BACHELORARBEIT

Der gestalterische Mehrwert der visuellen Ebene im Videopodcast

Entwicklung und Reflexion
des Interviewformats „Sidequest Social Media“

Name: Simon Zamorano Kranz
Matrikelnummer: 15490009
Hochschule: Technische Hochschule Ostwestfalen-Lippe
Fachbereich: Medienproduktion

Erstprüfer: Prof. Dr. phil. Frank Lechtenberg
Zweitprüfer: Dipl.-Kam. Dominik Moos
Abgabedatum: 17.08.26

Inhaltsverzeichnis

1. Einleitung **4**

2. Der Videopodcast als audiovisuelle Erweiterung des Audiopodcasts **7**

2.1 Der Audiopodcast als auditives Format 7

2.2 Der Videopodcast als audiovisuelle Erweiterung 9

2.3 Sichtbare Kommunikation im Interviewformat 11

2.4 Videomaterial als Grundlage für Social-Media-Clips 13

3. Gestalterische Anforderungen an einen Videopodcast im Interviewformat **14**

3.1 Bildgestaltung und Kameraperspektiven 14

3.2 Lichtgestaltung und Atmosphäre 16

3.3 Setgestaltung und visuelle Wiedererkennbarkeit 18

3.4 Personendarstellung und nonverbale Präsenz 20

3.5 Schnitt und visuelle Dynamik 21

3.6 Grafische Gestaltung und Formatidentität 23

3.7 Gestalterische Anforderungen an Social-Media-Clips 25

4. Entwicklung und Produktion der eigenen Videopodcast-Staffel **27**

4.1 Praxisbasierter Ansatz der Arbeit 27

4.2 Gestalterische Referenzanalyse ausgewählter Videopodcast-Formate 29

4.3 Ableitung der Gestaltungskriterien 38

4.4 Formatidee und Zielgruppe 41

4.5 Inhaltlicher Rahmen: Social Media 43

4.6 Visuelles Konzept 44

4.6.1 Grundidee des visuellen Konzepts 44

4.6.2 Setgestaltung und variable Gast-Elemente 45

4.6.3 Lichtgestaltung und Atmosphäre 46

4.6.4 Kameraperspektiven und Bildkomposition 46

4.6.5 Grafische Gestaltung und visuelle Formatidentität 50

4.6.6 Vorbereitung der Social-Media-Verwertung 50

4.7 Technisches Setup 51

4.8 Aufnahmeprozess 53

5. Reflexion und Fazit **59**

5.1 Zusammenfassung und Einordnung der zentralen Ergebnisse 59

5.2 Mehrwert der visuellen Ebene im eigenen Format 61

5.3 Reflexion der gestalterischen und technischen Umsetzung 62

5.4 Social-Media-Clips und Weiterverwertung 65

5.5 Grenzen der Arbeit 67
5.6 Beantwortung der Forschungsfrage 69
5.7 Ausblick 71

Literaturverzeichnis **73**

Abbildungsverzeichnis **78**

1. Einleitung

Die zunehmende Relevanz von Videopodcasts zeigt sich unter anderem an aktuellen Entwicklungen auf großen Podcastplattformen. Spotify gab im Juni 2023 bekannt, dass auf der Plattform mehr als 100.000 Video-Podcasts verfügbar waren. Ein Jahr später lag diese Zahl bereits bei mehr als 250.000 Video-Podcast-Shows (Spotify, 2023; Spotify, 2024). Diese Entwicklung verdeutlicht, dass Podcasts zunehmend nicht mehr ausschließlich als reine Audioformate gedacht werden, sondern verstärkt audiovisuelle Formen annehmen.

Podcasts haben sich als eigenständiges Medienformat etabliert und werden längst nicht mehr nur über ihre auditive Ebene verstanden. Während klassische Audiopodcasts vor allem über Stimme, Sprache, Klang und Gesprächsdynamik funktionieren, erweitert der Videopodcast dieses Format um eine sichtbare Ebene. Dadurch verändert sich das Format grundlegend: Ein Gespräch wird nicht mehr nur hörbar, sondern auch sichtbar. Zusätzlich treten Aspekte wie Kameraperspektiven, Lichtgestaltung, Set, Mimik, Gestik, Körperhaltung, Schnitt und visuelle Wiedererkennbarkeit in den Vordergrund.

Besonders bei Interviewformaten kann diese visuelle Ebene eine wichtige Rolle spielen. Während ein Audiopodcast allein über das Gesagte und die stimmliche Wirkung der Gesprächspartner funktioniert, macht ein Videopodcast weitere Ebenen der Kommunikation sichtbar. Reaktionen, Blickrichtungen, Gesten, räumliche Anordnung und die Atmosphäre des Sets können Teil der Gesamtwirkung werden. Die visuelle Ebene erweitert das Gespräch dadurch um gestalterische Möglichkeiten, bringt aber zugleich neue Anforderungen an die Produktion mit sich. Ein Videopodcast erzeugt nicht automatisch einen Mehrwert, nur weil eine Kamera mitläuft. Vielmehr muss die visuelle Gestaltung bewusst geplant und umgesetzt werden, damit sie das Gespräch unterstützt und nicht vom Inhalt ablenkt.

Die Relevanz des Themas zeigt sich außerdem in der Verbindung von Videopodcasts und Social Media. Gerade für Plattformen, auf denen kurze, schnell verständliche und visuell ansprechende Inhalte eine große Rolle spielen, kann das aufgenommene Videomaterial eine zusätzliche Nutzungsebene eröffnen. Der Videopodcast ist damit nicht nur eine audiovisuelle Erweiterung des Audiopodcasts, sondern auch ein Format, das stärker in plattformübergreifende Veröffentlichungsstrategien eingebunden werden kann.

Die persönliche Motivation für diese Arbeit entstand aus der vorherigen Auseinandersetzung mit Audioformaten im Studium. Im Rahmen eines Medienprojekts beim Uni-Radio wurde bereits ein Audiopodcast mit gestaltet. Diese Erfahrung führte zu der Frage, wie sich ein Podcast verändern würde, wenn er nicht nur auditiv, sondern auch visuell gestaltet wird. Daraus entstand das Interesse, einen eigenen Videopodcast zu entwickeln und zu untersuchen, welchen gestalterischen Mehrwert die sichtbare Ebene in einem Interviewformat bieten kann.

Thematisch grenzt sich die Arbeit bewusst ein. Sie untersucht nicht die allgemeine Geschichte des Podcasts, keine umfassende Marktanalyse von Podcastplattformen und keine empirische Wirkungsmessung beim Publikum. Auch die journalistische Gesprächsführung steht nicht im Mittelpunkt. Der Fokus liegt stattdessen auf dem gestalterischen Mehrwert der visuellen Ebene in einem Videopodcast im Interviewformat. Betrachtet werden vor allem Kamera, Licht, Setgestaltung, sichtbare Kommunikation, Schnitt, grafische Gestaltung und die spätere Aufbereitung des Materials für Social-Media-Clips.

Aus dieser Eingrenzung ergibt sich folgende Forschungsfrage:

Welchen gestalterischen Mehrwert bietet die visuelle Ebene eines Videopodcasts im Interviewformat im Vergleich zu einem reinen Audiopodcast?

Unter gestalterischem Mehrwert wird in dieser Arbeit kein messbarer Rezeptionserfolg im Sinne von Reichweite, Aufmerksamkeit oder Zuschauerbewertung verstanden. Gemeint ist vielmehr der zusätzliche funktionale Beitrag, den die visuelle Ebene gegenüber einem reinen Audiopodcast leisten kann. Ein gestalterischer Mehrwert liegt dann vor, wenn Bild, Raum, Licht, Set, sichtbare Kommunikation, Schnitt oder grafische Gestaltung das Gespräch nicht nur abbilden, sondern es erweitern, verständlicher machen, atmosphärisch einbetten, Personen sichtbarer machen, Wiedererkennbarkeit schaffen oder eine weitere Nutzung des Materials ermöglichen.

Der Begriff wird damit als gestalterische Arbeitskategorie verstanden. Bewertet wird nicht, ob ein Videopodcast grundsätzlich besser ist als ein Audiopodcast, sondern wie die sichtbare Ebene eine erkennbare Funktion übernimmt, die im reinen Audioformat nicht oder nur eingeschränkt vorhanden wäre. Ziel der Arbeit ist es, herauszuarbeiten, welche zusätzlichen gestalterischen Möglichkeiten durch die visuelle Ebene entstehen und wie diese in einer eigenen Produktion praktisch umgesetzt werden können Vielmehr geht es darum, zu

untersuchen, unter welchen Bedingungen die visuelle Ebene einen Mehrwert erzeugen kann und welche gestalterischen Anforderungen daraus entstehen.

Methodisch folgt die Arbeit einem theoriegestützten, gestalterisch-praktischen Vorgehen. Im ersten Schritt werden auf Grundlage bestehender Literatur zentrale Funktionen und Anforderungen der visuellen Ebene eines Videopodcasts herausgearbeitet. Dazu gehören insbesondere Bildgestaltung, Kameraperspektiven, Licht, Setgestaltung, nonverbale Präsenz, Schnitt, grafische Gestaltung und die Aufbereitung für Social-Media-Clips.

Im zweiten Schritt werden diese theoretisch herausgearbeiteten Gestaltungsbereiche durch eine exemplarische Referenzanalyse ausgewählter Videopodcast-Formate konkretisiert. Die Auswahl erfolgt gezielt anhand einzelner gestalterischer Merkmale, die für die Entwicklung des eigenen Formats relevant sind. Untersucht werden ausgewählte Sequenzen hinsichtlich Kamera und Bildgestaltung, Licht und Atmosphäre, Setgestaltung, Personendarstellung, Schnitt, grafischer Gestaltung, visueller Zusatzinhalte und Social-Media-Verwertung. Die Referenzanalyse stellt keine repräsentative Marktanalyse dar, sondern dient dazu, konkrete gestalterische Lösungsansätze sichtbar zu machen und die theoretischen Erkenntnisse an Praxisbeispielen zu ergänzen.

Im dritten Schritt werden die Erkenntnisse aus Literatur und Referenzanalyse zu konkreten Gestaltungskriterien für die eigene Videopodcast-Produktion zusammengeführt. Diese Kriterien bilden die Grundlage für die Entwicklung und Produktion der fünfteiligen Staffel “Sidequest Social Media”. Abschließend wird reflektiert, inwieweit sich die Kriterien praktisch umsetzen ließen, welche gestalterischen Funktionen die visuelle Ebene innerhalb des eigenen Formats übernahm und welche Grenzen und Herausforderungen dabei sichtbar wurden.

Die Arbeit verfolgt dabei keinen empirischen Wirkungsvergleich zwischen Audio- und Videopodcast. Es werden weder Zuschauerinnen und Zuschauer befragt noch unterschiedliche Rezeptionsvarianten experimentell miteinander verglichen. Aussagen über den gestalterischen Mehrwert beziehen sich deshalb auf die theoretische Auseinandersetzung, die exemplarische Gestaltungsanalyse und die Reflexion der eigenen Produktion. Auf dieser Grundlage wurde eine eigene Videopodcast-Staffel mit dem Titel “Sidequest Social Media” entwickelt und produziert. Die zuvor erarbeiteten Kriterien wurden dabei auf ein eigenes Interviewformat übertragen. Im Mittelpunkt standen fünf Gespräche mit Gästen, die unterschiedliche Perspektiven auf Social Media, Content Creation, Streaming, Foto- und Videoproduktion

sowie Influencer-Management einbringen. Die praktische Umsetzung dient dazu, die theoretischen und gestalterischen Überlegungen im Produktionsprozess anzuwenden und anschließend zu reflektieren.

Die Arbeit gliedert sich in fünf Kapitel. Nach der Einleitung behandelt Kapitel zwei den Videopodcast als Erweiterung des Audiopodcasts. Dabei wird zunächst der Audiopodcast als auditives Format betrachtet, bevor anschließend der Videopodcast als audiovisuelle Erweiterung beschrieben wird. Außerdem geht es um sichtbare Kommunikation im Interviewformat und um Videomaterial als Grundlage für Social-Media-Clips.

Kapitel drei beschäftigt sich mit den gestalterischen Anforderungen an einen Videopodcast im Interviewformat. Im Mittelpunkt stehen Bildgestaltung und Kameraperspektiven, Lichtgestaltung, Setgestaltung, Personendarstellung, Schnitt, grafische Gestaltung und Anforderungen an Social-Media-Clips. Dieses Kapitel bildet die theoretisch-gestalterische Grundlage für den praktischen Teil der Arbeit.

Kapitel vier widmet sich der Entwicklung und Produktion der eigenen Videopodcast-Staffel “Sidequest Social Media”. Zunächst wird der praxisbasierte Ansatz beschrieben. Anschließend folgt eine gestalterische Referenzanalyse ausgewählter Formate. Aus dieser Analyse werden Kriterien für das eigene Format abgeleitet, bevor Formatidee, Zielgruppe, inhaltlicher Rahmen, visuelles Konzept, technisches Setup, Aufnahmeprozess sowie Schnitt und Clip-Erstellung dargestellt werden.

Kapitel fünf fasst die zentralen Ergebnisse der Arbeit zusammen und reflektiert die praktische Umsetzung. Dabei wird betrachtet, welchen gestalterischen Mehrwert die visuelle Ebene im eigenen Videopodcast erzeugen konnte, welche Herausforderungen während der Produktion sichtbar wurden und welche Grenzen sich im Vergleich zu einem reinen Audiopodcast ergeben haben.

2. Der Videopodcast als audiovisuelle Erweiterung des Audiopodcasts

2.1 Der Audiopodcast als auditives Format

Der klassische Audiopodcast ist ein rein auditives Format. Seine Wirkung entsteht über Stimme, Sprache, Klang, Gesprächsrhythmus und Atmosphäre. Gerade im Interviewformat

folgen Hörerinnen und Hörer dem Gespräch ausschließlich über das Hören. Nähe, Anschaulichkeit und Aufmerksamkeit müssen deshalb mit akustischen Mitteln erzeugt werden.

Eine zentrale Stärke des Audiopodcasts liegt in der unmittelbaren Wirkung der Stimme. Klenk beschreibt, dass die Stimme „immer schon ein Mehr“ transportiert (Klenk, 2020, S. 4). Sie vermittelt also nicht nur den sachlichen Inhalt einer Aussage, sondern auch Stimmung, Persönlichkeit, Haltung und körperliche Präsenz. Im Audiopodcast wird die Stimme damit zum wichtigsten Gestaltungsmittel. Sie ersetzt keine sichtbare Ebene, erzeugt aber eine eigene Form von Nähe und Wahrnehmung.

Diese Nähe wird durch die typische Hörsituation unterstützt. Klenk beschreibt die Kopfhörersituation als intim und betont, dass sie ein intensives und genaues Hören ermöglicht (Klenk, 2020, S. 19). Stimmen können dadurch sehr direkt und nah wahrgenommen werden. Auch Knobloch hebt die besondere Nähe des Podcastformats hervor und beschreibt Podcasts als besonders wirksam, wenn sie dialogorientiert und intim bleiben, statt vollständig vorgeschrieben zu sein (Knobloch et al., 2020, S.13). Für Interviewformate ist deshalb nicht nur der Informationsgehalt entscheidend, sondern auch der Eindruck eines echten Gesprächs.

Die Reduktion auf das Hören ist somit nicht nur eine Einschränkung, sondern auch Teil der spezifischen Stärke des Audiopodcasts. Knobloch beschreibt die Nähe zwischen erzählender Person und Publikum im Podcast als besonders intim (Knobloch et al., 2020, S.13). Der Audiopodcast kann gerade durch diese Konzentration auf Stimme, Klang und Gespräch eine persönliche Wirkung entfalten, ohne visuelle Präsenz zu benötigen.

Gleichzeitig stellt das reine Audioformat besondere Anforderungen an die Gestaltung. Räume, Situationen und Stimmungen müssen über Sprache, Geräusche und Atmosphäre vermittelt werden. Preger macht am Beispiel dokumentarischen Erzählens deutlich, dass bereits einzelne akustische Hinweise eine räumliche oder situative Vorstellung bei Hörerinnen und Hörern auslösen können (vgl. Preger, 2020, S. 9). Audio kann dadurch bildhaft wirken, obwohl tatsächlich kein sichtbares Bild vorhanden ist.

Besonders Dialoge sind dafür wichtig. Preger beschreibt, dass Stimme und Dialog Charakter sichtbar beziehungsweise erfahrbar machen und Hörerinnen an Figuren einer Geschichte binden können (Preger, 2020, S. 9). Im Interviewformat werden Gesprächspartnerinnen und

Gesprächspartner deshalb vor allem über ihre Stimme, ihre Aussagen, ihren Sprechrhythmus und ihre Art zu reagieren wahrgenommen. Charakter, Haltung und Nähe entstehen akustisch.

Hinzu kommt die besondere Authentizität des Formats. Preger betont, dass Audio aufgrund seiner Intimität keine künstliche Dramatisierung oder Pseudonähe benötigt (Preger, 2020, S. 12). Ein Audiopodcast kann Nähe bereits durch eine glaubwürdige Gesprächssituation, durch Stimme und durch den Eindruck von Echtheit erzeugen.

Damit besitzt der Audiopodcast eine eigene mediale Qualität. Er wirkt über Stimme, Klang, Dialog und Vorstellungskraft. Zugleich bleiben bestimmte Informationen unsichtbar: Mimik, Gestik, Körperhaltung, Blickkontakt, Raumwirkung und sichtbare Reaktionen können im reinen Audioformat nicht unmittelbar wahrgenommen werden. Sie müssen entweder durch Stimme, Geräusche oder Sprache ersetzt werden oder bleiben für das Publikum offen.

Diese Begrenzung des reinen Audioformats bildet den Ausgangspunkt für die weitere Untersuchung. In den folgenden Kapiteln wird der Videopodcast deshalb als audiovisuelle Erweiterung des Audiopodcasts betrachtet. Im Mittelpunkt steht dabei die Frage, welche zusätzlichen gestalterischen Funktionen durch die sichtbare Ebene im Interviewformat entstehen können.

2.2 Der Videopodcast als audiovisuelle Erweiterung

Der Videopodcast erweitert das auditive Podcastformat um eine visuelle Ebene. Das Gespräch wird nicht mehr nur gehört, sondern auch gesehen. Damit verändert sich die Wahrnehmung des Formats: Stimme, Sprache und Klang bleiben zwar zentral, werden aber durch Bild, Raum, sichtbare Präsenz und visuelle Gestaltung ergänzt.

Eine grundlegende Einordnung findet sich bei Sutton-Brady et al. Sie unterscheiden zwischen „audio only, enhanced (with images, links and meta information) or vodcasts (video-on-demand podcasts)“ (Sutton-Brady et al., 2009, S. 220). Der Videopodcast kann damit als Podcastform verstanden werden, die über das rein Auditive hinausgeht, ohne den Bezug zum Podcastformat zu verlieren.

Wichtig ist dabei, den Videopodcast nicht als Ersatz des Audiopodcasts zu verstehen. Das auditive Fundament bleibt bestehen: Gespräch, Stimme, Rhythmus und Klang tragen weiterhin die inhaltliche Struktur. Die visuelle Ebene kommt ergänzend hinzu. Preger beschreibt in

Bezug auf Storytelling, dass neue Darstellungsformen bestehende Formen nicht ersetzen, sondern ergänzen können (Preger, 2020, S. 211). Dieser Gedanke lässt sich auf den Videopodcast übertragen: Die visuelle Ebene verdrängt den Audiopodcast nicht, sondern erweitert ihn um zusätzliche Wahrnehmungs- und Gestaltungsmöglichkeiten.

Diese Erweiterung besteht vor allem darin, dass eine zusätzliche Informationsebene entsteht. Sutton-Brady et al. zeigen im Bildungskontext, dass visuelle oder metabezogene Informationen die reine Tonspur ergänzen und unterschiedliche Rezeptionsweisen ansprechen können (Sutton-Brady et al., 2009, S. 220–221). Auch wenn sich diese Quelle nicht auf Videopodcast-Interviews bezieht, macht sie grundsätzlich deutlich, dass Video zusätzliche Informationen bereitstellen kann. Auf Videopodcasts übertragen heißt das: Das Publikum hört nicht nur, was gesagt wird, sondern sieht zusätzlich, wer spricht, wie gesprochen wird und in welcher Umgebung das Gespräch stattfindet.

Dadurch entstehen neue gestalterische Möglichkeiten. Personen, Räume, Blickrichtungen, Reaktionen und Atmosphäre können sichtbar werden. Gleichzeitig erhöht sich die Komplexität des Formats, weil Bildgestaltung, Raumwirkung, technische Umsetzung und dramaturgische Entscheidungen zusätzlich berücksichtigt werden müssen. Knobloch et al. weisen darauf hin, dass Bewegtbild-Erzählungen dramaturgisches Know-how erfordern (Knobloch et al., 2020, S. 27). Für Videopodcasts bedeutet das vorsichtig übertragen: Das Format wird nicht nur technisch, sondern auch gestalterisch anspruchsvoller.

Auch technische und räumliche Bedingungen beeinflussen die Wahrnehmung eines Gesprächs. Klenk beschreibt, dass technische Settings die körperliche Wahrnehmung des Sprechens mitprägen können, etwa durch Mikrofon, Kopfhörer, Sitzposition, Körperhaltung, Gegenüber und Raumklang (Klenk, 2020, S. 423). Im Videopodcast werden solche Bedingungen nicht nur akustisch, sondern auch visuell erfahrbar. Raum, Körperhaltung und Gesprächsanordnung werden Teil der sichtbaren Wirkung.

Die visuelle Ebene erzeugt ihren Mehrwert jedoch nicht automatisch. Eine mitlaufende Kamera reicht nicht aus, wenn das Bild das Gespräch nicht sinnvoll unterstützt. Klenk beschreibt, dass neue mediale Präsentationsformen erst entwickelt werden müssen, und zwar „über ein Ausprobieren und Erweitern von Möglichkeiten“ (Klenk, 2020, S. 442). Auch der Videopodcast benötigt daher bewusste gestalterische Entscheidungen, damit Bild und Ton zusammenwirken.

Zugleich besteht die Gefahr, dass das Visuelle zu dominant wird. Klenk beschreibt, dass visuelle Medien die Stimme in den Hintergrund drängen können, wenn das bildliche Zeigen stärker in den Vordergrund tritt (vgl. Knobloch et al., 2020, S. 60). Rime et al. beschreiben ähnlich, dass Podcasts, die zu stark auf visuelle Informationen angewiesen sind, ihren Audiofokus verlieren und „predominantly visual“ werden können (Rime et al., 2022, S. 15). Der Mehrwert des Videopodcasts liegt deshalb nicht darin, das Audioformat vollständig in ein Videoformat zu verwandeln. Er entsteht dort, wo die visuelle Ebene das Gespräch unterstützt, erweitert und vertieft, ohne den auditiven Kern zu verdrängen.

