Arbeit in der okularzentrischen Auseinandersetzung mit der *blickpunkt*abhängigen *Bildraum*generierung und der entrahmten *Bildlichkeit* widerspiegelt. Entgegen dem im Diskurs deutlich gewordenen Werbeversprechen von einer natürlichen, intuitiven und uneingeschränkt partizipativen Nutzung (»VR ist barrierefrei und inklusiv«, Stiftung Lesen 2018, 8) ist der Umgang mit VR damit nicht voraussetzungslos und universell zugänglich.²⁰ Insbesondere beim Einsatz entsprechender Technologien in Bildungskontexten gilt es, die »spezifische, materiell verbaute

20 Für eine Zusammenschau diverser Ausschlüsse durch die Gestalt und Gestaltung von VR-Systemen für Menschen mit Behinderungen vgl. Gerling/Spiel 2021.

Limitierung und damit die implizierte körperliche Zurichtung« (Spöhrer 2023, 73) der standardisierten Ein- und Ausgabekomponenten und -modalitäten zu reflektieren: Für welche Körper(-Eigenschaften) sind die als ›natürlich‹ ausgewiesenen Schnittstellen schlussendlich ausgelegt? Wem ermöglichen sie unmittelbare und störungsfreie Zugänge zu virtuellen Welten und wen behindern sie strukturell?

»There is a small literature – mostly in public writing and the field of Human-Computer Interaction – about VR, accessibility and disability. Several of these papers note that VR, despite being framed as having a more ›natural‹ UI [User Interface] – in the sense that it relies on movements and gestures of the body – it is not necessarily something that is intuitive or natural to *all* [Herv. i.O.] bodies.« (Carter/Egliston 2020a, 12)

Darüber hinaus sind in die Programmierung von Interaktionsmöglichkeiten immer Vorannahmen über zielführende und inakzeptable, effektvolle oder ineffektive, richtige und falsche Bewegungsformen eingeschrieben, die im Bewegungsvollzug eingeübt und einverleibt werden: »Code und Algorithmus [legen] den Funktionsumfang im Sinne binärer Einhegungen fest: Was kann man als Nutzerin oder Nutzer mit dem Gerät oder der Technologie tun und was ist damit nicht machbar?« (Mismelius 2022, 97). Daran anknüpfend heißt es in einem Blogpost auf Meta, dass zur Reduktion optischer Störfaktoren und Latenzen Künstliche Intelligenz eingesetzt wird, um mögliche Kopf-, Hand und daraus abgeleitete Körperbewegungen der User*innen vorherzusagen und vorauszurechnen (vgl. Reality Labs 2019). In diese algorithmisch prognostizierten Bewegungen sind diskursive Wissensbestände und – so zeigen die bisherigen Befunde – Normvorstellung von Bewegungen und Körperfunktionalitäten eingeschrieben, die in den Datenpraktiken fortlaufend reproduziert werden:

»Das Wissen um Bewegungsabläufe, um die Haltungs- und Verhaltensweisen realer Körper, das in solchen Programmen verbaut ist, hat seinen materialen Ermöglichungsgrund in der Geschichte der anthropologischen Datenverarbeitung, in den Beobachtungsreihen und Erhebungstechniken, in den Tabellen und Zahlenkolonnen, in statistischen Auswertungen und ihren Mittelwertbildungen, im Ableiten normaler und pathologischer Bewegungen. Im anthropologischen Atlas wird versammelt und systematisiert, was es vom Körper zu wissen gibt. [...] Und damit gibt der Atlas auf seine sehr eigene Weise konkrete Antworten darauf, wie es um die Geometrie des Körpers oder genauer noch, wie es um die Geometrie von Durchschnittskörpern bestellt ist.« (Rieger 2019b, 108)

Vor diesem Hintergrund ließen sich z.B. Sportanwendungen im Schulsportkontext dahingehend befragen, inwiefern am Sportartenlernen orientierte Prinzipien ›richtiger‹ und ›falscher‹ Bewegungen inkorporiert, performativ verfestigt und damit

pädagogische Perspektiven des sich körperlich Ausdrückens und der Gestaltung von Bewegungen vernachlässigt werden.

Nicht zuletzt befragt Felicitas Macgilchrist aus einer bildungstheoretischen Perspektive die laut Marketing angestrebte Reibungslosigkeit – die, wie deutlich wurde, aufgrund ihrer Schein-Intuitivität und Exklusivität keine ist – als (technisches) Gestaltungsideal eines Nicht-irritiert-Seins. Sind es doch die Irritationsmomente, die Reflexions- und Bildungsanlässe anstoßen können und die auch aus medienwissenschaftlicher Perspektive einer Medienreflexion zuträglich sind: »Sie [die Reibungslosigkeit] steht im Kontrast zu klassischen Zielen des Pädagogischen, wie die produktive Irritation, die Begegnung mit radikal Neuem« (Macgilchrist 2023, 62).

6.3 Kurze Schlussbetrachtung: Vermittelte Blickkontrolle

Es wurde aufgezeigt, wie die Hard- und Software von Medien der AR und VR und damit auch die ihrem Funktionieren zugrundeliegenden Algorithmen sowie Datenprozessierungen nicht nur Rezeptionsästhetiken, Teilhabe und Inhaltsgestaltung mitorganisieren, sondern auch Subjekte und deren Körper auf spezifische Weise positionieren. Grundlegend problematisiert wurde dabei, dass das ästhetische Verschwinden eines Rahmens als Marker für ein Außen des sichtbaren Bildfeldes (*hors-camp*) sowohl auf diskursiver als auch auf technischer Ebene mit einer Unsichtbarmachung des produktionsbezogenen, medienmateriellen Außen bzw. Dazwischen (*hors-cadre*) einhergeht. Entgegen der Vorstellung eines entrahmten, von fremdbestimmter Determiniertheit losgelösten und unmittelbaren Bildzugriffs bzw. Blicks ist dieser indes durchsetzt von Machtordnungen sowie soziotechnisch formatierten Kontrollregimen (»regime of control«, Kaerlein/Köhler 2018, 185). In diesem Sinne deutet auch Schröter die vermeintlich selbstbestimmte Blickwahl in eine gänzlich gegenteilige Richtung um:

»Ein perspektivisches Bild, welches sich mit der Kopfbewegung des Betrachters verändert [...], macht es dem Betrachter – im Unterschied zur Betrachtung statischer, perspektivischer Bilder – unmöglich, *nicht* [Herv. i.O.] den vorgesehenen Augenpunkt einzunehmen. Es gibt keine Möglichkeit, den von der apparativen Anordnung für das Subjekt vorgesehenen Platz gegenüber dem Bild zu verlassen [...]« (Schröter 2004, 192)

Anhand des 2024 erschienenen Head-Mounted Displays Apple Vision Pro wird diese Unfreiheit des Blicks umso eindrücklicher. Das Eyetracking des HMDs dient zum einen dazu, die Grafikleistung der computergenerierten Bilder dahingehend zu optimieren, dass nur jene Bereiche gerendert und scharfgestellt werden, die von den Rezipient*innen optisch fokussiert werden (»Perceptually Adaptive Graphics«). Zum

anderen fungiert der Blick als zentrales Steuerungselement, indem virtuelle Objekte fokussiert und mit einer pincettenartigen Fingergeste gegriffen und manipuliert werden. Es braucht daher den dauerhaften Blickkontakt, um mit den anvisierten Objekten umgehen und sich durch die virtuelle Welt gestisch navigieren zu können. Qua Interaktionsdesign werden die Augenbewegungen daher per Eyetracking dauerüberwacht.²¹ Der Journalist Nilay Patel stellt in seinem Review zur Apple Vision Pro auf dem Online-Portal *The Verge* heraus, dass diese permanente Überwachung des Blicks im Zusammenhang mit lizenzierten Inhalten zur Verhinderung von Screen-Aufnahmen führen kann:

»And unlike any other TV in your life, the Vision Pro can literally DRM [Digital Rights Management] your eyes — if you're watching a movie in the Apple TV app or Disney Plus and go to take a screen capture, the content blacks out. It's strange to experience a reality where big companies can block you from capturing what you see, even if all you're trying to do is show people how cool it looks in a review.«
(Patel 2024)

Sutherlands Überlegungen zum »head-mounted three-dimensional display« sind untrennbar mit seinem 1966 veröffentlichten Kongressbeitrag »The Ultimate Display« verknüpft. In diesem entwirft er die Vision eines ultimativen Displays, das auf einer verlustfreien, datentechnischen Rückkopplung der Mensch-Computer-Interaktion sowie einer vollständigen Realitätsillusion basiert. In dem Beitrag prognostizierte Sutherland bereits den Entwicklungsschritt hin zur Nutzung des menschlichen Blicks als Steuerungs- und Kontrollinstrument des Computers: »Machines to sense and interpret eye motion data can and will be built. It remains to be seen if we can use a language of glances to control a computer« (Sutherland 1966, 507). Anhand der Apple Vision Pro kann nun ernüchternd festgestellt werden, dass der den Computer kontrollierende Blick durch das ultimative Display vielmehr in eine ultimative Kontrolle des menschlichen Blicks gekippt ist. Ist es vor diesem Hintergrund nicht gerade umso perfider, unter der Maxime von Selbstbestimmtheit und Kontrolle dennoch – im Sinne Foucaults (1992 [1990], 52) – dermaßen regiert zu werden?

21 Anknüpfen ließe sich hier an die von Macele und Müggenburg nachgezeichnete Medienkulturgegeschichte der Blickerfassung und Augensteuerung. So wurde ausgehend von der psychologischen, behavioristisch ausgerichteten Verhaltensforschung im späten 19. Jahrhundert das Eyetracking zur barrierefreien Steuerungsmöglichkeit für die Nutzung von Computern und darauf aufbauend zur Optimierung von Rechenkapazitäten und Grafikleistungen in der Spielebranche weiterentwickelt: »The eyes were no longer conceptualized as a pure object of study but as human organs of control and eventually: play« (Macele/Müggenburg 2024, 118). Im Anschluss an die Diskursanalyse über den Einsatz von VR im Sport sei zudem auf die Anmerkung der Autoren zum prominenten Einsatz von Eyetracking als Analysetool im Sport verwiesen (vgl. ebd., 135).