Der Videopodcast ist damit eine audiovisuelle Erweiterung des Audiopodcasts. Er übernimmt dessen auditive Stärken und ergänzt sie um Bild, Raum, sichtbare Präsenz und visuelle Gestaltung. Entscheidend ist nicht das bloße Vorhandensein von Video, sondern dessen konkrete gestalterische Funktion.

2.3 Sichtbare Kommunikation im Interviewformat

Für Interviewformate ist die sichtbare Ebene besonders relevant, weil Gespräche nicht nur aus gesprochenen Aussagen bestehen. Während im Audiopodcast Stimme, Sprache und Klang die Wahrnehmung der Gesprächspartnerinnen und Gesprächspartner prägen, können im Videopodcast zusätzlich Mimik, Gestik, Blickrichtung, Körperhaltung und Reaktionen sichtbar werden. Damit erweitert sich die Wahrnehmung der Gesprächssituation um eine nonverbale und körperliche Ebene.

Eine zentrale Rolle spielt nonverbale Kommunikation. Croes et al. beschreiben nonverbale Hinweise als wesentlich für Beziehungsaufbau und Eindrucksbildung (Croes et al., 2019, S. 1212). Im Interviewformat werden Gesprächspartnerinnen und Gesprächspartner also nicht nur über ihre Aussagen wahrgenommen, sondern auch über ihr Auftreten, ihre Reaktionen und ihre sichtbare Beteiligung am Gespräch. Genau diese Ebene bleibt im reinen Audiopodcast weitgehend verborgen.

Zu den relevanten Signalen gehören Blickkontakt, Gesichtsausdruck, Gestik, Körperhaltung und sichtbare Aufmerksamkeit. Furlan beschreibt Blickkontakt und Gesichtsausdruck als wichtige Hinweise für Unmittelbarkeit (Furlan, 2021, S. 14). Im Videopodcast können solche Signale dazu beitragen, dass ein Gespräch persönlicher und präsenter wirkt. Das Publikum sieht nicht nur, wer spricht, sondern auch, wie eine Aussage körperlich begleitet wird.

Gerade sichtbare Reaktionen können die Wirkung eines Interviews erweitern. Nicken, Lächeln, Nachdenklichkeit oder Überraschung geben Hinweise darauf, wie ein Gesprächsmoment aufgenommen wird. Croes et al. verweisen darauf, dass aktive Gesprächsteilnahme mit körperlicher und stimmlicher Ausdruckskraft sowie Energie verbunden ist (Croes et al., 2019, S. 1213). Der Videopodcast kann diese Dynamik zusätzlich sichtbar machen.

Auch Gestik und Sprache sind eng miteinander verbunden. Furlan beschreibt, dass Sprache und Gesten gemeinsam zur Bedeutung einer Aussage beitragen (Furlan, 2021, S. 57). Eine Aussage wirkt demnach nicht nur durch ihren Wortlaut, sondern auch durch die körperliche Begleitung. Im Audiopodcast bleibt diese gestische Ebene unsichtbar; im Videopodcast kann sie Teil der Gesamtwirkung werden.

Dadurch verändert sich die Wahrnehmung der Gesprächspartnerinnen und Gesprächspartner. Sie erscheinen nicht nur als Stimmen, sondern als sichtbare Personen in einer konkreten Gesprächssituation. Blake beschreibt, dass gemeinsame Bedeutung durch das Zusammenspiel verbaler, nonverbaler, körperlicher und räumlicher Prozesse entsteht (Blake, 2026, S. 3). Auch wenn diese Quelle sich auf Video-Call-Technologie bezieht, lässt sich der grundlegende Zusammenhang auf Videopodcast-Interviews übertragen: Bedeutung entsteht nicht allein durch Sprache, sondern auch durch sichtbare und räumliche Aspekte der Kommunikation.

Furlan weist zudem darauf hin, dass videobasierte Gesprächssituationen nonverbale Hinweise wie Blick, Gesichtsausdruck, Aufmerksamkeit und einige Gesten übertragen können (Furlan, 2021, S. 18). Diese Aussage bezieht sich auf Videointerviews, ist aber für Videopodcasts anschlussfähig: Auch hier ermöglicht die Kamera, dass nonverbale Signale wahrnehmbar werden und die Gesprächswirkung erweitern.

Diese sichtbare Ebene darf jedoch nicht als eindeutig lesbarer Wahrheitsbeweis verstanden werden. Furlan betont, dass derselbe Gesichtsausdruck je nach Kontext unterschiedlich gedeutet werden kann (Furlan, 2021, S. 55). Mimik, Gestik und Körperhaltung eröffnen also eine zusätzliche Bedeutungsebene, bleiben aber interpretierbar und situationsabhängig.

Damit erklärt Kapitel 2.3 vor allem, warum sichtbare Kommunikation einen Mehrwert gegenüber reinem Audio darstellt: Der Videopodcast macht nonverbale Signale, Präsenz, Reaktionen und Interaktion sichtbar. Wie diese Ebene gestalterisch sichtbar gemacht werden muss, wird erst in Kapitel 3.4 behandelt.

2.4 Videomaterial als Grundlage für Social-Media-Clips

Ein zusätzlicher Nutzen des Videopodcasts liegt darin, dass bei der Produktion nicht nur Ton, sondern auch Bildmaterial entsteht. Dieses Material kann über die vollständige Podcastfolge hinaus weiterverwendet werden. Während ein Audiopodcast für visuelle Plattformen erst nachträglich bebildert oder grafisch ergänzt werden müsste, enthält ein Videopodcast bereits sichtbare Gesprächssituationen, Personen, Reaktionen und räumliche Eindrücke.

Damit entsteht aus einer Videopodcast-Produktion nicht nur ein einzelnes Langformat, sondern zugleich eine Grundlage für weitere audiovisuelle Inhalte. Aus dem vorhandenen Material können kurze Ausschnitte, Teaser oder Social-Media-Clips entwickelt werden. Der zusätzliche Nutzwert besteht darin, dass das aufgenommene Gespräch nicht nur hörbar, sondern auch sichtbar weiterverarbeitet werden kann. Host, Gast, Mimik, Gestik, Set und Gesprächsdynamik stehen als visuelle Bestandteile bereits zur Verfügung. Als Host wird im Folgenden die moderierende beziehungsweise gesprächsführende Person des Podcastformats bezeichnet.

Violot et al. beschreiben Kurzformate als Möglichkeit, Publikum mit prägnanten und attraktiven Inhalten zu erreichen (Violot et al., 2024, S. 2). Für Videopodcasts zeigt dieser Gedanke, dass vorhandenes Videomaterial grundsätzlich einen zusätzlichen Verwertungsraum eröffnet.

Auch Zhu et al. betonen, dass Video durch visuelle Anziehungskraft und die Fähigkeit, komplexe Inhalte zu veranschaulichen, wirksam sein kann (Zhu et al., 2023, S. 1). Da sich diese Quelle auf Gesundheitskommunikation bezieht, ist die Übertragung auf Videopodcasts vorsichtig zu verstehen. Der grundlegende Zusammenhang bleibt dennoch relevant: Sichtbare Personen, Reaktionen und visuelle Elemente können gesprochene Inhalte ergänzen und anschaulicher machen.

Der Mehrwert liegt damit nicht nur in der Veröffentlichung einer Podcastfolge als Video, sondern auch in der Weiterverwendbarkeit des Materials. Eine Aufnahme kann als vollständige Folge, kurzer Ausschnitt, thematischer Teaser oder visuelles Zitat genutzt werden. Das vorhandene Bildmaterial schafft jedoch nur die Grundlage. Welche gestalterischen Anforderungen Social-Media-Clips erfüllen müssen, wird in Kapitel 3.7 behandelt.

3. Gestalterische Anforderungen an einen Videopodcast im Interviewformat

Nachdem Kapitel 2 den Mehrwert der visuellen Ebene gegenüber dem reinen Audiopodcast beschrieben hat, geht es nun um die konkrete Gestaltung dieser Ebene. Bild, Licht, Set, Personendarstellung, Schnitt und grafische Elemente erzeugen ihren Mehrwert nicht automatisch. Sie müssen so eingesetzt werden, dass sie das Gespräch unterstützen, Orientierung schaffen, Präsenz ermöglichen und zur Wiedererkennbarkeit des Formats beitragen.

3.1 Bildgestaltung und Kameraperspektiven

Die Bildgestaltung ist ein zentrales Gestaltungsmittel des Videopodcasts im Interviewformat. Während der Audiopodcast ausschließlich über Stimme, Sprache und Klang wirkt, entscheidet beim Videopodcast auch das Bild darüber, wie ein Gespräch wahrgenommen wird. Die Kamera zeigt nicht nur, wer spricht, sondern lenkt Aufmerksamkeit, schafft Orientierung und beeinflusst die Nähe oder Distanz zu Host und Gast. Als Host wird im Folgenden die moderierende beziehungsweise gesprächsführende Person des Podcastformats bezeichnet.

Für Videopodcast-Interviews ist jedoch nicht die gesamte Bandbreite filmischer Einstellungsgrößen relevant. Anders als in fiktionalen Filmproduktionen steht nicht die szenische Inszenierung im Vordergrund, sondern die klare Darstellung eines Gesprächs. Ein Videopodcast arbeitet daher häufig mit einem reduzierten Einstellungsrepertoire: einer gemeinsamen Einstellung, die Host und Gast im Raum zeigt, sowie einzelnen halbnahen Einstellungen der Gesprächspartnerinnen und Gesprächspartner. Der Mehrwert entsteht nicht durch möglichst viele Perspektiven, sondern durch die bewusste Auswahl weniger funktionaler Kameraeinstellungen.

Grundlegende Prinzipien der filmischen Auflösung können dennoch hilfreich sein. Bruckert et al. zufolge können Bildkomposition, Framing, Kamerabewegung und Schnitt dazu beitragen, die Aufmerksamkeit des Publikums zu lenken (Bruckert et al., 2022, S. 1). Obwohl sich diese Aussage auf den Film bezieht, lässt sich der Gedanke auf Videopodcasts übertragen: Bildkomposition und Framing entscheiden darüber, ob ein Gespräch visuell klar, ruhig und nachvollziehbar wirkt.

Für Gesprächssituationen unterscheidet Cutting zwei zentrale Formen der visuellen Auflösung. Master Shots zeigen alle Gesprächsteilnehmenden innerhalb eines Bildes, während bei Shot/Reverse-Shot-Auflösungen zwischen den einzelnen Gesprächspartnerinnen und Gesprächspartnern gewechselt wird (Cutting, 2016, S. 1727–1728). In der vorliegenden Arbeit übernimmt die gemeinsame Totale die Funktion eines Master Shots, da sie Host und Gast innerhalb der gesamten Gesprächssituation zeigt und räumliche Orientierung schafft. Die Totale bezeichnet dabei die Einstellungsgröße, während der Begriff Master Shot die Funktion dieser Einstellung innerhalb der Gesprächsauflösung beschreibt. Halbnahe Einzelbilder ermöglichen dagegen eine stärkere Konzentration auf die sprechende oder reagierende Person.

Welche Einstellungsgröße gewählt wird, beeinflusst die Nähe, die Aufmerksamkeit und die Lesbarkeit des Gesprächs. Nach Bruckert et al. können Einstellungsgrößen Bedeutung und Spannung vermitteln, während weite Einstellungen zusätzlich die Umgebung etablieren (Bruckert et al., 2022, S. 5). Die gemeinsame Einstellung zeigt den Gesprächsraum, die Sitzordnung und das Verhältnis zwischen Host und Gast. Eine halbnahe Einstellung zeigt eine Person in dieser Arbeit ungefähr vom Oberkörper bis zum Kopf. Dadurch bleiben Gesichtsausdruck, Körperhaltung und Teile der Gestik sichtbar, während die Person stärker in den Mittelpunkt rückt.

Eine ähnliche Abfolge führen Sanz-Aznar et al. an: Zunächst schafft eine weite Einstellung den nötigen Kontext, anschließend rücken nähere Einstellungen die Figuren und ihre emotionalen Ausdrücke stärker in den Mittelpunkt (Sanz-Aznar et al., 2023, S. 3). Für Videopodcasts lässt sich daraus ableiten, dass ein Wechsel zwischen Überblick und Nähe sinnvoll sein kann, solange er dem Gespräch dient.

Perspektivwechsel müssen räumlich nachvollziehbar bleiben. Sanz-Aznar et al. führen die 180-Grad-Regel als Mittel an, um Verwirrung beim Publikum zu vermeiden (Sanz-Aznar et al., 2023, S. 3). In einem Videopodcast-Interview hilft diese Orientierung dabei, Host und Gast logisch zueinander zu positionieren und den Wechsel zwischen den Einstellungen verständlich zu halten.

Eine weitere Funktion übernehmen Reaktionsshots. Cutting bezeichnet einen Reaction Shot als Einstellung, in der eine Person nicht spricht, sondern dem Gegenüber zuhört (Cutting, 2016, S. 1728). Im Videopodcast müssen solche Reaktionsshots keine zusätzliche Kameraeinstellung sein. Sie können aus der halbnahen Einstellung entstehen, wenn eine Person zuhört, nachdenkt

oder auf das Gesagte reagiert. Dadurch wird sichtbar, dass ein Interview nicht nur aus Antworten besteht, sondern aus einem wechselseitigen Gespräch.

Nach der 30-Grad-Regel sollte sich der Kamerawinkel zwischen zwei aufeinanderfolgenden Einstellungen ausreichend verändern, damit der Wechsel als neue Perspektive erkennbar bleibt und die visuelle Kontinuität nicht gestört wird (Sanz-Aznar et al., 2023, S. 3). Bei Videopodcasts sollten Wechsel zwischen den Einstellungen deshalb eine klare Funktion erfüllen, etwa einen Sprecherwechsel, eine Reaktion oder die Hervorhebung eines wichtigen Moments.

Zur Klarheit trägt auch die Bildkomposition bei. Bruckert et al. greifen die Drittel-Regel auf, nach der wichtige Bildelemente entlang gedachter Drittellinien platziert werden können (Bruckert et al., 2022, S. 7). Eine ausgewogene Komposition kann dabei helfen, Personen ruhig und verständlich im Bild zu platzieren. Die visuelle Gestaltung darf den auditiven Kern jedoch nicht überlagern. Knobloch et al. zufolge haben visuelle Medien die Stimme in den Hintergrund gedrängt, indem das bildliche Zeigen stärker in den Vordergrund gerückt wurde (Knobloch et al., 2020, S. 60). Für die Bildgestaltung eines Videopodcasts heißt das, dass sie das Gespräch unterstützen sollte, ohne von der Stimme und dem Inhalt abzulenken.

Bildgestaltung und Kameraperspektiven sollten sich im Videopodcast deshalb vor allem an ihrer Funktion orientieren. Eine gemeinsame Einstellung schafft Orientierung, halbnahe Einstellungen erzeugen Nähe und Reaktionsshots machen das Zuhören sichtbar. Der gestalterische Mehrwert liegt in einer reduzierten, klaren und gesprächsorientierten Kameragestaltung.

3.2 Lichtgestaltung und Atmosphäre

Neben den Kameraperspektiven gehört auch das Licht zu den zentralen Gestaltungsmitteln eines Videopodcasts. Es sorgt nicht nur dafür, dass Host und Gast sichtbar sind, sondern beeinflusst auch, ob ein Gespräch professionell, nahbar oder atmosphärisch wirkt. Im Audiopodcast spielt Licht keine Rolle. Beim Videopodcast prägt es dagegen die visuelle Wahrnehmung.

Poland macht die Bedeutung des Lichts für Bewegtbilder mit folgender Aussage deutlich: „Just as in real life, light is everything for the moving image“ (Poland, 2015, S. 2). Die visuelle Ebene funktioniert nur, wenn Personen, Raum und Gesprächssituation zu erkennen sind.

Mimik, Haltung und Reaktionen verlieren an Wirkung, wenn sie durch die Beleuchtung nicht ausreichend sichtbar werden.

Zunächst erfüllt Licht eine praktische Funktion. Poland zufolge kann es Bilder dreidimensional erscheinen lassen und dem Publikum helfen, eine Szene zu deuten (Poland, 2015, S. 2). Auf Interviews in einem Videopodcast übertragen heißt das: Gesichter, Körperhaltung und Raum sollten so ausgeleuchtet sein, dass die Gesprächssituation klar und angenehm wahrgenommen werden kann.

Darüber hinaus bestimmt Licht die Atmosphäre. Poland greift hierzu eine Aussage des Kameramanns Henri Alekan auf, nach der Licht zur Stimmung wird und einem Film seinen Ton gibt (Alekan, zitiert nach Poland, 2015, S. 20). Die Aussage bezieht sich zwar auf den Film, lässt sich aber auch auf Videopodcasts übertragen. Durch Licht kann ein Gespräch warm, sachlich, ruhig, intim oder professionell wirken. Die Beleuchtung sollte deshalb zur Idee des Formats, zum Thema und zur gewünschten Wirkung passen.

Für ein Gesprächsformat ist oft eine Balance zwischen Natürlichkeit und professioneller Bildwirkung sinnvoll. Poland führt Available Light, also bereits am Drehort vorhandenes Licht, im Zusammenhang mit dokumentarischen und realitätsnahen Formaten an, in denen es den Eindruck von Wirklichkeit unterstützen kann (Poland, 2015, S. 18). Ein zu künstlicher Look passt deshalb nicht zu jedem Videopodcast. Natürlich wirkendes, aber kontrolliertes Licht kann dagegen Nähe und Glaubwürdigkeit unterstützen.

Vorhandenes Licht reicht allerdings nicht immer aus. Gerade bei einem Interview müssen Gesichtsausdruck und Blickrichtung erkennbar bleiben. Huttunens Untersuchung zeigt, dass verdeckte Augen oder verzerrte Gesichtszüge die emotionale Verarbeitung schon früh beeinflussen können (Huttunen, 2025, S. 1, 7–9). Damit die in Kapitel 2.3 behandelte sichtbare Kommunikation wirken kann, müssen Augen, Mund und Mimik deutlich zu erkennen sein.

Auch für die räumliche Wirkung des Bildes spielt Licht eine wichtige Rolle. Die Richtung des Lichts kann nach Huttunen Formen hervorheben und Objekte vom Hintergrund trennen (Huttunen, 2025, S. 2). In Videopodcast-Interviews können Hauptlicht, weiches Aufhelllicht oder dezentes Hintergrundlicht dabei helfen, Host und Gast vom Raum abzuheben und ein klareres, räumlicheres Bild zu schaffen.

Für Gesprächsformate bietet sich häufig eine ruhige und kontrollierte Ausleuchtung an. High-Key-Lighting bezeichnet eine überwiegend helle Ausleuchtung, bei der Personen und Raum deutlich sichtbar bleiben. Poland zufolge kann eine solche Lichtgestaltung positiv und sicher wirken (Poland, 2015, S. 16). Das heißt jedoch nicht, dass jeder Videopodcast gleichmäßig hell ausgeleuchtet sein muss. Wenn das Gespräch und die beteiligten Personen im Mittelpunkt stehen, kann eine weiche und klare Beleuchtung aber sinnvoll sein.

Die Beleuchtung sollte außerdem zum Inhalt und zur Idee des Formats passen. Poland weist darauf hin, dass Lichtstil und narrative Form aufeinander abgestimmt sein müssen, damit die gewünschte emotionale Wirkung entstehen kann (Poland, 2015, S. 69). Passen Licht und Handlung nicht zusammen, kann das ablenken und die Wirkung beeinträchtigen (Poland, 2015, S. 63). Für Videopodcasts heißt das, dass zu dramatisches, zu dunkles oder zu künstliches Licht vom Gespräch ablenken kann, wenn es nicht zur inhaltlichen Ausrichtung passt.

Licht erfüllt im Videopodcast mehrere Aufgaben. Es macht die Personen sichtbar, lässt Mimik und Reaktionen erkennen, schafft räumliche Tiefe und beeinflusst die Atmosphäre. Entscheidend ist deshalb nicht allein, dass das Bild ausreichend beleuchtet ist. Die Lichtgestaltung sollte auch zum Gespräch, zu den auftretenden Personen und zur Wirkung des Formats passen.

3.3 Setgestaltung und visuelle Wiedererkennbarkeit

Die Setgestaltung bildet den sichtbaren Rahmen des Videopodcasts. Während der Audiopodcast keinen Raum zeigt, macht der Videopodcast Hintergrund, Farben, Objekte und räumliche Anordnung sichtbar. Dadurch entsteht eine zusätzliche Bedeutungsebene, die beeinflussen kann, wie Host, Gast und Format wahrgenommen werden.

Setgestaltung ist nicht bloße Dekoration, sondern Teil der visuellen Kommunikation. Poland beschreibt im Zusammenhang mit der Mise en Scène, also der bewussten Gestaltung und Anordnung sichtbarer Elemente innerhalb einer Szene, dass die Darstellung von Raum die Lesart eines Films beeinflusst. Tiefe, Nähe, Größe und Proportionen von Orten und Objekten können durch Kamera, Licht und Dekor gestaltet werden und dadurch Stimmung oder Beziehungen zwischen Bildelementen prägen (Poland, 2015, S. 9). Auch wenn sich diese Aussage auf Film bezieht, lässt sie sich auf Videopodcast-Interviews übertragen: Der sichtbare Raum wirkt mit, weil er Atmosphäre, Nähe und Professionalität beeinflussen kann.

Der Hintergrund sollte deshalb nicht zufällig gewählt werden. Ein Set kann ruhig, professionell, persönlich, kreativ oder journalistisch wirken. Es kann Orientierung geben und die inhaltliche Ausrichtung eines Formats visuell unterstützen. Besonders bei einem wiederkehrenden Interviewformat sieht das Publikum nicht nur einzelne Personen, sondern auch ein wiedererkennbares Umfeld.

Auch Objekte im Set können die Wahrnehmung prägen. Cook et al. unterscheiden zwischen funktionalen Objekten, die mit Kompetenz verbunden werden, und symbolischen Objekten, die eher Wärme vermitteln können (Cook et al., 2023, S. 4). Da sich diese Studie auf Videokonferenzen bezieht, ist die Übertragung auf Videopodcasts vorsichtig zu verstehen. Der grundlegende Zusammenhang bleibt aber nutzbar: Sichtbare Gegenstände wie Möbel, Pflanzen, Bücher oder technische Elemente können Wirkung erzeugen, sollten aber bewusst und nicht überladen eingesetzt werden.

Setgestaltung trägt zudem zur räumlichen Wiedererkennbarkeit bei. Cui beschreibt, dass eine einfache visuelle Identität Konsistenz über verschiedene Plattformen hinweg unterstützen und Wiedererkennung erleichtern kann (Cui, 2024, S. 2). In Kapitel 3.3 wird dieser Gedanke auf den physischen Raum bezogen: Wiederkehrende Hintergründe, Farben, Raumdetails oder Objekte können dazu beitragen, dass einzelne Folgen als Teil desselben Formats wahrgenommen werden.

Auch Yu et al. betonen, dass visuelle Identität den Gesamtstil einer Marke prägen kann (Yu et al., 2024, S. 1). Für die Setgestaltung ist daran vor allem relevant, dass Raum, Farbe und Bildwirkung gemeinsam einen wiedererkennbaren Eindruck schaffen können. Die grafischen Bestandteile visueller Identität wie Logo, Typografie oder Bauchbinden werden hingegen erst in Kapitel 3.6 behandelt.

Das Set sollte zur Zielgruppe, zum Thema und zur gewünschten Wirkung passen. Yu et al. weisen darauf hin, dass visuelle Elemente strategisch gesteuert werden müssen, um mit Markenwerten und Wahrnehmungen übereinzustimmen (Yu et al., 2024, S. 24). Auf Videopodcasts übertragen bedeutet das: Der Raum soll das Gespräch visuell unterstützen, aber nicht dominieren.

Der gestalterische Mehrwert der Setgestaltung liegt darin, dass der Videopodcast die Gesprächssituation räumlich einbettet. Das Publikum sieht nicht nur Host und Gast, sondern

auch das Umfeld, in dem das Gespräch stattfindet. Dadurch kann das Format greifbarer, atmosphärischer und wiedererkennbarer werden.

3.4 Personendarstellung und nonverbale Präsenz

Kapitel 2.3 hat gezeigt, welchen Mehrwert sichtbare Kommunikation im Interviewformat gegenüber reinem Audio bietet. Kapitel 3.4 beschäftigt sich nun mit der Frage, wie diese nonverbale Präsenz gestalterisch umgesetzt werden kann. Im Mittelpunkt stehen dabei Host und Gast mit ihrem Gesichtsausdruck, ihrem Blick, ihrer Körperhaltung, ihrer Gestik und ihren Reaktionen. Raum und Set spielen an dieser Stelle keine Rolle.

Es reicht jedoch nicht, Host und Gast einfach nur im Bild zu zeigen. Gesicht, Oberkörper, Hände und Blickrichtung sollten durch Kamera, Licht und Bildausschnitt gut zu erkennen sein. Huttunen zufolge ist das Erkennen und Deuten von Gesichtszügen wichtig für die soziale Kommunikation. Es hilft dabei, Gedanken und Gefühle des Gegenübers wahrzunehmen (Huttunen, 2025, S. 5). Im Videopodcast sollten Augen, Mimik und Gesichtsausdruck deshalb klar sichtbar sein.

Auch die Gestik spielt eine wichtige Rolle. Fauviaux et al. zufolge können Kopf- und Handgesten dabei helfen, komplexe Inhalte zu verdeutlichen, die sich allein über die Stimme nur schwer vermitteln lassen (Fauviaux et al., 2024, S. 3). Bei einem Interview sollten Hände und Oberkörper deshalb nicht vollständig aus dem Bild fallen. Dafür eignet sich eine halbnahe Einstellung, in der Gesicht, Teile der Körperhaltung und Gestik gemeinsam zu sehen sind.

Die Körperhaltung beeinflusst ebenfalls, wie eine Person im Gespräch wirkt. Furlan zählt die Position von Kopf und Rumpf zu den sichtbaren Hinweisen auf Einstellung und Haltung. Wer sich seinem Gegenüber zuwendet, kann positiver wirken als jemand, der sich zurücklehnt oder wegdreht (Furlan, 2021, S. 110). Bei der Darstellung einer Person geht es also nicht nur um das Gesicht, sondern auch um ihre körperliche Präsenz im Gespräch.

Beim Bildausschnitt muss zwischen Nähe und Übersicht abgewogen werden. Ist eine Person zu weit entfernt, sind kleine mimische Reaktionen kaum zu erkennen. Ist der Ausschnitt zu eng, werden Haltung und Gestik abgeschnitten. Cutting zufolge lassen sich Reaktionen leichter einschätzen, wenn die Person größer im Bild zu sehen ist (Cutting, 2016, S. 1726). Für Videopodcast-Interviews heißt das: Die Person sollte groß genug gezeigt werden, damit ihre

Reaktionen sichtbar sind. Der Ausschnitt darf aber nicht so eng sein, dass ihre körperliche Präsenz verloren geht.

Croes et al. weisen darauf hin, dass Personen in videobasierter Kommunikation häufig nur von der Brust aufwärts zu sehen sind. Hinweise auf Körperhaltung oder den sozialen Kontext können dadurch fehlen (Croes et al., 2019, S. 1211). Die Quelle bezieht sich zwar auf videobasierte Kommunikation und nicht direkt auf Videopodcasts. Trotzdem zeigt sie, wie wichtig ein bewusst gewählter Bildausschnitt ist, wenn nonverbale Signale erkennbar bleiben sollen.

Zu einem Gespräch gehört auch sichtbares Zuhören. Furlan erklärt das häufigere Nicken und Lächeln von Interviewenden damit, dass sie auf diese Weise Rückmeldung geben und zeigen, dass sie zuhören (Furlan, 2021, S. 68). Diese sichtbare Aufmerksamkeit kann im Videopodcast das Gespräch unterstützen. Das Publikum sieht dadurch nicht nur die sprechende Person, sondern auch, wie das Gegenüber auf das Gesagte reagiert.

Im Vergleich zum Audiopodcast wird deutlich, was durch das Bild hinzukommt. Mondada verweist darauf, dass bei reinen Tonaufnahmen körperliche Bestandteile des Gesprächs wie Gesten, Haltung und Gesichtsausdruck übersehen werden können (Mondada, 2008, S. 3). Im Videopodcast sind diese Signale dagegen zu sehen. Ihre Bedeutung hängt jedoch von der jeweiligen Situation ab. Furlan weist darauf hin, dass derselbe Gesichtsausdruck je nach Zusammenhang unterschiedlich verstanden werden kann (Furlan, 2021, S. 12).

Die Darstellung von Host und Gast muss deshalb bewusst gestaltet werden. Es reicht nicht aus, beide einfach nur ins Bild zu setzen. Gesicht, Blick, Oberkörper, Gestik und Reaktionen sollten gut zu erkennen sein. Einzelne nonverbale Signale dürfen dabei nicht vorschnell gedeutet werden.

3.5 Schnitt und visuelle Dynamik

Neben Kamera, Licht, Set und Personendarstellung wirkt sich auch der Schnitt auf einen Videopodcast aus. Beim Audiopodcast wird das Gespräch durch Tonschnitt, Rhythmus und Pausen gestaltet. Beim Videopodcast kommt das Bild hinzu. Der Schnitt entscheidet, wann Host, Gast, eine gemeinsame Einstellung oder eine Reaktion zu sehen ist.

Der Schnitt verbindet nicht nur einzelne Einstellungen miteinander. Poland zufolge kann er auch räumliche Zusammenhänge herstellen und den Blick des Publikums auf die wichtigsten Bereiche einer Szene lenken (Poland, 2015, S. 9). Bruckert et al. zählen den Schnitt neben Bildkomposition, Framing und Kamerabewegung zu den Mitteln, mit denen sich die Aufmerksamkeit lenken lässt (Bruckert et al., 2022, S. 1). Der Schnitt lenkt also den Blick und hebt jeweils hervor, was gerade wichtig ist.

Gerade bei Gesprächen ist der Wechsel zwischen den beteiligten Personen wichtig. Die bereits in Kapitel 3.1 beschriebene Shot/Reverse-Shot-Auflösung ist laut Cutting und Armstrong eine häufig verwendete Form gefilmter Gespräche, bei der abwechselnd die Gesprächsbeteiligten gezeigt werden (Cutting & Armstrong, 2018, S. 1326). Bei Videopodcast-Interviews macht dieser Wechsel deutlich, wer gerade spricht, wie sich das Gespräch entwickelt und wie die andere Person darauf reagiert.

Reaktionsshots zeigen nicht nur die sprechende, sondern auch die zuhörende Person. Cutting und Armstrong beziehen sich auf den Kameramann Blain Brown. Ihm zufolge sind solche Aufnahmen wichtig, weil Mimik und Körpersprache der zuhörenden Person den emotionalen Gehalt einer Szene vermitteln (Brown, 2012, zit. nach Cutting & Armstrong, 2018, S. 1320). Im Videopodcast ist dadurch nicht nur zu sehen, wer gerade spricht, sondern auch, wie das Gegenüber zuhört, nachdenkt oder auf das Gesagte reagiert.

An derselben Stelle zitieren Cutting und Armstrong den Filmhistoriker David Bordwell. Er weist darauf hin, dass sich mimische Reaktionen durch den Schnitt hervorheben, schärfen und im passenden Moment zeigen lassen (Bordwell, 2008, zit. nach Cutting & Armstrong, 2018, S. 1320). Der Schnitt beeinflusst damit auch, welches Gewicht eine Reaktion innerhalb des Gesprächs erhält. Gegenüber dem Audiopodcast hat das den Vorteil, dass Zuhören und nonverbale Reaktionen nicht nur akustisch angedeutet, sondern auch sichtbar werden.

Auch der Rhythmus des Schnitts wirkt sich auf die Wahrnehmung aus. Cutting zufolge kommen in ruhigen, erklärenden Abschnitten eher längere Einstellungen mit weniger Schnitten vor. Bei Konflikten und Handlung finden sich dagegen häufiger kurze Einstellungen und mehr Schnitte (Cutting, 2016, S. 1720). Da Cutting Filme untersucht, lässt sich diese Aussage nicht direkt auf Videopodcasts übertragen. Der Schnittrhythmus sollte aber auch hier zur Gesprächssituation passen. Ruhige Erklärungen benötigen weniger Wechsel, während in lebhafteren Momenten häufigere Perspektivwechsel sinnvoll sein können.

Visuelle Dynamik darf dabei nicht mit beliebigem Schneiden verwechselt werden. Sanz-Aznar et al. warnen, dass übermäßige oder unpassende Schnitte die Aufmerksamkeit vom narrativen Inhalt ablenken können (Sanz-Aznar et al., 2023, S. 2). Für Videopodcast-Interviews ist diese Grenze besonders wichtig, weil der Schnitt das Gespräch unterstützen und nicht überdecken sollte.

Neben der Dynamik muss der Schnitt auch dafür sorgen, dass die räumliche Orientierung erhalten bleibt. Die 180-Grad-Regel soll verhindern, dass das Publikum den Überblick verliert (Sanz-Aznar et al., 2023, S. 3). Werden mehrere Kameraperspektiven verwendet, muss deshalb erkennbar bleiben, wo Host und Gast sitzen und wie sie zueinander ausgerichtet sind. Wechsel zwischen unterschiedlichen Einstellungsgrößen und Kamerawinkeln können räumliche Zusammenhänge zusätzlich verdeutlichen. Gleichzeitig lassen sich einzelne Einstellungen hervorheben, ohne dass der Zusammenhang der Szene verloren geht (Sanz-Aznar et al., 2023, S. 9). Im Videopodcast zeigt die gemeinsame Einstellung die gesamte Gesprächssituation. Eine halbnahe Einstellung richtet den Blick stärker auf die sprechende Person, während ein Reaktionsshot die emotionale oder soziale Reaktion des Gegenübers sichtbar macht.

Der Schnitt strukturiert den Videopodcast also nicht nur technisch, sondern gestalterisch. Er lenkt Aufmerksamkeit, zeigt Sprecherwechsel, macht Reaktionen sichtbar und erhält räumliche Orientierung. Der Mehrwert entsteht nicht durch möglichst viele Schnitte, sondern durch bewusste Bildwechsel, die das Interview verständlicher, lebendiger und visuell nachvollziehbarer machen.

3.6 Grafische Gestaltung und Formatidentität

Neben Kamera, Licht, Set und Schnitt trägt grafische Gestaltung zur Wirkung eines Videopodcasts bei. Dazu gehören beispielsweise Logo, Farben, Typografie, Titelgrafiken, Einblendungen und Bauchbinden. Als Bauchbinde wird dabei eine grafische Einblendung verstanden, die beispielsweise den Namen oder die Funktion einer gezeigten Person sichtbar macht. Diese Elemente können dem Format eine klare visuelle Identität geben und einzelne Folgen als zusammengehörig erkennbar machen.

Mehrere Gestaltungselemente prägen die visuelle Identität eines Formats. Yu et al. zählen dazu Name, Logo, Farbe, Typografie und Slogan. Zusammen bestimmen sie den visuellen Stil einer

Marke (Yu et al., 2024, S. 1). Werden diese Elemente in einem Videopodcast regelmäßig verwendet, sind die einzelnen Folgen leichter als Teil desselben Formats zu erkennen.

Wichtig ist dabei, dass die Gestaltung einheitlich bleibt. Lelis et al. zufolge sorgt ein Visual Identity System durch feste Vorgaben dafür, dass visuelle Elemente einheitlich eingesetzt und wiedererkannt werden (Lelis et al., 2020, S. 147). Cui kommt außerdem zu dem Ergebnis, dass eine reduzierte visuelle Identität die Wiedererkennung in digitalen Umgebungen stärken und Assoziationen mit einer Marke verbessern kann (Cui, 2024, S. 1). Farben, Schrift und Grafiken sollten daher nicht für jede Folge neu gewählt werden, sondern sich innerhalb des Formats wiederholen.

Grafische Elemente können außerdem Informationen vermitteln. Bauchbinden zeigen die Namen und Funktionen der Gäste, Titelgrafiken kennzeichnen Themen und Einblendungen ergänzen wichtige Begriffe oder Hinweise. Lelis et al. weisen darauf hin, dass visuelle Identitäten durch zusätzliche Wörter oder Textzeilen ergänzt werden können, um Produkte, Dienstleistungen oder Orte genauer zu benennen (Lelis et al., 2020, S. 162). Auch im Videopodcast können Texteinblendungen Informationen genauer benennen und verständlicher machen.

Typografie ist dabei besonders wichtig. Cui weist darauf hin, dass zu komplexe Schriftarten das Lesen erschweren, Verständlichkeit verringern und Wiedererkennbarkeit schwächen können, während einfache und klare Schriften Lesbarkeit und Wiedererkennung unterstützen (Cui, 2024, S. 10). Das gilt besonders für Bauchbinden, Titel, Untertitel oder Social-Media-Clips, die häufig auf kleineren Bildschirmen angesehen werden.

Auch Layout und kanalübergreifende Konsistenz spielen eine Rolle. Cui beschreibt, dass sorgfältige Typografie und Layoutgestaltung dazu beitragen können, über verschiedene Kanäle hinweg eine konsistente Markenidentität zu sichern und die Nutzererfahrung zu verbessern (Cui, 2024, S. 11). Für Videopodcasts ist das relevant, weil Inhalte als vollständige Folge, Teaser oder Clip erscheinen können. Eine konsistente grafische Linie verbindet diese Ausspielungsformen miteinander.

Gleichzeitig darf grafische Gestaltung nicht zu dominant werden. Cui warnt, dass übermäßig aufwendige visuelle Gestaltung zu visueller Ermüdung führen und sowohl Merkbarkeit als

auch Wiedererkennbarkeit schwächen kann (Cui, 2024, S. 9). Grafische Elemente sollten daher Orientierung, Lesbarkeit und Wiedererkennung unterstützen, ohne das Gespräch zu überlagern.

Der Mehrwert grafischer Gestaltung liegt darin, dass Informationen und Formatidentität im Videopodcast sichtbar eingebunden werden können. Im Audiopodcast müssen Namen, Themen oder Wiedererkennbarkeit ausschließlich akustisch vermittelt werden. Im Videopodcast entsteht zusätzlich eine visuelle Oberfläche, die das Format strukturieren und seine Wiedererkennbarkeit unterstützen kann.

3.7 Gestalterische Anforderungen an Social-Media-Clips

Kapitel 2.4 hat gezeigt, dass Videomaterial eine Grundlage für Social-Media-Clips bietet. Kapitel 3.7 behandelt nun die konkrete Gestaltung solcher Kurzclips. Social-Media-Clips sind nicht einfach beliebige Ausschnitte aus einer längeren Videopodcast-Folge. Sie müssen so bearbeitet werden, dass sie schnell verständlich sind, mobil funktionieren und den ursprünglichen Gesprächskontext nicht verfälschen.

Eine zentrale Anforderung ist ein schneller und klarer Einstieg. Zhang et al. beschreiben, dass die vertikale Scroll-Funktion von Kurzvideo-Apps Nutzerinnen und Nutzern nur sehr wenig Zeit lässt, um zu entscheiden, ob sie sich für den aktuellen Inhalt interessieren (Zhang et al., 2025, S. 5). Da sich diese Untersuchung auf Kurzvideo-Werbung auf TikTok bezieht, lässt sich der Zusammenhang nur vorsichtig auf Social-Media-Clips aus Videopodcasts übertragen. Er zeigt aber grundsätzlich, dass der Einstieg unmittelbar verständlich sein muss.

Damit hängt die sogenannte Hook zusammen, also ein aufmerksamkeitsstarker Einstieg, der früh Interesse am Clip wecken soll. Zhang et al. stellen fest, dass die ersten zehn Sekunden eines Kurzvideos einen besonders starken Einfluss auf das Engagement haben (Zhang et al., 2025, S. 7). Für Videopodcast-Clips bedeutet das: Der Anfang sollte eine relevante Aussage, eine prägnante Frage, einen starken Gesprächsmoment oder eine sichtbare Reaktion enthalten. Lange Herleitungen eignen sich für Kurzclips weniger, weil die Aufmerksamkeit früh gewonnen werden muss.

Da Clips nur wenig Zeit bieten, müssen die Inhalte kurz und klar aufbereitet werden. Nach Violot et al. eignen sich Kurzformate dazu, Inhalte prägnant und ansprechend zu präsentieren (Violot et al., 2024, S. 2). Die Studie beschäftigt sich mit YouTube Shorts. Für den Videopodcast ist hier vor allem der Gedanke wichtig, dass kurze Inhalte knapp aufbereitet sein

müssen. Ein Clip sollte sich deshalb auf einen Gedanken, eine Aussage oder eine Gesprächssituation konzentrieren und auch ohne die vollständige Folge verständlich sein.

Mobile Nutzung und vertikale Darstellung stellen zusätzliche Anforderungen. Zhang et al. beziehen sich auf vertikale Scroll-Logiken in Kurzvideo-Apps (Zhang et al., 2025, S. 5). Daraus lässt sich für Social-Media-Clips ableiten, dass horizontal aufgenommenes Videopodcast-Material häufig angepasst oder neu kadriert werden muss. Host und Gast sollten auch im vertikalen Ausschnitt gut erkennbar bleiben, ohne dass wichtige Mimik, Gestik oder Bildinformationen verloren gehen.

Die hohe Wechselbereitschaft auf Kurzvideo-Plattformen verstärkt den Anspruch an Klarheit. Zannettou et al. zeigen für TikTok, dass ein großer Teil empfohlener Videos nicht bis zum Ende angesehen wurde und die meisten davon bereits vor der Hälfte abgebrochen wurden (Zannettou et al., 2024, S. 3). Diese Quelle bezieht sich auf TikTok und darf nicht pauschal auf alle Plattformen übertragen werden. Sie stützt aber den allgemeinen Punkt, dass Social-Media-Clips schnell erfassbar und inhaltlich konzentriert sein müssen.

Untertitel und Texteinblendungen sind für solche Clips besonders relevant. Szarkowska et al. weisen darauf hin, dass viele Konsumentinnen und Konsumenten Videos in öffentlichen Räumen lieber ohne Ton ansehen und in diesen Fällen stark auf Untertitel angewiesen sind (Szarkowska et al., 2024, S. 2). Da Videopodcasts stark auf gesprochener Sprache basieren, können Untertitel entscheidend sein, damit zentrale Aussagen auch ohne Ton verständlich bleiben.

Ergänzende Texteinblendungen können zusätzlich helfen, den Clip schnell einzuordnen. Kurze Titel, hervorgehobene Begriffe oder Bauchbinden können zeigen, wer spricht, worum es geht oder welche Aussage im Mittelpunkt steht. Diese Elemente sollten jedoch klar und zurückhaltend eingesetzt werden. Sie unterstützen die Verständlichkeit, dürfen den Gesprächsausschnitt aber nicht überlagern.

Bei einem kurzen Clip müssen auch Schnitt und Tempo angepasst werden. Nach Liu et al. kann das Zusammenspiel von Rhythmus, Abfolge und Ton die emotionale Wirkung eines Kurzvideos beeinflussen und die Bereitschaft zum Weiterschauen erhöhen (Liu et al., 2018, zit. nach Gao et al., 2021, S. 3). Pausen, Wiederholungen und Nebensätze können in einem

Videopodcast-Clip daher stärker gekürzt werden als in der vollständigen Folge. Das Gespräch sollte trotzdem natürlich wirken.

Schon bei der Auswahl muss entschieden werden, welcher Gesprächsausschnitt für einen Clip geeignet ist. Preger weist darauf hin, dass zu viele Passagen ohne weitere Bedeutung Langeweile erzeugen können (Preger, 2020, S. 216). Das gilt auch für Social-Media-Clips. Nicht jeder Moment aus einer Podcastfolge eignet sich dafür. Ein Clip sollte einen klaren Inhalt vermitteln. Beim Kürzen darf eine Aussage weder verfälscht noch aus ihrem Zusammenhang gerissen werden. Sie kann zugespitzt werden, muss inhaltlich aber richtig bleiben.

Schließlich sollten Social-Media-Clips zur visuellen Identität des Videopodcasts passen. Wiederkehrende Farben, Schriften, Untertitelstile, Bauchbinden oder grafische Elemente können dafür sorgen, dass einzelne Clips als Teil desselben Formats erkennbar bleiben. Damit verbindet die Clipgestaltung die kurze Plattformversion mit der vollständigen Podcastfolge.

Social-Media-Clips stellen damit eigene gestalterische Anforderungen: Sie brauchen einen schnellen Einstieg, eine klare Hook, inhaltliche Verdichtung, mobile beziehungsweise vertikale Lesbarkeit, Untertitel, angepassten Schnittrhythmus, sorgfältige Ausschnittwahl und eine Verbindung zur visuellen Identität des Formats. Der Videopodcast liefert dafür das notwendige visuelle Ausgangsmaterial, dessen gestalterische Aufbereitung jedoch an die Anforderungen des jeweiligen Kurzformats angepasst werden muss.

4. Entwicklung und Produktion der eigenen Videopodcast-Staffel

4.1 Praxisbasierter Ansatz der Arbeit

Neben der theoretischen Auseinandersetzung mit dem gestalterischen Mehrwert von Videopodcasts umfasst diese Bachelorarbeit auch einen praktischen Produktionsteil. Im Mittelpunkt steht die Entwicklung und Umsetzung einer eigenen Videopodcast-Staffel mit dem Titel “Sidequest Social Media”. Das Format wurde als Interviewformat angelegt und besteht aus fünf Folgen mit jeweils einem Gast.

Ziel des praktischen Teils war es, die zuvor theoretisch herausgearbeiteten gestalterischen Anforderungen innerhalb einer eigenen Videopodcast-Produktion praktisch anzuwenden und ihre Umsetzung im konkreten Produktionsprozess zu untersuchen. Die Produktion diente damit als Erprobung der in Kapitel 2 und 3 erarbeiteten Überlegungen zur visuellen Ebene von

Videopodcasts. Im Vordergrund stand nicht nur die Erstellung einzelner Podcastfolgen, sondern die Frage, wie Kamera, Licht, Set, Personendarstellung, Schnitt, grafische Gestaltung und Social-Media-Verwertung in einem eigenen Format zusammengeführt werden können.

Inhaltlich beschäftigt sich “Sidequest Social Media” mit dem Thema Social Media. Die Gespräche mit den Gästen sollten unterschiedliche Perspektiven auf Plattformen, Content Creation, Produktion, Streaming, Influencer-Management und digitale Öffentlichkeit ermöglichen. Das Interviewformat bot dafür einen geeigneten Rahmen, weil es persönliche Erfahrungen, fachliche Einblicke und Gesprächsdynamik miteinander verbindet.

Der Videopodcast wurde bewusst als audiovisuelle Form umgesetzt. Neben dem gesprochenen Inhalt sollten auch Mimik, Gestik, Reaktionen, Blickrichtungen, Set, Licht und visuelle Zusatzinformationen Teil der Gesamtwirkung werden. Dadurch konnte die in der Forschungsfrage zentrale visuelle Ebene nicht nur theoretisch beschrieben, sondern im eigenen Produktionsprozess praktisch gestaltet werden.

Die fünf Folgen wurden an fünf unterschiedlichen Drehtagen aufgenommen. Die Aufnahmen verteilten sich auf zwei Produktionsblöcke beziehungsweise Studios, sodass das Set insgesamt zweimal vollständig aufgebaut werden musste.

Ein besonderer Bestandteil des Formats war die Verbindung aus wiederkehrender Grundstruktur und variablen Elementen. Die Gäste hatten die Möglichkeit, persönliche oder thematisch passende Gegenstände mitzubringen, die in das Set integriert wurden. Zudem wurden in ausgewählten Folgen Foto- und Videoeinblendungen genutzt, um Gesprächsinhalte visuell zu unterstützen. Für die Social-Media-Clips wurde bereits bei der Aufnahme darauf geachtet, dass das Material auch im Hochformat nutzbar bleibt.

Der praktische Teil bildet damit die Grundlage für die folgenden Unterkapitel. Nach der Referenzanalyse werden zunächst gestalterische Kriterien für “Sidequest Social Media” abgeleitet. Anschließend wird dargestellt, wie diese Kriterien in Formatidee, Zielgruppe, visuellem Konzept, technischem Setup, Aufnahmeprozess, Postproduktion und Clip-Erstellung umgesetzt wurden.

4.2 Gestalterische Referenzanalyse ausgewählter Videopodcast-Formate

Vor der Entwicklung von “Sidequest Social Media” wurden ausgewählte Videopodcast- und Podcast-Clip-Beispiele betrachtet. Ziel dieser Referenzanalyse war keine repräsentative Marktübersicht, sondern eine exemplarische und kriteriengeleitete gestalterische Betrachtung einzelner Formate. Die Analyse dient dazu, konkrete audiovisuelle Gestaltungslösungen sichtbar zu machen, die für die Entwicklung der eigenen Produktion relevant sind.

Die Auswahl der Referenzen erfolgte gezielt und nicht mit dem Anspruch, die Videopodcast-Landschaft vollständig oder repräsentativ abzubilden. Entscheidend war vielmehr, dass die jeweiligen Beispiele für einzelne Gestaltungsbereiche konkrete Lösungen zeigen, die für die Entwicklung des eigenen Formats relevant waren. Die Referenzen übernehmen damit unterschiedliche Funktionen innerhalb der Analyse und werden nicht als unmittelbar miteinander konkurrierende Formate bewertet.

Die Auswahl orientierte sich an Gestaltungsbereichen, die auch für “Sidequest Social Media” eine zentrale Rolle spielen. Der Tim Gabel Podcast wurde betrachtet, weil er eine klassische Interviewauflösung mit mehreren Kameraperspektiven, kontrollierter Lichtgestaltung und einem bewusst inszenierten Set zeigt. The Joe Rogan Experience wurde aufgrund der Einbindung zusätzlicher Videoelemente ausgewählt. Schatz, wir müssen reden wurde als Referenz für ein Setkonzept mit wiederkehrenden und variablen Elementen untersucht. Clare on Air wurde vor allem wegen der indirekten Lichtgestaltung und der atmosphärischen Farbwirkung betrachtet. Baywatch Berlin diente als Beispiel für die Weiterverwertung von Podcastmaterial in Social-Media-Clips.

Die Analyse orientiert sich an den in Kapitel 3 behandelten Gestaltungsbereichen. Betrachtet werden Kamera und Bildgestaltung, Licht und Atmosphäre, Setgestaltung, Personendarstellung, Schnitt, grafische Gestaltung, zusätzliche visuelle Inhalte sowie die Nutzung des Materials für Social Media. Nicht bei jeder Referenz werden alle Gestaltungsbereiche gleichermaßen untersucht. Der Fokus liegt jeweils auf denjenigen Merkmalen, aufgrund derer das jeweilige Format als Referenz ausgewählt wurde.

Die Analyse bezieht sich auf ausgewählte Sequenzen der jeweiligen Folgen beziehungsweise Clips. Die angegebenen Timecodes und Standbilder dokumentieren die untersuchten Stellen und sollen die gestalterischen Beobachtungen nachvollziehbar machen. Im Mittelpunkt steht

dabei die Frage, welche gestalterische Funktion das jeweils beobachtete Element innerhalb des Formats erfüllt und welche Überlegungen daraus für die eigene Produktion abgeleitet werden können.

Als erste Referenz wurde der Tim Gabel Podcast betrachtet. Grundlage der Analyse ist die Folge Nordkorea Insider: Die Wahrheit über das Leben in Nordkorea – Dr. Rüdiger Frank, erschienen am 9. August 2026 auf Spotify. Für die visuelle Detailbetrachtung wurden unter anderem Ausschnitte aus dem Bereich 44:10 bis 44:42 herangezogen.

In der analysierten Folge arbeitet das Format mit einer klassischen Gesprächsauflösung aus mehreren Kameraperspektiven. Sichtbar werden eine gemeinsame Einstellung, in der Host und Gast innerhalb des Raumes gezeigt werden, sowie halbnahe Einzelbilder der Gesprächspartner. Die gemeinsame Einstellung stellt die räumliche Beziehung zwischen den Personen her und zeigt das Gespräch als Gesamtsituation. Die halbnahen Einzelbilder konzentrieren sich stärker auf die jeweils sprechende Person und machen Gesichtsausdruck, Körperhaltung und Reaktionen sichtbarer. Die unterschiedlichen Funktionen der drei Perspektiven werden in Abbildung 1 beispielhaft gegenübergestellt.



Abbildung 1



Abbildung 2



Abbildung 3

Auch das Set ist deutlich als gestalteter Raum erkennbar. Steinwände, Regale, Pflanzen, Sessel, Beistelltische, sichtbare Mikrofone und indirekte Lichtakzente erzeugen einen visuell kohärenten Eindruck. Der Hintergrund wirkt nicht zufällig, sondern bewusst eingerichtet. Die

räumliche Tiefe, die dekorativen Elemente und die klare Positionierung der Personen tragen dazu bei, dass das Format eine erkennbare visuelle Umgebung erhält.

Die Lichtgestaltung unterstützt diese Wirkung. Die Personen sind gut erkennbar, während der Hintergrund durch einzelne Lichtquellen und Helligkeitsunterschiede Tiefe erhält. Dadurch entsteht ein kontrollierter Bildeindruck, bei dem sowohl die Gesprächspartner als auch der Raum sichtbar bleiben. Insgesamt zeigt die Referenz, wie Mehrkamera-Auflösung, Setgestaltung und Licht zusammenwirken können, um ein Interviewformat visuell ruhig, geordnet und wiedererkennbar erscheinen zu lassen.

Als zweite Referenz wurde die Folge 2527 – MrBeast aus The Joe Rogan Experience betrachtet. Die Folge erschien am 16. Juli 2026 auf Spotify. Für die Analyse waren insbesondere Passagen relevant, in denen die Gesprächspartner in Einzelperspektiven gezeigt werden und zusätzlich externes Videomaterial in das Podcastbild eingebunden wird.

In der untersuchten Folge arbeitet das Format in den betrachteten Passagen vor allem mit halbnahen Einzelperspektiven der Gesprächspartner. Eine gemeinsame Totale, die beide Personen gleichzeitig im Raum zeigt, steht in diesen Ausschnitten nicht im Vordergrund. Dadurch liegt der Fokus stärker auf der einzelnen Person als auf der räumlichen Gesamtsituation.

Die Bildgestaltung konzentriert sich auf Gesicht, Oberkörper, Mikrofon und Hintergrund. Mimik und Blickrichtung bleiben gut sichtbar, während die räumliche Beziehung zwischen den Gesprächspartnern weniger stark über eine gemeinsame Einstellung vermittelt wird. Der Hintergrund enthält ein deutlich erkennbares Leuchtelement beziehungsweise ein markantes Zeichen des Formats. Dieses Element unterstützt die Wiedererkennbarkeit, ohne dass der gesamte Raum ausführlich gezeigt werden muss.

Besonders auffällig ist die Einbindung eines externen Videos innerhalb des Podcastbildes. Das Gespräch bezieht sich sichtbar auf dieses Material. Im Unterschied zu einem reinen Audioformat muss der visuelle Gegenstand nicht nur sprachlich beschrieben werden, sondern kann vom Publikum gleichzeitig gesehen werden. Dadurch entsteht ein gemeinsamer visueller Referenzpunkt für Host, Gast und Publikum. Die visuelle Ebene übernimmt hier also nicht nur die Funktion, die Gesprächspartner zu zeigen, sondern erweitert das Gespräch um zusätzliches

Anschauungsmaterial. Abbildung 4 zeigt beispielhaft, wie dieses zusätzliche Videomaterial unmittelbar in die Gesprächssituation eingebunden wird.



Abbildung 4

Als dritte Referenz wurde der Videopodcast Schatz, wir müssen reden betrachtet. Das Format erscheint auf dem YouTube-Kanal SWMR_Podcast. Grundlage der Analyse sind unter anderem die Folge Healing Session mit Badmómzjay, erschienen am 2. Juli 2026, sowie die Folge Handy tauschen mit Mali, erschienen am 5. Februar 2026. Beide Folgen wurden herangezogen, um die Setgestaltung und die Veränderung einzelner Setelemente zwischen verschiedenen Episoden zu betrachten.

Das Format arbeitet mit mehreren Kameraperspektiven. In den analysierten Folgen werden sowohl gemeinsame Einstellungen genutzt, in denen beide Gesprächspartnerinnen im Set gezeigt werden, als auch nähere Einzelperspektiven. Dadurch wird einerseits räumliche Orientierung geschaffen, andererseits können einzelne Reaktionen, Blickrichtungen und Gesichtsausdrücke stärker hervorgehoben werden.

Besonders prägend ist das Setkonzept. Der Raum ist deutlich zweigeteilt: Eine Seite ist der Hostin zugeordnet und bleibt in ihrer Grundstruktur wiedererkennbar, während die Seite des Gastes stärker variiert. Im Vergleich der beiden betrachteten Folgen wird erkennbar, dass sich

die Gastseite je nach Folge sichtbar verändert. Dadurch entsteht ein Zusammenspiel aus konstanten und veränderbaren Setelementen.



Abbildung 5



Abbildung 6

Die wiederkehrende Seite der Hostin sorgt für visuelle Kontinuität. Gleichzeitig ermöglicht die variable Gastseite eine stärkere Anpassung an die jeweilige Folge beziehungsweise Person.

Der direkte Vergleich beider Folgen macht dieses Prinzip aus Konstanz und Variation besonders deutlich (vgl. Abb. 5,6)

Die Farbigkeit, die dekorativen Objekte und die klare räumliche Aufteilung tragen dazu bei, dass das Set nicht nur Hintergrund ist, sondern aktiv zur Formatwirkung beiträgt. Die Referenz zeigt damit, dass Wiedererkennbarkeit und Variation innerhalb eines Podcastsets miteinander verbunden werden können.

Als vierte Referenz wurde der Videopodcast Clare on Air betrachtet. Grundlage der Analyse ist die Folge Jermain über Therapie, Ozempic und Schönheitsideale, erschienen am 10. August 2026 auf Spotify. Die Folge wurde vor allem im Hinblick auf Lichtgestaltung, Atmosphäre und farbliche Akzentuierung betrachtet.

Das Format arbeitet in der untersuchten Folge mit Einzelperspektiven von Host und Gast. Eine gemeinsame Totale steht in den betrachteten Passagen nicht im Vordergrund. Die Bildgestaltung konzentriert sich dadurch stark auf die jeweilige Person, während der Hintergrund trotzdem als gestalteter Raum wahrnehmbar bleibt.

Besonders auffällig ist die indirekte Lichtgestaltung. Im Hintergrund sind warme Lichtquellen sichtbar, die nicht direkt auf die Personen gerichtet sind, sondern den Raum atmosphärisch strukturieren. Zusätzlich entstehen durch farbige Lichtakzente, insbesondere im rötlichen Bereich, einzelne visuelle Schwerpunkte im Bild. Dadurch wirkt das Set dunkler, kontrastreicher und stimmungsvoller als heller ausgeleuchtete Gesprächssituationen.

Die Lichtgestaltung erfüllt hier mehrere Funktionen gleichzeitig. Sie macht die Personen sichtbar, trennt sie vom Hintergrund und erzeugt eine bestimmte Atmosphäre. Der Raum wirkt nicht neutral, sondern bewusst inszeniert. Durch die Mischung aus warmem indirektem Licht, dunklen Flächen, punktuellen Farbakzenten und dekorativen Objekten entsteht ein stimmungsvoller Bildeindruck, der zur visuellen Identität des Formats beiträgt. Eine vergleichbare Funktion der Lichtgestaltung wird auch in der betrachteten Folge mit Jana sichtbar. Die beiden ausgewählten Einstellungen verdeutlichen, wie die Lichtstimmung über unterschiedliche Einzelperspektiven hinweg erhalten bleibt und die Personen gleichzeitig deutlich vom Hintergrund abgesetzt werden (vgl. Abb. 7,8).



Abbildung 7



Abbildung 8

Als fünfte Referenz wurde Baywatch Berlin betrachtet. Grundlage der Analyse ist der Instagram-Clip Wie würdet ihr das machen?, veröffentlicht am 16. Juli 2026. Anders als bei den zuvor betrachteten Beispielen steht hier nicht die vollständige Videopodcast-Folge im Mittelpunkt, sondern die Aufbereitung von Podcastmaterial als Social-Media-Clip.

Der analysierte Clip ist im Hochformat gestaltet und damit auf mobile Nutzung ausgerichtet. Die sprechenden Personen werden in halbnahen bis nahen Ausschnitten gezeigt. Gesicht, Mikrofon, Kopfhörer und Teile der Körperhaltung bleiben sichtbar, während der Raum nur ausschnitthaft erkennbar ist. Die Podcast-Situation wird dadurch schnell verständlich, ohne dass das gesamte Set gezeigt werden muss.

Auffällig ist außerdem der starke Einsatz von Untertiteln. Die gesprochenen Inhalte werden groß und gut lesbar eingeblendet. Dadurch wird der Clip auch in Nutzungssituationen verständlicher, in denen Ton nicht oder nur eingeschränkt wahrgenommen wird. Die Untertitel erfüllen dabei nicht nur eine funktionale Rolle, sondern prägen auch die grafische Erscheinung des Clips.

Die Referenz zeigt, dass Podcastmaterial für Social Media nicht einfach unverändert übernommen wird. Es wird ausgewählt, verdichtet, neu kadriert und durch Untertitel ergänzt. Dadurch entsteht aus einem längeren Gesprächsmoment ein eigenständiger kurzer Inhalt. Für die gestalterische Betrachtung ist deshalb vor allem relevant, wie aus vorhandenem Videomaterial ein clipfähiges Format entsteht. Die Kombination aus vertikaler Kadrierung, wechselndem Sprecherfokus und großformatigen Untertiteln lässt sich anhand der ausgewählten Einzelbilder des Clips nachvollziehen (vgl. Abb. 9,10,11).

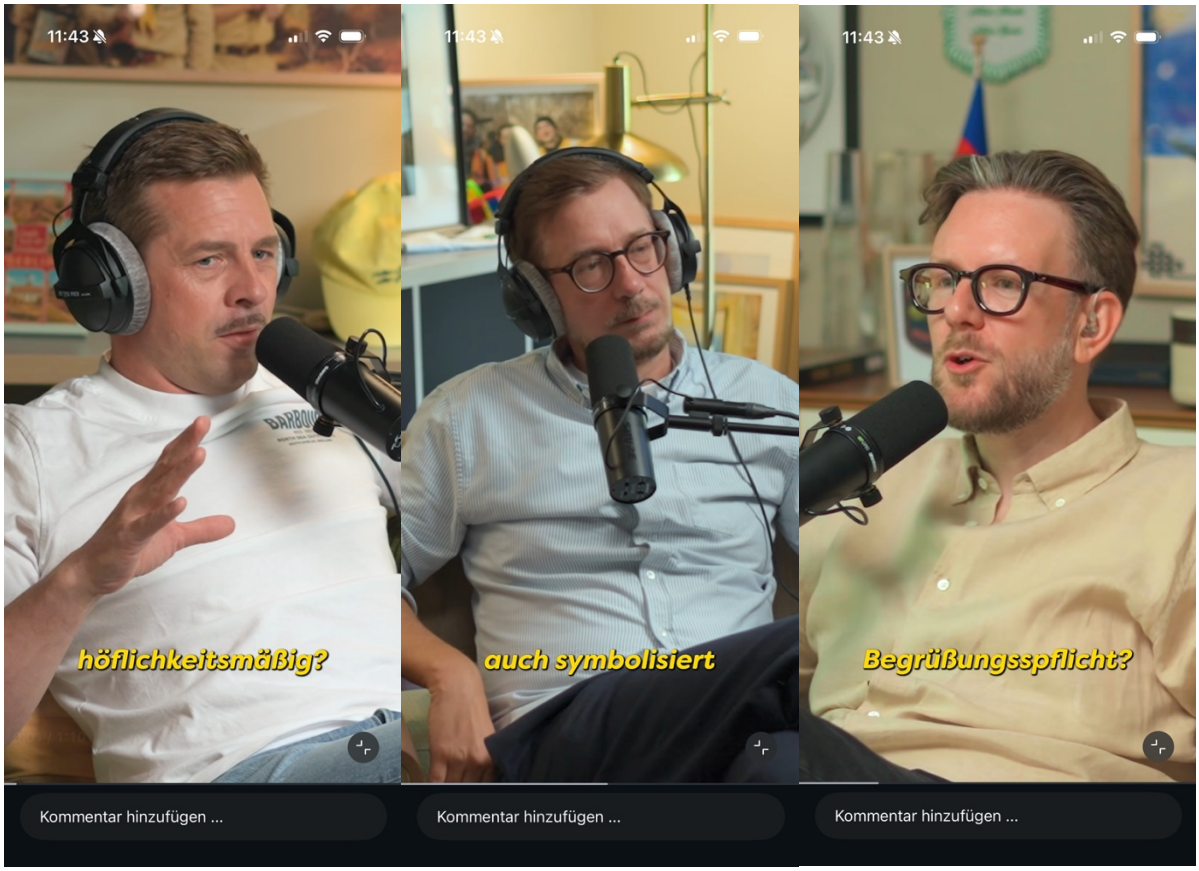


Abbildung: 9,10,11

Die fünf Referenzen zeigen unterschiedliche Schwerpunkte audiovisueller Podcastgestaltung. Der Tim Gabel Podcast steht für eine klassische Mehrkamera-Auflösung mit räumlicher Orientierung, Setgestaltung und kontrollierter Lichtwirkung. The Joe Rogan Experience zeigt die Einbindung zusätzlicher visueller Inhalte als gemeinsamen Referenzgegenstand innerhalb des Gesprächs. Schatz, wir müssen reden verdeutlicht ein Setkonzept, das konstante und variable Elemente miteinander verbindet. Clare on Air zeigt die Wirkung indirekter Lichtquellen und farbiger Akzente auf die Atmosphäre eines Videopodcasts. Baywatch Berlin zeigt schließlich die Transformation von Podcastmaterial in ein eigenständiges Social-Media-Format.

4.3 Ableitung der Gestaltungskriterien

Auf Grundlage der theoretischen Gestaltungsbereiche aus Kapitel 2 und 3 sowie ihrer Konkretisierung durch die Referenzanalyse in Kapitel 4.2 werden im Folgenden konkrete Gestaltungskriterien für “Sidequest Social Media” formuliert. Die Kriterien sind damit weder ausschließlich aus der Literatur noch ausschließlich aus den Referenzformaten abgeleitet. Sie entstehen aus der Zusammenführung theoretischer Anforderungen, beobachteter

Gestaltungslösungen und der Zielsetzung des eigenen Formats. Sie bilden die Grundlage für das visuelle Konzept und die anschließende technische und gestalterische Umsetzung des Formats.

Ein zentrales Kriterium ist eine klare Mehrkamera-Auflösung des Interviews. Sowohl die theoretischen Grundlagen als auch die betrachteten Referenzen zeigen, dass unterschiedliche Kameraperspektiven verschiedene Funktionen innerhalb einer Gesprächssituation übernehmen können. Für “Sidequest Social Media” wurde deshalb eine Auflösung vorgesehen, die sowohl eine gemeinsame Darstellung von Host und Gast als auch nähere Einzelperspektiven ermöglicht. Dadurch sollen räumliche Orientierung und eine stärkere Konzentration auf die jeweils sprechende oder reagierende Person miteinander verbunden werden.

Damit verbunden ist die Personendarstellung als weiteres Gestaltungskriterium. Da Host und Gast im Mittelpunkt des Interviewformats stehen, müssen Mimik, Gestik, Blickrichtung, Körperhaltung und sichtbare Reaktionen ausreichend erkennbar bleiben. Die theoretischen Überlegungen aus Kapitel 3.4 sprechen daher für Einstellungen, die einerseits genügend Nähe zur Person herstellen und andererseits noch Teile der körperlichen Präsenz und der Gesprächssituation sichtbar lassen.

Auch die Bildkomposition wurde als eigenes Kriterium berücksichtigt. Die Personen sollen klar im Mittelpunkt stehen, ohne vollständig vom gestalteten Raum getrennt zu werden. Eine ruhige und ausgewogene Kadrierung soll Orientierung schaffen und gleichzeitig genügend Raum für Mikrofon, Körperhaltung und Hintergrundelemente lassen. Gestaltungsregeln wie die Drittelregel dienen dabei als Orientierung, nicht als starre Vorgabe.

Ein weiteres Gestaltungskriterium betrifft den Schnitt und die visuelle Gesprächsdynamik. Wie in Kapitel 3.5 herausgearbeitet wurde, sollen Perspektivwechsel im Videopodcast nicht lediglich für visuelle Abwechslung sorgen, sondern eine konkrete Funktion innerhalb des Gesprächs übernehmen. Für „Sidequest Social Media“ wurde deshalb das Kriterium abgeleitet, den Schnitt am Gesprächsverlauf zu orientieren. Sprecherwechsel, Reaktionen und gemeinsame Gesprächsmomente sollen durch die Auswahl der jeweiligen Kameraperspektive unterstützt werden. Gleichzeitig soll auf zu häufige oder rein dekorative Bildwechsel verzichtet werden, damit der Schnitt den Gesprächsfluss strukturiert, ohne vom gesprochenen Inhalt abzulenken.

Ein weiterer Schwerpunkt ist die Verbindung aus Wiedererkennbarkeit und Variation innerhalb des Sets. Die Analyse von Schatz, wir müssen reden zeigte, dass ein wiederkehrendes Set nicht in jeder Folge vollständig identisch sein muss. Eine konstante visuelle Grundstruktur kann mit variablen Elementen kombiniert werden. Für “Sidequest Social Media” wurde daraus das Prinzip eines wiedererkennbaren Sets abgeleitet, das zugleich Möglichkeiten bietet, einzelne Gäste und Folgen visuell zu individualisieren.

Auch die Lichtgestaltung soll sowohl eine funktionale als auch eine atmosphärische Aufgabe erfüllen. Die theoretischen Grundlagen und die Analyse von Clare on Air zeigen, dass Licht nicht nur die Sichtbarkeit der Personen sicherstellt, sondern auch Raumwirkung und Atmosphäre beeinflussen kann. Für das eigene Format wurde deshalb das Kriterium abgeleitet, eine klare Erkennbarkeit von Gesicht, Mimik und Körperhaltung mit einer bewusst gestalteten Lichtstimmung zu verbinden.

Ein weiteres Gestaltungskriterium ist die gezielte Einbindung zusätzlicher visueller Inhalte. Die Analyse von The Joe Rogan Experience verdeutlichte, dass Bilder oder Videos innerhalb eines Gesprächs als gemeinsamer visueller Referenzpunkt dienen können. Für “Sidequest Social Media” sollen solche Einblendungen deshalb nur dort eingesetzt werden, wo sie einen konkreten Gesprächsinhalt ergänzen oder für das Publikum anschaulicher machen. Gerade bei Themen aus dem Bereich Social Media kann dies relevant sein, wenn über visuelle Inhalte, Fotos, Videos oder andere mediale Beispiele gesprochen wird.

Darüber hinaus wurde die spätere Social-Media-Verwertung bereits als Kriterium in die Entwicklung des Formats einbezogen. Die Analyse von Baywatch Berlin zeigte, dass Podcastmaterial für kurze Social-Media-Clips ausgewählt, verdichtet und für ein anderes Bildformat aufbereitet werden muss. Die Bildgestaltung von “Sidequest Social Media” soll deshalb grundsätzlich so angelegt sein, dass insbesondere die Einzelperspektiven auch als Ausgangsmaterial für vertikale Kurzclips genutzt werden können.

Damit verbunden ist eine konsistente grafische Gestaltung. Wiederkehrende Farben, Typografie und Untertitelgestaltung sollen dazu beitragen, die einzelnen Social-Media-Clips visuell miteinander zu verbinden und dem Format eine wiedererkennbare Gestaltung zu geben. Gleichzeitig sollen grafische Elemente zurückhaltend eingesetzt werden, damit sie die gesprochenen Inhalte unterstützen und nicht überlagern.

Aus Theorie und Referenzanalyse ergeben sich damit neun zentrale Gestaltungskriterien für „Sidequest Social Media“: eine klare Mehrkamera-Auflösung, eine auf nonverbale Präsenz ausgerichtete Personendarstellung, eine bewusste Bildkomposition, eine am Gespräch orientierte Schnittgestaltung, ein wiedererkennbares und zugleich variables Set, eine funktionale und atmosphärische Lichtgestaltung, die gezielte Einbindung zusätzlicher visueller Inhalte, eine bereits bei der Aufnahme berücksichtigte Social-Media-Verwertung sowie eine konsistente grafische Gestaltung.

Diese neun Kriterien bilden zugleich den Bezugsrahmen für die spätere Reflexion der eigenen Produktion. In Kapitel 5 wird daher betrachtet, in welchem Umfang sie umgesetzt werden konnten, welche Funktionen sie innerhalb des Formats übernahmen und an welchen Stellen praktische Grenzen oder Verbesserungspotenziale sichtbar wurden.

4.4 Formatidee und Zielgruppe

Das entwickelte Format trägt den Titel “Sidequest Social Media” und ist als videobasiertes Gesprächs- und Interviewformat angelegt. Im Mittelpunkt stehen Gespräche mit Gästen, die auf unterschiedliche Weise mit Social Media in Berührung stehen oder dort eigene praktische Erfahrungen gesammelt haben. Ziel des Formats ist es, verschiedene Perspektiven auf Social Media sichtbar zu machen und dabei persönliche Erfahrungen, berufliche Einblicke und medienpraktische Fragestellungen miteinander zu verbinden.

Die Entscheidung für ein Interviewformat ergab sich aus dem inhaltlichen Ansatz der Staffel. Social Media sollte nicht nur abstrakt oder theoretisch behandelt werden, sondern anhand konkreter Personen und ihrer Erfahrungen besprochen werden. Durch wechselnde Gäste konnte das Thema aus unterschiedlichen Blickwinkeln betrachtet werden, etwa aus der Perspektive von Content Creation, Foto- und Videoproduktion, Influencer-Management, Streaming, Schnitt oder Plattformnutzung. Das Format ist dabei nicht als klassisches journalistisches Interview angelegt, sondern als Mischung aus Interview, Gesprächspodcast und lockerem Talkformat. Dadurch sollte eine persönliche und zugängliche Gesprächsatmosphäre entstehen, ohne den informativen Charakter des Formats zu verlieren.

Die Tonalität von “Sidequest Social Media” ist bewusst locker und persönlich gehalten. Die Gespräche sollen nicht distanziert oder rein sachlich wirken, sondern nah an den Erfahrungen der Gäste bleiben. Gleichzeitig verfolgt das Format einen informativen Anspruch, da die

einzelnen Folgen Einblicke in verschiedene Formen von Social-Media-Arbeit und digitaler Selbstpräsentation geben. Die Verbindung aus persönlicher Gesprächssituation und thematischem Fokus bildet damit die Grundlage des Formats.

Die primäre Zielgruppe von “Sidequest Social Media” sind junge Erwachsene, insbesondere Studierende, angehende Medienschaffende sowie junge Creatorinnen und Creator, die sich für Social Media, Content Creation und digitale Selbstpräsentation interessieren. Diese Eingrenzung ergibt sich sowohl aus der thematischen Ausrichtung des Formats als auch aus der Auswahl der Gäste, die überwiegend selbst in einem jungen, social-media-affinen Umfeld verortet sind.

Der Fokus liegt dabei vor allem auf einer Zielgruppe im Umfeld der Gen Z und junger Erwachsener. Für diese Gruppe ist Social Media nicht nur ein Unterhaltungsmedium, sondern häufig auch ein Raum für Selbstdarstellung, berufliche Orientierung, kreative Produktion und öffentliche Sichtbarkeit. Das Format knüpft an diese Lebens- und Medienrealität an, indem es nicht abstrakt über Plattformen spricht, sondern konkrete Erfahrungen von Personen in den Mittelpunkt stellt, die selbst mit Social Media arbeiten oder darüber sichtbar geworden sind.

Darüber hinaus kann das Format auch für Personen interessant sein, die sich allgemein mit digitalen Medien, Influencer-Kultur, Plattformlogiken oder der Produktion von Social-Media-Inhalten beschäftigen. Die Zielgruppe ist damit nicht ausschließlich auf Studierende oder Creatorinnen und Creator beschränkt, bleibt aber bewusst auf ein junges, medieninteressiertes Publikum ausgerichtet.

Der Titel “Sidequest Social Media” greift die Idee auf, dass Social Media für viele Menschen zunächst nebenbei beginnt. Häufig entsteht Content nicht direkt als Haupttätigkeit, sondern als Begleitung zu anderen Interessen, Projekten oder beruflichen Wegen. Aus einer solchen Nebenbeschäftigung kann sich mit der Zeit eine größere Öffentlichkeit, ein eigenes Profil oder sogar eine berufliche Perspektive entwickeln. Der Begriff „Sidequest“ verweist dabei auf eine Nebenaufgabe oder einen zusätzlichen Weg, der zunächst nicht im Mittelpunkt steht, aber an Bedeutung gewinnen kann. Damit beschreibt der Titel sowohl den lockeren Zugang des Formats als auch die Entwicklung vieler Social-Media-Biografien: Was als Nebenprojekt beginnt, kann zu einem wichtigen Teil der eigenen Identität oder Arbeit werden.

“Sidequest Social Media” versteht Social Media daher nicht nur als Plattformnutzung, sondern als Raum für persönliche Entwicklung, kreative Produktion und mediale Sichtbarkeit. Das Format soll zeigen, wie unterschiedlich Menschen mit Social Media arbeiten, welche Chancen und Herausforderungen damit verbunden sind und wie sich digitale Präsenz im Laufe der Zeit verändern kann.

4.5 Inhaltlicher Rahmen: Social Media

Inhaltlich beschäftigt sich die erste Staffel von “Sidequest Social Media” mit verschiedenen Perspektiven auf Social Media und digitale Öffentlichkeit. Dafür wurden fünf Gäste ausgewählt, die jeweils unterschiedliche Erfahrungen mit Plattformen, Content-Produktion und medialer Selbstpräsentation mitbringen. Die Gästeauswahl sollte bewusst nicht nur eine einzige Social-Media-Perspektive abbilden, sondern mehrere Zugänge zum Thema eröffnen.

Ein Gast war Fabio KQZ, der als Influencer und Content Creator auf Instagram, TikTok und YouTube aktiv ist. Im Gespräch stand besonders sein Wandel vom früheren Gamer- und Streaming-Kontext hin zu einer sportlich geprägten Content-Erzählung im Mittelpunkt. Dabei ging es unter anderem um seine einjährige Entwicklung vom Gamer zum Ironman und darum, wie sich dadurch Inhalte, Selbstbild und Plattformauftritt verändern können.

Ein weiterer Gast war David Nickel, Foto- und Videograf sowie Content Creator mit einem Instagram-Account von etwa 5.000 Followern zum Zeitpunkt der Produktion. Er ist unter anderem in der Techno-Szene tätig und hat als Tour-Fotograf verschiedene Künstler begleitet, darunter 2hot2play. Durch ihn konnte die Perspektive der professionellen Foto- und Videoproduktion für Social Media eingebunden werden, insbesondere im Zusammenhang mit Event-, Musik- und Tourcontent.

Vincent Dreker brachte ebenfalls die Perspektive eines Foto- und Videografen ein. Seine Arbeit liegt unter anderem im Bereich Konzerte, Techno-Events und zunehmend Sportveranstaltungen. Da er auch im Umfeld von Fabio KQZ fotografisch und videografisch gearbeitet hat, ergab sich eine Verbindung zwischen Creator-Perspektive und Produktionsperspektive. Inhaltlich konnte dadurch thematisiert werden, wie Social-Media-Auftritte nicht nur durch die Person vor der Kamera, sondern auch durch die Bildproduktion dahinter geprägt werden.

Mit Lilli Meier wurde eine Perspektive aus dem Influencer-Management eingebunden. Sie arbeitete während ihres Studiums mehrere Jahre bei WeCreate. Im Gespräch ging es dadurch stärker um die professionelle Zusammenarbeit zwischen Influencerinnen, Influencern und Marken. Diese Folge erweitert den inhaltlichen Rahmen der Staffel, weil Social Media nicht nur aus Creator-Sicht, sondern auch aus Management- und Markenperspektive betrachtet wird.

Als fünfter Gast war Lauris Johnki Teil der Staffel. Er ist Content Creator und Cutter und sprach über seine Erfahrungen im Streaming- und Clip-Kontext. Dazu gehörte seine frühere Arbeit im Umfeld eines großen deutschen Streamers, bei der Streams für YouTube aufbereitet und Clips erstellt wurden. Gleichzeitig ging es um seinen eigenen Schritt, selbst stärker als Creator sichtbar zu werden. Sein aktueller Content verbindet unter anderem Laufen mit provokanten oder aufmerksamkeitsstarken Elementen und eröffnet damit eine Perspektive auf Inszenierung, Zuspitzung und Aufmerksamkeit in sozialen Medien.

Die fünf Gespräche decken damit unterschiedliche Bereiche von Social Media ab: persönliche Content-Entwicklung, Plattformwechsel, Foto- und Videoproduktion, Event- und Sportcontent, Influencer-Management, Streaming, Schnitt und Clip-Erstellung. Dadurch entsteht ein inhaltlicher Rahmen, der Social Media nicht nur als Plattformnutzung versteht, sondern als Zusammenspiel von Persönlichkeit, Produktion, Strategie, Sichtbarkeit und medialer Gestaltung.

4.6 Visuelles Konzept

4.6.1 Grundidee des visuellen Konzepts

Das visuelle Konzept von “Sidequest Social Media” wurde aus den zuvor abgeleiteten Gestaltungskriterien entwickelt. Ziel war es, ein Set zu schaffen, das einerseits eine wiedererkennbare Formatidentität besitzt und andererseits genug Offenheit bietet, um die einzelnen Gäste und Folgen individuell einzubinden. Die visuelle Gestaltung sollte dabei nicht nur als Hintergrund dienen, sondern die Gesprächssituation unterstützen und den Videopodcast als eigenständiges audiovisuelles Format erkennbar machen.

Das Format sollte visuell eine lockere, persönliche und zugleich bewusst gestaltete Gesprächsatmosphäre erzeugen. Da “Sidequest Social Media” thematisch im Bereich Social Media, Content Creation und digitaler Selbstpräsentation verortet ist, sollte auch das Bild nicht neutral oder rein funktional wirken. Stattdessen sollte das Set einen kreativen und mediennahen

Eindruck vermitteln, ohne vom Gespräch abzulenken. Die daraus entwickelte visuelle Grundgestaltung ist beispielhaft anhand der ersten Folge mit David in Abbildung 10 zu sehen.



Abbildung 10

4.6.2 Setgestaltung und variable Gast-Elemente

Ein zentrales Element des Sets war ein Regal im Hintergrund beziehungsweise zwischen den beiden Gesprächspositionen. Dieses Regal übernahm mehrere Funktionen. Zum einen strukturierte es den Raum und bildete einen wiederkehrenden visuellen Mittelpunkt des Formats. Zum anderen bot es Platz für variable Elemente, die pro Folge angepasst werden konnten.

Die Gäste hatten die Möglichkeit, bis zu drei persönliche oder thematisch passende Gegenstände mitzubringen, die in das Regal integriert wurden. Die meisten Gäste brachten drei Gegenstände mit, ein Gast brachte einen Gegenstand mit. Dadurch blieb die Grundstruktur des Sets gleich, während jede Folge eine individuelle Ebene erhielt. Das Set wurde somit nicht vollständig statisch angelegt, sondern als wiedererkennbarer Rahmen mit veränderbaren Details.

Die übrige Setgestaltung arbeitete mit einer dunkleren Grundstimmung, Pflanzen, Sitzmöbeln, dekorativen Elementen und sichtbaren Mikrofonen. Die dunkle Hintergrundgestaltung sorgte dafür, dass Personen, Lichtquellen und einzelne Objekte stärker hervortreten konnten. Pflanzen

und Dekorationselemente lockerten den Raum optisch auf und verhinderten, dass das Set zu technisch oder leer wirkte. Die sichtbaren Mikrofone machten die Podcast-Situation unmittelbar erkennbar und wurden nicht versteckt, sondern bewusst als Teil der visuellen Gestaltung akzeptiert.

4.6.3 Lichtgestaltung und Atmosphäre

Auch die Lichtgestaltung war ein wichtiger Bestandteil des visuellen Konzepts. Neben der grundsätzlichen Sichtbarkeit von Host und Gast sollte das Licht eine eigene Atmosphäre erzeugen. Dafür wurden zwei Stehlampen als indirekte Lichtquellen eingesetzt. Sie beleuchteten nicht direkt die Personen, sondern strukturierten den Hintergrund und gaben dem Set mehr Tiefe.

Zusätzlich wurde ein grüner Lichtakzent von unten eingesetzt. Dieser farbige Akzent sollte dem Set eine eigene visuelle Stimmung geben und zur Wiedererkennbarkeit des Formats beitragen. Die Farbigkeit war dabei nicht als dominantes Gestaltungselement gedacht, sondern als punktueller Akzent innerhalb einer insgesamt dunkleren und atmosphärischen Bildgestaltung.

Die Lichtgestaltung sollte einen Mittelweg zwischen Atmosphäre und Klarheit herstellen. Einerseits sollte das Bild nicht neutral oder rein funktional wirken, sondern eine bewusst gestaltete, leicht stimmungsvolle Umgebung erzeugen. Andererseits mussten Mimik, Gesichtsausdruck und Körperhaltung der Gesprächspartner weiterhin gut sichtbar bleiben. Das Licht wurde deshalb nicht nur als technische Ausleuchtung verstanden, sondern als gestalterisches Mittel, das die Wahrnehmung des Raumes und die Wirkung des Formats beeinflusst. Die Umsetzung dieser Lichtgestaltung zeigt sich beispielhaft in der Folge mit Jana. In beiden Kameraperspektiven bleiben die Personen klar erkennbar, während die Hintergrundbeleuchtung und die farbigen Akzente die gewünschte Atmosphäre erzeugen (vgl. Abb. 11).

4.6.4 Kameraperspektiven und Bildkomposition

Für die Bildgestaltung wurden drei Kameraperspektiven eingesetzt: eine Totale sowie zwei halbnahe Einzelperspektiven von Host und Gast. Die Totale zeigte beide Gesprächspartner gemeinsam im Set und machte die räumliche Beziehung zwischen ihnen sichtbar. Gleichzeitig wurde in dieser Einstellung das Regal als zentrales Setelement erkennbar. Die beiden

halbnahen Einzelperspektiven konzentrierten sich stärker auf die jeweilige Person. Dadurch konnten Gesicht, Oberkörper, Mikrofon, Gestik und Reaktionen besser sichtbar gemacht werden. Die drei Perspektiven bilden gemeinsam die grundlegende Mehrkamera-Auflösung des Formats und sind in Abbildung 11,12,13 am Beispiel der Folge mit Fabio gegenübergestellt.



Abbildung 11



Abbildung 12



Abbildung 13

Bei der Komposition der Einzelbilder wurde auf eine ruhige und ausgewogene Kadrierung geachtet. Die Drittelregel diente dabei als Orientierung für die Positionierung von Kopf, Blickrichtung und Oberkörper. Host und Gast sollten nicht zufällig im Bild stehen, sondern innerhalb einer geordneten Bildfläche erscheinen. Gleichzeitig blieb ausreichend Raum für Mikrofon, Körperhaltung und Hintergrundelemente. Dadurch sollten die Personen klar im Mittelpunkt stehen, ohne vom gestalteten Set getrennt zu wirken. Die Anwendung der Drittelregel wird in Abbildung 14,15 anhand zweier Einzelperspektiven aus der Folge mit Lilly durch ein nachträglich eingeblendetes Gestaltungsraster verdeutlicht.



Abbildung 14



Abbildung 15

Die Kameraperspektiven sollten zusammen eine klare visuelle Gesprächsstruktur erzeugen. Die Totale diente der Orientierung, während die Einzelperspektiven Nähe und Aufmerksamkeit auf die sprechende Person herstellten. Dadurch konnte im Schnitt zwischen Übersicht, Sprecherfokus und Reaktionen gewechselt werden.

4.6.5 Grafische Gestaltung und visuelle Formatidentität

Neben Set-, Licht- und Bildgestaltung wurde auch eine grafische Identität für „Sidequest Social Media“ entwickelt. Ein zentrales Element ist das Cover des Formats. Die dort eingesetzte Farbgestaltung greift die Farbwelt des Sets auf. Insbesondere der farbige Akzent, der bereits in der Lichtgestaltung verwendet wird, findet sich in der grafischen Gestaltung wieder. Dadurch sollten Set und Cover nicht als voneinander getrennte Gestaltungselemente erscheinen, sondern als Bestandteile einer gemeinsamen visuellen Formatidentität. Das entwickelte Cover und die darin verwendete Farbwelt sind in Abbildung 16 dargestellt.



Abbildung 16

Die Farbwelt wurde darüber hinaus für weitere Ausspielungsformen des Formats übernommen. Bei den Social-Media-Clips wird sie unter anderem in der Gestaltung der Untertitel aufgegriffen. Dadurch entsteht eine visuelle Verbindung zwischen dem Cover, der vollständigen Videopodcast-Folge und den daraus entwickelten Kurzformaten. Wiederkehrende Farben dienen somit als verbindendes Gestaltungselement über unterschiedliche Veröffentlichungsformen hinweg.

4.6.6 Vorbereitung der Social-Media-Verwertung

Ein weiterer Aspekt des visuellen Konzepts war die spätere Nutzbarkeit des Materials für Social-Media-Clips. Bereits bei der Aufnahme wurde darauf geachtet, dass die Einzelperspektiven auch in einem 9:16-Ausschnitt funktionieren können. Die Personen

mussten deshalb so positioniert sein, dass sie im Hochformat nicht abgeschnitten wirken und zentrale Bildelemente weiterhin erkennbar bleiben.

Die Clip-Erstellung wurde somit nicht erst nachträglich in der Postproduktion berücksichtigt, sondern bereits in der Bildgestaltung mitgedacht. Besonders die halbnahen Einzelperspektiven waren dafür relevant, weil sie sich besser für vertikale Ausschnitte eignen als eine reine Totale. Für die Social-Media-Clips konnten die Personen dadurch klar sichtbar bleiben, während Untertitel und wiederkehrende Farben später als grafische Elemente ergänzt wurden.

Zusammenfassend basiert das visuelle Konzept von “Sidequest Social Media” auf einer Verbindung aus wiedererkennbarer Setstruktur, variablen Gast-Elementen, atmosphärischer Lichtgestaltung und klarer Mehrkamera-Auflösung. Das Set sollte eine lockere und persönliche Gesprächsatmosphäre erzeugen, zugleich aber bewusst gestaltet und als Format wiedererkennbar bleiben. Die visuelle Ebene wurde dadurch nicht nur als Ergänzung zum Ton eingesetzt, sondern als zentraler Bestandteil der Formatentwicklung.

4.7 Technisches Setup

Das technische Setup von “Sidequest Social Media” wurde so gewählt, dass die zuvor entwickelte visuelle Struktur möglichst zuverlässig umgesetzt werden konnte. Da das Format als Videopodcast im Interviewstil angelegt ist, mussten Bild, Ton und Licht so aufgebaut werden, dass ein längeres Gespräch kontinuierlich aufgezeichnet und anschließend flexibel geschnitten werden konnte. Im Vordergrund stand dabei ein Setup, das sowohl eine klare Gesprächsauflösung als auch eine gute Tonqualität und eine wiedererkennbare Bildwirkung ermöglicht.

Für die Bildaufnahme wurde mit drei Kameras gearbeitet. Die beiden Einzelperspektiven von Host und Gast wurden jeweils mit einer Sony FX3 aufgenommen. Für die Totale wurde eine Sony Alpha 7 IV eingesetzt. Durch diese Aufteilung konnten drei unterschiedliche Bildebenen erzeugt werden: eine gemeinsame Einstellung, die Host und Gast im Set zeigt, sowie zwei halbnahe Einzelperspektiven, die stärker auf die jeweilige Person fokussiert sind. Das Drei-Kamera-Setup ermöglichte es, das Gespräch durchgehend aufzuzeichnen und im Schnitt zwischen Übersicht, Sprecherfokus und Reaktionen zu wechseln.

Für die halbnahen Einzelperspektiven von Host und Gast wurde mit einer Brennweite von ungefähr 55 mm gearbeitet. Dafür wurde ein Sony-G-Master-Zoomobjektiv mit durchgehender Blende 2.8 eingesetzt. Die Brennweite ermöglichte eine Bildwirkung, bei der Gesicht, Oberkörper, Mikrofon und Teile des Hintergrunds sichtbar bleiben, ohne dass die Person zu weit entfernt wirkt. Dadurch eigneten sich die Einzelperspektiven sowohl für die Langfolge als auch für spätere Social-Media-Ausschnitte im Hochformat.

Für die Totale wurde ein Sony-G-Master-Weitwinkel-Zoomobjektiv im Bereich von 16 bis 28 mm mit durchgehender Blende 2.8 verwendet. Diese Brennweite war für die gemeinsame Einstellung sinnvoll, da Host, Gast und zentrale Setelemente gemeinsam im Bild sichtbar gemacht werden konnten. Besonders das Regal als wiederkehrendes Element des Sets wurde dadurch in die Gesamtkomposition eingebunden. Die Totale diente somit nicht nur der räumlichen Orientierung, sondern machte auch die Setgestaltung als Teil des Formats sichtbar.

Die Tonaufnahme erfolgte mit zwei Shure SM7B-Mikrofonen. Diese wurden jeweils für Host und Gast eingesetzt und sichtbar im Bild platziert. Die Mikrofone waren an Mikrofonstativen von Gravity befestigt. Die bewusste Sichtbarkeit der Mikrofone hatte dabei nicht nur eine technische, sondern auch eine gestalterische Funktion: Sie machte die Podcast-Situation für das Publikum unmittelbar erkennbar und wurde als Teil des visuellen Erscheinungsbildes akzeptiert. Aufgezeichnet wurde der Ton über einen Zoom H6-Rekorder.

Das Lichtsetup bestand aus einer Hauptlichtquelle und mehreren Akzent- beziehungsweise Hintergrundlichtern. Als Hauptlicht wurde eine Aputure LS 600d verwendet, kombiniert mit einer Aputure Lantern 90 als Lichtformer. Durch die Lantern entstand ein weicheres und breiter gestreutes Licht, das sich für eine Gesprächssituation mit zwei Personen eignete. Zusätzlich wurden zwei Nanlite PavoTube II 6C eingesetzt, um farbige Lichtakzente im Set zu erzeugen. Diese Lichtspitzen unterstützten die atmosphärische Wirkung des Sets und griffen die zuvor entwickelte Idee eines farbigen Akzents auf. Ergänzend kamen zwei Stehlampen als indirekte Lichtquellen im Hintergrund zum Einsatz, ebenso wie eine Lavalampe.

Bild und Ton wurden getrennt aufgezeichnet und anschließend in der Postproduktion synchronisiert. Zur Synchronisation wurde beim Dreh geklatscht, sodass ein klarer Synchronisationspunkt zwischen Kamera- und Tonspuren entstand. Zusätzlich konnten die Tonspuren im Schnittprogramm automatisch aufeinander abgestimmt werden. Diese

Arbeitsweise ermöglichte eine saubere Verbindung der drei Kameraperspektiven mit der externen Tonaufnahme.

Die Postproduktion erfolgte in DaVinci Resolve. Dort wurden die Kameraperspektiven geschnitten, der Ton bearbeitet und die einzelnen Folgen weiterverarbeitet. Auch die Social-Media-Clips wurden nach der Auswahl geeigneter Ausschnitte in DaVinci Resolve final geschnitten. Zur Unterstützung der Clip-Erstellung wurde zusätzlich OpusClip eingesetzt. Die Software scannte die Podcastfolgen und schlug potenzielle Highlights vor. Diese Vorschläge wurden anschließend nicht unverändert übernommen, sondern in DaVinci Resolve geprüft, angepasst und gestalterisch weiterbearbeitet.

Das technische Setup war damit eng mit den gestalterischen Anforderungen des Formats verbunden. Die drei Kameraperspektiven ermöglichten eine strukturierte visuelle Gesprächsauflösung, die Mikrofone sorgten für eine klare und podcasttypische Tonsituation, und das Lichtsetup unterstützte die gewünschte Atmosphäre des Sets. Die Kombination aus Kamera, Ton, Licht und Schnittsoftware bildete die technische Grundlage dafür, dass “Sidequest Social Media” nicht nur als Gespräch, sondern als gestalteter Videopodcast umgesetzt werden konnte.

4.8 Aufnahmeprozess

Die Produktion von “Sidequest Social Media” umfasste fünf Folgen, die an fünf unterschiedlichen Drehtagen aufgenommen wurden. Die reine Aufnahmezeit variierte je nach Folge zwischen etwa 50 Minuten und eineinhalb Stunden. Hinzu kamen jeweils Vor- und Nachbereitungszeiten am Set, insbesondere für Kameras, Licht, Ton, Gastgegenstände und Datensicherung.

Die Aufnahmen fanden in zwei Produktionsblöcken statt. Die ersten drei Folgen wurden in einem Studio aufgenommen, die vierte und fünfte Folge in einem zweiten Studio. Dadurch musste das Set nicht für jede einzelne Folge vollständig neu aufgebaut werden, sondern insgesamt zweimal. Dennoch stellte gerade dieser zweimalige Auf- und Abbau eine gestalterische Herausforderung dar. Da das Format visuell möglichst einheitlich wirken sollte, mussten Sitzpositionen, Regal, Lichtquellen, Kameraperspektiven und Dekoration bei beiden Setaufbauten möglichst ähnlich eingerichtet werden. Eine vollständig identische

Rekonstruktion war durch die unterschiedlichen Studiosituationen nicht möglich, das Ziel war jedoch, die visuelle Grundstruktur des Formats so konstant wie möglich beizubehalten.

Bei den beiden größeren Studioaufbauten erhielt ich jeweils Unterstützung durch eine weitere Person, die vor allem bei der Kameraeinrichtung und der allgemeinen Setvorbereitung assistierte. Diese Unterstützung war besonders hilfreich, um den Aufbau effizienter umzusetzen und die technischen Komponenten schneller aufeinander abzustimmen. Da mit drei Kameras, separater Tonaufnahme und mehreren Lichtquellen gearbeitet wurde, war die Vorbereitung des Sets ein wichtiger Bestandteil des gesamten Produktionsprozesses.

Vor den eigentlichen Aufnahmen wurden Testdrehs durchgeführt. Dabei wurden Bild und Ton kontrolliert und anschließend Anpassungen vorgenommen. Besonders relevant waren dabei die Kameraperspektiven und Bildausschnitte. Die Testaufnahmen halfen dabei, die halbnahen Einstellungen von Host und Gast sowie die Totale so einzurichten, dass Personen, Mikrofone, Setelemente und Hintergrund stimmig im Bild sichtbar waren. Gleichzeitig konnte überprüft werden, ob die Bildgestaltung auch im Hinblick auf spätere Social-Media-Ausschnitte funktionierte.

Die Gäste wurden vorab grundsätzlich über das Thema und die grobe Richtung des Gesprächs informiert. Sie entschieden jedoch selbst, welche persönlichen oder thematisch passenden Gegenstände sie mitbringen wollten. Diese Gegenstände wurden entweder vor der Aufnahme direkt im Regal platziert oder während der Folge in das Set eingebunden. Dadurch wurden die Objekte nicht nur als Dekoration genutzt, sondern konnten teilweise auch Bestandteil des Gesprächs werden.

Für die Gesprächsführung wurden im Vorfeld Themengebiete und teilweise konkrete Fragen vorbereitet. Gleichzeitig sollte das Format nicht wie ein starres Interview wirken. Die Gespräche wurden daher offen geführt und konnten sich je nach Gast und Gesprächsverlauf entwickeln. Diese Arbeitsweise passte zur angestrebten Mischung aus Interviewformat, Gesprächspodcast und lockerem Talkformat. Der vorbereitete Rahmen sorgte für thematische Orientierung, während die offene Gesprächsführung persönliche Reaktionen und spontane Gesprächsmomente ermöglichte.

Während der Aufnahme liefen die drei Kameraperspektiven parallel. Bild und Ton wurden getrennt aufgezeichnet und durch einen Synchronisationsklatscher sowie über die Tonspuren

im Schnittprogramm zusammengeführt. Der parallele Einsatz mehrerer Perspektiven war nicht nur gestalterisch relevant, sondern auch praktisch hilfreich. Vereinzelt kam es während der Produktion zu technischen Problemen, etwa durch Überhitzung einer Kamera oder durch Akkuprobleme. Diese Schwierigkeiten konnten durch organisatorische Anpassungen, unter anderem durch den Einsatz eines Lüfters, reduziert werden. Da zusätzlich mehrere Kameraperspektiven vorhanden waren, konnten einzelne technische Ausfälle im späteren Schnitt besser abgefedert werden.

Nach jedem Dreh wurde das Material gesichtet und gesichert. Eine besondere Herausforderung bestand dabei in den großen Datenmengen. Da in 4K und Log aufgezeichnet wurde und pro Folge drei Kameraspuren entstanden, kamen teilweise Datenmengen von etwa 100 Gigabyte pro Folge zusammen. Die Datensicherung war deshalb ein wichtiger Teil der Nachbereitung und musste direkt im Anschluss an die Aufnahme berücksichtigt werden.

Der Aufnahmeprozess zeigte damit, dass die Umsetzung eines Videopodcasts deutlich über die eigentliche Gesprächsaufnahme hinausgeht. Neben der Moderation mussten Set, Licht, Bildausschnitt, Ton, technische Stabilität, Gastgegenstände und spätere Weiterverarbeitung mitgedacht werden. Besonders die Verbindung aus mehreren Drehtagen, zwei Studiosituationen und variablen Setelementen machte deutlich, dass visuelle Kontinuität aktiv hergestellt werden musste. Der Aufnahmeprozess bildete damit die praktische Grundlage dafür, dass “Sidequest Social Media” als zusammenhängendes Format erkennbar blieb.

4.9 Schnitt und Postproduktion

Nach der Aufnahme wurden die aufgezeichneten Bild- und Tonspuren in DaVinci Resolve weiterverarbeitet. Da jede Folge mit drei Kameraperspektiven und separater Tonaufnahme produziert wurde, bestand der erste Arbeitsschritt darin, das Material zu sichten, zu ordnen und zu synchronisieren. Die Synchronisation erfolgte mithilfe des Synchronisationsklatschers sowie über die Tonspuren im Schnittprogramm. Dadurch konnten die Totale sowie die beiden Einzelperspektiven von Host und Gast mit der externen Tonaufnahme verbunden werden.

Der Schnitt der Langfolgen orientierte sich an der Gesprächsdynamik. Die drei Kameraperspektiven wurden nicht willkürlich eingesetzt, sondern entsprechend der jeweiligen Situation im Gespräch ausgewählt. Die Totale diente vor allem dazu, die räumliche Situation zwischen Host und Gast sichtbar zu machen und das Set als gemeinsamen Gesprächsraum zu

zeigen. Die halbnahen Einzelperspektiven wurden genutzt, um die jeweils sprechende Person stärker in den Fokus zu rücken. Gleichzeitig konnten dadurch Mimik, Gestik, Blickrichtung und Reaktionen deutlicher sichtbar gemacht werden.

Ein wichtiger Bestandteil des Schnitts war der Wechsel zwischen Sprecherfokus und Reaktion. Wenn eine Person sprach, wurde häufig die entsprechende Einzelperspektive genutzt. In anderen Momenten konnten Reaktionen der zuhörenden Person sichtbar gemacht werden. Dadurch sollte die visuelle Ebene nicht nur den gesprochenen Inhalt begleiten, sondern auch Gesprächsatmosphäre, Aufmerksamkeit und nonverbale Kommunikation vermitteln. Der Schnitt hatte somit die Funktion, den Gesprächsverlauf visuell zu strukturieren und die Wirkung des Interviews zu unterstützen.

Die Folgen wurden nicht vollständig roh übernommen, sondern in der Postproduktion verdichtet. Längere Pausen, technische Unterbrechungen oder Stellen, die den Gesprächsfluss störten, konnten gekürzt werden. Gleichzeitig sollte der Charakter eines natürlichen Gesprächs erhalten bleiben. Ziel war daher kein stark fragmentierter oder übermäßig schneller Schnitt, sondern eine Bearbeitung, die die Verständlichkeit verbessert und die Gesprächsdynamik unterstützt, ohne den lockeren Podcastcharakter zu verlieren.

Auch die Tonspuren wurden in der Postproduktion bearbeitet. Dabei kam für Host und Gäste grundsätzlich eine ähnliche Bearbeitungskette zum Einsatz, die je nach Stimme leicht angepasst wurde. Zur Reduzierung von Hintergrundrauschen wurde der “NS1 Noise Suppressor” eingesetzt. Ein De-Esser diente dazu, auffällige S-Laute abzumildern, während ein leichter Kompressor Pegelunterschiede innerhalb der Sprachaufnahme reduzierte. Zusätzlich wurde der Frequenzgang mithilfe eines Equalizers angepasst, um die Sprachverständlichkeit zu verbessern und die einzelnen Stimmen ausgeglichener klingen zu lassen. Die Einstellungen wurden nicht unverändert auf alle Aufnahmen übertragen, sondern abhängig von der jeweiligen Stimme leicht angepasst. Bei der Folge mit Lilli waren insbesondere stärkere individuelle Anpassungen im Equalizer notwendig.

Für die visuelle Nachbearbeitung spielte neben dem Schnitt auch die Farbgestaltung eine wichtige Rolle. Da das Material im Log-Format aufgezeichnet wurde, wurde zunächst ein Log-to-Rec.709-LUT angewendet. Dieser diente als technische Grundlage, um das kontrastarme und entsättigte Log-Material in den für die Wiedergabe üblichen Rec.709-Farbraum zu überführen. Anschließend wurden die Aufnahmen weiter angepasst. Dabei wurde die Sättigung

leicht erhöht, der Kontrast verstärkt und die dunklen Bildbereiche beziehungsweise Schatten etwas abgesenkt, um dem Bild mehr Tiefe zu geben.

Die grundlegende Bearbeitung war bei den verschiedenen Kameraperspektiven und Folgen ähnlich. Trotzdem waren teilweise individuelle Anpassungen notwendig. Besonders durch den Wechsel des Studios beziehungsweise den erneuten Aufbau des Sets unterschieden sich Lichtverhältnisse und Bildwirkung geringfügig. Diese Unterschiede wurden in der Farbkorrektur ausgeglichen, damit die einzelnen Kameraperspektiven und Folgen möglichst einheitlich wirken. Ziel war dabei kein stark stilisiertes Colorgrading, sondern ein konsistenter Look, der die bereits durch Set- und Lichtgestaltung entwickelte visuelle Atmosphäre unterstützt.





(Abbildung 4: Folge 4 SQSM – Vincent Dreker – Color Grading)

In ausgewählten Folgen wurden zusätzlich Foto- und Videoeinblendungen eingesetzt. Diese Einblendungen wurden nicht in jeder Folge verwendet, sondern nur dort, wo sie den Gesprächsinhalt sinnvoll unterstützen konnten. Gerade beim Thema Social Media können konkrete Bilder, Videos oder visuelle Beispiele hilfreich sein, weil sie einen Gesprächsgegenstand für das Publikum unmittelbar sichtbar machen. Ein Beispiel dafür findet sich in der ersten Folge mit Fabio, in der während des Gesprächs über ein bestimmtes Social-Media-Reel gesprochen wurde. Dieses konnte in der Postproduktion direkt eingeblendet werden, sodass sich das Publikum auf denselben visuellen Inhalt beziehen konnte wie die Gesprächspartner. Die Einblendung erfüllte damit eine konkrete inhaltliche Funktion und diente nicht lediglich der visuellen Abwechslung.

Neben den Langfolgen wurden aus dem Material auch Social-Media-Clips erstellt. Die Clip-Erstellung wurde bereits während der Aufnahme mitgedacht, da die Einzelperspektiven so angelegt waren, dass sie auch in einem 9:16-Ausschnitt funktionieren können. In der Postproduktion wurden geeignete Gesprächsausschnitte ausgewählt, gekürzt und für das Hochformat aufbereitet. Dabei musste darauf geachtet werden, dass die sprechende Person im vertikalen Bildausschnitt klar sichtbar bleibt und wichtige Bildelemente nicht abgeschnitten werden.

Zur Unterstützung der Clip-Auswahl wurde OpusClip eingesetzt. Die KI-Software analysierte die Folgen und schlug potenzielle Highlights vor. Diese Vorschläge dienten jedoch nur als Ausgangspunkt. Die finale Auswahl und Bearbeitung der Clips erfolgte anschließend in DaVinci Resolve. Dadurch blieb die gestalterische Kontrolle über Schnitt, Bildausschnitt, Timing und Wirkung der Clips erhalten.

Die Social-Media-Clips erhielten ein eigenes grafisches Design. Dieses basierte nicht auf einem Logo, sondern auf wiederkehrenden Farben und einem einheitlichen Untertitelstil. Weitere Einblendungen wurden in den Clips nicht verwendet. Die Untertitel erfüllten dabei mehrere Funktionen: Sie machten die gesprochenen Inhalte auch ohne Ton verständlich, strukturierten den Clip visuell und trugen zur Wiedererkennbarkeit der Clipgestaltung bei. Durch den einheitlichen Stil der Untertitel und Farben entstand eine gestalterische Verbindung zwischen den einzelnen Clips. Die konkrete Umsetzung zeigt Abbildung 10 anhand eines fertig bearbeiteten Social-Media-Clips aus der eigenen Produktion.

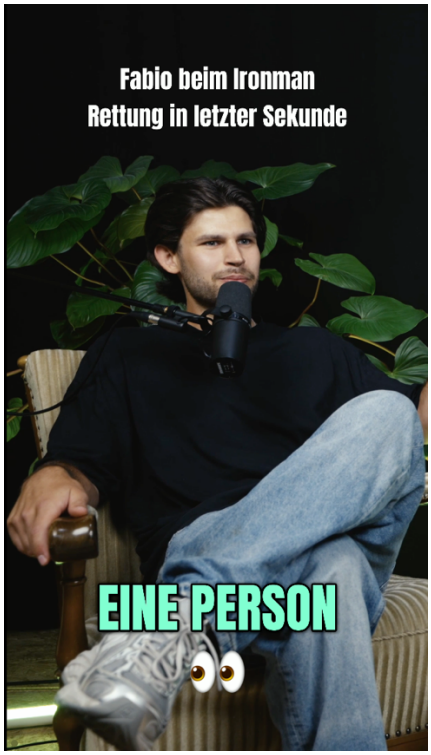


Abbildung 17

Der Schnitt- und Postproduktionsprozess war damit ein wesentlicher Teil der audiovisuellen Gestaltung von “Sidequest Social Media”. Erst durch die Auswahl der Kameraperspektiven, die Verdichtung der Gespräche, die Ton- und Bildbearbeitung, die gezielte Einbindung visueller Zusatzinhalte und die Erstellung von Social-Media-Clips wurde aus dem aufgenommenen Material ein gestaltetes Videopodcast-Format. Die Postproduktion erweiterte die visuelle Ebene damit nicht nur technisch, sondern auch gestalterisch.

5. Reflexion und Fazit

5.1 Zusammenfassung und Einordnung der zentralen Ergebnisse

Ziel dieser Arbeit war es, zu untersuchen, welchen gestalterischen Mehrwert die visuelle Ebene eines Videopodcasts im Interviewformat im Vergleich zu einem reinen Audiopodcast bieten kann. Der Videopodcast wurde dabei nicht als grundsätzlich besseres Format verstanden, sondern als audiovisuelle Erweiterung eines weiterhin wesentlich durch Stimme, Sprache und Gespräch getragenen Podcastformats.

Die theoretische Auseinandersetzung hat zunächst gezeigt, dass der Audiopodcast eine eigenständige mediale Qualität besitzt. Stimme, Klang, Gesprächsrhythmus und die

Konzentration auf das Hören können Nähe und eine persönliche Rezeptionssituation erzeugen. Gleichzeitig bleiben im reinen Audioformat visuelle Informationen wie Mimik, Gestik, Blickrichtung, Körperhaltung, sichtbare Reaktionen und die räumliche Gesprächssituation verborgen.

Daraus lässt sich als zentrale Erkenntnis ableiten, dass der gestalterische Mehrwert eines Videopodcasts nicht allein durch das Vorhandensein eines Bildes entsteht. Entscheidend ist vielmehr, ob die visuelle Ebene eine konkrete Funktion innerhalb des Gesprächs übernimmt. Kameraperspektiven, Personendarstellung, Bildkomposition, Licht, Setgestaltung, Schnitt und grafische Gestaltung können dazu beitragen, zusätzliche Informationen sichtbar zu machen, Aufmerksamkeit zu lenken, räumliche Orientierung zu schaffen und dem Format eine visuelle Identität zu geben.

Die Referenzanalyse ausgewählter Videopodcast-Formate konkretisierte diese theoretischen Gestaltungsbereiche anhand praktischer Beispiele. Dabei wurden unterschiedliche Lösungsansätze sichtbar. Der Tim Gabel Podcast zeigte eine klassische Mehrkamera-Auflösung mit räumlicher Orientierung und bewusst gestalteter Umgebung. The Joe Rogan Experience verdeutlichte die Möglichkeit, zusätzliche visuelle Inhalte direkt in ein Gespräch einzubinden. Schatz, wir müssen reden zeigte die Verbindung aus konstanten und variablen Setelementen. Clare on Air diente als Beispiel für atmosphärische Lichtgestaltung, während Baywatch Berlin zeigte, wie Podcastmaterial für Social-Media-Clips ausgewählt, verdichtet, neu kadriert und grafisch aufbereitet werden kann.

Aus Theorie und Referenzanalyse wurden für „Sidequest Social Media“ neun Gestaltungskriterien abgeleitet: eine klare Mehrkamera-Auflösung, eine auf nonverbale Präsenz ausgerichtete Personendarstellung, eine bewusste Bildkomposition, eine am Gespräch orientierte Schnittgestaltung, ein wiedererkennbares und zugleich variables Set, eine funktionale und atmosphärische Lichtgestaltung, die gezielte Einbindung zusätzlicher visueller Inhalte, eine bereits bei der Aufnahme berücksichtigte Social-Media-Verwertung sowie eine konsistente grafische Gestaltung.

Die Ergebnisse zeigen damit, dass der gestalterische Mehrwert eines Videopodcasts nicht auf einem einzelnen visuellen Element beruht. Er entsteht aus dem Zusammenspiel verschiedener Gestaltungsmittel und daraus, welche Funktion diese innerhalb des Gesprächs übernehmen.

Die neun abgeleiteten Kriterien bildeten deshalb den Bezugsrahmen für die Entwicklung und Umsetzung der eigenen Videopodcast-Staffel „Sidequest Social Media“.

5.2 Mehrwert der visuellen Ebene im eigenen Format

In der praktischen Umsetzung von „Sidequest Social Media“ zeigte sich der Mehrwert der visuellen Ebene besonders in der sichtbaren Kommunikation zwischen Host und Gästen, der räumlichen und atmosphärischen Gestaltung des Gesprächs, der Einbindung zusätzlicher visueller Inhalte sowie der Weiterverwendbarkeit des aufgenommenen Materials.

Ein wesentlicher Unterschied zum reinen Audiopodcast bestand in der sichtbaren Präsenz der Gesprächspartnerinnen und Gesprächspartner. Durch die halbnahen Einzelperspektiven konnten neben Stimme und Sprache auch Gesichtsausdruck, Gestik, Blickrichtung, Körperhaltung und Reaktionen wahrgenommen werden. Gerade beim Zuhören entstand dadurch eine zusätzliche Informationsebene. Reaktionen wie Nicken, Lachen oder Nachdenklichkeit konnten im Schnitt sichtbar gemacht werden, auch wenn die betreffende Person in diesem Moment nicht sprach.

Das Drei-Kamera-Setup unterstützte diese Funktion. Die Totale stellte die räumliche Beziehung zwischen Host und Gast her und zeigte beide Personen innerhalb des gemeinsamen Sets. Die beiden halbnahen Einzelperspektiven ermöglichten dagegen eine stärkere Konzentration auf die jeweils sprechende oder reagierende Person. Im Schnitt konnte dadurch zwischen räumlicher Orientierung, Sprecherfokus und Reaktionen gewechselt werden. Der Mehrwert entstand somit nicht aus der Anzahl der Kameras selbst, sondern aus den unterschiedlichen Funktionen, die die Perspektiven innerhalb des Gesprächs übernehmen konnten.

Auch Set- und Lichtgestaltung erweiterten die rein auditive Gesprächssituation. Das wiederkehrende Regal, die dunklere Grundgestaltung, Pflanzen, sichtbare Mikrofone, indirekte Lichtquellen und der farbige Lichtakzent bildeten einen visuellen Rahmen für die Gespräche. Gleichzeitig ermöglichten die persönlichen Gegenstände der Gäste eine Variation zwischen den Folgen. Besonders deutlich wurde diese Funktion, wenn ein Gegenstand nicht nur dekorativ im Hintergrund stand, sondern direkt in das Gespräch einbezogen wurde.

Ein konkretes Beispiel dafür zeigte sich bereits in der ersten Folge mit Fabio. Während des Gesprächs wurde über ein bestimmtes Social-Media-Reel gesprochen. Anstatt den Inhalt

ausschließlich sprachlich zu beschreiben, konnte das betreffende Reel direkt in die Folge eingeblendet werden. Dadurch erhielt das Publikum denselben visuellen Referenzpunkt, auf den sich das Gespräch bezog. Die visuelle Ebene ergänzte den gesprochenen Inhalt in diesem Fall unmittelbar und machte den Gesprächsgegenstand gleichzeitig sichtbar.

Auch weitere Foto- und Videoeinblendungen wurden dort eingesetzt, wo sie einen konkreten Bezug zum Gespräch hatten. Gerade beim Thema Social Media bot die visuelle Ebene die Möglichkeit, über Bilder oder Videos nicht nur zu sprechen, sondern diese gleichzeitig sichtbar zu machen. Die Produktion zeigte dabei auch, dass solche Einblendungen nicht in jeder Folge notwendig sind. Ihr gestalterischer Nutzen war vor allem dann erkennbar, wenn sie einen Gesprächsinhalt tatsächlich ergänzten.

Ein weiterer Nutzen entstand durch die bereits bei der Aufnahme berücksichtigte Social-Media-Verwertung. Die halbnahen Einzelperspektiven wurden so angelegt, dass sie auch für einen späteren 9:16-Ausschnitt genutzt werden konnten. Dadurch standen für die Clip-Erstellung bereits gefilmte Gesprächssituationen, sichtbare Personen, Reaktionen und Setelemente zur Verfügung. Das Material musste für das Kurzformat zwar neu ausgewählt, gekürzt, kadriert und mit Untertiteln versehen werden, die visuelle Grundlage war jedoch bereits Bestandteil der ursprünglichen Produktion.

Im eigenen Format zeigte sich der gestalterische Mehrwert der visuellen Ebene damit vor allem dort, wo das Bild eine konkrete zusätzliche Funktion übernahm. Es machte nonverbale Kommunikation sichtbar, strukturierte das Gespräch durch unterschiedliche Perspektiven, bettete Host und Gast räumlich und atmosphärisch ein, ermöglichte die direkte Darstellung visueller Gesprächsinhalte und stellte zugleich Ausgangsmaterial für weitere audiovisuelle Formate bereit.

5.3 Reflexion der gestalterischen und technischen Umsetzung

Die praktische Umsetzung von “Sidequest Social Media” zeigte, dass die zuvor entwickelten Gestaltungskriterien grundsätzlich auf ein eigenes Videopodcast-Format übertragen werden konnten. Gleichzeitig wurde deutlich, dass die Qualität der visuellen Ebene stark davon abhängt, wie konsequent einzelne Gestaltungsmittel aufeinander abgestimmt und unter realen Produktionsbedingungen umgesetzt werden.

Das Drei-Kamera-Setup erwies sich dabei als funktional. Die Kombination aus einer Totalen und zwei halbnahen Einzelperspektiven bot im Schnitt ausreichend Möglichkeiten, zwischen räumlicher Orientierung, Sprecherfokus und Reaktionen zu wechseln. Besonders die Einzelperspektiven waren wichtig, um nonverbale Kommunikation sichtbar zu machen. Gleichzeitig zeigte die Produktion, dass eine größere Zahl an Kameras allein keinen gestalterischen Mehrwert garantiert. Entscheidend war vielmehr, dass jede Perspektive eine nachvollziehbare Funktion innerhalb des Gesprächs erfüllte und Perspektivwechsel nicht nur zur visuellen Abwechslung eingesetzt wurden.

Eine zusätzliche Anforderung entstand durch die geplante Social-Media-Verwertung. Die Einzelbilder mussten sowohl für die horizontale Langfolge als auch für mögliche vertikale Ausschnitte funktionieren. Dadurch wurde die Kadrierung bereits während der Aufnahme eingeschränkt. Gesicht, Oberkörper, Mikrofon und relevante Gestik sollten auch nach einem späteren 9:16-Crop erhalten bleiben. Die frühe Berücksichtigung der Clip-Erstellung erwies sich als hilfreich, zeigte aber zugleich, dass die Gestaltung einer Langfolge und die Anforderungen eines Hochformat-Clips nicht vollständig identisch sind.

Das Setkonzept funktionierte besonders dort gut, wo die gleichbleibende Grundstruktur mit den individuellen Gastgegenständen verbunden werden konnte. Das Regal bildete einen wiederkehrenden visuellen Anker und unterstützte die Zusammengehörigkeit der Folgen. Die variablen Gegenstände ermöglichten gleichzeitig eine Individualisierung einzelner Episoden. Am stärksten war dieses Prinzip, wenn ein Objekt nicht nur im Hintergrund sichtbar blieb, sondern in das Gespräch eingebunden wurde. Dadurch erhielt die Setgestaltung eine inhaltliche Funktion.

Gleichzeitig zeigte sich hier eine Grenze des Konzepts. Die Gäste brachten unterschiedlich viele Gegenstände mit, und diese wurden nicht in jeder Folge gleich stark thematisiert. Dadurch fiel der Grad der visuellen Individualisierung zwischen den Episoden unterschiedlich aus. Für eine weitere Staffel wäre es daher sinnvoll, die Integration der Gastgegenstände noch stärker vorzubereiten und bereits im Vorfeld zu überlegen, ob und wie einzelne Objekte in das Gespräch eingebunden werden können.

Eine weitere Herausforderung bestand in der visuellen Kontinuität. Die fünf Folgen wurden an unterschiedlichen Drehtagen und in zwei Produktionsblöcken beziehungsweise Studios aufgenommen. Das Set musste dadurch zweimal vollständig aufgebaut werden. Trotz einer

möglichst ähnlichen Grundstruktur ließen sich Unterschiede in Raum, Kameraposition, Lichtwirkung und Anordnung nicht vollständig vermeiden. Für eine zukünftige Produktion wäre eine genauere Dokumentation von Kamerapositionen, Abständen, Lichtwerten und Setaufbau sinnvoll. Ein dauerhaft verfügbarer Produktionsort würde die visuelle Einheit zusätzlich erleichtern.

Auch die Lichtgestaltung erfüllte grundsätzlich ihre vorgesehene Funktion. Die Kombination aus Hauptlicht, indirekten Lichtquellen und farbigem Akzent unterstützte die angestrebte atmosphärische und mediennahe Bildwirkung. Gleichzeitig musste immer sichergestellt werden, dass die Atmosphäre nicht zulasten der Personendarstellung ging. Da gerade Gesichtsausdruck und Reaktionen einen wesentlichen Teil des visuellen Mehrwerts darstellen, hatte die ausreichende Sichtbarkeit von Host und Gast Vorrang vor einer möglichst starken Lichtstimmung.

Im Schnitt zeigte sich schließlich, wie stark die Wirkung des aufgenommenen Materials erst in der Postproduktion bestimmt wird. Die vorhandenen Kameraperspektiven ermöglichten zwar unterschiedliche Bildentscheidungen, ihre Funktion entstand jedoch erst durch die konkrete Auswahl im Schnitt. Sprecherwechsel, Reaktionen und gemeinsame Situationen konnten gezielt hervorgehoben werden. Gleichzeitig musste darauf geachtet werden, den Gesprächscharakter nicht durch zu häufige oder rein dekorative Bildwechsel zu überlagern. Die angestrebte visuelle Dynamik bestand deshalb weniger in einer hohen Schnittfrequenz als in funktionalen Perspektivwechseln.

Auch Foto- und Videoeinblendungen erwiesen sich vor allem dann als sinnvoll, wenn sie unmittelbar aus dem Gespräch heraus motiviert waren. Ihre Wirkung war geringer, wenn sie lediglich als zusätzliche visuelle Abwechslung eingesetzt worden wären. Deshalb wurden solche Elemente bewusst nicht in jeder Folge verwendet. Diese Entscheidung bestätigte, dass visuelle Gestaltung im Videopodcast nicht möglichst umfangreich, sondern funktional eingesetzt werden sollte.

Neben den gestalterischen Fragen erhöhte die audiovisuelle Produktion den technischen und organisatorischen Aufwand. Drei Kameras, separate Tonaufnahme, mehrere Lichtquellen sowie die Aufzeichnung in 4K und Log führten zu größeren Datenmengen und zusätzlichen Fehlerquellen. Während der Drehs traten unter anderem Probleme mit Kameraüberhitzung und Akkus auf. Die parallelen Kameraperspektiven boten in einzelnen Situationen zwar zusätzliche

Sicherheit für den Schnitt, gleichzeitig erforderte das gesamte Setup deutlich mehr Vorbereitung, Kontrolle und Datensicherung als eine reine Audioaufnahme.

Insgesamt konnte das entwickelte Gestaltungskonzept in der praktischen Produktion weitgehend umgesetzt werden. Besonders die Mehrkamera-Auflösung, das variable Set und die Verbindung zwischen Langfolge und Social-Media-Verwertung erwiesen sich als tragfähige Bestandteile des Formats. Gleichzeitig zeigte die Produktion Verbesserungsmöglichkeiten bei der visuellen Kontinuität, der systematischeren Einbindung der Gastgegenstände und der technischen Absicherung des Setups. Die praktische Umsetzung bestätigt damit nicht nur die Möglichkeiten der visuellen Ebene, sondern zeigt zugleich, dass ihr gestalterischer Mehrwert mit zusätzlichen Entscheidungen, Produktionsaufwand und einer konsequenten Abstimmung der einzelnen Gestaltungsmittel verbunden ist.

5.4 Social-Media-Clips und Weiterverwertung

Ein weiterer zentraler Bestandteil der praktischen Umsetzung war die Weiterverwertung des Videopodcast-Materials für Social-Media-Clips. Bereits in der theoretischen Auseinandersetzung wurde deutlich, dass ein Videopodcast nicht nur als vollständige Folge funktioniert, sondern zugleich visuelles Ausgangsmaterial für kürzere Formate bereitstellt. In der Produktion von “Sidequest Social Media” wurde dieser Gedanke bereits während der Aufnahme berücksichtigt.

Besonders relevant war dabei die Bildgestaltung der halbnahen Einzelperspektiven. Diese wurden nicht nur für die Langfolgen eingerichtet, sondern auch mit Blick auf eine spätere Hochformat-Nutzung gestaltet. Die Personen sollten so im Bild positioniert sein, dass sie auch in einem 9:16-Ausschnitt klar erkennbar bleiben. Dadurch konnte das Material in der Postproduktion für Social-Media-Clips neu kadriert werden, ohne dass wichtige Bestandteile wie Gesicht, Mikrofon, Gestik oder Körperhaltung vollständig verloren gingen.

Die Clip-Erstellung zeigte, dass Videopodcast-Material nicht einfach unverändert aus der Langfolge übernommen werden kann. Ein Gesprächsausschnitt, der innerhalb einer vollständigen Folge funktioniert, ist nicht automatisch als kurzer Clip verständlich. Für Social Media mussten die Ausschnitte daher ausgewählt, gekürzt, neu kadriert und mit Untertiteln versehen werden. Die Clips brauchten einen klaren Einstieg, eine verständliche Aussage und eine visuelle Gestaltung, die auch bei schneller mobiler Nutzung funktioniert.

Zur Unterstützung der Clip-Auswahl wurde OpusClip eingesetzt. Die Software schlug mögliche Highlights aus den Folgen vor. Diese Vorschläge waren hilfreich, um potenziell geeignete Stellen schneller zu finden, ersetzten aber keine gestalterische Entscheidung. Die finale Auswahl und Bearbeitung erfolgte weiterhin in DaVinci Resolve. Dadurch blieb die Kontrolle über Timing, Bildausschnitt, inhaltlichen Zusammenhang und visuelle Wirkung erhalten. Die KI-Unterstützung wurde somit nicht als automatischer Ersatz für Schnittentscheidungen genutzt, sondern als Hilfsmittel innerhalb des Postproduktionsprozesses.

Die Untertitel waren ein besonders wichtiger Bestandteil der Clipgestaltung. Sie machten die gesprochenen Inhalte auch ohne Ton verständlich und gaben den Clips zugleich eine wiedererkennbare grafische Form. Da auf zusätzliche Logos oder weitere Einblendungen verzichtet wurde, übernahmen wiederkehrende Farben und ein einheitlicher Untertitelstil eine wichtige Funktion für die Formatidentität. Die Clips sollten dadurch nicht überladen wirken, sondern klar, lesbar und wiedererkennbar bleiben.

Im Vergleich zu einem reinen Audiopodcast zeigte sich hier ein deutlicher Mehrwert der Videopodcast-Produktion. Für einen Audiopodcast müsste visuelles Material für Social Media erst nachträglich erstellt werden, etwa durch Grafiken, Standbilder oder animierte Audiogramme. Bei “Sidequest Social Media” standen dagegen bereits gefilmte Gesprächssituationen, sichtbare Personen, Reaktionen, Setelemente und teilweise zusätzliche visuelle Inhalte zur Verfügung. Dadurch konnte aus dem Material eine zweite Nutzungsebene entstehen, die eng mit der ursprünglichen Produktion verbunden blieb.

Gleichzeitig wurde deutlich, dass die Weiterverwertung für Social Media eigene gestalterische Anforderungen mit sich bringt. Die Langfolge und der Clip funktionieren nicht nach denselben Regeln. Während die vollständige Folge mehr Gesprächsverlauf, Kontext und längere Erklärungen zulässt, muss der Clip stärker verdichten. Dadurch entsteht immer die Herausforderung, kurze Ausschnitte aufmerksamkeitsstark zu gestalten, ohne Aussagen zu verkürzen oder aus dem Zusammenhang zu reißen. Die Clip-Erstellung verlangt deshalb eine Balance zwischen Zuspitzung, Verständlichkeit und inhaltlicher Genauigkeit.

Insgesamt zeigte die Arbeit an den Social-Media-Clips, dass die visuelle Ebene den Produktionswert eines Videopodcasts erweitert. Das Material konnte nicht nur als vollständige Folge genutzt, sondern für weitere audiovisuelle Inhalte aufbereitet werden. Dieser Mehrwert

entsteht jedoch nicht automatisch durch die Aufnahme selbst. Er setzt voraus, dass Social Media bereits bei Bildgestaltung, Aufnahme und Postproduktion mitgedacht wird.

5.5 Grenzen der Arbeit

Trotz der gewonnenen Erkenntnisse ist die Arbeit in mehreren Punkten begrenzt. Eine zentrale Grenze liegt darin, dass keine empirische Untersuchung der Zuschauerinnen und Zuschauer durchgeführt wurde. Es wurde also nicht untersucht, wie das Publikum die Videopodcast-Folgen tatsächlich wahrnimmt, ob die visuelle Ebene als Mehrwert empfunden wird oder welche Gestaltungselemente besonders stark wirken. Die Bewertung des Mehrwerts erfolgt daher auf Grundlage der theoretischen Auseinandersetzung, der Referenzanalyse und der eigenen gestalterischen Reflexion.

Eine weitere Grenze besteht darin, dass kein direkter Wirkungsvergleich zwischen einer Audio- und einer Videoversion derselben Folge durchgeführt wurde. Die Arbeit untersucht zwar den gestalterischen Mehrwert der visuellen Ebene im Vergleich zum reinen Audiopodcast, vergleicht aber nicht systematisch zwei unterschiedliche Rezeptionsformen mit Testpersonen oder Nutzungsdaten. Dadurch kann nicht abschließend gesagt werden, ob das Publikum einzelne Inhalte in der Videoversion tatsächlich besser, intensiver oder anders wahrnimmt als in einer reinen Audioversion.

Auch die Referenzanalyse ist begrenzt. Die ausgewählten Videopodcast-Formate und Podcast-Clips dienen als gestalterische Orientierung, bilden aber keine repräsentative Marktanalyse. Es wurden bewusst einzelne Beispiele betrachtet, die für bestimmte Gestaltungsbereiche relevant waren, etwa Mehrkamera-Auflösung, Setgestaltung, Licht, visuelle Einblendungen oder Social-Media-Verwertung. Die daraus abgeleiteten Kriterien sind daher nicht allgemein gültig für alle Videopodcast-Formate, sondern auf das eigene Format und dessen gestalterische Zielsetzung bezogen.

Eine weitere methodische Begrenzung ergibt sich daraus, dass Entwicklung, Produktion und Reflexion des Formats durch dieselbe Person durchgeführt wurden. Die Bewertung der gestalterischen Umsetzung basiert dadurch teilweise auf der eigenen Produktionsperspektive. Eine subjektive Einschätzung einzelner Gestaltungsentscheidungen kann deshalb nicht vollständig ausgeschlossen werden. Die theoretische Herleitung der Kriterien und die

Referenzanalyse dienen dazu, diese Reflexion nachvollziehbar zu strukturieren, ersetzen jedoch keine unabhängige externe Bewertung.

Darüber hinaus basiert der praktische Teil auf einer einzelnen eigenen Videopodcast-Staffel mit fünf Folgen. Dadurch konnten zwar verschiedene Gäste und Gesprächssituationen umgesetzt werden, dennoch bleibt der Produktionsrahmen begrenzt. Andere Themen, Zielgruppen oder Formate könnten andere gestalterische Anforderungen mit sich bringen. Ein politischer Interviewpodcast, ein Comedyformat oder ein journalistisches Nachrichtenformat würden möglicherweise andere Anforderungen an Set, Licht, Schnitt und visuelle Identität stellen als “Sidequest Social Media”.

Auch technische und organisatorische Rahmenbedingungen beeinflussten die Umsetzung. Die Produktion fand in zwei unterschiedlichen Studios statt, wodurch die visuelle Kontinuität des Sets nicht vollständig identisch hergestellt werden konnte. Obwohl die Grundstruktur beibehalten wurde, ergaben sich durch die unterschiedlichen Räume Abweichungen in Perspektive, Lichtwirkung und Setanordnung. Diese Begrenzung zeigt, dass die visuelle Wiedererkennbarkeit eines Videopodcasts stark von räumlicher und technischer Konstanz abhängt.

Weitere Grenzen ergaben sich aus dem Produktionsaufwand. Die Arbeit mit drei Kameras, separater Tonaufnahme, Lichtsetup, 4K-Log-Material und großen Datenmengen machte die Produktion deutlich komplexer als eine reine Audioaufnahme. Technische Probleme wie Kameraüberhitzung oder Akkuprobleme konnten zwar gelöst oder abgefedert werden, zeigen aber, dass Videopodcasts zusätzliche Risiken mit sich bringen. Der gestalterische Mehrwert ist deshalb immer auch mit einem höheren Produktionsaufwand verbunden.

Außerdem wurden visuelle Zusatzinhalte nicht in jeder Folge im gleichen Umfang eingesetzt. Foto- und Videoeinblendungen wurden nur dort verwendet, wo sie inhaltlich sinnvoll waren. Dadurch ist der Mehrwert solcher Einblendungen zwar an konkreten Beispielen sichtbar, aber nicht über alle Folgen hinweg gleich stark ausgeprägt. Ähnlich verhält es sich mit den Gastgegenständen: Sie waren ein wichtiges Element des Setkonzepts, wurden aber je nach Folge unterschiedlich stark in das Gespräch eingebunden.

Schließlich wurde die tatsächliche Performance der Social-Media-Clips nicht untersucht. Die Arbeit reflektiert, wie das Material für Clips gestaltet und aufbereitet wurde, bewertet aber

nicht, welche Reichweite, Aufmerksamkeit oder Interaktion diese Clips tatsächlich erzeugen. Aussagen über den gestalterischen Nutzwert der Clips beziehen sich daher auf ihre Produktionslogik und Gestaltung, nicht auf gemessene Plattformeffekte.

Diese Grenzen machen deutlich, in welchem Rahmen die Ergebnisse dieser Arbeit zu verstehen sind. Die Arbeit zeigt, wie die visuelle Ebene in einer eigenen Videopodcast-Produktion gestaltet, praktisch angewendet und reflektiert werden kann. Sie liefert jedoch keine allgemeingültige Aussage darüber, wie alle Videopodcasts wirken oder wie Zuschauerinnen und Zuschauer Video im Vergleich zu Audio tatsächlich bewerten.

5.6 Beantwortung der Forschungsfrage

Die Forschungsfrage dieser Arbeit lautete:

Welchen gestalterischen Mehrwert bietet die visuelle Ebene eines Videopodcasts im Interviewformat im Vergleich zu einem reinen Audiopodcast?

Auf Grundlage der theoretischen Auseinandersetzung, der exemplarischen Referenzanalyse und der gestalterischen Reflexion der eigenen Produktion lässt sich die Forschungsfrage im Rahmen dieser Arbeit wie folgt beantworten: Die visuelle Ebene bietet dann einen gestalterischen Mehrwert, wenn sie das Gespräch nicht nur abbildet, sondern funktional erweitert. Der Mehrwert entsteht nicht automatisch durch das Filmen eines Podcasts, sondern durch bewusste Entscheidungen in Bildgestaltung, Licht, Set, Personendarstellung, Schnitt, grafischer Gestaltung und Social-Media-Aufbereitung.

Im Vergleich zum reinen Audiopodcast macht der Videopodcast zusätzliche kommunikative Informationen sichtbar. Während der Audiopodcast vor allem über Stimme, Sprache, Klang und Gesprächsrhythmus wirkt, zeigt der Videopodcast zusätzlich Mimik, Gestik, Blickrichtung, Körperhaltung und sichtbare Reaktionen. Gerade im Interviewformat kann dadurch deutlicher werden, wie Gesprächspartnerinnen und Gesprächspartner aufeinander reagieren, zuhören oder eine Aussage körperlich begleiten. Der Videopodcast erweitert das Gespräch damit um eine nonverbale Ebene, die im Audiopodcast nur indirekt wahrnehmbar ist.

Ein weiterer Mehrwert liegt in der räumlichen und atmosphärischen Einbettung des Gesprächs. Durch Setgestaltung, Licht und Kameraperspektiven wird das Interview nicht nur gehört,

sondern als konkrete Gesprächssituation sichtbar. Das Publikum nimmt Host und Gast in einem gestalteten Raum wahr. Im eigenen Format wurde dies insbesondere durch das wiederkehrende Set, das Regal mit variablen Gastgegenständen, die kontrollierte Lichtgestaltung und den farbigen Akzent umgesetzt.

Auch der Schnitt trägt zum gestalterischen Mehrwert bei. Er entscheidet, wann die sprechende Person, die Reaktion des Gegenübers oder die gemeinsame Gesprächssituation gezeigt wird. Dadurch kann der Videopodcast das Gespräch visuell strukturieren und bestimmte Momente hervorheben. Im Audiopodcast entsteht Struktur vor allem über Ton und Gesprächsfluss. Im Videopodcast kommt eine zusätzliche visuelle Steuerung der Aufmerksamkeit hinzu.

Darüber hinaus ermöglicht die visuelle Ebene die Einbindung zusätzlicher Inhalte. Fotos, Videos, Gegenstände oder andere visuelle Referenzen können direkt gezeigt werden, anstatt ausschließlich sprachlich beschrieben zu werden. Besonders bei einem Thema wie Social Media ist dies relevant, weil viele Gesprächsinhalte selbst visuell geprägt sind. Der Videopodcast kann dadurch zusätzliche Anschauungsmöglichkeiten bieten, sofern die Einblendungen funktional aus dem Gespräch heraus eingesetzt werden.

Neben diesem unmittelbaren gestalterischen Mehrwert entsteht durch das produzierte Videomaterial ein weiterer produktions- und verwertungsbezogener Nutzen. Aus den Aufnahmen können nicht nur vollständige Folgen, sondern auch visuelle Ausschnitte für Social Media erstellt werden. Diese Clips können Personen, Reaktionen, Untertitel, Set und Gesprächsmomente nutzen. Ein reiner Audiopodcast müsste für solche Ausspielungen dagegen erst nachträglich visualisiert werden. Die Social-Media-Verwertung stellt damit nicht ausschließlich einen gestalterischen Mehrwert innerhalb der eigentlichen Podcastfolge dar, sondern erweitert die Möglichkeiten zur Weiterverwendung des produzierten Materials.

Die Forschungsfrage kann daher zusammenfassend so beantwortet werden: Die visuelle Ebene eines Videopodcasts bietet im Interviewformat einen gestalterischen Mehrwert, weil sie nonverbale Kommunikation sichtbar macht, das Gespräch räumlich und atmosphärisch einbettet, visuelle Orientierung ermöglicht und zusätzliche visuelle Inhalte in das Gespräch integrieren kann. Dieser Mehrwert ist jedoch an bewusste Gestaltung gebunden. Ohne klare Kameraperspektiven, passende Lichtsetzung, ein durchdachtes Set, sinnvollen Schnitt und eine gezielte Postproduktion kann die visuelle Ebene ihre Funktion nicht vollständig entfalten oder

vom Gespräch ablenken. Ergänzend schafft das produzierte Videomaterial Möglichkeiten zur plattformübergreifenden Weiterverwertung, insbesondere in Form von Social-Media-Clips.

Damit stützen die Ergebnisse die in Kapitel 2 entwickelte Einordnung des Videopodcasts als audiovisuelle Erweiterung des Audiopodcasts. Der Videopodcast ist nicht als Ersatz des Audiopodcasts zu verstehen. Der Audiopodcast bleibt ein eigenständiges Format mit eigenen Stärken. Der Videopodcast ergänzt diese dort, wo Sichtbarkeit, Raum, nonverbale Kommunikation und zusätzliche visuelle Gestaltungsmöglichkeiten eine relevante Rolle spielen.

5.7 Ausblick

Aus den Ergebnissen dieser Arbeit ergeben sich mehrere mögliche Anschlussfragen. Besonders interessant wäre eine Untersuchung der Zuschauerwahrnehmung. Dabei könnte erforscht werden, ob und wie stark Zuschauerinnen und Zuschauer die visuelle Ebene eines Videopodcasts tatsächlich als Mehrwert empfinden. Eine solche Untersuchung könnte zum Beispiel verschiedene Versionen derselben Folge vergleichen: eine reine Audioversion, eine einfache Videoaufnahme und eine bewusst gestaltete Videopodcast-Version.

Ein weiterer Ansatz wäre die Untersuchung einzelner Gestaltungselemente. So könnte genauer betrachtet werden, welchen Einfluss Kameraperspektiven, Lichtstimmung, Setgestaltung oder sichtbare Reaktionen auf die Wahrnehmung eines Interviewformats haben. Auch die Frage, ob ein variables Set mit persönlichen Gegenständen die Nähe zu den Gästen erhöht oder die Wiedererkennbarkeit des Formats beeinflusst, könnte in einer weiterführenden Arbeit gezielt untersucht werden.

Auch die Social-Media-Verwertung bietet weitere Untersuchungsmöglichkeiten. In dieser Arbeit wurde vor allem reflektiert, wie Videopodcast-Material für Clips aufbereitet werden kann. Eine anschließende Untersuchung könnte analysieren, welche Clipformen tatsächlich besser funktionieren, etwa im Hinblick auf Reichweite, Wiedergabedauer, Interaktion oder Verständlichkeit. Dabei könnten unterschiedliche Varianten verglichen werden, zum Beispiel Clips mit und ohne Untertitel, unterschiedliche Hooks oder verschiedene Bildausschnitte.

Für die praktische Weiterentwicklung von “Sidequest Social Media” wäre außerdem interessant, das Format über eine weitere Staffel hinweg zu überprüfen. Dabei könnten Set, Licht, grafische Gestaltung und Clipdesign weiter vereinheitlicht und auf Grundlage der ersten

Produktion angepasst werden. Besonders die visuelle Kontinuität über mehrere Folgen hinweg könnte durch einen festeren Produktionsort und genauere Setdokumentation weiter verbessert werden.

Zudem könnte das Verhältnis von Audio- und Videofassung stärker untersucht werden. Da viele Videopodcasts weiterhin auch auditiv funktionieren müssen, wäre es sinnvoll zu betrachten, welche Gestaltungselemente ausschließlich visuell wirken und welche auch im Audio erhalten bleiben. Dadurch könnte genauer bestimmt werden, welche Inhalte wirklich einen Videopodcast benötigen und welche auch als reines Audioformat funktionieren würden.

Insgesamt zeigt die Arbeit, dass Videopodcasts im Interviewformat ein großes gestalterisches Potenzial besitzen, dieses Potenzial aber bewusst entwickelt werden muss. Weitere Untersuchungen könnten daran anknüpfen, indem sie stärker die Wirkung beim Publikum, die Performance von Social-Media-Clips oder den Vergleich unterschiedlicher visueller Gestaltungskonzepte betrachten.

Literaturverzeichnis

Blake, S. (2026). Interviewing using video-call technology: informality, individuality, and incongruity. *International Journal of Social Research Methodology*, 1–17. https://doi.org/10.1080/13645579.2026.2618698

Bruckert, A., Christie, M., & Meur, O. L. (2022). Where to look at the movies: Analyzing visual attention to understand movie editing. *Behavior Research Methods*, *55*(6), 2940–2959. https://doi.org/10.3758/s13428-022-01949-7

Cook, A., Thompson, M., & Ross, P. (2023). Virtual first impressions: Zoom backgrounds affect judgements of trust and competence. *PLoS ONE*, *18*(9), e0291444. https://doi.org/10.1371/journal.pone.0291444

Croes, E. a. J., Antheunis, M. L., Schouten, A. P., & Krahmer, E. J. (2018). Social attraction in video-mediated communication: The role of nonverbal affiliative behavior. *Journal of Social and Personal Relationships*, *36*(4), 1210–1232. https://doi.org/10.1177/0265407518757382

Cui, C. (2024). Enhancing Brand Recognition through Simplified Visual Identity: Impact and Benefits. *SHS Web of Conferences*, *207*, 02019. https://doi.org/10.1051/shsconf/202420702019

Cutting, J. E. (2016). Narrative theory and the dynamics of popular movies. *Psychonomic Bulletin & Review*, *23*(6), 1713–1743. https://doi.org/10.3758/s13423-016-1051-4

Cutting, J. E., & Armstrong, K. L. (201). Facial expression, size, and clutter: Inferences from movie structure to emotion judgments and back. *Attention Perception & Psychophysics*, *78*(3), 891–901. https://doi.org/10.3758/s13414-015-1003-5

Fauviaux, T., Marin, L., Parisi, M., Schmidt, R., & Mostafaoui, G. (2024). From unimodal to multimodal dynamics of verbal and nonverbal cues during unstructured conversation. *PLoS ONE*, *19*(9), e0309831. https://doi.org/10.1371/journal.pone.0309831

Furlan, M. (2021). Non-verbal cues of engagement during video interviews: Third-party assessment, construction and validation of a training and automatic detection (Dissertation, Universität Padua).

Gao, P., Jiang, H., Xie, Y., & Cheng, Y. (2021). The triggering mechanism of short video customer inspiration – Qualitative analysis based on the repertory grid technique. *Frontiers in Psychology*, *12*, 791567. https://doi.org/10.3389/fpsyg.2021.791567

Huttunen, S. (2025). Faces in shadows: silhouette light, underlight and toplight elicit increased early posterior negativity. *Frontiers in Neuroscience*, *19*, 1553977. https://doi.org/10.3389/fnins.2025.1553977

Klenk, M. (2020). Stimme, Sprechen, hören. *Kommunikation Gesellschaft*, *21*(2). https://doi.org/10.15460/kommges.2020.21.2.623

Knobloch, H., Von Zur Mühlen, B., & Zimper, M. (2020). Wir müssen reden – Die Podcast-Story. *MedienWirtschaft*, *17*(2–3), 49–64. https://doi.org/10.15358/1613-0669-2020-2-3-49

Lelis, C., Leitão, S., Mealha, Ó., & Dunning, B. (2020). Typography: the constant vector of dynamic logos. *Visual Communication*, *21*(1), 146–170. https://doi.org/10.1177/1470357220966775

Mondada, L. (2008). Using video for a sequential and multimodal analysis of social interaction: videotaping institutional telephone calls. *Edoc (University of Basel)*, *9*(3), 35. https://doi.org/10.17169/fqs-9.3.1161

Poland, Jennifer Lee, "Lights, Camera, Emotion!: an Examination on Film Lighting and Its Impact on Audiences' Emotional Response" (2015). ETD Archive. 379. https://engagedscholarship.csuohio.edu/etdarchive/379

Preger, S. (2020). Dokumentarisches Erzählen in Podcasts. *Kommunikation Gesellschaft*, *21*(2). https://doi.org/10.15460/kommges.2020.21.2.634

Rime, J., Pike, C., & Collins, T. (2022). What is a podcast? Considering innovations in podcasting through the six-tensions framework. *Convergence the International Journal of Research Into New Media Technologies*, *28*(5), 1260–1282. https://doi.org/10.1177/13548565221104444

Sanz-Aznar, J., Bruni, L. E., & Soto-Faraco, S. (2023). Cinematographic continuity edits across shot scales and camera angles: an ERP analysis. *Frontiers in Neuroscience*, *17*, 1173704. https://doi.org/10.3389/fnins.2023.1173704

Sutton-Brady, C., Scott, K. M., Taylor, L., Carabetta, G., & Clark, S. (2009). The value of using short-format podcasts to enhance learning and teaching. *ALT-J*, *17*(3), 219–232. https://doi.org/10.1080/09687760903247609

Szarkowska, A., Ragni, V., Szkriba, S., Black, S., Orrego-Carmona, D., & Kruger, J. (2024). Watching subtitled videos with the sound off affects viewers' comprehension, cognitive load, immersion, enjoyment, and gaze patterns: A mixed-methods eye-tracking study. *PLoS ONE*, *19*(10), e0306251. https://doi.org/10.1371/journal.pone.0306251

Violot, C., Elmas, T., Bilogrevic, I., & Humbert, M. (2024). Shorts vs. regular videos on YouTube: A Comparative analysis of user engagement and content creation trends. *ACM Digital Libary*, 213–223. https://doi.org/10.1145/3614419.3644023

Yu, M., Abidin, S. B. Z., & Shaari, N. B. (2024). Effects of Brand Visual Identity on Consumer Attitude: A Systematic Literature review. *Preprints.org*. https://doi.org/10.20944/preprints202405.1109.v1

Zannettou, S., Nemeth, O., Ayalon, O., Goetzen, A., Gummadi, K. P., Redmiles, E. M., & Roesner, F. (2023). Analyzing User Engagement with TikTok's Short Format Video Recommendations using Data Donations. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2301.04945

Zhang, Z., Qiu, K., & Ye, Y. (2025). Influence of audiovisual features of short video advertising on consumer engagement behaviors: Evidence from TikTok. *Journal of Business Research*, *201*, 115662. https://doi.org/10.1016/j.jbusres.2025.115662

Zhu, Z., Liu, S., & Zhang, R. (2023). Examining the persuasive effects of health communication in short videos: Systematic review. *Journal of Medical Internet Research*, *25*, e48508. https://doi.org/10.2196/48508

Online- und Branchenquellen

Spotify. (2023, 29. Juni). Spotify tops 100,000 video podcasts. Spotify Newsroom. https://newsroom.spotify.com/2023-06-29/spotify-tops-100000-video-podcasts/

Spotify. (2024, 28. Juni). 250,000 video podcasts and counting: Creators and audiences are embracing video content on Spotify. Spotify Newsroom. https://newsroom.spotify.com/2024-06-28/250000-video-podcasts-and-counting-creators-and-audiences-are-embracing-video-content-on-spotify/

Medienquellen / analysierte Referenzen

Baywatch Berlin. (2026, 16. Juli). Wie würdet ihr das machen? [Social-Media-Clip]. Instagram. Abgerufen am 12.08.2026. https://www.instagram.com/reel/DaxLTtnjNgC/?igsh=YTl0YnNiMXcwZjk=&igsi=YTl0YnNiMXcwZjk=

Clare on Air. (2026, 10. August). Jermain über Therapie, Ozempic und Schönheitsideale - Videopodcast. Spotify. Timecodes: 28:57 und 29:39. Abgerufen am 12.08.2026. https://open.spotify.com/episode/0e6Cy3P9dyR5jR5SCTPYuH?si=bf5d79bea82c4aa4

The Joe Rogan Experience. (2026, 16. Juli). #2527 – MrBeast Videopodcast. Spotify. Timecodes: 05:26 und 48:31. Abgerufen am 12.08.2026. https://open.spotify.com/episode/5lokpznqvSrJO3gButgQvs?si=6eca23d4b8e84c17

SWMR_Podcast. (2026, 2. Juli). Schatz, wir müssen reden: Healing Session mit Badmómzjay Videopodcast. YouTube. Abgerufen am 12.08.2026. https://open.spotify.com/episode/6yWDKHfja56BsdQXjqN3pX?si=6583165ce53948c2

SWMR_Podcast. (2026, 5. Februar). Schatz, wir müssen reden: Handy tauschen mit Mali Videopodcast. YouTube. Abgerufen am 12.08.2026. https://open.spotify.com/episode/3V6fGhFOFOfheILEsVSr9v?si=d8148210c2a14aa3

Tim Gabel Podcast. (2026, 9. August). Nordkorea Insider: Die Wahrheit über das Leben in Nordkorea – Dr. Rüdiger Frank Videopodcast. Spotify. Timecodes: 44:10, 44:23 und 44:42. Abgerufen am 12.08.2026. https://open.spotify.com/episode/1P9b1lVTjXbB7xvU0bb9hM?si=68dec27b89a94a06

Abbildungsverzeichnis

Abbildung 1,2,3: Mehrkamera-Auflösung im „Tim Gabel Podcast“: gemeinsame Einstellung sowie halbnahe Einzelperspektiven von Host und Gast
Quelle: Screenshots aus *Tim Gabel Podcast: Nordkorea Insider: Die Wahrheit über das Leben in Nordkorea – Dr. Rüdiger Frank*, Spotify, 09.08.2026

Abbildung 4: Einbindung eines externen Videos in „The Joe Rogan Experience“
Quelle: Screenshot aus *The Joe Rogan Experience #2527 – MrBeast*, Spotify, 16.07.2026.

Abbildung 5,6: Vergleich der Setgestaltung in zwei Folgen von „Schatz, wir müssen reden“
Quelle: Screenshots aus *Healing Session mit Badmómzjay*, 02.07.2026, und *Handy tauschen mit Mali*, YouTube, 05.02.2026

Abbildung 7,8: Lichtgestaltung und atmosphärische Farbwirkung im Videopodcast „Clare on Air“
Quelle: Screenshots aus *Clare on Air: Jermain über Therapie, Ozempic und Schönheitsideale*, Spotify, 10.08.2026

Abbildung 9,10,11: Gestalterische Aufbereitung eines Podcastausschnitts als Social-Media-Clip bei „Baywatch Berlin“
Quelle: Screenshots aus Baywatch Berlin, *Wie würdet ihr das machen?*, Instagram, 16.07.2026.

Abbildung 10: Visuelle Grundgestaltung von „Sidequest Social Media“ am Beispiel der Folge mit David
Quelle: eigener Screenshot aus der Produktion.

Abbildung 11,12,13: Mehrkamera-Auflösung von „Sidequest Social Media“: Totale sowie halbnahe Einzelperspektiven von Host und Gast
Quelle: eigene Screenshots aus der Folge mit Fabio.

Abbildung 14,15: Anwendung der Drittelregel bei den halbnahen Einzelperspektiven in der Folge mit Lilly
Quelle: eigene Darstellung auf Grundlage eigener Aufnahmen.

Abbildung 16: Cover des Videopodcast-Formats „Sidequest Social Media“
Quelle: eigene Darstellung.

Abbildung 17: Gestaltung eines Social-Media-Clips von „Sidequest Social Media“ im 9:16-Format
Quelle: eigener Screenshot aus der Produktion.

X^